



RESEARCH ARTICLE

A COVID Net-predictor: A multi-head CNN and LSTM-based deep learning framework for COVID-19 diagnosis

Y. Mohammed Iqbal*, M. Mohamed Surputheen, S. Peerbasha

Abstract

COVID-19 pandemic alerts the necessity of preparing alternate respiratory health detective measures that improve time, expense, and prediction performance. Prevention of COVID-19 spread depends on early identification and precise diagnosis. Since the commonly used real-time Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) swab test is laborious and unreliable, radiography images are still advised for chest screening. Unfortunately, complexities in early detection using traditional approaches urge innovative research in this field. Intending to introduce a novel COVID-19 prediction scheme, this paper employed a COVIDNet Predictor. This model is built with various stages including preprocessing, segmentation, feature extraction, selection and fusion, prediction and monitoring. Initially, the input images are preprocessed to enhance image quality and noise reduction. A U-net segmentation is carried out to find the Region of Interest (ROI). Color, shape and textual features are extracted and are further optimally chosen by a hybrid optimizer EvoNSGA II. Besides, the optimal features are fused through a Hierarchical Attention Network (HAN) and given as input to the COVIDNet-Predictor. The proposed COVIDNet-Predictor is a combination of Multi-Head Convolutional Neural Network (MHCNN), and Long-Short-Term Memory (LSTM) architectures. Additionally, a monitoring and feedback loop is added to make the model fit the real-time applications based on patient data. The efficacy of the proposed COVIDNet -Predictor is evaluated via a comparison with SOTA models and proved its competence by attaining 95.04% accuracy.

Keywords: COVID-19 prediction, Deep learning, Convolutional neural network, Long-short-term memory, Attention mechanism, Hybrid optimizer.

Introduction

COVID has continued to have a significant worldwide impact more than 2 years ago resulting in numerous respiratory illnesses and a high number of fatalities. Because of COVID-19's high contagiousness, detection is still of utmost importance, Dash, T. K., Chakraborty, C., Mahapatra, S., & Panda, G. (2022). To limit the virus's transmission and lessen

its effects on the world's health, COVID-19 early diagnosis is crucial. To limit future transmission within communities, Health Authorities implement critical measures including isolation, contact tracing, and quarantine as soon as cases are identified. Essentially, early detection plays a critical role in managing global health during this unprecedented crisis by reducing the burden the pandemic has on healthcare systems, economies, and communities throughout the world. It also helps to contain the pandemic's spread, Dai, A., Harrou, F., & Sun, Y. (2021). Currently, using throat and nasopharyngeal swab samples for nucleic acid testing is the most practicable diagnostic approach. However, problems like low virus loads and sample mistakes might affect accurate diagnosis. Conversely, antigen tests have a lesser sensitivity but a quicker turnaround time. Chest X-ray (CXR), Zhang, X., Han, L., Sobeih, T., Han, L., Dempsey, N., Lechareas, S., ... & Zhang, D. (2022). and Computed Tomography (CT), Liu, J., Yu, H., & Zhang, S. (2020). imaging are essential for identifying contagions in patients, as well as pathological testing. To increase the precision of diagnosis, several DL models have been put out, particularly for differentiating between different types of infections using CT and CXR images. In order to treat and lessen the effects

PG & Research Department of Computer Science, Jamal Mohamed College (Autonomous), (Affiliated to Bharathidasan University), Tiruchirappalli, Tamil Nadu, India -620 020.

***Corresponding Author:** Y. Mohammed Iqbal, PG & Research Department of Computer Science, Jamal Mohamed College (Autonomous), (Affiliated to Bharathidasan University), Tiruchirappalli, Tamil Nadu, India -620 020, E-Mail: mohammedit2k15@gmail.com

How to cite this article: Iqbal, Y.M., Surputheen, M.M., Peerbasha, S. (2025). A COVID Net-predictor: A multi-head CNN and LSTM-based deep learning framework for COVID-19 diagnosis. The Scientific Temper, 16(3):3984-3995.

Doi: 10.58414/SCIENTIFICTEMPER.2025.16.3.16

Source of support: Nil

Conflict of interest: None.

of chronic diseases, early identification, accurate prognosis, and effective intervention are essential, Pradhan, M., Shah, K., Alexander, A., Ajazuddin, Minz, S., Singh, M. R., Singh, D., Yadav, K., & Chauhan, N. S. (2022).

Using DL techniques for COVID-19 detection has become a viable way to maximize the accuracy, speed, and ease of use of disease diagnosis procedures. Traditional diagnostic techniques, like RT-PCR tests, became bottlenecks with the rapid global spread of COVID-19 because of their limited availability, lengthy processing times, and need for specialized equipment, Ravikumar, K. K., Ishaque, M., Panigrahi, B. S., & Pattnaik, C. R. (2023). DL, a branch of Artificial Intelligence (AI), provides a potent substitute in this situation by using enormous databases of medical images for detecting COVID-19. Recently, medical imaging has undergone a revolution using DL, specifically, CNNs, which automate the feature extraction as well as classification process, Aslan, M. F., Unlarsen, M. F., Sabanci, K., & Durdu, A. (2021). Traditional Machine Learning (ML) techniques call for manual feature extraction but, DL models learn hierarchical data directly from raw images. This allows them to recognize complex patterns that may be challenging for human experts to discern. To stop the virus from spreading and make sure that patients receive treatment on time, early and accurate detection of COVID-19 is essential, Ji, T., Liu, Z., Wang, G., Guo, X., Lai, C., Chen, H., ... & Zhou, Q. (2020). Even though they are accurate, traditional diagnostic techniques like RT-PCR testing take a lot of time and may not always be available, particularly in environments with limited resources. Imaging methods like CT scans and CXR provide an additional diagnostic approach for COVID-19. These imaging modalities can show consolidations and ground-glass opacities, two lung abnormalities linked to COVID-19, Rai, P., Kumar, B. K., Deekshit, V. K., Karunasagar, I., & Karunasagar, I. (2021).

To detect COVID-19, DL models have been used to analyze CT and CXR images. The main objective is to create models that can discriminate between healthy lungs and other lung diseases like pneumonia in addition to COVID-19. Researchers and clinicians created several AI systems that aid in the quick and accurate diagnosis of COVID-19 from medical imaging by utilizing CNNs, Reshi, A. A., Rustam, F., Mehmood, A., Alhossan, A., Alrabiah, Z., Ahmad, A., Alsuwailam, H., & Choi, G. S. (2021)., Transfer Learning (TL), Perumal, V., Narayanan, V., & Rajasekar, S. J. S. (2021)., ensemble techniques, Dey, S., Bhattacharya, R., Malakar, S., Mirjalili, S., & Sarkar, R. (2021)., attention mechanisms, Ullah, Z., Usman, M., Latif, S., & Gwak, J. (2023)., and segmentation models, Voulodimos, A., Protopapadakis, E., Katsamenis, I., Doulamis, A., & Doulamis, N. (2021, June). Notwithstanding these difficulties, DL research and developments keep expanding the potential of these approaches, which makes them an invaluable addition to conventional diagnostic

methods, Sharma, N., Saba, L., Khanna, N. N., Kalra, M. K., Fouda, M. M., & Suri, J. S. (2022). With the advancements and application of DL algorithms, this paper presents a novel COVID-19 prediction framework. The key contributions are listed here.

