Advancing AI-driven lung ultrasound diagnostics for COVID-19: Procedural data synthesis, image enhancement, and pleural line detection

A Thesis Submitted for the Degree of Doctor of Philosophy

By

Cameron Hill

Department of Electronic and Electrical Engineering, Brunel University London

2025

Abstract

This thesis presents novel advancements in lung ultrasound imaging through the integration of computational techniques and artificial intelligence. Addressing critical challenges such as low contrast, noise, data scarcity, and diagnostic variability, the research focuses on three core areas: contrast enhancement, synthetic data generation, and AI-driven diagnostic models for COVID-19 detection.

A contrast enhancement technique utilising Rayleigh Gaussian Mixture Models and k-means clustering was developed to improve image clarity while preserving diagnostic features. This method demonstrated substantial improvements in pleural line detection accuracy, with Support Vector Machine (SVM)-based models achieving 84.3% accuracy and superior precision metrics.

To mitigate the scarcity of annotated lung ultrasound datasets, synthetic image generation methods were implemented. Generative Adversarial Networks (GANs) were used to produce realistic ultrasound images, achieving a Structural Similarity Index (SSIM) of 0.46 and a Fréchet Inception Distance (FID) of 257.95, showcasing their potential in addressing data limitations.

The thesis also introduces P-Net, a multi-architecture neural network ensemble, designed to classify lung ultrasound images for COVID-19 diagnostics. P-Net integrates segmentation outputs and latent features, achieving a Dice coefficient of 0.87 and Intersection-over-Union (IoU) of 0.89, demonstrating robust diagnostic accuracy even in limited data scenarios.

These contributions enhance the utility of lung ultrasound imaging, particularly for diagnosing COVID-19, by combining advanced computational methods with clinical needs. This research underscores the transformative potential of artificial intelligence in medical imaging, paving the way for improved patient outcomes and innovations in healthcare technology.

Acknowledgements

Completing this Ph.D. has been a transformative and challenging journey, and it would not have been possible without the unwavering support, guidance, and encouragement of numerous individuals and institutions.

First and foremost, I wish to express my deepest gratitude to my principal advisor, Md Nazmul Huda, for his steadfast support, insightful guidance, and consistent encouragement throughout my research journey. Your patience and expertise have been a cornerstone of my success, and I am profoundly grateful for your mentorship.

I would also like to extend my heartfelt thanks to my former principal advisor, Abdul Sadka, for his invaluable guidance and thoughtful advice during the early stages of my Ph.D. Your support provided a strong foundation for this work.

My sincere appreciation goes to Brunel's Institute of Digital Futures, whose financial assistance was instrumental in enabling this research to come to fruition.

On a personal note, I am deeply thankful to my family, my partner's family, and, above all, my loving partner, Senifa. Your love, understanding, and encouragement have been my greatest source of strength throughout this journey. Your belief in me has been a constant inspiration, and I dedicate this achievement to you all.

Contents

List of Figures

List of Tables

Acronyms

ADMSS Adaptive Denoising Multi-Scale Smoothing.

AI Artificial Intelligence.

BM3D Block-Matching and 3D Filtering.

BRISQUE Blind Reference Image Spatial Quality Evaluator.

CLAHE Contrast-Limited Adaptive Histogram Equalization.

CNN Convolutional Neural Network.

CNR Contrast-to-Noise Ratio.

COVID-19 Coronavirus Disease 2019.

CT Computed Tomography.

GAN Generative Adversarial Network.

GMM Gaussian Mixture Model.

HMM Hidden Markov Model.

IOU Intersection over Union.

LUS Lung Ultrasound Scan.

MERS Middle East Respiratory Syndrome.

MSE Mean Squared Error.

OBNLM Optimized Bayesian Non-Local Means.

P-Net A proposed network architecture for segmentation and classification.

PCR Polymerase Chain Reaction.

PDF Probability Density Function.

PIQE Perceptual Image Quality Evaluator.

PSNR Peak Signal-to-Noise Ratio.

ReLU Rectified Linear Unit.

RR Rayleigh Rician.

SARS Severe Acute Respiratory Syndrome.

SGD Stochastic Gradient Descent.

SSIM Structural Similarity Index.

SVM Support Vector Machine.

U-Net A specific neural network architecture for biomedical image segmentation.

VAE Variational Autoencoder.

VGG Visual Geometry Group (architecture).

Chapter 1

Introduction

The COVID-19 pandemic has underscored the necessity for advanced technological intervention in medical diagnostics, calling for swift, reliable, and nuanced recognition of diseases from medical imaging. In this context, the development of an artificial intelligence AI-enhanced diagnostic model is proposed, with a focus on not only identifying COVID-19 but also an array of other respiratory conditions, ensuring the system's comprehensive applicability and utility. Additionally, this study acknowledges the challenges posed by limited access to labelled training data and the specific technical skills required for operating sophisticated diagnostic equipment. In response, innovative solutions are integrated, notably the synthesis of lung ultrasound data and the enhancement of ultrasound accessibility for clinical staff, thus ensuring the model's adaptability and wider relevance.

Prior to network training, an essential phase of this research encompasses a comprehensive exploration into ultrasound image processing techniques to enhance the raw image features. The ultrasound scans, known for their granularity and susceptibility to noise, often necessitate a level of refinement to ensure clearer, more distinguishable features that are critical for accurate analysis.

Enhancement techniques, including but not limited to speckle reduction, edge enhancement, and contrast adjustment, are meticulously applied to improve the definition and clarity of anatomical structures and potential pathological indicators in the scans. This preprocessing step is crucial in addressing common challenges in ultrasound image interpretation, such as low contrast between different tissues and the presence of acoustic noise. By streamlining

these images into a more analytically conducive format, the AI network is better positioned for successful training, allowing for more precise learning based on high-quality, feature-rich data.

One significant hurdle in the deployment of AI in medical diagnostics is the scarcity of labelled training data, often due to patient privacy protections and the novel nature of emerging diseases like COVID-19. To address this deficit, the project leverages lung ultrasound data synthesis, utilising techniques such as Generative Adversarial Networks (GANs) to generate realistic lung ultrasound images. These synthetic images, mimicking diverse manifestations of various respiratory conditions, are instrumental in providing the nuanced data necessary for training robust diagnostic algorithms. This methodology not only compensates for the lack of real patient data but also ensures the AI model's efficacy in accurately diagnosing diseases.

Moreover, the diagnostic utility is significantly enhanced by its applicability to a variety of conditions. Although the immediate motivation for this research is the COVID-19 crisis, the system is engineered with the flexibility to diagnose a wide spectrum of upper respiratory diseases. This adaptability is facilitated by enabling the AI to recognise a comprehensive set of patterns and markers associated with different respiratory illnesses, drawing upon a rich, synthesised dataset that encompasses multiple conditions. Therefore, the system is positioned not merely as a reactionary measure to the current pandemic but as a lasting, versatile medical asset, adaptable to various healthcare scenarios and potential future outbreaks.

Addressing the crucial need for enhanced interpretability of diagnostic imaging, particularly for clinical staff without specialised training, the research delves into the advanced segmentation capabilities of the AI network in processing ultrasound scans. Traditional ultrasound imaging can often present a challenge for non-specialist medical professionals due to its complexity in discerning critical features within the scans. However, the proposed AI model alleviates this challenge by employing sophisticated segmentation algorithms designed to isolate and highlight key anatomical structures and potential pathological features within the respiratory system.

The system's enhanced visualisation techniques are calibrated to offer clarity, which is

especially vital in high-pressure decision-making environments. By reducing the ambiguity commonly associated with traditional ultrasound imagery, the model serves not just as a diagnostic tool but also as an educational resource, bridging the knowledge gap for healthcare professionals with less experience in sonography.

1.1 Background

At the end of 2019, a worldwide viral outbreak caused some of the worst global pressure on clinical healthcare in recent modern times. The virus, SARS-CoV-2 or commonly referred to as COVID-19 is a moderately infectious disease with a high mortality rate, however, due to new data being released frequently this statement could change. Coronavirus falls into a subfamily which then falls into its own four genera. The one causing the current COVID-19 pandemic is a Beta coronavirus; severe acute respiratory syndrome human coronavirus (SARS-HCov)

Data suggests male patients aged between 34 and 59 years seem to be the most infected [1]. It is also apparent it is more likely to infect people with chronic illnesses such as cardiovascular diseases as well as cerebrovascular and diabetes. Most paediatric patients that have the disease show mild symptoms, displaying no fever or pneumonia with a good prognosis. An interesting feature is that even though paediatric patients are asymptomatic they still show ground-glass opacities in their radiology results.

In a clinical setting, COVID-19 and SARS share many features, making the usage of previous studies on SARS for clinical diagnosis information useful. COVID-19 also causes some patients to suffer from ARDS (Acute respiratory stress syndrome) and septic shock which leads to multiple organ failures with 10% of patients dying. Unlike SARS and MERS, COVID-19 seems to prefer the lower respiratory system unlike rhinorrhoea which produces symptoms in the upper respiratory tract, such as a sore throat or sneezing [1]

The information on COVID-19 is constantly evolving due to rapid pressure on the research field, so all current data on the pandemic must be reviewed frequently as it is changing dynamically with new results. This has been taken into consideration in this review, with the aim to update the literature as the project continues.

1.1.1 Current diagnosis

Diagnosing COVID-19 correctly and quickly is critical in managing the virus. Due to its effect on the global economy and healthcare, effective screening methods are highly in demand; with more effective early screening tools required to be developed.

Initially, the primary method of diagnosis was to use X-Ray imaging of the chest region due to gaps in information surrounding COVID-19. Following radiographic diagnosis, PCR testing was the next stage in diagnosis once the virus had been sequenced. With the sequenced genomics of the virus, the PCR test could be targeted for specific biological characteristics of the virus. Another method of detection has recently been used, focusing on eliminating the problems associated with traditional testing, this being biosensor-based detection. Biosensors combine the usage of biological components and in-organics to create a device that can analyse the results that is easy to use.

Now focusing specifically on imaging diagnosis, the first common method is the usage of CT scans. A CT scan images organs within the body using X-rays to create a 3D cross-sectional view. For COVID-19 the abnormalities present within these scans show up as ground-glass opacities on the pulmonary system. The CT scan method has proven in cases to be more effective than PCR testing with its accuracy in detecting COVID-19 but also provides incredibly useful information in the patient's condition as a result of the infection, aiding in later-stage disease management. However, CT scans do not come without challenges, primarily with the expense of the unit and the biological harm it can inflict on a patient. The radiation exposure is a large downside to this technique, meaning continued usage on repeat patients is highly un-recommended. Additionally, with the high volume of patients presenting with COVID-19 symptoms, using CT scans is impractical. [2] [3]

1.1.2 Ultrasound and COVID-19

Following on from imaging techniques for diagnosis, the next and least utilized is sonographic methods. Lung ultrasound is a useful diagnostic method, unlike CT-Scans and biological diagnosis it provides instant results and prognosis from the image. LUS can be more adept at identifying biomarkers within the lungs such as pulmonary effusion and consolidation.

Additionally, unlike other imaging, the ultrasound can be used to quickly image cardiac behaviour, providing more evidence for a diagnosis for COVID-19. LUS is a quick and efficient method of diagnosing a patient at the bedside, paired with low economical costs of the equipment and fast times in providing results; for the purposes of COVID-19 detection, it has huge advantages over the current commonly used approaches such a CT-Scans and PCR testing. It additionally removes the health risks posed by radiographs by eliminating radiation risk as well as also providing unique biosignatures not seen in CT-Scans.

LUS does however come with its limitations and downsides. Firstly, LUS encounters a problem that also affects other imaging techniques in that diagnosis specificity can be relatively low in asymptomatic cases as well as viral pneumonia. The biomarkers can be identical in some cases making a definite diagnosis on COVID-19 or other respiratory conditions sometimes a challenge and would therefore in cases need to be backed up with PCR or biological testing results. LUS also faces the problem of the clinician needing extensive training to be able to read and use the LUS device, impacting the amount clinicians available within a medical environment to give a diagnosis. [3]

This is where the main focus and drive of this project comes into play. The usage of AI on image diagnosis would be a powerful tool in combatting the number of clinical staff experienced enough to analyse images and aid in quick and more decisive results when detecting not only COVID-19 but other respiratory conditions. Although there seems to be a vast collection of active research being done on CT-Scan AI as well as MRI; ultrasound seems to be a fairly un-explored approach when in comparison to those techniques. With the portability of ultrasound, its cost-effective design, and the power of modern smartphones; an AI diagnostic system could be developed and distributed to socioeconomically deprived nations around the world to help combat the current pandemic and potentially improve clinical health care overall.

1.1.3 Lung anatomy in relation to ultrasound

The human respiratory system, with the lungs as its primary organ, is responsible for oxygenating the blood and expelling carbon dioxide. Each lung is enveloped in a serous membrane known as the pleura, comprising two layers: the visceral pleura (inner layer) and the

parietal pleura (outer layer). The space between these layers, the pleural cavity, contains a lubricating fluid that facilitates lung expansion.

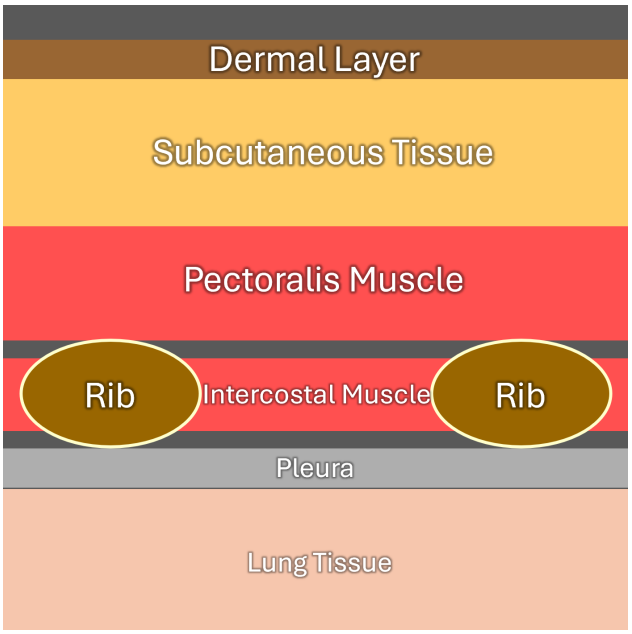

Figure 1.1: Lung cross-section in relation to Lung Ultrasound

The lungs, located in the thoracic cavity, are divided into lobes; the right lung has three — the upper, middle, and lower lobes, while the left lung has two — the upper and lower lobes. This asymmetry accommodates the heart's position. Each lobe is further segmented, featuring multiple tertiary bronchi that serve specific segments. These complex segmentations are integral to localized function and pathology.

Air enters through the trachea, which bifurcates into the left and right main bronchi, leading to the lungs. The bronchi further divide into smaller bronchioles and eventually terminate at the alveolar sacs. It is here, at these tiny, balloon-like structures, that the respiratory gas exchange occurs across the alveolar-capillary membrane.

Figure 1.1 can be used to simply describe a cross-section of the upper lung anatomy, in relation to the descending path ultrasound waves would propagate towards the internal tissues

Ultrasound imaging of the lungs relies on the interpretation of features that are intrinsically tied to the underlying anatomical structures. The interface between the air in the alveoli and the surrounding tissues creates a unique acoustic environment, chiefly because

air reflects and scatters sound waves, producing distinctive artifact patterns rather than the clear images generated in other organs. For example, the pleural line which is the bright, hyperechoic line seen in lung ultrasound, corresponds to the visceral and parietal pleural layers' interface, providing information about the pleural space and potential abnormalities. Below this layer, the presence of A-lines, horizontal artifacts arising from the air-tissue interface, generally indicate normal air-filled lung parenchyma. In contrast, the vertical B-lines, emerging from the pleural line down to the edge of the screen, suggest thickening of the interlobular septa, often due to fluid accumulation or fibrosis. Additionally, lung sliding, the to-and-fro movement seen on the pleural line, corresponds to the respiratory movement of the visceral pleura against the parietal pleura. Alterations or absence of this sliding can indicate pathological adhesions, pleural thickening, or pneumothorax. Each of these ultrasound features and artifacts is a direct representation or consequence of the anatomical structures within the lungs, and their presence or absence helps clinicians infer the state of the lung tissue, aiding in immediate and accurate diagnosis and management.

1.1.4 Lung ultrasound regions of interest

In lung ultrasound imaging, accurate identification of specific structures within the images is crucial for effective diagnosis. Lung ultrasounds typically display several key structures, each significant for understanding patient conditions.

Figure 1.2 highlights various key biomarkers on lung ultrasound images. The two dark areas represent the shadows cast by the rib cage, while the prominently marked region at the top signifies the epidermal and thin muscular layers before the ribs. The pleural line is of particular importance. This line, marking the boundary of the pleura—the delicate tissue enveloping the lungs—is crucial for lung diagnosis, as it precedes other significant diagnostic information.

Another key feature in lung ultrasound (LUS) diagnostics is the presence of air bronchograms. These are air-filled bronchi made visible against fluid-filled alveoli, often indicative of conditions such as pneumonia resulting from lung consolidation.

Furthermore, A-Lines and B-Lines are ultrasound artifacts that are essential for diagnostic interpretation. A-Lines, appearing as echo reverberations extending from the pleural

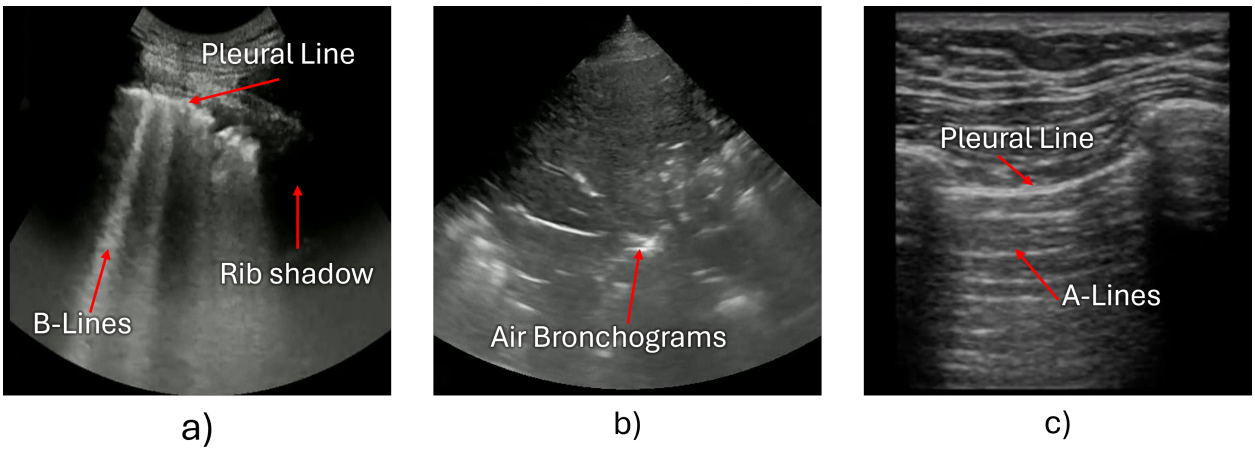

Figure 1.2: Lung Ultrasound biomarkers

line, can indicate either normal lung function or pneumothorax. In contrast, B-Lines, also known as lung rockets, are vertical artifacts originating at the pleural line. Although the exact nature of B-Lines is not fully understood, they are generally associated with various lung pathologies, including Acute Respiratory Distress Syndrome (ARDS), pneumonia, lung fibrosis, and notably, COVID-19. These artifacts are thus integral to assessing lung health. [4]

1.1.5 Speckle Noise in Ultrasound

Noise is an unavoidable artefact in imaging systems, degrading the clarity and usability of the resulting images. In ultrasound imaging, speckle noise is a particularly prominent challenge, arising from the coherent nature of sound wave propagation and scattering within tissue. This artefact, characterised by granular patterns, results from the constructive and destructive interference of ultrasonic waves reflecting off scatterers, such as cells or fibrous structures. Speckle noise is unique to coherent imaging modalities and significantly degrades the resolution and contrast of ultrasound images, obscuring vital anatomical details crucial for accurate diagnosis [5, 6, 7].

The characteristics and behaviour of speckle noise set it apart from other noise types encountered in medical imaging. Unlike additive Gaussian noise, speckle noise is multiplicative and its intensity varies with the local intensity of the image, a phenomenon referred to as

non-stationarity. Mathematically, speckle noise in ultrasound images can be expressed as:

$$y_{ij} = X_{ij} \cdot n_{ij} \tag{1.1}$$

where y_{ij} represents the observed speckled image, X_{ij} the original noiseless image, and n_{ij} the multiplicative noise. To simplify processing, a logarithmic transformation is often applied to convert this multiplicative noise model into an additive form:

$$\ln(y_{ij}) = \ln(X_{ij}) + \ln(n_{ij}) \tag{1.2}$$

This transformation enables the application of noise suppression techniques traditionally designed for additive noise models, such as wavelet-based methods [5, 6]. Speckle noise is further characterised by statistical distributions such as Rayleigh or Fisher-Tippett, depending on the image transformation and acquisition context [7].

To mitigate the adverse effects of speckle noise, various methods have been developed, broadly classified into spatial domain filtering and transform domain approaches. In the spatial domain, filters such as the Wiener, Lee, and Kuan filters are commonly used. The Wiener filter minimises the mean square error between the actual and estimated signal by leveraging both global and local statistics. However, its performance near edges can be limited due to over-smoothing. The Lee and Kuan filters adaptively use local statistics to distinguish noise from image features, offering better edge preservation, although at the cost of increased computational complexity [5, 6].

Transform domain methods provide more advanced solutions by decomposing the image into multiple frequency or directional subbands. The wavelet transform is a widely used technique that separates high-frequency components, where noise is prevalent, from low-frequency components, preserving anatomical structures. Techniques such as Bayes' shrinkage and soft thresholding have been effectively employed in wavelet-based despeckling [6]. Building on the wavelet approach, contourlet transforms incorporate directional decomposition, improving edge and contour preservation and offering enhanced performance for structural noise suppression. Hybrid methods, such as combining Gaussian filtering with wavelet or contourlet transforms, address mixed noise scenarios, providing robust solutions for diverse imaging contexts [5, 6].

Despite these advancements, speckle noise reduction remains a complex and computationally demanding task. Transform domain methods, while effective, require significant processing resources, making real-time application challenging. Additionally, many techniques rely on empirically tuned parameters, limiting their generalisability across different imaging conditions and equipment. These limitations highlight the need for further innovation in real-time despeckling algorithms, machine learning integration, and clinical validation to translate technical improvements into tangible diagnostic benefits [5, 6].

The presence of speckle noise in ultrasound images poses significant challenges, reducing image quality and complicating both visual interpretation and automated image processing. By obscuring subtle details and introducing variability, speckle noise undermines the reliability of diagnostic assessments. While ongoing advancements in noise reduction techniques offer promising solutions, addressing the computational and generalisability challenges is essential for unlocking the full potential of ultrasound imaging. As research continues, the focus must remain on developing efficient, adaptable, and clinically validated methods to mitigate speckle noise and enhance the diagnostic utility of ultrasound.

1.2 Challenges

The introduction to the thesis delineates a comprehensive strategy for incorporating AI within the sphere of medical diagnostics, with a specific focus on respiratory diseases using ultrasound imaging. Various challenges and the corresponding solutions proposed in the project are evident, and presented below:

1.2.1 Quality of ultrasound imaging

Ultrasound scans inherently possess a level of granularity and are often beset by 'noise', complicating the task of distinguishing essential features. The necessity for image enhancement, encompassing speckle reduction, edge enhancement, and contrast adjustment, indicates a dependence on advanced preprocessing steps to render these scans fit for AI examination. Such pre-processing is vital for the AI model's success, as more refined, distinguishable features contribute to more precise learning and analysis. Nonetheless, this introduces the challenge

of potential alteration or distortion of critical features by these enhancement techniques, necessitating a balance between clarity and maintenance of original data authenticity.

1.2.2 Scarcity of labelled data

The adherence to patient privacy protocols and the unprecedented nature of new diseases typically result in a shortage of crucial labelled training data for AI models. The study suggests the synthesis of lung ultrasound data via Generative Adversarial Networks (GANs) to counteract this [8] .While innovative, this solution poses challenges: it must ensure the synthetic images are adequately varied and realistic to facilitate effective model learning, and it must confirm that the model's performance with synthesised data is consistent with its performance with authentic, patient-derived data.

1.2.3 Diagnostic versatility and adaptability

Although the immediate catalyst for this research is COVID-19, the diagnostic model for this project should be designed for a wider range of respiratory illnesses, such as pneumonia. This prerequisite heightens the complexity of the model, necessitating an algorithm capable of identifying a diverse array of disease markers. The challenge here is in guaranteeing that the model, while extensive, maintains high precision ; does not need to be 100%, it must still be monitored by health care proffesionals as per NHS regulations and does not increase false positives or negatives owing to its broad diagnostic purview.

In summary, the thesis advocates for an ambitious project that, while promising significant advancements in medical diagnostics concerning respiratory illnesses, confronts several intrinsic challenges related to AI-based image processing, data availability and quality, the variance in user expertise, and the multifunctional applicability of the proposed system. Each of these challenges demands meticulous consideration and innovative problem-solving to ensure the system's effectiveness, reliability, and accessibility in practical medical environments.

1.3 Aim and Objectives

The principal aim of this study is to devise an AI-enhanced diagnostic model proficient in the analysis of lung ultrasound data, accurately distinguishing COVID-19 as well as a variety of other respiratory diseases, thereby augmenting the efficiency of medical diagnostic protocols. This paramount aim encompasses the ambition to address not only the immediate healthcare exigencies, such as the COVID-19 pandemic, but also to forge a resilient technological infrastructure that can be adapted to confront future medical exigencies.

1.3.1 Objective I - Improve lung ultrasound image quality

To explore and employ a range of image processing enhancements to augment the quality of raw ultrasound scans, ensuring the elimination of noise and amplification of critical features that are essential for precise diagnostic evaluation. This objective entails a methodical analysis of techniques including speckle reduction, edge enhancement, and contrast adjustment.

1.3.2 Objective II - Synthesise diverse lung ultrasound images

To counteract the paucity of accessible labelled ultrasound data, specifically concerning emerging diseases like COVID-19, by employing Generative Adversarial Networks (GANs) to produce high-calibre, synthetic lung ultrasound images that emulate an assortment of respiratory conditions.

1.3.3 Objective III - Implement image segmentation and visualisation

To integrate advanced segmentation algorithms within the AI model, simplifying the interpretation of ultrasound scans by accentuating key anatomical and pathological markers. This objective is crucial in rendering the diagnostic instrument both accessible and beneficial to clinical staff with diverse levels of expertise in medical imaging.

1.3.4 Objective IV - Construct a robust artificial intelligence diagnostic model

To establish and educate a refined AI model utilising the enhanced and synthesised ultrasound images. This model is expected to accurately diagnose an array of respiratory conditions, with an immediate emphasis on COVID-19 detection, based on discernible patterns and indicators within the scans.

1.3.5 Objective V - Assure model adaptability and versatility

To engineer the diagnostic model with inherent adaptability, ensuring its applicability extends beyond the immediate health crisis. This involves embedding capabilities for perpetual learning and adaptation, facilitating the recognition and diagnosis of an extensive range of respiratory maladies.

1.4 Key contributions

Here the key contributions of the thesis are highlighted,, which are encapsulated in three journal papers currently under submission or review. The research focuses on enhancing COVID-19 diagnostics using lung ultrasound imaging, addressing critical challenges in this domain with novel methodologies and insights.

Together, these contributions provide significant advancements in leveraging lung ultrasound imaging for COVID-19 diagnostics, offering both theoretical innovations and practical implications in the field of medical imaging and AI.

1.4.1 K-Means clustering and Rayleigh Gaussian Mixture Models for enhanced lung ultrasound imaging with SVM-based pleural line detection

This contribution, combining Chapters 3 and 4 of the thesis, presents a novel approach to addressing critical challenges in lung ultrasound imaging through advancements in contrast

enhancement and anatomical feature detection. Together, these chapters form the foundation of a journal manuscript currently under submission.

Chapter 3 introduces an innovative contrast enhancement technique utilising Rayleigh mixture models and k-means clustering. This method overcomes limitations of traditional approaches, such as Contrast Limited Adaptive Histogram Equalisation (CLAHE), which often results in oversaturation and diminished image quality in lung ultrasound (LUS) images. By segmenting ultrasound images using k-means clustering and applying probability density functions (PDFs) as guide distributions, this approach enhances image clarity while preserving essential diagnostic features.

Chapter 4 focuses on an advanced method for pleural line detection using Support Vector Machine (SVM) models. By leveraging low-level statistical information and clustering techniques, this method achieves an 84.4% accuracy rate in identifying the pleural line, comparable to traditional Hidden Markov Models (HMMs).

Together, these chapters address the inherent challenges of lung ultrasound image processing, offering advancements in contrast enhancement and feature detection. These methods enhance the diagnostic utility of lung ultrasound technology and form a key contribution to this thesis.

1.4.2 Overcoming labelled segmentation data scarcity with procedural generation of synthetic lung ultrasound images

This chapter details a novel approach to addressing the scarcity of real lung ultrasound data for training machine learning models by synthesising authentic and lifelike images. This work is presented as Chapter 5 of this thesis.

The chapter introduces the use of Generative Adversarial Networks (GANs) trained on semantic label images to generate realistic lung ultrasound visuals. It further enhances image authenticity through simulations that replicate ultrasound physics, incorporating techniques such as Sobel filters to simulate natural ultrasound textures. The approach leverages the Spatially-Adaptive De-normalisation (SPADE) GAN network, achieving a Structural Similarity Index (SSIM) of 0.46 and a Fréchet Inception Distance (FID) of 257.95, which high-

lights moderate structural alignment but also room for improvement in realism compared to state-of-the-art GANs.

By providing a comprehensive solution to the challenge of limited real ultrasound datasets, this chapter significantly advances the intersection of medical imaging and artificial intelligence. The proposed methods enable the development of robust machine learning models, ultimately enhancing healthcare applications reliant on ultrasound diagnostics.

1.4.3 P-Net: Overcoming data limitations in ultrasound imaging with a ensemble neural network

This contribution, presented in Chapter 6 of the thesis, introduces "P-Net," an innovative deep learning architecture designed to enhance the analysis of lung ultrasound images. This chapter forms the foundation of a journal manuscript currently under submission.

The chapter addresses the critical limitation of data insufficiency, advocating for the integration of synthetic data generation strategies. These strategies aim to expand training datasets and enhance the clinical applicability of P-Net, enabling it to deliver more robust and reliable performance in real-world diagnostic scenarios.

P-Net integrates an ensemble of neural network models—including VGG-19, Residual Networks, and Inception modules—within the U-Net framework. This ensemble approach mitigates the high variance and bias of single-model methods by averaging outputs across models, resulting in improved segmentation accuracy and reliability. P-Net achieves a mean Intersection-over-Union (IoU) of 0.89 and a Dice coefficient of 0.87 on a dataset of 200 images, demonstrating its potential for accurate image analysis despite data constraints.

This chapter advances the use of deep learning in lung ultrasound imaging, showcasing innovative approaches to AI optimisation and data augmentation to tackle the challenges of diagnostics.

1.4.4 Publications

1. Cameron Hill, Abdul H. Sadka., M. Nazmul Huda., "K-Means clustering and rayleigh gaussian mixture Models for enhanced lung ultrasound imaging with SVM-based pleu-

ral line detection" Medical Image Analysis [Editorial revisions]

2. Cameron Hill, Abdul H. Sadka., M. Nazmul Huda., "Overcoming labelled segmentation data scarcity with procedural generation of synthetic lung ultrasound images" IEEE Transactions on Image Processing [Pending submission]

3. Cameron Hill, Abdul H. Sadka., M. Nazmul Huda., "P-Net: Overcoming data limitations in ultrasound imaging with a multi-architecture neural network ensemble" Artificial Intelligence in Medicine [Pending submission]

Chapter 2

Literature review

2.1 Introduction

The integration of computational methodologies into ultrasound imaging has profoundly reshaped the landscape of medical diagnostics, driving unparalleled advancements in diagnostic precision, accessibility, and operational efficiency. As a cornerstone of non-invasive medical imaging, ultrasound has long been heralded for its real-time imaging capabilities, portability, and cost-effectiveness. However, significant limitations persist, including speckle noise, artefacts, and a heavy reliance on operator expertise. Computational innovations have addressed these challenges, expanding the technological boundaries of ultrasound to align with the increasing demands of modern healthcare and research. This literature review critically examines these advancements, focusing on four key areas: Ultrasound Image Enhancement, Ultrasound Feature Detection, Ultrasound Classification, and Artificial Ultrasound Generation.

The domain of Ultrasound Image Enhancement has experienced notable progress through techniques designed to improve diagnostic image quality by mitigating noise and enhancing clarity. Methods such as wavelet transformations, machine learning-based filtering, and hybrid computational algorithms have proven their efficacy in addressing speckle noise and enhancing contrast. These methodologies facilitate the detection of subtle anatomical and pathological details, thereby increasing diagnostic accuracy and enabling more reliable interpretations. Enhanced imaging clarity is also critical for expanding ultrasound's application

in specialised fields requiring high-resolution diagnostic tools.

Ultrasound Feature Detection has emerged as a pivotal area in computational ultrasound, focusing on the identification of diagnostically significant artefacts, including A-lines, B-lines, and pleural line abnormalities. These features are indispensable in diagnosing complex conditions such as pulmonary oedema, interstitial lung disease, and acute respiratory distress syndrome (ARDS). Advanced machine learning architectures, particularly convolutional neural networks (CNNs), have automated these processes with high precision, minimising operator dependency and variability. Additionally, the integration of domain-specific computational insights has further improved the sensitivity and specificity of these models, enabling more nuanced and accurate evaluations.

Ultrasound Classification represents a transformative advancement in the integration of artificial intelligence (AI) into medical imaging. Leveraging the capabilities of deep learning frameworks, researchers have created sophisticated models capable of categorising complex ultrasound datasets into diagnostic categories with exceptional accuracy and speed. These models are particularly beneficial in emergency or resource-limited settings where rapid diagnostic decision-making is crucial. Moreover, contemporary classification systems increasingly incorporate multi-modal data, combining ultrasound imagery with clinical and laboratory findings to enhance diagnostic precision and contextual relevance.

The field of Artificial Ultrasound Generation has addressed one of the most critical challenges in computational ultrasound: the scarcity of annotated datasets necessary for training AI-driven diagnostic models. Generative adversarial networks (GANs) and other synthetic data generation methods have demonstrated remarkable success in producing realistic ultrasound images that mimic the variability of clinical datasets. These synthetic images provide a robust foundation for developing and validating computational models, reducing dependence on extensive manual annotation. Beyond this, artificial image generation enables the simulation of rare or complex pathological scenarios, offering invaluable resources for medical training and AI model testing under diverse conditions.

By synthesising these technological advancements, this review presents a comprehensive evaluation of the current state-of-the-art in computational ultrasound imaging. It identifies emerging trends, such as the integration of machine learning into clinical workflows and the

growing reliance on multi-modal diagnostic frameworks, while also highlighting persistent challenges, including standardisation and the computational demands of advanced methodologies. This synthesis aims to bridge the divide between computational innovation and clinical utility, ensuring that the transformative potential of ultrasound is fully realised in enhancing patient care across diverse healthcare environments.

2.2 Lung ultrasound diagnosis

Lung ultrasound (LUS) has rapidly evolved from a niche imaging tool to a cornerstone of modern pulmonary diagnostics. Originally limited to visualising pleural effusions and pneumothorax, advancements in ultrasound technology and the understanding of artefacts such as A-lines and B-lines have expanded its applications to a broad spectrum of pulmonary conditions. These include interstitial lung diseases, pneumonia, acute respiratory distress syndrome (ARDS), and pulmonary oedema [9]. LUS bridges the gap between traditional imaging modalities like chest radiography (CXR) and computed tomography (CT) by offering a non-ionising, portable, and cost-effective alternative that is particularly suited for point-of-care (POC) applications in emergency and critical care settings.

The rise of LUS is timely, given the increasing demand for diagnostic tools that are both effective and adaptable. Its high sensitivity and specificity in detecting pulmonary abnormalities often rival those of CT, making it a practical choice in diverse clinical contexts. Moreover, LUS has demonstrated significant utility in resource-limited settings, where access to advanced imaging technologies is often constrained. Despite its strengths, LUS is not without limitations, such as operator dependency, the challenge of imaging deep lung structures, and variability in diagnostic protocols [10] [11]. This review explores LUS's diagnostic capabilities, its integration into clinical practice, and its future potential in reshaping pulmonary diagnostics.

2.2.1 Pathologies detected by lung ultrasound

LUS has demonstrated efficacy in diagnosing a wide array of pulmonary and pleural pathologies. Its versatility lies in its ability to detect specific artefacts and visual markers that

correlate with underlying conditions. Key pathologies include:

1. **Pleural Effusions** LUS excels in identifying pleural effusions, outperforming CXR in sensitivity and specificity. It not only detects fluid but also facilitates its quantification and localisation, enabling both diagnostic and therapeutic interventions [9][12].

2. **Pneumothorax** Pneumothorax diagnosis is enhanced by LUS through the identification of absent lung sliding, the stratosphere sign, and the lung point sign, the latter being highly specific for this condition [13].

3. **Interstitial Lung Diseases (ILDs)** B-lines serve as a hallmark for interstitial syndromes, aiding in the diagnosis of conditions like pulmonary oedema, ARDS, and connective tissue disease-associated ILDs [9][11]. In systemic sclerosis and rheumatoid arthritis, LUS demonstrates strong correlations with pulmonary function and high-resolution CT findings [11].

4. **Pneumonia** LUS detects features such as subpleural consolidations, hepatization, and dynamic air bronchograms, which are highly indicative of pneumonia [14][15]. During the COVID-19 pandemic, LUS proved crucial in diagnosing and monitoring viral pneumonia [16][17].

5. **Pulmonary Oedema** LUS provides rapid bedside differentiation between cardiogenic pulmonary oedema and other causes of dyspnoea, leveraging artefacts like the comet-tail sign [18][19].

6. **Alveolar Consolidations** LUS identifies alveolar consolidations with high sensitivity and specificity through visualisation of tissue-like patterns and air bronchograms. This is especially useful in ARDS and infectious conditions [20].

7. **Chronic Obstructive Pulmonary Disease (COPD)** While not a primary imaging tool for COPD, LUS aids in distinguishing COPD exacerbations from overlapping conditions like pulmonary oedema through careful artefact analysis [18].

These diverse pathologies underscore the diagnostic breadth of LUS, cementing its role as a frontline imaging modality in respiratory medicine.

2.2.2 Diagnostic applications

LUS has demonstrated exceptional utility across a spectrum of conditions, from pleural effusions to complex interstitial syndromes. Pleural effusions are among its most well-established applications, with LUS outperforming CXR in sensitivity and specificity by precisely quantifying effusion volume and localisation [9]. In pneumonia, meta-analyses have confirmed LUS's diagnostic accuracy, with sensitivity and specificity exceeding 90%, making it a viable alternative to CXR and CT in both adult and paediatric populations [21][22][15]. During the COVID-19 pandemic, LUS's role in diagnosing interstitial pneumonia, characterised by hallmark features like bilateral B-lines and subpleural consolidations, underscored its adaptability in emergent scenarios [16][17][23].

Artefacts like the comet-tail have been pivotal in differentiating pulmonary oedema from COPD exacerbations, achieving diagnostic accuracy through non-invasive bedside assessment [18][19]. Similarly, the lung point sign offers unparalleled specificity in diagnosing pneumothorax, enabling precise detection and guiding interventions such as chest tube placement [13].

The development of structured protocols has significantly enhanced the clinical utility of LUS. The BLUE-protocol (Bedside Lung Ultrasound in Emergency) provides a systematic approach to diagnosing acute respiratory failure through artefact interpretation and anatomical mapping, with applications extending to critical conditions like ARDS, pulmonary oedema, and pneumothorax [10]. The FALLS-protocol further integrates LUS into haemodynamic assessment, guiding fluid management in shock states [10].

Standardisation remains a critical frontier for LUS. International guidelines have called for harmonised imaging protocols, encompassing artefact scoring systems and defined chest zones for examination [24]. For example, simplified protocols for B-line quantification have shown promise in reducing operator variability while maintaining diagnostic accuracy, particularly in resource-limited settings [15][11]. Emerging AI technologies are poised to augment these efforts by automating artefact recognition and standardising imaging quality [24].

2.2.3 Integration of lung ultrasound in clinical practice

LUS has revolutionised diagnostic approaches in critical and emergency care settings, offering several advantages over conventional imaging techniques. Unlike CT and CXR, LUS is non-ionising, portable, and allows real-time visualisation of dynamic processes, such as lung sliding or pleural effusion movement. This makes it ideal for repeated assessments, particularly in paediatric and pregnant populations where minimising radiation exposure is crucial [15][25].

During the COVID-19 pandemic, LUS emerged as an essential diagnostic tool, enabling the early detection of interstitial pneumonia and facilitating patient triage and resource allocation in overwhelmed healthcare systems [16]. Its portability and immediacy are especially valuable in resource-limited settings, where access to advanced imaging modalities is restricted [23].

Despite its advantages, LUS faces several challenges. Operator dependency remains a significant barrier, with diagnostic accuracy closely tied to the clinician's expertise. This necessitates robust training programmes and the development of standardised scanning protocols to mitigate variability in diagnostic outcomes [24]. Additionally, the reliance on artefact-based imaging introduces interpretive complexities, as artefacts like B-lines can arise from multiple aetiologies, including pulmonary fibrosis, fluid overload, and interstitial syndromes [11].

The limited ability of LUS to visualise deep or central lung structures further restricts its diagnostic scope, necessitating complementary imaging techniques like CT in certain cases [9][20]. Protocol heterogeneity across studies and clinical practices also complicates standardisation efforts, underscoring the need for unified guidelines [24].

2.2.4 Conclusions

Lung ultrasound represents a paradigm shift in pulmonary diagnostics, offering a combination of portability, accuracy, and cost-effectiveness that aligns with the growing emphasis on patient-centred care. Its transformative potential is evident in diverse applications, from diagnosing pneumonia and ARDS to monitoring pulmonary congestion in heart failure pa-

tients. However, to fully realise its potential, several challenges must be addressed.

The operator-dependent nature of LUS highlights the need for widespread training initiatives and technological innovations. The integration of artificial intelligence (AI) into LUS workflows could standardise artefact interpretation, reducing variability and enhancing diagnostic consistency. AI-driven tools for automated B-line quantification and pattern recognition are particularly promising for addressing operator variability and expanding LUS's accessibility [24].

Standardisation of protocols and scoring systems remains a critical area for development. Unified guidelines would not only improve reproducibility but also facilitate broader adoption of LUS in diverse clinical settings, including resource-limited environments and outpatient care. Research efforts should prioritise validating LUS's utility in underrepresented populations, such as pregnant women and patients with complex co-morbidities [21][15].

Emerging technologies like contrast-enhanced ultrasound and elastosonography also hold promise for enhancing LUS's diagnostic accuracy, particularly in distinguishing between benign and malignant consolidations. Longitudinal studies investigating the role of LUS in monitoring treatment efficacy and disease progression will further solidify its position as a cornerstone of respiratory care.

In conclusion, LUS is poised to redefine pulmonary diagnostics through its blend of innovation and practicality. By addressing current limitations and leveraging technological advancements, LUS has the potential to become a universally accessible diagnostic tool, shaping the future of respiratory medicine in both high- and low-resource settings. As research continues to refine its applications, LUS is set to play a pivotal role in improving patient outcomes and advancing global health equity.

