Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 45 (2024), No. 2, pp. 235–245
DOI : 10.47974/JIOS-1543

Non-iterative fast and efficient HOG-KELM hybrid model for COVID-19 diagnosis using Chest X-ray images

Nikhat Ali *
Virendra P. Vishwakarma †
University School of Information, Communication and Technology
Guru Gobind Singh Indraprastha University
Dwarka 110078
New Delhi
India

Abstract

COVID-19 is one of the most widely spread infectious disease, which was declared as pandemic in year 2019 by WHO. The viral infection can be diagnosed using RT-PCR test and through radiological scans, which are time consuming and not 100% accurate. Researchers are trying to develop more reliable, accurate and faster diagnostic system with the help of advance machine learning techniques. Some of the advance models developed using deep learning techniques are efficient but dependent on larger dataset and high-end processing system. In this paper, unique features extracted from COVID-19 chest X-ray images are used for diagnosis with the help of Histogram of Oriented Gradients-Kernel based Extreme Learning Machine (HOG-KELM) model. The proposed model is fast, efficient and economic, which does not rely on heavy machinery or processing unit. The experimental analysis is performed on a dataset of 15,153 CX-ray images, in which 3616 are COVID-19 infected, 10,192 normal lung, and 1345 pneumonia infected images. This model produces 99.78% accuracy for binary classification which evoke the efficacy of the proposed model compared to the existing state-of-art models.

Subject Classification: *62H30, 68Q85, 68T10, 68T45, 68U07, 68U10, 92C55.*

Keywords: *Histogram of Oriented Gradients (HOG), Extreme learning machine, COVID-19, X-ray, Non-iterative, KELM.*

* *E-mail:* **tonikhatali@gmail.com (Corresponding Author)**
† *E-mail:* **virendravishwa@rediffmail.com and vpv@ipu.ac.in**

1. Introduction

The coronavirus infection caused by SAR-CoV-2 was first reported in the year end of 2019. The coronavirus disease is highly infectious and impacts respiratory system as well as complicates pre-existing health conditions of patients. The disease infection pattern and its symptoms were unknown in the beginning and before it could have been contained, it already marked its presence in entire world. Due to its wide spread throughout the world WHO declared it a pandemic in year 2020. The Healthcare system was not prepared for anything like this pandemic, it was highly infectious which caused lots of hospitalization. Lack of reliable and fast test kits, diagnostic tools and proper treatment lots of people lost their lives. Coronavirus outbreak not only impacted people's health but also their life and livelihood [1]. Some of the diagnostic tools available are Rapid antigen test and RT-PCR test. RT-PCR test takes almost 24 hours for producing any results, these test especially rapid antigen tests are non-reliable, time consuming and not 100% accurate. Imaging techniques like ultrasound, X-rays and CT-scans were used to detect any kind of infection in the lungs. The major task which radiologists faced during this COVID-pandemic is to correctly differentiate these lung infection from any other type of acute respiratory conditions which can be caused by any other pathogens. Recent trends in the field of Artificial intelligence in various sectors have influenced researchers to dive into development of automated AI based healthcare systems which can prove to be helpful in identifying various medical conditions. The dataset of COVID-19 infected chest scans can be used for developing above said advance diagnostic system, which are fast and less prone to errors.

The challenge faced by researchers is not only detection accuracy but also availability of open source database and clinical data. So far researchers have developed models based on deep learning and state-of-art convolutional neural network for analysis and diagnostic purpose [2, 3, 4, and 5]. In order to overcome these challenges many researchers adopted transfer learning based pre-trained model. Recent studies shows that these models have significantly improved overall performance of the designed system [6, 7, and 8]. Gayathri J.L. *et al*. [9] used ensemble of Xception and InceptionResnetV2 for feature extraction and a FFNN for classification. Bhattacharyya *et al*. [10] first segmented the lung scans using C-GAN and then used deep learning algorithm form feature extraction and classification. Nayak *et al*. [11] utilized and improvised a CNN model for COVID-19 classification. Kaya and Eris [12] developed a hybrid feature extraction network namely D3SENET which consists of DarkNet53, DarkNet19, DenseNet201, SqueezeNet, and EfficientNetb0. Godbin & Jasmine [13] designed model established on GLCM features and traditional machine learning classifiers which showed convincingly better performance compared to some deep learning model. Similarly Shankar *et al*. [14] applied deep learning and fusion based feature extraction technique for detection of COVID-19. ResNet50 and AlexNet convolutional neural network (CNN) model along with tunable Q-wavelet transform was proposed by Kumari *et al*. [15] for COVID-19 and non-COVID-19 classification. Ensemble bootstrap aggregation with neural network was implemented by Alshahrni *et al*. [16] where dataset was analyzed and clustered with the help of SPSS software. Ali *et al*. [17] developed COVID-19 classification model based on DWT and SVM. Tallapragada *et al*. [18] also explored DWT along with deep learning convolutional models like AgileNet (LeNet & AlexNet) and inception-ResNet (GoogLeNet & ResNet).

