

WE-BA: Covid-19 detection by Wavelet Entropy and Bat Algorithm

Wangyang Yu¹, Yanrong Pei², Shuihua Wang^{1,3}, Yudong Zhang^{1,4,*},

¹School of Computing and Mathematical Sciences, University of Leicester, Leicester, LE1 7RH, UK

²Huai'an Tongji Hospital, Huai'an, Jiangsu 223000, China

³Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, China

⁴Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Abstract

INTRODUCTION: Covid-19 is a kind of fast-spreading pneumonia and has dramatically impacted human life and the economy.

OBJECTIVES: As early diagnosis is the most effective method to treat patients and block virus transmission, an accurate, automatic, and effective diagnosis method is needed.

METHODS: Our research proposes a machine learning model (WE-BA) using wavelet entropy for feature extraction to reduce the excessive features, one-layer FNNs for classification, 10-fold cross-validation (CV) to reuse the data for the relatively small dataset, and bat algorithm (BA) as a training algorithm.

RESULTS: The experiment eventually achieved excellent performance with an average sensitivity of $75.27\% \pm 3.25\%$, an average specificity of $75.88\% \pm 1.89\%$, an average precision of $75.75\% \pm 1.06\%$, an average accuracy of $75.57\% \pm 1.21\%$, an average F1 score of $75.47\% \pm 1.64\%$, an average Matthews correlation coefficient of $51.20\% \pm 2.42\%$, and an average Fowlkes–Mallows index of $75.49\% \pm 1.64\%$.

CONCLUSION: The experiments showed that the proposed WE-BA method yielded superior performance to the state-of-the-art methods. The results also proved the potential of the proposed method for the CT image classification task of Covid-19 on a small dataset.

Keywords: Covid-19 diagnosis, wavelet entropy, bat algorithm, feedforward neural network, K-fold cross-validation

Received on 28 April 2022, accepted on 05 August 2022, published on 17 October 2023

Copyright © 2023 W. Yu *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [Creative Commons Attribution license](#), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.9.711

1. Introduction

Covid-19 is a kind of pneumonia caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The World Health Organisation (WHO) defined the COVID-19 outbreak as a public health emergency of international concern (PHEIC) in January 2020. Covid-19 has

dramatically impacted human life and caused significant damage to the economy. Over 100 million people are facing a return to extreme poverty because of Covid-19 [1, 2].

COVID-19 can cause cough, fever, taste loss, smell, dry cough, and other symptoms [3] and may develop into severe disease and lead to rapid death. Some infected people are asymptomatic but can also be a source of infection. At the same time, early diagnosis is the most

*Corresponding authors. Email: yudongzhang@ieeee.org

effective method to treat patients and block virus transmission [4], so it is necessary to have an accurate and effective diagnosis method. The two most widely used detection methods are Polymerase Chain Reaction (PCR) and CT screening. However, the PCR test is time-consuming and has a high ratio of false negatives. The CT scanning method has higher sensitivity, which can assess the level of the disease [5].

However, manual diagnosing CT scanning needs a lot of human resources, and subjective factors may affect the results. Therefore, it is essential to develop an accurate computer-aided diagnosis system.

Machine learning technology has been a popular field in recent years. It has made significant progress and has been widely used in many areas, such as computer vision [6-9] and real-time systems [10-13]. At present, many researchers have used machine learning methods for the diagnosis of CT images, some of which perform well and surpass human radiologists in many indicators. Ahuja, S. *et al.* (2021) [14] proposed a method based on deep learning. It used the pre-trained architectures ResNet18, with 30% data as a test set and the rest as a validation set, achieving 92.21% sensitivity and 98.50% specificity.

Yildiz, A. *et al.* (2009) [15] proposed a new feature extraction method that combined discrete wavelet transform (DWT) with Shannon entropy. The obtained signal will be decomposed to its sub-bands by DWT, and then calculate the Shannon entropy of each sub-bands. Utilizing wavelet entropy can reduce the dimension of the feature, which can help the model get better generalization performance. Yao, X. *et al.* (2021) [16] proposed a model for COVID-19 CT image classification by using a wavelet entropy (WE) and biogeography-based optimization (BBO) based method (WE-BBO). They found that the combination of optimisation algorithm and Wavelet Entropy can improve performance. Based on the WE-BBO, Wang, W. *et al.* (2022) [17] proposed a method combining wavelet entropy and the self-adaptive Jaya algorithm and getting 85.47% sensitivity, 87.23% precision, and 86.23% accuracy. However, the performance of these models is not good enough to assist the clinical diagnosis. Because of the high fake positive rate, applying these models in the diagnosis would cause an unnecessary bad impact on the patient's quality of life.

