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.
Novel ensemble of optimized CNN and dynamic selection techniques for accurate Covid-19 screening using Chest CT images

Sameena Pathan, P.C. Siddalingaswamy, Preetham Kumar, Manohara Pai M M, Tanweer Ali, U.Rajendra Acharya

PII: S0010-4825(21)00629-6
DOI: https://doi.org/10.1016/j.compbiomed.2021.104835
Reference: CBM 104835

To appear in: Computers in Biology and Medicine

Received Date: 16 May 2021
Revised Date: 30 August 2021
Accepted Date: 1 September 2021

Please cite this article as: S. Pathan, P. Siddalingaswamy, P. Kumar, M. Pai M M, T. Ali, U.R. Acharya, Novel ensemble of optimized CNN and dynamic selection techniques for accurate Covid-19 screening using Chest CT images, Computers in Biology and Medicine, https://doi.org/10.1016/j.compbiomed.2021.104835.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 Elsevier Ltd. All rights reserved.
Novel ensemble of optimized CNN and dynamic selection techniques for accurate Covid-19 screening using Chest CT images

Sameena Pathan¹, P C Siddalingaswamy², Preetham Kumar¹, Manohara Pai M M¹, Tanweer Ali³, U Rajendra Acharya⁴,⁵,⁶,⁷

¹Dept. of Information and Communication Technology, Manipal Institute of Technology, MAHE, Manipal
²Dept. of Computer Science and Engineering, Manipal Institute of Technology, MAHE, Manipal
³Dept. of Electronics and Communication Engineering, Manipal Institute of Technology, MAHE, Manipal
⁴Dept. of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
⁵Ngee Ann Polytechnic, Department of Electronics and Computer Engineering, 599489, Singapore
⁶Department of Biomedical Engineering, School of Science and Technology, SUSS University, Singapore
⁷Department of Biomedical Informatics and Medical Engineering, Asia University, Taichung, Taiwan

Corresponding author: tanweer.ali@manipal.edu

Abstract — The world is significantly affected by infectious coronavirus disease (covid-19). Timely prognosis and treatment are important to control the spread of this infection. Unreliable screening systems and limited number of clinical facilities are the major hurdles in controlling the spread of covid-19. Nowadays, many automated detection systems based on deep learning techniques using computed tomography (CT) images have been proposed to detect covid-19. However, these systems have the following drawbacks: (i) limited data problem poses a major hindrance to train the deep neural network model to provide accurate diagnosis, (ii) random choice of hyperparameters of Convolutional Neural Network (CNN) significantly affects the classification performance, since the hyperparameters have to be application dependent and, (iii) the generalization ability using CNN classification is usually not validated. To address the aforementioned issues, we propose two models: (i) based on a transfer learning approach, and (ii) using novel strategy to optimize the CNN hyperparameters using Whale optimization-based BAT algorithm + AdaBoost classifier built using dynamic ensemble selection techniques. According to our second method depending on the characteristics of test sample, the classifier is chosen, thereby reduced the risk of overfitting and simultaneously produced promising results. Our proposed methodologies are developed using 746 CT images. Our method obtained a sensitivity, specificity, accuracy, F-1 score, and precision of 0.98, 0.97, 0.98, 0.98, and 0.98, respectively with five-fold cross-validation strategy. Our developed prototype is ready to be tested with huge chest CT images database before its real-world application.

Index Terms — BGWO, Covid-19, CNN, DST, Ensemble, GWO, WOA, hyperparameters
1. INTRODUCTION

Covid-19, the disease caused by the SARS-CoV-2 was officially declared as a pandemic by World Health Organization (WHO), if not detected early, can be fatal and evolve as Acute Respiratory Disease Syndrome (ARDS). Thus, early prediction of risk factors and screening of patients is crucial to curb the spread in society thereby preventing mortality [1]. Limited access to health monitoring facilities hinders the development of sustainable cities. With the global pandemic swaying its way in the lives of people in a capricious manner, causing unprecedential damage to health, economy. It is necessary to combat the spread of covid-19 in early stages and provide proper treatment. Currently, RT-PCR is the most common screening method for the detection of covid-19. However, the method is laborious and also has low detection sensitivity in the initial stages. At the onset of covid-19, few changes are observed in the CT-scans of patients [2]. For instance, consolidation and dilation were observed by Zhao et al., [3], in covid affected patients.

A CT scan provides information regarding the size, shape, texture, and density of various anatomical regions of the body. Figure 1 illustrates the sample CT images of covid affected and non-covid affected images. In contrast, to the conventional X-Rays, much information is embedded in a CT scan image. This information is usually analyzed by clinicians to provide the diagnosis. However, due to the subjectivity involved and also due to the scarcity of clinicians in fighting this global pandemic, it is necessary to develop an objective diagnostic mechanism, that could speed up the diagnosis and aid in the early detection.

![Sample chest CT images](image1)

Fig. 1. Sample chest CT images: covid-19 affected patients (first row images), and normal subjects (second row) [4].
Off late, deep learning tools have been used to automate and speed up the diagnosis [4-15]. A self-trans approach using transfer learning and feature extraction methods was proposed by He et al., [5] achieving an F1-score of 0.85. A precision of 0.84, recall of 0.95, was obtained by Mobiny et al., [6] using a detailed oriented capsule network. Polsinelli et al., [7] achieved an accuracy of 0.83 by using SqueezeNet Convolution Neural Networks (CNN) architecture. A voting-based scheme using certain deep learning models was proposed by Silva et al., [4] achieving an accuracy of 0.87. A weakly-supervised method using deep learning was utilized by Hu et al., [8], achieving covid classification accuracy of 0.89. An accuracy of 0.9 and sensitivity of 0.8 was obtained by Wang et al., [9] by employing U-net architecture on 3D CT volumes. However, the method was trained on imperfect ground truth masks thereby resulting in misleading results [9]. Ardakani et al., [10] developed a CAD tool termed as CovidDiag by extracting a set of twenty radiological features using CT images and further classifying into covid and non-covid classes using five different supervised classifiers. Ozakya et al., [11] achieved a sensitivity of 92% using a convolutional support vector machine. Further few other pre-trained models were also evaluated for the CT covid images. A sensitivity of 68% was obtained by Ozturk et al., [12] using a set of 495 shrunken features obtained by employing stacked auto encoder and principal component analysis. Maghdid et al., [13] obtained an accuracy of 94% using pre-trained AlexNet architecture, additionally, a dataset of images was also created [13]. A detailed review of methods for detection of covid19 using deep learning techniques and lung segmentation is given in [14]. A set of ten prominent CNN’s were investigated by Ardakani et al., [15], for distinguishing covid and non-covid CT images, it was observed that ResNet 101 and Xception resulted in good classification performance [15]. The Table I provides a summary of the studies pertaining to CT image based covid-19 detection techniques.

