

RESEARCH ARTICLE



Automated System for Prediction and Prognosis of Infection Diseases Using Deep Learning-Based Approaches

OPEN ACCESS**Received:** 30-05-2023**Accepted:** 07-08-2023**Published:** 14-09-2023Kavita Thakur¹, Navneet Kaur Sandhu², Yogesh Kumar^{3*}¹ Research Scholar, Desh Bhagat University, Mandi Gobindgarh, Punjab, India² Assistant Professor, Desh Bhagat University, Mandi Gobindgarh, Punjab, India³ Department of CSE, School of Technology, Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Citation: Thakur K, Sandhu NK, Kumar Y (2023) Automated System for Prediction and Prognosis of Infection Diseases Using Deep Learning-Based Approaches. Indian Journal of Science and Technology 16(34): 2730-2739. <https://doi.org/10.17485/IJST/v16i34.1306>

* **Corresponding author.**yogesh.arora10744@gmail.com**Funding:** None**Competing Interests:** None

Copyright: © 2023 Thakur et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](https://www.isee.org/))

ISSN

Print: 0974-6846

Electronic: 0974-5645

Abstract

Objectives: This study explores the potential of deep learning-based techniques to improve disease management and intervention by focusing on their use in infectious disease prediction and prognosis. **Methods:** The research used deep learning models EfficientNetB0, NASNetLarge, DenseNet169, ResNet152V2, and InceptionResNetV2. For this study, a dataset comprising 29,252 images of different diseases such as COVID-19, MERS, Pneumonia, SARS, and tuberculosis. To visualize pixel intensity, exploratory data analysis was performed on the pictures. Preprocessing eliminated disruptive signals via image augmentation and contrast enhancement. After that, Otsu thresholding and contour feature morphological values retrieved relevant features. **Findings:** The best successful model was found to be EfficientNetB0. During training, it obtained a 90.22% accuracy rate, a loss of 0.279, having an RMSE value of 0.578. However, InceptionResNetV2 showed the best accuracy, loss, and RMSE values throughout model testing. The precise accuracy, loss, and RMSE results were 88%, 0.399, and 0.631, respectively. **Novelty:** The novelty resides in exploring methods based on deep learning for predicting and prognosticating infectious diseases, with the potential for handling diseases, strategies for intervention, and public health decisions.

Keywords: Tuberculosis; Pneumonia; Infectious diseases; Deep learning; InceptionResNetV2

1 Introduction

Recent investigations have shed light on the nature, transmission, and health effects of infectious diseases. These illnesses, caused by pathogenic microorganisms such as bacteria, viruses, fungi, and parasites, can be transmitted through casual contact, airborne droplets, contaminated food and drink, and insect vectors. Understanding the epidemiology of infectious illnesses is of paramount importance as it involves examining trends, causes, and consequences within populations⁽¹⁾. Genomics has brought transformative changes to infectious disease research, enabling the genetic identification and study of pathogens, facilitating insights into disease genesis,

evolution, and dissemination. This data aids in disease prediction, targeted therapy, and vaccine development. In recent years, immunology has significantly contributed to our comprehension of infectious diseases. As infectious disease prediction and management become increasingly critical, artificial intelligence (AI) has become indispensable in healthcare⁽²⁾. AI systems excel at rapidly and effectively processing massive volumes of data, enabling the anticipation of disease outbreaks, analysis of transmission patterns, and swift identification of high-risk groups. By analyzing signs, medical backgrounds, and diagnostic test results, AI systems can assist in the early detection and diagnosis of infectious diseases. Moreover, AI's ability to assess individual and population risk variables can prioritize infectious disease prevention and resource allocation. Additionally, machine learning and deep learning techniques have accelerated medication and vaccine development⁽³⁾.