This paper's contributions are: 1. We discovered a wavelet entropy and Bat Algorithm combined method (WE-BA) for COVID-19 diagnosing; 2. We explored the possibility of training a model to classify COVID-19 on small datasets; 3. We further demonstrated the potential of combining wavelet entropy and optimization algorithms in COVID-19 diagnosis.

In this paper, the second part introduces the information about the dataset. The third part introduces the methods used in this paper. The fourth part presents the experiment and discusses the results. The fifth part is the conclusion.

2. Dataset

The database comprises CT images taken by the Fourth People's Hospital of Huai'an, China. The database contains 132 subjects, aged from 24 to 91, of whom 77 are men and 55 are women. The database includes two parts: the observation group and the control group. Each group contains 148 selected CT slices. The observation group included 66 COVID-19-infected subjects, including 39 men and 27 women, with an average age of 48. In the control group, we randomly selected 66 subjects from 159 healthy people, including 38 men and 28 women, with an average age of 38.44. In Figure 1, we show a set of photos in the database.

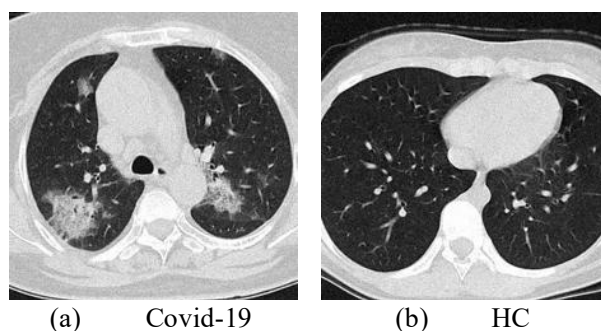


Figure 1. Samples in the database

3. Methodology

3.1 Wavelet Entropy

The Fourier transform is a standard signal processing method that converts the signal from the time domain to the frequency domain. The Fourier transform has been widely used in many fields, and the formula of the transform is shown in Formula (1).

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt, \quad (1)$$

where t refers to the time, ω refers to the frequency, F refers to the function in the frequency domain, and f refers to the function in the time domain.

Although the Fourier transform can obtain the spatial frequency information of the image, it also has disadvantages. In the actual signal in nature, it is often not stationary. For non-stationary signals, the performance of the Fourier transform is not very good. The Fourier transform can only determine which frequencies are contained in the signal but not when a specific frequency component exists. It will lose time characteristics, making it possible for two non-stationary signals that are different in the time domain to appear the same in the frequency domain. Therefore, the Fourier transform cannot solve the medical image signal processing problem.

The short-time Fourier transform (STFT) improves the problem of poor time resolution. It decomposes the whole time-domain signal into an infinite number of short-time signals. This process of signal decomposition is called adding windows. Then, it is assumed that the signal is smooth in the short-term interval, and each smooth part will respectively use the Fourier transform. In this way, we can get the frequency and time information simultaneously. On this basis, we can further do time-frequency analysis and get a spectrogram of the signal. However, STFT also has its shortcomings. The window's width is fixed and will not change in the transformation process. According to the uncertainty principle, as shown in Formula (2), there is no way to determine the frequency at a certain moment, but we can know the frequency at a certain period.

$$\Delta t \Delta f \geq \frac{\pi}{4}, \quad (2)$$

where Δt refers to the interval of the time, and Δf refers to the frequency at the time interval.

Therefore, a too wide window will lead to low time resolution, and a too-narrow window will lead to low frequency resolution. Therefore, there is no way to have good resolution for high-frequency and low-frequency signals simultaneously. To achieve good resolution of both high-frequency and low-frequency signals, the window's width needs to be variable.