In the proposed work, two covid-19 classification architectures are proposed, the first architecture is built using features extracted from five standard CNN architectures and classified using AdaBoost of decision stump trees created using ensemble selection techniques. The second architecture is a novel concept, which is built by optimizing the hyperparameters of a CNN using Whale Optimization and Bat (WOA-BAT) swarm heuristics algorithm, to obtain features from CT images, and the extracted features are further classified using a robust classification set-up. A CNN hyperparameter optimization methodology is proposed by Pathan et al. [16], however, the proposed method is different from the method in [16], in the following ways, (i) An optimized CNN is used for extraction of features and further dynamic ensemble of classifiers are used to classify CT images, into covid-19 and non covid-19, whereas in [16], a GWO optimized CNN is used for three class classification of images, (ii) The proposed method addressed the limited dataset and overfitting problem, since a dynamic and robust
A classifier set-up is used for classification, rather than a complex CNN, which requires a larger dataset for achieving good classification accuracy.

Table I. Overview of the studies conducted on automated detection of covid-19 using chest CT images.

| Ref. | Dataset | Feature types | Feature Selection | Classification | Computational Complexity | Runtime | Accuracy | Sensitivity |
|------|---------|----------------|--------------------|----------------|--------------------------|---------|----------|-------------|
| [5]  | Chest Dataset 746 images | Self-trans approach for feature representation | - | ResNet-50, Densenet-16 | Computational expensive | - | 0.86 | - |
| [6]  | Chest Dataset 746 images | - | - | Decapsulated architecture built using ResNet | - | - | - | 0.91 |
| [7]  | Chest Dataset 746 images and Italian dataset | - | - | CNN based on SqueezeNet | 7.8s for a single image | 0.85 | - |
| [8]  | Real time Hospital dataset 480 patient scans | - | - | Inspired by VGG | Time consuming manual labelling process | - | 0.96 | 0.94 |
| [9]  | Real time Hospital dataset 630 patient scans | - | - | 3D deep CNN | 1.93s | 0.90 | 0.84 |
| [10] | Real time Hospital dataset 612 patient scans | Pattern location distribution | - | Decision Trees, KNN, Naive Bayes, SVM, Ensemble | - | - | 0.91 | 0.93 |
| [11] | SARS-CoV2 2482 scans | - | - | Convolutional SVM | 25min 36s | 0.96 | 0.92 |
| [12] | 126 images | GLCM, LBGLCM, GLRLM, SFTA | PCA, SAE-method | SVM | - | - | 0.94 | 0.91 |
| [13] | 339 CT images collected from various sources | - | - | One convolutional layer with 10 filters | - | - | 0.94 | - |
| [15] | Real time Hospital dataset | - | - | AlexNet, VGG-16, SqueezeNet | - | - | 0.99 (ResNet-), 0.99 (ResNet-) |
A. *Motivation*

CNNs possess excellent capability to analyze the images with very high level of semantics by learning from abstract representations. The filter bank approach adopted in a CNN aids in exploiting the texture content in the images, in contrast to the handcrafted filter bank approaches. One of the major bottlenecks associated with the studies reported in the literature (Table 1) is the availability of huge data. Since deep learning models perform better with larger datasets, the application of the above-reported methods in clinical scenario lacks reliability. Also, these methods have employed standard parameters to develop the models. The classification performance of a CNN is mainly influenced by the hyperparameters chosen and dataset [26]. Since the choice of hyperparameters is application dependent and may lead to low-performance metrics. Rather than randomly choosing the hyperparameter values, application specific values derived from an optimization methodology is adopted. In this regard, the WOA-BAT optimization algorithm is adopted, in contrast to other swarm meta heuristic optimization algorithms. Also, the WOA avoids local optima and reaches a global optimum solution without any structural reformation [28-29]. Additionally, we aim to improve the classification performance of covid-19 images, thus we created an ensemble of classifiers, and based on the test sample characteristics, a classifier is chosen in a dynamic fashion using ensemble selection techniques. This not only addresses the limited data problem without increasing the computation time and cost, but also improves the classification accuracy, which are needed in the current scenario.

B. *Novelty and Contributions*

The novelty of the methodology is the application of WOA-BAT algorithm to determine the hyperparameter values of optimized CNN architecture. This will help to obtain reduced and efficient feature set for image classification. The major contributions of the paper are given below:

1. Presented a compact deep learning-based dynamic ensemble classifier for prediction of covid-19 using chest CT images.
2. Developed an optimized CNN using WOA-BAT algorithm to extract features. These features are classified using robust ensemble of dynamic classifiers built using decision stump trees. Appropriate choice of hyperparameters determine the accuracy of the classification. In most of the CNN based classification algorithms, the hyperparameters are selected randomly, irrespective of the application, thereby yielding poor classification performance. In the proposed design, hyperparameters are selected based the features extricated from the chest CT images. To the best of our knowledge, the proposed optimized set-up is the first of its kind in the literature to date for the prediction of covid-19 using chest CT images.

