Available online at: www.metrology-journal.org

RESEARCH ARTICLE

OPEN 3 ACCESS

Accurate COVID-19 detection using full blood count data and machine learning

Richard Yang¹, Ding Chen², Qingping Yang^{1,*}, Yang Qiu², and Fang Wang¹

¹ College of Engineering, Design and Physical Sciences, Brunel University London, Uxbridge, UK

² Wuhan Union Hospital Affiliated with Tongji Medical College, Huazhong University of Science and Technology, Wuhan, PR China

Received: 17 August 2023 / Accepted: 25 June 2024

Abstract. COVID-19 has spread rapidly worldwide in the past three years, triggering partial and full lockdowns globally. The successful control of the COVID-19 pandemic on a global scale depended heavily upon the accurate detection of COVID-19. However, the main diagnostic tests for COVID-19 have some significant limitations, e.g. the major nucleic acid (RT-PCR) tests while having a high sensitivity are time-consuming and require expensive equipment with the shortage of test kits in many countries. Antigen lateral flow tests have a lower sensitivity and they cannot be used during the early pandemic as well as usually more expensive than the full or complete blood count test used in this paper which can be potentially performed using a finger blood sample. The last decade has seen rapid growth of AI, particularly deep learning, which has found wide applications in medical image analysis, with results comparable to and even surpassing human expert performance. There have been several machine learning models reported for COVID-19 diagnostics or prognosis predictions, most of them based on CT and X-ray images. In this paper we have applied traditional machine learning and convolutional neural networks (CNNs) based deep learning to the blood test data obtained from hematology analyzers and demonstrated that the AI models can be used to detect COVID-19 with a high degree of accuracy (>97%). The performance of different classifiers will be compared and discussed. The work should have potential applications in current COVID-19 and future pandemics.

Keywords: COVID-19 / full blood count / machine learning / deep learning / convolutional neural networks

1 Introduction

COVID-19 was initially reported in Wuhan, China, then quickly spread worldwide and was declared a global pandemic by the World Health Organization (WHO) on the 11th of March 2020. It is a type of SARS-CoV2 virus that produces various symptoms (e.g. acute respiratory failure, acute respiratory distress syndrome (ARDS), and COVID-19 pneumonia) in humans that can cause death. This has led many countries to enforce strict procedures such as lockdowns and the closure of borders, schools, and other sectors.

The successful control of the pandemic is heavily dependent on the accurate detection of COVID-19. However, the accurate detection of the virus is a challenging task, with current methods of testing having significant limitations, e.g. major nucleic acid (RT-PCR) tests while having a high sensitivity (>90%) [1] are time-consuming and require expensive equipment with the shortage of test kits in many countries as well. Antigen lateral flow tests have a lower sensitivity [2] and they cannot be used during the early pandemic as well as usually more expensive than the full blood count test used in this paper which can be potentially performed using a finger blood sample.

Machine learning can provide very powerful methods to detect COVID-19 [3]. In general, there are many different machine learning methods that can be used for the detection of COVID-19 such as Logistic Regression, Naive Bayes, k-nearest neighbours (KNN), Support Vector Machine (SVM), Random Forest, shallow Artificial Neural Networks (ANN), etc.

There have been many published papers employing the use of machine learning to detect COVID-19. Many of these papers use image processing based on CT and X-ray images [4,5] whereas other papers use traditional machine learning classifiers on numerical blood report data to detect COVID-19. Jiang et al. [6] evaluated six machine learning methods used to diagnose COVID-19 (SVM, Random Forest, KNN, Logistic Regression, and two different

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

^{*}Corresponding author: qingping.yang@brunel.ac.uk

decision trees). These were applied over a real dataset obtained from Wenzhou Central Hospital and Cangnan People's Hospital in Wenzhou (China). The performance of SVM was strong compared to other methods with an accuracy of 80%. Batista et al. [7] examined the performance of five machine learning methods (SVM, Logistic Regression, Random Forest, ANN, and gradient boosted trees (GBT)) over a real dataset obtained from Hospital Israelita Albert Einstein at Sao Paulo, Brazil. The SVM and Random Forest methods outperformed other methods with an AUC (area under curve) value of 0.847. Alakus and Turkoglu [8] reviewed the performance of five deep learning methods (i.e. CNN, Long-Short Term Memory (LSTM), Recurrent Neural Networks (RNN), CNNLSTM, and CNNRNN) and one shallow neural network (i.e. ANN) to detect COVID-19 based on laboratory findings. The CNNLSTM with an accuracy of 92.30% outperforms other methods. Despite generating impressive results, these models [8] require nine blood test components in addition to the full blood count data.

