



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



A novel classifier architecture based on deep neural network for COVID-19 detection using laboratory findings

Volkan Göreke^a, Vekil Sarı^b, Serdar Kockanat^{b,*}

^a Sivas Vocational School of Technical Sciences, Sivas Cumhuriyet University, 58140, Sivas, Turkey

^b Department of Electrical and Electronics Engineering, Sivas Cumhuriyet University, 58140, Sivas, Turkey

ARTICLE INFO

Article history:

Received 8 January 2021

Received in revised form 15 February 2021

Accepted 7 March 2021

Available online 19 March 2021

Keywords:

COVID-19 disease

Deep neural network

Blood findings

ABC algorithm

ABSTRACT

Unfortunately, Coronavirus disease 2019 (COVID-19) is spreading rapidly all over the world. Along with causing many deaths, it has substantially affected the social life, economics, and infrastructure worldwide in a negative manner. Therefore, it is very important to be able to diagnose the COVID-19 quickly and correctly. In this study, a new feature group based on laboratory findings was obtained considering ethnical and genetic differences for interpretation of blood data. Then, using this feature group, a new hybrid classifier architecture based on deep learning was designed and COVID-19 detection was made. Classification performance indicators were obtained as accuracy of 94.95%, F1-score of 94.98%, precision of 94.98%, recall of 94.98% and AUC of 100%. Achieved results were compared with those of the deep learning classifiers suggested in literature. According to these results, proposed method shows superior performance and can provide more convenience and precision to experts for diagnosis of COVID-19 disease.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a novel zoonotic coronavirus that causes acute respiratory disease in humans. The disease known as COVID-19, has affected the whole world not only by being very contagious and causing many deaths but also affecting the social life, economics, and infrastructure [1,2].

Therefore, it is clinically very important to be able to predict the progress of the disease. Recent studies have shown that the interpretation of blood findings have a substantial clinical application value in the progress of contagious diseases [3,4]. At the same time, Guan et al. have detected anomalies in the blood findings by examining the clinical attributes of 1000 COVID-19 patients [5]. The studies have also determined that the blood values were able to have ethnical and genetic differences. Different types of blood findings may be more determinant in the detection of the disease in societies in different locations of the world [6]. Therefore, considering the genetic differences related to this disease, for which there is insufficient scientific data yet, is important in increasing the success in detection.

For the doctors, the lack of experience about a very recent disease of which properties are not completely known, has led to

difficulties in diagnosis. Due to all these limitations, an artificial intelligence (AI) system is needed to make right decisions in the diagnosis. AI has been actively used in healthcare systems to provide clinical decision support [7].

Especially in real world classification problems which medical datasets are used, data sets are called unbalanced data sets, where the number of samples between classes are different from each other. Studies to be carried out on these unbalanced training data sets negatively affect classifier performance [8]. In studies based on deep learning, eliminating this problem on data sets as a pre-process, is a more professional approach that helps to obtain more accurate and high-performance results. The most widely used class imbalance measure in the literature is calculated as the ratio of the sample numbers of the largest majority class and the smallest minority class and is called the imbalance ratio. The higher this ratio is the greater the imbalance scope of the dataset and causes over fitting problem in the classification process and decreases performance [9,10]. A widely used method to eliminate the imbalance between data classes encountered by Deep Learning classifier models is SMOTE (Synthetic Minority Over-sampling Technique) method. In this method, the number of samples of the class with a small number of samples is increased with synthetic data [11]. In addition to this, the training data increased by this method provides important objective gains for deep learning. These gains are expressed as increased stability and reduced over fitting [12].

In this study, firstly, a new feature group was obtained from blood findings. COVID-19 prediction was realized with three deep

* Corresponding author.

E-mail addresses: vgoreke@cumhuriyet.edu.tr (V. Göreke), vsari@cumhuriyet.edu.tr (V. Sarı), skockanat@cumhuriyet.edu.tr (S. Kockanat).

learning classifiers with different architectures using this feature group. The results were compared with the results in literature and classic laboratory findings. Contributions of the proposed model can be listed as follows:

- (1) With the proposed two-layer deep learning-based classifier architecture, it has achieved a superior performance than classical deep learning approaches. In particular, the pre-weighting vector, which gives the maximum classification accuracy, has been optimized using artificial bee colony (ABC) algorithm and the performance has been improved.
- (2) Ethnic and genetic differences affect the diagnosis of the disease based on blood data. Using the proposed feature acquisition method, it is ensured to obtain the features related to the value with higher discrimination in each different region. So, the performance of the classifier is also increased.
- (3) As a result, the proposed model can be used to help doctors in early diagnosis of the COVID-19 disease.
- (4) A novel classifier architecture based on deep neural network using for COVID-19 detection gives better results for classification.

The paper is organized as follows. In Section 2, deep learning concept was explained. In Section 3, the studies about COVID-19 prediction using laboratory findings with artificial intelligence methods, were reviewed. In Section 4, firstly, an algorithm was developed to obtain a new feature set based on these findings and COVID-19 prediction was realized using three deep learning classifiers with different architectures. Accuracy, precision, recall, F1-score, and AUC parameters were compared with the results obtained by the classical findings. Afterwards, a hybrid classifier architecture based on deep learning was designed and classification performance was measured with the same measurement parameters mentioned above. In Section 5, the conclusions of the study were interpreted, and future work is presented.

2. Deep learning and artificial bee colony algorithm

Deep learning has gained popularity because of the performance improvements of hardware units such as graphics cards and the decrease in the unit prices. The increase in the number of training data and studies about machine learning and information processing have also contributed to deep learning. Different deep learning architectures such as artificial neural network (ANN), convolutional neural network (CNN) and recurrent neural network (RNN) have been widely used in many fields such as image classification, natural language processing and speech recognition [13,14]. ANN is an information processing approach inspired by the human biological nervous system. It consists of neurons, activation functions, input, output, and hidden layers. Every layer in ANN contains a set of neurons that form a hierarchy. The output of the previous layer is the input to the next layer. Each of the layers learn more complicated relationships from the input data. CNN is a deep learning algorithm developed for visual data processing such as images and videos. CNN has various layer types fulfilling various duties. These are called convolution layer, activation function layer, pooling layer, fully connected layer, and dropout layer [13]. In RNN structures, the result depends not only on the current inputs but also on the other inputs. These networks generate their output with the combination of current and previous data [15].