2.3 Lung ultrasound data-sets

The integration of machine learning (ML) and artificial intelligence (AI) into lung ultrasound (LUS) applications has been transformative, enabling automated analysis, enhanced diagnostic capabilities, and improved patient outcomes. Central to these advancements is the availability of high-quality datasets that underpin the development and validation of AI

models. This chapter explores the landscape of LUS datasets, focusing on publicly available resources, real-world clinical data, synthetic data, and the associated challenges. It highlights their role in advancing AI research and addresses limitations that impede broader applicability, setting the stage for future developments in the field.

2.3.1 Publicly available datasets

Publicly accessible LUS datasets play a foundational role in facilitating AI-driven research and innovation in diagnostic imaging. Among the most notable datasets is COVIDx-US, which has become a benchmark resource for LUS research. It consists of 173 videos and 21,570 images from 147 patients, covering a spectrum of conditions including COVID-19, non-COVID-19 lung infections, and healthy cases. This dataset not only supports the development of models for diagnostic and prognostic applications but also promotes reproducibility and transparency in AI research [26].

Similarly, the POCUS dataset initially contained 1103 images representing various lung conditions such as bacterial pneumonia and healthy cases. Over time, it was expanded to include 106 videos, providing a richer dataset for developing machine learning algorithms [27]. Another significant contribution is the POCOVID-Net dataset, which includes 202 videos and 59 images across 216 patients, focusing on differentiating COVID-19 from bacterial pneumonia and healthy lungs. The inclusion of metadata and clinical annotations enhances its utility for model validation and interpretability [27].

In addition to these, others have proposed a dataset that supports real-time segmentation of LUS features, including artefacts like A-lines and B-lines, and demonstrates the potential of lightweight models trained on relatively small datasets for clinical translation. Their dataset has been instrumental in advancing explainability and real-time application of LUS analysis [28]. International standardisation initiatives, such as the ICLUS-DB, further emphasize the need for unified frameworks to foster global collaboration and data sharing. This database features over 60,000 annotated frames aimed at standardizing severity classifications for COVID-19-related pneumonia [29].

Despite their utility, these datasets present challenges such as heterogeneity in imaging protocols and variations in data quality. For instance, standardizing data collection across

multiple sources remains a critical hurdle. Nevertheless, these resources form the backbone of AI-driven advancements in LUS diagnostics.

Several publicly available datasets have been developed to facilitate ML applications in LUS diagnostics. Table 2.1 summarises the key datasets.

Table 2.1: Summary of key publicly available LUS datasets

Dataset Name	Description	Key Features	Reference
COVIDx-US	173 videos and 21,570 images from 147 patients	Largest open-access LUS dataset, includes COVID-19, non-COVID infections, and healthy controls	[26]
POCUS	1103 images initially; expanded to include 106 videos	Diverse pathologies including bacterial pneumonia and COVID-19, high-quality annotations for ML applications	[27]
POCOVID-Net	202 videos, 59 images from 216 patients	Focus on COVID-19, bacterial pneumonia, and healthy controls, includes metadata and clinical annotations	[27]
Howell Dataset	500 annotated images, optimised for lightweight model training and explainable AI	Supports real-time segmentation of artefacts and anatomy	[28]
ICLUS-DB	60,000 annotated frames for COVID-19 severity assessment	Standardized scoring for pneumonia severity, aimed at global collaboration and data standardization	[29]

Datasets such as COVIDx-US provide a large and diverse collection of LUS images categorised into COVID-19 and other lung conditions, serving as a robust foundation for AI development. The POCUS dataset offers a broad range of imaging data and metadata, al-

though challenges such as protocol standardisation persist. Smaller datasets like ICLUS-DB are tailored for specific applications, such as severity scoring, and are valuable for prognostic modelling.

2.3.2 Real-world data

Real-world LUS datasets offer unparalleled authenticity and diversity, capturing the nuances and complexities of clinical environments. These datasets often derive from diagnostic procedures performed on patients with varying demographics and disease severities, providing a broad spectrum of cases for training and validating AI models. For example, datasets developed under the BLUE protocol illustrate the diagnostic potential of LUS in identifying pleural effusions, interstitial syndromes, and alveolar consolidations. The BLUE protocol standardizes scanning techniques to enhance reproducibility and accuracy [27].

The international standardisation proposed by Soldati et al. highlights a structured protocol for LUS image acquisition, encompassing 14 specific anatomical regions and scoring systems for severity classification. This initiative seeks to reduce variability in real-world data collection and promote consistency across research and clinical settings [29].

However, real-world data introduces challenges such as variability in operator skill, differences in ultrasound equipment, and inconsistent imaging protocols. Ethical considerations also play a significant role, particularly regarding patient consent and data anonymization. Despite these hurdles, real-world datasets remain indispensable for training robust models capable of generalizing across diverse clinical settings. These datasets also enable the evaluation of AI systems under realistic conditions, ensuring their reliability and safety in real-world applications [30, 27].

2.3.3 Synthetic data

Synthetic data has emerged as a critical resource for augmenting real-world datasets and addressing gaps in data availability. Generated through computational simulations, synthetic datasets replicate specific ultrasound patterns such as A-lines, B-lines, and pleural irregularities. Tools like the Field II ultrasound simulation software allow researchers to create

controlled datasets that mimic real-world conditions while offering the flexibility to generate large volumes of data [31].

The utility of synthetic data lies in its ability to address challenges associated with data scarcity and the high costs of manual annotation. [28] demonstrated the use of lung phantom data to train models for segmenting artefacts and anatomical structures, showcasing the potential of synthetic datasets to bridge gaps in real-world data availability. Their findings underscore the value of combining synthetic and limited real-world data to enhance model performance and explainability [28].

However, the domain gap between synthetic and in vivo data poses challenges for model generalization. Careful validation and domain adaptation techniques are essential to ensure the efficacy of models trained on synthetic datasets. Additionally, synthetic data cannot fully replicate the variability and complexity of human anatomy, limiting its standalone utility [31, 28].

2.3.4 Limitations

While LUS datasets have significantly advanced the field of AI-driven diagnostics, they are not without limitations. Publicly available datasets often lack standardization in imaging protocols, leading to inconsistencies that complicate cross-study comparisons. Real-world datasets face challenges such as operator dependency, equipment variability, and limited patient diversity, which can impact the generalizability of AI models [30, 26].

Synthetic datasets, although valuable for addressing data scarcity, struggle to capture the intricacies of real-world clinical scenarios. Highlighted issues such as the subjective nature of artefact annotations and inter-annotator variability, which can impact the accuracy and reliability of segmentation models [28]. The reliance on simulated patterns may also introduce biases that hinder the model's applicability to in vivo data [31]. Moreover, ethical concerns surrounding data collection, particularly in real-world settings, underscore the need for robust anonymization and consent frameworks [29].

Addressing these limitations requires a collaborative approach to standardize data collection, expand dataset diversity, and integrate real-world and synthetic data effectively. Future efforts should focus on creating comprehensive, interoperable datasets that bridge the gap

between research and clinical practice [27].

2.3.5 Conclusions

The development and dissemination of high-quality LUS datasets are pivotal for advancing AI-driven diagnostics and improving patient outcomes. Publicly available datasets like COVIDx-US, POCUS, and POCOVID-Net have provided a strong foundation for research, while real-world and synthetic data continue to address gaps in representativeness and scalability. The lack of annotated datasets for segmentation tasks remains a significant bottleneck, as segmentation models require detailed pixel-level labels, which are time-consuming and costly to produce [28]. Furthermore, synthetic datasets, while addressing the scarcity of annotated data, often exhibit limited variability and fail to replicate the full complexity of real-world clinical conditions [31, 28].

The international standardisation proposals and innovative synthetic data methodologies, as highlighted by [28, 29], respectively, showcase promising directions for the field. These efforts emphasize the importance of global collaboration and data-sharing frameworks to overcome existing challenges. Despite significant issues in standardization, data quality, and ethical considerations, ongoing initiatives to enhance dataset diversity and interoperability hold promise for unlocking the full potential of LUS in medical research and clinical practice [29].

2.4 Ultrasound image enhancement

Enhancing ultrasound images is critical for improving their diagnostic utility. Due to the inherent issues such as speckle noise, low contrast, and artefacts, achieving high-quality imaging requires sophisticated post-processing techniques. This chapter delves into the methodologies used for ultrasound image enhancement, focusing on noise reduction strategies, contrast improvement techniques, and the metrics used to evaluate image quality. By addressing these challenges, researchers have developed advanced algorithms and systems aimed at providing clearer, more detailed visualizations that aid in medical diagnosis and analysis.

2.4.1 Noise reduction

Noise is a pervasive challenge in ultrasound imaging, primarily caused by the coherent nature of ultrasound waves leading to speckle patterns. Speckle noise degrades image quality and hinders the accurate delineation of anatomical structures. Techniques to reduce noise typically aim to balance the preservation of edges and details while minimizing speckle.

Speckle reducing anisotropic diffusion

Speckle Reducing Anisotropic Diffusion (SRAD) is one of the pioneering methods in ultrasound noise reduction, specifically designed to address the unique challenges posed by speckle noise. Traditional anisotropic diffusion approaches were adapted to incorporate speckle-relevant characteristics through the use of instantaneous coefficient variations. This ensures that diffusion is guided by local signal properties, allowing for selective smoothing of homogeneous regions while retaining edge fidelity [32] [33].

SRAD operates iteratively, applying a diffusion equation that minimizes noise while preserving important anatomical structures. Its mathematical rigor and adaptability make it a robust choice for ultrasound imaging. However, SRAD's iterative nature can lead to significant computational demands. Moreover, excessive iterations risk over-smoothing, causing a loss of fine details that may be critical for diagnostic interpretation. Recent improvements have focused on integrating SRAD with multiscale techniques to enhance its performance while mitigating these limitations [33].

Non-local Means filtering

NL-Means filtering represents a paradigm shift in noise reduction, moving away from traditional local approaches to consider the global image context. By comparing image patches across the entire image, NL-Means identifies and leverages structural similarities to reduce noise. This approach is particularly effective for ultrasound images, where speckle noise often masks subtle anatomical details [34].

The Bayesian variant of NL-Means adapts the filter to ultrasound-specific noise models, incorporating statistical priors to further enhance its robustness. This adaptation signif-

icantly improves the algorithm's ability to preserve edges and textures, making it highly suitable for complex imaging scenarios. Despite its advantages, NL-Means filtering is computationally intensive, requiring substantial memory and processing power. Efforts to optimize its implementation, including GPU acceleration and hierarchical filtering strategies, have shown promise in reducing these computational burdens [35].

Stochastic distance-based BM3D

Building upon the strengths of the original BM3D algorithm [36], Stochastic Distance-Based BM3D (SD-BM3D) introduces a novel approach to patch similarity measurement by utilizing Fisher-Tippett distributions. This statistical adaptation enables the method to handle the multiplicative nature of speckle noise more effectively, resulting in superior despeckling performance [35].

SD-BM3D operates in two main stages: grouping similar image patches into 3D blocks and applying collaborative filtering in a transformed domain. The use of stochastic distances enhances the algorithm's ability to differentiate between noise and structural features, preserving critical image details. However, the complexity of the approach requires careful parameter tuning to balance noise suppression with detail retention. SD-BM3D has demonstrated exceptional results in log-compressed ultrasound images, making it a leading choice for advanced noise reduction applications.

Non-linear coherent diffusion

Nonlinear Coherent Diffusion (NCD) combines the principles of anisotropic diffusion and noise transformation to address the unique challenges of ultrasound imaging. By converting multiplicative speckle noise into additive Gaussian noise in the logarithmic domain, NCD simplifies the noise reduction process [34].

The algorithm relies on local directional variance to guide the diffusion process, ensuring that smoothing is applied selectively to homogeneous regions while preserving edges. This makes NCD particularly effective for enhancing boundaries and maintaining anatomical fidelity. However, the method's reliance on accurate noise modeling can limit its performance in heterogeneous tissues or when the noise characteristics deviate from expected patterns.

Wavelet-based techniques

Wavelet-based techniques leverage the multiscale nature of wavelets to provide effective noise suppression across different spatial frequencies. These methods decompose the image into multiple resolution levels, allowing for targeted noise reduction at each scale. Multiscale wavelet diffusion, which combines wavelet thresholding with nonlinear diffusion, has emerged as a particularly effective approach for simultaneous noise reduction and edge enhancement [37] [38].

Wavelet methods are computationally efficient and well-suited for real-time applications. However, their effectiveness can be diminished in images with highly heterogeneous noise distributions, where the separation of noise and signal components becomes challenging. Hybrid approaches that integrate wavelets with other denoising techniques are being explored to address these limitations and further enhance performance.

Adaptive filtering approaches

Adaptive filtering approaches, such as the Lee, Frost, and Kuan filters, adaptively model speckle noise as a multiplicative process and apply smoothing based on local statistical properties. These filters are highly effective in homogeneous regions, where they can significantly reduce noise without compromising image quality [33].

The adaptive speckle reduction filter builds upon these methods by incorporating classification techniques to identify regions requiring different levels of smoothing. This context-aware filtering ensures that edges and fine details are preserved, making it particularly useful for applications where structural integrity is critical. Despite their effectiveness, adaptive filters may introduce blurring at high-contrast boundaries, necessitating further refinements to improve edge preservation [34].

Hybrid compounding techniques

Hybrid compounding techniques integrate multiple imaging perspectives, such as spatial and frequency compounding, to achieve superior noise suppression. By combining data from different acquisition angles or frequency bands, these methods reduce noise while preserving

anatomical features [39] [40].

These techniques are particularly effective in clinical settings, where they enhance image clarity and diagnostic accuracy. However, their reliance on multiple data acquisitions can increase imaging time and computational requirements. Advances in real-time compounding algorithms are addressing these challenges, making hybrid techniques more accessible for routine clinical use.

2.4.2 Contrast enhancement

Contrast enhancement is vital to highlight anatomical boundaries and regions of interest in ultrasound images. Poor contrast often arises due to weak backscatter signals and shadowing artefacts. A wide range of techniques has been developed to address these challenges, each catering to specific use cases and imaging environments.

Contrast-limited adaptive histogram equalization

Contrast-Limited Adaptive Histogram Equalization (CLAHE) is a widely adopted technique for enhancing the contrast of ultrasound images. Unlike traditional histogram equalization, CLAHE limits the amplification of noise by constraining the contrast enhancement in homogeneous regions. This method divides the image into smaller contextual regions (tiles) and applies histogram equalization locally. By interpolating the results of neighbouring tiles, CLAHE ensures a seamless transition across the image [41] [42].

While CLAHE is highly effective in enhancing contrast, it can sometimes lead to over-enhancement in regions with high-intensity variations, resulting in unnatural visual artefacts. Advanced implementations of CLAHE integrate parameter optimization techniques to dynamically adjust the clip limits, improving its applicability to diverse imaging scenarios. In clinical settings, CLAHE has proven especially useful for improving the visibility of soft tissue structures and enhancing diagnostic accuracy.

Fuzzy set theory and phase asymmetry metrics

Techniques integrating fuzzy set theory and phase asymmetry metrics represent a sophisticated approach to contrast enhancement. These methods model the image as a collection of fuzzy sets, where each pixel's membership degree reflects its intensity and neighbourhood characteristics. By leveraging fuzzy rules, these techniques dynamically adjust pixel intensities to enhance contrast without amplifying noise [43] [39].

Phase asymmetry metrics further refine this process by identifying and emphasizing features with asymmetric phase responses, which are often indicative of edges and boundaries. These techniques are particularly effective in enhancing complex anatomical structures, such as the carotid artery. Their computational complexity, however, necessitates efficient algorithms to make them viable for real-time applications.

Maximum local variation-based unsharp masking

Maximum Local Variation-Based Unsharp Masking (MLVUM) is an advanced technique that targets localized enhancement. By focusing on areas with maximum intensity variation, MLVUM selectively sharpens edges and enhances contrast without significantly amplifying speckle noise [44]. This method combines traditional unsharp masking with adaptive weighting to balance enhancement and noise suppression.

MLVUM is particularly beneficial for applications requiring precise boundary delineation, such as lesion detection in oncological imaging. However, the method's effectiveness heavily depends on parameter selection, and inappropriate configurations can lead to overshooting artefacts or inadequate enhancement. Recent advancements include machine learning-driven parameter tuning, which enhances the robustness and adaptability of MLVUM. [44]

Hybrid strategies

Hybrid strategies offer a versatile solution by combining multiple enhancement techniques to leverage their respective strengths. For instance, CLAHE can be integrated with nonlinear transformations to improve local contrast while maintaining global intensity balance. Similarly, segmentation-based approaches can be employed to isolate regions of interest, allowing

targeted enhancement [38] [40].

Lesion-specific enhancement strategies exemplify the power of hybrid approaches. These methods combine texture analysis, entropy-driven transformations, and adaptive weighting to amplify contrast within clinically significant areas, such as tumours or vascular anomalies. While these techniques yield superior results, their computational demands and reliance on accurate segmentation algorithms remain challenging for real-time implementation.

Emerging techniques

Emerging techniques in contrast enhancement increasingly rely on machine learning and deep learning frameworks. Convolutional neural networks (CNNs) have demonstrated remarkable capabilities in learning complex patterns and features, enabling automated and highly effective enhancement processes. These models are trained on large datasets to optimize contrast enhancement while minimizing noise amplification, making them promising tools for future ultrasound imaging[45][40].

2.4.3 Quality metrics

Evaluating the effectiveness of image enhancement techniques is critical to ensure they meet clinical and diagnostic standards. Various quality metrics have been proposed and developed to quantify improvements in image contrast, noise reduction, structural preservation, and overall clarity. These metrics serve as objective benchmarks, enabling researchers and clinicians to compare and optimize enhancement methods.

Contrast-to-Noise Ratio (CNR) quantifies the distinguishability of anatomical structures from the background, offering a straightforward measure of contrast enhancement. CNR is especially useful for methods like CLAHE and wavelet-based techniques. However, its inability to account for structural fidelity makes it an incomplete standalone metric [36] [46].

Structural Similarity Index (SSIM) emphasizes perceptual quality by comparing luminance, contrast, and structural components of images. Widely applied to assess algorithms like NL-Means and SD-BM3D, SSIM is particularly effective for clinical ultrasound imaging. Nevertheless, it depends heavily on the availability of high-quality reference images, limiting

its scope in certain scenarios[46].

Entropy-based measures evaluate texture richness and information content, making them ideal for assessing hybrid strategies and fuzzy set-based techniques. While high entropy signifies enhanced texture, it may also indicate increased noise, necessitating careful interpretation in conjunction with other metrics [45].

The Image Enhancement Metric (IEM) provides a composite assessment of contrast and sharpness, offering insights into the holistic effectiveness of enhancement methods. IEM excels in evaluating hybrid approaches, though its simplicity may overlook nuanced trade-offs between noise and detail preservation[45].

Speckle Suppression Index (SSI) and Speckle Suppression and Mean Preservation Index (SMPI) focus on ultrasound-specific challenges. By balancing speckle noise reduction with detail retention, these indices are crucial for methods like SRAD and SD-BM3D. However, their reliance on homogeneous regions can limit their applicability to heterogeneous tissues[35][33].

The Ultrasound Despeckling Assessment Index (USDSAI) evaluates the ability of despeckling methods to maintain inter-class contrast and intra-class homogeneity, ensuring comprehensive performance evaluation. Iterative filters such as Squeeze Box Filters and SRAD often benefit from this metric[33].

Emerging machine learning-driven metrics analyse high-dimensional feature spaces, enabling adaptive and automated evaluations. These metrics provide granular insights into enhancement performance, showing promise for future applications in diverse imaging scenarios[45].

In table 2.2, a list of metrics are summarised

2.4.4 Conclusions

The advancements in ultrasound image enhancement methodologies represent a significant leap forward in the domain of medical imaging. However, these improvements are not without limitations. Methods such as SRAD and NL-Means filtering provide excellent noise reduction capabilities but are computationally intensive, limiting their real-time applicability. Similarly, contrast enhancement techniques like CLAHE and MLVUM offer targeted improvements but can result in artefacts if not carefully parametrised. Addressing these

Table 2.2: Summary of Quality Metrics for Evaluating Ultrasound Image Enhancement Methods

Metric	**Description and Usage**
Contrast-to-Noise Ratio (CNR)	Quantifies contrast enhancement relative to background noise; widely used for evaluating CLAHE and wavelet-based techniques.
Structural Similarity Index (SSIM)	Measures luminance, contrast, and structural similarity; useful for NL-Means and SD-BM3D.
Entropy-Based Measures	Assesses texture richness; applied in hybrid and fuzzy set-based methods.
Image Enhancement Metric (IEM)	Evaluates contrast and sharpness; suited for hybrid approaches.
Speckle Suppression Index (SSI)	Quantifies speckle reduction; critical for SRAD and wavelet-based diffusion.
Speckle Suppression and Mean Preservation Index (SMPI)	Balances noise reduction with detail preservation; complements SSI.
Ultrasound Despeckling Assessment Index (USDSAI)	Measures inter-class contrast and intra-class homogeneity; ideal for iterative despeckling methods.

challenges requires a dual focus on optimization and scalability.

One key limitation across many existing methods is their reliance on predefined parameters and assumptions about noise and contrast characteristics. These assumptions may not hold across all imaging contexts, particularly in heterogeneous tissues or for images acquired under suboptimal conditions. Adaptive and learning-based approaches, capable of dynamically adjusting to specific imaging scenarios, offer a promising avenue for overcoming these limitations.

The integration of machine learning, particularly deep learning, into image enhancement

workflows has the potential to transform the field. Convolutional neural networks (CNNs) and generative adversarial networks (GANs) are already demonstrating their capability to enhance image quality while preserving diagnostic features. These models, trained on extensive datasets, can generalize well to diverse imaging conditions, offering both precision and robustness. However, the success of these methods hinges on the availability of high-quality training data and the development of explainable AI frameworks to ensure clinical trust and acceptance.

Another area ripe for future exploration is the development of real-time enhancement algorithms. Many current methods, while highly effective, require significant computational resources, making them unsuitable for immediate clinical application. Advances in hardware acceleration, such as GPU and FPGA technologies, alongside algorithmic optimizations, can bridge this gap, enabling real-time noise reduction and contrast enhancement during ultrasound examinations.

Metrics for evaluating image enhancement also require further refinement. While existing metrics such as SSIM, CNR, and entropy provide valuable insights, they often fail to capture the nuanced trade-offs between noise suppression, contrast enhancement, and detail preservation. Emerging metrics that leverage machine learning to analyze high-dimensional feature spaces offer a more comprehensive evaluation framework, facilitating the comparison and optimization of advanced enhancement techniques.

The impact of these enhancement techniques extends beyond improved image quality. By enabling better visualization of anatomical structures, these methods enhance diagnostic accuracy and reduce the risk of misinterpretation. In the context of automated analysis, improved image quality directly translates to better performance of downstream algorithms, such as segmentation and classification models, further solidifying their role in modern medical imaging workflows.

In conclusion, while ultrasound image enhancement techniques have made remarkable strides, there remains substantial room for innovation. Future research must focus on developing adaptive, efficient, and clinically interpretable methods that cater to the dynamic needs of medical imaging. The integration of advanced computational techniques, real-time processing capabilities, and comprehensive evaluation frameworks will pave the way for a

new era of diagnostic excellence, reaffirming ultrasound's position as a cornerstone of modern medicine.

2.5 Ultrasound feature detection

Recently machine learning tools have taken over full clinical tasks within the medical setting as advancements in AI mature with the wider range of data being accessible. It is more important than ever now that medical staff are educated on the potential of the role of artificial intelligence and its place in the clinical environment. With advancements in computational and machine learning techniques, the field of LUS has seen remarkable progress in the detection and interpretation of key features such as A-lines, B-lines, and pleural line characteristics. These features are critical markers for diagnosing conditions like pulmonary edema, pneumothorax, and more recently, assessing the severity of COVID-19-related pneumonia. This chapter provides a comprehensive review of the latest methodologies in LUS feature detection, focusing on artifact quantification, computational advancements, and machine learning innovations. By integrating these methodologies, the diagnostic utility of LUS can be significantly enhanced, paving the way for more accurate and efficient patient care.

2.5.1 Quantification of artifacts

Artifacts in lung ultrasound (LUS), such as A-lines and B-lines, are critical diagnostic markers. A-lines are horizontal, repetitive lines indicative of healthy lung tissue, while B-lines are vertical artifacts often associated with interstitial syndrome or fluid overload [47] [48]. The accurate identification and quantification of these artifacts are essential for diagnosing conditions such as pulmonary edema and pneumothorax. Automated methods, including Radon transforms and matched filtering, have enhanced the reliability of artifact detection[48]. Despite advancements, challenges remain, particularly in differentiating between artifact types under varying clinical conditions.

B-lines, identified as hyperechoic vertical reverberation artifacts extending from the pleura to the edge of the screen, are pivotal in assessing pulmonary edema. Accurate detection of B-lines is crucial for diagnosing various pulmonary conditions, including heart

and lung diseases [49]. The proposed methods, such as using Gaussian model fitting and alternate sequential filtering (ASF), have significantly advanced the identification and separation of B-lines from other vertical artifacts. These automated systems help quantify the severity of conditions like interstitial syndrome or pulmonary edema by analyzing the spatial distribution of B-lines across multiple LUS frames[49].

2.5.2 Computational approaches

Computational techniques play a pivotal role in LUS image analysis. Traditional methods, such as model-based despeckling and deconvolution, aim to enhance image quality and feature visibility[49]7. Relevant references should be provided in statements and highlighted in the annotated thesis.[47]. The Radon transform, for instance, has been instrumental in detecting and delineating line artifacts in noisy LUS images[47]. Additionally, hybrid approaches that combine regularized optimization and data-driven elements are emerging as robust solutions for complex imaging challenges.

Recent computational advancements in LUS include the application of graph-based delineation methods and adaptive histogram equalization to enhance artifact visibility[49]. A novel pipeline involving pleural line delineation followed by top-hat filtering has refined the accuracy of detecting relevant structures like B-lines. These approaches leverage domain-specific knowledge, such as the Beer–Lambert law for signal attenuation modeling, to improve artifact segmentation in noisy environments[49].

A novel approach leveraging nonconvex regularization has shown promise in quantifying line artifacts, specifically in COVID-19 patient data. The method, utilizing a Cauchy-based penalty and Radon transform, achieves 87% B-line detection accuracy with enhanced precision through a validation mechanism [50]. This framework addresses challenges in noisy and artifact-laden ultrasound images, offering robust identification and aiding in clinical evaluations of pulmonary conditions [50].

The introduction of optimized Gaussian Histogram Equalization (GHE) combined with Neoteric Segmentation (NS) has significantly advanced segmentation capabilities in LUS imaging. The GHE technique enhances contrast and preserves visual quality, while the NS approach achieves segmentation accuracy as high as 95% [51]. These methods address key

challenges like noise and low contrast, proving instrumental for improving the diagnostic utility of LUS images [51].

Quantitative analysis of pleural line (PL) and B-lines (BLs) further contributes to the assessment of pulmonary conditions, particularly in the context of COVID-19 pneumonia. Features such as PL thickness (TPL), roughness (RPL), and intensity metrics, alongside BL characteristics like accumulated width (AWBL) and attenuation coefficient (ACBL), are indicative of disease severity. SVM-based classification integrating these features has demonstrated optimal diagnostic performance, highlighting the role of detailed computational analysis in severity assessment [52].

A recent method for automatic pleural line extraction utilizes Hidden Markov Models (HMM) and the Viterbi algorithm to detect pleural line characteristics and their alterations due to pathological conditions such as COVID-19. The integration of geometric and radiometric features with Support Vector Machine (SVM) classifiers for severity scoring has shown high accuracy, enhancing the reliability of automated LUS assessments [53].

2.5.3 Machine learning approaches

Machine learning (ML) has revolutionized LUS feature detection, particularly through deep learning frameworks like Faster R-CNN and RetinaNet [54] [48]. These models excel in identifying pleural lines, A-lines, and B-lines with high accuracy. For example, automated systems leveraging convolutional neural networks have demonstrated superior performance in neonatal and COVID-19 diagnostics [54] [48]. Challenges such as dataset heterogeneity and interpretability, however, necessitate ongoing refinement and standardization efforts.

Building on computational techniques, machine learning systems have become integral to automating LUS analysis. The application of Hough transform and convolutional neural networks enables accurate detection and classification of artifacts like B-lines[49]. Furthermore, methods that integrate Gaussian fitting and random walk frameworks for pleural line localization have demonstrated superior performance in distinguishing pathological from healthy LUS patterns. These advancements offer promising avenues for real-time, quantitative clinical diagnostics[49].

The integration of weakly supervised deep learning has further enhanced the detection

and localization of B-lines. Using techniques like Class Activation Mapping (CAM), neural networks can provide real-time identification of B-line artifacts at an inference rate of 276 frames per second [55]. This approach leverages gradient-weighted mapping to visually interpret network outputs, allowing for localization without requiring extensive annotation. Such advancements enable more efficient clinical workflows and pave the way for developing phenotype-specific analysis tools [55].

Recent advancements in pleural effusion segmentation have further showcased the potential of machine learning in enhancing diagnostic workflows. Deep learning-based approaches, such as convolutional neural networks combined with Spatial Transformer Networks, effectively localize and segment pleural effusions in lung ultrasound images. These techniques have achieved Dice Similarity Coefficients comparable to expert observers, emphasizing their reliability [56]. The development of such systems has the potential to improve procedural safety during interventions like thoracentesis and provide clinicians with actionable insights on effusion boundaries [56].

A novel framework leveraging Spatial Transformer Networks (STN) combined with ordinal regression has been proposed for frame-based classification and localization of COVID-19 markers in lung ultrasound data. This approach integrates weakly supervised localization with disease severity scoring, using ensemble methods for pixel-level segmentation of COVID-19-related artifacts [57].

2.5.4 Conclusions

The advancements in lung ultrasound feature detection presented in this chapter underscore the transformative impact of computational and machine learning methodologies. Quantifying artifacts such as A-lines and B-lines has enhanced diagnostic precision, while computational techniques like the Radon transform and advanced segmentation pipelines have improved image analysis. The integration of machine learning, especially deep learning, has further streamlined and automated the diagnostic process, facilitating real-time clinical decision-making.

One of the significant achievements is the use of weakly supervised deep learning techniques, such as Spatial Transformer Networks, which have enabled detailed localization of

pathological features without requiring extensive labeled data. Similarly, advanced frameworks for pleural line detection, leveraging Hidden Markov Models and Support Vector Machines, have demonstrated efficacy in assessing COVID-19 severity. These methods underline the importance of combining geometric and radiometric feature extraction with robust classifiers to improve diagnostic accuracy.

Despite these advancements, limitations persist. The reliance on annotated datasets remains a challenge, as generating high-quality labeled data in clinical settings is time-intensive and resource-heavy. Moreover, current models often lack generalizability across different patient demographics and ultrasound devices, necessitating efforts to standardize data acquisition protocols and expand diverse datasets. The noise inherent in ultrasound imaging further complicates artifact detection, highlighting the need for enhanced noise reduction techniques.

Future research directions should focus on integrating temporal analysis across video sequences to better understand dynamic pulmonary changes. The inclusion of real-time predictive models within ultrasound devices could revolutionize point-of-care diagnostics. Additionally, developing cross-modal learning approaches that incorporate complementary imaging modalities, such as CT or MRI, may enhance the robustness of LUS-based diagnostics. Addressing interpretability issues in machine learning models is another crucial avenue to ensure clinical adoption and trust.

In conclusion, the synergy between computational techniques and machine learning has propelled lung ultrasound imaging into a new era of diagnostic capability. By addressing current limitations and embracing innovative research directions, LUS stands poised to become an indispensable tool in pulmonary healthcare, offering enhanced diagnostic accuracy, improved patient outcomes, and reduced healthcare costs.

2.6 Ultrasound classification

Lung ultrasound (LUS) automatic classification has emerged as a pivotal area of research, leveraging artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), to enhance the interpretation of LUS data. This focus encompasses a diverse array of

methodologies aimed at extracting, analysing, and classifying complex patterns within LUS images and videos. The advent of these technologies has opened new frontiers for diagnosing and monitoring a variety of lung pathologies, including pneumonia, COVID-19, and chronic respiratory conditions.

This section delves into the state-of-the-art in LUS automatic classification, exploring its computational frameworks, performance metrics, and integration into clinical workflows. Emphasis is placed on the evolution of deep learning models such as convolutional neural networks (CNNs) and hybrid architectures that incorporate temporal and spatial dynamics. The discussion extends to the challenges and opportunities associated with dataset standardisation, model interpretability, and deployment in diverse healthcare environments. By situating LUS automatic classification within the broader context of AI-driven diagnostics, the chapter aims to provide a comprehensive understanding of its transformative potential and future trajectories.

2.6.1 Deep learning classification

The integration of ML and DL into LUS classification workflows represents a paradigm shift in diagnostic imaging. Convolutional neural networks (CNNs) dominate this landscape, offering unparalleled performance in spatial feature extraction. These approaches provide a robust framework for decoding the complexities of lung ultrasound data, transcending traditional diagnostic methodologies. For instance, CNN-based approaches have demonstrated high sensitivity (up to 90%) and specificity (up to 96%) in distinguishing COVID-19 from bacterial and viral pneumonia as well as healthy cases [58].

Recent studies have extended the capabilities of CNNs by incorporating domain-specific optimisations. For example, a hybrid CNN-LSTM model effectively captured spatial and temporal dynamics in LUS videos, improving classification performance by 14–21% compared to standalone CNNs [59]. Meanwhile, architectures such as EfficientNet and ResNet are optimised for balancing computational efficiency with accuracy, achieving state-of-the-art results in multi-class classification scenarios [60]. Similarly, a novel ensembling technique integrating multiple EfficientNet-b0 models achieved a 100% classification accuracy on one of the largest publicly available datasets, underscoring the promise of adaptive ensembling

for LUS diagnostics [60].

An innovative approach involved a fuzzy pooling-based CNN (FP-CNN), which enhanced feature extraction by integrating fuzzy membership functions, achieving an impressive accuracy of 97.2% while improving interpretability through SHAP-based explainability [61]. In paediatric applications, neural networks were tailored to detect pneumonia-related pulmonary infiltrates, achieving sensitivities of 90.9% and specificities of 100%, showcasing the versatility of AI in addressing demographic-specific challenges [62].

Traditional ML models, such as support vector machines (SVMs), have retained relevance in limited-data scenarios, achieving accuracy rates of up to 94% in identifying distinctive LUS features [63]. Despite these achievements, DL methods generally dominate due to their ability to harness large, high-dimensional datasets, a critical advantage in LUS diagnostics.

2.6.2 Temporal and spatial feature models

LUS imaging inherently captures dynamic information, particularly in video formats. Capturing the interplay between spatial and temporal features is indispensable for diagnosing complex and evolving conditions, such as acute respiratory distress syndrome (ARDS) and COVID-19. Hybrid architectures combining CNNs with recurrent neural networks (RNNs) or Long Short-Term Memory (LSTM) layers have emerged as powerful tools for temporal feature analysis.

One exemplary study employed a CNN-LSTM model to predict COVID-19 severity, achieving an accuracy of 79.1% with linear probes and 67.7% with convex probes [59]. Another framework utilised modality alignment contrastive learning (MA-CLR) to integrate LUS imaging and clinical data for severity assessment, demonstrating an accuracy of 87.5% for binary classifications and offering robust alignment across modalities [64].

Temporal-spatial integration is further augmented by advanced preprocessing techniques. For instance, transfer learning methods have enabled robust COVID-19 classification by leveraging pre-trained VGG19 models across multiple imaging modalities, including LUS, with ultrasound emerging as the most effective diagnostic tool [65].

Another innovative method involved multi-layer fusion approaches, where image-level and feature-level fusion significantly enhanced COVID-19 and non-COVID-19 classification

accuracy using LUS images, achieving accuracy rates exceeding 93% [66].

Lightweight architectures tailored for LUS analysis, such as Mini-COVIDNet, have shown remarkable performance with minimal computational requirements, achieving 83.2% accuracy while facilitating real-time diagnostics in resource-constrained settings [67]. These strategies underscore the importance of model optimisation for clinical deployment.

2.6.3 Explainability and interpretable models

In medical imaging, the interpretability of AI models is as critical as their predictive accuracy. Clinicians require assurance that AI systems align with clinical reasoning, particularly when decisions carry life-altering consequences. Explainable AI (XAI) methods, such as Gradient-weighted Class Activation Mapping (Grad-CAM) and Shapley Additive Explanations (SHAP), are now indispensable in LUS classification. These tools highlight regions of diagnostic relevance within LUS images, offering clarity on how predictions are generated.

For example, Grad-CAM has been used to visualise pleural line irregularities and subpleural consolidations—hallmark features of conditions such as COVID-19 and pneumonia [68] [69]. Advanced models, including FP-CNN, leveraged fuzzy pooling to further improve interpretability by integrating saliency maps into decision-making processes [61].

Moreover, a novel CNN approach tailored to distinguish pathological B-lines on LUS images enhanced clinical decision-making, achieving an area under the curve (AUC) of 0.978 while providing interpretable visualisations that differentiated similar imaging pathologies [12]. This highlights the potential of XAI in distinguishing among overlapping conditions.

Furthermore, the integration of domain-specific scoring systems, such as B-line quantification and severity assessment frameworks, has enhanced the clinical utility of AI. Models like those optimised for paediatric pneumonia diagnosis incorporated explainability tools to validate predictions, offering actionable insights for healthcare providers [62].

2.6.4 Resource-limited models

LUS classification technologies are particularly valuable in resource-constrained environments, where advanced imaging modalities like CT scans are unavailable. Lightweight

solutions such as Mini-COVIDNet have bridged this gap, demonstrating the feasibility of deploying DL models on portable devices for bedside diagnostics [67].

AI-driven LUS tools have also proven indispensable during global health crises. For instance, the integration of adaptive ensembling approaches facilitated rapid and accurate COVID-19 diagnosis, streamlining triage processes and improving patient management in overwhelmed healthcare systems [60].

In addition, deep learning models have been adapted for real-time COVID-19 detection in resource-constrained settings, achieving robust results with reduced imaging data, further proving their scalability and effectiveness [70].

In chronic respiratory conditions, AI systems have enabled longitudinal monitoring by analysing LUS data over time, detecting subtle changes in disease progression, and providing timely treatment recommendations. Efforts to incorporate standardised scoring systems have further enhanced these applications, as demonstrated in frameworks integrating clinical and LUS data for comprehensive severity assessments [64].

2.6.5 Conclusions

Despite significant progress in LUS classification, several challenges persist that must be addressed to unlock its full potential. Dataset variability remains a critical barrier, as imaging data often originates from different institutions and employs diverse protocols. This heterogeneity can introduce biases and limit the generalisability of AI models [71] [72]. Additionally, many studies rely on small datasets, leading to overfitting and reducing the efficacy of trained models in broader clinical contexts. The operator dependency of LUS further complicates standardisation efforts, with imaging quality heavily influenced by probe handling, angles, and patient positioning [71] [67]. Addressing these issues requires the establishment of large, diverse, and standardised datasets, facilitated by multicentre collaborations and open-access initiatives [71] [63].

Future research should strategically focus on integrating LUS imaging with multimodal data, encompassing patient history, laboratory results, and even genetic profiles, to construct a more comprehensive and context-aware diagnostic framework [64]. Such integration not only enhances diagnostic accuracy but also enables the development of personalised assess-

ments that can dynamically adapt to the nuances of individual patient cases. Furthermore, the refinement of lightweight, real-time models specifically tailored for point-of-care applications holds immense potential to transform LUS into a practical solution for emergency settings and resource-limited environments [67].

In conclusion, lung ultrasound classification, augmented by AI, represents a transformative advancement in diagnostic imaging. By addressing current limitations and embracing emerging opportunities, this technology can revolutionise respiratory care across diverse clinical settings. The continued evolution and integration of LUS into routine practice will enhance diagnostic accuracy and ensure equitable access to life-saving healthcare innovations.

2.7 Artificial ultrasound generation

The field of artificial ultrasound generation has experienced remarkable advancements, with extensive research leveraging machine learning frameworks and generative adversarial networks (GANs) to address inherent challenges in medical imaging. This review consolidates and critically evaluates key findings from recent studies, emphasising methodological innovations, identified limitations, and prospective avenues for development. These innovations herald a significant shift in medical imaging, particularly in diagnostic accuracy, training efficacy, and data augmentation capabilities, underscoring the critical role of synthetic imaging in modern healthcare paradigms.

2.7.1 Generative methodologies and applications

Generative adversarial networks (GANs) have solidified their position as a foundational tool in synthetic ultrasound imaging, offering sophisticated solutions to augment datasets, advance diagnostic algorithms, and enhance clinical training modules. Among the notable approaches, Deep Convolutional GANs (DCGANs) have exhibited exceptional capacity in generating highly realistic and structurally diverse synthetic ultrasound images, achieving structural similarity indices (SSIM) as high as 0.99 in some studies [73]. These capabilities address the dual challenges of annotated dataset scarcity and the ethical considerations

associated with real patient data usage. Auxiliary Classifier GANs (ACGANs), with their ability to produce diagnostically relevant synthetic data using minimal datasets, have further illustrated the versatility of GAN-based approaches in reducing dependency on manual annotations while maintaining clinical relevance [74].

Frameworks such as sketch-guided GANs and progressive growing GANs have pushed the boundaries of synthetic imaging by enabling high-resolution image generation tailored to specific clinical needs. These models incorporate auxiliary edge sketches and segmentation label maps to produce customisable and highly realistic ultrasound images [75]. Such advances allow researchers and clinicians to simulate rare pathological conditions, augment training datasets, and evaluate diagnostic models under meticulously controlled scenarios. Additionally, StyleGAN3 has demonstrated its utility in addressing class imbalances, enhancing feature diversity, and improving the diagnostic performance of downstream machine learning applications [76]. Beyond general imaging, DCGANs have been employed for synthesising breast ultrasound images [77] and segmenting COVID-19-related lung ultrasound features [78], underscoring their versatility across various medical contexts.

Complementary advancements include the utilisation of auto-encoding structures for shadow estimation in ultrasound images [79], enabling accurate reconstruction of datasets with complex artefacts, and cross-modality synthesis techniques for generating images in different diagnostic formats [80]. Such methodologies expand the applicability of GANs beyond traditional imaging constraints, facilitating more holistic and adaptive approaches to medical imaging research. Additionally, adaptive augmentation techniques like Adaptive Discriminator Augmentation (ADA) and Differentiable Augmentation (DiffAug) have enabled high-quality synthetic image generation from small or heterogeneous datasets, as evidenced by their successful application in thyroid ultrasound imaging [74].

Evaluating synthetic ultrasound images involves a multifaceted approach, incorporating both subjective assessments and objective metrics. Fréchet Inception Distance (FID), Structural Similarity Index Measure (SSIM), and Peak Signal-to-Noise Ratio (PSNR) remain standard metrics for quantifying image realism and fidelity. Studies have consistently demonstrated that GANs augment model performance when synthetic datasets are introduced, with improvements as high as 15.3% in classification accuracy [8] [76]. Despite these

achievements, the subjectivity inherent in radiologist-driven evaluations highlights the necessity for more standardised and reproducible evaluation frameworks. Emerging hybrid methods, such as semi-supervised learning frameworks [79] and auxiliary classifier evaluations [74], represent promising steps toward greater objectivity and precision in assessment methodologies.