The main contribution of this paper is listed as follows:

- An advanced machine learning algorithm along with statistical feature extraction technique using non-iterative HOG-KELM is applied for developing COVID-19 diagnosis system.
- The aim is to perform statistical analysis on X-ray images and differentiate between Covid-19 disease and other conditions. The feature extracted using histogram oriented gradient technique along with kernel based extreme learning machine provide enhanced diagnosis.
- The performance evaluation made on current system and compared with some of the latest state-of-art deep learning models.
- The experiment result shows that proposed model out-performs most of the latest high end deep learning models and is more efficient in terms of accuracy and processing time.

The paper is organized as follows. In section 2, proposed model architecture is demonstrated, followed by section 3 in which necessary preliminaries required are explained. Section 4 lay out experimental outcomes of the proposed model. Finally, in section 5 conclusions and future work is discussed.

2. Proposed Methodology

The classification model developed is based on single layer feed forward network. In order to enhance system performance and reduce computational complexity Kernel based extreme learning machine for classification of COVID-19 chest images with non-COVID-19 images is adopted. The proposed model consists of three phases: Preprocessing of raw images, extraction of relevant features using HOG feature extraction, and lastly the extracted features are used for training non-iterative KELM based model for further analysis as shown in Figure 1. The detailed explanations of each phase is carried out in following subsections.

3. Preliminaries

3.1 *Mathematical analysis of KELM*

G.B. Huang [19, 20] developed a simple and faster single layer neural network called Extreme learning machine. The input weights are randomly chosen and output weights determined using Moore Penrose generalized inverse method.

The output function given by following equation (1),

$$f(x) = h(x)\beta \tag{1}$$

$$\beta = H^{\dagger}T \tag{2}$$

where, for N distinct input samples (x_i, t_i), there are L hidden nodes with output $h(x) = [h_1(x), \ldots, h_L(x)]$, output weight $\beta = [\beta_1, \beta_2, \ldots, \beta_L]$, and

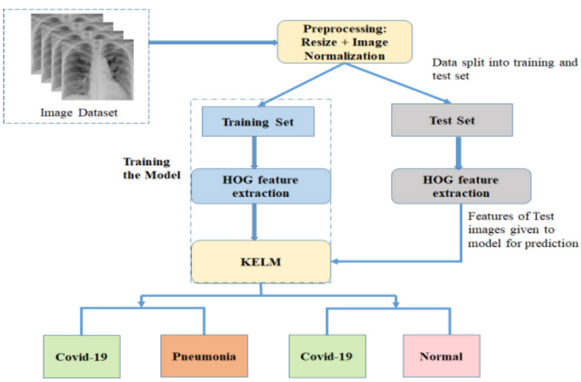


Figure 1

Proposed binary HOG-KELM classification model block diagram

$H^{\dagger}$ is Moore Penrose inverse of h(x). The hidden layer feature maps input space into L dimensional feature space. Since, the hidden function in ELM is unknown and most of the non-linear piecewise continuous function can be used as hidden output function [21], thus different extensions of ELM variants can be implemented. One such extension is Kernel based ELM, in which Mercer's condition is applied. The hidden output function and number of hidden nodes are unknown, with the given kernel matrix (equation (3)):

$$\Omega_{ELM} = HH^T : \Omega_{ELM_{i,j}} = h(x_i)h(x_j) = k(x_i, x_j) \tag{3}$$

The output function of KELM classifier can be resolved into following equation:

$$f(x) = h(x)H^T \left(\frac{I}{C} + HH^T\right)^{-1} T \tag{4}$$

$$f(x) = \begin{bmatrix} K(x,x_1) \\ \vdots \\ K(x,x_N) \end{bmatrix}^T \left(\frac{I}{C} + \Omega_{ELM}\right)^{-1} T \tag{5}$$