Swarm intelligence is a research area that models the behaviour of self-organizing populations such as flock of birds, ant colonies or bee colonies mathematically. In recent years, the development of swarm intelligence algorithms and the effective results of them in solving engineering problems in many research

areas has attracted the attention of many researchers and practitioners. Especially Artificial Bee Colony (ABC) algorithm came to the fore with its successful, easy and effective use. Karaboga introduced the ABC algorithm that is modelled food search behaviour of honey bees. Using literature benchmark functions, the performance of the ABC algorithm has been compared with those of the well-known meta-heuristic algorithms such as differential evolution (DE) and particle swarm optimization (PSO) [16,17]. Also, ABC algorithm has been efficiently employed in different research areas such as signal and image processing [18–20]. Especially in recent times, it has been successfully applied in hybridized approaches based deep learning and machine learning techniques [21,22].

3. Literature review

Initially, computer aided diagnosis systems (CADs) that are implemented via deep learning classifiers with laboratory findings as inputs, were reviewed.

Literature review was realized in Science Direct, IEEE, PubMed, Google Scholar, and arxiv.org with keywords such as COVID-19 blood artificial intelligence/deep learning, COVID-19 laboratory data deep learning, and the results were summarized below.

Schwab et al. compared the performances of machine learning and artificial neural network models such as support vector machine (SVM), Neural Network (NN), Gradient Boosting (XGB), Random Forest (RF) using 5644 blood analysis data in the prediction of COVID-19 [23].

Mei et al. combined computer tomography (CT) findings with laboratory data to predict COVID-19 using CNN and machine learning and obtained 84.3% sensitivity, 82.8% specificity and 92% AUC [24].

Banerjee et al. implemented machine learning algorithms and artificial neural networks to predict COVID-19 disease. They presented the performances of ANN, RF and glmnet classifiers using 14 different laboratory data. Additionally, they measured the performance of LF classifiers with fewer laboratory data [25].

Jiang et al. presented the classification performances of machine learning algorithms using a combination of clinical and laboratory findings [26].

Batista et al. tested the predictive performance of COVID-19 positive diagnosis using neural networks, gradient boosted trees, random forests, logistic regression and support vector machines with a combination of laboratory findings and clinical data such as age and gender [27].

Brinati et al. tested the accuracies of machine learning classifiers (Decision Tree, Extremely Randomized Trees, K-nearest neighbours, Logistic Regression, Native Bayes, Random Forest) that use laboratory findings and gender clinical data. They reached 82% and 86% accuracy values with the proposed machine learning classifiers [28].

Alakus and Turkoglu used a laboratory data set that contains 600 samples to compare deep learning approaches for prediction of COVID-19. The experimental results were obtained as 92.30% accuracy, 93% F1-score, 92.35% precision, 93.68% recall and 90% AUC [29].

4. Material and method

The aim of this study is to obtain a new feature group based on laboratory findings using the developed algorithm and to detect COVID-19 with a new hybrid classifier based on deep learning that uses the new feature group as input parameters.

In this study, the free-access dataset provided by Alakus and Turkoglu (<https://github.com/burakalakus/COVID-19-Clinical>) was used. This dataset includes 600 samples and 18 laboratory

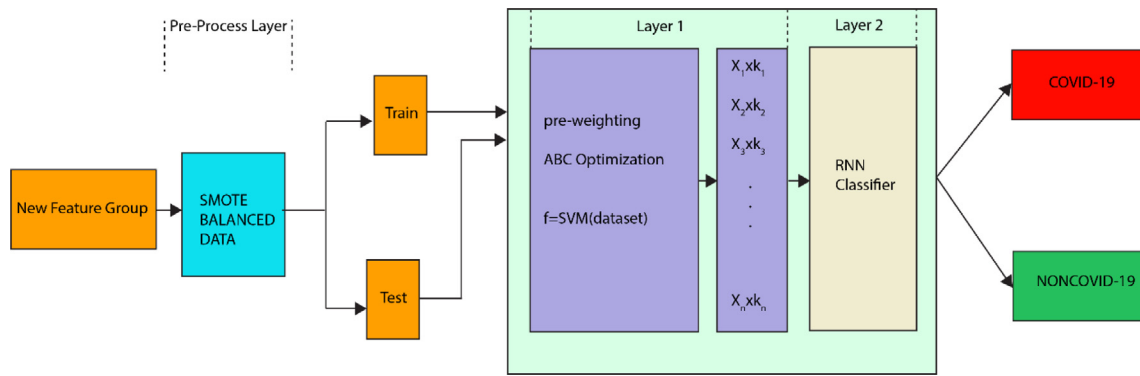


Fig. 1. Proposed classifier architecture.

findings. Table 1 shows the names of the findings. 520 samples belong to the patients diagnosed as COVID-19, and the other 80 samples belong to the healthy individuals. This dataset does not include clinical data of patients such as age and gender. Samples of healthy individuals are labelled as 1, and samples of patients are labelled as 0.

Laboratory findings (blood values) are decisive parameters used by the medical experts for the diagnosis of COVID-19. Sun et al. examined statistically the discrimination of the findings such as Leucocytes, Monocyte, and Platelet in the prediction of the disease using SPSS. Additionally, it has also been encountered that the medical experts have rated some of these laboratory findings together such as Neutrophils-to-lymphocytes ratio [30].

According to the researches, blood findings may give different results in different ethnic and genetic groups [6].

Therefore, the parameters that depend on the blood findings can have differences in discrimination due to the ethnical and genetic factors when used in deep learning systems. For example, while Monocyte-to-lymphocytes ratio [30] is a more important parameter in the diagnosis of COVID-19 using artificial intelligence blood findings sampled from a European country, Neutrophils-to-lymphocytes ratio is more significant in Asian genetics.