The COVID-19 detection method we developed only requires the full or complete blood count (CBC) data. Based on the CBC numerical and flow cytometry image data we collected from the first wave of COVID-19 in Wuhan in early 2020, we have been carrying out their systematic studies and developing a range of machine learning models, including both traditional machine learning and deep learning models. We have demonstrated that the use of flow cytometry images and deep learning models can achieve an accuracy comparable to a PCR test, thus offering a novel, low-cost, fast and accurate COVID-19 detection solution. This paper presents some of our key results of the application of machine learning to CBC numerical and image data.

2 Theoretical background

2.1 Machine learning classifiers

There are a number of machine learning methods that can be used as classification algorithms. In this paper, we have selected nine different classification methods, namely Convolutional Neural Network (CNN), shallow ANN, Decision trees, Random Forest, SVM, Discriminant Analysis, Logistic Regression, Naïve Bayes and KNN. These classifiers have been used successfully in various domains.

2.2 Deep learning

Deep learning is an extension of machine learning based on algorithms that attempt to model high-level abstractions in data using multiple processing layers that are comprised of complex structures or multiple non-linear transformations. It has an important emphasis on replacing handcraft features with efficient algorithms, which allow for unsupervised or partly supervised feature learning and feature extraction [9]. Various deep learning architectures such as convolutional neural networks, deep belief networks and recurrent neural networks have recently produced state-ofthe-art results in image recognition.

2.3 Convolutional neural networks

The human visual system is very efficient at recognizing objects or images even in cluttered scenes. For artificial systems, this is still very difficult due to view-dependent object variability as well as the high variability of the objects themselves. Deep neural networks such as CNNs roughly mimic the visual cortex of mammals. This makes CNNs one of the most promising architectures to be used to develop image recognition systems [10].

CNNs are constructed from several layers with each layer fulfilling a specific function. The majority of CNNs used for image recognition are based on the following basic components:

Convolutional layer: The core component of any CNN. The parameters are a set of learnable filters or kernels. The filters have a receptive field that extends the input. In the forward pass, every filter is convolved across the height and width of the input and the dot product between the filter entries is computed producing a 2-D activation map corresponding to it. Hence the CNN learns the filters that activate when they see a specific type of feature in the input.

Pooling layer: Used for non-linear down-sampling. The most used function for pooling is max pooling. It splits the image into non-overlapping partitions and, for each of these regions, outputs the maximum. The purpose of this is to progressively reduce the size of the representation, therefore reducing the number of parameters as well as computation in the network.

Fully connected layer/classification layer: After the convolution layers and the max-pooling layers, the high-level classification is done using the fully connected layer/ classification layer in a neural network.

A CNN consists of multiple convolutional layers, the set of kernels within the layer scan the pixels of the input image outputting data as a set of matrices called a feature map. The convolution layer at the front of the network captures the local and detailed information. There is a small receptive field where each pixel of the output image only uses a small range of the input image. The receptive field of each subsequent convolution layer is increased in every layer to capture more complex and abstract information. After the computations of multiple convolution layers, the abstract representations of the image at various scales are obtained.

3 Dataset

The clinical and laboratory data were collected from medical records of laboratory-confirmed COVID-19 cases from Wuhan Union Hospital and Wuhan Mobile Cabin Hospital between 25th January 2020 and 11th March 2020, under the approval by the ethics committee of the Wuhan Union Hospital of Tongji Medical College.