This study conducts an exhaustive review of research articles to assess and analyze factors such as summarization, improvements, challenges, categorization, and possible effects. Convolutional neural networks, also known as CNN, were used in the research by Cinar et al.⁽⁴⁾ to distinguish between parasite and healthy malaria pictures. Initially, the dataset was filtered using a medium and a Gaussian one. Using the Gaussian-filtered dataset, the DenseNet201 architecture identified malaria data with an impressive accuracy rate of 97.83%. Irmak et al.⁽⁵⁾ suggested a unique CNN-based technique for categorizing COVID-19 severity. Researchers hypothesized that a computer network (CNN) model fed chest X-ray pictures could reliably categorize COVID-19 patients into the following categories: moderate, severe, mild, and critical. Cross-testing was used by Leblic et al.⁽⁶⁾ for predicting COVID-19 diagnosis using machine learning techniques. Loss of taste and smell was associated with higher positive probability COVID-19 values of 6.21 and 2.50, accordingly. The ML algorithms' COVID-19 diagnostic prediction accuracy, sensitivity, and specificity were 80%, 82%, and 78%, respectively. Feng et al.⁽⁷⁾ devised a method for automatically diagnosing COVID-19 by listening for coughs in audio recordings. Essential parts of an audio stream are extracted and labelled using artificial intelligence-based models like support vector machines, k-nearest neighbours, and recurrent neural networks. The highest accuracy (81.25%; AUC = 0.79) was attained when the approach was tested on data kept separate from the training data. Similar results were shown by Rahman et al.⁽⁸⁾, who segmented and enhanced chest X-ray images to detect tuberculosis (TB) accurately. The findings of their method for diagnosing TB were reliable. The study includes two alternative approaches X-ray segmentation and X-ray classification. DenseNet201 obtained 98.6% accuracy, 98.56% F1-score, 98.56% sensitivity, 98.57% precision, and 98.54% specificity for segmented lung images. To mimic cholera epidemics linked to seasonal weather oscillations, Leo et al.⁽⁹⁾ studied machine learning techniques and addressed the problem of data imbalance. To evaluate the performance of seven designs, were analyzed. XGBoost classifiers were chosen based on the model features as the research's best option. The results provide new insight into the essential roles that machine learning plays in healthcare data processing. Tran et al.⁽¹⁰⁾ provided an educational overview of the potential applications of artificial intelligence and machine learning in infectious disease testing. The authors covered supervised ML approaches commonly used in laboratory medicine which included diseases such as COVID-19, sepsis, and tuberculosis. The paper also discussed the concept of "data fusion" and how AI/ML could integrate multiple data streams to provide actionable clinical knowledge. The belief rule-based expert system (BRBES), according to Hossain et al.⁽¹¹⁾, can model ambiguous information and conclude under ambiguous circumstances. Furthermore, BRBES results were compared to those obtained using other data-driven machine-learning methods. Wang et al.⁽¹²⁾ developed a deep learning model called as multiple infectious diseases diagnostic model (MIDDM) to classify seven infectious diseases by using 10,000 data of unstructured electronic medical records and obtained better results in predicting and diagnosing the diseases. Taresh et al.⁽¹³⁾ evaluated various deep learning models for the diagnosis of Covid -19 which had been taken from chest x-ray images. They used the dataset of 1200 CXR covid -19 images, 1341 images of healthy patients, and 1345 images of patients suffering from viral pneumonia. During experimentation, it had been found that MobileNet and VGG16 computed the highest accuracy of 98.28%.

While several studies have explored the use of deep learning and machine learning techniques for infectious disease diagnosis and prediction, there are still areas that require further investigation and improvement:

- Limited scope of diseases: Many studies focus on specific infectious diseases such as COVID-19 or tuberculosis. This research claims to conduct an exhaustive review of research articles across a broader range of infectious diseases, which could provide a more comprehensive understanding of the applications of deep learning in infectious disease management.
- Dataset size and diversity: Some studies may have smaller datasets or focus on a limited patient population, leading to potential limitations in generalizability. This research aims to address this by analyzing various datasets encompassing multiple infectious diseases, which could enhance the generalizability of the findings.
- Addressing data imbalance: Imbalanced datasets can affect the performance of machine learning models, particularly in diseases with low prevalence. The study claims to address the issue of data imbalance in some cases, which can contribute to more accurate predictions and classifications.