Wavelet transform uses finite decaying wavelet to replace the infinitely long trigonometric basis of the Fourier transform, so it can get the time of a certain frequency when obtaining the frequency. The formula of wavelet transform is shown in Formula (3).

$$W(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-\tau}{a}\right) dt, \quad (3)$$

where a refers to the scale, τ refers to the translation, t refers to the time, ψ refers to the parent wavelet function, and W refers to the function of wavelet transformation.

Wavelet transform realizes scaling and translation (changing frequency and time) by changing a and τ [18]. The large a value applies to low frequency, and small a value is applicable to high frequency. In wavelet transform, there are multiple resolutions according to different frequencies. Therefore, the wavelet transform has a good resolution in both high and low frequencies, which solves the problem of unsatisfactory non-stationary signal processing in the Fourier transform.

Although wavelet transform contains multiple resolutions and retains rich information, it contains a lot of redundant information, which requires a lot of time and space [19]. We can use entropy to reduce the dimension of features. Entropy represents the uncertainty of image texture features—the more uniform the image distribution, the lower the entropy [20]. Wavelet entropy is a new tool for analyzing the instantaneous characteristics of non-stationary signals. Wavelet entropy combines wavelet decomposition and entropy. The original Shannon entropy can be used as a standard to compare and quantify the

energy distribution in wavelet transform. The definition of Shannon entropy is shown in Formula (4).

$$S(\vec{g}) = -\sum_i P(g_i) \log_b P(g_i), \quad (4)$$

where \vec{g} refers to the grey levels, P refers to the probability of grey levels, and S refers to the energy distribution in wavelet subbands according to the probabilities of grey levels occurring.

3.2 Feedforward Neural Network

Feedforward neural network (FNN) is a classical deep learning model that has been mature in application and development. Structurally, FNN consists of an input layer, an output layer, and one or several hidden layers between the two layers.

In this paper, a one-hidden-layer FNN is adopted, and its structural diagram is shown in Figure 2. FNN has a strong expression ability. FNN can approximate all functions at any degree with only one hidden layer according to the universal approximation theory [21]. The extracted feature of the image will be sent to the input layer, so the number of neurons in the input layer is equal to the number of features.

The grid-searching method can determine the number of neurons in the hidden layer. The number of neurons in the output layer equals the number of classifications. The classification result can be obtained using the argmax function in the output layers. We use FNN because the shallow architecture can effectively prevent the overfitting problem when training with a small dataset. The training method is backpropagation. After initializing the weights, the cost function will be calculated, and then the cost function will be backpropagated to update the weights. The cost function will gradually become smaller and get our final model by repeating this process. This article uses the mean squared error (MSE) as the cost function.

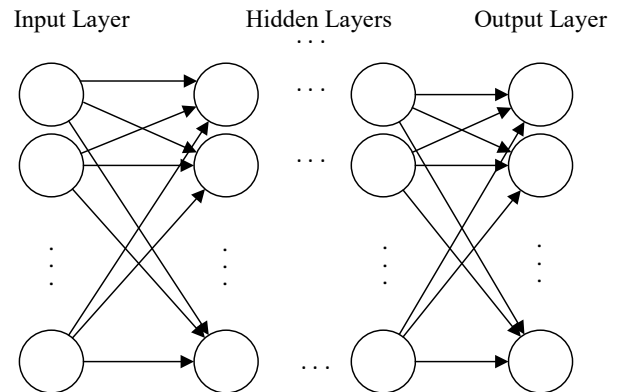


Figure 2. Structural diagram of FNN

3.3 Bat Algorithm

As the weights and biases of the model may not be optimal, we use the Bat Algorithm to optimize them. It is a metaheuristic algorithm based on the echolocation behavior of bats proposed by Yang, X.-S. (2010) [22]. The BA performs better than genetic algorithms [23] and particle swarm optimization (PSO) [24]. For simplicity, it sets several idealized rules:

1. All bats sense distance by echolocation, and they can distinguish food from the barriers.
2. Bats randomly fly at location \vec{x}_i with a speed of \vec{v}_i , and search for prey using a fixed frequency f_{min} , wavelength λ , and loudness A_0 . They can adapt the wavelength of their emitted pulses and adjust the rate of pulse emission r while approaching the target.
3. We assume that the loudness varies from the maximum value A_0 to the minimum A_{min} .
4. The range of frequency is from f_{min} to f_{max} .