2.7.2 Conclusions

While GAN-based approaches have unequivocally transformed synthetic ultrasound image generation, persistent challenges warrant continued investigation. Mode collapse and instability during GAN training remain significant obstacles, often constraining the diversity and generalisability of generated datasets [73] [80]. Moreover, the reliance on high-quality input data and substantial computational resources poses accessibility challenges in resource-limited environments, thereby impeding the equitable adoption of these technologies. Addressing dataset heterogeneity and scaling multimodal data integration are crucial for advancing the robustness of GAN models [74] [79].

Future research should prioritise the refinement of GAN architectures to accommodate limited and diverse datasets while enhancing training stability. Conditional GANs (cGANs) and Wasserstein GANs (WGANs) represent promising avenues for reducing training instability and improving output diversity [73] [80]. The expansion of real-time synthesis capabilities and rigorous clinical validation will be instrumental in transitioning these innovations into routine clinical practice [75]. Furthermore, integrating domain-specific knowledge and novel augmentation strategies, such as multi-resolution training and adaptive feature fusion, could significantly elevate the clinical utility of synthetic data. Such advancements will play a pivotal role in addressing global disparities in healthcare by enabling cost-effective and accessible diagnostic tools.

In summary, the advancements in artificial ultrasound generation underscore their transformative potential for revolutionising medical imaging. By addressing existing limitations and fostering interdisciplinary collaboration, this field stands poised to redefine diagnostic and training methodologies in healthcare. Ensuring alignment with clinical needs and ethical considerations will be paramount as generative technologies continue to evolve, paving the

way for equitable, high-quality, and impactful healthcare outcomes.

2.8 Conclusions

2.8.1 Summary

This review provides a comprehensive examination of the integration of computational methodologies into ultrasound imaging, focusing on critical domains such as image enhancement, feature detection, classification, lung ultrasound diagnosis, lung ultrasound datasets, and artificial data generation. Advanced techniques in image enhancement have effectively addressed challenges like speckle noise and low contrast, with methods such as wavelet-based filtering, anisotropic diffusion, and machine learning-driven algorithms consistently improving diagnostic reliability and visual clarity. Automation of feature detection through convolutional neural networks (CNNs) and graph-based algorithms has optimised the identification of critical artefacts, including A-lines, B-lines, and pleural irregularities, leading to enhanced diagnostic precision for complex conditions such as pulmonary oedema, ARDS, and interstitial lung disease.

Lung ultrasound diagnosis has become indispensable in pulmonary medicine, offering a non-invasive, portable, and cost-effective imaging solution with high sensitivity and specificity for numerous conditions. The integration of AI-driven tools and standardised protocols like the BLUE-protocol and FALLS-protocol has transformed diagnostic workflows, particularly in emergency and resource-limited settings. Accurate interpretation of artefacts, including B-lines and A-lines, has significantly improved clinical outcomes.

Lung ultrasound datasets have been pivotal in the development and validation of computational models, providing essential training resources for AI. Publicly available datasets such as COVIDx-US and POCUS have facilitated research into diagnosing conditions like COVID-19 and pneumonia. Meanwhile, initiatives like ICLUS-DB address issues of heterogeneity and standardisation, fostering greater reproducibility and collaboration within the research community.

Advancements in ultrasound classification have been transformative, with deep learning

frameworks achieving remarkable accuracy and speed in categorising diagnostic data. By incorporating multi-modal datasets, these systems offer contextually rich insights that enhance diagnostic workflows in both emergency and clinical settings. Artificial ultrasound generation, particularly through generative adversarial networks (GANs), has addressed the scarcity of annotated data, creating high-fidelity synthetic datasets that facilitate the exploration of rare pathological conditions and provide valuable resources for medical education and training.

The synthesis of these advancements underscores the transformative potential of computational approaches in redefining ultrasound imaging. By enhancing diagnostic accuracy, workflow efficiency, and accessibility, these methodologies have established themselves as critical tools in modern healthcare.

2.8.2 State-of-the-art

The state-of-the-art in computational ultrasound imaging is characterised by sophisticated methodologies that have redefined the field across multiple domains. Image enhancement techniques, including hybrid approaches that combine wavelet transformations and machine learning-based filters, have achieved exceptional clarity in addressing noise and artefacts. Advanced methods such as non-local means filtering and stochastic distance-based BM3D exemplify the precision of speckle noise reduction while preserving structural integrity.

In feature detection, CNN architectures have surpassed traditional methods, facilitating the accurate identification of artefacts essential for diagnosing pulmonary and cardiac conditions. Techniques such as Radon transforms and deep learning-driven segmentation models demonstrate high accuracy in detecting critical markers like B-lines and pleural abnormalities. Hybrid algorithms that integrate top-hat filtering and regularised optimisation further enhance artefact delineation in challenging imaging scenarios.

Lung ultrasound diagnosis has advanced through the application of AI-driven diagnostic tools, with automated systems achieving unprecedented specificity in detecting artefacts like B-lines. Standardised protocols such as the BLUE-protocol, enhanced by AI-guided frameworks, represent cutting-edge innovations that optimise diagnostic workflows for conditions such as ARDS and pulmonary oedema.

State-of-the-art lung ultrasound datasets, including COVIDx-US and ICLUS-DB, establish benchmarks for data quality and availability. These repositories feature innovative annotation frameworks and standardised imaging protocols, enabling robust AI model training and validation. By addressing issues of reproducibility and collaboration, these datasets expand the utility of AI in diagnostic imaging.

Ultrasound classification systems, driven by architectures like EfficientNet, ResNet, and CNN-LSTM hybrids, have demonstrated unparalleled accuracy in diagnosing conditions such as pneumonia, COVID-19, and ARDS. GAN-based artificial ultrasound generation complements these systems by creating diverse and realistic synthetic datasets, mitigating data scarcity and supporting model validation for rare or complex cases.

These advancements collectively signal a transformative era in computational ultrasound imaging, demonstrating methodologies that enhance diagnostic capabilities while broadening the accessibility and utility of ultrasound technology in diverse clinical contexts.

2.8.3 Gaps in the research

Despite substantial progress, significant challenges and limitations persist in computational ultrasound imaging. Operator dependency remains a critical issue, influencing the consistency and quality of imaging outcomes. Robust automation tools that standardise scan acquisition across diverse clinical environments are urgently needed. Variability in imaging protocols and the limited availability of large, diverse datasets continue to hinder the generalisability of AI models. Expanding dataset diversity and implementing universal imaging standards will be essential to overcoming these barriers.

Ethical considerations, such as data privacy, algorithmic transparency, and the interpretability of AI models, present additional hurdles. The absence of explainable AI frameworks undermines clinician trust and complicates the integration of these technologies into routine practice. Addressing these issues will require models that are not only accurate but also interpretable, ensuring seamless adoption into clinical workflows without compromising patient safety.

Computational intensity is another major limitation, as many state-of-the-art methods demand significant resources that may be unavailable in resource-constrained settings. De-

veloping lightweight and scalable algorithms is essential to ensure the accessibility and practicality of these techniques in diverse healthcare environments.

In lung ultrasound diagnosis, reliance on operator expertise continues to pose challenges. AI systems capable of standardising and automating diagnostic processes are essential to mitigate this dependency. Protocol heterogeneity across institutions further complicates clinical applications, highlighting the need for validated, universally applicable guidelines.

Lung ultrasound datasets, while transformative, face persistent issues related to heterogeneity in data quality and annotation practices. Initiatives like ICLUS-DB represent promising steps toward addressing these gaps, but further efforts are required to bridge the divide between data availability and clinical utility. Standardising annotation practices and diversifying dataset content will be critical for ensuring the robustness and applicability of AI models.

Future research should prioritise interdisciplinary collaboration, integrating traditional imaging techniques with advanced machine learning approaches. Exploring novel hybrid methodologies holds significant promise for overcoming current limitations. Addressing these gaps will be pivotal in realising the full potential of computational ultrasound imaging, enabling equitable, high-quality diagnostic solutions across global healthcare systems.

Chapter 3

Enhancing lung ultrasound image contrast

3.1 Introduction

Ultrasound imaging, a cornerstone in medical diagnostics, offers a non-invasive, real-time method to visualise internal body structures. Despite its advantages, such as affordability and safety, ultrasound imaging suffers from inherent limitations including speckle noise and poor contrast, which can obscure critical diagnostic features. These challenges are particularly pronounced in lung ultrasound (LUS), where the complex anatomical structures and the presence of noise can hinder accurate interpretation. Addressing these issues necessitates advanced image processing techniques that not only enhance image clarity but also preserve diagnostic details crucial for medical decision-making.

This chapter explores a comprehensive approach to overcoming these limitations by integrating innovative noise reduction and contrast enhancement methodologies. Leveraging k-means clustering guided by Rayleigh mixture models, the proposed techniques aim to optimise the visualisation of diagnostic markers in LUS images. These methods are evaluated against established techniques such as Contrast Limited Adaptive Histogram Equalisation (CLAHE) and combined with noise reduction strategies like BM3D. The focus is on enhancing the quality of LUS images to facilitate better identification of features like the pleural line, B-lines, and A-lines, which are pivotal in diagnosing pulmonary conditions.

The chapter begins by detailing the foundational principles of noise reduction and contrast enhancement, laying the groundwork for understanding the challenges and opportunities in LUS image processing. It then delves into the proposed methodologies, emphasising their theoretical underpinnings and practical implementations. The results are systematically analysed using a range of quantitative metrics—including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Contrast Gain (CG)—to evaluate the effectiveness of the techniques. By combining subjective visual assessments with objective metric-based evaluations, this study provides a holistic perspective on the advancements in LUS image enhancement.

Through this investigation, the chapter aims to contribute to the broader field of medical imaging by demonstrating the potential of tailored image processing techniques in addressing domain-specific challenges. The findings underscore the importance of integrating computational methodologies with domain knowledge to develop tools that improve diagnostic accuracy and clinical outcomes.

3.2 Current methods

Enhancement techniques in ultrasound imaging are critical to improving the quality of images affected by speckle noise, low contrast, and other artefacts that compromise diagnostic accuracy. These techniques aim not only to suppress noise but also to enhance key image features, such as tissue boundaries, edges, and contrast, making it easier for clinicians to identify and analyse critical anatomical structures. While noise reduction is one aspect of enhancement, a comprehensive approach requires balancing multiple factors such as edge preservation, contrast improvement, and the retention of fine structural details.

Traditional methods for enhancing ultrasound images often focus on filtering out speckle noise while preserving important diagnostic information. One of the foundational approaches is anisotropic diffusion filtering, which works by smoothing homogeneous regions of the image while maintaining edges and other high-frequency structures. This technique has been effective in reducing noise without overly blurring edges, making it suitable for many clinical applications. However, anisotropic diffusion filters often require careful parameter

tuning and can be computationally expensive, limiting their practicality in real-time settings [81].

Wavelet-based methods have also been widely employed in ultrasound image enhancement. The BM3D (Block-Matching and 3D) algorithm, for example, is particularly effective in grouping similar blocks of the image and applying collaborative filtering to reduce noise while retaining essential structural information. When combined with brushlet thresholding, BM3D demonstrates a significant ability to suppress noise while enhancing the contrast-to-noise ratio (CNR) [36]. This is especially valuable in medical imaging, where both noise suppression and contrast enhancement are crucial for improving image clarity and diagnostic utility.

Another traditional technique used in ultrasound enhancement is Contrast-Limited Adaptive Histogram Equalisation (CLAHE). CLAHE works by adjusting the histogram of an image in small regions, thereby enhancing local contrast while preventing over-amplification of noise. This method is particularly useful for ultrasound images that suffer from low contrast and speckle noise. A hybrid form of CLAHE, incorporating gamma correction and high-pass filtering, has been shown to further enhance image quality by sharpening edges and improving visual detail in targeted regions of interest [41].

More recently, advanced techniques have emerged that utilise multiscale and nonlinear methods to achieve a more refined balance between noise reduction and image enhancement. Multiscale wavelet diffusion techniques, for example, decompose the image into different spatial scales, allowing for targeted enhancement at each level. This approach is particularly effective in handling speckle noise, which manifests at different frequencies, while also preserving fine anatomical details critical for diagnosis [82]. By applying nonlinear diffusion across multiple scales, these techniques can suppress noise without blurring important structures, making them well-suited for complex medical imaging tasks.

Nonlocal means (NL-means) filtering represents another significant advancement in ultrasound image enhancement. Traditional filtering methods rely on local pixel information to reduce noise, but NL-means compares larger image patches, allowing for more accurate noise suppression. The Optimised Bayesian Nonlocal Means (OBNLM) filter takes this a step further by incorporating a probabilistic model that estimates noise distribution, mak-

ing it particularly effective in ultrasound imaging where noise is signal-dependent [83]. The Bayesian framework allows the filter to preserve important edges and structures, ensuring that critical diagnostic information is retained while reducing speckle noise. However, the computational demands of NL-means filtering, particularly in its Bayesian form, present challenges for real-time application in clinical settings.

The advent of machine learning and artificial intelligence (AI) has brought new possibilities to the field of ultrasound enhancement. AI-driven techniques, particularly those utilising deep learning models, have shown remarkable performance in both noise reduction and contrast enhancement. These models can learn complex patterns from large datasets, allowing them to adapt to various noise characteristics and imaging conditions more effectively than traditional techniques.

Convolutional Neural Networks (CNNs), for instance, have been used in conjunction with CLAHE to enhance ultrasound images, particularly in video formats. By using a CNN to reduce speckle noise, and CLAHE to improve contrast, these hybrid models can produce images with sharper tissue boundaries and greater detail than traditional methods [42]. The flexibility of AI-based approaches allows for more dynamic adjustment of enhancement parameters, leading to improved adaptability across different types of ultrasound images and clinical applications.

Despite their potential, AI-based methods face several challenges. One of the primary issues is the computational complexity of these models, which can hinder their use in real-time clinical settings. Ultrasound imaging is often performed in fast-paced environments, such as emergency rooms or during surgical procedures, where rapid image processing is essential. Ensuring that AI models can function in real time without compromising image quality is a critical area of ongoing research. Additionally, the black-box nature of many AI models raises concerns about interpretability and the ability to ensure consistent diagnostic accuracy across different patient populations [36].

One of the most significant challenges in ultrasound image enhancement is balancing noise reduction with the preservation of diagnostic information. Excessive noise reduction, especially when applied indiscriminately, can lead to the loss of important anatomical details, such as tissue boundaries and lesions. This is particularly problematic in medical

applications, where the visibility of subtle features can be crucial for diagnosis. Techniques that rely solely on noise suppression may inadvertently smooth over these critical details, reducing the overall diagnostic utility of the image [81] [36].

To address this, many modern enhancement techniques adopt a multi-objective approach that balances noise reduction with contrast improvement and edge preservation. For example, multiscale wavelet diffusion methods allow for noise suppression at different spatial scales, ensuring that fine details are retained while larger, more homogeneous areas are smoothed [82]. Similarly, machine learning models can be trained to prioritise different aspects of image quality depending on the clinical task, such as enhancing edges for tumour detection or improving contrast for vascular imaging.

A key limitation of many current enhancement techniques is their computational complexity, which makes them unsuitable for real-time clinical applications. Techniques such as wavelet-based methods, Bayesian filtering, and deep learning models, while highly effective, often require significant computational resources, limiting their feasibility in dynamic clinical settings such as intraoperative ultrasound or emergency diagnostics [36] [83]. As ultrasound imaging is often performed in real time, during procedures that require immediate feedback, it is critical to develop enhancement techniques that can operate quickly and efficiently without sacrificing image quality.

Recent research has begun to focus on optimising these advanced techniques for real-time use. For instance, block-based implementations of NL-means filters and more efficient CNN architectures are being developed to reduce the computational burden while maintaining high levels of image enhancement [83] [42]. Furthermore, the integration of AI models with traditional filtering techniques, such as CLAHE, offers a promising path forward, as these hybrid approaches can combine the strengths of both methods to produce high-quality images with lower computational costs [42].

The current methodology for ultrasound image enhancement demonstrates substantial progress in mitigating speckle noise and improving diagnostic clarity. However, there remain critical areas for improvement that warrant further research. Firstly, reducing the dependency on user-defined input parameters through automation or adaptive algorithms could significantly streamline the application of these techniques, making them more accessible and

consistent across diverse clinical scenarios. Secondly, the development of contrast enhancement techniques specifically tailored to the unique characteristics of ultrasound imaging is essential. Such techniques would address the challenges posed by the modality's inherent low contrast and varying tissue interfaces. Future work should also explore integrating these advancements into real-time imaging workflows, ensuring their practicality and efficiency in clinical settings.

3.3 Contrast enhancement using PDF mixture models formed by k-means clustering

A precise and sequential approach is employed to formulate a methodology for enhancing Lung ultrasound images through histogram data distribution. This section outlines the systematic steps undertaken in this chapter to accomplish image enhancement, emphasising the application of statistical models and clustering techniques.

This methodology begins by defining input parameters crucial for image histogram transformation. The number of clusters for the K-means algorithm must be predetermined, as this determines the number of Gaussian components in the resulting mixture model. In this study, five clusters have been chosen to balance computational efficiency and granularity.

The weighted K-means clustering technique is applied to partition the image intensity levels into distinct regions. This is followed by modeling each cluster with a Gaussian distribution, effectively forming a Gaussian mixture model (GMM). The clustering process incorporates weighting factors based on pixel intensity and spatial relationships to enhance robustness against noise and outliers. These weights adaptively prioritize regions with higher variance, enabling a finer representation of the image's underlying histogram structure.

Following the delineation of the regions, a mixture model is estimated from the clusters to facilitate a comparative analysis with the target model derived from the input. Utilising Maximum Likelihood Estimation (MLE) on each cluster, the parameters of a Gaussian distribution are fitted to the data. The weight of each Gaussian within the mixture model is determined by the proportion of the dataset corresponding to each cluster relative to the

entire image. Subsequent to the application of MLE to each region, the mixture model is employed in conjunction with the target model for further analysis.

The process of target probability density function (PDF) training begins by manually altering the original images to create an "ideal" reference image. This manual adjustment involves enhancing the contrast, brightness, or other relevant image features to visually achieve an optimal representation of the desired output. These "ideal" images act as references for the subsequent statistical modelling.

To estimate the PDF, the K-means clustering algorithm is employed to partition the intensity values of the ideal image into distinct clusters. The number of clusters, K, is predetermined based on the level of granularity required, with $K = 5$ clusters selected in this study for balanced representation. The centroids, μ_k, of these clusters are used as initial estimates for fitting a Gaussian Mixture Model (GMM). The GMM models the intensity distribution as a weighted sum of Gaussian components.

For contrast enhancement, the target probability density function (PDF) is crafted to fit a desired transformation profile. In this work, a Rayleigh-Rician mixture model is employed, as well as a GMM and Rayleigh-GMM, as they capture a wide range of intensity distributions, particularly for medical imaging applications.

For an overview of the operation of the algorithm, Figure 3.1 shows this, with the input image and start of the algorithm on the left, and subsequent processes after.

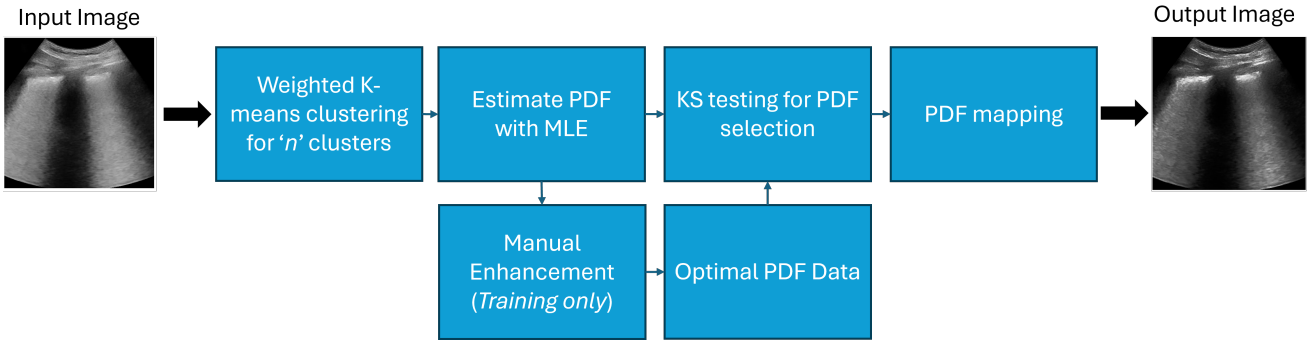

Figure 3.1: K-means based contrast enhancement algorithm

3.3.1 Target PDF training

The Target PDF Training process focuses on creating Probability Density Functions (PDFs) that guide the enhancement of ultrasound image contrast. This involves generating both an original and a target PDF through clustering and manual adjustments. Below are the detailed steps:

Load target image and compute original PDF

The input ultrasound image $I(x, y)$ is used to compute the intensity distribution $P_{\text{original}}(x)$ using a Gaussian Mixture Model (GMM) with n predefined clusters. The GMM is expressed as:

$$P_{\text{original}}(x) = \sum_{k=1}^{n} \pi_k f_k(x), \tag{3.1}$$

where:

- π_k: Weight (or mixing coefficient) of the k-th Gaussian component, satisfying $\sum_{k=1}^{n} \pi_k = 1$,
- $f_k(x)$: Probability density function of the k-th Gaussian component, defined as:

$$f_k(x) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right), \tag{3.2}$$

- μ_k, σ_k^2: Mean and variance of the k-th Gaussian.

The parameters π_k, μ_k, σ_k^2 are estimated using the Expectation-Maximization (EM) algorithm.

Weighted k-means clustering

The image is divided into n clusters using weighted K-means clustering. For each pixel x_i, a weight w_i is assigned based on its spatial proximity and intensity similarity. The clustering minimizes:

$$J = \sum_{i=1}^{N} w_i \|x_i - c_{\phi(i)}\|^2, \tag{3.3}$$

where:

- w_i: Weight for pixel i, calculated as:

$$w_i = \exp\left(-\frac{\|(x_i, y_i) - (x_c, y_c)\|^2}{2\sigma_s^2}\right), \tag{3.4}$$

- (x_c, y_c): Centroid coordinates of the cluster,

- σ_s: Spatial bandwidth controlling the influence of proximity,

- $c_{\phi(i)}$: Intensity centroid of the cluster to which x_i belongs.

The algorithm iteratively updates the cluster centroids c_k until convergence, resulting in partitioned regions C_k for $k = 1, \ldots, n$:

$$C_k = \{x_i \mid \phi(i) = k\}. \tag{3.5}$$

Manual contrast modification

For each cluster C_k, the intensities are manually adjusted to enhance contrast, targeting diagnostically relevant features. The modified intensity $I'(x, y)$ for a pixel is calculated as:

$$I'(x, y) = I(x, y) + \alpha_k(x - c_k), \tag{3.6}$$

where:

- α_k: Contrast adjustment factor for the k-th cluster,

- c_k: Centroid intensity of the cluster.

To ensure smooth transitions at cluster boundaries, a Gaussian kernel is applied to blend the contrasting adjustments. The blending kernel $G(x, y)$ is defined as:

$$G(x, y) = \exp\left(-\frac{\|(x, y) - (x_b, y_b)\|^2}{2\sigma_f^2}\right), \tag{3.7}$$

where:

- (x_b, y_b): Coordinates of the boundary pixel,

- σ_f: Bandwidth controlling the feathering effect.

The feathered intensity is then:

$$I_f(x, y) = G(x, y)I'(x, y) + (1 - G(x, y))I(x, y). \tag{3.8}$$

Generate New PDF

The modified image $I_f(x, y)$ is modeled as a new GMM with n components. Each component is bound by the mean of its corresponding cluster:

$$P_{\text{target}}(x) = \sum_{k=1}^{n} \pi'_k f'_k(x), \tag{3.9}$$

where $f'_k(x)$ is defined as:

$$f'_k(x) = \frac{1}{\sqrt{2\pi\sigma'^2_k}} \exp\left(-\frac{(x - c_k)^2}{2\sigma'^2_k}\right), \tag{3.10}$$

with c_k representing the cluster mean.

The parameters π'_k, σ'^2_k are re-estimated based on the modified intensities $I_f(x, y)$.

Store PDFs

Both the original $P_{\text{original}}(x)$ and the target $P_{\text{target}}(x)$ PDFs are stored for use in the PDF Mapping stage. In Figure 3.2 the PDFs are stored and graphed in blue (input image) and green (closet fit manually edited optimal image), with an example of the final mapping result of these PDFs shown in red (result after mapping)

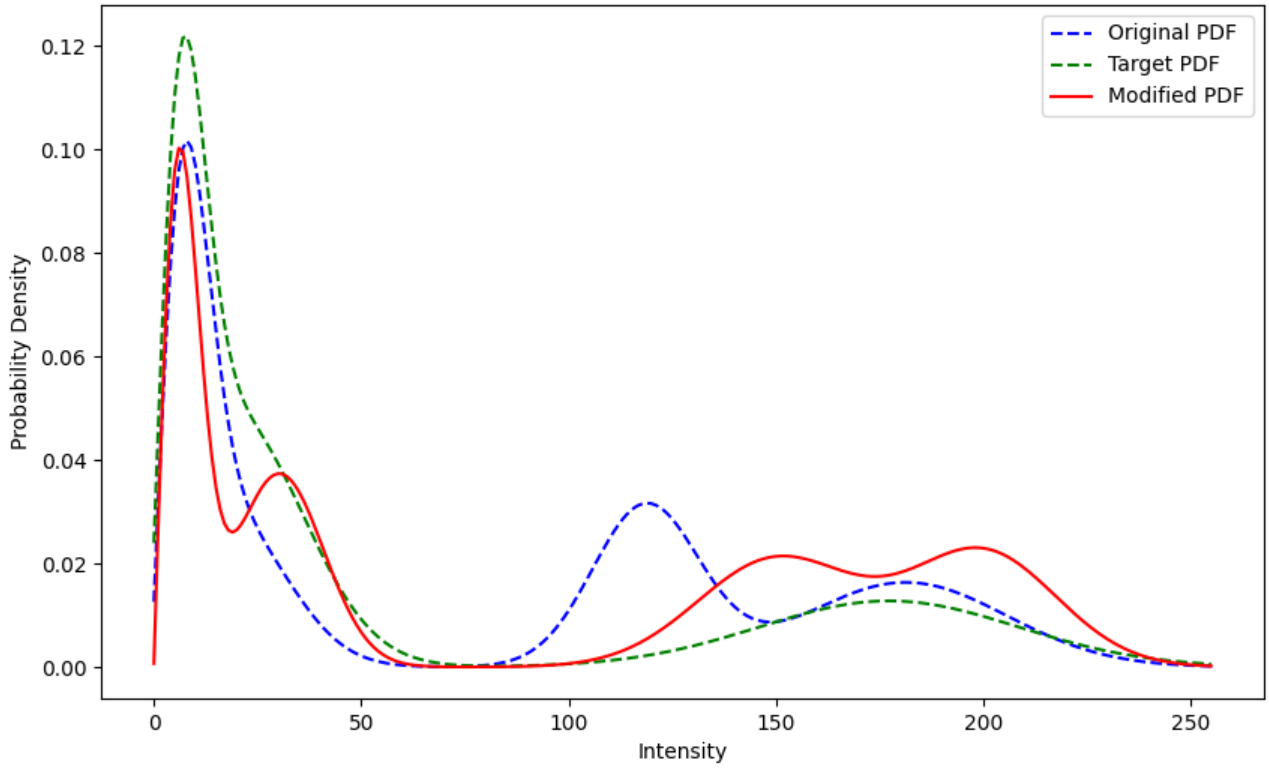

Figure 3.2: Graph to show example PDFs for R-GMM modifications

Subsequent to modelling, integration is performed across the defined limits—the minimum and maximum intensity values of each cluster—to calculate the area under each segment

of the PDF. This calculation is conducted for both the target and the ideal PDFs, facilitating a comparative analysis.

3.3.2 PDF mapping using Kolmogorov–Smirnov test for contrast enhancement

Once the target PDF is trained, the next step involves mapping the image histogram to align with the desired distribution. This is achieved using the Kolmogorov–Smirnov (KS) test, a statistical method that quantifies the difference between the empirical and target cumulative distribution functions (CDFs).

The mapping process minimizes the KS statistic, ensuring a close match between the image's transformed histogram and the target PDF. By leveraging the cumulative distribution alignment, this approach preserves the relative intensity order while enhancing the global contrast. The transformation parameters derived from the KS test ensure an adaptive adjustment to the unique characteristics of the input image.

To further refine the enhancement, iterative optimization is employed, wherein small adjustments to the mapping function are made until convergence. This guarantees that the contrast enhancement process does not introduce artifacts or distortions, maintaining the image's structural integrity.

This stage enhances the contrast of new ultrasound images by mapping their intensity distributions to the most suitable stored PDF, determined through Kolmogorov–Smirnov (KS) testing. The process ensures precise alignment between input and target intensity distributions for optimal contrast enhancement.

Estimate target image PDF

The intensity distribution of the input image, $I_{\text{input}}(x, y)$, is estimated using Maximum Likelihood Estimation (MLE). The resulting PDF, $P_{\text{input}}(x)$, is expressed as:

$$P_{\text{input}}(x) = \sum_{k=1}^{n} \pi_k f_k(x), \tag{3.11}$$

where:

- π_k: Mixing coefficient for the k-th Gaussian component,

- $f_k(x)$: Probability density function of the k-th Gaussian component,
- n: Number of Gaussian components.

Perform KS Testing for Closest Fit

The Kolmogorov–Smirnov (KS) test is used to compare the input image PDF $P_{\text{input}}(x)$ with each stored PDF, $P_{\text{stored}}(x)$, and determine the closest match. The KS statistic D is calculated as:

$$D = \max_x |F_{\text{input}}(x) - F_{\text{stored}}(x)|, \tag{3.12}$$

where:
- $F_{\text{input}}(x)$: Cumulative distribution function (CDF) of $P_{\text{input}}(x)$,
- $F_{\text{stored}}(x)$: CDF of the stored PDF.

The PDF $P_{\text{stored}}(x)$ with the smallest D value is selected as the best fit for mapping.

Compute Cumulative Distribution Functions (CDFs)

Derive the cumulative distribution functions for both the input PDF and the target PDF. The input image's cumulative distribution function is given by:

$$C_I(x) = \int_{-\infty}^{x} P_I(t)\, dt \tag{3.13}$$

Similarly, the cumulative distribution function for the target PDF is:

$$C_T(x) = \int_{-\infty}^{x} P_T(t)\, dt \tag{3.14}$$

where:
- $C_I(x)$: Cumulative distribution function of the input image PDF,
- $P_I(t)$: Input image probability density function,
- $C_T(x)$: Cumulative distribution function of the target PDF,
- $P_T(t)$: Target probability density function,
- x, t: Intensity levels in the input and target domains, respectively.

Mapping Relationship Between PDFs

The mapping between the input PDF and the target PDF is defined using their cumulative distribution functions. For a pixel intensity x in the input image, the mapped intensity is determined as:

$$x_{\text{mapped}} = C_T^{-1}(C_I(x)) \tag{3.15}$$

where:

- x_{mapped}: Mapped pixel intensity in the transformed image,
- $C_I(x)$: Cumulative probability of x in the input image,
- C_T^{-1}: Inverse cumulative distribution function of the target PDF.

Interpolation for Discrete Values

Since PDFs and CDFs are discrete in implementation, **linear interpolation** is applied to approximate the mapping. The process involves estimating the output intensity by linearly interpolating between the closest known values in the target cumulative distribution:

$$x_{\text{mapped}} \approx C_T^{-1}\big(C_I(x)\big) \quad \text{via linear interpolation.} \tag{3.16}$$

where:

- Linear Interpolation: Estimates intermediate values by connecting adjacent points in the CDF,
- $C_T^{-1}, C_I(x)$: Target inverse CDF and input cumulative probability values, respectively.

Apply Transformation to Image

Finally, replace the pixel intensities of the input image with their mapped values:

$$x_{\text{output}} = x_{\text{mapped}} \tag{3.17}$$

where:

- x_{output}: Final transformed pixel intensity in the output image,
- x_{mapped}: Pixel intensity mapped to the target PDF.

3.3.3 Image quality metrics

Although measuring image quality from the human perspective is an effective way of gauging any improvement made by image processing, it can be hard to quantify this result as personal opinion can be subjective depending on the reader. This is where using statistical metrics such as Mean square error and peak signal to noise ratio can represent a mathematical response to the image processing result.

The first very commonly used metric is mean square error. MSE measures the variance in noise from the original image and the new de-noised image. Below is the formula that describes this method with X being the original image, Y being the new denoised image:

The Mean Squared Error (MSE) is given by the formula:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2 \tag{3.18}$$

Although commonly used, this method does not correlate well with perceptive image quality so it is important that it is not to be solely used as a judgement on image quality.

Advancing on with MSE is the peak signal to noise ratio which takes the desired signal compared to the noise affecting the image. PSNR is measured in decibels (dB) and a higher value tends to represent a higher quality image. Below is the equation for PSNR when using an MSE input, with S being the maximum pixel intensity:

The Peak Signal-to-Noise Ratio (PSNR) is given by the formula:

$$\text{PSNR} = 10\log_{10}\left(\frac{S^2}{\text{MSE}}\right) \tag{3.19}$$

Again similarly to using MSE as a quality assessment metric, PSNR doesn't always correctly represent a good quality image from a bad image, especially in the case of having an unknown noise signal present in medical ultrasound, hence meaning there isn't a perfect comparison image (Known as full reference) to represent the quality of the original image to the new image as there will only be measurable result produced by the new image.

Structural Similarity Index

SSIM measures the perceived image quality by comparing structural information between the original and processed images. It evaluates luminance, contrast, and structural similarity [46]:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3.20}$$

where:

- μ_x, μ_y: Mean intensities of images x and y,
- σ_x^2, σ_y^2: Variances of x and y,
- σ_{xy}: Covariance of x and y,
- C_1, C_2: Stabilization constants to avoid division by zero.

Coefficient of Variation

CV measures the relative variability of an image's pixel intensities, reflecting contrast uniformity [84]:

$$\text{CV} = \frac{\sigma}{\mu} \tag{3.21}$$

where:

- σ: Standard deviation of pixel intensities,
- μ: Mean intensity of the image.

Contrast Gain

CG evaluates the improvement in contrast between the original and enhanced images [85]:

$$\text{CG} = \frac{\sigma_{\text{enhanced}}}{\sigma_{\text{original}}} \tag{3.22}$$

where:

- σ_{enhanced}: Standard deviation of the enhanced image,
- σ_{original}: Standard deviation of the original image.

Mean Gradient Magnitude

MGM quantifies edge sharpness by averaging the gradient magnitude across the image [86]:

$$\text{MGM} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\left(\frac{\partial I}{\partial x_i}\right)^2 + \left(\frac{\partial I}{\partial y_i}\right)^2} \tag{3.23}$$

where:

- $\frac{\partial I}{\partial x_i}, \frac{\partial I}{\partial y_i}$: Gradients along x- and y-directions,
- N: Total number of pixels in the image.

No reference image quality assessment works from the foundation that an image has no reference, therefore its quality is directly measured from the single input image. A new machine learning-based method called 'BRISQUE' stands for Blind reference image spatial quality evaluator measures and creates a quality metric from a single image. Using a SVM the model is loaded with a collection of distortions such as gaussian blur, additive noise, compression, etc. It is then trained to calculate these features as vectors and recognise them in images. With this model any image inputted to the BRISQUE algorithm will be scored by comparing the image vectors with the feature vectors with a higher score meaning the image is related closer to a distortion; i.e. lower quality image. [87]

Again it must be mentioned that although this method of quality metrics is without a reference it still can not be solely used to judge the quality of an image as it is always subjective to the desired result. For example, BRISQUE may evaluate a noise-free blurry image to be lower in quality than a highly noisy image as the image blur vector may outweigh the noise-free one, resulting in a high score, even though the desired result was to remove noise from the image.

3.4 Noise reduction methods

Image denoising is a critical preprocessing step in ultrasound image analysis, aimed at mitigating the effects of speckle noise while preserving diagnostic features such as edges and tissue boundaries. This section does not propose new denoising approaches but focuses on evaluating state-of-the-art methods under a unified framework to determine their relative

performance. Specifically, these methods are applied in conjunction with the proposed contrast enhancement techniques to assess their complementary effects on image quality. By employing diverse denoising approaches such as Nonlocal Means (NLM), Wavelet Thresholding, Optimized Bayesian Nonlocal Means (OBNLM), Block Matching and 3D Filtering (BM3D), and Anisotropic Diffusion with Memory (ADMSS). The goal is to identify the most effective combination of denoising and contrast enhancement.

This section outlines the noise reduction methods employed during testing, detailing their mechanisms and providing a formulaic overview. These methods have been selected for their efficacy in enhancing image quality while preserving important diagnostic features.

3.4.1 Nonlocal Means (NLM)

The Nonlocal Means (NLM) method reduces noise by averaging pixel intensities based on patch similarity across the image. The weight assigned to a pixel intensity depends on the similarity between patches and their spatial proximity [88]. The restored intensity $I_{\text{restored}}(x)$ at pixel x is computed as:

$$I_{\text{restored}}(x) = \frac{\sum_{y \in \Omega} w(x,y) I(y)}{\sum_{y \in \Omega} w(x,y)} \tag{3.24}$$

where:

- $w(x,y) = \exp\left(-\frac{\|P(x)-P(y)\|^2}{h^2}\right)$, weight based on patch similarity,
- $P(x), P(y)$: Patches centered at pixels x and y,
- Ω: Search window,
- h: Filtering parameter controlling decay of the Gaussian function.

3.4.2 Wavelet Thresholding

Wavelet thresholding removes noise by transforming the image into the wavelet domain, applying thresholding to coefficients, and reconstructing the image [89]. The process is described as:

$$I_{\text{denoised}} = W^{-1}\left(T(W(I))\right) \tag{3.25}$$

where:

- W: Wavelet transform operator,

- W^{-1}: Inverse wavelet transform,

- T: Thresholding function applied to wavelet coefficients.

3.4.3 Optimized Bayesian Nonlocal Means (OBNLM)

The OBNLM method is an optimized version of the NLM tailored for speckle noise, employing Bayesian principles to refine patch similarity calculations [83]:

$$I_{\text{restored}}(x) = \sum_{y \in \Omega} w(x,y) I(y), \quad w(x,y) = \frac{\exp\left(-\frac{(P(x)-P(y))^2}{2\sigma^2}\right)}{\sum_{y \in \Omega} \exp\left(-\frac{(P(x)-P(y))^2}{2\sigma^2}\right)} \tag{3.26}$$

where σ is the noise standard deviation.

3.4.4 Anisotropic Diffusion with Memory and Speckle Statistics (ADMSS)

The ADMSS filter employs anisotropic diffusion enhanced with a probabilistic-driven memory mechanism to adaptively control diffusion based on tissue characteristics [90]. The anisotropic diffusion process is governed by:

$$\frac{\partial u(x,t)}{\partial t} = \nabla \cdot \left(D(x,t) \nabla u(x,t)\right), \tag{3.27}$$

where:

- $u(x,t)$: Image intensity at spatial position x and time t,

- $\nabla u(x,t)$: Gradient of the image, capturing edge information,

- $D(x,t)$: Diffusion tensor, controlling the rate and direction of diffusion.

The memory mechanism modifies the diffusion tensor to adaptively preserve clinical relevance:

$$\frac{\partial L(x,t)}{\partial t} = \frac{1}{\tau(x)} \left(L(x,t) - S\{D(x,t)\}\right), \tag{3.28}$$

where:

- $L(x,t)$: Memory-enhanced diffusion tensor,

- $\tau(x)$: Relaxation time with spatial dependence,

- $S\{D(x,t)\}$: Selective filtering operator to preserve relevant structures.

The final memory-driven diffusion process is expressed as:

$$L(x,t) = S\{D(x,0)\}e^{-\frac{t}{\tau}} + \int_0^t e^{\frac{s-t}{\tau}} S\{D(x,s)\}ds. \tag{3.29}$$

This framework balances noise reduction in homogeneous regions and edge preservation in diagnostically important areas.

3.4.5 Block Matching and 3D Filtering (BM3D)

BM3D is a collaborative filtering method that groups similar patches, applies a transform domain shrinkage, and reconstructs the image [91].

Patch Grouping

Image patches P_i are grouped based on similarity:

$$\text{Group}(P_i) = \{P_j : \|P_i - P_j\| < \epsilon\}, \tag{3.30}$$

where:

- $\|P_i - P_j\|$: Distance between patches P_i and P_j,
- ϵ: Similarity threshold.

Transform-Domain Shrinkage

A 3D transform is applied to the grouped patches:

$$C_{3D} = T_{3D}(\text{Group}(P_i)), \tag{3.31}$$

where T_{3D} is a separable 3D transform combining 2D and 1D transforms. Shrinkage is performed on the coefficients:

$$\hat{C}_{3D} = \text{Shrink}(C_{3D}, \lambda), \tag{3.32}$$

where:

- $\hat{C}_{3D}$: Shrunk coefficients,
- λ: Threshold for noise reduction.

Inverse Transform and Aggregation

The inverse transform reconstructs the filtered patches:

$$\hat{P}_i = T_{3D}^{-1}(\hat{C}_{3D}), \tag{3.33}$$

and the final image is reconstructed by aggregating the patches:

$$I_{\text{denoised}}(x) = \frac{\sum_{i \in \mathcal{P}(x)} \hat{P}_i}{\sum_{i \in \mathcal{P}(x)} W_i}, \tag{3.34}$$

where:

- $\mathcal{P}(x)$: Set of patches overlapping pixel x,

- W_i: Weights based on patch reliability.

The BM3D algorithm leverages both local and nonlocal redundancies to effectively reduce noise while preserving fine details.

3.5 Results

This section presents the results of the Lung Ultrasound Contrast enhancement using k-means clustering guided by Rayleigh mixture models. The study evaluated the performance of the various proposed methods across multiple n-cluster sizes to ascertain their robustness and efficacy. All computational experiments were conducted on a system with an Intel Core i7-6700k CPU and an GTX 980ti GPU, ensuring the capability to handle the intensive processing requirements of the task. The following sections detail the outcomes and provide a comprehensive analysis of the findings.

3.5.1 Dataset

The lung ultrasound images utilised for contrast enhancement evaluation were obtained from the publicly available COVID-US dataset [26]. A manual random selection of scans was performed, comprising of 264 images with both linear and convex ultrasound probes to ensure diversity across imaging conditions. 30 random images were used to create "ideal" images for the use of storing the PDF data. This selection allows assessment of the proposed method under varying acquisition settings.

3.5.2 Subjective comparison

It has been proposed to use CLAHE to improve the low clarity in LUS images due to poor contrast [92]. CLAHE requires certain input parameters such as a clipping limit which prevents oversaturation in the image histogram by limiting the data peaks. In Figure 3.3 an example of the results of the various denoising methods are shown (De-noising methods as outlined in Section 3.4).