- To develop a COVIDNet-Predictor to efficiently predict COVID-19-prone patients based on CXR and CT images.
- To present an effective module to resolve misclassification issues in COVID-19 prediction and a continuous monitoring and feedback module to provide a robust and reliable network.
- To develop a U-Net segmentation module to find the ROI such as lungs, surrounding features and pathological features.
- To introduce a novel hybrid optimizer EvoNSGA II to pick the best features and fusion module to combine the features through a HAN.

A rigorous comparison analysis to validate the efficacy of the proposed COVIDNet-Predictor with baseline models. This article is structured as a literature study on the current research regarding COVID-19 detection in Section II. Section III explains the implemented DL network. Experimentation and results are given in Section IV. Section V ends the research.

Literature Study

Recent research

The applications of acoustics feature in COVID-10 detection using Logistic Regression (LR), Support Vector Machine (SVM) and LSTM. In addition, symptoms like sneezing, coughing, and breathing data were utilized. In order to do this, a web application was used to gather data on speech signals as well as a record of health symptoms over 20 months, Chetupalli, S. R., Krishnan, P., Sharma, N., Muguli, A., Kumar, R., Nanda, V., Pinto, L. M., Ghosh, P. K., & Ganapathy, S. (2023).

A CNN- Categorical Boosting (CNN-CatBoost) to classify the test samples to produce predictions with the best predictive performance. Also, a low-cost sensing module with a variety of sensors was developed to measure the amounts of various particles in exhaled breath for detecting COVID-19, Bhaskar, N., Bairagi, V., Munot, M. V., Gaikwad, K. M., & Jadhav, S. T. (2023).

A two-stage classification system to characterize the changes in speech-producing organs like lungs, and larynx. Moreover, the suggested two-stage system utilized novel modulation spectral data and a linear forecasting approach. The investigation results revealed that the 2-stage prediction model attained better performance, Xue, Y., Zhu, H., & Neri, F. (2023).

A Hierarchical Spectrogram Transformer (HST) using spectrogram depictions of respiratory resonances. HST embodied the spectrograms' self-attention mechanisms, and to capture context from local to global, window size was

gradually increased over model stages, Aytekin, I., Dalmaz, O., Gonc, K., Ankishan, H., Saritas, E. U., Bagci, U., Celik, H., & Çukur, T. (2023).

A Dual Sampling dilated Pre-activation residual Attention CNN (DSPANN) to attain accurate COVID-19 detection. A progressive split deformable field fusion module integrated features for both visual and sound. By using augmented Snake Optimization (SO), the proposed DSPANN reduced computational complexity, Thandu, A. L., & Pradeepini, G. (2024).

A style distribution transfer generative adversarial network (SD-GAN) to diagnose COVID-19. The created model successfully performed COVID-19 detection by transforming the style distribution of imaging among various datasets containing individuals of various racial backgrounds, Kausar, T., Lu, Y., Kausar, A., Ali, M., & Yousaf, A. (2023).

Deep Reinforcement Learning (RL) with class activation maps to forecast patients' disease progression. The suggested RL scheme searched the approach respective to higher rewards for all epochs. Besides, it obtained the initial model for the next epoch using validation set accuracy as a reward, Chen, S., Liu, M., Deng, P., Deng, J., Yuan, Y., Cheng, X., Xie, T., Xie, L., Zhang, W., Gong, H., & Wang, X. (2022).

Recent research has applied machine learning models like Naïve Bayes and SVM for predicting bipolar disorder cases. Studies highlight that deep learning methods enhance accuracy in mental health diagnosis, Peerbasha, S., Mohamed Surputheen, M. (2021).

A segmentation-based COVID-19 classification network (SC2Net) using CXR imaging. The spatial attention network (SANet) as well as COVID-19 lung segmentation (CLSeg) block were 2 subnets that make up the SC2Net. Firstly, lungs were segmented in CXR using the CLSeg in order to suppress background interference. After segmenting the lung region, the SANet was fed input to the classifier, Deng, W., Shang, S., Cai, X., Zhao, H., Song, Y., & Xu, J. (2021).

Problem statement

Table 1 portrays the recent advancements in COVID-19 detection via various approaches. Generally, DL, RL and ensemble models exposed promising performance in detecting COVID-19 using CXR and CT images. Also, the self-attention mechanism boosts the accuracy performance of the classifiers. However, some of the issues still restrict the effective and trustworthy implementation of COVID-19 detection. Usually, COVID-19 datasets were frequently of limited size and quality, particularly in the early phases of the pandemic. Since many models were trained on tiny datasets, there was overfitting and restricted generalizability. Furthermore, the performance of the model was skewed in many cases by the datasets being highly unbalanced, with a notably higher proportion of non-COVID-19 cases than positive cases. It takes experience and a lot of work to accurately label COVID-19 cases, which increases the risk

of mistakes and inconsistent training data. The inability to easily interpret the problematic medical context limits the classifier's performance. Besides, models for RL were frequently intricate and demanded a large amount of processing power. When DL models were applied to unseen data from different populations, scanners, or hospitals, they performed poorly due to overfitting the unique features of training samples. This restricts the model's capability to be applied to different areas or contexts and to be generalized. Thereby, new ideas and methodologies are mandatory in attaining expected accuracy and performance in COVID-19 diagnosis and pave the way for researchers to devise novel detection networks.

Proposed architecture

Fig. 1 delivers the overview of the proposed COVIDNet-Predictor. Before feeding the network with the input image, some pre-processing stages are involved to enhance the input data to attain the expected results. Firstly, a pre-processing module to enhance the image quality using contrast limited adaptive histogram equalization (CLAHE) and noise reduction via non-local means (NLM). Secondly, a segmenting module using U-Net to segment the ROI such as lungs, surrounding features and pathological features. Thirdly, the color features from color histograms, texture features from the gray-level co-occurrence matrix (GLCM), and shape features from Contour Analysis are extracted. Fourthly, a hybrid optimizer Evolution-based NSGA II (EvoNSGA II) is developed using non-dominated sorting genetic algorithm II (NSGA-II) and differential evolution (DE). Fifthly, a feature fusion module using HAN. Sixthly, the fused features are given as input to COVIDNet-Predictor, which is a combination of MHCNN, and LSTM. Finally, a monitoring and feedback module is added to continuously monitor the patient data to make the system fit for real-time settings.