By focusing on these gaps and conducting an exhaustive review of research articles, this study aims to develop a deep learning based model for the detection and classification of various infectious diseases in order to contribute to the advancement of AI applications in healthcare. The following are the ways that the paper contributes:

- Images from the dataset are visually visualized using exploratory data analysis, with the color channels of the photos being extracted using an RGB histogram.
- Before performing image augmentation alongside contrast enhancement procedures, it is necessary to convert the source images into grayscale.
- Utilized are feature extraction methods like Otsu thresholding and contour features. Afterwards, the collected data is partitioned into training and test sets based on retrieved characteristics.
- Different metrics are calculated throughout deep learning models' training and testing phases, including accuracy, loss, F1 score, recall, precision, and root-mean-squared error (RMSE).

The research study is organized into various sections. Section 1 acts as an introduction, summarizing infectious diseases, their effect, and the use of AI approaches to combat them. Section 2 provides background knowledge and examines current research attempts to identify numerous infectious diseases. It also describes the dataset, processes, methodologies, and parameters used to obtain the findings. Section 3 provides a wrap-up and final observations on this paper, and section 4 provides recommendations for further study and a preview of potential difficulties.

2 Methodology

This article examines the methods used at every investigation stage and is diagrammatically shown via the proposed system i.e. Figure 1 to depict the order of these stages.

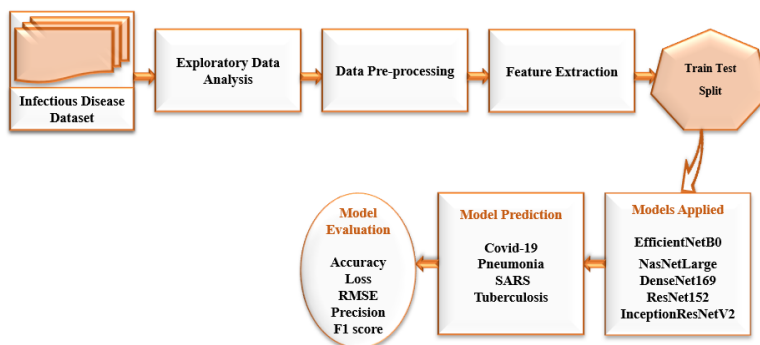


Fig 1. Proposed classification framework for infectious diseases

2.1 Dataset

Several different illnesses are represented in the datasets used in this investigation, from COVID-19 and lung opacity to MERS and pneumonia to TB and SARS to normal lung imaging and viral pneumonia. These data sets provide a compilation of information from many different sources. A large COVID-19 radiography dataset created a chest X-ray picture collection that reflects diverse diseases. The dataset consists of 1,341 images depicting chest X-rays (CXRs) and 1,345 impressions displaying CXRs exhibiting viral pneumonia. Additionally, the dataset includes the corresponding lung masks for each image. Furthermore, the dataset comprises 3,616 cases of COVID-19-positive individuals, 10,192 cases of individuals without any respiratory abnormalities, and 6,012 points of individuals with lung opacity (specifically, non-COPD lung infection). Each patient in the dataset is accompanied by its corresponding lung mask, as referenced in⁽¹⁴⁾. The chest X-ray images used to show pneumonia were taken from images of patients with conditions connected to pneumonia. Each image type (pneumonia/normal) is stored in its subdirectory under one of three main directories (train, test, val). It includes 5,863 X-ray pictures in JPEG format that have been labelled as pneumonia or normal⁽¹⁵⁾. Sputum samples were used to train the classifier on pictures related to TB for the corresponding dataset. This collection has 928 sputum photos with bounding boxes for 3,734 bacilli. The image's bounding boxes are specified in the XML file⁽¹⁶⁾. Table 1 provides the number of photos used for training and testing the model from each of the aforementioned illness groups, for 29,252. Jupyter Notebook was used to create the model, and several Python libraries were imported to make it work.

Table 1. Infectious disease classification by image count

Dataset classes	Count of images
COVID 19	3615
Pneumonia	3876
SARS	1144
Tuberculosis	2520

2.2 EDA

Histograms are first used to graphically represent the intensity distribution of an image before enabling the counting and displaying of RGB values. This visual aid may be used to investigate the image’s pattern and spot irregularities. Figure 2 shows how missing information or signal noise manifests in the original images. Histograms are created using the hist() function. After extraction, the hist() method receives its input from a flattened version of the color channel’s associated array. Two hundred fifty-six bins are selected to represent the whole range of values for pixels (0–255), where 0 is completely black and 255 is completely white.