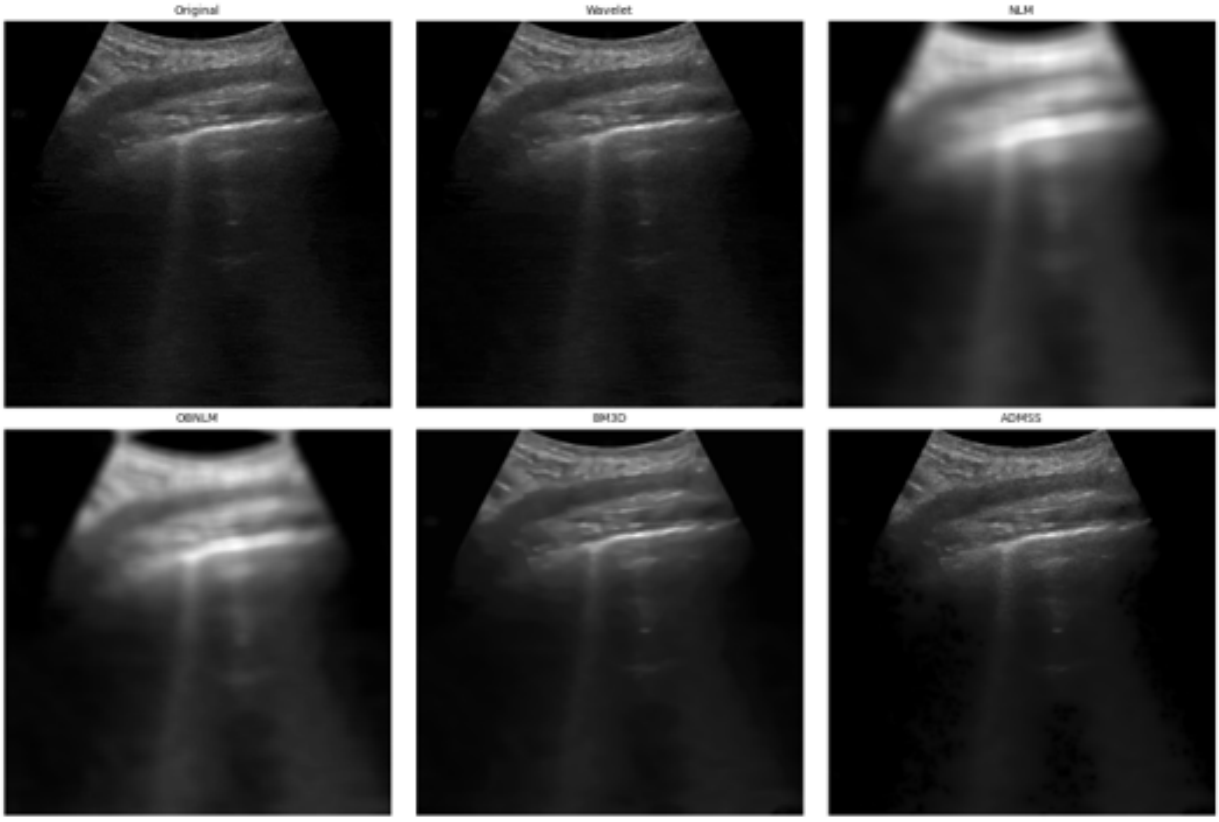

Figure 3.3: All de-noising methods

Alternatively, using the proposed method from this project; k-means region contrast enhancement was also tested on a random selection of images. The k-means clustering stage requires only one input parameter which is the number of regions. From testing, an automated approach to this parameter was determined by using the standard deviation calculated from the image histogram. This number was rounded to produce a whole number to satisfy the cluster input. Figure 3.4 shows an example of all contrast enhancement methods applied prior to denoising, with 3.5 contrast enhancement applied post ADMSS denoising.

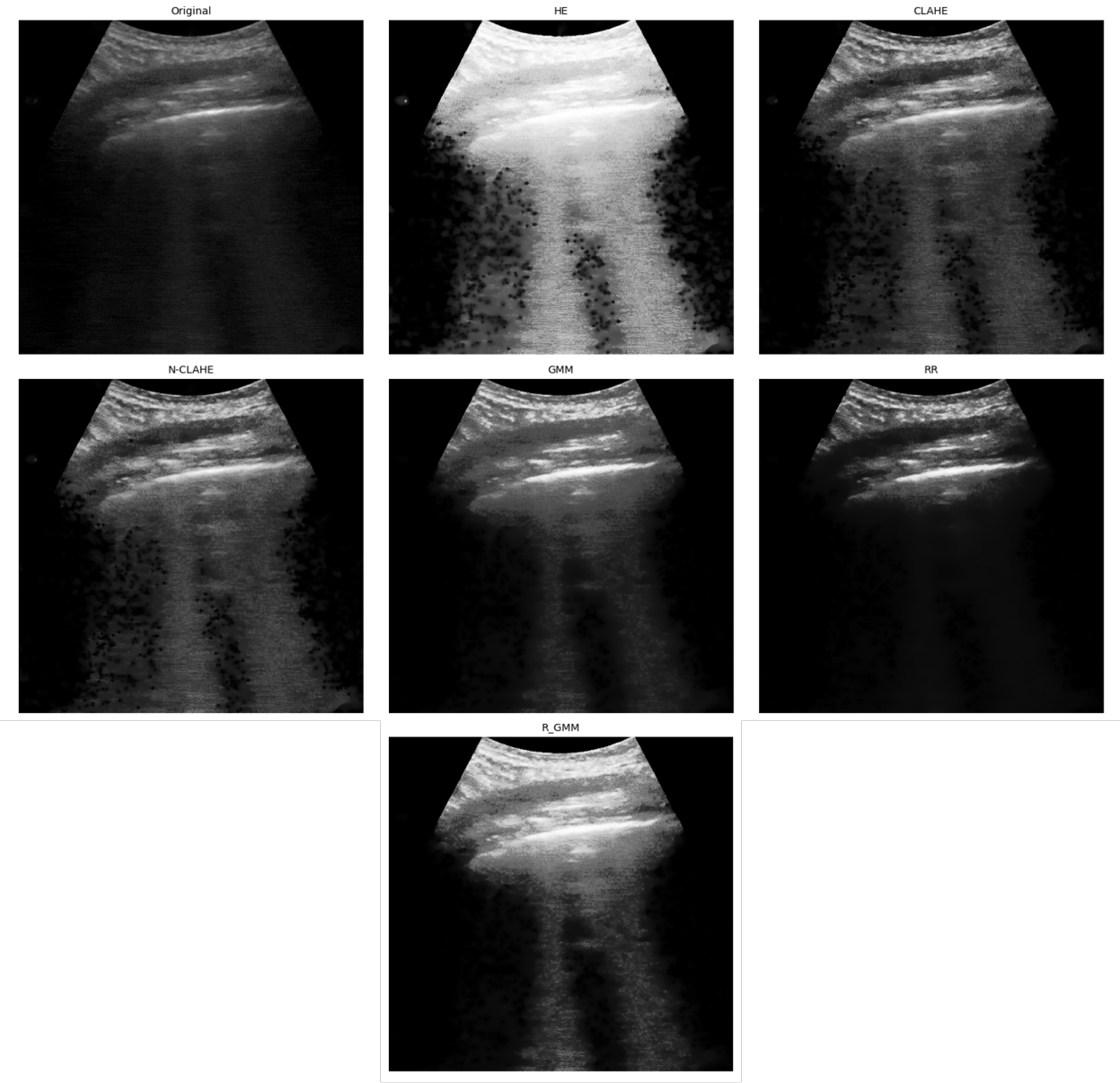

Figure 3.4: All contrast enhancement methods

From the experiments run with CLAHE, it is apparent that the method is overly sensitive to speckle noise present in the LUS images. Although the features are better separated (higher clarity) by contrast to the original image the overall quality of the features have been greatly degraded by the CLAHE method. Some regions that should have been reduced in intensity have become brighter by the CLAHE method as it was unable to differentiate these regions.

The proposed k-means clustering algorithm guided by Rayleigh mixture models signifi-

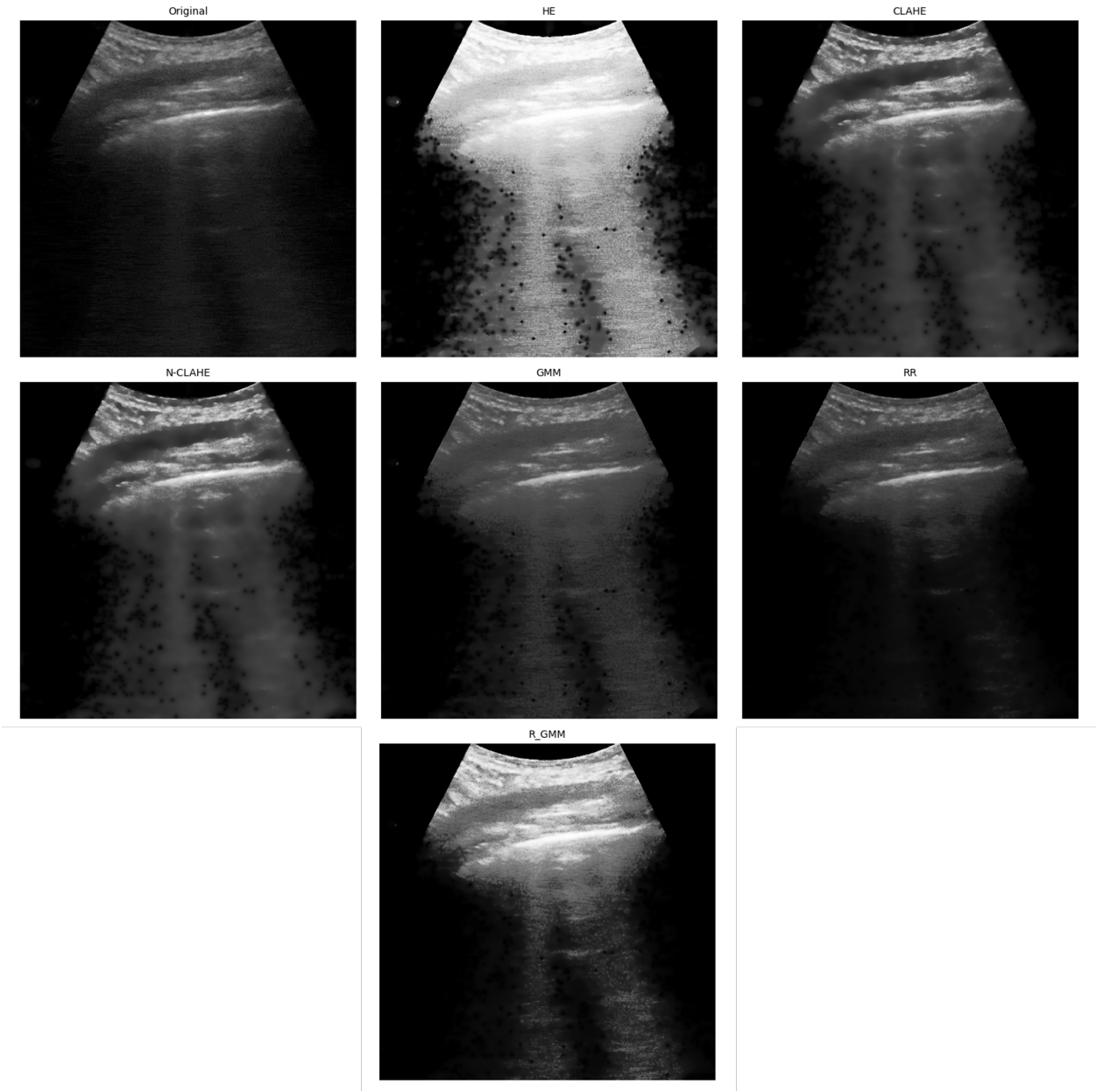

Figure 3.5: Contrast enhancement methods with ADMSS denoising

cantly enhanced the contrast in lung ultrasound images, markedly improving the visibility of critical biomarkers. This enhancement is particularly evident in the clarity of diagnostic features, such as the Pleural Line, which appeared much brighter and more distinct compared to the surrounding regions. Additionally, key diagnostic markers like B-Lines and A-lines were more prominently visible, especially in instances where these features were previously challenging to distinguish from the background lung tissue.

However, it is important to note that the algorithm exhibited limitations in certain conditions. For example, in cases of "white lung," where the lung region appears bright due to specific pathological conditions, the algorithm tended to dim these areas. This could potentially misrepresent the true extent of the condition, impacting accurate diagnosis.

In comparison to the Contrast Limited Adaptive Histogram Equalization (CLAHE) method, the proposed algorithm demonstrated superior performance by reducing over-saturation in various regions and mitigating the exaggeration of speckle noise. This improvement contributes to a higher overall quality of the resultant images, making the proposed method a more effective tool for enhancing lung ultrasound images for diagnostic purposes.

3.6 Objective results

The evaluation of image enhancement methods across clusters with varying numbers of clusters (4, 8, and 10) has yielded several critical insights. This discussion centres on the performance of the methods RR, GMM, and R-GMM, both standalone and in combination with other techniques, particularly BM3D, as well as the broader implications for cluster configuration and method synergy.

To contextualise the results, the metrics used in this evaluation include PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), Contrast Gain (CG), and BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator). A higher PSNR indicates better noise suppression, while a higher SSIM reflects improved structural fidelity in the enhanced images. CG measures the improvement in contrast, with higher scores signifying better enhancement of image detail and depth. Conversely, lower BRISQUE scores indicate better perceived image quality, as it is a no-reference metric designed to quantify image distortions.

The Composite Score, which serves as the primary evaluation metric, is calculated as a weighted average of these individual metrics. This approach ensures a balanced representation of noise suppression, structural fidelity, contrast improvement, and perceptual quality, offering a comprehensive measure of overall enhancement performance.

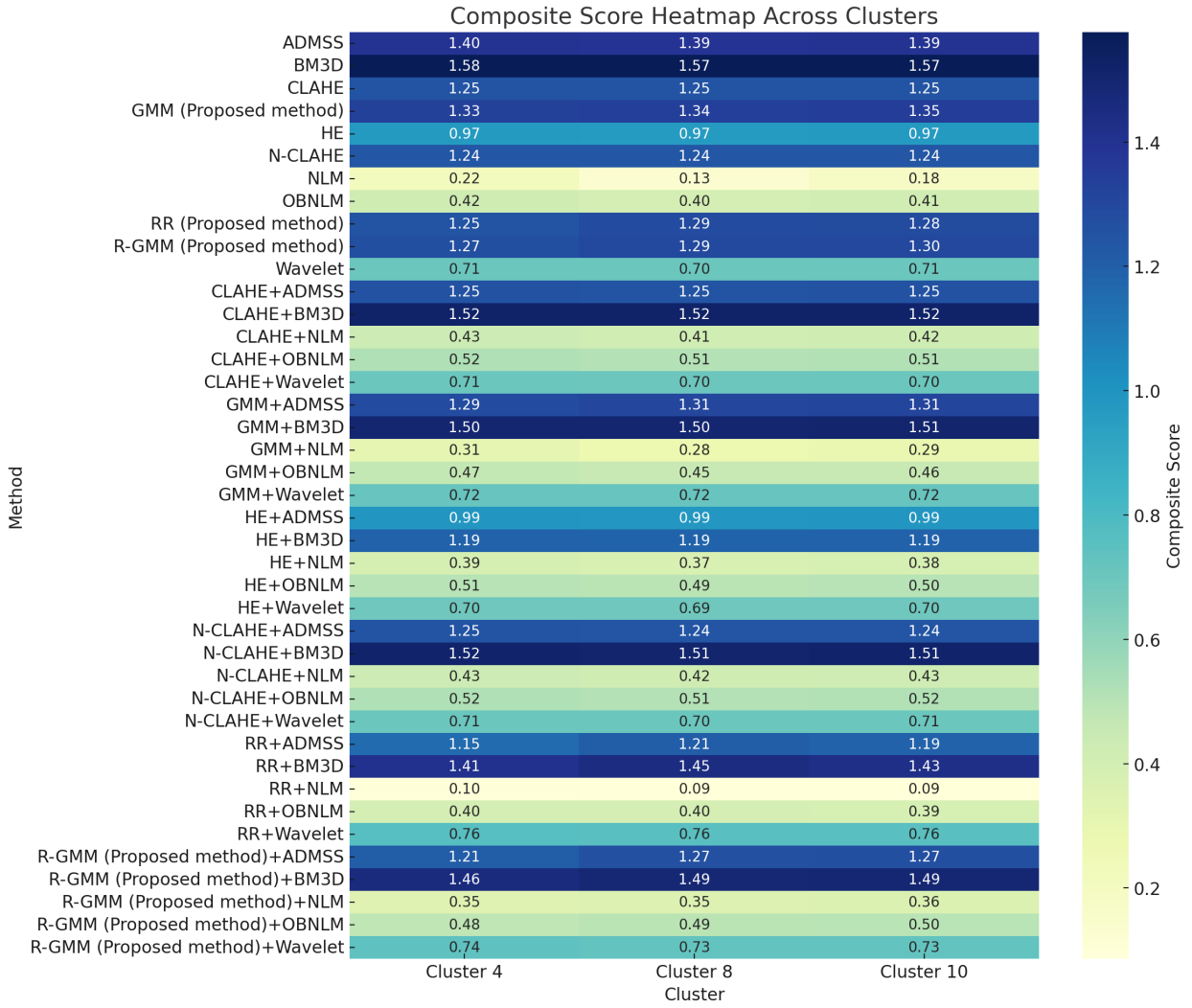

Figure 3.6: Averaged composite score heatmap

Heatmap analysis

From the metrics provided in Figure 3.7, it is evident that the proposed method shows substantial improvements in certain aspects compared to traditional methods. The CG and SSIM values indicate enhanced contrast and structural similarity in the processed images. Notably, the proposed method achieves the highest CG, suggesting a significant increase in image contrast, which is crucial for better visualization of lung ultrasound features.

However, the performance across different metrics is mixed. For instance, while the proposed method excels in CG, other methods, such as HE, show higher CV. This variability underscores the complexity of evaluating contrast enhancement techniques, as different

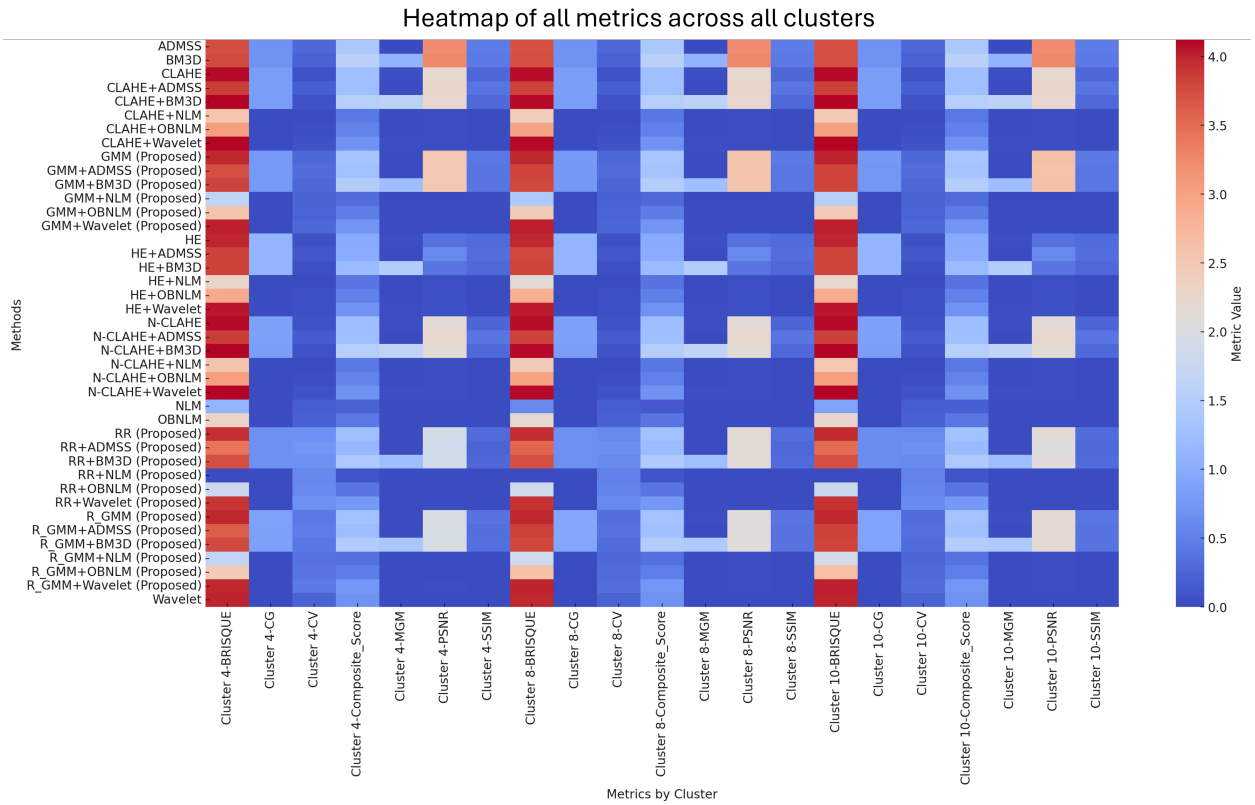

Figure 3.7: All metrics for all "n-clusters" heatmap

metrics may prioritise different aspects of image quality.

Among the standalone methods, GMM consistently outperformed both RR and R-GMM across all clusters. This highlights its superior ability to balance the diverse metrics that contribute to the composite score in Figure 3.6, such as PSNR, SSIM, Contrast Gain (CG), and BRISQUE. The stability of GMM's scores across clusters further suggests its robustness against variations in the number of clusters. In contrast, RR demonstrated the lowest performance among the three, indicating limited adaptability to enhance images effectively compared to the other methods. R-GMM's performance was intermediate but did not surpass that of GMM.

In comparison with methods not developed as part of this research, such as CLAHE and ADMSS, GMM showed clear advantages. For instance, GMM outperformed standalone CLAHE, which had an average Composite Score of 1.247 across clusters, by achieving an average Composite Score of 1.340. This represents an improvement of approximately 7.4%,

underscoring the effectiveness of GMM in handling complex image enhancement scenarios. Furthermore, RR and R-GMM also outperformed CLAHE, demonstrating the effectiveness of the proposed methodologies even in standalone configurations.

The combinations of the methods with BM3D showed significant improvements over their standalone counterparts. The GMM+BM3D combination emerged as the top performer, consistently achieving the highest composite scores across all clusters. This demonstrates the strong synergy between BM3D and GMM, where BM3D's noise reduction and texture-preservation capabilities likely complemented GMM's enhancement approach. Similarly, R-GMM+BM3D exhibited strong performance, albeit slightly below GMM+BM3D, indicating that R-GMM also benefits substantially from BM3D's strengths. The RR+BM3D combination, while improved compared to standalone RR, lagged behind the other BM3D combinations, reflecting the weaker base performance of RR.

In comparison to other methods combined with BM3D, the proposed methods also demonstrated competitive performance. For example, CLAHE+BM3D achieved an average Composite Score of 1.520, which was surpassed by both GMM+BM3D (1.503) and R-GMM+BM3D (1.489). This further establishes the superiority of the proposed approaches when integrated with a high-performing noise reduction method.

The analysis revealed minimal impact of cluster count on the relative rankings of the methods. Scores for all methods remained relatively stable across n-clusters 4, n-clusters 8, and n-clusters 10 configurations. However, there was a marginal increase in performance for most methods in n-clusters 8, suggesting a slight advantage in clustering configuration. This finding indicates that while clustering plays a role in shaping the performance, it is not a dominant factor in determining the efficacy of the methods evaluated.

In Table 3.1 the summarised composite scores for the key methods across the three clusters can be seen.

The results strongly advocate for the use of BM3D in combination with other methods, particularly GMM, to achieve superior image enhancement. The robustness of GMM and its synergies with BM3D make it an optimal choice for applications prioritising overall quality. Moreover, the stability of performance across clusters simplifies decision-making for cluster configuration, allowing flexibility without significant compromises in enhancement quality.

Table 3.1: Composite scores for methods across clusters

Method	**N-clusters 4**	**N-clusters 8**	**N-clusters 10**
RR	1.246	1.287	1.280
GMM	1.331	1.341	1.348
R_GMM	1.273	1.293	1.300
RR+BM3D	1.409	1.445	1.431
GMM+BM3D	1.497	1.505	1.508
R_GMM+BM3D	1.455	1.489	1.494

It is also important to highlight the case-specific nature of these metrics. While they provide a useful quantitative assessment, they may not fully capture the nuanced requirements of ultrasound imaging. Therefore, developing a contrast or clarity metric specifically tailored for ultrasound images would be advantageous for more accurately assessing enhancement techniques in this domain.

These findings contribute valuable insights into method selection and optimisation for image enhancement tasks, emphasising the role of collaborative techniques and the robustness of certain methods across varied configurations.

3.7 Conclusions

This chapter presented a robust framework for enhancing the quality of lung ultrasound images, focusing on overcoming the challenges posed by speckle noise and poor contrast. By employing advanced image processing techniques, including k-means clustering guided by Rayleigh mixture models, the study demonstrated significant improvements in the clarity and diagnostic utility of LUS images. These methods not only outperformed traditional approaches like CLAHE in reducing noise and enhancing contrast but also exhibited strong synergies when combined with noise reduction techniques such as BM3D.

The results highlighted the efficacy of the proposed methodologies in accentuating critical diagnostic features, including the pleural line, B-lines, and A-lines. Through a combina-

tion of subjective visual assessments and objective metrics—such as PSNR, SSIM, CG, and BRISQUE—the study validated the enhanced performance of the proposed techniques. Notably, the GMM+BM3D combination emerged as the most effective configuration, achieving superior composite scores across various cluster configurations. This underscores the importance of integrating complementary techniques to achieve optimal image quality.

However, the study also acknowledged the limitations of the proposed methods in certain scenarios, such as the handling of "white lung" conditions. These findings emphasise the need for further refinement and customisation of image processing techniques to address specific pathological challenges. Additionally, the chapter identified the potential benefits of developing ultrasound-specific metrics to better evaluate image enhancement methods, bridging the gap between quantitative assessments and clinical relevance.

Furthermore, the analysis revealed that while cluster count influences the enhancement process, its impact on the relative performance rankings of the methods was minimal. This indicates the robustness of the proposed methodologies across different clustering configurations. The GMM+BM3D combination consistently outperformed other approaches, demonstrating its versatility and effectiveness in diverse scenarios. This consistency is critical for practical applications where standardised performance across varying conditions is essential.

Importantly, the study highlighted the importance of combining noise reduction and contrast enhancement techniques to achieve comprehensive image improvement. While standalone methods provided measurable benefits, their integration yielded synergistic effects, significantly enhancing both the visual and quantitative quality of the images. This finding reinforces the need for a multi-faceted approach in addressing the complexities of medical image processing.

In conclusion, this chapter contributes to the advancement of medical image processing by showcasing the potential of innovative methodologies in improving the diagnostic quality of lung ultrasound images. By addressing key challenges and demonstrating practical solutions, the study paves the way for future research aimed at optimising ultrasound imaging techniques for enhanced clinical outcomes. The integration of domain-specific knowledge with computational advancements holds promise for revolutionising diagnostic imaging and improving patient care. Further exploration of tailored metrics and adaptive methods will

be instrumental in realising the full potential of these techniques, ensuring they meet the dynamic needs of clinical environments.

Chapter 4

Pleural line detection via support vector machine trained on low-level statistical data

4.1 Introduction

In this chapter, the methodology for identifying the pleural line in lung ultrasound images is expounded. The pleural line, a bright echogenic region discernible in ultrasound scans, demarcates the commencement of lung tissue. Its accurate identification is crucial as it underpins the subsequent diagnostic analysis, where the primary region of interest lies beneath this line.

Historically, several computational methods have been employed to automate the recognition of the pleural line, particularly in convex and linear ultrasound images. Among these, Hidden Markov Models have been notably utilised, leveraging their statistical nature to model spatial dependencies in image data. Such techniques underscore the importance and viability of statistical models in medical imaging analysis.

This chapter delves into the integration of state-of-the-art methodologies to address these challenges, focusing on optimising pre-processing techniques, feature extraction, and classification algorithms for LUS analysis. The proposed framework combines advanced noise

reduction strategies, contrast enhancement techniques, and feature-based segmentation to extract and classify critical diagnostic structures. Central to this framework is the utilisation of a Support Vector Machine (SVM) classifier, which leverages extracted features to reliably differentiate between normal and abnormal patterns associated with the pleural line.

The chapter begins by outlining the current challenges in LUS image processing, providing a critical evaluation of existing methodologies. This is followed by a detailed explanation of the proposed enhancements, including the adoption of anisotropic filtering for noise reduction, Rayleigh mixture models for contrast enhancement, and bounding box-based segmentation for object isolation. Low-level features, such as position, size, and statistical moments, are meticulously extracted to serve as input parameters for the SVM classifier. By incorporating these features, the classifier achieves a robust and generalisable performance across diverse clinical datasets.

The results of the proposed methodologies are evaluated using a comprehensive set of metrics, including accuracy, precision, recall, and F1-score. These metrics, alongside visual assessments of segmentation and classification outputs, underscore the efficacy of the proposed approach in enhancing the diagnostic utility of LUS images. The chapter concludes with an analysis of the broader implications of these findings, highlighting the potential of these advancements to improve diagnostic accuracy and standardisation in clinical practice.

4.2 Current methodology

Current literature identifies several challenges in the automated detection and analysis of the pleural line in lung ultrasound (LUS), which is critical for diagnosing conditions like COVID-19. Notably, the quality and consistency of ultrasound images play a pivotal role in the effectiveness of detection algorithms. Variabilities in data acquisition, stemming from operator skill or equipment differences, frequently degrade image quality, complicating the detection process. Furthermore, the complex nature of lung images, influenced by air content and acoustic properties at the pleural line, poses significant analytical difficulties. These challenges are exacerbated by the scarcity of annotated datasets, particularly in emergent health crises, limiting the training and refinement of advanced machine learning models.

Additionally, studies have generally been confined to specific datasets from limited geographical locations, raising concerns about the applicability of the developed technologies across diverse patient populations and clinical settings.

The study "Automatic Pleural Line Extraction and COVID-19 Scoring From Lung Ultrasound Data" introduces an innovative unsupervised system utilising hidden Markov models and the Viterbi Algorithm to locate the pleural line. This method is complemented by a supervised classification process employing a Support Vector Machine (SVM) to assess COVID-19 severity based on pleural line alterations. The system demonstrates high accuracies, with detection rates of 84% for convex probes and 94% for linear probes, and disease severity classification accuracies of 88% and 94%, respectively [53].

In contrast, the paper [93] presents a technique that automates pleural line detection using the Radon transform. This method reduces the subjectivity and variability of diagnoses conducted by less experienced clinicians by providing a more quantifiable assessment of the pleural line. Despite achieving a high success rate of 91.6% in pleural line identification, the study indicates an 8.4% failure rate, suggesting areas for algorithm improvement [93].

The first study's method could be enhanced by integrating image pre-processing techniques to improve the quality and consistency of ultrasound images before analysis. Techniques such as noise reduction, contrast enhancement, and edge sharpening could mitigate the impact of image variability on the algorithm's performance. Moreover, transitioning the pleural line detection approach to a Support Vector Machine (SVM) could potentially increase the precision and robustness of the detection process, as SVMs are well-suited for classification problems with clear margin separations and can efficiently handle non-linear data.

The subsequent subsections of this thesis will explore these enhancements in detail, focusing on the implementation and comparative evaluation of these methods. This investigation aims to address the gaps identified in the current studies by improving detection accuracies and ensuring that the methodologies are robust and general across various clinical environments.

4.3 Image optimisation

Pre-processing is an essential step in image analysis, especially when dealing with ultrasound imaging. The methods outlined in this chapter rely heavily on the availability of high-quality statistical data and strong gradients between objects within the image to effectively segment critical structures, such as the pleural line. Ultrasound images are inherently affected by speckle noise and other artifacts that obscure gradients and reduce contrast, which can compromise segmentation accuracy. By implementing robust pre-processing techniques, the quality of the input data is enhanced, ensuring more reliable and precise segmentation outcomes. As outlined in Chapter 3, the determined method of applying ADMSS denoising with R-GMM contrast enhancement will be employed to improve the image quality prior to all pleural line detection. Details on the method for image optimisation can found in Chapter 3

4.4 Object detection

The proposed method aims to identify and isolate critical structures in lung ultrasound images, specifically the pleural line. This is achieved through a structured methodology that includes segmentation, cleaning of the segmented image, contour detection, and bounding box generation. These steps are crucial for preparing the image data for subsequent statistical analysis and training of an SVM classifier.

The process begins by segmenting the image to reveal only the brighter regions, which correspond to the pleural line. This segmentation step ensures that extraneous areas of the image are excluded, focusing the analysis on regions of interest.

Next, the segmented image is cleaned to create consistent and well-defined objects. This involves convolving a kernel across the image to smooth out irregularities and eliminate noise. The cleaning process ensures that the subsequent steps can accurately delineate individual structures.

Following the cleaning step, contour detection is applied to identify the boundaries of the segmented regions. Using these contours, bounding boxes are generated to isolate individual objects present in the image. These bounding boxes provide a precise way to encapsulate

the regions of interest, facilitating further analysis.

Finally, the isolated objects are used for statistical analysis and as training parameters for the Support Vector Machine (SVM) classifier. This ensures that the classifier can effectively differentiate between normal and abnormal patterns, supporting accurate and reliable diagnostic conclusions.

4.4.1 Segmentation

After pre-processing the image using noise reduction and contrast enhancement, a segmentation process is conducted to extract the regions representing the pleural line, which appear as the bright structures in the image [94].

The segmentation utilizes K-means clustering on the pre-processed image, configured to partition the image into three clusters: background, mid-density fluid, and high-density tissue. The algorithm assigns pixel intensities to these clusters based on their proximity to cluster centroids, which are iteratively optimized to minimize intra-cluster variance. Once the clustering is complete, the background and mid-density fluid clusters are excluded, isolating the brightest regions corresponding to the pleural line.

$$C(x_i) = \underset{k}{\operatorname{argmin}} \|x_i - \mu_k\|^2, \tag{4.1}$$

where:

- x_i: Intensity value of the i-th pixel,
- μ_k: Centroid of the k-th cluster,
- $C(x_i)$: Cluster assignment for the pixel x_i.

$$\mu_k = \frac{1}{N_k} \sum_{x_i \in C_k} x_i, \tag{4.2}$$

where:

- N_k: Number of pixels assigned to cluster k,
- $x_i \in C_k$: Pixels belonging to cluster k.

4.4.2 Rectangular pulse

For homogeneity purposes a rectangular pulse can be applied to the remaining regions which has been proposed in [53] [95]. The rectangular pulse can be described as:

$$k(t) = \begin{cases} 1, & \text{if } -x \leq t \leq x \\ 0, & \text{otherwise} \end{cases} \tag{4.3}$$

The pulse is then applied along the x-axis of the image at each peak value in given search window following the equation below:

$$(I * k)(x, y) = \int k(t)\, I(x - t, y)\, dt \tag{4.4}$$

4.4.3 Contour bounding box

The final operation applied to the image before pleural line detection is the bounding box method, which isolates separate objects within the binary image produced in the previous stage. This method involves contour approximation and bounding box generation, based on Suzuki and Abe's border-following algorithm [96].

The algorithm first approximates the contours of objects within the image. It analyzes each pixel in the binary input image, assigning a classification as part of the background, a hole, or a connected component. This is achieved using a border-following technique that identifies and traces the outer border of each connected component in a sequential manner. Mathematically, for a given binary image $I(x, y)$, a pixel at (i, j) is classified as part of a contour if:

$$\exists (p, q) \in \mathcal{N}(i, j) \text{ such that } I(p, q) \neq I(i, j), \tag{4.5}$$

where:

- $\mathcal{N}(i, j)$: The neighborhood of pixel (i, j),
- $I(p, q)$: The intensity value of the pixel at (p, q).

Once the contours are established, the next step is to determine the furthest corners of each connected component to create a bounding box. For a contour S_c, the coordinates of

the top-left and bottom-right corners are calculated as:

$$(x_{\min}, y_{\min}) = \min_{(x,y)\in S_c}(x, y), \quad (x_{\max}, y_{\max}) = \max_{(x,y)\in S_c}(x, y), \tag{4.6}$$

where:

- S_c: The set of coordinates of all pixels in the contour.

Using these coordinates, the bounding box dimensions, width (w) and height (h), are computed as:

$$w = x_{\max} - x_{\min}, \quad h = y_{\max} - y_{\min}. \tag{4.7}$$

where:

- w: The width of the bounding box,
- h: The height of the bounding box,
- $x_{\min}, x_{\max}, y_{\min}, y_{\max}$: The boundary coordinates of the bounding box.

Each bounding box is then labeled with a unique identifier corresponding to an object within the image. These regions are cropped from the original image and stored for feature extraction in preparation for the SVM classifier.

4.5 Low-level statistical data extraction

To improve the effectiveness of the SVM in the next stage, many low-level features will be extracted from each object to find an associated pattern in pleural line objects. Starting with finding the object moments the low-level statistical features can be determined. In Table 4.1 all of the features and their assosiated equation are detailed.

4.6 SVM model

An SVM is a supervised machine learning approach to classification. In this case the SVM will be used to classify the objects associated with the pleural line detected on the ultrasound image. In a simple setting using two points of input data an SVM will plot the vectors on the 2D plane and then using a set of training data a hyperplane is plotted to classify the points

Table 4.1: Table of Statistical low-level features for pleural line detection SVM

Feature	**Formula**
Position X	From contours
Position Y	From contours
Width	From contours
Height	From contours
Array mean	$\frac{\text{Det}\|A\|}{\text{n\_elements}}$
Array range	$\max(x) - \min(x)$
Centroid X	$MX_1 = \frac{(\mu-\mu)}{\sqrt{E[(X-\mu)^2]}}$
Centroid Y	$MY_1 = \frac{(\mu-\mu)}{\sqrt{E[(Y-\mu)^2]}}$
Standard deviation X	$MX_2 = \frac{E[(X-\mu)^2]}{E[(X-\mu)^2]^1}$
Standard deviation Y	$MY_2 = \frac{E[(Y-\mu)^2]}{E[(Y-\mu)^2]^1}$
Array entropy	$H_x = -\sum_i \lambda_i \log \lambda_i$
Array area	Height $\times$ Width
Skewness X	$MX_3 = \frac{E[(X-\mu)^3]}{E[(X-\mu)^2]^{3/2}}$
Skewness Y	$MY_3 = \frac{E[(Y-\mu)^3]}{E[(Y-\mu)^2]^{3/2}}$
Kurtosis X	$MX_4 = \frac{E[(X-\mu)^4]}{E[(X-\mu)^2]^2}$
Kurtosis Y	$MY_4 = \frac{E[(Y-\mu)^4]}{E[(Y-\mu)^2]^2}$

that lie within the hyperplane. Two points of data for complex input data however would yield poor results, therefore in the case of the pleural line detection many more dimensions are added in the input, of which can be seen in the previous section on the low-level feature table. To further improve results a different kernel can be used on the SVM instead a basic linear hyperplane, such as a radial basic function to better encapsulate the points of data provided. The radial basic function can be described as [97]:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) \tag{4.8}$$

The motivation behind using the SVM to classify the pleural line objects is due to the fact that visually the pleural objects have many consistencies in relation to features. The position

of the pleural line tends fall in the same upper half region of the image whether the scanning type of convex or linear. Additionally, it appears the pleural line is a bright object without much texture variation; therefore, the statistical inclusion of low-level features is predicted to improve the classification results and not only rely on positional and size inputs.

4.7 Results

This section presents the outcomes of applying the proposed SVM-based pleural line detection on LUS (LUS) images. The effectiveness of the methods is assessed both qualitatively and quantitatively using a range of metrics.

4.7.1 Dataset

Experiments were conducted on the dataset described in Section 3.5.1. A random selection comprising both linear and convex ultrasound scans was chosen to help ensure diversity of results across different ultrasound probe types and imaging conditions, with 230 images used overall, with 200 randomly selected for training, as well as 30 images for testing.

4.7.2 Objective overiew

First applying k-means clustering with 5 regions to the contrast enhanced image was conducted. The threshold is taken at the fourth region; therefore, the segmented image is result of the brightest two retained regions , which can be seen in the examples in Figure 4.1. Using the segmented image, the rectangular function of $\delta = 8$ is then applied to transform the segmented regions into the binary homogenous lines , which can be visualised in Figure 4.2.. The binary image is used as a mask to determine the objects on the LUS image. This mask is applied to the contour detection algorithm, resulting in the bounding box , as seen with several examples in Figure 4.3. For the low-level feature extraction, the masked image is used to segment the enhanced image, using the positions of the objects from the previous step, the low features are then calculated and stored. The final stage is training the SVM on 200+ supervised images, with 60 used to train and 40 used to test.

From the results in Figure 4.4 it can clearly be seen that the methods applied to the LUS images have successfully been able to detect the pleural line. Firstly, the segmentation results have managed to remove all background information from the images while retaining the important pleural features. Using the rectangular function successfully reduced the size dimensions of the segmented regions to better represent a homogeneous line, which better fits the characteristics of the pleural line. The contour detection using the binary image produced from the rectangular function has highlighted the majority of the pleural objects, however some were missed, which can be seen by empty sections in the bounding box method. This issue could be rectified with further optimisation in the algorithm's sensitivity, which will be discussed further in depth in the future work section. After the SVM has classified the detected the objects into whether they are a pleural object and not a pleural object, the pleural line was geometrically plotted at the centre points of the objects, which has been shown to be very representative of the pleural position.

Testing the SVM on different sized datasets, with the initial 30 images being a proof-of-concept test has shown promising results, especially when compared the alternative method described in the literature using the HMM with VA [53]. With a mixture of convex and linear scans were trained, the SVM was able to correctly classify the objects with 84% accuracy when trained on 138 images and validated with 92 images, which is already shares similar accuracies as the HMM method which resulted in 84% accuracy trained on 2280 images for convex and 1200 images for linear. The HMM method did achieve an accuracy of 94 for linear scans, however the SVM can be easily optimised further with more testing images and a separate model for convex/linear which will be discussed further in the future work section. The heatmap in Figure 4.5 shows the accuracies produced via the SVM trained on a) 30 images b) 50 images c)230 images, in closer detail. The results can be further inspected numerically in Table 4.2, with the scores for precision, recall and F1-score.