2. METHODOLOGY

The proposed classification architectures are briefed as follows:

1. The first classification model deploys an ensemble of five standard CNN architectures for feature extraction and features are further selected using Binary Grey Wolf Optimizer (BGWO).

2. In the second classification model, the CNN hyperparameters are optimized through WOA-BAT algorithm, for the extraction of features.

In both the classification models, the same classifier set-up is used for determining the class of the CT images using the features extracted from the respective CNN architectures.

A. Deep Learning Ensembles for feature Extraction

The overview of the model 1 is given in Figure 2. The first step includes the extraction of features from the CT images. Transfer learning is useful when the dataset is small. In this study, five standard CNN architectures namely ResNet-50, AlexNet, VGG19, Densenet, and Inception V3 are used to extract the features.

- **AlexNet** - This architecture can classify more than 1000 different classes and built using 60 million parameters [17]. It comprises of 5 Convolutional Layers (CL) with three pooling, with two fully connected layers, and softmax layer. The first convolution layer has 96 kernels of size 11x11x3 and takes an input image of dimension 227x227x3. The two CL layers are followed by overlapping max-pooling layers. The Rectified Linear Unit (ReLu) is applied as the activation function. In this work, from the last layer, a set of 1000 features are extracted.

- **ResNet** - It is introduced to solve the vanishing gradient and degradation problem in CNN’s [18]. Depending on the number of layers, it has different versions like ResNet 18, 50, and 101.
During training, it learns from the residual features. In our work, we have used ResNet-50 architecture and generated 1000 features.

- **VGG19** - It is a 19 layer deep CNN similar to AlexNet from the visual geometry group and can classify images into 1000 classes [19]. The input dimensions are 224x224x3 and VGG19 consists of 16 CL with 3 fully connected layers. We have extracted 1000 features from the last fully connected layer (fc8).

- **DenseNet** - It is an extension of ResNet, it exploits the potential of network through feature reuse [20]. Dense nets are divided into dense blocks, wherein feature map dimensions remain constant within a block, with variations in the number of filters.

- **Inception V3** - It is a 48 layer deep CNN that takes an input image of size 299x299x3 [21]. The model consists of techniques such as factorized convolutions, regularizations, dimension reduction, and parallelized computations to obtain optimum network performance. Thus from the five standard architectures, 5000 features are extracted.

In this work, we have obtained large number of features, hence redundant features are to be eliminated. Thus, a swarm-based binary grey wolf optimization algorithm is applied to each of the individual feature vectors from each architecture, leading to 3320 feature vectors. The feature vectors are then concatenated and subjected to the feature selection process, resulting in final set of 2185 features. The procedure for feature selection is described below.
**B. Feature Selection using Binary Grey Wolf Optimizer (GWO)**

Over the recent years swarm based optimization methods have been used for several applications [22-24]. Mirjalili et al., [25] proposed a GWO optimization algorithm that mimics the hunting mechanism of wolves. The wolves are named as alpha (α), beta (β), delta (δ), and omega (ω). The initial three wolves are the fittest wolves and ω wolves are considered as subordinates. The nature of prey-hunt is given in (1-2).

\[
D = |CX_p(t) - X(t)| \quad (1)
\]
\[
X(t + 1) = X_p - AD \quad (2)
\]

where, \( t \) indicates the iteration, the distance is given as D, C and A are the coefficients computed as given in (3). The prey’s position is denoted as \( X_p \), \( X \) is the current location of wolf.

\[
A = 2ar_1 - a, C = 2r_2 \quad (3)
\]

\( r_1 \) and \( r_2 \) are assigned random values in the range [0-1], and ‘a’ varies from 2 to 0. The initial solutions are stored as alpha, beta and delta, the subordinate wolf position (ω) is iterated as given in (4-7).

\[
\overline{D_\alpha} = |C_1X_\alpha - \bar{X}| \quad (4)
\]
\[
\overline{D_\beta} = |C_2X_\beta - \bar{X}| \quad (5)
\]
\[ D_\delta = |C_3 X_\delta - \bar{X}| \]  

(6)

\[ \bar{X}(t + 1) = \bar{X}_1 + \bar{X}_2 + \bar{X}_3 / 3 \]  

(7)

Here, \( \bar{X}_1, \bar{X}_2 \) and \( \bar{X}_3 \) are approximate distances as given in [25]. Random candidate solutions are initially created and updated over the iterations considering probable positions of the prey. The value of \( \bar{A} > 1 \) and \( \bar{A} < 1 \) describe the divergence and convergence of solutions. Upon reaching the maximum iterations, optimum solution is obtained using GWO.

For the selection of the features, the binary version of GWO termed BGWO is used [25, 27]. Considering a feature vector of size \( N \), \( 2^N \), feature sets are searched exhaustively, such that the best feature combination is obtained. In this work, best feature combination refers to the feature set which gives the best classification performance. The objective function is the minimization of fitness function as given in (8).

\[ F = \alpha E_R(D) + \beta \frac{|R|}{|C|} \]  

(8)

where, \( E_R(D) \) is the classifier error rate. \( R \) is the feature subset length selected. \( C \) is the sum of all the features. The constants \( \alpha \in [0,1] \), \( \beta = 1 - \alpha \), control the accuracy of classification and feature reduction. After applying BGWO to the entire deep learning-based features, a set of 2185 features are obtained for the corresponding class.