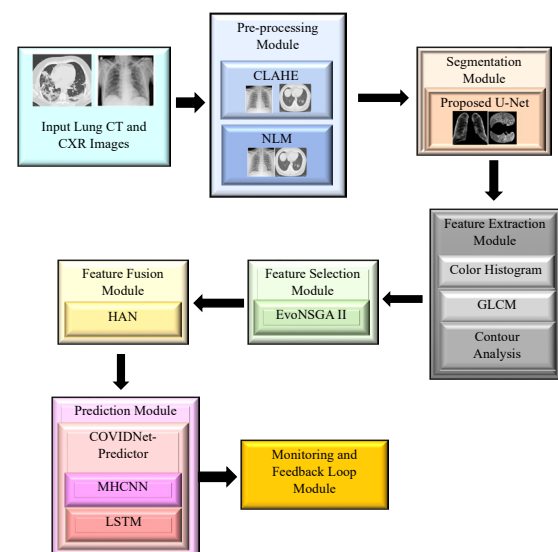


Figure 1: Overview of proposed COVID net-predictor model

Table 1: advantages and limitations of covid-19 detection via various methods

<i>Authors/Year</i>	<i>Methods</i>	<i>Aim</i>	<i>Advantages</i>	<i>Limitations</i>
Chetupalli <i>et al.</i> in 2023	LR, SVM, and LSTM	To know the significance of acoustic features in COVID-19 detection	96% Area-Under-Curve (AUC)	Low accuracy and sensitivity performance were achieved
Bhaskar <i>et al.</i> in 2023	CNN-CatBoost	To detect COVID-19 via exhaled breath features	96% accuracy	Exposed high-level computational complexity for implementing sensing modules
Zhu <i>et al.</i> in 2023	2-stage Prediction Model	To exhibit the role of spectral features in COVID-19 detection	Attained low-cost system	Exposed minimum sensitivity and precision performance
Aytekin <i>et al.</i> in 2023	HST	To detect COVID-19 using spectrogram representations of respiratory sounds	90% AUC	Accuracy performance needed to be improved
Thandu & Pradeepini in 2023	DSPANN	To detect COVID-19 via both visual and sound features	98% accuracy	Exposed complexity in implementing attention mechanism
Kausar in 2023	SD-GAN	To diagnose COVID-19 using style variations	98% accuracy	Exposed more misclassified results
Chen <i>et al.</i> in 2022	RL	To forecast disease progression using class activation maps	91% accuracy	Need to enhance accuracy performance
Zhao <i>et al.</i> in 2022	SC2Net	To classify COVID-19 using CXR images	84% accuracy	Accuracy performance further needs to be improved

A Novel COVID-19 Prediction Framework

Pre-processing Module

Pre-processing images before feeding them into a DL network is crucial. It improves model performance by normalizing pixel values and correcting distortions, ensuring consistent input quality. Removing noise and artifacts helps the classifier focus on relevant features. Overall, effective pre-processing leads to more accurate and robust predictions.

CLAHE

In general, CLAHE, Hayati, M., Muchtar, K., Maulina, N., Syamsuddin, I., Elwirehardja, G. N., & Pardamean, B. (2023) limits the contrast improvement, making it effective in handling noise. Additionally, applying histogram equalisation to small tiles improves the image. Histogram equalization is assessed by computing histogram for each tile and normalizing the histogram as stated in Eq. (1), in which H_T denotes the overall histogram, and P represents pixel count.

$$H_T = \frac{h_T}{a \times b} \quad (1)$$

Now, compute the Cumulative Distribution Function (CDF) C_T based on Eq. (2), in which H'_T signifies the modified histogram and is achieved based on Eq. (3), and D represents histogram bins count.

$$C_T = \sum_{i=0}^j H'_T(i) \quad (2)$$

$$H''_T(i) = H'_T(i) + \frac{P}{D} \quad (3)$$

Here, $H'_T(i)$ points to clip histogram and is estimated based on Eq. (4), in which G refers to clip limit.

$$H'_T(i) = \min(H_T(i), G) \quad (4)$$

Eq. (5) determines the excess pixels P , in which

$$P = \sum_{i=0}^{D-1} \max(H_T(i) - G, 0) \quad (5)$$

The equalized tiles are combined using bilinear interpolation to prevent the need for arbitrary tile borders. The quality-enhanced images are given in Eq. (6).

$$I_{clahe}(i) = H''_T(i) \quad (6)$$

NLM

It is a technique used for improving the clarity of medical imaging through noise reduction while preserving important details. The core idea of NLM is to average similar patches from different parts of the image to denoise a given pixel through similarity of patches rather than their spatial proximity (Zhang, 2022). The value of a pixel $I_p(i)$ is assessed as similar pixels' weighted average from entire images, where weight is calculated via similarity of patches around each pixel as given in Eq. (7).

$$I_{nlm}(i) = \frac{1}{G(i)} \sum_{j \in \mathcal{G}} \exp\left(-\frac{I_p(i) - I_p(j)}{f^2}\right) I_p(j) \quad (7)$$

Here, $G(i)$ refers to normalization term ensuring the weighted sum to 1, $I_p(g_i)$ and $I_p(g_j)$ points to image patches centered at pixels i and j , respectively, f denotes parameter controlling the filter's sensitivity to patch differences, and $\mathcal{S}$ addresses search window.

Segmentation module

Segmentation is crucial in medical images as it isolates relevant regions (such as lungs or bronchi) from irrelevant backgrounds, enhancing the classifier's focus on critical features. This process improves model accuracy by reducing noise and irrelevant data, leading to more precise and meaningful predictions.

Proposed U-Net

The U-Net model, Al Qurri, A., & Almekkawy, M. (2023) is a broadly utilized architecture for image segmentation, mainly ROI identification. It is specifically significant in segmenting biomedical imaging due to its ability to capture both local and global features through its encoder-decoder structure and skip connections. Normally, U-Net comprises 2 key structures such as an encoder (contracting path) to capture context and minimize the spatial dimensions of the image via consecutive convolutional and pooling layers. Decoder (expansive path) to reconstruct spatial resolution of an image via transposed convolutions (upsampling) and integrates fine-grained data from the encoder through skip connections.

In U-Net, the prediction for each pixel is obtained by applying a series of convolutions and deconvolutions, incorporating data from both encoder and decoder paths. Basic operation at each layer involves convolutions followed by non-linear activation functions. The implemented U-Net layer operation is given in Eq. (8), and the final segmented image is given in Eq. (9).

$$ROI(i) = ReLU\left(Conv\left(Conc\left(F_{enc}(i), F_{dec}(i)\right)\right)\right) \quad (8)$$

$$I_{seg}(i) = ROI(i) \quad (9)$$

Here, $ReLU$ specifies non-linear activation, $Conv$ addresses convolution, $Conc$ means concatenation, $F_{enc}(i)$ indicates features extracted from the encoder path and $F_{dec}(i)$ signifies features from the corresponding decoder layer, which are concatenated with $F_{enc}(i)$ through skip connections.