By applying the algorithm given below, first the data set was divided into patient and healthy groups in terms of each laboratory finding (parameters). Distribution test was applied for each parameter. *p* values were calculated using Student's test for the data with normal distribution, and Mann-Whitney U test for the others [31]. As shown in Table 1, the finding with the highest *p* value was serum glucose. Different parameters may be more determinant in populations belonging to different locations. Thanks to this method, genetic and ethnic differences are also taken into account.

In this study, first the rating method of Sun et al. [30] that was applied for some of the findings, was applied to all findings using Serum Glucose which has the highest *p* value among classes. To the best of our knowledge, there is no study dealing with this aspect of the subject for computer-aided diagnostic systems. The algorithm to obtain a new feature group is given below.

A new parameter group was obtained by dividing the parameter group with the highest *p* value to the other parameters. A value of 0 can be encountered simultaneously both for *Pmax* and other parameter groups, in this division process. To deal with this zero-division problem, the maximum value of *Pmax* group was used as a constant division value instead of the zero value of *Pmax* group. Data_set_1 refers to the new feature group.

The block diagram of the proposed architecture is given in Fig. 1. In architecture, the pre-processing layer functions as the balancing layer using the SMOTE technique.

Table 1
P values of the laboratory findings.

| No | Laboratory data | <i>P</i> Value | No | Laboratory data | <i>P</i> Value |
|----|-----------------|----------------|-----|--------------------------|----------------|
| 1. | Hamatocrit | 0.001 | 10. | Serum Glucose | 0.995 |
| 2. | Hamoglobin | 0.008 | 11. | Neutrophils | 0.632 |
| 3. | Platelets | 0.000 | 12. | Urea | 0.88 |
| 4. | Red blood Cells | 0.004 | 13. | C-reactive Protein mg/dL | 0.001 |
| 5. | Lymphocytes | 0.630 | 14. | Creatinine | 0.038 |
| 6. | Leukocytes | 0.000 | 15. | Potassium | 0.142 |
| 7. | Basophils | 0.156 | 16. | Sodium | 0.126 |
| 8. | Eosinophils | 0.000 | 17. | Alanine transaminase | 0.717 |
| 9. | Monocytes | 0.000 | 18. | Aspartate transaminase | 0.399 |

The data set based on classical findings includes 520 samples diagnosed with COVID-19 and 80 healthy samples. Here, the ratio of the majority sample number to the minority sample is 6.5. By the SMOTE technique, this ratio was reduced to 1 [9,10]. The SMOTE method is based on the algorithm which is given below [32].

In the literature, multi-layer architectures have been used in various areas in the artificial intelligence-based classification problems. In a study, the system architecture was designed, as each ANFIS output is an input to the next ANFIS layer [33].

Uzunhisarcıklı and Göreke proposed a two-layer fuzzy inference system for mass detection in ultrasound images [34].

Liu et al. used a CNN neural network to identify diabetic retinopathy from eye fundus images. In their work, they used a pre-weighting coefficient for the convolution layer outputs of the neural network [35].

In this study, pre-weighting vector was applied for the input layer of the deep neural network and a two-layer classifier in hybrid structure was proposed.

In Fig. 1, $X_1 \dots X_n$ represent the new feature group in the pre-weighting layer which is called as Layer 1. In Layer 1, the weights of the input parameter were determined by calculating a pre-weighting coefficients. Calculation of these coefficients ($k_1 \dots k_n$) is an optimization problem and the ABC algorithm was efficiently used to optimize them. The pseudo-code structure of the basic ABC algorithm is given in Algorithm 3.

In this study, the accuracy parameter (*Acc*) of SVM classifier was used as the cost function and ABC algorithm tries to minimize the objective function (*J*) given in Eq. (2).

$$J = \min\left(\frac{1}{Acc_{SVM}}\right), Acc_{SVM} = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

TP and *TN* show the number of correctly predicted positive and negative samples, whereas *FP* and *FN* represent to the number of incorrectly predicted positive and negative samples. These values were calculated using new feature group for SVM classifier. A nested software was developed to realize it. In the software,

Algorithm 1: Obtaining a new feature group

Data_set=600×18 matrix and 600×1 label (0 or 1).
 1: Group the data set as patient/healthy for each parameter.
 2: Apply the distribution test for each parameter pair.
 3: Choose Student’s test/Mann-Whitney test for normal/abnormal groups.
 4: Calculate the p values for data groups.
 5: Save the index (idx) of the parameter with highest p value ($Pmax$).
 6: Value_max=maximum ($Pmax$)
 7: **for** i =row **do**
 8: **for** j = column **do**
 9: **if** Data_set ($index==idx$) **then**
 10: Data_set (i, j) = $Pmax(idx)$
 11: **end if**
 12: **if** Data_set (i, j) == 0 **then**
 13: Data_set_1 (i, j) = Value_max
 14: **else**
 15: Data_set_1 (i, j) = Data_set (i, j) / $Pmax(idx)$
 16: **end if**
 17: **end for**
 18: **end for**

Algorithm 2: SMOTE method

1: Minority class set X for each x in X select r nearest neighbours.
 2: The sampling rate SR , for each x in X randomly selected from its r -nearest neighbours construct the set X_j .
 3: For each x_r in X_j ($r=1, 2, 3, \dots, SR$) use Equation 1 and generate new sample.

$$x_{smote} = x + rand(0,1) \times |x - x_r| \tag{1}$$

Table 2
 Comparative results of competitor algorithms for pre-weighting vector optimization in Layer 1.

| Algorithm | Best | Worst | SD | Mean | Time (s) |
|-----------|---------|---------|--------|---------|----------|
| ABC | 95.8333 | 93.3333 | 0.6922 | 94.2500 | 640.1677 |
| GA | 95.0000 | 93.3333 | 0.4487 | 94.0833 | 697.1724 |
| PSO | 95.0000 | 93.3333 | 0.5000 | 94.0000 | 673.7480 |

SVM code was run inside the ABC algorithm code. The developed software calculates the optimum pre-weighting vector that will provide the best accuracy in the SVM classifier.