3.2 *Histogram of Oriented Gradient (HOG)*

HOG feature descriptor is mostly used for object detection. It is based on existence of gradient orientation in image regions. In this method an image window is formed to get desired image segment or image region. The image segment is obtained, then gradient magnitude and angle matrices are calculated. With the help of these two matrices cells are formed, which is used for finding gradient histogram. These cells are then grouped to form blocks, from each block feature vector is derived. Feature vector of all the blocks are concatenated to get final HOG feature vector [22]. For the feature extraction, cell size [88] is used, it contains more descriptive information, but bigger cell size also increases dimensionality of feature vector [23].

3.3 *Dataset*

The chest X-ray images of COVID-19 and other conditions required for training of the proposed model is acquired from 'Kaggle' website [24]. The metadata consists of total 21165 CX-ray images of lung, which include 10192: Normal, 3616: COVID-19 infected, 1345: Viral pneumonia, and remaining 6012: Lung opacity. As the proposed model is investigated for binary classification only, the dataset is divided into two set using combination of COVID, pneumonia and normal CX-ray images, the dataset distribution can be viewed in Table 1.

Table 1

Dataset of Chest X-ray

COVID-19	Normal	Viral Pneumonia	Lung Opacity	Dataset used for the experiment
3,616	10,192	1,345	6,012	**Set-1**: [COVID-19: 3616, Normal: 10192] **Set-2**: [COVID-19 : 3616, Pneumonia: 1345]

4. Experimental results and Discussion

The effectiveness of proposed model is validated by performing experimental analysis on two subsets of above mentioned dataset. One set consists of COVID-19 & Normal (set-1) chest X-ray images, whereas other set have COVID-19 & Viral pneumonia (set-2) images. The results are shown in Table 2 and 3. For two-way classification the performance of proposed model is better than some of the deep learning model as shown in Table 4. The HOG-KELM based model achieves 99.79% accuracy and 100% sensitivity.

Table 2

COVID Vs NORMAL

Parameters	60% training	70% training	80% training	90% training
Accuracy	94.858	96.283	96.849	97.176
Precision	94.9	95.772	95.954	96.275
Sensitivity	84.924	89.77	91.84	92.818
Specificity	98.381	98.594	98.626	98.724
F-score	89.635	92.674	93.852	94.515
Percentage Test Error	5.1421	3.7171	3.151	2.824

Table 3

COVID Vs Pneumonia

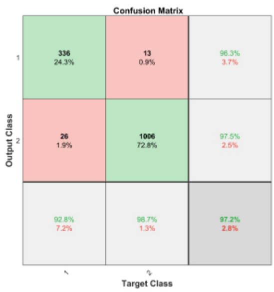
Parameters	60% training	70% training	80% training	90% training
Accuracy	99.345	99.194	99.294	99.798
Precision	99.312	99.175	99.448	99.725
Sensitivity	99.793	99.724	99.585	100
Specificity	98.141	97.772	98.513	99.254
F-score	99.552	99.449	99.516	99.862
Percentage Test Error	0.65524	0.80591	0.70565	0.20161

Table 4

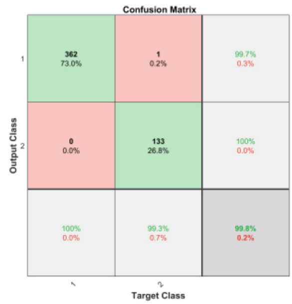
Comparative Analysis

Authors, year and References	Methodology/ Technique	Accuracy
Gayathri J.L. *et al*. [2022], [9]	CNN with sparse autoencoder and FFNN	95.78%
Soumya Ranjan Nayak *et al*. [2022], [11]	LightWeight CNN model called LW-CORONet	Dataset 1: 99%, Dataset 2: 96.25%
Ibrahim Al-Shourbaji *et al*. [2023], [25]	CNN model based on VGG	99.27%
Kumari Jyoti *et al*. [2023], [15]	MCA inspired TQWT model	Dataset 1: 98.82% Dataset 2: 94.64%
Sourabh Singh Verma *et al*. [2022], [26]	VGG16 with SVM as classifier	95%
K. Shankar *et al*. [2022], [14]	Feature Fusion: GLCM, LBP, and Salp Swarm algorithm(SSA) with ANN	95.10%
Proposed Approach	**Non-iterative HOG-KELM**	Dataset 1: **97.176**% Dataset 2: **99.79**%