The CBC tests were performed by Mindray series haematology analyzers, and the measurement thresholds were given by the hospital laboratory. The CBC test results were routinely generated as a blood panel of numerical readings and also the associated flow cytometry image used to derive the numerical readings. The flow cytometry used



Fig. 1. Sample image from SFL-SSC dataset.

for the CBC tests hydrodynamically focuses each cell to pass through the laser and collects the scattered light and fluorescence emission. There are three common channels in the CBC tests. The forward scatter channel (FSC) reflects the cell size and can be used to pick out cellular debris. The side scatter channel collects the granular content information within cells. The side fluorescence (SFL) with DNA staining can measure the amount of nucleic acid. The unique channel information of each cell is used to differentiate cell types in the CBC test. There were variations in the types of actually reported images and CBC numerical components in different hospitals or in their different departments. In this work, the image data used is the SFL-SSC type (as shown in Figs. 1 and 2) and the CBC numerical readings have 22 components (with the names of these components listed in Tab. 1). Both types of data need to be accurately labelled according to the confirmed diagnosis by the doctors. The CBC SFL-SSC images were manually cropped from the test reports and the CBC numerical reading collection was semi-automated. Both types of data were manually checked and verified to identify and correct possible error entries or incorrect image types.

For the CBC numerical data, there were 662 records for COVID-19 patients (outpatients and inpatients) and 659 records for COVID-19 negative patients. For the CBC image data, there were 799 CBC images for Covid-19 patients, and 945 CBC images for the control group, with a total of 1744 patients. The CBC numerical and image data have allowed us to develop a range of machine learning models. To our knowledge, we are the first to apply deep learning to the CBC flow cytometry images for COVID-19 detection or diagnostics since we collected these data from the very first wave of COVID-19 in Wuhan in early 2020.



Fig. 2. Labelled regions of dataset image.

4 Experiments and results

4.1 Experiment design and methods

Initially we performed some exploratory data analysis on the CBC numerical data, including correlations, T-test, Analysis of variance (ANOVA) and principal component analysis (PCA). Pearson's correlation coefficients were calculated between each pair of the 22 features. For COVID-19 detection, two-sided *t*-test was carried out to test the significance of the differences in the features of the study and control groups, assuming unequal variances due to the significant variations between the variances of the two groups. These statistical analyses could assist the understanding and explanation of the machine learning models.

We have also applied traditional machine learning (Discriminant Analysis, KNN, Decision trees, Naïve Bayes, SVM, Logistic Regression and Random Forest). For each type of model below, a full model and a partial model are trained. The partial model is trained by sequential feature selection, i.e. sequentially selecting the important feature for training. The full model is trained using 10-Fold training and the partial model using 7-Fold training. The performance metrics are then calculated for ROC (receiver operating characteristics) and AUC using the full model. We chose these metrics as ROC and AUC are generally seen as a more important measure of how well an algorithm performs as it considers trade-offs between precision and recall.

A shallow neural network of ANN was also trained (Fig. 3). This model was trained 10 times with the 10 models being used to calculate the ROC and AUC metrics.

For the CBC image data, we also performed some exploratory image analysis to understand the key image features, including edge detection, image PCA and SIFT (Scale-Invariant Feature Transform).

Table 1. *T*-test results (left half: significance of all 22 CBC components; right half: significant CBC components in the order of *p*-value; * $10^{-10} ; ** <math>10^{-20} , *** <math>10^{-30} , **** <math>p \le 10^{-30}$).

All components	p-value(Covid-	-19 vs Control)	Significant components	<i>p</i> -value (Covid-19 vs Control)		
WBC	1.43E-12	**	LYMPHn	2.70E-66		
RBC	5.23E-04	*	EOn	1.13E-42		
HGB	7.22E-15	**	EOp	1.10E-32	****	
HCT	1.49E-11	**	BASOp	3.60E-22	***	
MCV	2.24E-10	*	BASOp	2.43E-21	***	
MCH	4.80E-15	**	LYNPHp	4.34E-17	**	
MCHC	5.48E-09	*	MCH	4.80E-15	**	
PLT	7.62E-01		NEUTp	5.49E-15	**	
NEUTp	5.49E-15	**	HGB	7.22E-15	**	
LYNMPHp	4.34E-17	**	WBC	1.43E-12	**	
MONOp	9.59E-08	*	HCT	1.49E-11	**	
EOp	1.10E-32	****	MCV	2.24E-10	*	
BASOp	2.43E-21	***	MCHC	5.48E-09	*	
NEUTn	2.76E-03	*	MONOp	9.59E-08	*	
LYMPHn	2.70E-66	****	RBC	5.23E-04	*	
MONOn	5.83E-03	*	RDW-CV	8.26E-04	*	
Eon	1.13E-42	****	PWD	1.64E-03	*	
BASOn	3.60E-22	***	NEUTn	2.76E-03	*	
RDW-CV	8.26E-04	*	MONOn	5.83E-03	*	
PDW	1.64E-03	*				
MPV	1.85E-01					
PCT	7.78E-01					



Fig. 3. Model Design of ANN and CNN.