In the pre-weighting vector optimization, ABC algorithm was compared with particle swarm optimization (PSO) and genetic algorithm (GA). For a fairly comparison, population size and maximum iteration number were selected 20 and 500. Lower and upper bound were determined as 1 and 3. Also, competitor algorithms were run 20 times with different initial values. In Table 2, the performances of competitor algorithms were shown and accuracy parameter Acc_{SVM} were analysed when objective function calculated $1/Acc_{SVM}$.

As seen Table 2, for the ABC algorithm, mean and best accuracy values are obtained 94.2500 and 95.8333, and they are bigger than those of the other algorithms. So, ABC algorithm shows good performance for pre-weighting vector optimization and improves accuracy of SVM classifier. Also, GA achieves the best SD values for multiple runs. From Table 2, the best pre-weighting vector are achieved by ABC algorithm and it is given by Eq. (3).

$$k = [2.72539285, 1.56784067, 2.68382047, 1.1796389, 2.9169504, 1.92253802, 1.89977199, 1.49214547, 1.52678425, 2.33359363, 1.33939422, 1.49647264, 2.8662631,$$

$$2.03975598, 1.54127335, 1.86499559, 2.54004318, 1.08099064] \tag{3}$$

The layer called as Layer 2 represents the three deep learning classifiers with different basic architectures. In this study, these three deep learning classifiers were designed separately and their performance was measured. The RNN type that provides the best performance was selected.

4.1. Used deep learning architectures

In this study, three deep learning classifiers with basic architectures were used. The hyper-parameter values of each architectures are given in Table 3. Normally, in the artificial intelligence studies with relatively small datasets, cross-validation method is mostly preferred. But especially in medical clinical application, this cross-validation approach gives fewer clear results [29]. In this study, to have a clearer result, the tests were realized using train-test split approach and cross-validation approach was not preferred. Also, split ratio was determined as 20.

4.2. Test criteria

The performances of deep learning classifiers with new features were measured with accuracy, recall, precision, and F1-score metrics. The mathematical representations of these metrics are given with Eqs. (4)–(7).

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{4}$$

$$Recall = TP/(TP + FN) \tag{5}$$

$$Precision = TP/(TP + FP) \tag{6}$$

Algorithm 3: (Pseudo-code structure of the basic ABC algorithm)

-
- 1: Initialize the population of food sources k_i by using $k_i^j = k_{\min}^j + rand[0,1](k_{\max}^j - k_{\min}^j)$
($i = 1 \dots SN$; SN is number of food sources. $j = 1 \dots D$; D is number of optimization parameters)
 - 2: Evaluate the population of solution k_i .
 - 3: cycle=1
 - 4: **repeat**
 - 5: In the Employed bee phase, produce new solution v_i (neighbor food source) by using
 $v_i^j = k_i^j + \phi_j^j(k_i^j - k_m^j)$ and evaluate the new solution. (ϕ_j : Random weighting value within the range [-1, 1].)
 - 6: Apply greedy selection process between k_i and v_i for the employed bees.
 - 7: Calculate the fitness of the solution k_i using $fit(k_i) = \begin{cases} 1/(1+f(k_i)), & f(k_i) \geq 0 \\ 1+abs(f(k_i)), & f(k_i) < 0 \end{cases}$
 - 8: For the Onlooker bee phase, calculate the probability of the solution k_i using $prb_i = \frac{fit(k_i)}{\sum_{i=1}^{SN} fit(k_i)}$

Produce new solution v_i by using $v_i^j = k_i^j + \phi_j^j(k_i^j - k_m^j)$ depending on prb_i values and evaluate the new solution.
 - 9: Apply greedy selection process between k_i and v_i for the onlooker bees.
 - 10: Determine the abandoned solution and produce randomly new solution k_i using
 $k_i^j = k_{\min}^j + rand[0,1](k_{\max}^j - k_{\min}^j)$
 - 11: Memorize the best solution
 - 12: Cycle=cycle+1
 - 13: **until** cycle=maximum cycle number
-

Table 3
Deep learning classifier architecture and hyper parameters.

| Parameter | ANN | CNN | RNN |
|---------------------|----------------------|----------------------|----------------------|
| Units | 32-16-8 | 512-256 | - |
| Layers | 1-2-3 | 1-2 | 1 |
| Activation function | ReLU | ReLU | ReLU |
| Learning rate | 1e-3 | 1e-3 | 1e-3 |
| Loss function | Binary cross entropy | Binary cross entropy | Binary cross entropy |
| Epochs | 250 | 250 | 250 |
| Optimizer | SGD | SGD | SGD |
| Decay | 1e-5 | 1e-5 | 1e-5 |
| Momentum | 0.3 | 0.3 | 0.3 |
| Fully Conn.units | - | 2048-1024 | 2048-1024 |
| Fully Conn.layer | - | 1-2 | 1-2 |
| RNN units | - | - | 512 |
| Dropout | - | - | 0.25 |

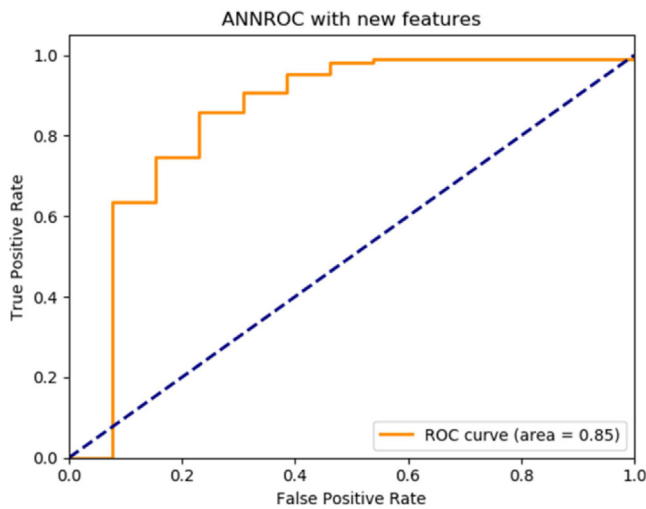


Fig. 2. ANN deep learning architecture ROC graph.