Feature extraction module

Feature extraction is essential in processing images as it minimizes data dimensionality, highlighting the most relevant patterns and structures. This process enhances the classifier's efficiency and accuracy by focusing on critical features, leading to better generalization and faster convergence during training.

Color histogram

A color histogram can help in identifying and analyzing different tissues, and other features based on their color distribution. The histogram captures the frequency of color intensities across the image, making it a useful descriptor for distinguishing between different regions in the image, Mohseni, S. A., Wu, H. R., Thom, J. A., & Bab-Hadiashar, A. (2020). Eq. (10) gives the color histogram $I_{col_{hist}}^i(c)$ estimation, in which c represents color, $I_{seg}^i(x, y)$ means for color value of the pixel at position (x, y) in a segmented image, δ points to Kronecker delta function in $[0, 1]$ and X and Y specifies image dimensions.

$$I_{col_{hist}}^i(c) = \sum_{x=1}^X \sum_{y=1}^Y \varphi\left(I_{seg}^i(x, y) - c\right) \quad (10)$$

GLCM

A GLCM, Aouat, S., Ait-hammi, I., & Hamouchene, I. (2021) is a matrix in which row and column count is equal to possible gray level amount in images. Each component (a, b) in the GLCM represents the frequency where pixels with value a occurs adjacent to pixels with value b in a specified direction and distance. The contrast $I_c(i)$ features are extracted based on Eq. (11), in which $P(a, b)$ denotes normalized probability of a pixel with gray level a being adjacent to a pixel with gray level b within a specified spatial relationship in the image.

$$I_c(i) = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} (a-b)^2 \cdot P(a, b) \quad (11)$$

Eq. (12), (13), (14) and (15) shows the energy feature $I_e(i)$, homogeneity $I_h(i)$, correlation $I_{cor}(i)$, and entropy $I_{ent}(i)$.

$$I_e(i) = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} P(a, b)^2 \quad (12)$$

$$I_h(i) = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} \frac{P(a, b)}{1 + |a - b|} \quad (13)$$

$$I_{cor}(i) = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} \frac{(a \cdot b \cdot P(a, b)) - \mu_m \cdot \mu_n}{\sigma_m \cdot \sigma_n} \quad (14)$$

Here, μ_m and μ_n refers to mean, and σ_m and σ_n specifies standard deviation for distributions of $P(a, b)$.

$$I_{ent}(i) = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} P(a, b) \cdot \log(P(a, b)) \quad (15)$$

The final GLCM texture features are given in Eq. (16)

$$I_{GLCM}(i) = \{I_c(i), I_e(i), I_h(i), I_{cor}(i), I_{ent}(i)\} \quad (17)$$

Contour analysis

Contours are curves that join continuous points along the boundary of an object with the same intensity or color. In binary images, contours can be detected using edge detection methods or by finding the boundaries of

connected components, Elder, J. H. (2018). Several shape features can be extracted from contours. Here, perimeter, area, eccentricity, circularity, convex hull and solidity aspect ratio are utilized.

Feature selection module

Selecting substantial features is important as it minimizes the data dimensionality, helping to eliminate irrelevant and redundant features. This leads to more efficient model training, improves generalization by preventing overfitting, and enhances the total accuracy and interpretability of the network's predictions.

Hybrid optimizer evo NSGA II

Proposed EvoNSGA II is a combination of DE, Deng, W., Shang, S., Cai, X., Zhao, H., Song, Y., & Xu, J. (2021) and NSGA II, Xue, Y., Zhu, H., & Neri, F. (2023). Generally, DE is an optimization technique that uses differential mutation and crossover for exploring and exploiting search space efficiently. Besides, NSGA II is a renowned Multi-Objective (MO) evolutionary technique that utilizes non-dominated sorting as well as crowding distance for maintaining diversity and convergence towards the Pareto front. Initialize a population pop of size n randomly. Create mutant vectors $M_v(i)$ using the differential mutation strategy as given in Eq. (18), in which x_a^1 , x_a^2 , and x_a^3 represents randomly selected individuals, and S denotes the scaling factor.

$$M_v(i) = x_a^1 + S \cdot (x_a^2 - x_a^3) \quad (18)$$

Generate trial vectors $T_v(i)$ by combining mutant vectors with current vectors as stated in Eq. (19) where cr addresses crossover rate, and z_{rand} defines a randomly chosen index to ensure diversity.

$$T_v(i) = \begin{cases} M_v(i) & \text{if } rand < cr \text{ or } z = z_{rand} \\ x_i & \text{o.w} \end{cases} \quad (19)$$

Apply non-dominant sorting by sorting pop into various fronts using Pareto dominance. Eq. (20) gives the Pareto dominance which determines one solution if x_i dominates another solution x_z .

$$x_i \text{ dominates } x_z \text{ if } \begin{cases} f_i \leq f_z & \forall Fit \\ f_i < f_z \text{ for at least one fit} \end{cases} \quad (20)$$

Evaluate the diversity of solutions within each front by calculating crowding distance CD_i as addressed in Eq. (21), in which $k_{i,h}$ denotes a distance between the i^{th} individual and its neighbors in the h^{th} objective, and K_h signifies range of h^{th} objective.

$$CD_i = \sum_{h=0}^H \frac{k_{i,h}}{K_h} \quad (21)$$

After generating offspring $pop_{offspring}$ using DE operations (mutation and crossover), combine the parent population

pop with the offspring population $pop_{offspring}$ to form pop_{new} . The combined population $pop_{combined}$ size is $2n$ (where n specifies size of the original population) and is estimated based on Eq. (22). Estimate final population based on Eq. (23).

$$pop_{combined} = pop \cup pop_{new} \quad (22)$$

$$pop_{final} = best(pop_{combined}, pop) \quad (23)$$

Select individuals from the combined parent and offspring populations to form the next generation using non-dominated sorting and crowding distance. Here, the proposed EvoNSGA II is a MO algorithm and fitness is based on fit_1 for accuracy maximization and fit_2 as a cost function for selected features and cost as given in Eq. (24) and (25), in which ω denotes a weighting factor to balance the cost.

$$fit_1 = \max(Acc) \quad (24)$$

$$fit_2 = \omega \cdot n(selected\ features(f) + cost) \quad (25)$$

The fitness of proposed EvoNSGA II is determined based on Eq. (26).

$$fit = [fit_1, fit_2] \quad (26)$$

At this point, the proposed EvoNSGA II considers maximizing accuracy with a more compact optimal feature set. Algorithm 1 gives the pseudocode of developed EvoNSGA II.