Fig. 4. Correlation matrix of 22 CBC features.

Subsequently, a convolutional neural network was designed and trained to classify the CBC SFL-SSC images. In order to determine the optimal network architecture, more than 20 networks were designed with different hyperparameters, including different learning algorithms, number of convolution layers, number of filters at each layer, filter size, type of pooling layers, pooling strides and number of hidden neurons in the fully connected layer. Each configuration was trained five times and the average performance was then compared and the best model architecture was determined as shown in Figure 3, which has two convolutional layers (each has $16\ 2 \times 2\ 10\ 2$, trained by Adam algorithm. This model was also trained 10 times to calculate the ROC and AUC metrics.

The average performances of all the classifiers (including traditional machine learning, shadow neural network and CNN) were compared using the same data set under similar training conditions (i.e. sample balancing and splitting, and validation), with the performance metrics including sensitivity, specificity, accuracy, AUC, and ROC plots.

4.2 Experiment results and discussions

Some results of exploratory data analyses including correlation, T-test, PCA and ANOVA are shown below in Figures 4–6, and Table 1.

The correlation matrix (Fig. 4) shows how the features of the CBC components correlate with each other and whether the dimensionality could be potentially reduced.

In order to see which of the 22 CBC components are significantly affected by COVID-19, a *t*-test is performed to compare the means of two samples corresponding to the patient group and the control for each CBC component, assuming unequal variances for each pair of the samples. The *t*-test results in Table 1 have shown that for COVID-19 detection there is a significant difference in 19 of the 22 features between the means of the disease and control groups, with the most significant components for COVID-19 detection being lymphocyte counts, eosinophil counts and eosinophil percentage, followed by basophil counts and basophil percentage. These indicate that lymphocyte, eosinophil and basophil cells are closely linked to the infection of COVID-19.

Figure 5 shows part of the PCA results. Typically, the first 3–5 principal components could explain about 80% of the variance and further modelling could just be based on them. But for the CBC readings, the differences seem to be gradual among the components as seen in the PCA results. This means the models will need to use most of the 22 CBC components.

The ANOVA plots shown in Figure 6 indicate the significant differences in the CBC components between COVID-19 positive, negative, and recovered patients.

After the exploratory data analysis, traditional machine learning classifiers were applied to the data, this includes Discriminant Analysis, KNN, Decision trees, Naïve Bayes, SVM, Logistic Regression and Random Forest. These models produced a similar performance with accuracies of around 80% on the test set, with KNN giving a loss value of 0.1785, Discriminant Analysis 0.2173, Naïve Bayes 0.2163, decision trees 0.2353, SVM 0.2163 and Random Forest 0.1649. The test prediction confusion matrix (using 22% of the dataset) for the Random Forest model, the best performing traditional classifier, is shown in Figure 7.

These results showed promising performances for all traditional classifiers. However, the accuracies are too low to be recommended for practical use. For KNNs an assumption is made that nearby neighbours are similar.



Fig. 5. PCA results with percentage of the total variance explained by top 10 principal components.



Fig. 6. ANOVA of CBC numerical components with the most significant differences (Covid0 = negative; Covid1 = COVID-19 positive; CovidR = negative; Covid1 = COVID-19 positive; CovidR = negative; Covid1 = nega

But for a dataset with 22 features the 'Curse of dimensionality' [11] should occur, i.e. when the number of dimensions increases, the distribution of the neighbouring points also increases, which breaks the above fundamental assumption, with KNNs losing all predictive power. While Naive Bayes works well with high dimensionality data, Naïve Bayes also has the assumption that all the features are mutually independent, which, according to the correlation analysis, does not really hold since the correlation analysis has shown clear correlations in some of the features in this dataset.

The shallow ANN was then developed, and the results show performance similar to or slightly better than the traditional machine learning classifiers. The typical test prediction confusion matrix (using 15% of the CBC readings) for the ANN is shown in Figure 8. The best performance metrics of the ANN numerical CBC model are shown in Table 2.