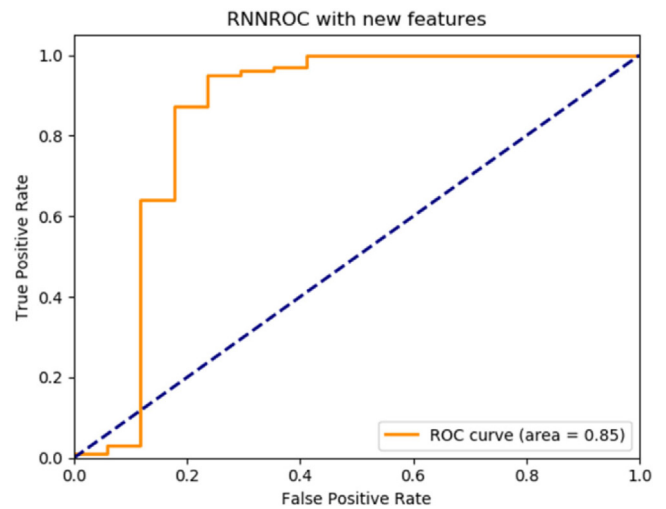


Fig. 4. RNN deep learning architecture ROC graph.

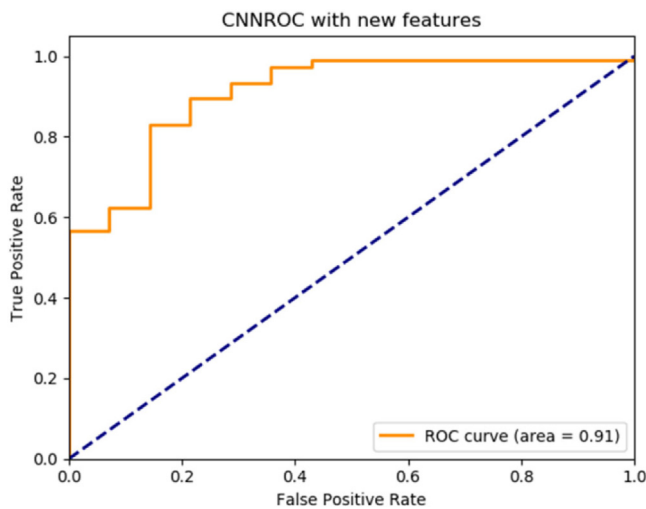


Fig. 3. CNN deep learning architecture ROC graph.

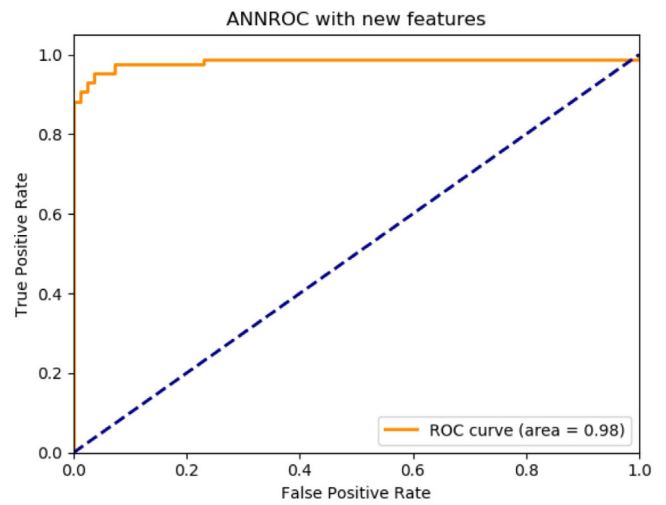


Fig. 5. Proposed method for ANN deep learning architecture ROC graph.

$$F1 - score = 2 \times precision \times recall / (precision + recall) \quad (7)$$

Here TP (true positive) is the number of patients diagnosed as patient, TN (true negative) is the number of healthy individuals diagnosed as healthy, FP (false positive) is the number of healthy individuals diagnosed as patient and, FN (false negative) is the number of patients diagnosed as healthy [36].

Additionally, receiving operating characteristic (ROC) curve gives the change between the recall and $FP/(TN+FP)$ ratio [37], and it is used in the calculation of AUC (area under the curve). AUC is interpreted as the possibility of a higher test measure of a random patient than a random healthy individual [38]. The reliability and the validity of the model is proven by measuring the accuracy of the model [39]. F1-score is used for determining that the classifier has small FP and FN values [26].

4.3. Performance comparison

Tests have been realized on an i7 2.6 GHz computer using Microsoft Visual Studio Community 2017 IDE and Python Anaconda 5.2.0 environment.

In the first stage of the test process, the performance of the new feature group against classical laboratory findings was measured with classifiers in standard architecture. In the second step,

the new feature group was classified using the classifiers in the proposed architecture and its performance was compared with the studies in the literature.

By the implemented algorithm, the new feature set was applied as the input to three deep learning classifiers with basic architectures. The performance of each classifier according to the label information in the data set was measured by accuracy, recall, precision and F1-score metrics and the ROC graphs were plotted. The results are given in Table 4. ROC graphs obtained using novel features are shown in Figs. 2–4 for ANN, CNN and RNN architectures.

The performance results of deep learning classifiers using laboratory findings for the detection of COVID-19 are given in Table 5 [29].

When the results of Tables 4 and 5 are compared, it is seen that results in Table 4 outperform the results in Table 5. The averages of each performance measure of three different classifiers using both classical and the novel features are given in Table 6.