Multi-Modal Feature Fusion Module

Multi-modal feature fusion enhances diagnostic accuracy by integrating complementary information from various imaging modalities such as CT, and CXR. This comprehensive approach improves the robustness and reliability of the classification by capturing diverse features and patterns that individual modalities might miss. Ultimately, it enables more precise and informed medical decision-making.

HAN

HAN, Tao, H., & Duan, Q. (2024) is normally designed to handle multi-modal feature fusion by incorporating attention mechanisms at different levels of a hierarchical structure. Features from different modalities are extracted and optimized $F_{optimal}$ and given as input to HAN. Compute attention scores $A(i)$ for each image feature in $F_{optimal}$ as given in Eq. (27), in which W_v denotes attention weight vector for the image modality.

$$A(i) = \frac{\exp(W_v^T F_{optimal}(i))}{\sum_{j=0}^n \exp(W_v^T F_{optimal}(j))} \quad (27)$$

Compute the context vector $C_v(i)$ for $F_{optimal}$ as stated in Eq. (28),

Algorithm 1: Pseudocode of Developed EvoNSGA II

Begin

Initialize pop , n , maximum and current iteration m_{it} and t

Fitness estimation

While $(t < m_{it})$

Generate Offspring

For $i = 1 : pop$ Select three random individuals x_a^1 , x_a^2 , and x_a^3

Apply DE Mutation as per Eq. (18)

Apply DE Crossover as per rand

Evaluate the fitness of trial vector using Eq. (19)

End for

Combine Parent and Offspring Population

Create a pop_{new} by combining pop and the newly generated offspring $pop_{offspring}$

Apply Non-Dominated Sorting as per Eq. (20)

Sort the combined population $pop_{combined}$ into different non-dominated frontsCalculate CD_i for individuals in each front as per Eq. (21)

Apply Selection

Select the next generation population P based on non-dominated fronts in Eq. (20) and CD_i to maintain diversity $t = t + 1$

End while

Return best features in pop_{final} End

$$C_v(i) = \sum_{i=0}^n A(i) F_{optimal}(i) \quad (28)$$

Combine the context vectors from each modality as expressed in Eq. (29).

$$f_{fusion}(i) = [C_v(i)] \quad (29)$$

Apply a dense layer to merge these features as shown in Eq. (30), in which W_{fusion} represents weight matrix, and b_{fusion} indicates bias.

$$F_{fused} = ReLU(W_{fusion} \bullet f_{fusion} + b_{fusion}) \quad (30)$$

COVID Net-Predictor

COVIDNet-Predictor is an integration of MHCNN and LSTM architectures. This name highlights the model's advanced and multi-faceted approach to achieving robust and accurate predictions. By combining the powerful Neural Network (NN) components, the model utilizes strengths of feature extraction, temporal sequence analysis, and image classification. This integration aims to provide a comprehensive solution for predicting COVID-19 from medical imaging such as CT and CXR, enhancing diagnostic accuracy and reliability. Fig. 2 displays the architecture of proposed COVIDNet-Predictor.

MHCNN-LSTM

The MHCNN-LSTM component captures spatial and temporal dependencies, particularly useful if processing sequences of images, Kumar, A., Vishwakarma, A., & Bajaj, V. (2024) and Hasib, K. M., Azam, S., Karim, A., Al Marouf, A., Shamrat, F. J. M., Montaha, S., & Rokne, J. G. (2023).

CNN Layers

The CNN applies convolutional filter (kernels) to input images. Every filter finds particular features in images. The convolution operation is given in Eq. (31), in which $F_{out}(m, n)$ indicates output feature map at position (m, n) , C specifies convolutional kernel, and $F_{Darknet}$ is the input feature from DarkNet.

$$F_{out}(m, n) = \sum_{x=1}^X \sum_{y=1}^Y F_{Darknet}(m+x, n+y) \cdot C_{x,y} \quad (31)$$

After convolution, activation functions like Rectified Linear Unit (ReLU) are used element-wise for introducing non-linearity as stated in Eq. (32)

$$F_{relu}(m, n) = \max(0, F_{out}(m, n)) \quad (32)$$

Pooling layers lessen feature map dimensionality without disturbing important features by selecting the higher values from a feature map area as expressed in Eq. (33)

$$F_{pool}(m, n) = \max_{x \in \text{pooling region}} F_{relu}(m+x, n_y) \quad (33)$$

The final output F_{cm} is the feature map after applying all convolutional, activation, and pooling operations. This feature map captures various hierarchical features from the input features $F_{Darknet}$ as shown in Eq. (34).

$$F_{cm} = CNN(F_{Darknet}) \quad (34)$$

LSTM Layer

In LSTM (Hasib et al., 2023) layer, the sequence of feature vectors from the CNN to capture temporal dependencies

are processed as given in Eq. (35), F_{cm} indicates input feature vector at time t .

$$Y_{lstm} = H_t, C_t = LSTM(F_{cm}, H_{t-1}, C_{t-1}) \quad (35)$$

Here, H_{t-1} represents a hidden state from preceding time step $t-1$, carrying short-term memory, C_{t-1} specifies cell state from preceding time step $t-1$, maintaining long-term memory. Besides, H_t signifies updated hidden state for the current time step t , combining the relevant information processed so far, and C_t means for updated cell state for the current time step t , integrating new information while preserving long-term dependencies. Y_{lstm} is the final feature set from MHCNN-LSTM layer. Now, pass the final features Y_{lstm} through Fully Connected (FC) layers to produce the final prediction as stated in Eq. (36), in which W_{fc} and b_{fc} addresses weight and bias of FC layer, and P refers to probability distribution over classes.

$$P = \text{Softmax}(W_{fc} \cdot Y_{lstm} \cdot b_{fc}) \quad (36)$$

Table 2 shows the hyperparameter settings of proposed COVIDNet-Predictor.

Monitoring and Feedback Loop

A feedback loop is implemented that continuously monitors the model's performance in real-time settings and updates it based on new patient data and outcomes. This process involves systematically tracking the model's performance, gathering new data, and updating the model to adapt to any changes in the underlying data distribution or clinical practices. Implementing a feedback loop in a DL model for medical applications involves continuously monitoring the model's performance in real-world settings, such as hospitals or clinics. The model's predictions are tracked against actual patient outcomes, with key performance indicators like accuracy, sensitivity, and specificity being evaluated regularly. If the model's performance falls below predefined thresholds, alerts are triggered, prompting further analysis. New patient data, particularly cases where the model's predictions were incorrect, is labeled and added to the training dataset. This new data helps to address any biases or changes in patient demographics or clinical practices that can have impacted the model's accuracy. The feedback loop is completed by retraining the model with the updated dataset and redeploying the improved version back into the clinical setting. This continuous process ensures that the model adapts to new data and remains accurate over time. By regularly incorporating real-world feedback and new patient data, the model evolves to better meet the needs of healthcare providers, improving its reliability and effectiveness in making accurate predictions.