In simulations, each bat's frequency is randomly assigned by

$$f_i = f_{min} + (f_{max} - f_{min})\beta, \quad (5)$$

where $\beta \in [0,1]$ is a random variable, f_i refers to the pulse frequency of a specific bat. We update the solutions and velocities at time step t as

$$\vec{v}_i^{t+1} = \vec{v}_i^t + (\vec{x}_i^t - \vec{x}_*)f_i, \quad (6)$$

$$\vec{x}_i^{t+1} = \vec{x}_i^t + \vec{v}_i^{t+1}, \quad (7)$$

where \vec{v}_i^t refers to the speed of a specific bat at time step t , \vec{x}_i^t refers to the solution of a specific bat at time step t and \vec{x}_* is the current global best solution. For the local search, once a solution is chosen from the current best solutions, a new solution for each bat is generated locally using the random walk.

$$\vec{x}_{new} = \vec{x}_{old} + \varepsilon A^t, \quad (8)$$

where \vec{x}_{old} is selected among the best solutions, and $\varepsilon \in [-1,1]$ is a random variable. If the new solution is accepted, the A_i will reduce and r_i will increase according to

$$A_i^{t+1} = \alpha A_i^t, \quad (9)$$

$$r_i^{t+1} = r_i^0(1 - e^{-\gamma t}), \quad (10)$$

where γ, α are constants, $0 < \alpha < 1, 0 < \gamma$. r_i^0 is the initial emission rate and $A^t = \langle A_i^t \rangle$ is the mean loudness of all the bats at this time step. When $t \rightarrow \infty$, we have $A_i^t \rightarrow 0, r_i^t \rightarrow r_i^0$. The flow chart shown in Figure 3 illustrates the basic steps of BA.

3.4 K-fold Cross-Validation

The dataset in machine learning is usually divided into a training set used for training the model and a test set used for testing the model's performance. Machine learning is a data-driven science. The size of the dataset can greatly affect the performance of the model. Dividing the dataset into a training set and a test set can greatly reduce the data available for training the model.

Cross-Validation reuses the data and increases the data used for training, which can further improve the model's performance. CV can also effectively avoid overfitting. The cross-validation uses each sample data as both training data and test data, and the overlearning and underlearning can be avoided, so the results obtained are more convincing than just using hold-out validation. K-fold CV is a widely used CV method. It divides the dataset into pre-specified K groups and takes one group of data as the test set without repeating and the other data as the training set to train the model. The training will repeat K times and ensure each data group is test set once. After calculating the performance of each model, we can get the final performance. In this paper, we use the most commonly used 10-fold CV.

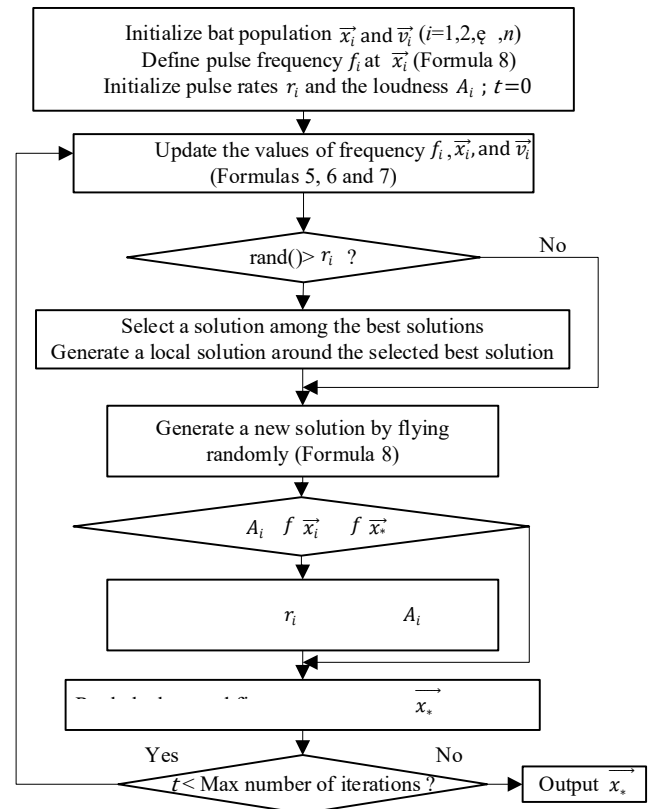


Figure 3. Flow chart of bat algorithm, where $f(\)$ is the objective function, and $\text{rand}(\)$ is a function generating random variable $\in [0,1]$.