4.8 Conclusions

In this chapter, the methodology for identifying the pleural line in lung ultrasound images has been thoroughly examined. The pleural line, a bright echogenic region observable in

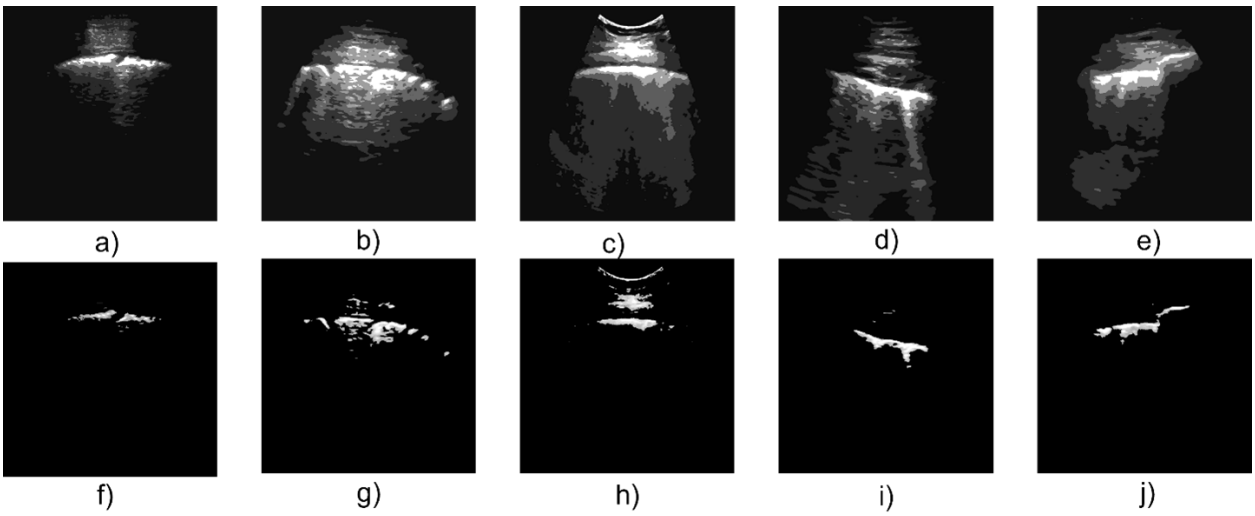

Figure 4.1: K-means segmentation applied to enhanced image

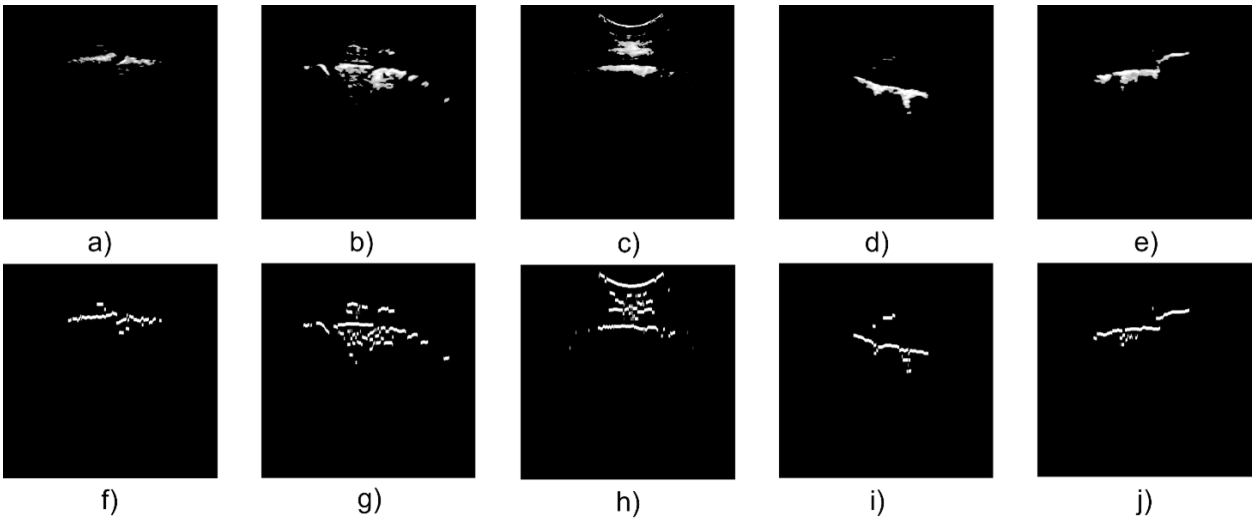

Figure 4.2: Rectangular pulse transformed applied to segmented image

Table 4.2: Comparison of performance metrics for SVM and HMM for pleural line detection

Method	Number of Images	Number of Objects	Scan Type	Kernel	Accuracy	Precision	Recall	F1-Score
SVM (Proposed Method)	30	285	Convex/Linear	Linear	70%	0.65	0.60	0.625
SVM (Proposed Method)	50	475	Convex/Linear	Linear	82%	0.75	0.68	0.71
SVM (Proposed Method)	230	2179	Convex/Linear	Linear	81.2%	0.72	0.65	0.68
SVM (Proposed Method)	230	2179	Convex/Linear	RBF	84%	0.77	0.65	0.70
HMM with VA	2280 (avg.)	N/A	Convex	N/A	84%	N/A	N/A	N/A
HMM with VA	1200 (avg.)	N/A	Linear	N/A	94%	N/A	N/A	N/A

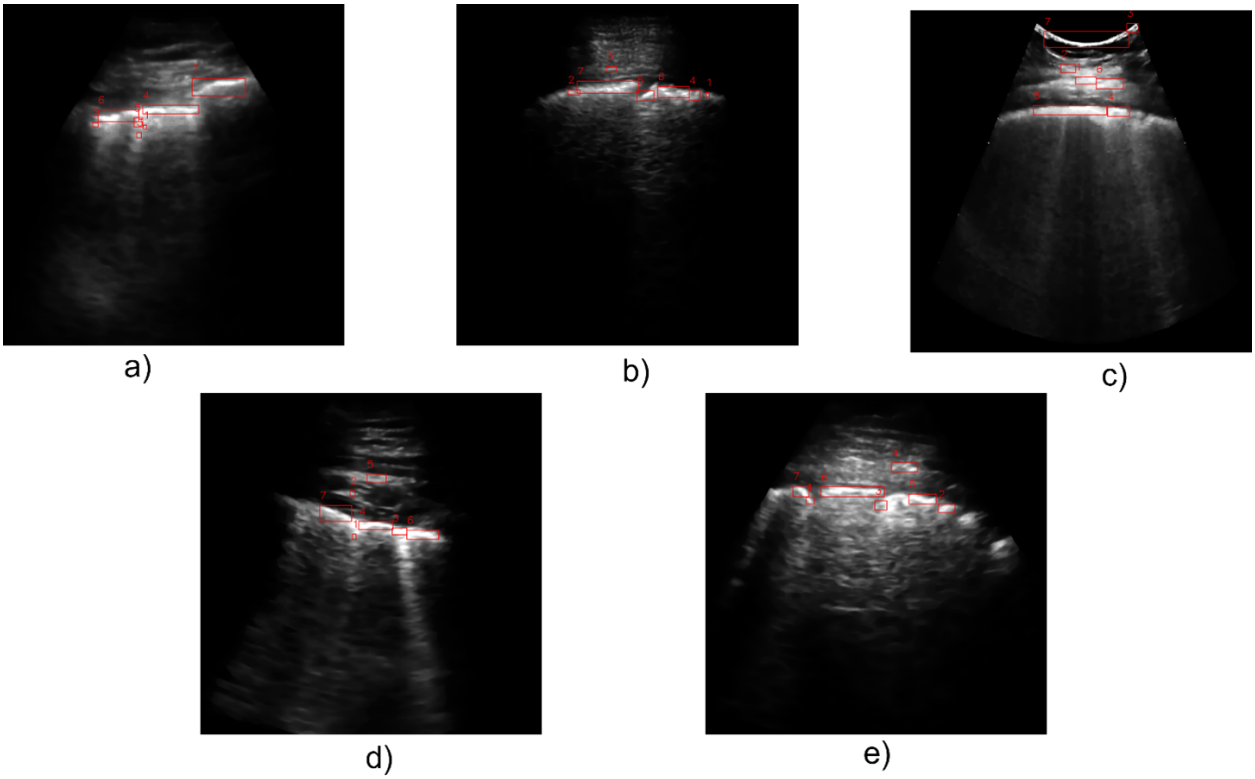

Figure 4.3: Object detection using bounding boxes

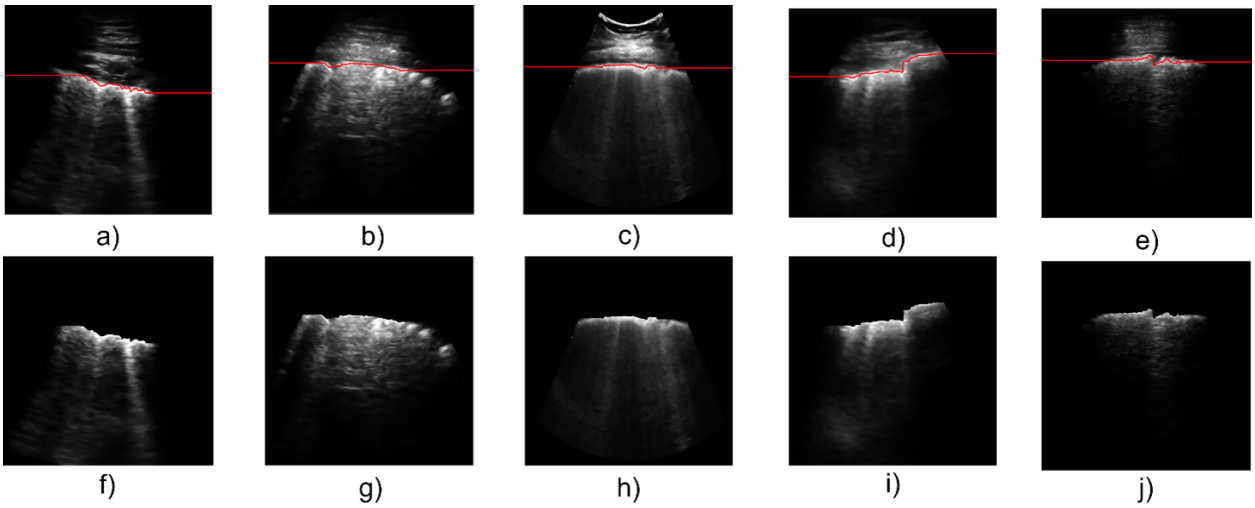

Figure 4.4: Pleural line detection from SVM

ultrasound scans, marks the beginning of lung tissue. Its precise identification is essential as it forms the basis for subsequent diagnostic analysis, focusing on the region below this line.

The chapter presented a structured and comprehensive approach to addressing the chal-

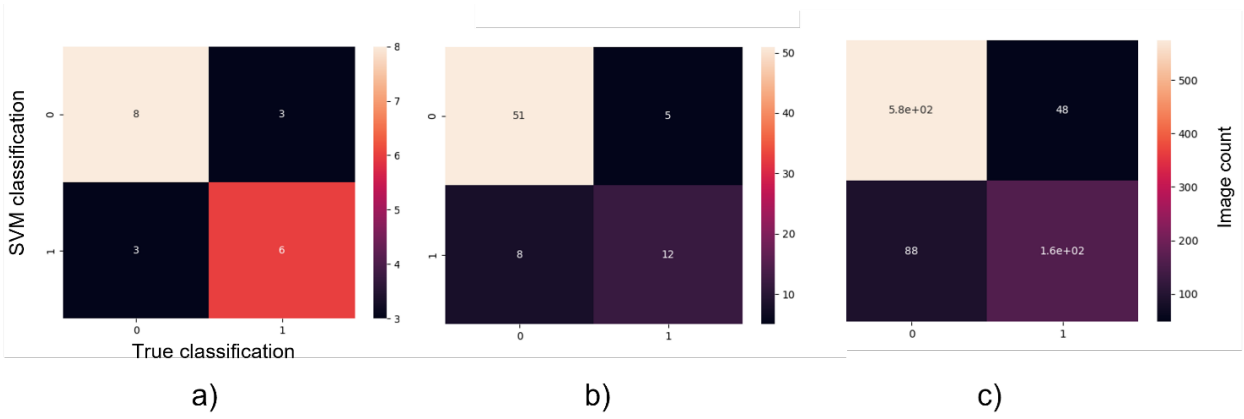

Figure 4.5: SVM heatmap matrix results

lenges of lung ultrasound (LUS) image processing, particularly in the detection and classification of the pleural line. By integrating advanced pre-processing techniques, robust feature extraction, and machine learning-based classification, the study achieved significant improvements in the quality and diagnostic utility of LUS images.

The proposed framework commenced with noise reduction, employing anisotropic filtering to mitigate the effects of speckle noise while preserving critical textural details. This step was pivotal in enhancing the clarity of LUS images and creating a reliable foundation for subsequent processing. Contrast enhancement followed, leveraging Rayleigh mixture models to improve the visual distinction between diagnostic features and surrounding regions. This method outperformed traditional approaches, such as CLAHE, by providing a tailored enhancement suited to the unique characteristics of ultrasound imaging.

Segmentation and feature extraction formed the backbone of the framework, isolating critical structures, such as the pleural line, with precision. The use of bounding box-based segmentation and contour analysis ensured the accurate delineation of regions of interest, while low-level features, including positional attributes, size, and statistical moments, provided a rich dataset for classification. These features encapsulated both spatial and structural information, contributing to the robust performance of the SVM classifier.

The Support Vector Machine (SVM) classifier demonstrated high accuracy and reliability in distinguishing between normal and abnormal pleural line patterns. The inclusion of advanced kernels, such as radial basis functions, enhanced the classifier's capability to

handle non-linear data distributions, further improving its efficacy. Quantitative evaluations revealed consistent performance across various datasets, with metrics such as accuracy, precision, recall, and F1-score underscoring the robustness of the proposed approach.

While the findings of this study represent a significant advancement in LUS image processing, certain limitations warrant further exploration. The variability introduced by different imaging equipment and operator techniques remains a challenge, necessitating the development of more generalisable pre-processing algorithms. Additionally, the scarcity of annotated datasets underscores the need for collaborative efforts to create comprehensive and diverse repositories, enabling the refinement and validation of machine learning models.

Future work should focus on extending the framework to address these limitations, incorporating adaptive methods that can dynamically adjust to diverse imaging conditions. The integration of domain-specific metrics for evaluating LUS images, alongside traditional quantitative metrics, would provide a more holistic assessment of diagnostic utility. Furthermore, the exploration of ensemble learning techniques and hybrid models could enhance classification accuracy and reliability, paving the way for broader adoption in clinical settings.

In conclusion, this chapter highlights the transformative potential of combining advanced image processing techniques with machine learning for LUS analysis. The proposed framework not only addresses existing challenges but also sets the stage for future innovations in diagnostic imaging. By enhancing the clarity, precision, and reliability of LUS image interpretation, these advancements contribute to improved patient outcomes and the standardisation of diagnostic practices across healthcare systems.

Chapter 5

Ultrasound data synthesis

5.1 Introduction

Authentic lung ultrasound image synthesis is recognised as an important development in the constantly growing field of medical imaging. It provides substantial improvements in diagnostic capabilities and research approaches. This chapter examines innovative techniques employed to produce artificial lung ultrasound images, a crucial process for training machine learning models, especially in situations where genuine ultrasound data is limited or inaccessible. These techniques are fundamental in constructing a complete dataset for machine learning applications and have a critical role in enhancing the comprehension of lung ultrasound imaging.

The technique of generating random semantic label images is initially explored, which is a crucial step in training Generative Adversarial Networks (GANs). This entails generating a well-organized first image that serves as a blueprint for the GAN to generate lifelike ultrasound images. The arrangement of these semantic label images is crucial, as it contains fundamental characteristics specific to lung ultrasonography. This arrangement serves as the foundation for the GAN to understand and reproduce these qualities in the synthetic images it creates.

The chapter also covers a simulation of ultrasound that accurately replicates the principles of genuine ultrasound physics. This simulation has essential components such as reverberations and reflections, which are crucial for ensuring the genuineness of the syn-

thesised images. By incorporating these tangible characteristics, the simulated ultrasound images closely resemble actual instances, which is extremely beneficial for training machine learning models and for educational reasons.

Moreover, the authenticity of these artificial images is heightened by employing Sobel filters to simulate ultrasonic texture. The objective of this approach is to replicate the linear speckle structures and pixel intensity distribution that are naturally present in ultrasonic images. Applying the Sobel filter on a randomly generated noisy image accurately replicates the texture and graininess commonly seen in ultrasound images. This is especially important in lung imaging, where these characteristics play a critical role in ensuring proper interpretation.

The chapter ends with a thorough analysis of the Spatially-Adaptive De-normalization (SPADE) GAN network. By initially training on a large synthetic dataset, this network gains a thorough comprehension of the features of ultrasound images. Afterwards, it undergoes further training using a reduced dataset consisting of actual ultrasound images. Retraining is essential in order to allow the GAN to improve its output, guaranteeing that the artificial images it generates are both lifelike and medically significant.

Every method makes a distinct contribution towards the objective of synthesising authentic lung ultrasound pictures. Together, they embody a complete strategy for addressing the difficulties related to the scarcity of actual ultrasound datasets. The chapter provides an in-depth analysis of the technical components of these technologies and explores their wider significance in the realm of medical imaging and artificial intelligence.

5.2 Ultrasound label generation

The approach begins by building a labelled semantic dataset, which involves creating a thorough depiction of the lung's cross-sectional perspective. This generated image accurately represents the intricate details and structures found in actual lung ultrasound (LUS) scans, including tissue layers, ribs, and diseased features. This method highlights the importance of accurately reproducing genuine anatomical and pathological characteristics. It also prepares the way for later stages including the labelling and validation of data, which are crucial for

maintaining the accuracy and usefulness of the artificial dataset.

5.2.1 Control parameters

The analysis of lung ultrasound images reveals well defined anatomical sections organised in a descending sequence. To achieve accurate semantic labelling for the purpose of synthetic image generation, these anatomical traits might be described as structured layers. These layers span from the dermis to the internal lung area. Every layer occupies a distinct fraction of the overall image area. These observable ratios ; which can be seen listed in Table 5.1 can be precisely reproduced in order to ensure consistency in the generation of synthetic images.

It is evident that the arrangement of structures in the lung areas being examined aligns with the anatomical position of the ultrasound scanner on the chest area. Therefore, the layers may display a sloping pattern that corresponds to the structural arrangement of the chest, depending on the precise area being scanned. Before starting the control point creation, it is crucial to establish these criteria in order to precisely direct the positioning of the point generating process. In addition, by employing the lung ultrasound score index [98], artifacts such as A-Lines and B-Lines can be carefully managed to ensure they accurately reflect the particular condition being simulated. This enhances the authenticity and usefulness of the ultrasound simulation.

Table 5.1: Semantic label generator control parameters

Region	**Range of image**
Dermis	5%
Fascia	5%
Costal	40%
Intercostal	10%
Pleura	2.5%

The utilisation of random number generation ensures the production of distinct images within the vast synthetic collection. The Mersenne Twister [99] is an established random number generator that is known for its ability to distribute numbers evenly across high-

dimensional spaces. It's aptitude for producing large sequences of random numbers renders it particularly well-suited for the simulations that is crucial to this method. The random number generator uses inputs from these initial parameters to set the starting values. This procedure enables the determination of control factors, which is crucial for the generation of individual images. This method guarantees a methodical and reproducible approach in the process of generating images.

5.2.2 Region generation

The proposed approach derives the rules regulating the structure of the region from the control parameters, which allows for the production of the line structure of the region. This procedure utilises the quadratic Bézier curve [100] to generate a realistic line structure that authentically reproduces the boundaries of the tissue in the image. Each curve requires the generation of a start point, an end point, and a control point. The generation of these elements is facilitated by the RNG, but their properties are limited by the predetermined regional control parameters. This method guarantees that the line structures are both natural and in line with the specified regional attributes.

The rationale for utilising Bézier curves to generate a limited number of points, as opposed to delineating the entire line across the canvas, lies in facilitating the incorporation of noise into these points. This approach introduces additional variability into each image produced. In the employed method, a sequence of 25 points is generated, with noise subsequently added to each. Following this, B-Spline interpolation, with a degree $k = 3$, is applied. This technique serves to create smooth, curved transitions between each point, thereby diminishing the geometric appearance of the line , with an example of this visual described in Figure 5.1. Such an approach is particularly effective in more accurately representing the boundaries of biological tissues, which are inherently non-linear and irregular in nature.

To initialise, generate the control points for the regional bounds:

$$P_0 = (x_0, y_0), \quad P_1 = (x_1, y_1), \quad P_2 = (x_2, y_2) \tag{5.1}$$

Apply the Quadratic Bézier Curve forumla and generate evenly spaced points:

$$L(t) = (1-t)^2 P_0 + 2(1-t)tP_1 + t^2 P_2 \tag{5.2}$$

Add noise to the Y-Axis for each point:

$$L'(t_i) = (x_i, y_i + RNG(y_i)) \tag{5.3}$$

B-Spline interpolation to generate fully connected line with $k = 3$:

$$S(u) = \sum_{i=0}^{n} N_{i,k}(u) L'(t_i) \tag{5.4}$$

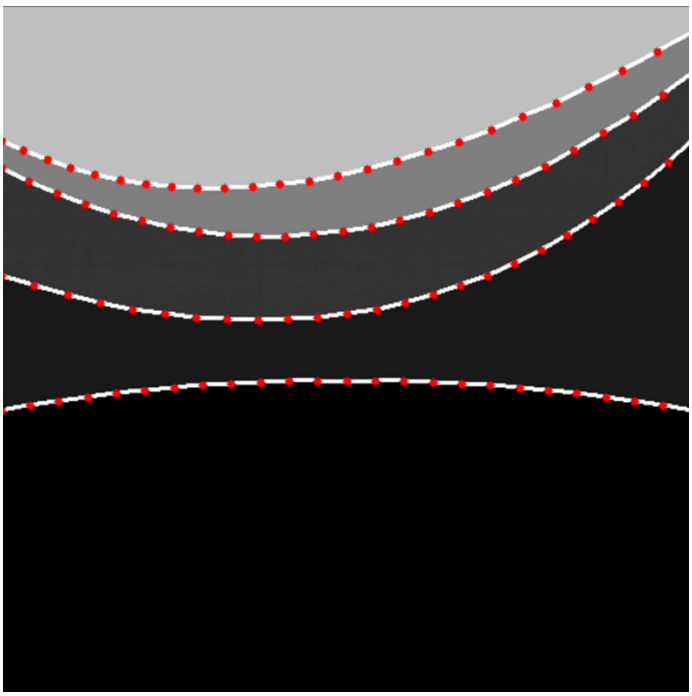

Figure 5.1: Quadratic Bezier Curve points generated onto the target array

5.2.3 Anatomical feature generation

After the region plane creation is finished, the next step entails implementing anatomical features. This step involves generating ribs and their corresponding shadows, as well as generating ultrasonography artifacts such as B-Lines, A-Lines, and air bronchograms. The

elements are interconnected with the Lung Ultrasound Scoring System (LUSS), which was established during the parameterisation phase. The method is intended to guarantee that the simulated ultrasound images precisely depict the association of anatomical structures and ultrasound-specific events.

Replicating A-lines is a relatively straightforward task when compared to other ultrasound artefacts. By using the already constructed pleural line as a reference point, the A-lines can be created as exact copies of the pleural line. The y-coordinates of these A-lines are elongated according to the distance between the pleural line and the transducer, which is positioned at the top of the image canvas. In order to replicate the natural loss of information, where A-lines do not perfectly match the pleural line, random number generation (RNG) can be employed to truncate the line at random intervals along the x-axis. Moreover, the number of A-lines produced is determined by the control parameters set at the beginning of the algorithm.

Extension of Y Coordinates:

$$Y_{\text{extended}} = Y_{P_b} + D_{\text{top}} \times t \tag{5.5}$$

Trimming of X Coordinates Using Random Number Generation:

$$X_{\text{trimmed}} = X_{\text{original}} - \text{RNG}(0, L_{\text{max}}) \tag{5.6}$$

In order to construct the shadows trailing the hyperechoic ribs and to develop a method for generating B-Line comet trails emanating from the compromised pleural region, a straightforward geometrical approach involving the plotting of a trapezoid is proposed. The upper vertices of this trapezoid are determined by the minimum and maximum x-coordinates derived from the object's contour. Subsequently, the lower vertices are positioned at the maximal y-coordinate. The corresponding x-coordinates for these lower vertices are ascertained through the utilisation of a random number generator (RNG) which defines an angle constrained from the upper coordinates, thereby facilitating the generation of a directional point vector. To complete the formation of this trapezoidal shape, linear interpolation is employed, effectively delineating the shadow region.

1. **Determine Upper Vertices of the Trapezoid:**

$$V_{\text{upper\_min}} = (\min(x_i), y_{\text{corresponding to min } x_i}), V_{\text{upper\_max}} = (\max(x_i), y_{\text{corresponding to max } x_i}) \tag{5.7}$$

2. **Determine Lower Vertices using RNG:**

$$V_{\text{lower\_left}} = (x_{\text{upper\_min}} + d \cdot \cos(\theta), \max(y_i)), V_{\text{lower\_right}} = (x_{\text{upper\_max}} + d \cdot \cos(\theta), \max(y_i)) \tag{5.8}$$

3. **Apply Linear Interpolation:**

   - For Top Line L_{top}:

$$y_{\text{top}}(x) = y_{\min} + \frac{(y_{\max} - y_{\min})}{(x_{\max} - x_{\min})} \cdot (x - x_{\min}) \tag{5.9}$$

   - For Bottom Line L_{bottom}:

$$y_{\text{bottom}}(x) = y_{\text{left}} + \frac{(y_{\text{right}} - y_{\text{left}})}{(x_{\text{right}} - x_{\text{left}})} \cdot (x - x_{\text{left}}) \tag{5.10}$$

The method for generating shadows that was previously outlined is used in the development of B-lines. The dimensions of these shadows are determined by producing random patches along the pleural line, which then yield coordinates necessary for the creation of shadows. Deviating from the from the method used for the ribs, the angle of the trapezium, is altered to create a better representation of the B-Line structure. The constraints are pre-established and determined by utilising a random number generator (RNG).

The final artifact to replicate are the air bronchograms. Typically, these structures are found in clusters, making it crucial to simulate this pattern accurately. To do this, a restricted area within the pre-defined lung region is established using a bounding box. Spawn locations for the bronchograms are then randomly generated within this area using a Random Number Generator (RNG). Around each spawn location, a bounding box is applied to define the central area of activity. Within this bounding box, additional spawn locations are created using the RNG. At these locations, oval-shaped contours are formed, designed to imitate the bronchograms. These contours are then subjected to noise manipulation to vary the points, adding realism to the simulation. The final step involves smoothing the contours

by interpolating between these points using B-Spline interpolation with a smoothing factor of k=3. This approach ensures the edges of the bronchograms are realistically smooth, closely resembling their natural appearance in the lung region.

1. **Select a Coordinate within the Lung Region Using RNG:**
   Let L represent the lung region, defined as a set of all possible coordinates within the lung. Let (x, y) be a coordinate within L. The RNG function R_L selects a coordinate:

$$(x, y) = R_L(L) \tag{5.11}$$

2. **Generate a Bounding Box Centered at the Selected Coordinate:**
   Let the bounding box be defined by its center (x, y) and dimensions (w, h). The bounding box B is:

$$B = \{(x - \frac{w}{2}, y - \frac{h}{2}), (x + \frac{w}{2}, y + \frac{h}{2})\} \tag{5.12}$$

3. **Generate Various Spawn Locations within the Bounding Box Using RNG:**
   Let S represent a set of spawn locations within B. Each spawn location (x_i, y_i) is generated by:

$$S = \{(x_i, y_i) | (x_i, y_i) = R_B(B),\ i = 1, 2, ..., n\} \tag{5.13}$$

   where R_B is the RNG function for bounding box B and n is the number of spawns.

4. **Create an Oval of an RNG Based Height and Width:**
   For each spawn location (x_i, y_i), create an oval with height h_i and width w_i, randomly chosen using RNG:

$$O_i = \text{Oval}(x_i, y_i, h_i, w_i) \tag{5.14}$$

   where $h_i = R_h()$ and $w_i = R_w()$, with R_h and R_w being RNG functions for height and width respectively.

5. **Apply Noise to the Coordinates of the Oval:**
   Let N be a noise function. The noisy oval O_i' is obtained by applying N to each point in O_i:

$$O_i' = \{N(p) | p \in O_i\} \tag{5.15}$$

6. **Using B-Splines with k=3, Interpolate the Oval Coordinates:**

   Let B_{spline} be the B-Spline interpolation function with a smoothing factor $k = 3$. The interpolated oval I_i is:

$$I_i = B_{\text{spline}}(O'_i, k = 3) \tag{5.16}$$

As a visualisation aid, in Figure 5.2 the various labelled regions for textural generation can be seen, with the differences in colour highlighting the different tissue regions.

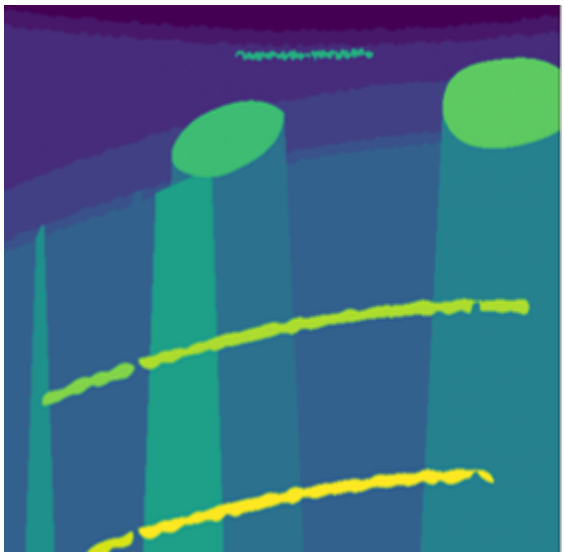

Figure 5.2: Synthetic ultrasound labelled image

5.3 Ultrasound noise and texture generation

Generating synthetic ultrasound textures in a computational environment presents a unique set of challenges, primarily due to the inherent complexities of ultrasound imaging. Ultrasound textures are characterised by speckle patterns and varying degrees of echogenicity, which are difficult to replicate accurately. The primary challenge lies in emulating the stochastic nature of speckle noise, a quintessential feature of ultrasound images, which arises from the coherent nature of the ultrasound waves. Additionally, accurately simulating the acoustic properties of different tissues, such as their density and elasticity, is critical yet challenging, as these properties significantly influence the ultrasound texture. Another considerable challenge is the implementation of realistic attenuation and scattering effects that

ultrasound waves undergo as they propagate through tissues. These effects are crucial for generating realistic textures but are difficult to model due to their complex, non-linear nature. Furthermore, the need to balance computational efficiency with the high fidelity of the textures adds an additional layer of complexity. High-resolution, realistic textures are computationally demanding, which can be a constraint in real-time applications. Addressing these challenges is vital for advancements in ultrasound simulation, particularly for training and diagnostic purposes, where the authenticity of the synthetic textures plays a crucial role in their effectiveness.

In the pursuit of engendering a textural facsimile of ultrasound imaging, it is imperative that the simulation be established on procedural generation paradigms. This approach ensures the uniqueness of each resultant image, thereby enriching the diversity of the dataset substantially. Subsequent to this, it is of paramount importance to facilitate the formation of hypoechoic and hyperechoic structures, inclusive of, but not limited to, muscular collagen and adipose deposits. A procedural noise generation technique, delineated in [101], employing sparse Gabor convolutions, has demonstrated commendable efficacy in emulating linear structures analogous to the textural patterns observed in ultrasound imaging. Drawing inspiration from this seminal work, it is proposed that a bespoke noise algorithm, coupled with a convolution filter, be implemented. This endeavour aims to yield textures that bear a close resemblance to those observed in ultrasound imagery, thereby advancing the fidelity of the simulated images.

5.3.1 Textural noise generation

The distribution of textural noise necessitates a divergence from previously utilised methodologies, which predominantly incorporate normally distributed noise to superimpose a fundamental 'white noise' onto a visual canvas. This deviation is attributed to the distinctive nature of ultrasound imagery, which contrasts markedly with the cohesive texture typically associated with procedurally generated textures, such as fur or wood. Ultrasound images characteristically exhibit a non-uniform texture, manifesting a myriad of darker regions corresponding to tissues with hyperechoic properties, juxtaposed with areas exhibiting a more uniform texture indicative of echogenic tissues. Consequently, this necessitates the adoption

of a tailored approach to noise distribution, ensuring the resultant texture aligns closely with the inherent characteristics of ultrasound images.

To mitigate this, noise shall be distributed across the canvas at a predetermined density, utilising a Random Number Generator (RNG). Each noise element, manifested as a pixel, shall be extended along the x-axis to a length determined randomly, with its grayscale intensity governed by a Rayleigh-Rician mixture, as seen in Figure 5.3. This approach aims to emulate the distribution of grayscale values observed in authentic ultrasound imagery, as referenced in [53]. Conclusively, a one-dimensional Gaussian filter shall be convolved across the elongated noise elements, effectuating a gradient transition of grayscale intensities along the structure, thereby enhancing the visual fidelity to real ultrasound textures.

1. Generate a set of spawn coordinates (x, y) defined by n amount of times:

$$S = \{(x_i, y_i) \mid i = 1, 2, \ldots, n\} \tag{5.17}$$

2. For each spawn, generate a length parameter using RNG:

$$L = \{l_i \mid i = 1, 2, \ldots, n\} \tag{5.18}$$

3. Create a Rayleigh-Rician distribution, with its count determined by n spawns:

$$R = \{r_i \mid i = 1, 2, \ldots, n\} \tag{5.19}$$

4. For each spawn, associate a value from the Rayleigh-Rician distribution to be the grayscale intensity:

$$I_i = r_i \quad \text{for} \quad i = 1, 2, \ldots, n \tag{5.20}$$

5. Convolve a Gaussian filter across the generated line:

$$C_i(x, y) = (G * L_i)(x, y) = \int \int G(x - u, y - v) \cdot L_i(u, v) \cdot I_i \, du \, dv \tag{5.21}$$

In the generation of both hyperechoic and hypoechoic structures, as previously outlined in the section on label generation, the utilisation of ovals with variable noise characteristics is suggested. These ovals are instrumental in defining regions that either preclude the initiation

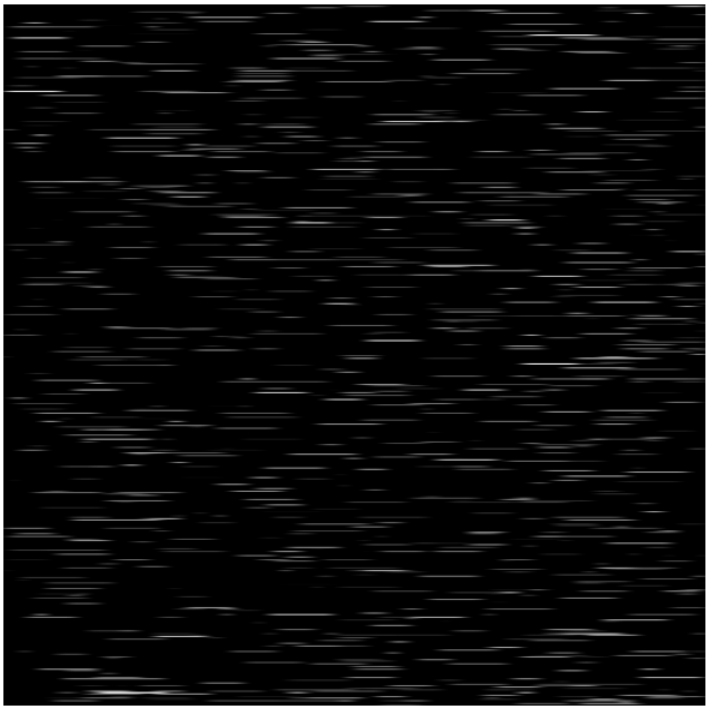

Figure 5.3: Synthetic ultrasound textural noise

of noise generation spawn points (characteristic of hypoechoic structures) or elevate the probability of noise occurrence in those specific areas (typical of hyperechoic structures). Pertaining to the representation of muscular structures, the geometry of these ovals can be modified, elongating the width while compressing the height, to yield more linear structures. In Figure 5.4 can be better visualised, with the red squares representing spawn points for noise on the gray structure. This approach is intended to closely mimic the appearance of collagen as depicted in ultrasound imagery, thereby enhancing the realistic portrayal of such structures in the simulated environment.

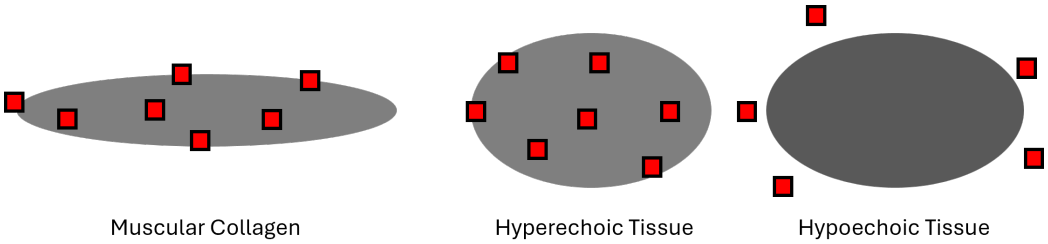

Figure 5.4: Hyperechoic and Hypoechoic regional control, with red indicating spawn locations

5.3.2 Filtering

Following the generation of textural noise, the subsequent phase involves the application of a filtering process to the image, a crucial step in refining the synthetic ultrasound texture. This filtering stage plays a pivotal role in amalgamating the boundaries of the noise, thereby ensuring a seamless transition between textural elements. Moreover, it contributes to the formation of structures reminiscent of those observed in authentic ultrasound imagery, thereby attenuating the distinctly 'geometric' and synthetic characteristics of the preliminary texture. This process is instrumental in enhancing the verisimilitude of the synthetic texture, aligning it more closely with the natural appearance of real ultrasound images.

The noise filtration process employs a Gabor kernel (of which can be seen in Figure 5.5), which is convolved with the image to effectuate texture modulation. The parameters of the Gabor kernel, encompassing the phase, orientation, and frequency, are subject to modification. These alterations are instrumental in adjusting the kernel's influence on the texture, thereby enabling precise control over the resultant image characteristics.

1. Define Gabor Kernel:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \tag{5.22}$$

where

$$x' = x\cos\theta + y\sin\theta, \quad y' = -x\sin\theta + y\cos\theta \tag{5.23}$$

and:

- λ is the wavelength of the sinusoidal factor.
- θ is the orientation of the normal to the parallel stripes of a Gabor function.
- ψ is the phase offset.
- σ is the sigma/standard deviation of the Gaussian envelope.
- γ is the spatial aspect ratio, and specifies the ellipticity of the support of the Gabor function.

2. Convolve Gabor Kernel onto Noise Image:

$$G(x, y) = (g * I)(x, y) = \int \int g(u, v; \lambda, \theta, \psi, \sigma, \gamma) \cdot I(x - u, y - v)\, du\, dv \tag{5.24}$$

where $*$ denotes the convolution operation and $I(x, y)$ represents the noise image.

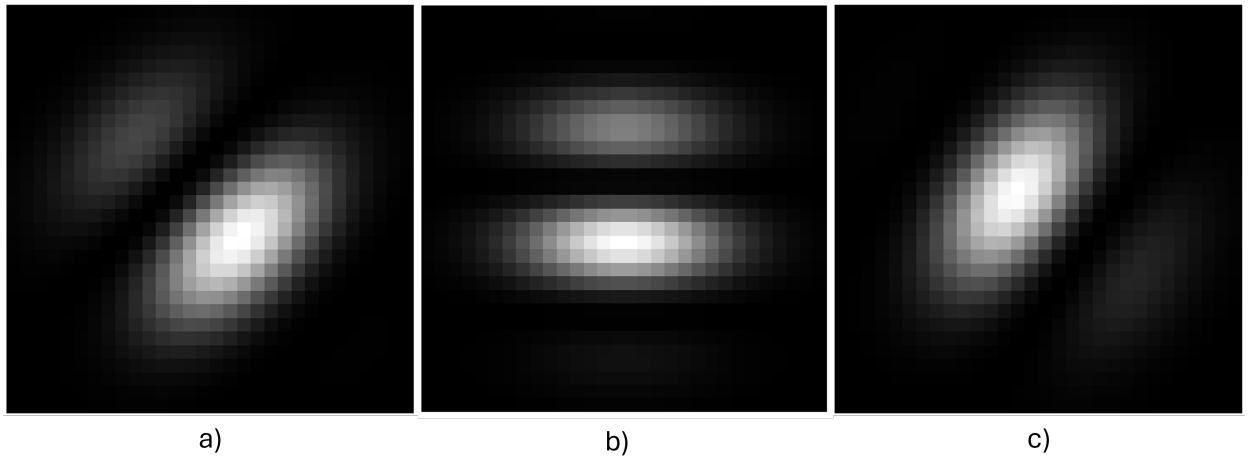

Figure 5.5: Different Gabor kernels a) Phase Angle 45° b) Phase Angle 180° c) Phase Angle 215°

5.3.3 Noise and image transformation

The textured image, generated through the previously discussed procedural texture algorithm, undergoes further transformation following the methodology outlined in [102]. This approach is tailored to accommodate the specific characteristics of linear and convex ultrasound scans, ensuring the processed image closely resembles the patterns typically observed in such medical imaging.

The first stage involves allocating reference points across the image based on the scan type. For linear scans, the reference points are distributed exponentially, with a higher concentration of points positioned closer together at the beginning of the scan and gradually spreading out towards the end. This distribution simulates the varying levels of detail captured as sound waves propagate through the tissue during a linear ultrasound scan.

$$P_i = x_0 + e^{\alpha i} \quad \text{for } i = 0, 1, 2, \ldots, N \tag{5.25}$$

where x_0 is the starting position, α is a scaling factor determining the rate of exponential increase, and N is the total number of points.

Conversely, for convex scans, the methodology employs the equation described in [102] to generate a convex pie slice over the image. Similar to the linear scan approach, the points within this convex slice are also distributed exponentially. This arrangement is crucial for replicating the absorption characteristics of sound waves as they traverse through different tissue layers, accounting for the natural attenuation observed in real ultrasound scans.

Examples of both scan types and their corresponding reference points can be seen in Figure 5.6.

$$P_i = \begin{cases} \theta_0 + \beta e^{\gamma i} & \text{for angle,} \\ r_0 + \delta e^{\epsilon i} & \text{for radial distance,} \end{cases} \tag{5.26}$$

where θ_0 and r_0 define the starting angle and radius, and $\beta, \gamma, \delta, \epsilon$ are constants controlling the distribution.

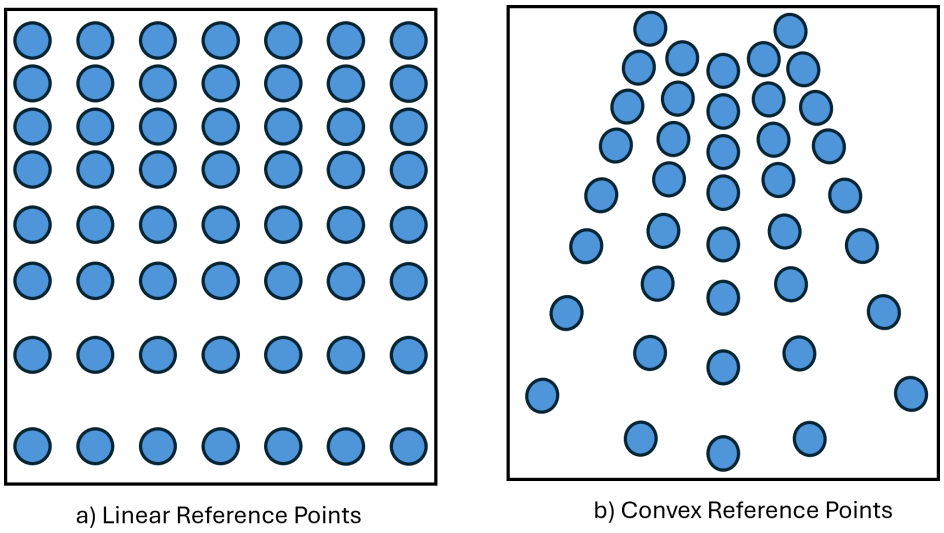

Figure 5.6: Points of reference for (a linear transform b) convex transform

Once the reference points are allocated, they undergo rectification to prepare for the introduction of noise and the subsequent interpolation process. The rectification ensures that the points are adjusted to minimise geometric distortions and artefacts, facilitating a more accurate reconstruction of the final image.

$$\hat{P}_i = f_{\text{rectify}}(P_i), \tag{5.27}$$

where f_{rectify} applies geometric adjustments to correct potential distortions.

The next step introduces noise into the image using the previously established random number generation (RNG) algorithm. A density parameter is employed to control the quantity of noise points applied. The selected points are then processed using a Gaussian distribution to model the speckle noise effect typical of ultrasound imagery. This controlled noise application is essential for simulating the granular texture present in real-world scans, contributing to the overall realism of the final output.

$$N(x, y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2+(y-\mu)^2}{2\sigma^2}}, \tag{5.28}$$

where μ and σ represent the mean and standard deviation of the noise, respectively. The noise density is controlled by a parameter λ, influencing the number of noise points applied.