The above experiments were all conducted using the numerical CBC data from Wuhan Union Hospital. Based on the literature review and the study of the CBC data it is believed that to achieve better performance the use of deep learning techniques and more specifically convolutional neural networks would be needed. Using the CBC SFL-SSC images a convolutional neural network was developed. The model produced much better results achieving very high sensitivity and specificity, 98.3% and 97.5%, respectively as well as high accuracies 97.8%, based on the test data set (15% of the CBC images), as shown in Figure 9. Table 3 shows the results of the CNN model with the training repeated 10 times.

The metrics including the ROC plots and AUC for all the machine learning models are shown in Table 4 and Figure 10. It can be seen that whilst all the classifiers have performed quite well, the CNN, Random Forest and ANN are the best three classifiers.

The ANN COVID-19 detection using CBC numerical readings have sensitivity and specificity of about 86% for all the data, and 85.7% and 83.3% (Tab. 2), respectively, for the test data set with an average AUC of 0.881.

In particular, the COVID-19 detection using CNN with the CBC images has achieved very high sensitivity and specificity, 99.1% and 98.4% (Fig. 9), respectively, for all the data directly used in the development of the model and an AUC of 0.981 for the test data. These performances are also consistent with other performance metrics, namely ROC and AUC.



Table 2. The best performance metric of ANN numerical CBC model.										
Validation				Test		All				
Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy		
87.4%	86.3%	86.9%	85.7%	83.3%	84.3%	86.3%	86.1%	86.2%		



Fig. 9. CNN confusion matrix.

We have tested the deep learning model specially with the 13 influenza A patients, all with correct predictions.

5 Conclusions

In this work, we examined the performance of nine machine learning classifiers and their ability to detect COVID-19. These classifiers were performed on a real dataset of CBC blood test data on both numerical data and SFL-SSC images from Wuhan Union Hospital in China. The results show that all the classifiers produced strong results with all of them producing accuracies over 70%. The best performance was produced by the deep learning CNN model with the best test accuracy of over 97% which is corroborated by the AUC values and ROC plots. This has demonstrated that our CNN model has achieved an accuracy comparable to a PCR test, thus offering a novel, low-cost, fast and accurate COVID-19 detection solution.

Table 3.	CNN	results	(10)	simu	lations)).
----------	-----	---------	------	------	----------	----

No.	Validation			Training			Test			All		
	Sensitivity	Specificity	Accuracy									
1	96.1%	96.4%	96.3%	97.0%	99.0%	98.3%	92.2%	95.3%	94.1%	96.2%	98.1%	97.3%
2	98.0%	97.6%	97.8%	98.3%	99.5%	99.1%	94.1%	94.1%	94.1%	97.6%	98.4%	98.1%
3	90.2%	96.4%	94.1%	96.6%	98.0%	97.5%	92.2%	98.8%	96.3%	95.0%	97.9%	96.8%
4	90.2%	92.9%	91.9%	97.5%	98.2%	97.9%	94.1%	92.9%	93.4%	95.9%	96.6%	96.3%
5	100.0%	92.9%	95.6%	99.6%	99.7%	99.7%	98.0%	97.6%	97.8%	99.4%	98.4%	98.8%
6	92.2%	98.8%	96.3%	98.7%	99.2%	99.1%	88.2%	97.6%	94.1%	96.2%	98.9%	97.9%
7	96.1%	94.0%	94.8%	97.9%	98.7%	98.4%	98.0%	96.5%	97.1%	97.6%	97.7%	97.7%
8	98.0%	97.6%	97.8%	99.2%	99.7%	99.5%	92.2%	97.6%	95.6%	97.9%	99.1%	98.7%
9	96.1%	91.7%	93.3%	96.6%	98.0%	97.5%	96.1%	96.5%	96.3%	96.5%	96.8%	96.7%
10	94.1%	95.2%	94.8%	97.9%	99.2%	98.7%	94.1%	98.8%	97.1%	96.8%	98.6%	97.9%
Average	95.1%	95.4%	95.3%	97.9%	98.9%	98.6%	93.9%	96.6%	95.6%	96.9%	98.1%	97.6%

Table 4. Performances of all classifiers.