4. Experimental Results and Discussions

4.1 WE Results

Figure 4 shows the wavelet decomposition of a sample image at four levels. In wavelet transform, the low-frequency signal will be further decomposed, and the high-frequency part of signal details will be reserved so the wavelet transform can characterize the groups with low-frequency information as the main components of the signal.

Figure 4(a) is the first level wavelet decomposition image, which contains both high-frequency sub-bands (lower right corner, lower left corner, and upper right corner) and low-frequency (upper left corner) sub-band of the input image. Figure 4(b) is the second level wavelet decomposition, which reserved the high-frequency sub-bands and did the same transformation to the low-frequency sub-band of the first level result. Figure 4(c) and (d) are the third and fourth-level wavelet decomposition obtained by a similar recursive operation. The sample image is painted with pseudo-color for a clear view and is represented in greyscale.

After 4-level wavelet decomposition, we get thirteen sub-bands. We can calculate the entropy of each sub-band to form the feature vector according to Formula(4). Thus, we extract the feature and reduce the feature number to only thirteen. Reducing feature numbers can efficiently solve the overfitting problem caused by a small training set.

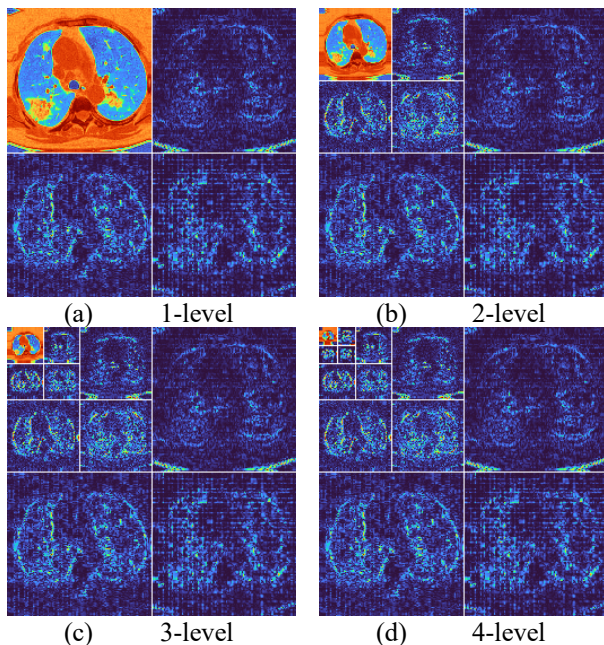


Figure 4. Wavelet decomposition of a sample image

4.2 Statistical Results

The experiment used the wavelet entropy to extract the feature, the one-hidden-layer feedforward neural network as the classifier, the 10-fold cross-validation to validate the model's performance and the bat algorithm as the training algorithm. The experiment eventually achieved great performance (shown in Table 1) with an average sensitivity of $75.27\% \pm 3.25\%$, an average specificity of $75.88\% \pm 1.89\%$, an average precision of $75.75\% \pm 1.06\%$, an average accuracy of $75.57\% \pm 1.21\%$, an average F1 score of $75.47\% \pm 1.64\%$, an average Matthews correlation coefficient of $51.20\% \pm 2.42\%$, and an average Fowlkes–Mallows index of $75.49\% \pm 1.64\%$.

Table 1. Results of 10 runs of 10-fold cross-validation.

Run	Sen	Spc	Prc	Acc	F1	MCC	FMI
1	78.38	72.97	74.36	75.68	76.32	51.43	76.34
2	73.65	77.03	76.22	75.34	74.91	50.70	74.93
3	75.68	77.03	76.71	76.35	76.19	52.71	76.19
4	73.65	77.70	76.76	75.68	75.17	51.39	75.19
5	72.97	77.03	76.06	75.00	74.48	50.04	74.50
6	72.30	77.03	75.89	74.66	74.05	49.38	74.07
7	70.27	75.68	74.29	72.97	72.22	46.01	72.25
8	79.73	72.30	74.21	76.01	76.87	52.17	76.92
9	76.35	77.03	76.87	76.69	76.61	53.38	76.61
10	79.73	75.00	76.13	77.36	77.89	54.79	77.91
MSD	75.27	75.88	75.75	75.57	75.47	51.20	75.49
	± 3.25	± 1.89	± 1.06	± 1.21	± 1.64	± 2.42	± 1.64