**C. Feature extraction from CNN Optimized using Ensemble of WOA-BAT algorithm.**

The images are initially divided into a ratio of 80:20, such that each set consists of images from both the classes. The first step includes feature extraction using 80% of the image data from ResNet-50 architecture, and second step involves determining the hyperparameters of CNN using WOA-BAT optimization methodologies. As illustrated in Figure 3, a fully connected, soft-max layer, and classification layer form the last three layers of ResNet-50 architecture. The transfer learning approach is used to extract the features, by replacing last three layers of pre-trained ResNet-50 without freezing. The WOA algorithm [28], mimics the humpback whales hunting pattern. The prey is initially chased and a bubble net strategy is simulated. Prey encircling and exploitation are the two main phases in the algorithm. In the proposed work, the weights and bias of the MLP are randomly initialized. For a set of features, appropriate choice of bias and weight results in achieving good performance of the MLP. Thus MLP parameters are optimized to tune the CNN hyperparameters.

At each iteration stage, the search agents are randomly chosen, or the current best solution at the preceding iteration is retained. The hunting nature for the best solution is given as (9-10)

\[ \bar{X}(t + 1) = \bar{X}^*(t) - \bar{A} \bar{D} \]  

(9)
\[ \vec{D} = |\vec{C}\vec{X}^*(t) - \vec{X}(t)| \]  

(10)

\( \vec{X}^*(t) \) is the preceding prey position, \( \vec{X}(t + 1) \) is the present position of the whale. Similar to (3), \( \vec{D} \), \( \vec{C} \) and \( \vec{A} \) are computed as (11).

\[ \vec{C} = 2r, \vec{A} = 2ar + a \]  

(11)

The convergence speed is increased by incorporating few modifications in the nature of bats mechanism. Initially, there is 50% of chance to employ either simulating model or spiral mechanism to update the current position [28-29]. The present iteration is given in (12).

\[ \vec{X}(t + 1) = \vec{D} e^{bk} \cos(2\pi k) + \vec{X}(t) \]  

(12)

For simplicity we consider, \( b = k = 1 \), hence,

\[ \vec{X}(t + 1) = 2.7 \times \vec{D} + \vec{X}(t) \]  

(13)

If \( p < 0.5 \), then the value of the present iteration is updated as given in (14-16).

\[ f_i = f_{min} + (f_{max} - f_{min}) \]  

(14)

\[ v_{i}^{t+1} = v_{i}^t + (x_{i}^t - x_{*}) f_i \]  

(15)

\[ x_{i}^{t+1} = (x_{i}^t + v_{i}^{t+1}) \]  

(16)

Here, \( x_i \) is position of bat, and \( v_i \) is bats velocity. Lower and upper bound the algorithm corresponds to the frequency of waves. The bats position is updated using the computed velocity, because as per the law of nature when bat detects the food or prey, the rate of loudness is inversely proportional to instinct direction [30].
Fig. 3. Proposed Covid-19 detection system using hyperparameter optimization with WOA-BAT.

Fig. 4. Graphical illustration of parameter and objective space using WOA-BAT.
Based on the new velocity, the bat location is changed. The loudness rate and emission rate are inversely related. The best positions which minimize the MSE objective function are chosen as the hyperparameters of CNN. The graphical illustration of parameter space and objective space pertaining to the average MSE is depicted in Figure 4. The steps followed to compute the best values of parameters are described in Algorithm I.

**ALGORITHM I**

1. The CT images are segregated into training and test data in a ratio of 80:20
2. A set of features are extracted from the training data
3. The lower and upper bound \(\{lb, ub = -1 \times 10^{-4}, 4 \times 10^{-4}\}\), search agents=30, \(t_{\text{max}} = 30\) are chosen and the population is randomly initialized.
4. Objective function: The average MSE is computed
   \[
   \text{MSE} = \{\text{Features}\}(\text{MLP}) = \sum_{i=1}^{m} (o_i^k - d_i^k)^2
   \]
   \[
   \text{where } m=2 \text{ (Number of outputs), } d_i^k \text{ is the desired output and } o_i^k \text{ is the actual output}
   \]
5. The average MSE:
   \[
   \text{Avg}\{\text{MSE}\} = \left(\frac{\sum_{k=1}^{s} \sum_{i=1}^{m} (o_i^k - d_i^k)^2}{s}\right)
   \]
   
   Here, \(s\) is the training samples.
6. For a given set of (weights, bias) \(\rightarrow V=\{W, B\}\) the average MSE is computed.
7. Minimize the \(\text{Avg}\{\text{MSE}\}\) for the set \(V=\{W, B\}\)
8. Compute \(\tilde{A}\) using equation (11) and chose the value for \(p\)
9. If \(p \geq 0.5\), update \(X(t+1)\)
10. If \(p < 0.5\), and \(\tilde{A} \geq 1\), update \(x_i^{t+1}\), as given in (16)
11. If \(p < 0.5\), and \(\tilde{A} < 1\), update \(x_i^{t+1}\), as given in (16)
12. The above steps are repeated until the best value of MSE or a constant value of MSE or until \(t_{\text{max}} = 30\).

Our proposed CNN model is illustrated in Figure 5. At each offset, the CNN extracts the image features. Using the hyperparameter values obtained during WOA-BAT optimization, CNN is trained. The extracted features are mapped into the feature space using ReLu. The normalization of gradients and activations is performed using batch normalization. The most relevant information from the image is retained by the max-pooling layer. The image is classified using fully connected and softmax layers.
The details of filter sizes are given in Figure 5. Adam solver or stochastic gradient is used to tune the hyperparameters. However, the rate of convergence and performance of CNN, mainly rely on the appropriate choice of hyperparameters. If the learning rate is very small, the training time increases. The CNN model is updated at every epoch depending on the error, and the change is decided by the learning rate hyperparameter. Similarly, L2 regularization prevents the model from overfitting. These hyperparameters are computed by best positions obtained using WOA-BAT optimization. The hyperparameters obtained by optimization are given in Table II.

Once the values of hyperparameters are obtained, the CT images are fed to CNN to extract the features. A set of 2 features are obtained from the proposed model 2. Further, the feature sets obtained from two models are subjected to classification using AdaBoost in conjunction with dynamic ensemble selection techniques. To depict the separability between the two classes of data (covid & non-covid), the scatter plot distribution of features is shown in Figure 6. A high degree of separability can be observed in two features extracted from proposed method 2, compared to the features extracted from model 1.