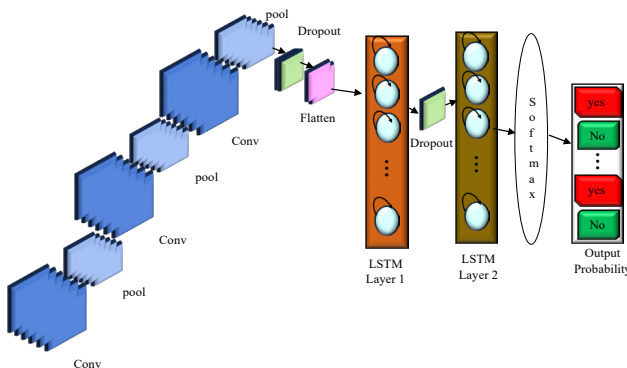


Figure 2: Architecture of proposed COVIDNet-Predictor

Table 2: Hyperparameter Settings of COVIDNet-Predictor

<i>Module</i>	<i>Hyper parameter</i>	<i>Description</i>	<i>Values</i>
MHCNN-LSTM	Number of CNN layers	Before LSTM	3
	Filter Size	Convolutional filters	3×3
	Number of LSTM units	In LSTM layer	50
	Number of LSTM Layers	In LSTM module	2
COVIDNet-Predictor	Learning Rate	For optimizer	0.001
	Batch Size	Number of samples per batch	32
	Epochs	For training	200
	Optimizer	Type of optimizer	Adam
	Dropout Rate	To prevent overfitting	0.5

Simulation Results

Simulation Setup

The proposed COVID-19 prediction model through COVIDNet-Predictor is developed in Python on an Intel core® i5 processor with 2.3 GHz, 16 GB RAM, and 64-bit OS. Here, Chest CT and CXR Datasets are utilized for prediction and is accessible via <https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images> and <https://www.kaggle.com/datasets/nih-chest-xrays/data>. In order to analyze the competence of developed COVIDNet-Predictor, a comparative evaluation is carried out with baseline models such as CNN, MHCNN, and LSTM, and SOTA models such as DenseNet, Sanghvi, H. A., Patel, R. H., Agarwal, A., Gupta, S, Sawhney, V., & Pandya, A. S. (2023), and Darknet, Mahrishi, M., Morwal, S., Muzaffar, A. W., Bhatia, S., Dadheech, P., & Rahmani, M. K. I. (2021). The performance measures including accuracy, specificity, recall, precision, and F1-score are employed to estimate the efficacy of proposed COVIDNet-Predictor.

Algorithmic Analysis

The proposed COVIDNet-Predictor model and the attained results are addressed in this section. Table 3 shows the performance of COVIDNet-Predictor over other baseline and SOTA models. The accuracy of developed COVIDNet-Predictor is high when compared with existing models, which are 1.67%, 0.95%, 2.67%, 3.59%, and 2.76% better than

CNN, DenseNet, Darknet, MHCNN, and LSTM, respectively. For recall, the proposed methods reached 5.5% improved than CNN, 6.95% better than DenseNet, 2.69% better than Darknet, 2.68% improved than MHCNN, and 4.66% better than LSTM. As for specificity, precision, and F1-score, the proposed model attained better performance and outperformed other models. Thereby, the proposed COVIDNet-Predictor is efficient in predicting COVID-19-prone patients using CXR and chest CT images.

Fig. 3 shows the Receiver Operating Curve (ROC) of proposed COVIDNet-Predictor. The ROC curve for the COVID-19 prediction using the proposed COVIDNet-Predictor demonstrates excellent model performance, with an Area Under the Curve (AUC) of 0.9504. This indicates that the model has a very high true positive rate (sensitivity) while maintaining a low false positive rate. The curve's proximity to the top left corner of the graph reflects the model's ability to accurately distinguish between COVID-19 positive and negative cases. The near-perfect AUC score highlights the effectiveness of COVIDNet-Predictor in optimizing the model for superior prediction accuracy in COVID-19 diagnosis, ensuring reliable and robust classification in practical applications.

Fig. 4 represents the validation loss and accuracy plot of proposed COVIDNet-Predictor. Validation accuracy typically increases as the model learns from the data, indicating improved performance over time. If the validation loss decreases alongside increasing accuracy, this recommends

Table 3: Performance of COVIDNet-Predictor over other Baseline and SOTA Models

<i>Methods</i>	<i>Accuracy</i>	<i>Recall</i>	<i>Specificity</i>	<i>Precision</i>	<i>F1-score</i>
CNN	93.45	92.4	93.21	91.21	92.14
DenseNet	94.14	90.98	94.68	93.6	93.84
Darknet	92.51	95.14	93.11	90.15	92.21
MHCNN	91.62	95.15	90.32	91.25	91.34
LSTM	92.41	93.22	91.01	90.01	90.18
Proposed COVIDNet-Predictor	95.04	97.78	96.91	97.85	95.88

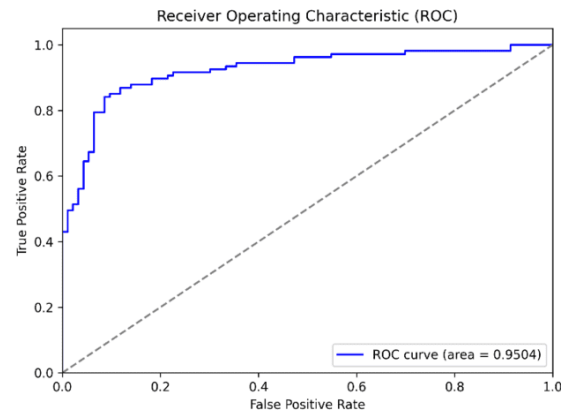


Figure 3: ROC of Proposed COVIDNet-Predictor

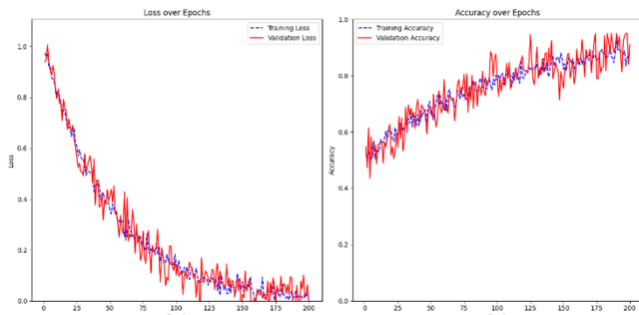


Figure 4: Validation Loss and Accuracy Plot of Proposed COVIDNet-Predictor

that the DL network is effectually learning and generalizing to unseen data. However, if the validation loss maximizes over epochs while accuracy plateaus or minimizes, it could specify overfitting, where the model is learning noise rather than the underlying patterns. For the COVIDNet-Predictor, a consistent decrease in validation loss with a corresponding increase in accuracy would reflect the model’s robustness in COVID-19 prediction tasks, confirming its effectiveness in generalizing from the training data to real-world applications.