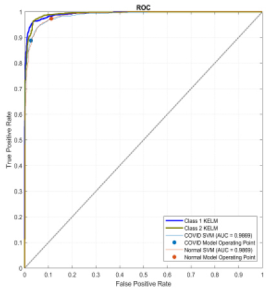
The confusion matrix is employed for performance visualization of the model, Figure 2 shows confusion matrix for the experiments performed using two set of dataset. The ROC graph is plotted for comparing efficacy



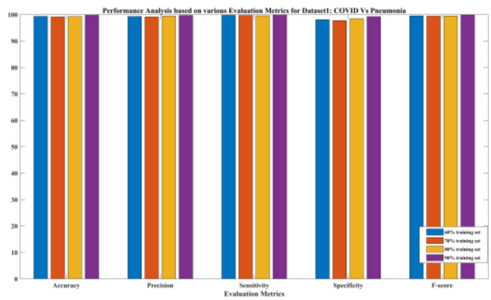
(a) COVID Vs Normal



(b) COVID Vs Pneumonia



(c) ROC



(d) Performance plot

Figure 2

Confusion matrices for binary classification HOG-KELM model: (a)-(b), (c) ROC of KELM classifier, and (d) Performance plot of the model

of KELM over SVM classifier, from the graph it is clear that KELM performs better than SVM and also require less time for classification.

5. Conclusion

The work aimed to introduce hybrid classification model which can be used for COVID-19 detection using chest X-ray images. The proposed HOG-KELM model shows efficient performance (accuracy 99.79%) based on more than one evaluation metrics. The model is cost effective and does not require heavy GPU based infrastructure for its implementation. The processing time is significantly less than the pre-trained DNN models. The only constraint faced in this work is acquiring dataset of COVID-19 chest scans. The available X-ray images of COVID-19 which can be used for implementation of such model are limited in public domain and highly imbalanced, also they are not first choice of recommendation for early COVID-19 diagnosis. Low dose CT imaging is better choice for early diagnosis and severity analysis of infection. For the future work, HOG-KELM model needs to be investigated for imbalanced multi-classification.

References

[1] World Health Organization (WHO). Health Topics. "Impact of COVID-19 on people's livelihoods, their health and our food systems". Available online: https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems.

[2] Tsuneki, Masayuki. "Deep learning models in medical image analysis." *Journal of Oral Biosciences* 64, no. 3 : 312-320 (2022).

[3] Celik, Gaffari. "Detection of Covid-19 and other pneumonia cases from CT and X-ray chest images using deep learning based on feature reuse residual block and depthwise dilated convolutions neural network." *Applied Soft Computing* 133 : 109906 (2023).

[4] Al-Shourbaji, Ibrahim, Pramod H. Kachare, Laith Abualigah, Mohammed E. Abdelhag, Bushra Elnaim, Ahmed M. Anter, and Amir H. Gandomi. "A deep batch normalized convolution approach for improving COVID-19 detection from chest X-ray images." *Pathogens* 12, no. 1 : 17 (2022).

[5] Ali, Nikhat, and Jyotsna Yadav. "Computer-Aided Detection and Diagnosis of Lung Nodules Using CT Scan Images: An Analytical Review." In Proceedings of Second Doctoral Symposium on Computational Intelligence*: DoSCI 2021*, pp. 545-558. Singapore: Springer Singapore (2021).

[6] Pathak, Yadunath, Prashant Kumar Shukla, Akhilesh Tiwari, Shalini Stalin, and Saurabh Singh. "Deep transfer learning based classification model for COVID-19 disease." *Irbm* 43, no. 2 : 87-92 (2022).

[7] Khan, Muhammad Attique, Tallha Akram, Yu-Dong Zhang, and Muhammad Sharif. "Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework." *Pattern Recognition Letters* 143 : 58-66 (2021).

[8] Salehi, Ahmad Waleed, Preety Baglat, and Gaurav Gupta. "Review on machine and deep learning models for the detection and prediction of Coronavirus." Materials Today: Proceedings 33 : 3896-3901 (2020).