(Sen = Sensitivity; Spc = Specificity; Prc = Precision; Acc = Accuracy; F1 = F1 Score; MCC = Matthews correlation coefficient; FMI = Fowlkes–Mallows index)

Figure 5 shows the ROC curve (Receiver Operating Characteristic curve) and the AUC (Area Under Curve) value of the WE-BA model. Each point on the ROC curve corresponds to a classifier with different thresholds.

When the threshold reaches the maximum, TPR (True Positive Rate, known as sensitivity) = FPR (False Positive Rate, which equals $1 - \text{specificity}$) = 0, which corresponds to the origin of the coordinate system. When the threshold is the minimum, TPR = FPR = 1, corresponding to (1,1) in the coordinate system. The TPR and FPR will decrease as the threshold increases. The AUC value is the size of the area under the ROC curve. The AUC value of WE-BA is 0.8361.

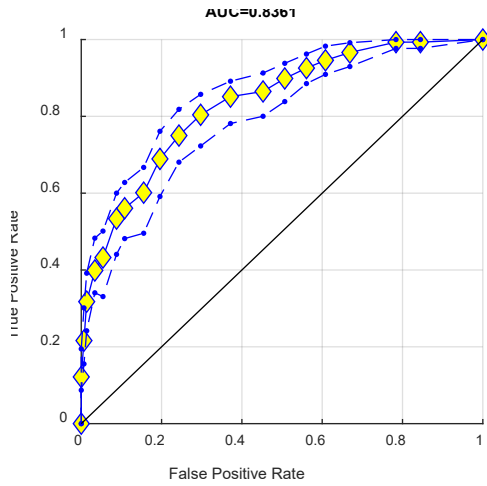


Figure 5. ROC curve and the AUC value of WE-BA

4.3 Comparison to State-of-the-art Approaches

Compared to other state-of-the-art models classifying COVID-19 by CT images, WE-BA greatly improves in many aspects. The numerical comparison between WE-BA and other approaches is shown in Table 2.

Figure 6 shows the overview of performance comparison to other algorithms. Except for specificity and precision, our model achieved the best performance in all the other metrics. The results also proved the potential of the proposed method for the CT image classification task of Covid-19 on a small dataset.

Table 2. Performance comparison to other algorithms.

Approach	Sen	Spc	Prc	Acc	F1	MCC	FMI
RBFNN [25]	66.89 ±2.43	75.47 ±2.53	73.23 ±1.48	71.18 ±0.80	69.88 ±1.08	42.56 ±1.61	69.97 ±1.04
K-ELM [26]	58.78 ±3.36	60.27 ±3.89	59.73 ±1.68	59.53 ±1.42	59.19 ±1.86	19.09 ±2.83	59.22 ±1.79
WE+BBO [27]	72.84 ±3.00	75.00 ±1.99	74.47 ±1.20	73.92 ±1.18	73.61 ±1.57	47.89 ±2.34	73.63 ±1.56
GLCM-SVM [28]	72.03 ±2.94	78.04 ±1.72	76.66 ±1.07	75.03 ±1.12	74.24 ±1.57	50.20 ±2.17	74.29 ±1.53
WE-BA (Ours)	75.27 ±3.25	75.88 ±1.89	75.75 ±1.06	75.57 ±1.21	75.47 ±1.64	51.20 ±2.42	75.49 ±1.64

(Sen = Sensitivity; Spc = Specificity; Prc = Precision; Acc = Accuracy; F1 = F1 Score; MCC = Matthews correlation coefficient; FMI = Fowlkes–Mallows index)