Fig. 5 portrays the convergence of proposed EvoNSGA II over other models. The graph shows the convergence accuracy of various optimization algorithms, including DE, NSGA II, GA, PSO, and the proposed EvoNSGA-II, over iterations 10 to 50 for optimal feature selection. As iterations progress, the proposed EvoNSGA-II consistently outperforms the other algorithms, reaching the highest accuracy of 99.05%. This demonstrates the superior capability of EvoNSGA-II in refining the model’s predictive accuracy through iterative optimization. The other algorithms show steady improvements, but none reach the performance level of the proposed method, highlighting EvoNSGA-II’s effectiveness in handling complex optimization tasks for COVID-19 optimal feature selection.

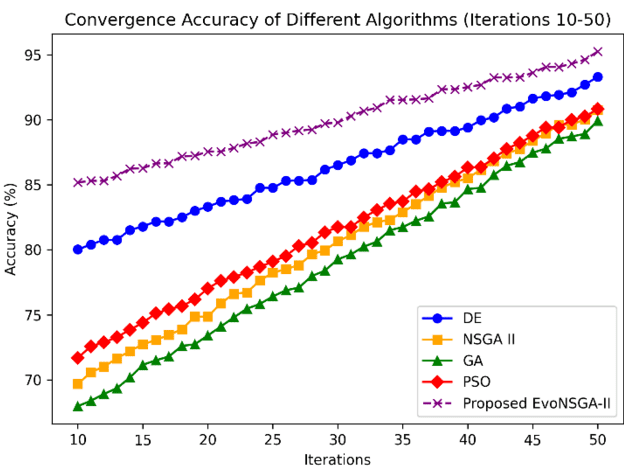


Figure 5: Convergence of proposed EvoNSGA II over other models

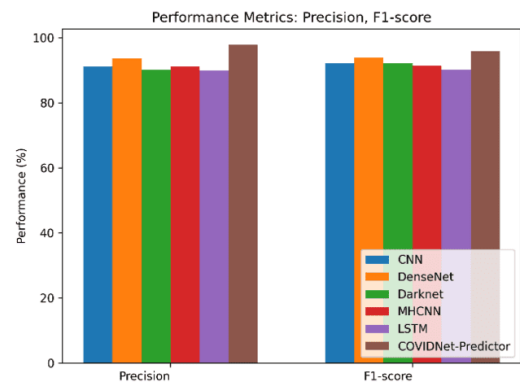
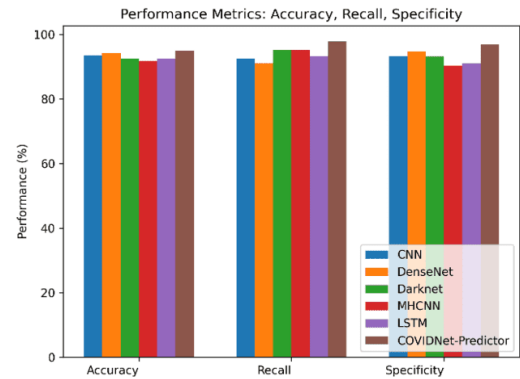


Figure 6: Performance of proposed COVIDNet-predictor over other models

Fig. 6 depicts the graphical representation of the performance of the proposed COVIDNet-Predictor over other models. In terms of accuracy, the proposed COVIDNet-Predictor outperforms the other models, demonstrating that it is more reliable in correctly identifying both COVID-19 positive and negative cases. Similarly, the sensitivity, precision, specificity and F1-score of the graph highlight the superior performance of the COVIDNet-Predictor across all four metrics, demonstrating its robustness and effectiveness

in the task of COVID-19 prediction. Thus, the proposed COVIDNet-predictor is an effective DL network for predicting COVID-19 and aids in early diagnosis.

Conclusion

In this paper, a novel COVIDNet-Predictor was introduced to predict the possibility of COVID-19-prone patients based on chest CT and CXR images. Certain pre-processing steps were applied to improve the input data in order to achieve the desired outcomes before feeding the network with the input image. First, a pre-processing module that used CLAHE to improve image quality and NLM to reduce noise was implemented. Second, U-Net, a segmentation module that segmented the ROI for the lungs, surrounding features, and pathological features. Thirdly, contour analysis for shape features, GLCM for texture features, and color histograms for color features were extracted. Fourth, NSGA-II and DE were used to create a hybrid optimizer named EvoNSGA II for feature selection was used. Besides, a HAN-based feature fusion module, and a sixth, COVIDNet-Predictor, which combined Multi Head CNN, and LSTM for prediction were employed. In order to make the system suitable for real-time settings, a monitoring and feedback module was added at the end to continuously monitor the patient data. The proposed COVIDNet-Predictor achieved 95.04% accuracy and outran all other baseline and SOTA models. In the future, COVIDNet-Predictor will be expanded with capabilities to detect a broader range of respiratory diseases, integrating with electronic health records for personalized patient analysis, and enhancing its adaptability to new COVID-19 variants and emerging diseases. It could also be deployed in low-resource settings, to assist in early detection and monitoring. Continuous updates and integration of multimodal data sources like lab results and clinical symptoms could further improve its diagnostic accuracy and clinical utility.

Acknowledgments

I am deeply grateful for the successful completion of this research article titled "COVIDNet-Predictor: A Multi-Head CNN and LSTM-Based Deep Learning Framework for COVID-19 Diagnosis". I would like to extend my sincere gratitude to my Research Supervisor, Dr. M. Mohamed Surputheen, for his valuable guidance and continuous support throughout this research. Special thanks to Dr. S. Peerbasha for his expertise and motivation, which greatly enhanced this work and appreciate the contributions of all collaborators and the support from our institution and colleagues, which made this research possible.