No.	Classifiers	Data type	Training			Test			Average
			Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	test AUc
1	KNN	CBC readings	98.1%	98.3%	98.2%	80.7%	82.6%	81.7%	0.901
2	Discriminant Analysis	CBC readings	76.0%	82.0%	79.0%	70.1%	85.1%	77.6%	0.854
3	SVM	CBC readings	77.4%	81.6%	79.5%	73.4%	85.7%	79.6%	0.857
4	Naive Bayes	CBC readings	79.9%	81.4%	80.6%	82.2%	80.1%	81.1%	0.860
5	Decision Tree	CBC readings	82.5%	81.2%	81.8%	73.4%	75.4%	74.4%	0.832
6	Logistic Regression	CBC readings	76.6%	81.9%	79.2%	77.3%	83.3%	80.3%	0.866
7	Random Forests	CBC readings	98.2%	98.2%	98.2%	82.7%	82.1%	82.4%	0.899
8	ANN	CBC readings	81.3%	84.4%	82.8%	78.7%	83.0%	80.9%	0.881
9	CNN1	SFL-SSC images	97.9%	98.9%	98.6%	93.9%	96.6%	95.6%	0.981



Fig. 10. ROC plot of all classifiers.

In future work, we will analyse whether the models have a similar performance on the data obtained from later stages in the pandemic as well as perform experiments on the explainability of the models to support their practical deployment.

Funding

This work was partly funded by Brunel University London.

Conflict of interest

The authors declare there is no conflict of interest.

Data availability statement

The raw dataset generated or analyzed during this study is not publicly available due to them containing information that could compromise patient privacy. The models generated by this dataset are available at request.

Author contribution statement

Conceptualization, D.C. and Q.Y; Data collection, D.C. and Y.Q., Methodology, R.Y., D.C. and Q.Y.; Validation, R.Y., D.C. and Q.Y.; Resources, D.C., Q.Y. and F.W.; Writing—original draft, R.Y., D.C. and Q.Y.; Writing—review & editing, Y.Q., F.W. and R.Y.; Supervision, Q.Y. and F.W.

References

- R. Pu, S. Liu, X. Ren, D. Shi et al., The screening value of RT-LAMP and RT-PCR in the diagnosis of COVID-19: systematic review and meta-analysis, J. Virolog. Methods **300**, 114392 (2022)
- V.T. Chu, N.G. Schwartz, M.A. Donnelly et al., Comparison of home antigen testing with RT-PCR and viral culture during the course of SARS-CoV-2 infection, JAMA Intern. Med. 182, 701–709 (2022)
- L. Wynants, B. Van Calster, G.S. Collins et al., Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal, BMJ 369, 1328 (2020)
- R. Kumar, R. Arora, V. Bansal et al., Accurate prediction of COVID-19 using chest X-ray images through deep feature learning model with SMOTE and machine learning classifiers, MedRxiv 2020–04 (2020)

- S.H. Kassania, P.H. Kassanib, M.J. Wesolowskic, K.A. Schneidera, R. Detersa, Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach, Biocybern. Biomed. Eng. 41, 867–879 (2021)
- X. Jiang, M. Coffee, A. Bari et al., Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity, Comput. Mater. Continu. 63, 537–51 (2020)
- A.F. de Moraes Batista, J.L. Miraglia, T.H.R. Donato, A.D.P. Chiavegatto Filho, COVID-19 diagnosis prediction in emergency care patients: a machine learning approach, medRxiv 2020–04 (2020)
- T.B. Alakusv, I. Turkoglu, Comparison of deep learning approaches to predict COVID-19 infection, Chaos Solitons Fract. 140, 110120 (2020)
- C. Farabet, C. Couprie, L. Najman, Y. LeCun, Learning hierarchical features for scene labelling, IEEE Trans. Pattern Anal. Machine Intell. 35, 1915–1929 (2012)
- D.C. Ciresan, U. Meier, J. Masci et al., Flexible, high performance convolutional neural networks for image classification. Switzerland, in Twenty-second International Joint Conference on Artificial Intelligence (2011)
- N. Kouiroukidis, G. Evangelidis, The effects of dimensionality curse in high dimensional kNN search, in 15th *Panhellenic Conference on Informatics*, Kastoria, Greece (2011). pp. 41–45

Cite this article as: Richard Yang, Ding Chen, Qingping Yang, Yang Qiu, Fang Wang, Accurate COVID-19 detection using full blood count data and machine learning, Int. J. Metrol. Qual. Eng. 15, 17 (2024)