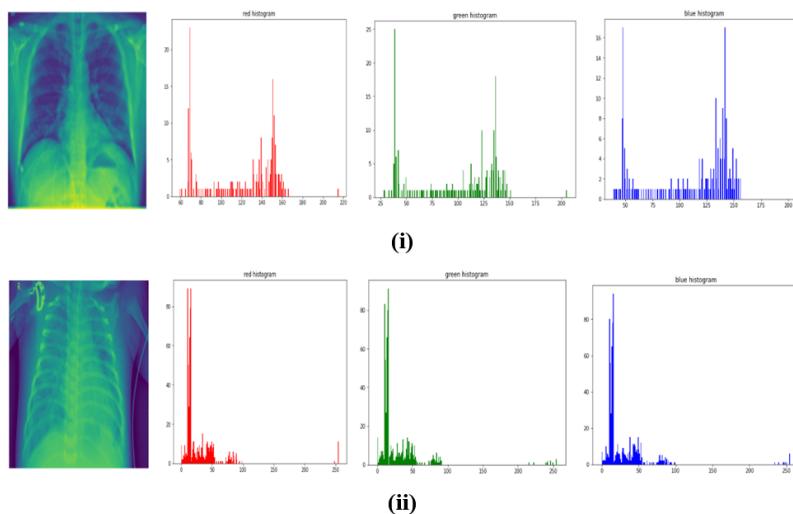


Fig 2. Visualizations of few images using histograms (i) CoVID-19, (ii) Pneumonia

2.3 Pre-processing

Preprocessing infectious illness images improves quality for analysis after visually inspecting pixel intensities. Figure 3 shows horizontal and vertical flip image augmentation in an illness image collection. Images and their annotations may be flipped to improve a model’s performance in surprising ways. The techniques described involve reversing pixel order, either row-wise or column-wise. The NumPy library offers np.flipud() and np.fliplr() that facilitate horizontal flipping. Vertical flipping entails rotating the rows and columns, whereas horizontal flipping involves reversing the rows and columns. The initial color images undergo a conversion process to transform them into grayscale during this procedure. The method of converting a color image to grayscale involves the utilization of a function known as G. The purpose of this function is to convert an $R^{m \times n \times 3}$ color image into an $R^{m \times n}$ representation, ensuring that all image values are within the range of 0 and 1. Figure 4 demonstrates how the original illustrations representing four distinct classes were converted to grayscale. The major benefit of converting RGB images to grayscale is the resulting algorithm’s simplicity and decrease in processing requirements.

The contrast of the images has been improved once the RGB images are converted to grayscale (as shown in Figure 5). The goal of this improvement is to make items in the images more visible by adjusting the brightness and darkness of such objects. Histogram equalization is a method used to further improve the contrast. The sklearn library’s equalize_adapthist () module has been imported to carry out this action. By utilizing this module, the images’ contrast is effectively increased, allowing for enhanced visual analysis.