According to the proposed method classifier performances are as follows: ANN, accuracy 92.08%, F1-score 91.89%, precision 91.89%, recall 91.89%, AUC 98%, CNN, accuracy 94.79%, F1-score 94.79%, precision 94.79%, recall 94.99%, AUC 100%, RNN, accuracy 94.95%, F1-score 94.98%, precision 94.98%, recall 94.98%,

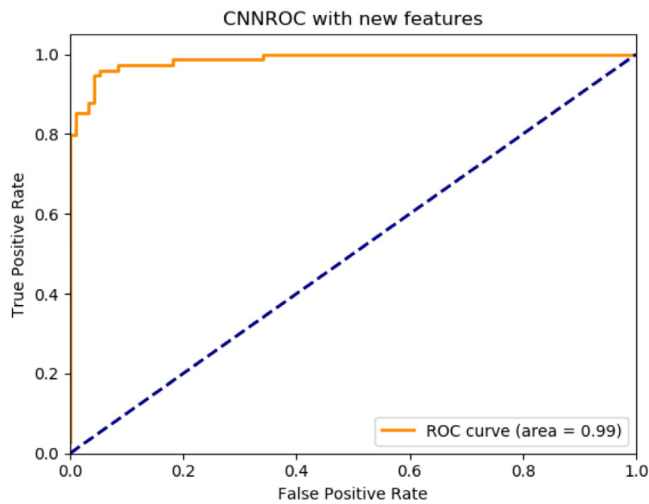


Fig. 6. Proposed method for CNN deep learning architecture ROC graph.

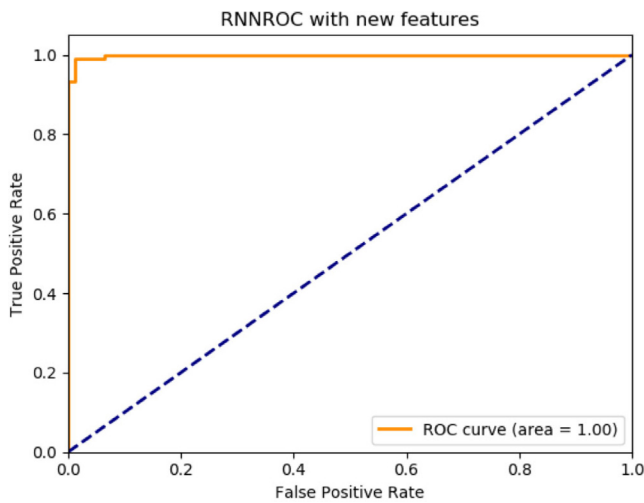


Fig. 7. Proposed method for RNN deep learning architecture ROC graph.

Table 4
Classification performance with the new features based on the laboratory findings (with standard deep neural network architecture).

| Classifier | Accuracy | F1-score | Precision | Recall | AUC |
|------------|----------|----------|-----------|--------|------|
| ANN | 0.8910 | 0.8910 | 0.8910 | 0.8910 | 0.85 |
| CNN | 0.8946 | 0.8950 | 0.8950 | 0.8950 | 0.90 |
| RNN | 0.9224 | 0.9220 | 0.9220 | 0.9220 | 0.85 |

Table 5
Classification performance with the classical laboratory findings (with standard deep neural network architecture).

| Classifier | Accuracy | F1-score | Precision | Recall | AUC |
|------------|----------|----------|-----------|--------|------|
| ANN | 0.8690 | 0.8713 | 0.8713 | 0.8713 | 0.85 |
| CNN | 0.8735 | 0.8856 | 0.8847 | 0.8867 | 0.80 |
| RNN | 0.8400 | 0.8427 | 0.8428 | 0.8427 | 0.83 |

AUC 100%. The best performance belongs to the deep learning classifier in RNN structure.

The ROC graphic obtained with the proposed method for ANN, CNN and RNN architecture are shown in Figs. 5–7. The performance comparison of the proposed method against the studies in the literature is given in Table 7.

Table 6
Average classification performance of classical and new features.

| Feature | Accuracy | F1-score | Precision | Recall | AUC |
|---------|----------|----------|-----------|--------|--------|
| Classic | 0.8608 | 0.8665 | 0.8662 | 0.8669 | 0.8266 |
| New | 0.9173 | 0.9171 | 0.9171 | 0.9171 | 0.9266 |

Table 7
The proposed method and literature performance comparison.

| Reference | Accuracy | F1-score | Precision | Recall | AUC |
|--------------------------|---------------|---------------|---------------|---------------|-------------|
| Schwab et al. [23] | – | – | – | 0.8200 | 0.98 |
| Mei et al. [24] | – | – | – | 0.8430 | 0.92 |
| Banerjee et al. [25] | 0.9100 | – | – | 0.9200 | 0.95 |
| Jiang et al. [26] | 0.8000 | – | – | – | – |
| Batista et al. [27] | 0.8420 | 0.7800 | 0.7800 | 0.8000 | 0.85 |
| Brinati et al. [28] | 0.8600 | – | – | 0.9300 | 0.85 |
| Alakus and Turkoglu [29] | 0.9230 | 0.9300 | 0.9235 | 0.9368 | 0.90 |
| Proposed method | 0.9495 | 0.9498 | 0.9498 | 0.9498 | 1.00 |

5. Conclusion and discussion

In the first stage of the study, the data set containing the laboratory findings were obtained from the link <https://github.com/burakalakus/COVID-19-Clinical>. Initially, the distribution of each finding was tested for normality using SPSS software. After the normality distributions were determined, the predictive values of each group were calculated and the finding with the highest value was chosen as the rating parameter. In the next stage, a new feature set was obtained with the developed algorithm using this parameter. COVID-19 was predicted using the new features and three deep learning classifiers with basic architectures (ANN, CNN, RNN), and the prediction performances of each classifier were measured with accuracy, F1-score, precision, recall and AUC metrics.

Alakus and Turkoglu used the laboratory findings of their dataset to test and compare the classification performance of deep learning classifiers with different architectures. In this study, COVID-19 prediction was implemented using new features and three deep learning classifiers with basic architectures which were tested by Alakus and Turkoglu. The results of this study were compared with the results of Alakus and Turkoglu for the classifiers with the same architecture. When Tables 4 and 5 are compared, the new features that depend on the laboratory findings have better results than the classical findings for the prediction of COVID-19.

Considering the results given in Table 7, it can be stated that the proposed method was obtained more successful than in literature works.

This study depends on the samples taken from Hospital Israelita Albert Einstein. In the future work, using data sets obtained from the groups with different genetics (different countries etc.), the effect of the biological and genetic differences in the prediction of COVID-19 will be searched.