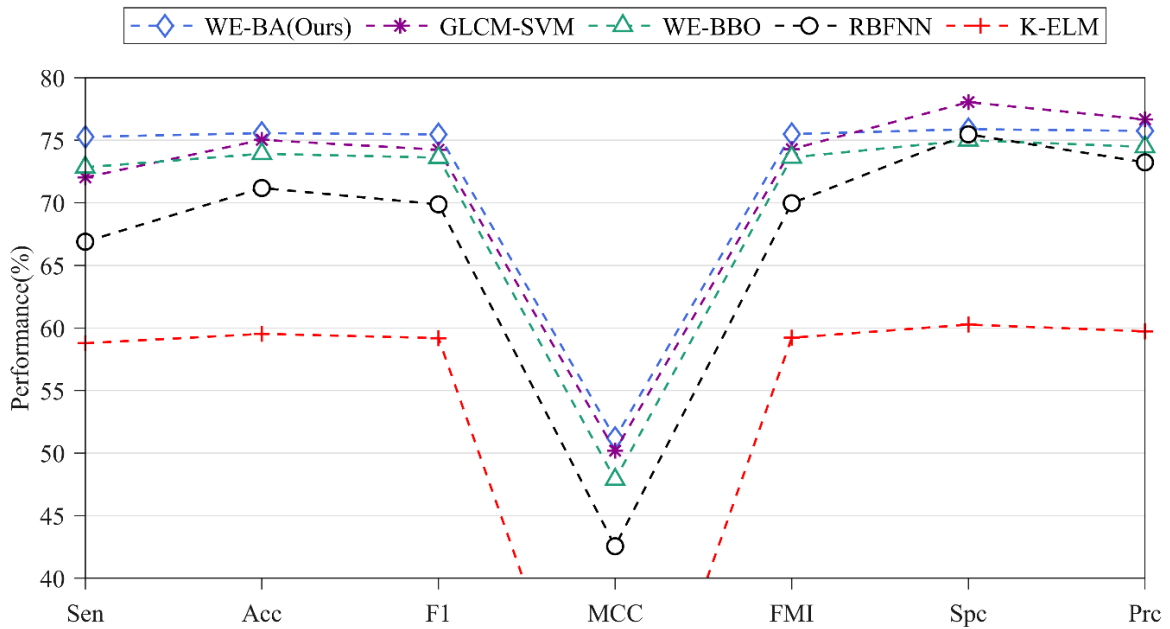


Figure 6. Model performance overview

5. Conclusions

To develop an accurate, efficient, and automatic COVID-19 classifier of CT chest slices, we used wavelet entropy for feature extraction to reduce the excessive features. We introduced the 10-fold cross-validation to reuse the data for the relatively small dataset. Afterwards, we proposed to use the one-hidden-layer FNN classifier and BA as the training algorithm. The experiments showed that the proposed WE-BA method yielded superior performance to the state-of-the-art methods.

The results also proved the potential of the proposed method for the CT image classification task of Covid-19 on a small dataset. At the same time, the specificity and precision of our model are lower than that of the SOTA algorithms. We believe that using more optimization and improvement methods for the wavelet entropy feature can achieve better performance in the future.

Acknowledgment

This paper is partially supported by MRC (MC_PC_17171); Royal Society (RP20G0230); BHF (AA/18/3/34220); Hope Foundation for Cancer Research (RM60G0680); GCRF (P202PF11); Sino-UK Industrial Fund (RP20G0289); LIAS (P202ED10, P202RE969); Data Science Enhancement Fund (P202RE237); Fight for Sight (24NN201); Sino-UK Education Fund (OP202006); BBSRC (RM32G0178B8).