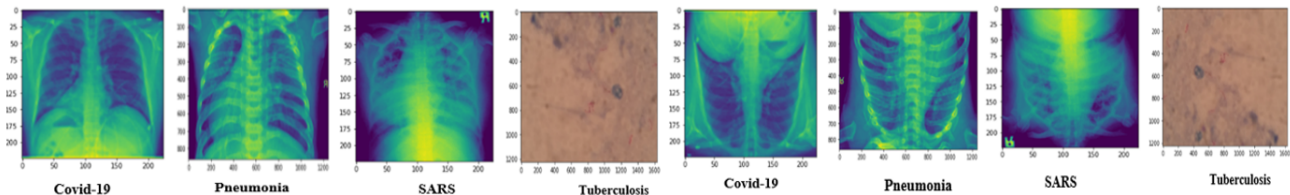


Fig 3. Original images flipped horizontally and vertically

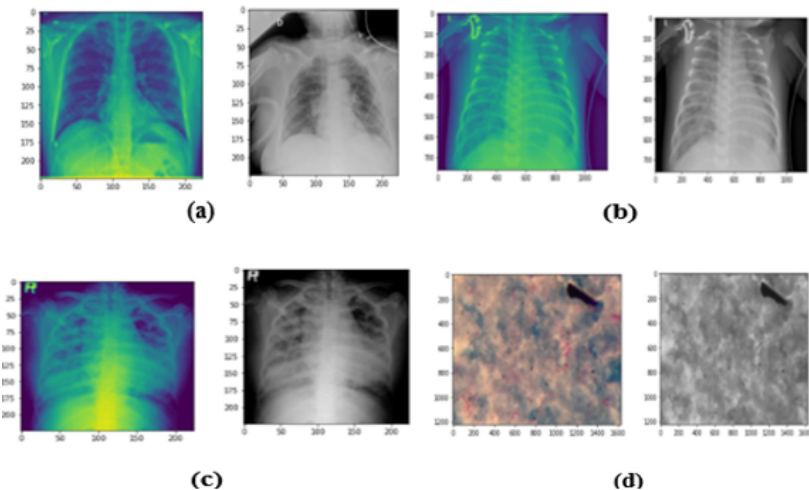


Fig 4. Conversion of RGB images to grayscale ((a) covid 19, (b) Pneumonia, (c) SARS, and (d) Tuberculosis)

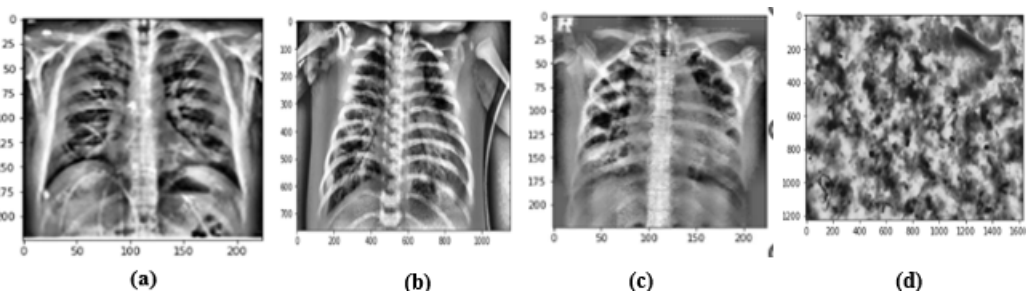


Fig 5. Images with enhanced contrast ((a) covid 19, (b) Pneumonia, (c) SARS, and (d) Tuberculosis)

2.4 Feature Extraction

This research segmented images and contour features were derived from feature extraction and target area retrieving. Images from eight categories were utilized in the contour features stage to get parameter morphological values. Figure 6 shows segmentation using the Otsu thresholding technique used to extract features from pre-processed images after morphological values were determined. Otsu’s approach, often known as the binarization process, is a straightforward automated thresholding technique with excellent results. The method determines each candidate cutoff value’s within-class (σ_{wc}) and between-class (σ_{bc}) variances. To create a binary or black-and-white image, the pixel values are switched out for either 0 (black) or 1 (white), depending on whether or not their brightness is below or over a predetermined threshold.

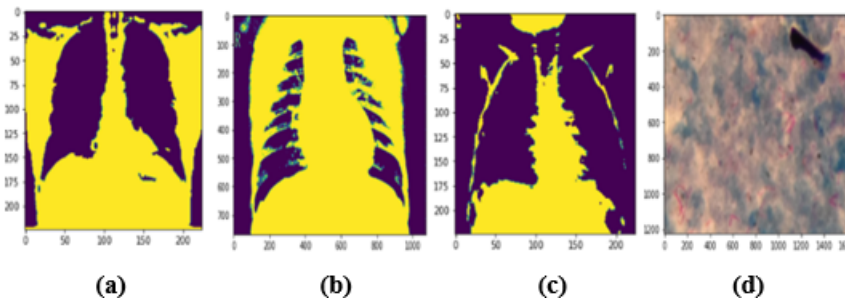


Fig 6. Images after using the Otsu method ((a) covid 19, (b) Pneumonia, (c) SARS, and (d) Tuberculosis)

2.5 Models Applied

In deep learning, EfficientNetB0 is a cutting-edge convolutional neural network, also called a CNN, architecture that has become quite well-known. EfficientNetB0, created by Google’s research group, excels in image categorization. Using compound scaling to optimize depth, breadth, and resolution, the network architecture balances model size and accuracy. Scaling these dimensions makes EfficientNetB0 computationally efficient and ideal for resource-constrained applications⁽¹⁷⁾.