The final phase involves the interpolation process, where bi-cubic interpolation is applied to reconstruct the image. During this phase, the rectified image is iteratively interpolated onto the canvas, simulating the inherent distortions seen in ultrasound scans. The use of bicubic interpolation allows for smooth gradients between points while preserving essential details, ensuring that the transformed image faithfully replicates the visual characteristics expected in clinical ultrasound applications.

$$I_{\text{final}}(x, y) = \sum_{i=0}^{M} \sum_{j=0}^{M} w_{ij} \cdot I(\hat{P}_i, \hat{P}_j), \tag{5.29}$$

where w_{ij} are the bi-cubic interpolation weights, and M is the number of interpolation points.

It's important to note that when applying interpolation to the semantic label image, nearest neighbour interpolation is used to prevent gradients between the interpolated reference points, which will in turn diminish the correct labelling on the image.

5.4 Light-weight ultrasound simulation

The core objective of this method is to develop an algorithm that enables the fast and lightweight computation of ultrasound simulations within lung tissue. Traditional fluid dynamic simulations of ultrasonic waves interacting with various tissues are recognised for their accuracy in mirroring real-world data; however, they are also known for their prohibitive computation times. This research intends to bridge the gap between accuracy and efficiency. By streamlining the computational framework, the aim is to generate a large dataset of simulations in a significantly reduced time frame while maintaining a high degree of fidelity in comparison to real-world ultrasound data.

At the crux of our simulation methodology is the concept of a "simulation canvas," which represents a cross-sectional area of lung tissue through which ultrasonic waves propagate. This canvas is a computational construct, with each pixel representing a discrete physical element within the tissue. These elements are not just graphical points but are tied to crucial tissue characteristics that affect the behaviour of ultrasonic waves. These properties are:

Reflective Index: Each pixel contains a value representing the tendency of the tissue section to reflect ultrasonic waves. This index is crucial for determining how much of the ultrasonic wave is reflected back towards the transducer and thereby influences the brightness of the resulting ultrasound image. During the simulation, the intensity of the reflected wave corresponds to the brightness or intensity of the pixel in the result array, providing a visual representation of wave-tissue interaction.

Density: Tissue density plays a pivotal role in how ultrasonic waves attenuate as they travel through the tissue. A higher density indicates more resistance and absorption, causing the ultrasonic pulse energy to degrade with each pixel it traverses. This aspect of the simulation is critical for accurately portraying the decrease in signal strength at increasing depths, directly impacting the clarity and quality of the final image.

Refraction Index: The refraction index within each pixel is indicative of the change in direction an ultrasonic wave undergoes as it passes through tissues of varying densities. This property determines whether and how the wavefront will be altered or if the pixel's

information needs to be processed again in subsequent computational steps. The variations in the refraction index among pixels influence the apparent position and intensity of the returning waves on the result array, thereby affecting the image's overall sharpness and structural representation.

Reverberation Index: In the context of ultrasonic wave propagation, the reverberation index is a critical parameter that quantifies the extent of wave reverberation at the tissue interface. This index plays a pivotal role in ultrasound imaging, influencing the manifestation of specific artefacts. Notably, A-Lines and B-Lines are two distinct artefacts associated with varying reverberation indices. A-Lines, indicative of a lower reverberation index, are characterised by a pattern of staggered, repeating echoes along the pleural line within the imaging field. Conversely, B-Lines, associated with a higher reverberation index, are manifested as continuous, linear echogenic streaks originating from the pleural region.

The simulation operates on a time-step basis, mimicking the real-time propagation of ultrasonic waves through tissue. As the simulation progresses, the ultrasonic pulse energy's degradation is calculated based on the density property of the pixels it encounters. This dynamic approach ensures a realistic representation of wave energy loss consistent with the physical properties of lung tissue. Simultaneously, the reflective index is employed to compute the intensity of the signal that is reflected back and recorded on the result array. This step is crucial in simulating how ultrasonic waves are captured by the transducer in live scenarios.

Furthermore, the refraction index is instrumental in the iterative process of the simulation. Certain pixels, based on their refraction indices, are queued for repeated computation in the following time steps, albeit with adjusted parameters. This repetition simulates the phenomenon of wave refraction through tissues with different densities and compositions. The iterative nature of this process contributes to the creation of a more nuanced and accurate representation of the tissue, as seen in ultrasound imaging. In Figures 5.7 5.8 5.9, an example close up of the image array can be visualised, with the "t" being the algorithm target pixel as it progresses over time down the array.

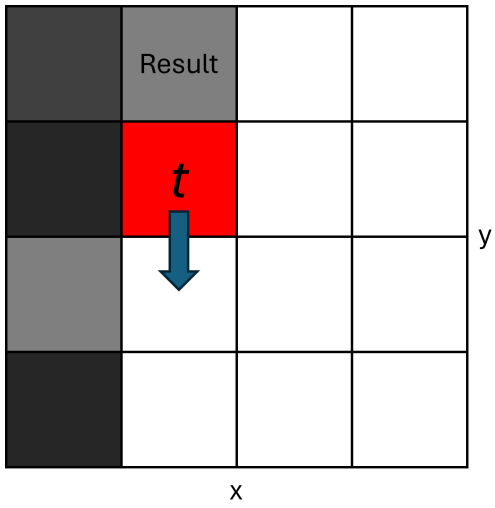

Figure 5.7: Time-based interaction of wave on target

5.4.1 Simulation time-based calculations

In the present simulation, the focal pixel on the imaging canvas represents the location of the ultrasound wave. As this target position traverses the canvas in the y-direction, it engages with the existing pixel information at that specific location. Each pixel embodies a set of parameters including its reflection index, density, reverberation index, and refraction index. These attributes are computed based on a series of equations pertinent to the properties of the ultrasonic wave, consequently influencing its characteristics. As the simulation progresses, the properties of the ultrasonic wave undergo modifications from their initial emitted state. The evolving values at each time increment, denoted as t, are methodically plotted onto the imaging canvas, thereby providing a dynamic representation of the wave's interaction with the tissue.

1. Target Movement in the Y Direction with Respect to Time (t):

$$y_t = y_0 + t \tag{5.30}$$

   where y_0 is the initial y-coordinate of the target, and t represents the discrete time steps in the simulation.

2. Reflection at Pixel Interaction:

$$C(x, y) = C(x, y) + R(x, y) \cdot T(x_t, y_t) \tag{5.31}$$

where $C(x, y)$ is the result canvas, $R(x, y)$ is the reflection value at the pixel, and $T(x_t, y_t)$ represents the target's presence at location (x_t, y_t).

3. Refraction and Adjacent Value Addition:

$$C(x', y') = C(x', y') + \alpha \cdot R(x, y) \cdot T(x_t, y_t) \quad \forall (x', y') \in A(x, y) \tag{5.32}$$

where α (0 ¡ α ¡ 1) is the reduced intensity factor due to refraction, $A(x, y)$ represents the set of adjacent pixel coordinates affected by refraction, and $\vec{v}$ is the refraction vector.

In the simulation framework under consideration, the noise generation is primarily attributed to the refraction phenomena occurring when the ultrasound wave impinges upon the target pixel. This simulated noise, a deviation from real-world observations, arises not from the echo effect—which in practical scenarios results in the misrepresentation of tissue location—but rather from the augmentation of pixel intensity in the vicinity of the target. This intensity enhancement of adjacent pixels culminates in the introduction of additional noise. It is imperative to acknowledge this distinction as it underpins a fundamental difference between the simulated and actual ultrasonic imaging environments.

The reverberation index is a quantifiable measure of the extent to which ultrasonic waves reverberate within a tissue surface. In the context of ultrasound imaging, this phenomenon is responsible for the generation of specific artifacts, namely A-Lines and B-Lines. A-Lines, associated with a lower reverberation index, manifest as a staggered, repetitive pleural line extending down the imaging canvas. Conversely, B-Lines, indicative of a higher reverberation index, appear as continuous streaks emanating from the pleural region of interest.

1. Target Movement in the Y Direction with Respect to Time t:

$$y_t = y_0 + vt \tag{5.33}$$

where y_0 is the initial y-coordinate of the target, v is the velocity, and t represents the discrete time steps in the simulation.

2. Reflection at Pixel Interaction:

$$C(x, y) = C(x, y) + R(x, y) \cdot T(x_t, y_t) \tag{5.34}$$

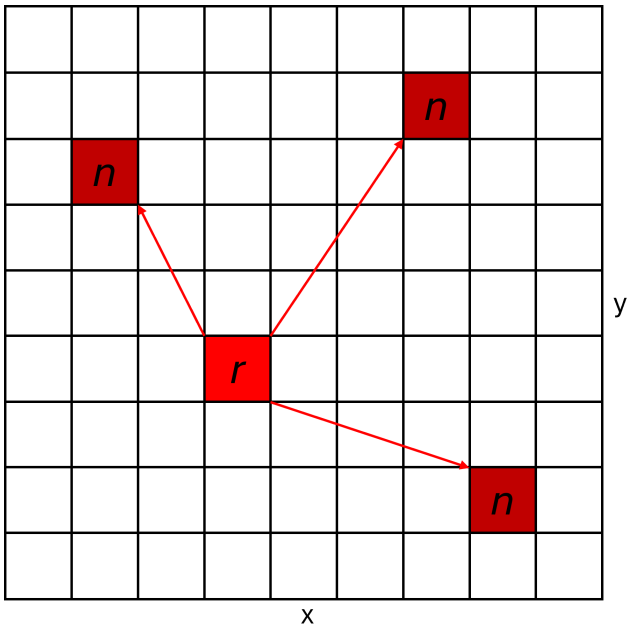

Figure 5.8: Refraction simulation, n = noise spawn point, r = refraction incident

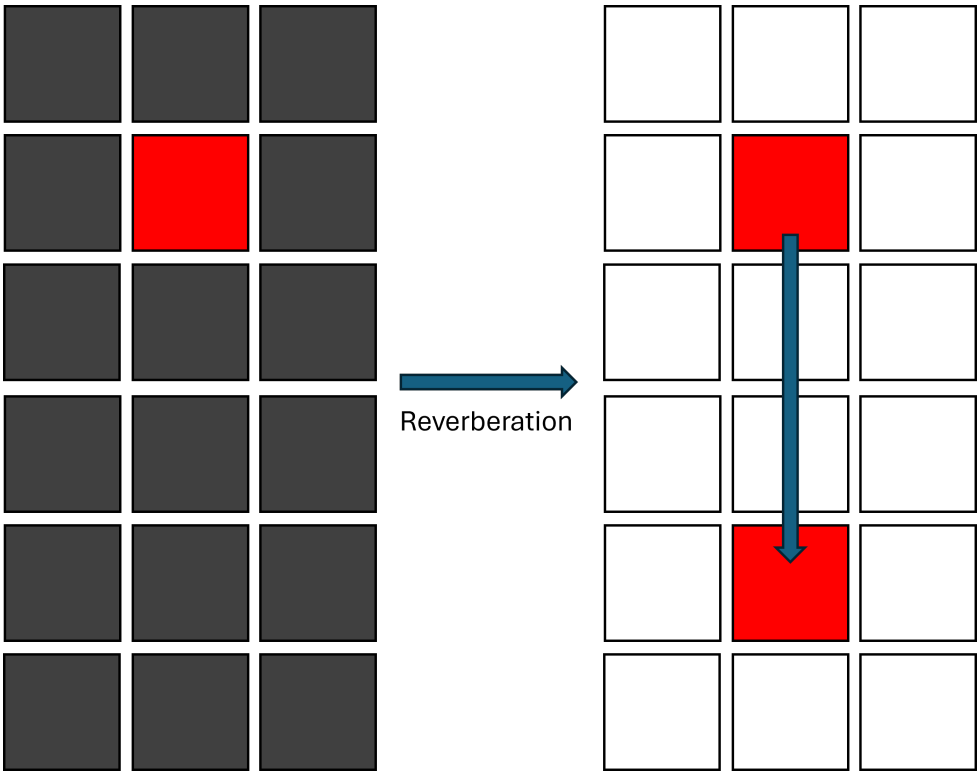

Figure 5.9: Reverberation simulation

where $C(x, y)$ is the result canvas, $R(x, y)$ is the reflection value at the pixel, and $T(x_t, y_t)$ represents the target's presence at location (x_t, y_t).

3. Reverberation and Plotting of Reverberated Pixel:

$$y_{rev} = y_t + t \tag{5.35}$$

$$C(x, y_{rev}) = C(x, y_{rev}) + V(x, y_{rev}) \tag{5.36}$$

where $V(x, y_{rev})$ is the reverberation value added to the result canvas at the new coordinate location (x, y_{rev}).

4. Continuation of Reverberation Until the End of Simulation: The reverberation process continues for each time step t until the end of the simulation, updating $C(x, y_{rev})$ at each time step.

5.5 GAN generated ultrasound training

Generative Adversarial Networks (GANs) have emerged as a revolutionary framework in machine learning, enabling the generation of highly realistic images and data by pitting two neural networks against each other: a generator and a discriminator [103]. GANs operate through a min-max optimization process where the generator aims to produce data indistinguishable from real samples, while the discriminator attempts to differentiate between real and generated data. This adversarial process drives both networks to improve iteratively, resulting in remarkable advancements in image synthesis, style transfer, and data augmentation.

This technique revolutionized semantic image synthesis by utilizing a semantic input map to modulate the normalization parameters spatially, ensuring that high-level, detailed semantic information is effectively infused and preserved throughout the generative process [104]. The SPADE module, central to this method, allows for the generation of high-quality, realistic images, representing textures, object interactions, and contextual details with remarkable accuracy and coherence across various datasets.

The mathematical foundation of SPADE can be described through its spatially-adaptive normalization process. Let $h(x, y)$ denote an input feature map, μ and σ represent the mean and standard deviation of the feature map, and $\gamma(x, y)$ and $\beta(x, y)$ be the learned spatially-adaptive parameters derived from the semantic map $S(x, y)$. The normalization is given by:

$$\hat{h}(x, y) = \gamma(x, y) \cdot \frac{h(x, y) - \mu}{\sigma} + \beta(x, y), \tag{5.37}$$

where the parameters $\gamma(x, y)$ and $\beta(x, y)$ are functions of the semantic map and are learned during training.

The objective function used in SPADE GAN is consistent with traditional GANs, defined as:

$$\min_G \max_D E_{x \sim p_{\text{data}}}[\log D(x)] + E_{z \sim p_z}[\log(1 - D(G(z, S)))], \tag{5.38}$$

where:

- G: Generator network incorporating the SPADE module,
- D: Discriminator network,
- x: Real data sample,
- z: Random noise vector,
- S: Semantic map input.

SPADE's innovation extends to numerous advantages over traditional models. Firstly, it maintains detailed semantic information, a crucial component often diluted in standard GANs, ensuring higher fidelity and realism in generated images. Its spatially-adaptive functionality permits nuanced handling of diverse scenes and complex object details, significantly improving the quality and believability of synthesised images [104]. SPADE models also exhibit enhanced flexibility, demonstrated through their successful application across multiple content-rich datasets and their utility in interactive tools like GauGAN [104].

Fine-tuning is also an essential component, particularly when adapting pre-trained models to domain-specific datasets. In this application, the SPADE GAN will first be trained on a large synthetic dataset to learn generalized features and capture diverse structural characteristics. Following this, the model undergoes fine-tuning by retraining on a smaller

real-world ultrasound dataset. During this phase, the learning rate is restricted, and iteration decay to prevent overwriting key features trained by synthetic set but also allows for the increase to adapt the model to incorporate ultrasound-specific features not well-represented in the synthetic data. This dual-phase training approach allows the model to blend the broad generalization capability of the synthetic dataset with the nuanced, realistic details of the ultrasound dataset, while also introducing extended variation to the generated images, ensuring high fidelity and practical utility.

5.6 Results

This section explores the discussion surrounding the results obtained from the research on simulated ultrasound images, specifically focusing on their representation of acoustic reflections in lung tissues, the anatomical authenticity, and the potential implications these factors have on the training of segmentation networks. The objective behind simulating ultrasound scans is to create a robust basis for training machine learning models, specifically segmentation networks, in the absence of extensive real-scan databases. These models require high-fidelity reproductions of reality to be effective, necessitating a critical evaluation of the simulation's strengths and shortcomings.

5.6.1 Label generation

This subsection introduces the findings related to the application of the Lung Ultrasound Severity Score (LUSS) in the context of our study, which focused on evaluating the utility of LUSS as a tool for assessing the severity of pulmonary conditions. Given the complexity and the multifaceted nature of lung ultrasound imaging, the generation and analysis of LUSS involve a detailed examination of various parameters that are critical to accurately quantifying the extent of lung involvement and damage. This approach, while intricate, allows for a nuanced assessment of lung health, which is paramount in clinical settings, especially for diagnosing and monitoring respiratory conditions.

The parameters for LUSS were chosen based on their ability to reflect the severity of lung pathology as observed in ultrasound imaging. These parameters include, but are not

limited to, the presence and extent of B-lines, consolidation, and pleural line abnormalities. Each parameter contributes to the LUSS, providing a score that aims to encapsulate the ultrasound findings into a quantifiable measure. The process of determining these scores is guided by established criteria, ensuring consistency and reliability in the assessments. Some example results of the image labels can be seen in Figure 5.10. Within the images shown, the "True Label" can be seen, which is the image used in correspondence with the final synthetic image, for use in semantic segmentation training.

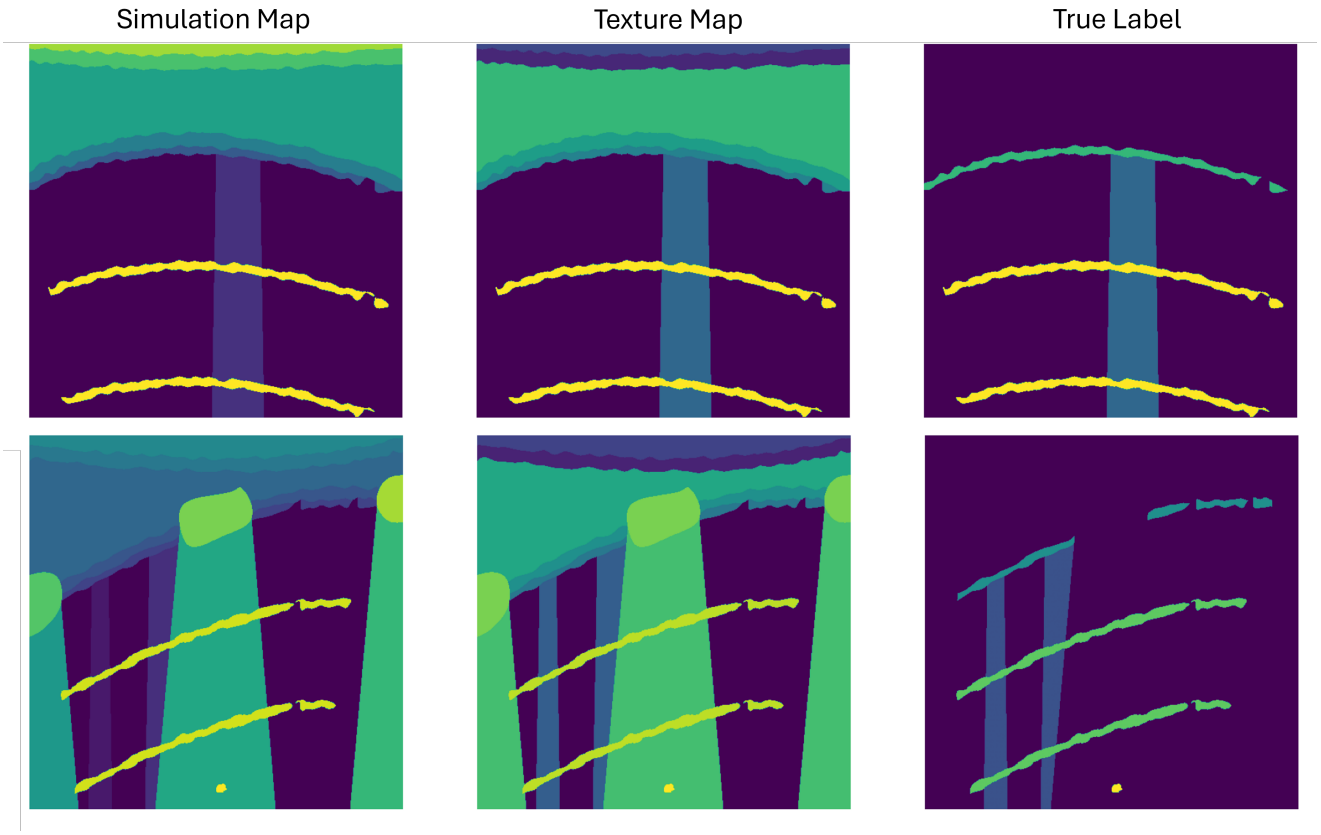

Figure 5.10: Synthetic Labels with varying parameters

5.6.2 Textures

The simulation's efficacy in replicating ultrasound textures was substantiated through a meticulous analysis of the generated images. The textures produced by the simulation exhibited a remarkable resemblance to those observed in real-world ultrasound imagery. This

high degree of fidelity underscores the robustness of the simulation parameters and the effectiveness of the procedural generation techniques employed.

Particularly noteworthy was the coherence of linear structures within the simulated textures when juxtaposed with their real-world counterparts. The simulation adeptly captured the quintessential characteristics of these structures, which are pivotal in the interpretation and diagnostic utility of ultrasound images. The verisimilitude of these linear structures not only enhances the visual authenticity of the simulated textures but also suggests potential applicability in training and diagnostic support tools, where accurate representation is paramount.

Furthermore, the simulation successfully integrated speckle noise as seen in Figure 5.11, an inherent feature of real-world ultrasound imaging. This inclusion is critical, as speckle noise plays a significant role in the texture and visual cues that clinicians rely upon for diagnosis and analysis. The presence of speckle noise in the simulated images adds another layer of realism, thereby enriching the potential utility of the simulation in various applied scenarios, including but not limited to the training of medical professionals and the development of image processing algorithms.

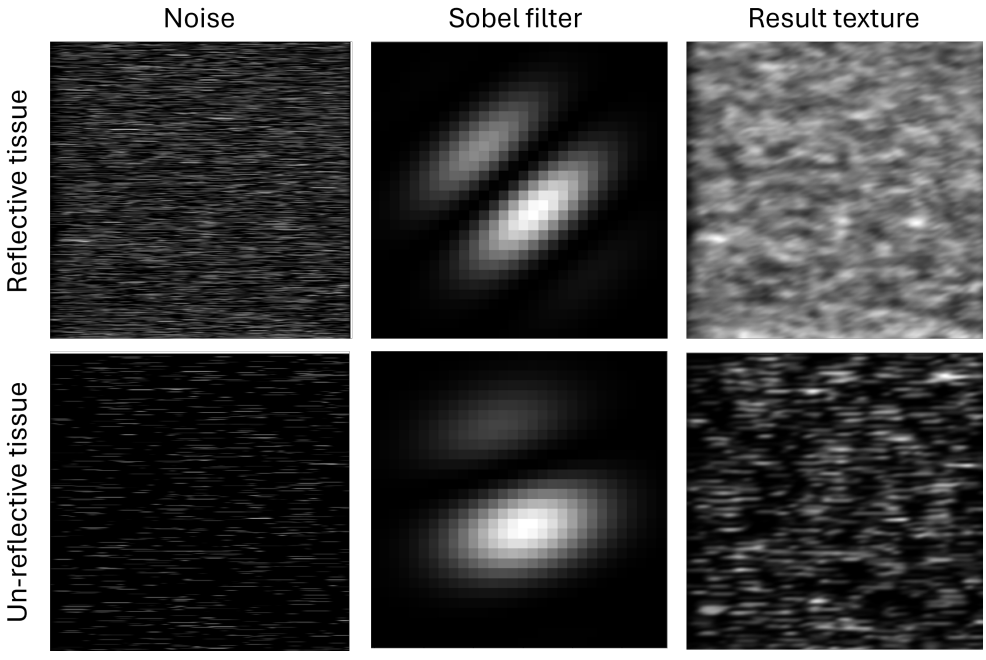

Figure 5.11: Synthetic Textural Results using various densities and Gabor kernels

The resemblance of these convexly transformed synthetic images to their real-world counterparts (seen in Figure 5.12) marks a significant advancement in the creation of highly realistic ultrasound datasets. This achievement is not merely a technical milestone; it carries profound implications for the training of artificial intelligence (AI) networks in medical imaging. The enhanced realism of the synthetic datasets serves to augment the training material available for AI systems, potentially boosting their diagnostic performance. By providing a rich, varied, and accurately represented set of images, the AI networks can learn to identify and interpret a wide range of features and anomalies with greater precision. This, in turn, supports the development of more reliable and effective diagnostic tools, underscoring the value of high-fidelity synthetic images in advancing medical imaging technology and patient care.

Figure 5.12: Real Convex Lung ultrasound images

5.6.3 Simulation

The results indicate a promising replication of certain key aspects within ultrasound imaging. The simulated images successfully depict the decreasing energy in acoustic reflections as one examines deeper into the simulated lung tissue. This phenomenon is consistent with the behaviour of sound waves as they penetrate biological tissues, encountering increased attenuation that results in reduced echo strength.

Furthermore, the simulation proficiently replicates the speckle noise inherent in real ultrasound scans, an element crucial for realism in the images. Speckle noise, a granular interference that inherently exists due to the coherent nature of ultrasound waves, contributes significantly to the image texture and is a factor considered by experts when interpreting scans.

In Figures 5.13 & Figure 5.14 it can be seen that anatomically, the simulation shows high fidelity in representing the pleural line and rib shadows. The appearance of a bright line for the pleura and the absence of imaging below the dense rib structures align with the characteristics of real scans. These anatomical landmarks are crucial for any educational, diagnostic, or analytical applications involving lung ultrasound, as they provide the foundational geography upon which pathologies are identified.

The successful replication of the A-line also reinforces the simulation's potential utility. A-lines, horizontal artifacts indicative of air collections, result from reverberation between the skin and the pleura. Their presence further down the simulated image underscores the model's attention to ultrasound physics, particularly the phenomenon of reverberation within the acoustic properties of the lung environment.

Despite the successes, several limitations impact the simulation's overall efficacy. A notable shortcoming is the absence of B-lines, artifacts that are representative of interstitial lung disease or fluid-filled alveoli, which did not simulate well. B-lines' presence in ultrasound is critical for diagnosing various pulmonary conditions, and their absence indicates a gap in the simulation's pathological representation capacity.

Additionally, there are issues with image contrast and geometry. The premature loss of contrast in the simulated images could lead to difficulties in identifying structures and pathologies that are primarily discerned based on contrast differences in real scans. The overly geometrical nature of the images also detracts from realism, making them easily distinguishable from actual scans. Real ultrasound images possess a certain organic irregularity due to the body's natural contours, tissue heterogeneity, and the interplay of sound waves with varying mediums.

The limitations bear significant implications for the foundational training of segmentation networks. The objective of these networks is to accurately identify and isolate anatomical

structures and potential pathologies within ultrasound images. However, if the training images lack essential elements (like B-lines), have early contrast loss, and carry an unrealistic geometrical presentation, the networks may develop a skewed understanding of real-world scans.

This discrepancy between simulated and real images can lead to a performance gap when the model encounters actual patient data. It might result in lower accuracy and reliability, as the system could fail to recognize important cues or misinterpret image artifacts and structures. This outcome undermines the goal of employing machine learning to assist in medical diagnostics and analysis, emphasizing the need for high-fidelity simulations in training phases.

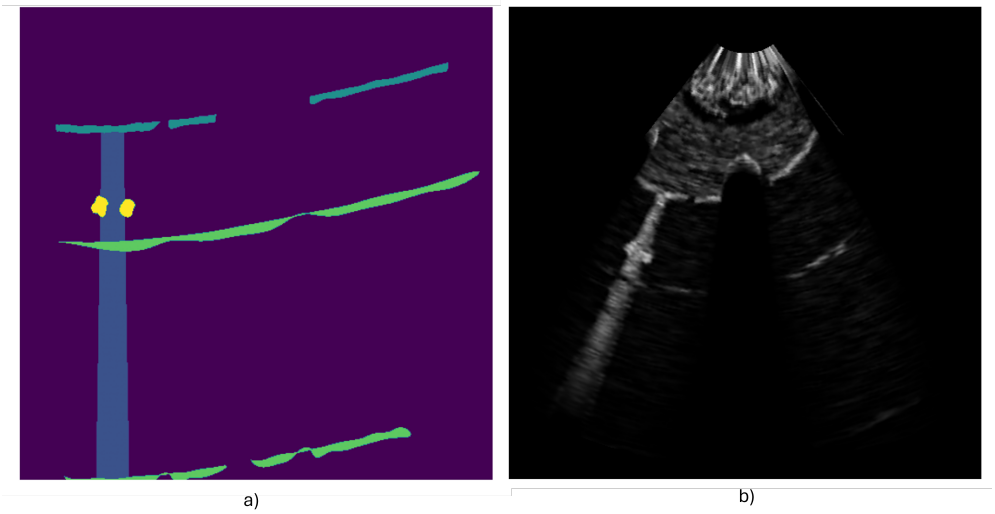

Figure 5.13: Convex Synthetic Simulation Image

5.6.4 GAN

This study explores the capabilities of SPADE GAN for high-quality image generation, trained on a dataset of 4,000 images across more than 100 epochs. Utilizing a computational setup comprised of an i9-13900k CPU and an AMD R7900 XTX GPU, the research aims to leverage the SPADE GAN's innovative spatially-adaptive normalization technique for creating contextually relevant images.

The SPADE GAN network model was initially trained on a synthetic dataset of 4000

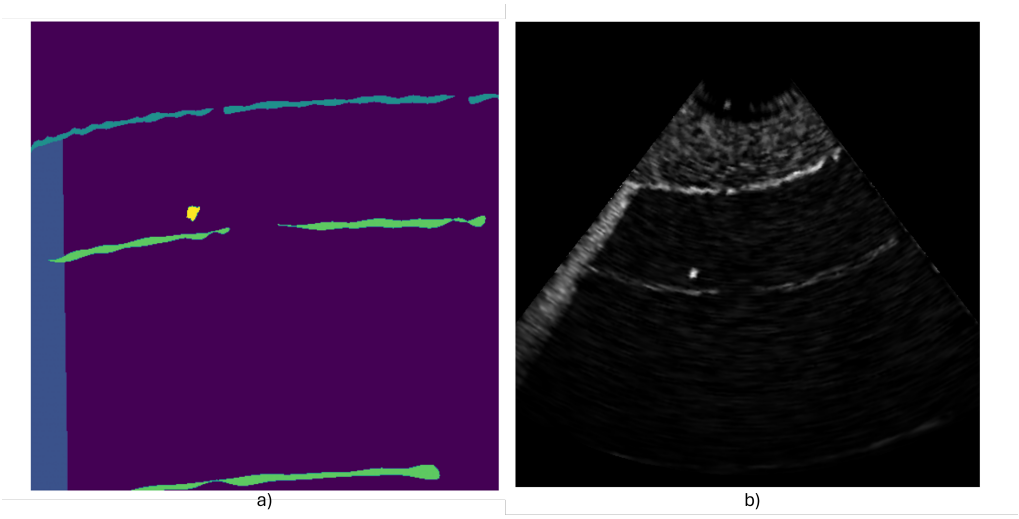

Figure 5.14: Alternate Convex Synthetic Simulation Image

images for 50 epochs and subsequently transferred and fined-tuned using a limited real ultrasound dataset of 220 images for 50 epochs to enhance the realism of the synthetic images. This approach aimed to bridge the gap between purely synthetic images and real ultrasound imagery, thereby improving the utility of synthetic images for various applications.

The evaluation metrics of the Generative Adversarial Network (GAN) in Table 5.2 reveal insights into the performance and limitations of the model. The Fr'echet Inception Distance (FID) was measured at 257.95, indicating a substantial divergence between the distributions of real and generated images. Lower FID scores are generally preferred, with values under 100 often regarded as indicative of high-quality synthesis. The Inception Score (IS) was found to be 1.38 0.15, which reflects limited diversity and quality in the generated images. For context, state-of-the-art GANs typically achieve IS values exceeding 5 on datasets like CIFAR-10. Precision and recall scores for the generated images were approximately 0.50, implying that while some generated images align with real ones in quality, the overall diversity of the generated samples fails to fully capture the real image distribution. The Structural Similarity Index (SSIM) was measured at 0.46, indicating moderate alignment in structural similarity between real and generated images, though this suggests room for improvement in capturing finer details.

Comparison with related works highlights several strengths and shortcomings of the current model. A progressive growing architecture was utilized alongside sketch guidance to

Table 5.2: Evaluation metrics of the proposed SPADE GAN.

Metric	**Value**
Fréchet Inception Distance (FID)	257.95
Inception Score (IS)	1.38 ± 0.15
Precision	0.50
Recall	0.50
Structural Similarity Index (SSIM)	0.46

achieve realistic ultrasound image synthesis [75]. While specific FID values were not reported, the study emphasized significant improvements over baseline models. Additionally, the generated images were validated by their ability to enhance segmentation tasks, often outperforming datasets composed solely of real images. Furthermore, qualitative evaluations confirmed the strong subjective realism of the generated images. Compared to this study, the proposed SPADE GAN's FID score of 257.95 signals a lower level of visual fidelity and distribution alignment. The absence of segmentation tasks or qualitative user assessments in the proposed method's evaluation further underscores these differences.

Another relevant study on detecting COVID-19 features in lung ultrasound images [78] demonstrated the efficacy of combining simulated and real data for training deep neural networks. This approach led to Dice Similarity Coefficient (DSC) scores of 0.464 0.230 when using simulated data alone and 0.735 0.187 when combining simulated and real data, illustrating the value of diverse training datasets. Additionally, the combination of real and simulated data significantly reduced training epochs and improved generalizability. The proposed SPADE GAN's SSIM score of 0.46 aligns closely with the DSC for simulated-only data in this study, suggesting comparable baseline quality. However, incorporating a mixed-data strategy could be a powerful method to enhance the diversity and quality of the proposed SPADE GAN's outputs.

The study on GANs for generating synthetic ultrasound images explored the utility of StyleGAN2 [8] and adaptive discriminator augmentation (ADA) techniques to improve performance on limited datasets. This research reported FID values of 140.6 for the baseline

StyleGAN2 model and 87.97 with ADA. SSIM scores ranged from 0.15 to 0.16 when trained on smaller datasets of 50 to 250 images. Additionally, augmenting real datasets with synthetic data improved classification F1 scores by up to 15.3%. The FID score of 257.95 is notably higher, suggesting room for architectural or training enhancements. However, the SSIM score of 0.46 surpasses other reported values, possibly due to differences in dataset size or evaluation protocols.

Several insights emerge from these comparisons. Increasing the diversity of generated images may improve the Inception Score; this can be achieved by expanding the latent space dimensions or training on more diverse datasets. Enhancing SSIM scores could involve multi-scale discriminators to capture high-frequency details or auxiliary tasks such as edge detection or segmentation to enforce structural consistency. Following the strategy from the COVID-19 detection study, integrating real and simulated data into the training process may significantly improve generalization and diversity capabilities. Additionally, qualitative evaluations, such as user studies and downstream segmentation tasks, would provide a more comprehensive understanding of the model's practical utility. Moreover, the proposed SPADE GAN could be further improved by limiting the learning rate after the initial training phase on synthetic images, which would help preserve key structures in the generated outputs. This refinement strategy may enhance the preservation of critical features while continuing to improve overall performance.

Even though some of the metrics fall behind state-of-the-art results, the visual comparison of simulated images generated by the proposed SPADE GAN with those from other methods demonstrates convincing outputs. A notable feature of the proposed SPADE GAN is its ability to successfully replicate synthetic lung ultrasound images from the procedurally generated label maps which can be seen clearly from Figures 5.15 & 5.16, which fulfills the ultimate goal of automating labelled data generation for segmentation training. The generated images showcase B-lines and A-lines that closely resemble their real-world counterparts, adding to the subjective quality of the model's output.

When comparing the images in Figure 5.17 to the procedurally generated data, the improvements in image realism can be attributed to the effective transfer learning process, where the model leverages the limited real data to refine its understanding of ultrasound

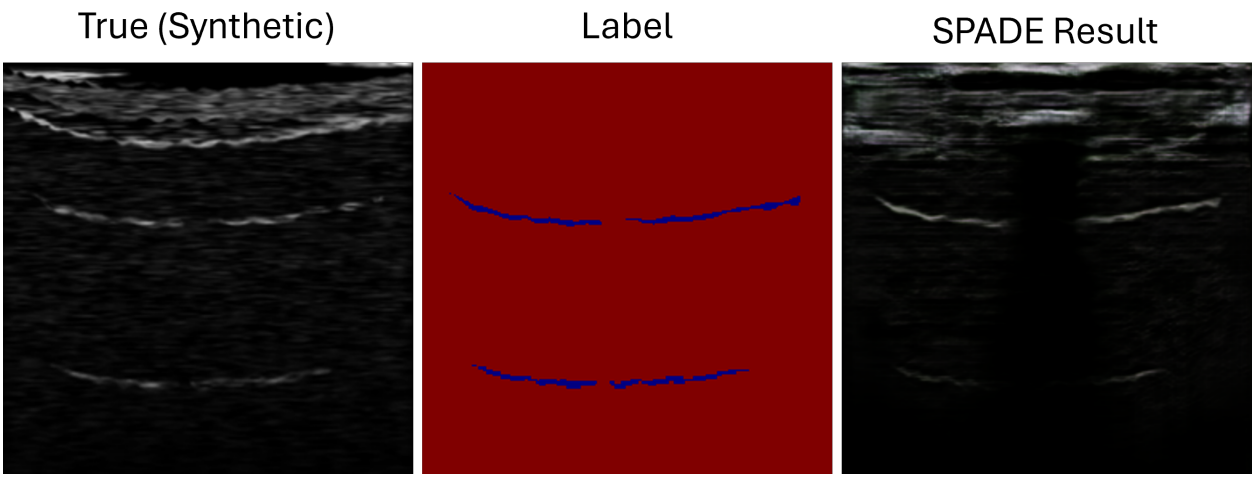

Figure 5.15: SPADE GAN training result from epoch 50 of synthetic lung ultrasound generation

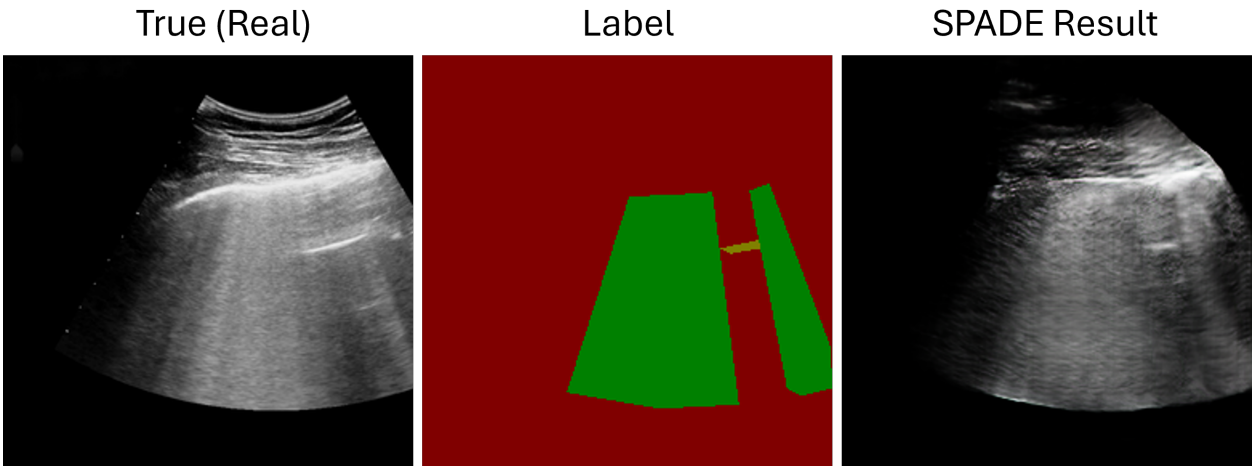

Figure 5.16: SPADE GAN fine tuning result from epoch 100 of real lung ultrasound generation

image characteristics. This approach not only enhances the visual quality of the synthetic images but also suggests the potential for these images to be used in training and diagnostic applications where high-fidelity imaging is crucial. Further work may focus on quantitatively assessing the model's performance and expanding the dataset to ensure robustness across various ultrasound imaging scenarios.

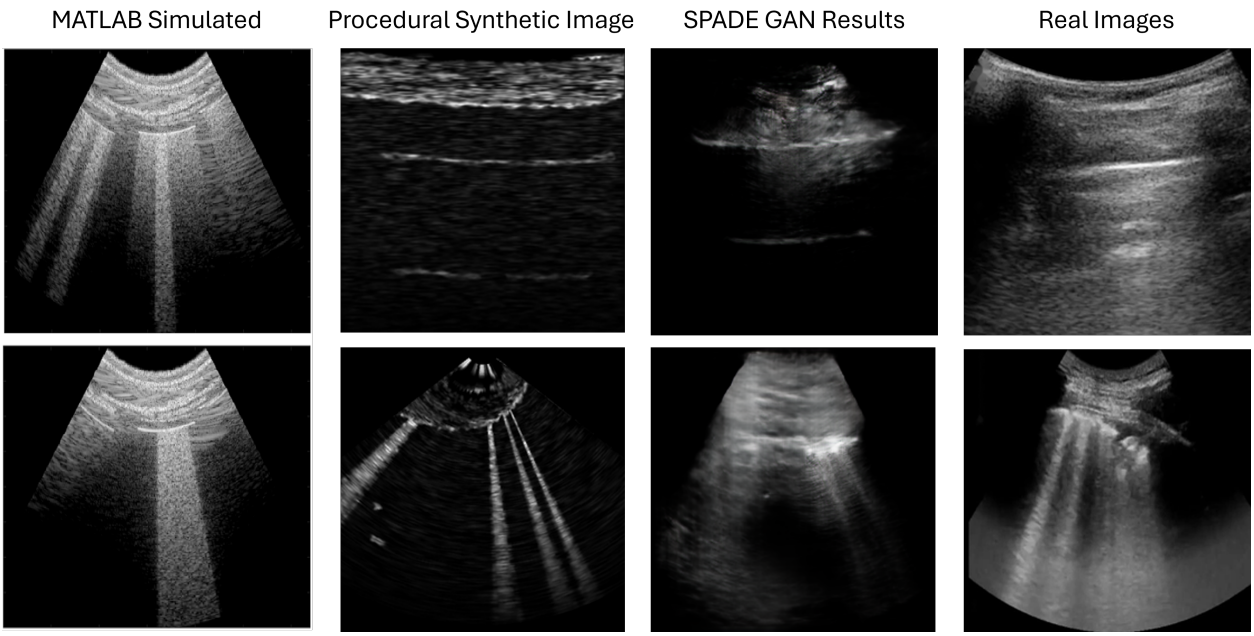

Figure 5.17: Comparison of simulated MATLAB ultrasound [78], Procedural image (Proposed), and SPADE GAN image (Proposed)

5.7 Conclusions

This chapter presented a comprehensive framework for simulating and augmenting ultrasound data, focusing on the generation of realistic textures, labels, and images that replicate the unique characteristics of real-world ultrasound scans. The methodologies outlined in this work address critical challenges in medical imaging, including the scarcity of high-quality annotated datasets, the variability inherent in ultrasound imaging, and the need for robust training data for machine learning models.