Table II: Hyperparameters obtained from WOA-BAT.

| Training Solver | Initial learning Rate | L2 regularization | Max Epochs | Gradient Decay Factor | Validation Frequency |
|-----------------|-----------------------|-------------------|------------|-----------------------|----------------------|
| WOA-BAT         | 0.0004                | 0.0005            | 20         | 0.9                   | 35                   |

Fig 5. Illustration of CNN architecture.
Fig. 6. Distribution of features obtained for the two class problem: (a) Features extracted from optimized CNN (model 2), (b-d) Features obtained after applying BGWO to the features extracted using model 1.

D. AdaBoost of stump trees built from Dynamic ensemble selection techniques.

The last step in computer-aided detection of any disease is classification [31]. Usually, in an ensemble approach to test the samples, several classifiers are used, and most suitable classifier is chosen. However, in dynamic ensemble selection technique, depending on the characteristics of test sample, a different classifier is dynamically chosen. We created a dynamic ensemble of stump trees and by adopting dynamic classifier selection technique, we chose a classifier for every test pattern. For every test sample, a set of training samples are created, and the most accurate classifier is chosen. Intuitively, depending on the diversity of classifier ensemble better performance is obtained. The major advantage of this method is that the risk of overfitting and generalization is reduced.

The classifier is constructed by ensemble of AdaBoost of stump trees, and applying the dynamic classification selection techniques namely, Static Ensemble (EN), Local Class Accuracy (LCA), Overall Local Accuracy (OLA), A-Posterior (A-Po), A-Priori (A-Pr), KNORA-U(K-U), KNORA-E
(K-E), Ensemble refers to a set of stump trees are formed by boosting with a maximum tree depth chosen as 2. Initially, five neighbors of a particular test sample are selected using the training data to identify the most appropriate classifier among the classifier pool to predict the test sample. Further, the classifiers are tested on the K-NN samples from the training set and the corresponding accuracy is computed. Whereas for OLA, the K nearest training samples may not belong to the similar group, compared to OLA, the LCA, takes into account the class of K nearest samples (similar to the test sample). Further, the classifier that achieved the maximum good accuracy during classification of training samples is selected. If two classifiers have the same accuracy then K is increased, the detailed explanation can be found in [32, 37-38]. KNORA-techniques in contrast to the above mentioned dynamic ensemble selection techniques create an ensemble using the pool of classifiers to decide the test label, using the majority voting scheme, and the nearest training samples are selected in a similar fashion.

3. RESULTS AND DISCUSSION

A. Dataset

The study is carried out using a publicly available dataset [33], containing 347 COVID-19 and 397 normal (non-COVID) images associated with various kinds of pathologies. The images are in portable network graphics format.

B. Training and Testing

The proposed methodology was designed and tested in MATLAB 2020a. For model 1, the features are extracted from the entire dataset. For model 2, initially, the features are extracted using 80% of the data for obtaining the optimized hyperparameters. Once, the hyperparameter values are obtained, then the features are obtained from the entire dataset. The images are subjected to resizing to reduce the processing time. The average difference between the height and width of the images is small (approximately ±87), hence irrespective of the aspect ratio, the images were resized to 448x448x1. The optimized CNN is proposed for extracting features from the two categories (covid-19, no-covid-19). Further, five-fold cross-validation (CV) is carried for validating the proposed methodologies. The performance is computed as given in (17-21).

\[
\text{Accuracy (ACC)} = \frac{(TP + TN)}{TP + TN + FP + FN} \tag{17}
\]

\[
\text{Sensitivity (SE)} = \frac{TP}{TP + FN} \tag{18}
\]

\[
\text{Specificity (SP)} = \frac{TN}{TN + FP} \tag{19}
\]
\[ \text{Precision (PR)} = \frac{TP}{TP + FP} \]  
\[ F1 \text{ score (F-1)} = 2 \left( \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \]

Accuracy quantifies the average predictive ability into the covid and non-covid classes. Sensitivity indicates the rate of correctly identified positive instances into two classes, whereas the rate of correctly identified negative instances is quantified using specificity. Precision determines the rate of accurately identified true cases. Similarly, precision and recall are combined using F-1 score. The following subsections provide the quantitative results of the proposed methodologies.

C. Results of classification using Deep Learning based Features and Dynamic AdaBoost Classification Ensemble

The features extracted from the standard CNN architectures are divided into five distinct groups, with one set used for testing, and four sets used for training at each iteration. The Table III shows the average results of classification obtained using model 1, and the ROC corresponding to the best results using KNORA-U dynamic ensemble is given in Figure 7. The confusion matrices for each of the five folds obtained using various dynamic ensemble selection techniques with deep learning based features and dynamic AdaBoost classification ensemble is illustrated in Figure 8.

| DES Technique | SE     | SP     | ACC    | F-1    | PR     |
|---------------|--------|--------|--------|--------|--------|
| Ensemble      | 0.89±0.042 | 0.85±0.052 | 0.87±0.037 | 0.87±0.037 | 0.87±0.037 |
| OLA           | 0.78±0.024 | 0.76±0.039 | 0.77±0.030 | 0.77±0.030 | 0.77±0.030 |
| LCA           | 0.95±0.026 | 0.79±0.046 | 0.80±0.020 | 0.87±0.019 | 0.87±0.019 |
| A-Pr          | 0.83±0.024 | 0.78±0.069 | 0.80±0.035 | 0.80±0.035 | 0.80±0.035 |
| A-Po          | 0.95±0.026 | 0.79±0.046 | 0.88±0.020 | 0.87±0.019 | 0.89±0.018 |
| K-E           | 0.91±0.048 | 0.84±0.040 | 0.88±0.027 | 0.88±0.268 | 0.88±0.027 |
| K-U           | 0.89±0.042 | 0.85±0.048 | 0.87±0.032 | 0.87±0.032 | 0.87±0.032 |
Fig. 7. ROC obtained for model 1 using five-fold CV.
Fig. 8. Confusion matrices obtained for each fold for model-1 using various ensemble techniques: (a) Ensemble, (b) OLA, (c) LCA, (d) A-Pr, (e) A-Po, (f) K-E, (g) K-U [Column 1: First fold, Column 2: Second fold, Column 3: Third fold, Column 4: Fourth fold, Column 5: Fifth fold].
* N-C refers to non covid samples, and C refers to covid samples