[9] Gayathri, J. L., Bejoy Abraham, M. S. Sujarani, and Madhu S. Nair. "A computer-aided diagnosis system for the classification of COVID-19 and non-COVID-19 pneumonia on chest X-ray images by integrating CNN with sparse autoencoder and feed forward neural network." *Computers in Biology and Medicine* 141 : 105134 (2022).

[10] Bhattacharyya, Abhijit, Divyanshu Bhaik, Sunil Kumar, Prayas Thakur, Rahul Sharma, and Ram Bilas Pachori. "A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images." *Biomedical Signal Processing and Control* 71 : 103182 (2022).

[11] Nayak, Soumya Ranjan, Deepak Ranjan Nayak, Utkarsh Sinha, Vaibhav Arora, and Ram Bilas Pachori. "An efficient deep learning method for detection of COVID-19 infection using chest X-ray images." *Diagnostics* 13, no. 1 : 131 (2022).

[12] Kaya, Mustafa, and Mustafa Eris. "D3SENet: A hybrid deep feature extraction network for Covid-19 classification using chest X-ray images." *Biomedical Signal Processing and Control* 82 : 104559 (2023).

[13] Godbin, A. Beena, and S. Graceline Jasmine. "Screening of COVID-19 based on GLCM features from CT images using machine learning classifiers." *SN Computer Science* 4, no. 2 : 133 (2022).

[14] Shankar, K., Eswaran Perumal, Prayag Tiwari, Mohammad Shorfuzzaman, and Deepak Gupta. "Deep learning and evolutionary intelligence with fusion-based feature extraction for detection of COVID-19 from chest X-ray images." *Multimedia Systems* 28, no. 4 : 1175-1187 (2022).

[15] Jyoti, Kumari, Sai Sushma, Saurabh Yadav, Pawan Kumar, Ram Bilas Pachori, and Shaibal Mukherjee. "Automatic diagnosis of COVID-19 with MCA-inspired TQWT-based classification of chest X-ray images." *Computers in Biology and Medicine* 152 : 106331 (2023).

[16] Alshahrni, Mohammad M., Mostafa A. Ahmad, Monir Abdullah, Nadir Omer, and Muzzamil Aziz. "An intelligent deep convolutional network based COVID-19 detection from chest X-rays." *Alexandria Engineering Journal* 64 : 399-417 (2023).

[17] Ali, Nikhat, and Jyotsna Yadav. "Coronavirus disease identification using Multi-subband feature analysis in DWT domain." *Procedia Computer Science* 218 : 574-584 (2023).

[18] Tallapragada, VV Satyanarayana, N. Alivelu Manga, and GV Pradeep Kumar. "A novel COVID diagnosis and feature extraction based on discrete wavelet model and classification using X-ray and CT images." *Multimedia Tools and Applications* : 1-42 (2023).

[19] Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme learning machine: a new learning scheme of feedforward neural networks." In 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541), vol. 2, pp. 985-990. Ieee (2004).

[20] Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme learning machine: theory and applications." *Neurocomputing* 70, no. 1-3 : 489-501 (2006).

[21] Huang, Guang-Bin, Hongming Zhou, Xiaojian Ding, and Rui Zhang. "Extreme learning machine for regression and multiclass classification." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 42, no. 2 : 513-529 (2011).

[22] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol. 1, pp. 886-893. Ieee (2005).

[23] Freeman, William T., and Michal Roth. "Orientation histograms for hand gesture recognition." In International workshop on automatic face and gesture recognition, vol. 12, pp. 296-301 (1995).

[24] Rahman, Tawsifur, Amith Khandakar, Yazan Qiblawey, Anas Tahir, Serkan Kiranyaz, Saad Bin Abul Kashem, Mohammad Tariqul Islam et al. "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images." *Computers in biology and medicine* 132 : 104319 (2021). Available online: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

[25] Al-Shourbaji, Ibrahim, Pramod H. Kachare, Laith Abualigah, Mohammed E. Abdelhag, Bushra Elnaim, Ahmed M. Anter, and Amir H. Gandomi. "A deep batch normalized convolution approach for improving COVID-19 detection from chest X-ray images." *Pathogens* 12, no. 1 : 17 (2022).

[26] Verma, Sourabh Singh, Ajay Prasad, and Anil Kumar. "CovXmlc: High performance COVID-19 detection on X-ray images using Multi-Model classification." *Biomedical signal processing and control* 71 : 103272 (2022).