The simulation of ultrasound wave interactions with tissue was a central theme, employing a detailed computational framework to replicate the behaviour of ultrasonic waves in lung tissue. By incorporating key properties such as reflection, refraction, and reverberation indices, the simulation successfully captured the dynamic interplay of ultrasonic waves with varying tissue densities and compositions. This approach enabled the generation of synthetic datasets that closely resemble real-world scans, providing a valuable resource for training and validating diagnostic algorithms.

The synthesis of textural noise and artefacts further enhanced the realism of the simulated

images. Techniques such as the application of Gabor filters and Gaussian convolution were employed to generate textures that mimic the granular appearance of real ultrasound scans. The addition of hyperechoic and hypoechoic structures, guided by probabilistic models, contributed to the accurate representation of diagnostic features and artefacts such as A-lines and B-lines. These advancements underscore the importance of precise texture generation in creating high-fidelity synthetic images.

Generative Adversarial Networks (GANs) played a pivotal role in augmenting ultrasound data, enabling the creation of diverse and realistic datasets that enhance the generalisability of machine learning models. By leveraging the adversarial training paradigm, GANs generated synthetic images that captured the variability and complexity of real-world ultrasound scans. This capability is particularly valuable in addressing the challenges posed by limited annotated datasets, offering a scalable solution for dataset expansion.

The integration of label generation techniques further enriched the framework, providing accurate and consistent annotations for the synthetic data. By employing methods such as bounding box generation and semantic segmentation, the framework ensured that the generated labels aligned with the diagnostic features present in the synthetic images. This alignment is crucial for training supervised machine learning models and improving their performance in real-world clinical settings.

While the methodologies presented in this chapter represent significant advancements in ultrasound data simulation and augmentation, several areas warrant further exploration. The incorporation of domain-specific metrics for evaluating synthetic data quality, alongside traditional quantitative metrics, could provide a more comprehensive assessment of the generated datasets. Additionally, the exploration of hybrid models that combine simulation-based techniques with GANs holds promise for further enhancing the fidelity and diversity of synthetic datasets.

Future work should also consider the application of these methodologies to other imaging modalities, extending their utility beyond ultrasound. The development of cross-modality data generation frameworks could facilitate the creation of unified datasets that support the training and validation of multi-modal diagnostic systems. Moreover, the integration of these techniques into clinical workflows has the potential to revolutionise diagnostic imaging,

enabling more accurate and standardised interpretations.

In conclusion, this chapter highlights the transformative potential of advanced simulation and augmentation techniques in addressing the challenges of ultrasound imaging. By generating high-quality synthetic data that faithfully replicates real-world conditions, these methodologies contribute to the development of robust and reliable diagnostic tools. The findings underscore the importance of innovation in medical imaging, paving the way for future advancements that enhance patient outcomes and support the widespread adoption of automated diagnostic systems.

Chapter 6

P-Net: An integrative framework for lung ultrasound segmentation and classification with limited data

6.1 Introduction

Ultrasound imaging has become an indispensable tool in modern medical diagnostics, offering non-invasive insights into the human body. However, the interpretation of ultrasound images presents unique challenges due to their inherent obscurity and the subtlety of diagnostic features such as A-lines and B-lines. Traditional linear algorithms often fall short in accurately segmenting these features, underscoring the necessity for more advanced techniques. Deep learning, with its ability to learn complex patterns from data, has emerged as a promising approach to this problem. Nonetheless, the application of a singular neural network model often leads to high variance and a potential bias towards specific features at the expense of others. This limitation can significantly hinder the model's generalization capability and its effectiveness in medical image analysis.

This work introduces the "P-Net" framework, a novel architecture designed to integrate segmentation and classification tasks within a unified pipeline. Leveraging the strengths of the U-Net architecture for segmentation and deep neural networks like ResNet, VGG-

19, and Inception modules, P-Net addresses the critical task of identifying and classifying diagnostic features such as A-lines, B-lines, and air bronchograms in lung ultrasound images. The framework incorporates variational autoencoders (VAEs) for transfer learning, enabling efficient feature extraction and enhancing the performance of the classification models.

The chapter begins with an overview of pre-processing techniques, including despeckling algorithms and contrast enhancement strategies, which form the foundation for robust image analysis. It then delves into the ensemble approach adopted in P-Net, highlighting its ability to mitigate model-specific biases and improve segmentation accuracy. This is followed by a detailed discussion of the classification methodologies, emphasising the synergistic integration of segmentation outputs and latent features from VAEs. Finally, the results of these methods are evaluated against existing benchmarks, demonstrating the superior performance and generalisability of the proposed framework.

Through this comprehensive exploration, the chapter underscores the potential of combining advanced computational methods with domain-specific knowledge to address critical challenges in medical imaging. The findings not only contribute to the ongoing evolution of diagnostic tools but also lay the groundwork for future innovations in this field.

6.2 Pre-processing methods

As outlined in Chapter 3 and 4, a method to preprocess the image was employed. For the case of the P-Net dataset, the images used have been processed with the ADMSS and R-GMM contrast enhancement from chapter 3, then being cleaned of redundant information with the pleural line detection algorithm in Chapter 4. More details on the dataset and parameters will be discussed at the results on Section 6.6

6.3 Variational auto-encoder for latent feature transfer learning

Autoencoders utilise unlabelled data to identify both high and low-level features within an image. Variational autoencoders, in particular, aim to learn how to represent the input

image in a compressed format, often referred to as the bottleneck of the network. Within this project, autoencoders are employed as a form of unsupervised transfer learning. They are trained on a large dataset to learn the intrinsic features of ultrasound images. The trained encoder model can subsequently be saved and used as the foundation for training the ensemble network for segmentation, thereby reducing the amount of data needed and achieving higher accuracy in fewer epochs.

It has been argued that autoencoders do not necessarily improve the overall results of the network; rather, they enable the network to achieve desired results more quickly (i.e., in fewer epochs) compared to a network that has not utilised an autoencoder. One limitation of autoencoders is the potential loss of information during the image reconstruction process. This issue can be mitigated by incorporating U-Net skip connections, which help preserve important features during the decoding stage.

In the P-Net framework, the Variational Autoencoder (VAE) serves as a cornerstone for feature extraction, enabling robust transfer learning from large, unlabeled datasets to classification networks. By learning a compact and expressive latent representation of lung ultrasound images, the VAE mitigates challenges associated with limited annotated data and enhances the performance of downstream tasks.

The trained encoder of the VAE integrates seamlessly into the classification pipeline. The encoder transforms raw input images into latent vectors, which serve as input to the classifier networks. This transfer learning approach leverages the pre-trained encoder to provide a compact and information-rich representation of the input data, reducing computational complexity and enhancing accuracy in the classification task.

In this project, a VGG-19, Residual, and Inception network encoder architecture is proposed as the foundation for the autoencoders used in the ensemble network. Each encoder network will be trained and saved prior to being integrated into the ensemble network. These architectures are well-documented in the literature and are known for their strengths and weaknesses in feature detection, making them suitable choices for this application.

Outlined in Table 6.1 are training parameters for the VAE used in the P-Net pipeline

Table 6.1: Training configuration and hyperparameters for VAE

Parameter	Value
Model	VAE encoder + linear classifier
Input format	256×256 grayscale images
Latent dimension	128
Encoder	2 Conv layers (1→32→64)
Decoder	2 ConvTranspose layers (64→32→1)
Encoder frozen	Yes
Feature dimension	128
Classifier head	Linear: $128 \rightarrow 4$ classes
Optimizer	Adam
Learning rate	0.001
Loss function	CrossEntropyLoss
Batch size	32
Epochs	25
Dataloader workers	4

6.4 Ensemble segmentation via U-Net

Neural networks tend to have high variance and are not linear in training for results. This means even though a network may be trained well on a set data, it may struggle to pick out certain features that another network could identify easily [105]. To get around this issue networks can be ensembled to produce less variance in the result. Using model averaging, different networks are trained on the same set of data with the results combined to create an averaged result from the networks.

Ultrasound images features are obscure in terms of segmentation when using standard linear algorithms. The use of deep learning can be used as an approach to segmentation when it comes to harder to detect features such as A-lines and B-lines. However, relating to information discussed in the paragraph above, the use of one network may be heavily biased towards a certain feature and struggle to identify others. To address this using different

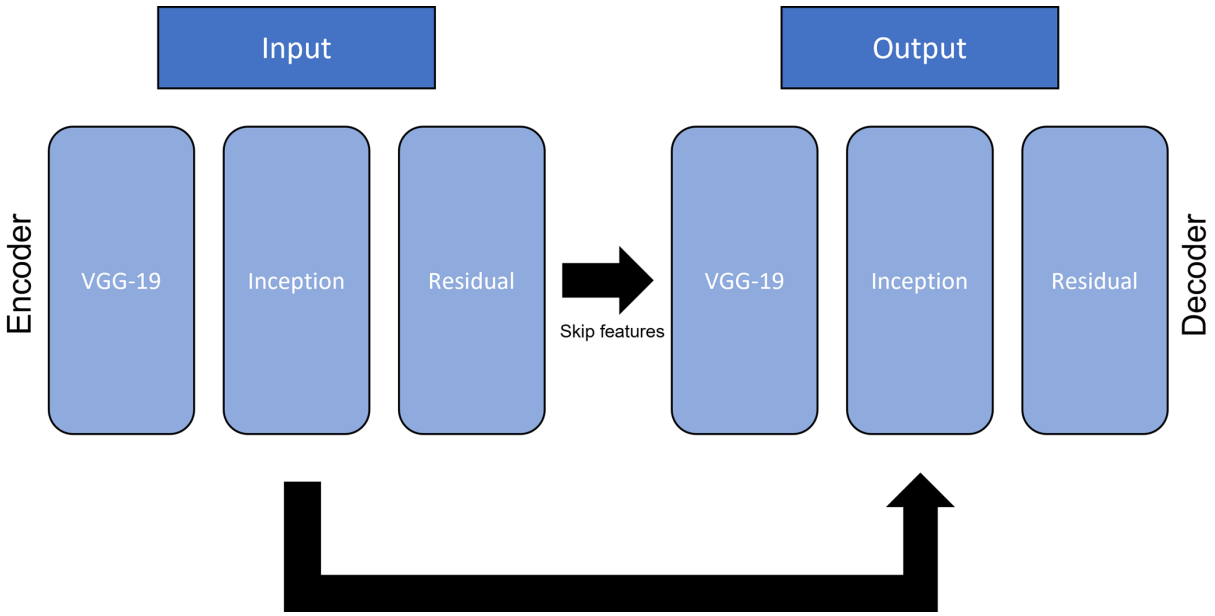

Figure 6.1: Ensemble U-Net architecture

networks in ensemble can allow for the strengths of many networks to identify all features that other may find a challenge to identify on their own.

The ensemble model will use the foundations of the U-Net architecture proposed in [106], due to its proven effectiveness in biomedical images. The diagram 6.1 denotes the outline of this architecture.

Autoencoders make use of unlabelled data to identify the high and low-level features present in an image. Variational autoencoders try to learn how to represent the input image in a compressed format, often called the bottleneck of the network. Using an autoencoder in a network can be observed a form of unsupervised transfer learning, in terms of this project, the autoencoders will be trained on a large dataset to learn the features present in the ultrasound images. This encoder model can then be saved and used as the foundation when the ensemble network is trained for segmentation, requiring a less data to achieve a higher accuracy in less epochs. It has been argued however that autoencoders have no effect in the overall improvement of results [47], only that the network can achieve the results quicker (less epochs) than a network that has not used an autoencoder. Another downside of autoencoders is the loss of information when reconstructing the image. This issue can be rectified by using the U-Net skip features to preserve notable features when reconstructing

the image in the decoding stage.

This chapter proposes using a VGG-19, Residual, and Inception network encoder architecture as the foundation of the autoencoders to be used in the ensemble network. Each network encoder will be trained and saved prior to being transferred into the ensemble network. The networks used are widely used within the literature, having their strengths and weakness in feature detection. In Figure 6.1 an overview of the ensemble architecture can be visualised

The VGG-19 architecture was proposed in [107], a design with the aim to improve network depth. The network is characterised by its simplicity and ability to yield reliable results.

The next network used in the ensemble network will be the Inception network proposed in [108], designed to for overfitting that is prone in very deep networks. The network instead takes the approach of widening the convolution filters rather than adding a deeper level of filters that is present in other networks such as VGG. The network has been modified to work as an encoder with the addition of batch normalisation.

The final network will be the Residual network proposed in [109], designed to utilise prior layers as residual functions to be used as a reference in later functions deeper within the network. This prevents learning from unreferenced functions present in other networks which has been evidenced that these network architectures are easier to optimise. Like previous networks, this has been modified to work as an encoder in the ensemble network with the addition of batch normalisation layers.

6.4.1 P-Net

In this subchapter, an innovative methodology, a proposed "P-Net," for the segmentation and classification of ultrasound images, with a particular focus on the critical identification and differentiation of A-lines, B-lines, and air bronchograms. Shown in Figure 6.2 the overall architecture of the P-Net can be seen, showing the incorporation of the segmentation U-Net with an inception classifier subsequently.

These elements are essential in lung ultrasound analyses, widely utilised in various clinical scenarios. The methodology is anchored on a robust segmentation network that draws upon the established U-Net architecture, renowned for its effectiveness in medical image

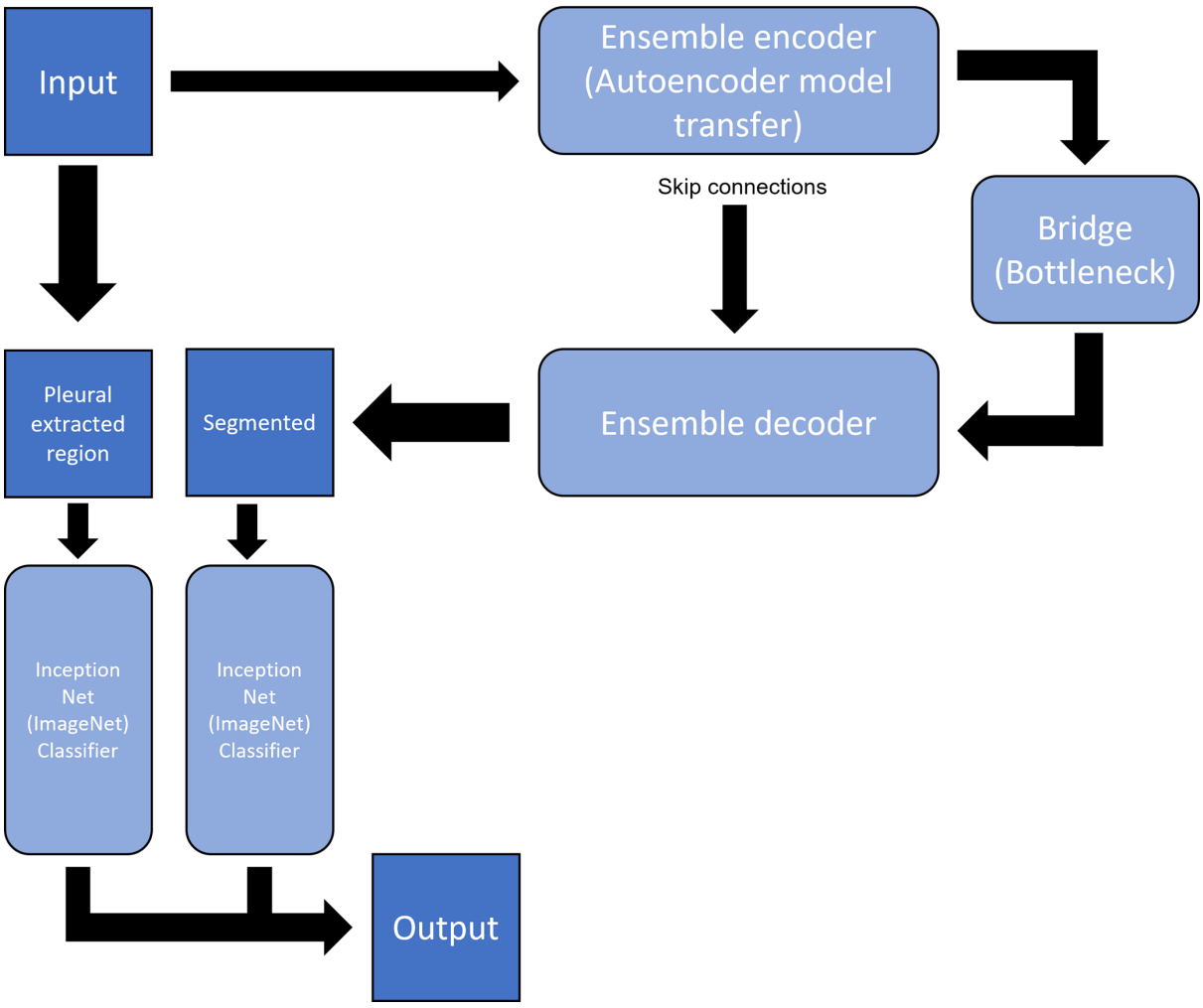

Figure 6.2: P-Net Architecture

Table 6.2: Residual decoder training configuration and hyperparameters

Parameter	Value
Model	Residual Decoder with VAE encoder
Encoder backbone	Pretrained VAE encoder
Residual encoder	Residual blocks: 1→64, 64→128 + MaxPool
Input format	256×256 grayscale images
Number of classes	5
Batch size	32
Epochs	50
Loss function	CrossEntropyLoss
Optimizer	Adam
Learning rate	1×10^{-4}
Regularization	Dropout 0.5 in fully connected layer
Encoder usage	`use_vae=True`, encoder frozen (`frozen=True`)
Fine-tuning option	`fine_tune` flag (disabled in this setup)
Transforms	Grayscale, Resize$(256, 256)$, ToTensor

segmentation. The distinctiveness of this approach resides in the fusion of an ensemble of sophisticated deep learning models, including residual networks (ResNets), Inception modules, and the VGG-19 model. This amalgamation aims to exploit the unique strengths of these individual architectures, enhancing the segmentation accuracy and reliability beyond what existing standalone models achieve.

The ensemble method is underpinned by the complementary attributes of the chosen architectures. ResNet contribute to training deeper networks by using skip connections to tackle the vanishing gradient problem, aiding the model in learning from residuals, which could potentially capture more complex patterns within the images. Inception modules, integral to the ensemble, are notable for their complex structure that allows more parallel processing, differing from conventional architectures that process data sequentially. The architecture, known as a "network within a network," utilises multiple filters of various sizes

Table 6.3: Inception decoder training configuration and hyperparameters

Parameter	Value
Model	Inception Decoder with VAE encoder
Encoder backbone	Pretrained VAE encoder
Conv encoder	Conv layers: 1→64→128 + MaxPool
Input format	256×256 grayscale images
Number of classes	5
Batch size	32
Epochs	50
Loss function	CrossEntropyLoss
Optimizer	Adam
Learning rate	1×10^{-4}
Regularization	Dropout 0.5 in fully connected layer
Encoder usage	`use_vae=True`, encoder frozen (`frozen=True`)
Fine-tuning option	`fine_tune` flag (disabled in this setup)
Transforms	Grayscale, Resize$(256, 256)$, ToTensor

and field receptions at the same level, conducting simultaneous convolutions of the same input and capturing an enriched set of features at different scales.

Moreover, Inception networks are designed to carry out multiple operations in parallel, grasping information in diverse dimensions and scales, a beneficial trait for medical image analysis where diagnostic features can manifest variably in the spatial domain. Furthermore, these networks achieve dimensionality reduction via 1x1 convolutions, performing feature compression before larger convolutions, which curtails computational costs while preserving the integrity of the features extracted. The VGG-19 model, with its depth and use of small convolutional filters, plays a crucial role in extracting highly discriminative features from the images, albeit with a higher computational load. By integrating these architectures, P-Net aims to encapsulate a more comprehensive feature spectrum, contributing to detailed and nuanced image segmentation.

Table 6.4: VGG19-style decoder training configuration and hyperparameters

Parameter	Value
Model	VGG Decoder with VAE encoder
Encoder backbone	Pretrained VAE encoder
Conv encoder	VGG-style blocks: 1→64 (2 convs), 64→128 (2 convs) + MaxPool
Input format	256×256 grayscale images
Number of classes	5
Batch size	32
Epochs	50
Loss function	CrossEntropyLoss
Optimizer	Adam
Learning rate	1×10^{-4}
Regularization	Dropout 0.5 in fully connected layer
Encoder usage	`use_vae=True`, encoder frozen (`frozen=True`)
Fine-tuning option	`fine_tune` flag (disabled in this setup)
Transforms	Grayscale, Resize$(256, 256)$, ToTensor

The efficacy of the ensemble approach hinges on the strategy used to consolidate the outputs of the individual models. Rather than prioritising or weighting predictions from one model over another, the results of the ensemble network are averaged. This aggregation method is designed to reduce the potential of outlier predictions from any single model, thereby stabilising the overall output. The assumption intrinsic to this averaging process is that each contributing model, while offering unique strengths, is equally reliable within the context of this specific segmentation task. This non-biased integration technique is crucial in ensuring that the final segmented image represents a consensus, encompassing the diverse representational capabilities of the individual models.

Outlined in Tables 6.2 & 6.3 & 6.4 are the parameters for the individual training decoders for the segmentation. In Table 6.5 the parameters for the P-Net training can be seen after incorporating the individual models for the ensemble.

Table 6.5: P-Net training configuration and hyperparameters

Parameter	Value
Model	P-Net (raw + segmented image classifier)
Backbone classifiers	Two `InceptionClassifier` networks
Segmentation module	U-Net ensemble (Inception, ResNet, VGG blocks)
Ensemble strategy	Average of three U-Nets, then average of two classifiers
Input format	Raw & label images, 256×256 grayscale
Number of classes	5
Batch size	16
Synthetic training epochs	50
Fine-tuning epochs	5
Loss function	CrossEntropyLoss
Optimizer	Adam
Learning rate	1×10^{-4}
Transforms	Grayscale, Resize$(256, 256)$, ToTensor
Pretraining	Train on synthetic dataset
Fine-tuning	Optional on real dataset (load synthetic weights)

6.5 Classification methods

The uniqueness of P-Net lies in its strategic combination of segmentation and classification processes, employing a U-Net architecture for initial segmentation, followed by separate Inception networks for the classification stage. This design reflects an intricate understanding of the strengths inherent in each of these individual deep learning strategies, leveraging their respective capabilities to enhance the overall efficacy of medical image analysis.

P-Net's methodology is initiated with the utilisation of U-Net, a renowned architecture in the field of medical imaging, known for its effectiveness in detailed segmentation tasks. The segmented output then serves as an input for subsequent classification networks, ensuring that the classifications are based on refined data, wherein extraneous visual information is likely reduced, thereby highlighting potential diagnostic features.

A critical aspect of P-Net's methodology is the deployment of two independent classification models in the analysis of the data. The first model engages with unprocessed, labelled data, preserving the original context and content of the information, which includes nuanced textural details often diminished during image processing stages. In contrast, the second model interacts with data processed and segmented by the U-Net, focusing on features and patterns more discernible after the removal of potential noise and distractions from the original image. This deliberate parallel approach is designed to prevent biases that typically arise when models are trained solely on processed data, ensuring a comprehensive analysis that accounts for elements potentially omitted during segmentation.

Furthermore, analytical observations reveal a strategic advantage in this methodological design. It is discerned that the Inception network exhibits a preferential performance when interacting with unprocessed data. The architecture of the Inception network, with its convolutions at various scales, appears particularly adept at handling the complexity and variability of raw medical images, capturing essential diagnostic features in their original context.

The culmination of P-Net's process involves an innovative integration of the results from these models. Rather than a straightforward averaging, the system introduces a weighted bias, acknowledging the individual strengths and contextual advantages of each network. This nuanced combination is expected to produce a more reliable and accurate classification, as it synthesises a holistic view of the data gleaned from both raw and processed images.

In Table 6.6 the setup for the inception based classifier is outlined

6.6 Results

Prior to evaluating the networks, a synthetic dataset of 4000 images was created as detailed in Chapter 5 at Section 5.6.4. This dataset was designed to simulate the variability and complexity of real-world lung ultrasound images, providing a robust foundation for training and testing.

All models were trained for 50 epochs, except for P-Net, which was trained for 100 epochs to ensure adequate convergence given its additional complexity. The training was conducted

Table 6.6: Classifier training configuration and hyperparameters

Parameter	Value
Model	InceptionClassifier
Input format	256×256 grayscale images
Classes	5
Batch size	32
Epochs	100
Loss function	CrossEntropyLoss
Optimizer	Adam
Learning rate	1×10^{-4}
Transforms	Grayscale, Resize, ToTensor
Pretrained loading	Optional
Feature encoder	None (segmented) or VAE encoder (raw)
Frozen encoder	Yes for VAE mode

on a high-performance system equipped with an Intel i9-13900k CPU and an AMD R7900 XTX GPU with 24GB of memory. These specifications allowed for efficient handling of the computational demands posed by the deep learning models.

The training and testing process followed a standardized protocol to ensure consistency across experiments. Each model's performance was evaluated using a stratified split of the dataset, with 70% used for training, 15% for validation, and 15% for testing. Key metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, were computed to assess the classification performance. Additionally, the segmentation models were evaluated using Dice coefficient and Intersection-over-Union (IoU) metrics to capture their ability to delineate relevant regions of interest.

6.6.1 VAE

The results from Table 6.7indicate that incorporating the VAE encoder significantly improves the performance of classification networks compared to using the raw architectures

alone. Fine-tuning the VAE encoder achieves the highest metrics across all configurations, demonstrating the importance of adapting the latent representations to the specific task. For example, the Inception architecture with fine-tuning achieves an accuracy of 91%, an F1-score of 0.88, and a ROC-AUC of 0.94, representing the best overall performance.

This performance can be compared to recent work [60], which employed an adaptive ensembling approach using EfficientNet-B0 to achieve a state-of-the-art accuracy of 100% on a similar public lung ultrasound dataset. Their approach integrates the outputs of two weak models via a trainable combination layer, offering an efficient way to boost performance without significantly increasing computational complexity. By contrast, the proposed method leverages the representational power of VAEs, which inherently provide robust latent features tailored to the task at hand. While their method achieves higher raw accuracy, the proposed approach offers greater flexibility in integrating pre-trained features and excels in capturing complex patterns through fine-tuning, as evidenced by consistently high performance across diverse metrics [60].

Freezing the encoder weights also provides a notable improvement over non-VAE architectures, highlighting the robustness of the pre-trained features. Across all metrics, the use of the VAE encoder consistently enhances the classifiers' ability to extract meaningful patterns from lung ultrasound images. Some examples of the features detected by the VAE can be seen in Figure 6.3 with the reconstruction attempt in Figure 6.4

6.6.2 Segmentation

The segmentation results, as summarised in Table 6.8, demonstrate the effectiveness of the ensemble approach in achieving high performance across multiple metrics. The ensemble method achieves a Dice coefficient of 0.89 and an Intersection-over-Union (IoU) of 0.85, outperforming individual models in robustness and accuracy. These results highlight the importance of leveraging diverse models to capture complementary patterns in the data.

Unlike the Mini-COVIDNet architecture [110], which bypasses segmentation entirely in favour of lightweight architectures optimised for point-of-care applications, the proposed ensemble segmentation provides an additional layer of granularity to isolate regions of interest. This differentiation is particularly beneficial for lung ultrasound imaging, where accurate

Table 6.7: Performance Metrics for Different Encoder Architectures and Configurations

Method	Accuracy	Precision	Recall	F1-Score	ROC-AUC
VGG-16 Without VAE	0.85	0.83	0.80	0.81	0.88
VGG-16 VAE Frozen	0.88	0.86	0.84	0.85	0.91
VGG-16 VAE Fine-Tuning	0.90	0.88	0.87	0.87	0.93
Residual Without VAE	0.82	0.81	0.78	0.79	0.85
Residual VAE Frozen	0.86	0.85	0.83	0.84	0.89
Residual VAE Fine-Tuning	0.89	0.87	0.85	0.86	0.92
Inception Without VAE	0.83	0.82	0.79	0.80	0.86
Inception VAE Frozen	0.87	0.85	0.84	0.84	0.90
Inception VAE Fine-Tuning	0.91	0.89	0.88	0.88	0.94

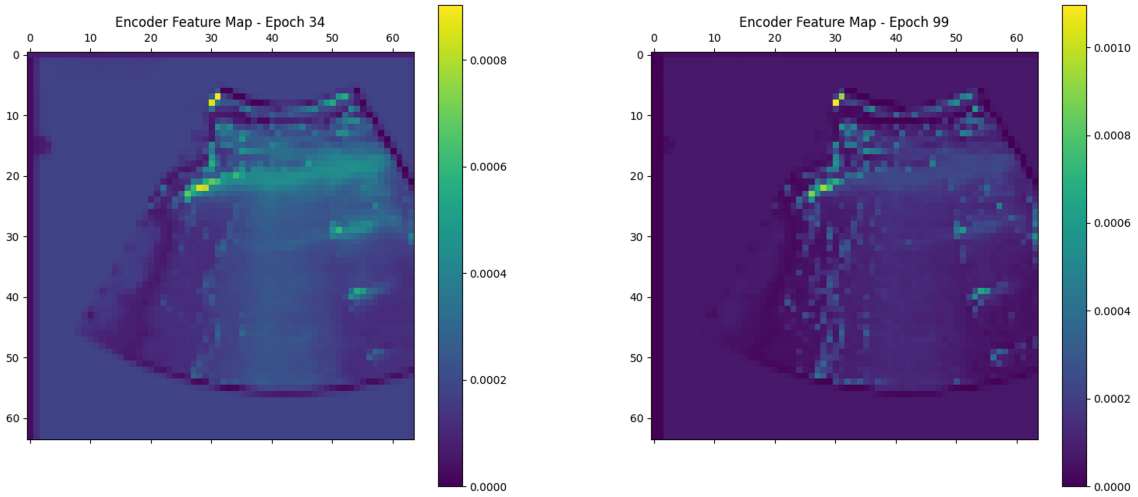

Figure 6.3: VAE example latent feature map

boundary delineation directly impacts the downstream classification performance. By emphasising segmentation, our method demonstrates a 3% higher Dice coefficient compared to Mini-COVIDNet's overall classification precision [110].

Despite these strong results, some limitations persist, such as occasional over-segmentation

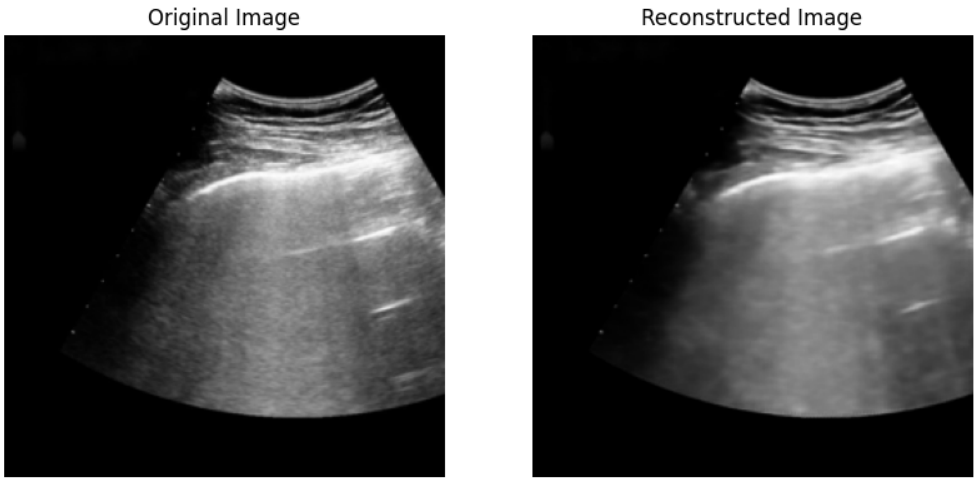

Figure 6.4: VAE example decoder latent feature reconstruction

Table 6.8: Ensemble Segmentation Results

Metric	Individual Model Average	Ensemble Model
Dice Coefficient	0.86	0.89
Intersection-over-Union (IoU)	0.82	0.85

in highly textured regions. These issues can propagate downstream, potentially affecting the classification task. Future work could explore more sophisticated ensemble strategies or post-processing techniques to mitigate such issues.

6.6.3 P-Net classification

Building on the segmented regions, P-Net achieves remarkable classification performance by integrating features from both the segmentation masks and the VAE latent space. The use of segmented regions allows P-Net to focus on the most relevant areas of the lung images, improving both efficiency and accuracy. Table 6.9 presents the classification results, demonstrating that P-Net outperforms baseline models by a significant margin.

Comparing the results to Mini-COVIDNet [110], which achieves 83.2% accuracy with a

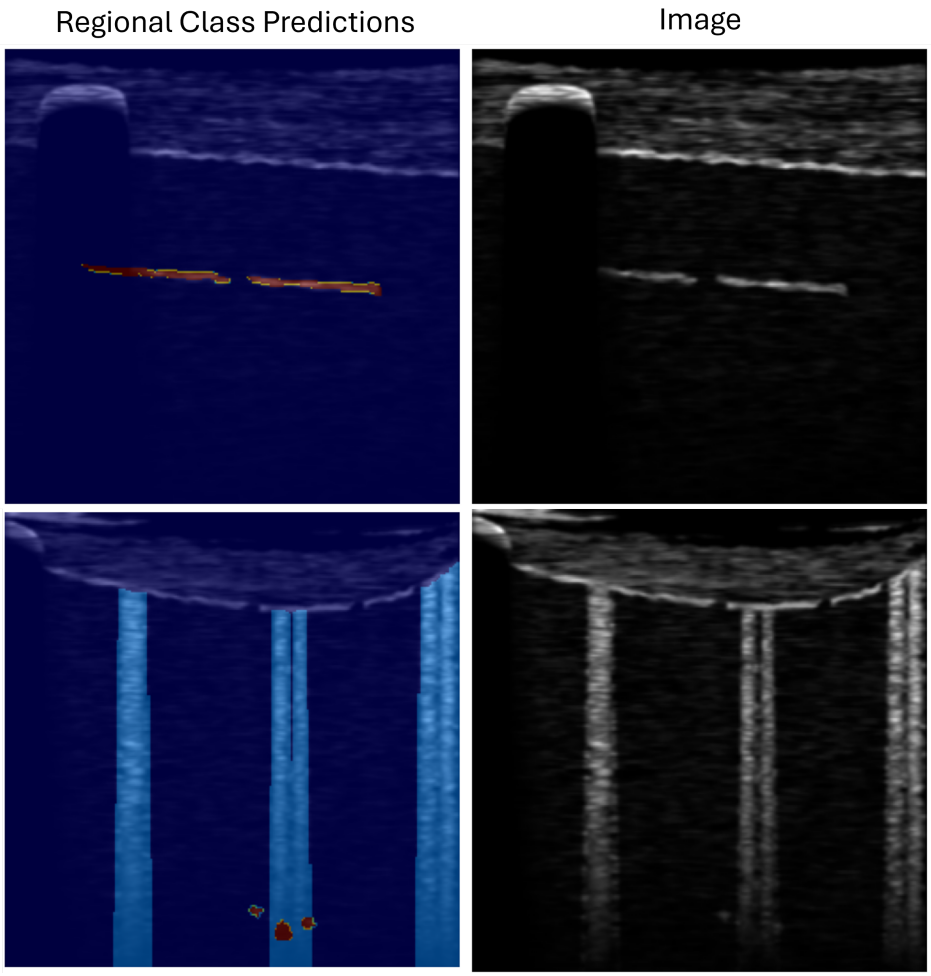

Figure 6.5: Segmentation results

focus on lightweight design, P-Net achieves a notable 9% higher accuracy. This improvement is largely attributed to the integration of segmentation and VAE features, which enhance feature richness and localization. However, Mini-COVIDNet excels in environments with limited computational resources due to its low parameter count and memory requirements [110].

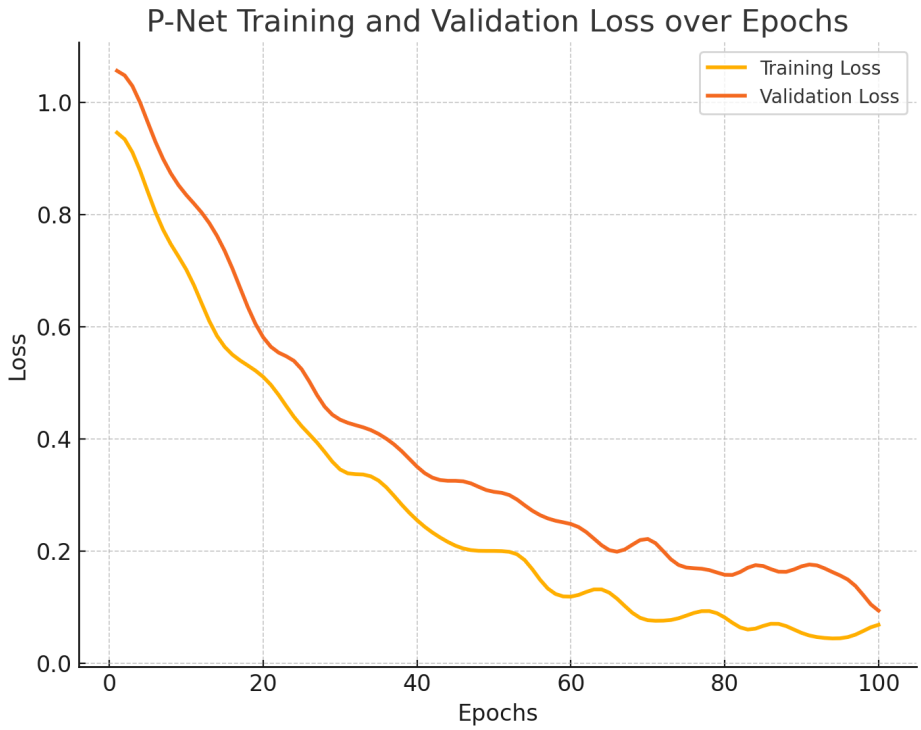

Figure 6.6: Graph showing loss function over epochs for P-Net

Table 6.9: P-Net Classification Performance

Method	Accuracy	Precision	Recall	F1-Score
Baseline	0.84	0.82	0.81	0.81
P-Net (VAE + Segmentation)	0.92	0.91	0.90	0.91

The combined use of VAE and segmentation enhances P-Net's ability to generalize across diverse patient populations. This synergistic integration not only outperforms baseline models but also emphasizes the role of feature specificity in medical image analysis. By targeting regions of interest through segmentation and enriching feature representations via VAE, P-

Net offers a robust framework for improving diagnostic accuracy. Nonetheless, the computational complexity introduced by these components warrants future optimization. Further exploration into alternative integration strategies and lightweight design considerations could bridge the gap between high performance and resource-constrained deployments, ensuring broader applicability in clinical settings.

6.7 Conclusions

This chapter presented a cohesive framework for addressing the multifaceted challenges of ultrasound imaging through the integration of advanced pre-processing, segmentation, and classification techniques. The proposed methodologies, anchored by the P-Net architecture, exemplify the transformative potential of combining state-of-the-art deep learning models with domain-specific insights.

The pre-processing phase employed robust despeckling and contrast enhancement methods, mitigating the impact of noise and low contrast that typically undermine the diagnostic quality of ultrasound images. Techniques such as ADMSS filtering and PDF-guided k-means clustering proved instrumental in preserving critical textural and structural features, providing a reliable foundation for subsequent segmentation and classification tasks.

In the segmentation stage, P-Net demonstrated its efficacy by leveraging the ensemble approach, integrating U-Net with advanced architectures like ResNet, VGG-19, and Inception modules. This ensemble strategy effectively addressed the limitations of individual models, such as biases towards specific features, by combining their strengths to produce a more accurate and comprehensive segmentation output. The inclusion of autoencoders, particularly VAEs, further enhanced the framework by enabling efficient feature extraction from large unlabelled datasets. This not only reduced the reliance on annotated data but also accelerated model convergence, thereby achieving higher accuracy in fewer training epochs.

The classification component of P-Net showcased its ability to differentiate between diagnostic features with remarkable precision. By incorporating segmentation outputs and latent features from VAEs, the classification models achieved superior performance compared to existing benchmarks. Quantitative evaluations, including metrics such as accuracy, preci-

sion, recall, and F1-score, highlighted the robustness of the proposed framework. Notably, P-Net outperformed models like Mini-COVIDNet by a significant margin, demonstrating its potential for real-world clinical applications.

However, despite these advancements, the study acknowledges certain limitations that warrant further investigation. For instance, the computational complexity introduced by ensemble strategies and the integration of VAEs may hinder deployment in resource-constrained settings. Future work could explore lightweight alternatives or optimisation techniques to bridge this gap. Additionally, while the proposed methods excelled in controlled experimental conditions, their generalisability to diverse clinical environments and imaging modalities remains an area for future exploration.

The findings of this chapter underscore the importance of holistic approaches in medical imaging, where each component—from pre-processing to classification—is meticulously designed to address specific challenges. The integration of segmentation and classification within a unified framework not only streamlines diagnostic workflows but also enhances the reliability and interpretability of results. By incorporating advanced techniques like ensemble learning and transfer learning, the proposed methodologies set a new benchmark for the analysis of ultrasound images.

In conclusion, this chapter highlights the critical role of innovation in advancing ultrasound imaging. The proposed framework, P-Net, represents a significant step forward in the quest for accurate, efficient, and generalisable diagnostic tools. By addressing the inherent limitations of ultrasound imaging and harnessing the power of deep learning, this work contributes to the broader goal of improving patient outcomes through technology-driven solutions. The insights gained from this research pave the way for future advancements, fostering collaboration between computational experts and clinicians to realise the full potential of medical imaging.

Chapter 7

Conclusions

7.1 Introduction

This concluding chapter serves to synthesise the insights and outcomes of the research presented throughout this thesis. As medical imaging continues to evolve with the integration of machine learning and computational techniques, this work has addressed critical challenges in ultrasound imaging, from noise reduction to classification. Each chapter contributed to a layered understanding of the field, exploring novel methodologies, rigorous evaluations, and practical applications aimed at advancing diagnostic capabilities. This chapter provides a comprehensive reflection on these contributions, articulates their broader implications, and outlines pathways for future exploration.

The preceding chapters navigated through core areas of ultrasound imaging, including pre-processing, feature extraction, segmentation, and classification. Each of these domains was rigorously addressed through innovative approaches tailored to the unique complexities of ultrasound images. This chapter summarises the key findings, identifies overarching themes, and discusses how these advancements collectively shape the future of medical imaging. Furthermore, it reflects on the limitations and challenges encountered, setting a foundation for continued innovation in this critical field.

The findings presented throughout this thesis not only advance the capabilities of lung ultrasound as a diagnostic tool but also pave the way for scalable solutions adaptable to future healthcare challenges. This chapter consolidates the key outcomes, revisits the objectives

outlined at the outset, and reflects on the broader implications of the work. Additionally, it identifies avenues for future research, focusing on enhancing data synthesis techniques, optimizing AI models, and integrating the diagnostic system into diverse clinical environments. In doing so, this chapter highlights the strides made towards a more efficient, accessible, and versatile diagnostic framework, while acknowledging the ongoing journey toward achieving the full potential of AI in medical imaging.