References

- Hotez, P.J., et al., *The new covid-19 poor and the neglected tropical diseases resurgence*. Infectious Diseases of Poverty, 2021. **10**(1): p. 10
- Saqib, M., et al., *Covid-19 detection from radiographs: Is deep learning able to handle the crisis?* Signals, 2022. **3**(2): p. 296-312
- Struyf, T., et al., *Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has covid - 19*. Cochrane Database of Systematic Reviews, 2021(2)
- Yu, X., et al., *Resgnet-c: A graph convolutional neural network for detection of covid-19*. Neurocomputing, 2021. **452**: p. 592-605
- Fang, Y., et al., *Sensitivity of chest ct for covid-19: Comparison to rt-pcr*. Radiology, 2020. **296**(2): p. E115-E117
- Liu, S., et al., *Human inertial thinking strategy: A novel fuzzy reasoning mechanism for iot-assisted visual monitoring*. IEEE Internet of Things Journal, 2022: p. 1-1
- Liu, S., et al., *Fuzzy detection aided real-time and robust visual tracking under complex environments*. IEEE Transactions on Fuzzy Systems, 2021. **29**(1): p. 90-102
- Shi, J., et al. *Top-down driver head orientation detection method based on monocular camera and virtual scene generation*. in *International Forum on Digital TV and Wireless Multimedia Communications*. 2022. Springer. p. 349-360
- Yao, C., et al. *Claw u-net: A unet variant network with deep feature concatenation for scleral blood vessel segmentation*. in *CAAI International Conference on Artificial Intelligence*. 2021. Springer. p. 67-78
- Liu, S., et al., *Human memory update strategy: A multi-layer template update mechanism for remote visual monitoring*. IEEE Transactions on Multimedia, 2021. **23**: p. 2188-2198
- Guan, H., et al., *Encoding multiple information into european article number (ean) - 13 codes with various chromic sensing materials to construct intelligent barcodes*. Advanced Intelligent Systems, 2022. **4**(1): p. 2100113
- Tang, J., et al. *Special cane with visual odometry for real-time indoor navigation of blind people*. in *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*. 2020. IEEE. p. 255-255
- Tang, J., et al., *Design and optimization of an assistive cane with visual odometry for blind people to detect obstacles with hollow section*. IEEE Sensors Journal, 2021. **21**(21): p. 24759-24770
- Ahuja, S., et al., *Deep transfer learning-based automated detection of covid-19 from lung ct scan slices*. Applied Intelligence, 2021. **51**(1): p. 571-585
- Yildiz, A., et al., *Application of adaptive neuro-fuzzy inference system for vigilance level estimation by using wavelet-entropy feature extraction*. Expert Systems with Applications, 2009. **36**(4): p. 7390-7399
- Yao, X., et al., *Covid-19 detection via wavelet entropy and biogeography-based optimization*, in *Covid-19: Prediction, decision-making, and its impacts*, K.C. Santosh and A. Joshi, Editors. 2021, Springer Singapore: Singapore. p. 69-76.
- Wang, W., et al., *Covid-19 diagnosis by we-saj*. Systems Science & Control Engineering, 2022: p. 1-11
- Saritha, M., et al., *Classification of mri brain images using combined wavelet entropy based spider web plots and probabilistic neural network*. Pattern Recognition Letters, 2013. **34**(16): p. 2151-2156
- El-Dahshan, E.-S.A., et al., *Hybrid intelligent techniques for mri brain images classification*. Digital Signal Processing, 2010. **20**(2): p. 433-441
- Zhang, Y., et al., *Multivariate approach for alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization*. J Alzheimers Dis, 2018. **65**(3): p. 855-869
- Nguyen, H.D., et al., *A universal approximation theorem for mixture-of-experts models*. Neural Comput, 2016. **28**(12): p. 2585-2593
- Yang, X.-S., *A new metaheuristic bat-inspired algorithm*, in *Nature inspired cooperative strategies for optimization (nisco 2010)*, J.R. González, et al., Editors. 2010, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 65-74.
- Srinivas, M., et al., *Genetic algorithms: A survey*. Computer, 1994. **27**(6): p. 17-26
- Zhang, Y., et al., *A comprehensive survey on particle swarm optimization algorithm and its applications*. Mathematical Problems in Engineering, 2015. **2015**: p. 931256
- Lu, Z., *A pathological brain detection system based on radial basis function neural network*. Journal of Medical Imaging and Health Informatics, 2016. **6**(5): p. 1218-1222

26. Yang, J., *A pathological brain detection system based on kernel based elm*. *Multimedia Tools and Applications*, 2018. **77**(3): p. 3715-3728
27. Yao, X., *Covid-19 detection via wavelet entropy and biogeography-based optimization*, in *Covid-19: Prediction, decision-making, and its impacts*, K.C. Santosh and A. Joshi, Editors. 2020, Springer. p. 69-76.
28. Chen, Y., *Covid-19 classification based on gray-level co-occurrence matrix and support vector machine*, in *Covid-19: Prediction, decision-making, and its impacts*, K.C. Santosh and A. Joshi, Editors. 2020, Springer Singapore: Singapore. p. 47-55.