References

Al Qurri, A., & Almekkawy, M. (2023). Improved UNet with attention for medical image segmentation. *Sensors*, 23(20), 8589.
 Aouat, S., Ait-hammi, I., & Hamouchene, I. (2021). A new approach

for texture segmentation based on the Gray Level Co-occurrence Matrix. *Multimedia Tools and Applications*, 80(16), 24027-24052.
 Aslan, M. F., Unlarsen, M. F., Sabanci, K., & Durdu, A. (2021). CNN-based transfer learning-BiLSTM network: A novel approach for COVID-19 infection detection. *Applied Soft Computing*, 98, 106912.
 Aytekin, I., Dalmaz, O., Gonc, K., Ankishan, H., Saritas, E. U., Bagci, U., Celik, H., & Çukur, T. (2023). Covid-19 detection from respiratory sounds with hierarchical spectrogram transformers. *IEEE Journal of Biomedical and Health Informatics*.
 Bhaskar, N., Bairagi, V., Munot, M. V., Gaikwad, K. M., & Jadhav, S. T. (2023). Automated COVID-19 Detection From Exhaled Human Breath Using CNN-CatBoost Ensemble Model. *IEEE Sensors Letters*.
 Chetupalli, S. R., Krishnan, P., Sharma, N., Muguli, A., Kumar, R., Nanda, V., Pinto, L. M., Ghosh, P. K., & Ganapathy, S. (2023). Multi-modal point-of-care diagnostics for COVID-19 based on acoustics and symptoms. *IEEE Journal of Translational Engineering in Health and Medicine*, 11, 199-210.
 Chen, S., Liu, M., Deng, P., Deng, J., Yuan, Y., Cheng, X., Xie, T., Xie, L., Zhang, W., Gong, H., & Wang, X. (2022). Reinforcement learning based diagnosis and prediction for COVID-19 by optimizing a mixed cost function from CT images. *IEEE Journal of Biomedical and Health Informatics*, 26(11), 5344-5354.
 Dash, T. K., Chakraborty, C., Mahapatra, S., & Panda, G. (2022). Gradient boosting machine and efficient combination of features for speech-based detection of COVID-19. *IEEE Journal of Biomedical and Health Informatics*, 26(11), 5364-5371.
 Dairi, A., Harrou, F., & Sun, Y. (2021). Deep generative learning-based 1-SVM detectors for unsupervised COVID-19 infection detection using blood tests. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-11.
 Deng, W., Shang, S., Cai, X., Zhao, H., Song, Y., & Xu, J. (2021). An improved differential evolution algorithm and its application in optimization problem. *Soft Computing*, 25, 5277-5298.
 Dey, S., Bhattacharya, R., Malakar, S., Mirjalili, S., & Sarkar, R. (2021). Choquet fuzzy integral-based classifier ensemble technique for COVID-19 detection. *Computers in Biology and Medicine*, 135, 104585.
 Elder, J. H. (2018). Shape from contour: Computation and representation. *Annual Review of Vision Science*, 4(1), 423-450.
 Hasib, K. M., Azam, S., Karim, A., Al Marouf, A., Shamrat, F. J. M., Montaha, S., ... & Rokne, J. G. (2023). Mccnn-lstm: Combining cnn and lstm to classify multi-class text in imbalanced news data. *IEEE Access*.
 Hayati, M., Muchtar, K., Maulina, N., Syamsuddin, I., Elwirehardja, G. N., & Pardamean, B. (2023). Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning. *Procedia Computer Science*, 216, 57-66.
 Ji, T., Liu, Z., Wang, G., Guo, X., Lai, C., Chen, H., ... & Zhou, Q. (2020). Detection of COVID-19: A review of the current literature and future perspectives. *Biosensors and Bioelectronics*, 166, 112455.
 Kausar, T., Lu, Y., Kausar, A., Ali, M., & Yousaf, A. (2023). SD-GAN: A style distribution transfer generative adversarial network for Covid-19 detection through X-ray images. *IEEE Access*, 11, 24545-24560.

- Kumar, A., Vishwakarma, A., & Bajaj, V. (2024). Multi-headed CNN for colon cancer classification using histopathological images with tikhonov-based unsharp masking. *Multimedia Tools and Applications*, 1-20.
- Liu, J., Yu, H., & Zhang, S. (2020). The indispensable role of chest CT in the detection of coronavirus disease 2019 (COVID-19). *European Journal of Nuclear Medicine and Molecular Imaging*, 47, 1638-1639.
- Mahrishi, M., Morwal, S., Muzaffar, A. W., Bhatia, S., Dadheech, P., & Rahmani, M. K. I. (2021). Video index point detection and extraction framework using custom YoloV4 Darknet object detection model. *IEEE Access*, 9, 143378-143391.
- Mohseni, S. A., Wu, H. R., Thom, J. A., & Bab-Hadiashar, A. (2020). Recognizing induced emotions with only one feature: A novel color histogram-based system. *IEEE Access*, 8, 37173-37190.
- Perumal, V., Narayanan, V., & Rajasekar, S. J. S. (2021). Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. *Applied Intelligence*, 51(1), 341-358.
- Peerbasha, S., Mohamed Surputheen, M. (2021). A predictive model to identify possible affected bipolar disorder students using Naïve Bayes, Random Forest, and SVM machine learning techniques of data mining and building a sequential deep learning model using Keras. *IJCSNS International Journal of Computer Science and Network Security*, 21(5), 267-275.
- Pradhan, M., Shah, K., Alexander, A., Ajazuddin, Minz, S., Singh, M. R., Singh, D, Yadav, K, & Chauhan, N. S. (2022). COVID-19: clinical presentation and detection methods. *Journal of Immunoassay and Immunochemistry*, 43(1), 1951291.
- Rai, P., Kumar, B. K., Deekshit, V. K., Karunasagar, I., & Karunasagar, I. (2021). Detection technologies and recent developments in the diagnosis of COVID-19 infection. *Applied Microbiology and Biotechnology*, 105, 441-455.
- Ravikumar, K. K., Ishaque, M., Panigrahi, B. S., & Pattnaik, C. R. (2023). Detection of COVID-19 using AI application. *EAI Endorsed Transactions on Pervasive Health and Technology*, 9.
- Reshi, A. A., Rustam, F., Mehmood, A., Alhossan, A., Alrabiah, Z., Ahmad, A., Alsuwailam, H., & Choi, G. S. (2021). An efficient CNN model for COVID-19 disease detection based on x-ray image classification. *Complexity*, 2021(1), 6621607.
- Sharma, N., Saba, L., Khanna, N. N., Kalra, M. K., Fouda, M. M., & Suri, J. S. (2022). Segmentation-based classification deep learning model embedded with explainable AI for COVID-19 detection in chest X-ray scans. *Diagnostics*, 12(9), 2132.
- Tao, H., & Duan, Q. (2024). Hierarchical attention network with progressive feature fusion for facial expression recognition. *Neural Networks*, 170, 337-348.
- Thandu, A. L., & Pradeepini, G. (2024). Privacy-Centric Multi-Class Detection of COVID 19 through Breathing Sounds and Chest X-Ray Images. *IEEE Access*.
- Ullah, Z., Usman, M., Latif, S., & Gwak, J. (2023). Densely attention mechanism based network for COVID-19 detection in chest X-rays. *Scientific Reports*, 13(1), 261.
- Xue, Y., Zhu, H., & Neri, F. (2023). A feature selection approach based on NSGA-II with ReliefF. Elsevier 134.109-987.
- Zhang, X. (2022). Two-step non-local means method for image denoising. *Multidimensional Systems and Signal Processing*, 33(2), 341-366.