These differences can determine more clearly that the different findings are more effective on different groups for the detection of COVID-19.

Therefore, the parameter detection process of CAD systems that implement deep learning can be shortened for different regions, and it can also have a positive contribution on the CAD performance. It helps the doctors to have more accurate answers in the decision process.

Factors such as the environment of the swab sample from the mouth and the qualification of the sample affects the accuracy of the RT-PCR test used worldwide for the diagnosis of the disease [40,41]. In this regard, using computer aided diagnosis tools such as deep neural networks implemented in this study can help doctors in the early diagnosis of the disease. With the early diagnosis of the disease, more successful results can be achieved in the treatment and the mortality rate can be decreased.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] N. Dey, R. Mishra, S.J. Fong, K.C. Santosh, S. Tan, R.G. Crespo, COVID-19: Psychological and psychosocial impact, fear, and passion, *Digit. Gov.: Res. Pract.* 2 (1) (2020) 1–4, <http://dx.doi.org/10.1145/3428088>.
- [2] S. Fong, G. Li, N. Dey, R.G. Crespo, E. Herrera-Viedma, Finding an accurate early forecasting model from small dataset: A case of 2019-nCoV novel coronavirus outbreak, *Int. J. Interact. Multimedia Artif. Intell.* 6 (1) (2020) 1–10, <http://dx.doi.org/10.9781/ijimai.2020.02.002>.
- [3] H.C. Jan, W.H. Yang, C.H. Ou, Combination of the preoperative systemic immune-inflammation index and monocyte-lymphocyte ratio as a novel prognostic factor in patients with upper-tract urothelial carcinoma, *Ann. Surg. Oncol.* (26) (2019) 669–684, <http://dx.doi.org/10.1245/s10434-018-6942-3>.
- [4] T. Demirdal, P. Sen, The significance of neutrophil-lymphocyte ratio, platelet-lymphocyte ratio and lymphocyte-monocyte ratio in predicting peripheral arterial disease, peripheral neuropathy, osteomyelitis and amputation in diabetic foot infection, (144) 2018, 118–125, <http://dx.doi.org/10.1016/j.diabres.2018.08.009>.
- [5] W. Guan, Z. Ni, Y. Hu, W. Liang, C. Ou, G.J. He, et al., Clinical characteristics of coronavirus disease 2019 in China, *New Engl. J. Med.* 382 (18) (2019) 1708–1720, <http://dx.doi.org/10.1056/NEJMoa2002032>.
- [6] S. Sarnaik, P.M. Simpson, M. Hamre, J.A. Williams, Y. Ravindranath, Distributions of ethnic differences in blood counts are suggestive of genetic variability, *Pediatr. Res.* (45) (1999) 152, <http://dx.doi.org/10.1203/00006450-199904020-00903>.
- [7] T. Davenport, R. Kalakota, The potential for artificial intelligence in healthcare, *Future Healthc. J.* 6 (2) (2019) 94–98, <http://dx.doi.org/10.7861/futurehosp.6-2-94>.
- [8] M.A. Mazurowski, P.A. Habas, J.M. Zurada, J.Y. Lo, J.A. Baker, G.D. Tourassi, Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance, *Neural Netw.* (21) (2008) 427–436, <http://dx.doi.org/10.1016/j.neunet.2007.12.031>.
- [9] R. Zhu, Y. Guo, J.H. Xue, Adjusting the imbalance ratio by the dimensionality of imbalanced data, *Pattern Recognit. Lett.* (133) (2020) 217–223, <http://dx.doi.org/10.1016/j.patrec.2020.03.004>.
- [10] P. Shamsolmoali, M. Zareapoor, L. Shen, A.H. Sadka, J. Yang, Imbalanced data learning by minority class augmentation using capsule adversarial networks, *Neurocomputing* (2020) <http://dx.doi.org/10.1016/j.neucom.2020.01.119>.
- [11] G. Douzas, F. Bacao, Effective data generation for imbalanced learning using conditional generative adversarial networks, *Expert Syst. Appl.* (91) (2018) 464–471, <http://dx.doi.org/10.1016/j.eswa.2017.09.030>.
- [12] E. Lashgari, D. Liang, U. Maoz, Data augmentation for deep-learning-based electroencephalography, *J. Neurosci. Methods* (346) (2020) 108885, <http://dx.doi.org/10.1016/j.jneumeth.2020.108885>.
- [13] F. Auro, H. Chiroma, A.Y. Gital, M. Almutairi, S.M. Abdulhamid, J.H. Abawajy, Deep learning architectures in emerging cloud computing architectures: Recent development, challenges and next research trend, *Appl. Soft Comput.* (96) (2020) 106582, <http://dx.doi.org/10.1016/j.asoc.2020.106582>.
- [14] E. Uzunhisarcıklı, V. Göreke, V. Sari, Comparison of type-2 fuzzy inference method and deep neural networks for mass detection from breast ultrasonography images, *Cumhur. Sci. J.* 41 (4) (2020) 968–975, <http://dx.doi.org/10.17776/cs.691683>.
- [15] A. Sherstinsky, Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network, *Physica D* (404) (2020) 132306, <http://dx.doi.org/10.1016/j.physd.2019.132306>.
- [16] D. Karaboga, B. Akay, On the performance of artificial bee colony (ABC) algorithm, *Appl. Soft Comput.* 8 (1) (2008) 687–697, <http://dx.doi.org/10.1016/j.asoc.2007.05.007>.
- [17] D. Karaboga, B. Akay, A comparative study of artificial bee colony algorithm, *Appl. Math. Comput.* 214 (1) (2009) 108–132, <http://dx.doi.org/10.1016/j.amc.2009.03.090>.
- [18] S. Kockanat, Acceleration harmonics estimation and elimination with MABC-RLS algorithm: Simulation and experimental analyses on shaking table, *Appl. Soft Comput.* (92) (2020) 106377, <http://dx.doi.org/10.1016/j.asoc.2020.106377>.