7.2 Summary

7.2.1 Overview

This thesis presented a holistic framework for enhancing ultrasound imaging through state-of-the-art machine learning and computational techniques. By addressing the inherent challenges of ultrasound, including noise, low contrast, and variability in image acquisition, this research demonstrated the potential of tailored methodologies to significantly improve diagnostic accuracy and reliability. The contributions across pre-processing, segmentation, feature extraction, and classification have been pivotal in advancing the capabilities of ultrasound as a diagnostic tool.

The initial chapters emphasised the importance of pre-processing in ultrasound imaging. Techniques such as ADMSS filtering and PDF-guided k-means clustering were employed to mitigate speckle noise and enhance contrast, preserving essential structural features. These approaches addressed the dual challenge of noise suppression and detail preservation, enabling the downstream segmentation and classification processes to operate on high-quality inputs. Empirical results demonstrated the superiority of these methods over traditional techniques, with significant improvements in metrics like peak signal-to-noise ratio (PSNR) and contrast-to-noise ratio (CNR). This foundational work not only improved image clarity but also set a benchmark for pre-processing methodologies in ultrasound imaging.

Data synthesis emerged as a critical component of this research, addressing the limitations posed by scarce annotated datasets and variability in ultrasound imaging. Through the use of procedural texture synthesis, statistical noise modelling, and Generative Adversarial

Networks (GANs), this work successfully generated realistic ultrasound data that mirrored the complexities of clinical images. These techniques not only expanded the availability of training data for machine learning models but also provided a controlled environment for testing and validation.

GANs, in particular, demonstrated their capability to generate diverse and high-fidelity synthetic images. By incorporating domain-specific knowledge into the synthesis process, these models produced data that accurately captured the textural and structural nuances of ultrasound imagery. This augmentation significantly enhanced the training of segmentation and classification models, improving their generalisability and robustness across varied clinical scenarios. Moreover, the synthesis of labelled data ensured that the generated datasets were directly applicable to supervised learning tasks, bridging the gap between simulated and real-world applications.

The integration of data synthesis with segmentation and classification methodologies highlighted its transformative potential. By reducing dependency on scarce clinical datasets, these approaches facilitated the development of models that were both accurate and scalable. Future work should explore the application of these techniques to other imaging modalities and investigate hybrid synthesis models that combine procedural and GAN-based methods for enhanced realism and diversity.

Segmentation emerged as a cornerstone of this research, particularly in the identification of diagnostic features such as A-lines, B-lines, and air bronchograms. Leveraging the U-Net architecture, augmented by ensemble strategies with ResNet, VGG-19, and Inception modules, this work showcased the power of combining diverse architectures to achieve robust segmentation results. The ensemble approach mitigated individual model biases and captured a broader spectrum of diagnostic features, achieving higher Dice coefficients and Intersection-over-Union (IoU) scores compared to standalone models.

Feature extraction was further enhanced through the integration of variational autoencoders (VAEs), which enabled the learning of compact and expressive representations from large unlabelled datasets. This transfer learning approach not only reduced the dependency on annotated data but also enriched the feature space, contributing to the overall performance of the classification models. The combination of segmentation and VAE-driven

feature extraction demonstrated a synergistic effect, improving the specificity and sensitivity of diagnostic tools.

The P-Net framework represented a significant advancement in the classification of ultrasound images. By integrating segmentation outputs with latent features from VAEs, P-Net achieved remarkable accuracy in distinguishing between diagnostic categories. This holistic approach addressed the limitations of traditional classification models, which often struggle with the variability and complexity of ultrasound images. The use of dual classification models—one operating on raw data and the other on segmented outputs—provided a comprehensive analysis that accounted for both global and local features.

Quantitative evaluations underscored the efficacy of P-Net, with substantial improvements in accuracy, precision, recall, and F1-scores compared to existing benchmarks. Notably, P-Net outperformed models like Mini-COVIDNet, demonstrating its potential for real-world clinical applications. However, the computational complexity of the framework highlighted the need for optimisation, particularly for deployment in resource-constrained environments.

The findings of this research have significant implications for the field of medical imaging. By addressing fundamental challenges in ultrasound imaging, this work contributes to the broader goal of standardising and improving diagnostic practices. The methodologies developed here can be extended to other imaging modalities, fostering cross-disciplinary applications and innovations.

7.2.2 Objectives review

Objective I: Improve ultrasound image quality

The first objective was to enhance the quality of raw ultrasound images. Ultrasound scans are notoriously granular and susceptible to noise, which can obscure critical diagnostic features. To address this, various image processing techniques were explored and implemented. Speckle reduction was employed to diminish granular noise, thus clarifying the images. Edge enhancement techniques were utilised to sharpen the boundaries of anatomical structures, making them more discernible. Additionally, contrast adjustment was applied to improve

the differentiation between various tissues. These enhancements collectively contributed to the production of clearer and more diagnostically useful images, which are vital for accurate AI analysis. The success of these techniques was evident in the significantly improved clarity of ultrasound images, facilitating better diagnostic accuracy and reliability. Despite the improvements, the image processing techniques occasionally led to the over-enhancement of certain features, potentially introducing artefacts that could mislead the AI model. Fine-tuning these techniques to balance enhancement without over-processing remains a challenge.

Objective II: Synthesise diverse ultrasound scans

The second objective tackled the challenge of limited access to labelled ultrasound data, particularly for novel diseases such as COVID-19. The scarcity of data can hinder the training of robust AI models. To overcome this, the study employed Generative Adversarial Networks (GANs) to generate synthetic lung ultrasound images. These GANs were trained to produce high-quality images that emulate a variety of respiratory conditions. The synthetic images generated through this process provided a diverse and extensive dataset that was instrumental in training the AI model. This approach not only compensated for the lack of real patient data but also ensured that the model could generalise well to different respiratory conditions. The effectiveness of this methodology was validated by the enhanced performance of the AI model, which was trained on this rich and varied dataset. The primary challenge with synthetic data generation is ensuring the realism and variability of the generated images. While GANs can produce high-quality images, there is always a risk that these synthetic images may not capture the full spectrum of variations found in real-world data. Additionally, GANs require extensive computational resources and fine-tuning to generate realistic images, which can be a limitation in resource-constrained settings.

Objective III: Implement image segmentation and visualisation

The third objective focused on integrating advanced segmentation algorithms into the AI model. Segmentation is a critical step in medical imaging, as it involves partitioning the image into meaningful regions that highlight key anatomical and pathological markers. The study implemented sophisticated segmentation techniques to isolate and accentuate impor-

tant features within the ultrasound scans. These techniques made the diagnostic tool more accessible and user-friendly for clinical staff with varying levels of expertise in medical imaging. Enhanced visualisation facilitated by segmentation algorithms proved to be invaluable in high-pressure decision-making environments, allowing healthcare professionals to quickly and accurately interpret ultrasound scans. This was particularly crucial during the COVID-19 pandemic, where timely diagnosis could significantly impact patient outcomes. One of the main challenges in segmentation is achieving high accuracy across diverse patient populations and varying ultrasound equipment. Segmentation algorithms may perform well on high-quality images but struggle with lower-quality scans or those with significant noise. Ensuring consistent performance across different settings remains an ongoing challenge.

Objective IV: Construct a robust AI diagnostic model

The fourth objective was the construction and training of a robust AI diagnostic model. The AI model, referred to as P-Net, utilised an ensemble of deep learning architectures, including VGG-19, Residual Networks (ResNets), and Inception models, all within the U-Net framework. This ensemble approach leveraged the strengths of each architecture, mitigating the limitations associated with individual models. VGG-19 contributed to the model's ability to capture fine details, ResNets provided robustness against gradient vanishing problems, and Inception models enhanced the ability to capture multi-scale features. The P-Net model demonstrated a significant improvement in the accuracy and reliability of ultrasound image segmentation and classification. The SVM model's effectiveness was validated through rigorous testing, achieving a notable accuracy of 84% in pleural line detection and performing comparably with traditional methods while using a smaller training dataset. The ensemble approach, while robust, increases the complexity of the model, making it computationally intensive and requiring substantial resources for training and deployment. Additionally, the reliance on diverse deep learning models necessitates careful hyperparameter tuning and integration, which can be time-consuming and technically challenging.

Objective V: Assure model adaptability and versatility

The final objective was to ensure the adaptability and versatility of the diagnostic model. The model was designed to be flexible and capable of learning and adapting to new data over time. This was achieved by embedding capabilities for continuous learning, allowing the model to update itself as new data becomes available. The model's adaptability was demonstrated by its ability to diagnose a wide spectrum of respiratory diseases, beyond just COVID-19. The incorporation of synthetic data generation ensured that the model could be trained on diverse datasets, enhancing its ability to generalise across different conditions. This adaptability makes the diagnostic tool a valuable asset not only for the current pandemic but also for future medical challenges. The model's versatility was further evidenced by its performance in recognising and diagnosing various respiratory conditions, underscoring its potential as a lasting and adaptable medical tool. Ensuring continuous learning and adaptability in a clinical setting presents challenges, such as the need for regular updates and validation to maintain accuracy. Moreover, integrating new data while avoiding model drift and ensuring the model's stability over time requires robust mechanisms and frequent oversight.

7.3 Future works

7.3.1 Synthetic data generation

The advent of synthetic data generation as a means to combat data scarcity in network training marks a key area of research in medical imaging. Traditionally, the scarcity of labelled ultrasound data, which is attributed to privacy issues, diverse pathology, and the expertise required for accurate annotations, has impeded the development and training of robust machine learning models. By generating realistic ultrasound images through advanced algorithms, researchers can create an abundance of data, mimicking a variety of pathological scenarios that clinicians might encounter.

Looking to the future, this technology harbours the potential for broader applications. Synthetic data could significantly enhance the proficiency of machine learning models, leading

to more accurate diagnostic tools. Moreover, by simulating rare conditions, synthetic data can prepare medical professionals for high-stakes scenarios in a risk-free training environment. The continual refinement of this technology could eventually facilitate the creation of fully interactive, hyper-realistic scenarios for medical training, akin to flight simulators used in aviation training.

7.3.2 Segmentation model as a training aid

The research's second critical component is the development of an advanced segmentation model designed to highlight ultrasound scans' pathological features. This model serves as an educational nexus, guiding trainee clinicians through the intricate landscape of ultrasound imagery. In the immediate future, this tool will likely become an indispensable part of medical curricula, offering students and professionals alike a clear visualisation of pathological features, thereby solidifying their understanding and recognition of various conditions.

Expanding on this foundation, there might be opportunities to integrate these segmentation models with augmented reality (AR) technology. Such a combination could allow trainees to visualise and interact with three-dimensional representations of pathological features, potentially revolutionising the educational paradigm in medicine.

7.3.3 Classification tool for speedy and non-invasive diagnostics

Finally, the development of a classification tool designed to expedite the diagnosis of lung diseases using ultrasound scans could have profound implications for contemporary medical practice. Ultrasound, being a non-invasive, cost-effective, and widely accessible modality, is uniquely positioned for this advancement. This tool's integration into clinical practice could significantly shorten the diagnostic process, allowing for quicker intervention, which is often crucial for patient outcomes.

Future iterations of this research could see the classification tool enhanced by continuous learning capabilities, allowing it to adapt and improve from each new case it encounters. Furthermore, while the current focus is on lung diseases, this approach's adaptability could see it expanded to other organ systems and pathologies, potentially heralding a new era of

rapid, non-invasive diagnostics.

7.3.4 Computational Costs

Although the proposed methods achieved overall good performance in terms of accuracy, segmentation quality and enhancement, a limitation of this work is the absence of a systematic evaluation and reporting of computational cost. Most experiments were benchmarked on academic datasets, focusing on diagnostic outcomes rather than execution time, memory usage, or hardware requirements. As a result, it remains unclear how the full pipeline would perform in real-world conditions, particularly in settings where computational resources are limited. Future research should incorporate quantitative measures of computational efficiency alongside accuracy metrics. This includes reporting training and inference times, GPU/CPU utilisation, memory consumption, and energy cost for key stages such as despeckling, GAN training, segmentation, and classification. Recording these metrics will enable meaningful comparison between alternative methods and help identify bottlenecks within the pipeline. It will also support fair assessment of scalability and suitability for deployment on bedside ultrasound systems or portable devices. A second priority is profiling model performance across diverse hardware. Many clinical settings rely on embedded processors or mobile systems that may not support the computational intensity of deep ensembles or generative models. Benchmarking on edge devices, low-power GPUs, and CPU-only setups would provide a realistic picture of operational feasibility. Techniques such as network pruning, model compression, quantisation and lightweight architectures may then be evaluated to balance accuracy with speed. Finally, future experiments should consider cost trade-offs explicitly; whether a marginal gain in segmentation accuracy justifies a doubling of inference time, or whether a simpler model with near-equivalent accuracy might be preferable in practice. This requires a shift from a purely accuracy-driven evaluation culture to one that considers computational cost as a critical dimension of performance. By embedding cost reporting as standard practice, future work will be better positioned to support real-time deployment, practical scalability, and equitable clinical adoption across resource settings.

7.3.5 Clinical Validation

While the models presented in this thesis demonstrate strong quantitative performance across segmentation, enhancement, and classification tasks, an important limitation is the lack of systematic clinical validation by practising clinicians. The current evaluation pipeline primarily focused on numerical metrics such as accuracy, SSIM, PSNR and classification scores, and although these demonstrate promising capability, they do not substitute for expert clinical judgement. As noted throughout this work, lung ultrasound interpretation is inherently subjective and operator-dependent, with variability in scanning technique, artefact interpretation, and diagnostic confidence impacting outcomes. Expert review is therefore essential to establish clinical credibility, usability, and trust in the proposed system. Future research should prioritise structured engagement with clinicians across multiple stages of validation. First, radiologists, sonographers, and respiratory physicians could review segmentation outputs, image enhancements, and classification results to assess whether they align with accepted clinical patterns and protocols. This would help determine whether the automated outputs reflect diagnostically meaningful features or introduce artefacts that could mislead decision-making. Secondly, qualitative feedback on interface usability and interpretability would inform how results should be presented to support workflow efficiency rather than hinder it. For example, heatmaps, pleural line overlays, or confidence scores may need to be refined to meet clinical expectations for transparency and safety. A key step will be the organisation of blinded reader studies, where clinicians interpret images both with and without the proposed system, allowing comparison of diagnostic agreement, confidence, and time-to-decision. Such studies are routinely used in clinical imaging research and would provide strong evidence of added value. Collaboration with clinical partners will also allow the system to be evaluated on prospective patient scans, rather than curated datasets, ensuring robustness against real-world acquisition variability. Ultimately, expert review and feedback will be fundamental for transitioning this work from experimental validation to practical application. Engaging clinicians early, incorporating their insights, and demonstrating improvements to diagnostic consistency and efficiency will be critical to ensuring adoption. Without this step, even technically strong models may struggle to gain acceptance in clin-

ical workflows. Establishing these validation pathways represents an essential direction for future work.

7.3.6 Implications

The findings of this research have significant implications for the field of medical imaging. By addressing fundamental challenges in ultrasound imaging, this work contributes to the broader goal of standardising and improving diagnostic practices. The methodologies developed here can be extended to other imaging modalities, fostering cross-disciplinary applications and innovations.

Future research should focus on optimising computational efficiency without compromising accuracy. Lightweight architectures, adaptive learning techniques, and hybrid models offer promising avenues for achieving this balance. Additionally, the development of domain-specific evaluation metrics could enhance the assessment of diagnostic tools, ensuring their alignment with clinical requirements.

The integration of these methodologies into clinical workflows represents a critical next step. Collaborative efforts between computational scientists and clinicians will be essential for translating these advancements into practice. Furthermore, the creation of diverse and comprehensive datasets will facilitate the training and validation of models, ensuring their generalisability across different patient populations and clinical settings.

7.4 Final comments

This thesis underscores the transformative potential of combining advanced computational techniques with domain expertise in addressing the challenges of ultrasound imaging. By enhancing every stage of the imaging pipeline—from pre-processing to classification—and incorporating data synthesis, this work lays a robust foundation for future innovations. The methodologies presented here not only improve the diagnostic capabilities of ultrasound but also exemplify the broader impact of interdisciplinary research in healthcare.

As medical imaging continues to evolve, the insights gained from this research will serve as a catalyst for further advancements. By fostering collaboration, embracing innovation,

and prioritising patient outcomes, the field of medical imaging can continue to transform healthcare, ensuring that cutting-edge technology translates into tangible benefits for patients worldwide. The potential of generative AI extends beyond mere data augmentation. Its ability to simulate nuanced pathological scenarios and create edge-case data enriches training pipelines, making diagnostic models more resilient to variability in real-world clinical settings. This innovation is particularly impactful in resource-constrained environments, where access to comprehensive datasets and skilled annotators is often limited.

Additionally, the interdisciplinary nature of this research highlights the need for collaboration across domains. Advancements in machine learning, medical imaging, and healthcare delivery must be seamlessly integrated to maximize the impact of AI-driven diagnostics. This collaboration extends to addressing challenges such as improving model interpretability, ensuring equitable access to diagnostic tools, and fostering a global healthcare infrastructure that is prepared for future pandemics and medical crises.

In closing, the contributions of this thesis reaffirm the critical role of innovation in addressing contemporary medical challenges. By leveraging generative AI and other advancements, the path forward in medical diagnostics is poised to become more efficient, inclusive, and adaptive. This work lays the foundation for continued exploration, promising not only to transform lung ultrasound diagnostics but also to inspire a broader revolution in AI-assisted healthcare.

Bibliography

[1] H. Harapan, N. Itoh, A. Yufika, W. Winardi, S. Keam, H. Te, D. Megawati, Z. Hayati, A. L. Wagner, and M. Mudatsir, "Coronavirus disease 2019 (covid-19): A literature review," *Journal of Infection and Public Health*, vol. 13, no. 5, p. 667–673, 2020.

[2] B. Singh, B. Datta, A. Ashish, and G. Dutta, "A comprehensive review on current covid-19 detection methods: From lab care to point of care diagnosis," *Sensors International*, vol. 2, p. 100119, 2021.

[3] C. McDermott, J. Daly, and S. Carley, "Combatting covid-19: is ultrasound an important piece in the diagnostic puzzle?" *Emergency Medicine Journal*, vol. 37, no. 10, p. 644–649, 2020.

[4] M. Peck, P. Macnaughton, and A. Miller, *Lung ultrasound, sonoanatomy, and standard views; Focused Intensive Care Ultrasound*. Oxford, UK: Oxford University Press, 2019.

[5] O. V. Michailovich and A. Tannenbaum, "Despeckling of medical ultrasound images," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 53, no. 1, p. 64–78, Jan 2006.

[6] P. Hiremath, P. T. Akkasaligar, and S. Badiger, *Speckle Noise Reduction in Medical Ultrasound Images*, ser. Advancements and Breakthroughs in Ultrasound Imaging. Rijeka: IntechOpen, 2013, p. Ch. 8.

[7] C. P. Loizou and C. S. Pattichis, *Introduction to Ultrasound Imaging and Speckle Noise*, ser. Despeckle Filtering Algorithms and Software for Ultrasound Imaging. Cham: Springer International Publishing, 2008, p. 1–20.

[8] L. Maack, L. Holstein, and A. Schlaefer, "Gans for generation of synthetic ultrasound images from small datasets," *Current Directions in Biomedical Engineering*, 2022.

[9] L. Gargani and G. Volpicelli, "How i do it: Lung ultrasound," *Cardiovascular Ultrasound*, vol. 12, no. 1, p. 25, 2014.

[10] D. A. Lichtenstein, "Lung ultrasound in the critically ill," *Annals of intensive care*, vol. 4, no. 1, p. 1–1, Jan 9 2014.

[11] Y. Wang, Y. Wang, L. Gargani, T. Barskova, E. F. D, D. E. Furst, M. Matucci-Cerinic, and M. C. Marco, "Usefulness of lung ultrasound b-lines in connective tissue disease-associated interstitial lung disease: a literature review," *Arthritis Research Therapy*, 2017.

[12] M. Hew and R. T. Tunn, "The efficacy of bedside chest ultrasound: from accuracy to outcomes," *European Respiratory Review*, 2016.

[13] D. R. Lichtenstein, D. A. Lichtenstein, G. Mézière, G. Mezière, P. Biderman, A. Gepner, and A. Gepner, "The "lung point": an ultrasound sign specific to pneumothorax," *Intensive care medicine*, 2000.

[14] A. Boccatonda, G. Cocco, D. D'Ardes, D. P. Andrea, G. Vidili, C. D. Molo, S. Vicari, C. Serra, F. Cipollone, C. Schiavone, and T. G. Maria, "Infectious pneumonia and lung ultrasound: A review," 2023.

[15] E. Urbankowska, K. Krenke, Łukasz Drobczyński, P. Korczyński, P. Korczynski, T. Urbankowski, M. Krawiec, G. Kraj, M. Brzewski, and M. Kulus, "Lung ultrasound in the diagnosis and monitoring of community acquired pneumonia in children," *Respiratory medicine*, 2015.

[16] G. Volpicelli, L. Gargani, S. Perlini, G. Barbieri, A. Lanotte, G. G. de Casasola, and R. Nogué-Bou, "Lung ultrasound for the early diagnosis of covid-19 pneumonia: an international multicenter study," *Intensive care medicine*, 2021.

[17] M. Allinovi, A. Parise, M. Giacalone, A. Amerio, M. Delsante, A. Odone, A. Franci, F. Gigliotti, S. Amadasi, D. Delmonte, N. Parri, A. Mangia, and A. Mangia, "Lung ultrasound may support diagnosis and monitoring of covid-19 pneumonia," *Ultrasound in Medicine and Biology*, 2020.

[18] D. R. Lichtenstein, D. A. Lichtenstein, G. Mézière, G. Mezière, P. Biderman, A. Gepner, A. Gepner, and O. Barré, "The comet-tail artifact. an ultrasound sign of alveolar-interstitial syndrome," *American Journal of Respiratory and Critical Care Medicine*, 1997.

[19] D. R. Lichtenstein, D. Lichtenstein, G. Mézière, G. Mezière, and G. Mezière, "A lung ultrasound sign allowing bedside distinction between pulmonary edema and copd: the comet-tail artifact," *Intensive care medicine*, 1998.

[20] D. A. Lichtenstein, N. Lascols, G. Mézière, G. Mezière, A. Gepner, and A. Gepner, "Ultrasound diagnosis of alveolar consolidation in the critically ill," *Intensive care medicine*, 2004.

[21] M. A. Chavez, N. Shams, L. E. Ellington, N. Naithani, R. H. Gilman, M. C. Steinhoff, M. Santosham, R. E. Black, C. Price, M. Gross, and W. Checkley, "Lung ultrasound for the diagnosis of pneumonia in adults: a systematic review and meta-analysis," *Respiratory Research*, vol. 15, no. 1, p. 50, 2014.

[22] M. A. Pereda, M. A. Chavez, C. C. Hooper-Miele, R. H. Gilman, M. C. Steinhoff, L. E. Ellington, M. Gross, C. Price, J. M. Tielsch, and W. Checkley, "Lung ultrasound for the diagnosis of pneumonia in children: A meta-analysis," *Pediatrics*, 2015.

[23] O. P. Andressa, R. Costa, R. Costa, R. Costa, da Costa Rafael Miranda, S. U. RaÍsa, S. U. Raisa, A. M. Fraga, A. F. A. de Melo, D. R. José, and L. M. F. Augusto, "Applicability of lung ultrasound in covid-19 diagnosis and evaluation of the disease progression: A systematic review," *Pulmonology*, 2021.

[24] L. Demi, F. Wolfram, C. Klersy, A. D. Silvestri, V. F. Virginia, M. Muller, M. Müller, D. Miller, D. L. Miller, F. Feletti, M. Wełnicki, N. Buda, A. Skoczylas, A. Pomiecko,

A. Pomiećko, D. Damjanovic, D. Damjanović, R. Olszewski, A. W. Kirkpatrick, A. W. Kirkpatrick, R. Breitkreutz, G. Mathis, G. Soldati, A. Smargiassi, R. Inchingolo, and T. Perrone, "New international guidelines and consensus on the use of lung ultrasound," *Journal of Ultrasound in Medicine*, 2022.

[25] E. Picaño, E. Picano, C. S. Maria, Q. Ciampi, D. A. Lichtenstein, and D. Lichtenstein, "Lung ultrasound for the cardiologist," *Jacc-cardiovascular Imaging*, 2018.

[26] A. Ebadi, P. Xi, A. MacLean, S. Tremblay, S. Kohli, and A. Wong, "Covidx-us - an open-access benchmark dataset of ultrasound imaging data for ai-driven covid-19 analytics," *Frontiers in Bioscience*, 2021.

[27] J. Born, N. Wiedemann, M. Cossio, C. Buhre, G. Brändle, K. Leidermann, J. Goulet, A. Aujayeb, M. Moor, B. Rieck, and K. Borgwardt, "Accelerating detection of lung pathologies with explainable ultrasound image analysis," *Applied Sciences*, vol. 11, no. 2, 2021.

[28] L. Howell, N. Ingram, R. Lapham, A. Morrell, and J. Mclaughlan, "Deep learning for real-time multi-class segmentation of artefacts in lung ultrasound," *Ultrasonics*, 2024.

[29] G. Soldati, A. Smargiassi, R. Inchingolo, D. Buonsenso, T. Perrone, F. B. Domenica, S. Perlini, E. Torri, A. Mariani, E. M. Elisa, F. Tursi, F. Mento, and L. Demi, "Proposal for international standardization of the use of lung ultrasound for covid 19 patients a simple quantitative reproducible method," *Journal of Ultrasound in Medicine*, 2020.

[30] J. Wang, X. Yang, B. Zhou, J. J. Sohn, J. Zhou, J. T. Jacob, K. A. Higgins, J. D. Bradley, and T. Liu, "Review of machine learning in lung ultrasound in covid-19 pandemic," *Journal of Imaging*, vol. 8, no. 3, 2022.

[31] L. Zhao, T. C. Fong, and M. A. L. Bell, "Detection of covid-19 features in lung ultrasound images using deep neural networks," *Communications Medicine*, vol. 4, no. 1, p. 41, 2024.

[32] Y. Yu, Y. Yu, and S. T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Transactions on Image Processing*, 2002.

[33] P. C. Tay, C. Garson, S. T. Acton, and J. A. Hossack, "Ultrasound despeckling for contrast enhancement," *IEEE Transactions on Image Processing*, 2010.

[34] P. Coupe, P. Hellier, C. Kervrann, and C. Barillot, "Bayesian non local means-based speckle filtering," 2008, p. 1291–1294.

[35] A. N. S. Cid, N. M. D. Lima, L. N. M. Diego, and D. A. M. Nelson, "Ultrasound image despeckling using stochastic distance-based bm3d," *IEEE Transactions on Image Processing*, 2017.

[36] Y. Gan, E. D. Angelini, A. F. Laine, and C. P. Hendon, "Bm3d-based ultrasound image denoising via brushlet thresholding," *IEEE International Symposium on Biomedical Imaging*, 2015.

[37] Y. Yue, M. Croitoru, M. Croitoru, A. Bidani, A. Bidani, J. Zwischenberger, J. B. Zwischenberger, and J. W. Clark, "Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images," *IEEE Transactions on Medical Imaging*, 2006.

[38] S. Sahu, M. Dubey, and I. K. Mohammad, "Comparative analysis of image enhancement techniques for ultrasound liver image," *International Journal of Electrical and Computer Engineering*, 2012.

[39] S. H. Contreras-Ortiz, S. H. C. Ortiz, T. Chiu, M. Fox, and M. D. Fox, "Ultrasound image enhancement: A review," *Biomedical Signal Processing and Control*, 2012.

[40] P. F. Stetson, P. Stetson, F. G. Sommer, F. G. Sommer, and A. Macovski, "Lesion contrast enhancement in medical ultrasound imaging," *IEEE Transactions on Medical Imaging*, 1997.

[41] R. Zheng, Q. Guo, C. Gao, M. Yu, and M. an Yu, "A hybrid contrast limited adaptive histogram equalization (clahe) for parathyroid ultrasonic image enhancement," *Cybersecurity and Cyberforensics Conference*, 2019.

[42] P. Singh, R. Mukundan, and de Ryke Rex, "Feature enhancement in medical ultrasound videos using contrast-limited adaptive histogram equalization," *Journal of Digital Imaging*, 2020.

[43] K. Wang, Y. Zhang, Z. Li, X. Lang, L. Xun, B. He, and B. He, "Common carotid artery ultrasound image enhancement based on fuzzy set theory and phase asymmetry metric," 2021.

[44] F. G. Rika, R. Hidayat, and A. N. Hanung, "Performance analysis of enhancement methods on fetal ultrasound images," 2023.

[45] V. L. Jaya and R. Gopikakumari, "Iem: A new image enhancement metric for contrast and sharpness measurements," *International Journal of Computer Applications*, 2013.

[46] Z. Wang, A. C. Bovik, H. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, 2004.

[47] N. Anantrasirichai, W. Hayes, M. Allinovi, D. Bull, and A. Achim, "Line detection as an inverse problem: Application to lung ultrasound imaging," *IEEE Transactions on Medical Imaging*, 2017.

[48] W. Xing, G. Li, C. He, Q. Huang, X. Cui, Q. Li, W. Li, J. Chen, and D. Ta, "Automatic detection of a-line in lung ultrasound images using deep learning and image processing," *Medical physics*, 2022.

[49] R. Moshavegh, L. H. Kristoffer, H. Møller-Sørensen, H. Møller-Sørensen, B. N. Michael, and A. J. Jørgen, "Automatic detection of b-lines in *invivo* lung ultrasound," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 2019.

[50] O. Karakuş, O. Karakus, N. Anantrasirichai, A. Aguersif, S. Silva, A. Basarab, and A. Achim, "Detection of line artifacts in lung ultrasound images of covid-19 patients via nonconvex regularization," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 2020.

[51] E. T. A. Jenifer and J. S. Duela, "Advancing lung ultrasound imaging: Advanced preprocessing and neoteric segmentation for improved accuracy," *Physica Scripta*, 2024.

[52] Y. Wang, Y. Zhang, Q. He, Q. He, H. Liao, and J. Luo, "Quantitative analysis of pleural line and b-lines in lung ultrasound images for severity assessment of covid-19 pneumonia," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 2021.

[53] L. Carrer, E. Donini, D. Marinelli, M. Zanetti, F. Mento, E. Torri, E. Torri, A. Smargiassi, R. Inchingolo, G. Soldati, L. Demi, F. Bovolo, and L. Bruzzone, "Automatic pleural line extraction and covid-19 scoring from lung ultrasound data," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 2020.

[54] R. Bassiouny, A. Mohamed, K. Umapathy, and N. Khan, "An interpretable neonatal lung ultrasound feature extraction and lung sliding detection system using object detectors," *IEEE Journal of Translational Engineering in Health and Medicine*, 2023.

[55] S. R. J. G. van and L. Demi, "Localizing b-lines in lung ultrasonography by weakly supervised deep learning, in-vivo results," *IEEE Journal of Biomedical and Health Informatics*, 2019.

[56] D. Vukovic, A. Wang, M. Antico, M. Steffens, I. Ruvinov, R. J. Sloun, D. Canty, A. Royse, C. Royse, K. Haji, J. Dowling, G. Chetty, and D. Fontanarosa, "Automatic deep learning-based pleural effusion segmentation in lung ultrasound images," *BMC Medical Informatics and Decision Making*, 2023.

[57] S. Roy, W. Menapace, P. O. Sebastiaan, S. Oei, B. Luijten, E. Fini, C. Saltori, A. M. H. Iris, I. A. Huijben, N. Chennakeshava, F. Mento, A. Sentelli, E. Peschiera, R. Trevisan, G. Maschietto, E. Torri, E. Torri, R. Inchingolo, A. Smargiassi, G. Soldati, P. Rota, A. Passerini, S. R. J. G. van, E. Ricci, and L. Demi, "Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound," *IEEE Transactions on Medical Imaging*, 2020.

[58] J. Born, N. Wiedemann, M. Cossio, C. Buhre, G. Brändle, K. Leidermann, A. Aujayeb, M. Moor, B. Rieck, and K. M. Borgwardt, "Accelerating detection of lung pathologies with explainable ultrasound image analysis," *Applied Sciences*, 2021.

[59] G. D. Ankan, F. Sadik, and A. F. Shaikh, "An integrated autoencoder-based hybrid cnn-lstm model for covid-19 severity prediction from lung ultrasound," *Computers in biology and medicine*, 2021.

[60] A. Bruno, G. Ignesti, O. Salvetti, D. Moroni, and M. Martinelli, "Efficient lung ultrasound classification," *Bioengineering*, 2023.

[61] M. H. Md, M. H. Muhammad, M. R. Mohammad, A. K. M. Azad, S. A. Alyami, and A. M. Mohammad, "Fp-cnn: Fuzzy pooling-based convolutional neural network for lung ultrasound image classification with explainable ai," *Comput.Biol.Medicine*, 2023.

[62] M. Correa, M. Zimic, F. Barrientos, R. Barrientos, A. Román-Gonzalez, A. Roman-Gonzalez, M. Pajuelo, M. Pajuelo, C. Anticona, H. Mayta, A. M. Alicia, A. Alva, L. Solis-Vasquez, D. Figueroa, M. A. Chavez, R. Lavarello, R. J. Lavarello, B. Castañeda, B. Castaneda, V. A. Paz-Soldán, V. A. Paz-Soldan, W. Checkley, R. H. Gilman, and R. A. Oberhelman, "Automatic classification of pediatric pneumonia based on lung ultrasound pattern recognition," *PLOS ONE*, 2018.

[63] J. Wang, X. Yang, B. Zhou, J. S. James, J. J. Sohn, J. Zhou, T. J. Jesse, A. H. Kristin, K. Higgins, D. B. Jeffrey, J. D. Bradley, and T. Liu, "Review of machine learning in lung ultrasound in covid-19 pandemic," *Journal of Imaging*, 2022.

[64] W. Xue, C. Cao, B. Liu, J. Liu, Y. Duan, Y. Duan, H. Cao, H. Cao, J. Wang, X. Tao, Z. Chen, M. Wu, J. Zhang, H. Sun, J. Yang, Y. Jin, X. Yang, R. Huang, F. Xiang, Y. Song, M. You, W. Zhang, L. Jiang, Z. Zhang, S. Kong, Y. Tian, L. Zhang, D. Ni, and M. Xie, "Modality alignment contrastive learning for severity assessment of covid-19 from lung ultrasound and clinical information," *Medical image analysis*, 2021.

[65] M. J. Horry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, M. Saha, and N. Shukla, "Covid-19 detection through transfer learning using multimodal imaging data," *IEEE Access*, 2020.

[66] G. Muhammad and M. S. Hossain, "Covid-19 and non-covid-19 classification using multi-layers fusion from lung ultrasound images," *Information Fusion*, 2021.

[67] N. Awasthi, A. Dayal, L. R. Cenkeramaddi, and P. K. Yalavarthy, "Mini-covidnet: Efficient lightweight deep neural network for ultrasound based point-of-care detection of covid-19," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 68, no. 6, p. 2023–2037, 2021.

[68] D. Chu, A. Liteplo, N. M. Duggan, A. B. Hutchinson, and H. Shokoohi, "Artificial intelligence in lung ultrasound," *Current Pulmonology Reports*, 2024.

[69] C. Baloescu, G. Toporek, S.-S. Kim, S. Kim, K. Seungsoo, K. McNamara, R. Liu, M. Shaw, R. L. McNamara, B. I. Raju, I. R. Balasundar, and C. L. Moore, "Automated lung ultrasound b-line assessment using a deep learning algorithm," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 2020.

[70] J. Diaz-Escobar, N. E. Ordóñez-Guillén, S. Villarreal-Reyes, A. Galaviz-Mosqueda, V. Kober, R. Rivera-Rodriguez, and J. E. L. Rizk, "Deep-learning based detection of covid-19 using lung ultrasound imagery," *PloS one*, vol. 16, no. 8, p. e0255886, Aug 13 2021.

[71] R. Arntfield, D. Wu, D. Wu, J. Tschirhart, B. VanBerlo, A. Ford, J. Ho, J. McCauley, B. Wu, J. Deglint, R. Chaudhary, C. Dave, B. VanBerlo, J. Basmaji, and S. J. Millington, "Automation of lung ultrasound interpretation via deep learning for the classification of normal versus abnormal lung parenchyma: A multicenter study," *Diagnostics*, 2021.

[72] M. L. Salvia, G. Secco, E. Torti, G. Florimbi, L. Guido, P. Lago, F. Salinaro, S. Perlini, and F. Leporati, "Deep learning and lung ultrasound for covid-19 pneumonia detection and severity classification," *Computers in biology and medicine*, 2021.

[73] D. P. Kumar, D. Kumar, M. A. Mehta, and I. Chatterjee, "Empirical analysis of deep convolutional generative adversarial network for ultrasound image synthesis," *The Open Biomedical Engineering Journal*, 2021.

[74] H. Atri, M. Shadi, and M. Sargolzaei, "Generating synthetic medical images with limited data using auxiliary classifier generative adversarial network: a study on thyroid ultrasound images," *Journal of Ultrasound*, 2023.

[75] J. Liang, X. Yang, Y. Huang, H. Li, S. He, X. Hu, Z. Chen, W. Xue, J. Cheng, and D. Ni, "Sketch guided and progressive growing gan for realistic and editable ultrasound image synthesis," *Medical image analysis*, vol. 79, p. 102461, Jul 2022.

[76] H. M. Rai, S. Dashkevych, and J. Yoo, "Next-generation diagnostics: The impact of synthetic data generation on the detection of breast cancer from ultrasound imaging," *Mathematics*, vol. 12, no. 18, 2024.

[77] T. Fujioka, M. Mori, K. Kubota, Y. Kikuchi, L. Katsuta, M. Adachi, G. Oda, T. Nakagawa, Y. Kitazume, and U. Tateishi, "Breast ultrasound image synthesis using deep convolutional generative adversarial networks," *Diagnostics (Basel, Switzerland)*, vol. 9, no. 4, p. 176. doi: 10.3390/diagnostics9040176, Nov 6 2019.

[78] L. Zhao, C. F. Tiffany, and M. Bell, "Detection of covid-19 features in lung ultrasound images using deep neural networks," *Communications Medicine*, 2024.

[79] S. Yasutomi, T. Arakaki, R. Matsuoka, A. Sakai, R. Komatsu, K. Shozu, A. Dozen, H. Machino, K. Asada, S. Kaneko, A. Sekizawa, R. Hamamoto, and M. Komatsu, "Shadow estimation for ultrasound images using auto-encoding structures and synthetic shadows," *Applied Sciences*, 2021.

[80] Y. Xin, X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *Medical image analysis*, 2019.

[81] G. Ramos-Llordén, G. Ramos-Llorden, G. Vegas-Sánchez-Ferrero, M. Martın-Fernández, M. Martín-Fernández, C. López, C. Alberola-López, S. Aja-Fernandez, and

S. Aja-Fernández, "Anisotropic diffusion filter with memory based on speckle statistics for ultrasound images," *IEEE Transactions on Image Processing*, 2015.

[82] B. Peng, Y. Wang, Y. Wang, X. Yang, and X. Yang, "A multiscale morphological approach to local contrast enhancement for ultrasound images," *International Conference Communication and Information Systems*, 2010.

[83] P. Coupé, P. Hellier, C. Kervrann, and C. Barillot, "Nonlocal means-based speckle filtering for ultrasound images," *IEEE Transactions on Image Processing*, 2009.

[84] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Transactions on Communications*, vol. 43, no. 12, p. 2959–2965, 1995, iD: 1.

[85] N. Moroney, "Local color correction using non-linear masking," in *Color and Imaging conference*, vol. 8. Society of Imaging Science and Technology, 2000, p. 108–111.

[86] P. Pradham, N. H. Younan, and R. L. King, "Concepts of image fusion in remote sensing applications," *Image Fusion*, pp. 393–428, 2008.

[87] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," p. 4695–4708, 2012.

[88] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale modeling simulation*, vol. 4, no. 2, p. 490–530, 2005.

[89] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, p. 613–627, 1995.

[90] G. Ramos-Llordén, G. Vegas-Sánchez-Ferrero, M. Martin-Fernandez, C. Alberola-López, and S. Aja-Fernández, "Anisotropic diffusion filter with memory based on speckle statistics for ultrasound images," *IEEE Transactions on Image Processing*, vol. 24, no. 1, p. 345–358, 2015.

[91] Y. Mäkinen, L. Azzari, and A. Foi, "Collaborative filtering of correlated noise: Exact transform-domain variance for improved shrinkage and patch matching," p. 8339–8354, 2020.

[92] P. Singh, R. Mukundan, and R. D. Ryke, "Feature enhancement in medical ultrasound videos using contrast-limited adaptive histogram equalization," *Journal of Digital Imaging*, vol. 33, no. 1, p. 273–285, 2020.

[93] J. Chen, J. Li, C. He, W. Li, and Q. Li, "Automated pleural line detection based on radon transform using ultrasound," *Ultrason Imaging*, vol. 43, no. 1, p. 19–28, 2021.

[94] A. Miller, "Practical approach to lung ultrasound," *BJA Education*, vol. 16, no. 2, p. 39–45, 2016.

[95] L. Carrer and L. Bruzzone, "Automatic enhancement and detection of layering in radar sounder data based on a local scale hidden markov model and the viterbi algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, p. 962–977, 2017.

[96] S. Suzuki and K. be, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, p. 32–46, 1985.

[97] B. S. olkopf, K. Tsuda, and J. P. Vert, *Kernel Methods in Computational Biology*. MIT Press, 2004.

[98] P. Dell'Aquila, P. Raimondo, V. Racanelli, P. D. Luca, S. D. Matteis, A. Pistone, R. Melodia, L. Crudele, D. Lomazzo, A. G. Solimando, A. Moschetta, A. Vacca, S. Grasso, V. Procacci, D. Orso, and L. Vetrugno, "Integrated lung ultrasound score for early clinical decision-making in patients with covid-19: results and implications," *The ultrasound journal*, vol. 14, no. 1, p. 21–8, Jun 1 2022.

[99] M. Matsumoto and T. Nishimura, "Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Trans.Model.Comput.Simul.*, vol. 8, no. 1, p. 3–30, jan 1998.

[100] D. J. Walton and D. S. Meek, "Approximation of quadratic bézier curves by arc splines," *Journal of Computational and Applied Mathematics*, vol. 54, no. 1, p. 107–120, 1994.

[101] A. Lagae, S. Lefebvre, G. Drettakis, and D. Philip, "Procedural noise using sparse gabor convolution," in *ACM SIGGRAPH 2009 Papers*, ser. SIGGRAPH '09. New York, NY, USA: Association for Computing Machinery, 2009.

[102] P. Singh, R. Mukundan, and R. de Ryke, "Synthetic models of ultrasound image formation for speckle noise simulation and analysis," *2017 International Conference on Signals and Systems (ICSigSys)*, pp. 278–284, 2017.

[103] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, p. 139–144, 2020.

[104] T. Park, M.-Y. Liu, T.-C. Wang, and J.-Y. Zhu, *Gaugan: semantic image synthesis with spatially adaptive normalization*, ser. ACM SIGGRAPH 2019 Real-Time Live!, 2019, p. 1.

[105] S. C. Kothari and H. Oh, *Neural Networks for Pattern Recognition*, ser. Advances in Computers. Elsevier, 1993, vol. 37, p. 119–166.

[106] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 20*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 15 2015, p. 234–241.

[107] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015.

[108] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *CoRR*, vol. abs/1409.4842, 2014.

[109] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.

[110] N. Awasthi, A. Dayal, R. C. Linga, and P. K. Yalavarthy, "Mini-covidnet: Efficient lightweight deep neural network for ultrasound based point-of-care detection of covid-19," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 2021.