D. Results obtained using WOA-CNN and Dynamic AdaBoost Classification Ensemble

The features extracted from the standard CNN architectures are divided into five distinct groups, with one set used for testing, and four sets used for training at each iteration. The Table IV presents the classification results in the form of mean±standard deviation using model 2, and the ROC corresponding to the best results using KNORA-U dynamic ensemble is given in Figure 9. The confusion matrices obtained for each of the five folds, for various dynamic ensemble selection techniques using WOA-CNN and dynamic AdaBoost classification ensemble is shown in Figure 10.
Table IV. Results of classification obtained using five-fold cross-validation for model 2.

| DES Technique | SE         | SP         | ACC         | F-1         | PR         |
|---------------|------------|------------|-------------|-------------|------------|
| Ensemble      | 0.96±0.038 | 0.98±0.027 | 0.96±0.038  | 0.96±0.038  | 0.96±0.038 |
| OLA           | 0.96±0.038 | 0.95±0.031 | 0.95±0.031  | 0.95±0.031  | 0.95±0.031 |
| LCA           | 0.96±0.038 | 0.96±0.038 | 0.96±0.038  | 0.96±0.038  | 0.96±0.038 |
| A-Pr          | 0.96±0.038 | 0.96±0.038 | 0.96±0.038  | 0.96±0.038  | 0.96±0.038 |
| A-Po          | 0.96±0.038 | 0.96±0.038 | 0.96±0.038  | 0.96±0.038  | 0.96±0.038 |
| K-E           | 0.96±0.038 | 0.96±0.038 | 0.96±0.038  | 0.96±0.038  | 0.96±0.038 |
| K-U           | 0.97±0.020 | 0.96±0.038 | 0.96±0.038  | 0.96±0.038  | 0.96±0.038 |

Fig. 9. ROC obtained for model 2 using five-fold CV.
Fig. 10. Confusion matrices obtained for each fold for model-2 using various ensemble techniques: (a) Ensemble, (b) OLA, (c) LCA, (d) A-Pr, (e) A-Po, (f) K-E, (g) K-U [Column 1: First fold, Column 2: Second fold, Column 3: Third fold, Column 4: Fourth fold, Column 5: Fifth fold].

A pairwise statistical test has been performed for the 21 combinations to check the statistical significance of each of the dynamic selection method with respect to the other. It was observed that the p-values for OLA and A-priori with respect to the other dynamic selection techniques were lesser.
than 0.05 (significance level) [35-36]. Hence these methods (OLA and A-priori) are statistically
different from the other methods. But, Ensemble, A-posterior, LCA, KNORA-E and KNORA-U are
found to be statistically equivalent. Similarly, for method 2, (results reported in Table IV), statistical
equivalence is observed. Since, the sensitivity of covid-19 detection using KNORA-U is higher, as
compared to the other dynamic selection techniques, the ROC for the same is given in Fig. 7 & 9 for
the two methods, respectively.

E. Comparison between the proposed optimized Classification methodology and the Unoptimized
CNN architecture

Unoptimized CNN architecture refers to the standard hyperparameter values. The comparison is
performed to provide an insight regarding the proposed classification set-up and the normal layered
CNN with standard values of hyperparameters. The Figure 11, provides the graphical illustration of
performance of optimized and unoptimized CNNs. It can be observed from Figure 11 that,
classification performance using proposed optimized CNN outperformed the standard CNN
performance using all the evaluation metrics.

![Graphical illustration of performance parameters obtained using optimized and unoptimized CNNs.](image)

A student’s t-test was also performed to compare the means of the two population. The first population
consists of performance parameters obtained using unoptimized CNN and the second population
consists of performance parameters obtained using optimized CNN. The t-test assumes normal
distribution of population for testing the hypothesis which states that, “Optimized CNN improves the
accuracy of covid-19 detection”. The significance value considered was $\alpha = 0.05$. The statistical test
resulted in with a p-value of 0.0012, which is lesser than 0.05. Hence the hypothesis is accepted,
which proves that optimized CNN significantly performs better in contrast to unoptimized CNN.
F. Comparative Analysis

Summary of comparison with literature based approaches developed for covid-19 detection obtained from the same dataset is shown in Table V. He et al [5], used 25% of the images for testing. The self-trans approach was based on ResNet-50 and Densenet architecture with Adam optimizer default hyperparameter values, resulting in classification accuracy of 86%. Mobiny et al [6], proposed the detailed oriented capsule network derived from ResNet with three residual blocks with Adam optimizer default hyperparameter values, 20% data was used for testing. In [7], the values for the learning rate, momentum and L2 regularization were tuned experimentally, depending on the performance of SqueezeNet model. Cruz et al [35], proposed an ensemble of deep learning methods, 27% of the image data was used for testing the model, and data augmentation was also performed to enlarge the dataset using rotation and vertical flips. Inspite of the data augmentation and robust ensemble, an accuracy of 86.7% was obtained, with an inference time of 14s. The methods reported in the comparative analysis were based on the same dataset, with variation in train and test ratio with standard values for hyperparameters. The comparative analysis performed is fair, since in the proposed work a five-fold cross validation is performed, which indicates that each sample in the dataset has been a part of the test set in either of the five iterations, and moreover the average performance of the five iterations is reported. It can be observed from the comparative study that, the proposed design comparatively performed better. It can be inferred from Table V, that the method proposed overcomes the approaches reported in literature in terms of all the evaluation parameters.