- [19] S. Kockanat, T. Koza, N. Karaboga, A. Logoglu, Adaptive FIR filtering using ABC algorithm: a noise reduction application on mitral valve doppler signal, *Elektron. Elektrotech.* 24 (5) (2018) 62–68, <http://dx.doi.org/10.5755/j01.eie.24.5.21845>.
- [20] S. Kockanat, N. Karaboga, A novel 2d-ABC adaptive filter algorithm: A comparative study, *Digit. Signal Process.* (40) (2015) 140–153, <http://dx.doi.org/10.1016/j.dsp.2015.02.010>.
- [21] A. Banharsakun, Towards improving the convolutional neural networks for deep learning using the distributed artificial bee colony method, *Int. J. Mach. Learn. Cybern.* (10) (2019) 1301–1311, <http://dx.doi.org/10.1007/s13042-018-0811-z>.
- [22] A. Kumar, G. Kabra, E.K. Mussada, M.K. Dash, P.S. Rana, Combined artificial bee colony algorithm and machine learning techniques for prediction of online consumer repurchase intention, *Neural Comput. Appl.* (31) (2019) 877–890, <http://dx.doi.org/10.1007/s00521-017-3047-z>.
- [23] P. Schwab, A.D. Schütte, B. Dietz, S. Bauer, Clinical predictive models for COVID-19: Systematic study, *J. Med. Internet Res.* 22 (10) (2020) e21439, <http://dx.doi.org/10.2196/21439>.
- [24] X. Mei, H.C. Lee, K. Diao, M. Huang, B. Lin, M.C. Liu, et al., Artificial intelligence-enabled rapid diagnosis of patients with COVID-19, *Nat. Med.* (26) (2020) 1224–1228, <http://dx.doi.org/10.1038/s41591-020-0931-3>.
- [25] A. Banerjee, S. Ray, B. Vorselaars, J. Kitson, M. Mamalakis, S. Weeks, M. Baker, L.S. Mackenzie, Use of machine learning and artificial intelligence to predict SARS-CoV-2 infection from full blood counts in a population, *Int. Immunopharmacol.* (86) (2020) 106705, <http://dx.doi.org/10.1016/j.intimp.2020.106705>.
- [26] X. Jiang, M. Coffee, A. Bari, J. Wang, X. Jiang, J.J. Huang, et al., Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity, *Comput. Mater. Continua* 63 (1) (2020) 537–551, <http://dx.doi.org/10.32604/cmc.2020.010691>.
- [27] A.F.de M. 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, 2020, <http://dx.doi.org/10.1101/2020.04.04.20052092>.
- [28] D. Brinati, A. Campagner, D. Ferrari, M. Locatelli, G. Banfi, F. Cabitza, Detection of COVID-19 infection from routine blood exams with machine learning: A feasibility study, *J. Med. Syst.* 44 (8) (2020) 135, <http://dx.doi.org/10.1007/s10916-020-01597-4>.
- [29] T.B. Alakus, I. Turkoglu, Comparison of deep learning approaches to predict COVID-19 infection, *Chaos Solitons Fractals* (140) (2020) 110120, <http://dx.doi.org/10.1016/j.chaos.2020.110120>.
- [30] S. Sun, X. Cai, H. Wang, G. He, Y. Lin, B. Lu, et al., Abnormalities of peripheral blood system in patients with COVID-19 in Wenzhou, China, *Clin. Chim. Acta* (507) (2020) 174–180, <http://dx.doi.org/10.1016/j.cca.2020.04.024>.
- [31] C. Ai, L. Huang, Z. Zhang, A mann-whitney test of distributional effects in a multivalued treatment, *J. Statist. Plann. Inference* (209) (2020) 85–100, <http://dx.doi.org/10.1016/j.jspi.2020.03.002>.
- [32] N. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: Synthetic minority over-sampling technique, *J. Artificial Intelligence Res.* (16) (2002) 321–357, <http://dx.doi.org/10.1613/jair.953>.
- [33] M.S. Iraj, Multi-layer architecture for adaptive fuzzy inference system with a large number of input features, *Cogn. Syst. Res.* (42) (2017) 23–41, <http://dx.doi.org/10.1016/j.cogsys.2016.11.006>.
- [34] E. Uzunhisarcıklı, V. Göreke, A novel classifier model for mass classification using BI-RADS category in ultrasound images based on type-2 fuzzy inference system, *Sadhana* (43) (2018) 1–12, <http://dx.doi.org/10.1007/s12046-018-0915-x>.
- [35] Y.P. Liu, Z. Li, C. Xu, J. Li, R. Liang, Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network, *Artif. Intell. Med.* (99) (2019) 101694, <http://dx.doi.org/10.1016/j.jartmed.2019.07.002>.
- [36] W. Mumtaz, A. Qayyum, A deep learning framework for automatic diagnosis of unipolar depression, *Int. J. Med. Inform.* (132) (2019) 103983, <http://dx.doi.org/10.1016/j.ijmedinf.2019.103983>.
- [37] K. Hajian-Tilaki, Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation, *Casp. J. Intern. Med.* 4 (2) (2013) 627–635, <https://pubmed.ncbi.nlm.nih.gov/24009950>.
- [38] S. Safari, A. Baratloo, M. Elfil, A. Negida, Evidence based emergency medicine; part 5 receiver operating curve and area under the curve, *Emergency* 4 (2) (2016) 111–113, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4893763>.
- [39] R. Pierce, Evaluating information: validity, reliability, accuracy, triangulation, *Res. Methods Polit.* (2008) 79–99, <http://dx.doi.org/10.4135/9780857024589>.
- [40] R. Wölfel, V.M. Corman, W. Guggemos, M. Seilmaier, S. Zange, W.M.A. Müller, et al., Virological assessment of hospitalized patients with COVID-2019, *Nature* 581 (7809) (2019) 465–469, <http://dx.doi.org/10.1038/s41586-020-2196-x>.
- [41] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, F.P. Pang, et al., Sensitivity of chest CT for COVID-19: Comparison to RT-PCR, *Radiology* 296 (2) (2020) 115–117, <http://dx.doi.org/10.1148/radiol.20200432>.