Table V. Summary of comparison with state-of-the-art approaches developed for automated prediction of Covid-19 using CT images obtained from the same dataset.

| Ref.        | No. of images | Methodology                                | SE  | SP  | ACC    | F-1    | PR   |
|-------------|---------------|--------------------------------------------|-----|-----|--------|--------|------|
| He et al [5]| 746           | Self-trans approach                        | -   | -   | 0.86   | 0.85   | -    |
| Mobiny et al [6]| 746          | Detailed oriented capsule network           | -   | -   | 0.84   | -      | -    |
| Polsinelli et al [7]| 746        | CNN based on SqueezeNet                     | 0.87| 0.87| 0.85   | 0.86   | 0.85 |
| Cruz et al [35]| 746          | Transfer learning ensemble                  | 0.83| 0.89| 0.86   | 0.85   | -    |
| Proposed    | 746           | Optimized CNN + Dynamic Ensemble selection technique | 0.97±0.020| 0.96±0.038| 0.96±0.038| 0.96±0.038| 0.96±0.038 |
4. DISCUSSION AND REMARKS

The major objective of the proposed study is the classification of covid-19 and non-covid-19 classes using chest CT images with higher classification performance. To avoid problems pertaining to overfitting and underfitting, a computer-aided diagnostic experts need to address three major issues: (i) sufficient data is required to prevent underfitting, (ii) achieving a balance between sensitivity and specificity, and (iii) hyperparameters used to use build the CNN architecture should be carefully chosen to avoid overfitting. The proposed study aims to address the aforementioned limitations. At present, the number of publicly available covid-19 CT images is limited. Hence, to improve the classification performance, a robust set of AdaBoost of stump trees are built using dynamic selection techniques. Additionally, to avoid overfitting the training parameters are chosen until the best values obtained from a WOA-BAT optimizer with MSE of the MLP as objective function. To prove the robustness and reliability of the proposed design, the features extracted from the standard CNN architectures and the features extracted from the optimized CNN architectures are compared. As it can be seen from Figure 6, the separability between the two classes of the data is quite high using WOA-BAT optimized CNN as compared to the features extracted from the standard CNN architectures (AlexNet, ResNet, VGG-19, Densenet, and Inception). Also, the classification performance of covid-19 images improved to 98% using the proposed method compared to the standard CNN architecture. The salient features of the proposed study are as follows:

1. Hyperparameters of the classifier play a major role in classification. Thus, an appropriate choice of hyperparameters is crucial in improving the classifier performance.
2. The features extracted from the optimized CNN, although few in number (2), exhibit a great degree of separability between the two classes.
3. AdaBoost of ensemble classifiers built using dynamic ensemble selection techniques produced considerably better performance in the classifying covid-19 and non-covid-19 CT images.

The proposed system was implemented on a 64-bit CPU based system with 8GB RAM, and processor speed of 2.3GHz. Hence the processing time required to test a single image was 30s, which is a limitation of the proposed system. However, a higher end GPU based system can reduce the processing time.
5. CONCLUSION

In this study, an optimized CNN-based architecture is proposed for detection of covid-19 cases from non-covid CT images accurately. The hyperparameters of CNN are application dependent and play a major role in the performance of the classifier. Hence, we have incorporated an ensemble of WOA-BAT technique to optimize CNN parameters. The optimized CNN is used for feature extraction. In this work, we have observed that, in contrast to the features extracted from five standard CNN architectures (2185), the features extracted from the optimized CNN (2 features), showed better discriminative ability. Although a BGWO, technique is used to eliminate the redundant features, from the original 5000 features obtained from the well-known CNN designs, the two features extracted from optimized CNN outperformed, and significantly improved the classification accuracy. We have obtained the highest classification accuracy of 96% for predicting covid-19. Further, the optimized CNN is also compared with performance obtained from the unoptimized CNN built using standard values of hyperparameters. Our proposed method yielded better performance with standard CNN architecture irrespective of the limited dataset. The summary of comparison with the state-of-the-art techniques developed for automated detection of covid-19 using chest CT images obtained from the same dataset given in Table V proves the superiority of the proposed method [5-6, 34]. Our developed prototype can be used as a pre-screening software for covid-19, especially in remote villages with limited access to health monitoring facilities. The proposed diagnosis method also eliminates the need for the CNN model to be trained using large datasets. The developed system can also be used for the automated detection of various pathologies using medical images of followed by the optimization methodology, however with a variation in the hyperparameter values based on the type of dataset.

In the future, we intend to evaluate the robustness of our developed design using more chest covid-19 CT images collected from local hospitals and polyclinics. This generated economical user-friendly system can be accessed by clinicians to obtain a second opinion about their manual diagnosis. Also, we intend to extend this work with more classes like pneumonia in addition to covid-19 and normal classes.

REFERENCES

1. Alizadehsani, R., Alizadeh Sani, Z., Behjati, M., Roshanzamir, Z., Hussain, S., Abedini, N., ... & Islam, S. M. S. (2021). Risk factors prediction, clinical outcomes, and mortality in COVID-19 patients. Journal of medical virology, 93(4), 2307-2320.
2. Kwee, T. C., & Kwee, R. M. (2020). Chest CT in COVID-19: what the radiologist needs to know. Radiographics, 40(7), 1848-1865.

3. Silva, P., Luz, E., Silva, G., Moreira, G., Silva, R., Lucio, D., & Menotti, D. (2020). COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis. Informatics in Medicine Unlocked, 20, 100427.

4. He, X., Yang, X., Zhang, S., Zhao, J., Zhang, Y., Xing, E., & Xie, P. (2020). Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans. medRxiv.

5. Mobiny, A., Cicalese, P. A., Zare, S., Yuan, P., Abavisani, M., Wu, C. C., … & Van Nguyen, H. (2020). Radiologist-Level COVID-19 Detection Using CT Scans with Detail-Oriented Capsule Networks. arXiv preprint arXiv:2004.07407.

6. Polsinelli, M., Cinque, L., & Placidi, G. (2020). A Light CNN for detecting COVID-19 from CT scans of the chest. arXiv preprint arXiv:2004.12837.

7. Hu, S., Gao, Y., Niu, Z., Jiang, Y., Li, L., Xiao, X., … & Ye, H. (2020). Weakly supervised deep learning for covid-19 infection detection and classification from ct images. IEEE Access, 8, 118869-118883.

8. Wang, X., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., … & Zheng, C. (2020). A Weakly-supervised Framework for COVID-19 Classification and Lesion Localization from Chest CT. IEEE Transactions on Medical Imaging.

9. Ardakani, A. A., Acharya, U. R., Habibollahi, S., & Mohammadi, A. (2021). COVIDiag: A clinical CAD system to diagnose COVID-19 pneumonia based on CT findings. European radiology, 31(1), 121-130.

10. Özkaya, U., Öztürk, Ş., Budak, S., Melgani, F., & Polat, K. (2020). Classification of COVID-19 in Chest CT Images using Convolutional Support Vector Machines. arXiv preprint arXiv:2011.05746.

11. Öztürk, Ş., Özkaya, U., & Barstuğan, M. (2020). Classification of Coronavirus (COVID-19) from X-ray and CT images using shrunken features. International Journal of Imaging Systems and Technology.

12. Maghdid, H. S., Asaad, A. T., Ghafoor, K. Z., Sadiq, A. S., & Khan, M. K. (2020). Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms. arXiv preprint arXiv:2004.00038.

13. Shoeibi, A., Khodatars, M., Alizadehsani, R., Ghassemi, N., Jafari, M., Moridian, P., … & Srinivasan, D. (2020). Automated detection and forecasting of covid-19 using deep learning techniques: A review. arXiv preprint arXiv:2007.10785.

14. Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., & Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. Computers in Biology and Medicine, 121, 103795.

15. Pathan, S., Siddalingaswamy, P. C., & Ali, T. (2021). Automated Detection of Covid-19 from Chest X-ray scans using an optimized CNN architecture. Applied Soft Computing, 104, 107238.

16. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

17. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

18. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

19. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).
20. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90.
21. Eberhart, R., & Kennedy, J. (1995, October). A new optimizer using particle swarm theory. In MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science (pp. 39-43). IEEE.
22. Dorigo, M., Birattari, M., & Stutzle, T. (2006). Ant colony optimization. IEEE computational intelligence magazine, 1(4), 28-39.
23. Kirkpatrick, S. (1984). Optimization by simulated annealing: Quantitative studies. Journal of statistical physics, 34(5-6), 975-986.
24. Mirjalili, S. (2015). How effective is the Grey Wolf optimizer in training multi-layer perceptrons. Applied Intelligence, 43(1), 150-161.
25. Julie, E. G., Robinson, Y. H., & Jaisakthi, S. M. (Eds.). (2021). Handbook of Deep Learning in Biomedical Engineering and Health Informatics. CRC Press.
26. Emary, E., Zawbaa, H. M., & Hassanien, A. E. (2016). Binary grey wolf optimization approaches for feature selection. Neurocomputing, 172, 371-381.
27. Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. Advances in engineering software, 95, 51-67.
28. Mohammed, H. M., Umar, S. U., & Rashid, T. A. (2019). A systematic and meta-analysis survey of whale optimization algorithm. Computational intelligence and neuroscience, 2019.
29. Yang, X. S. (2011). Bat algorithm for multi-objective optimisation. International Journal of Bio-Inspired Computation, 3(5), 267-274.
30. Britto Jr, A. S., Sabourin, R., & Oliveira, L. E. (2014). Dynamic selection of classifiers—a comprehensive review. Pattern recognition, 47(11), 3665-3680.
31. Cruz, R. M., Sabourin, R., & Cavalcanti, G. D. (2014, October). Analyzing dynamic ensemble selection techniques using dissimilarity analysis. In IAPR Workshop on Artificial Neural Networks in Pattern Recognition (pp. 59-70). Springer, Cham.
32. Jinyu Zhao, Yichen Zhang, Xuehai He, and Pengtao Xie, “Covid-ct-dataset: A ct scan dataset about covid19,” arXiv preprint arXiv:2003.13865, 2020.
33. Amyar, A., Modzelewski, R., & Ruan, S. (2020). Multi-task Deep Learning Based CT Imaging Analysis For COVID-19: Classification and Segmentation. medRxiv.
34. Santa Cruz, J. F. H. (2021). An ensemble approach for multi-stage transfer learning models for COVID-19 detection from chest CT scans. Intelligence-Based Medicine, 5, 100027.
35. García, S., Fernández, A., Luengo, J., & Herrera, F. (2010). Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. Information sciences, 180(10), 2044-2064.
36. García, S., & Herrera, F. (2008). An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons. Journal of machine learning research, 9(12).
37. Cruz, R. M., Sabourin, R., & Cavalcanti, G. D. (2018). Dynamic classifier selection: Recent advances and perspectives. Information Fusion, 41, 195-216.
38. Ko AH, Sabourin R, Britto Jr AS. From dynamic classifier selection to dynamic ensemble selection. Pattern recognition. 2008 May 1;41(5):1718-31.
HIGHLIGHTS

- A novel ensemble architecture built using CNN and dynamic selection techniques is proposed for detection of covid-19.
- The CNN hyperparameters are optimized using WOA-BAT optimization techniques.
- Limited data problem is solved by using AdaBoost of decision stump trees.
- A classification accuracy of 96% is obtained in detection of covid-19 using Chest CT scans.
Conflict of interests

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.