NASNetLarge is a NAS model for large-scale image recognition. NASNetLarge uses reinforcement learning and neural architecture search algorithms. Its deep, expansive structure has several levels and characteristics. NASNetLarge has cell-based construction blocks, skip connections & path-level regularization. These aspects help the model collect and represent complex image patterns and characteristics, improving accuracy and performance. NASNetLarge’s flexible architecture makes it suited for many image identification jobs⁽¹⁸⁾. NASNetLarge is useful for handling difficult image recognition issues since neural architecture search automatically discovers and designs appropriate deep learning models.

In computer vision, the DCNN architecture known as DenseNet169 has generated interest. DenseNet169’s stellar results have made it a popular option for various image-related applications, including object identification and classification⁽¹⁹⁾. It’s notable for its thick connection pattern, in which all its layers are interconnected directly to improve feature reuse and data flow. By using this topology, DenseNet169 is better able to make use of the learnt representations from previous layers, which in turn improves gradient flow and helps to alleviate the vanishing gradient issue. DenseNet169 is a formidable CNN architecture with impressive accuracy and durability regarding visual identification tasks.

ResNet152V2, a 152-layer ResNet variation, has medicinal potential. ResNet152V2 excels in illness diagnosis, tumor identification, and segmentation because of its complex architecture. Because of the network’s depth, it can learn complicated patterns & extract high-level characteristics from medical images, resulting in more accurate and dependable predictions. ResNet152V2’s residual connections reduce the vanishing gradient issue, enabling stable and efficient training. This architecture’s performance and resilience help healthcare professionals improve medical imaging analysis and patient care⁽²⁰⁾.

InceptionResNetV2, which combines Inception and ResNet designs, has medicinal promise. Medical image analysis is enhanced by its multi-scale feature extraction and residual connections. InceptionResNetV2’s capacity to capture fine-grained details and hierarchical features improves illness classification, lesion identification, and segmentation⁽²¹⁾.

2.6 Performance Metrics

Accuracy: After training and testing, a model’s ability to accurately identify images is measured by a parameter called accuracy⁽²²⁾.

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}}$$

Loss: A machine learning model’s projected output is compared to the actual target value using a mathematical function called a loss function, which helps the model perform better while it is being trained⁽²³⁾.

$$\text{Loss} = \frac{(\text{Actual} - \text{Predicted})^2}{\text{Total number of observations}}$$

RMSE: The accuracy of prediction for infectious diseases is evaluated using the metric Root Mean Squared Error (RMSE), which indicates the typical discrepancy between anticipated and observed values⁽²³⁾.

$$RMSE = \sqrt{\frac{(\text{Actual}-\text{Predicted})^2}{\text{Total number of observations}}}$$

Recall: It is a deep learning statistic that quantifies the percentage of genuine positive predictions that a model accurately identifies out of all real positive cases⁽²⁴⁾.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Precision: A performance indicator called precision assesses how well a model predicts the future. Precision aims to reduce false positives, which makes it beneficial in situations like medical diagnosis when the cost of incorrectly identifying negative cases is considerable⁽²⁵⁾.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

F1 Score: A typical statistic for predicting medical disorders is the F1 score. It offers a fair assessment of a model's performance by combining recall and accuracy into a single rating⁽²⁶⁾.

$$F1 \text{ score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

3 Results and Discussion

Various models are compared in this section, with their results for different disease datasets such as Covid-19, SARS, pneumonia, and TB. Precision, loss, recall, accuracy, F1 score, & root mean square error (RMSE) are only a few assessment measures. The performance and applicability of the models is evaluated during training and testing phase for recognizing and classifying different illnesses by yielding essential findings.

Table 2. Model evaluation for several disease dataset classifications

Diseases	Models	Training			Testing		
		Accuracy	Loss	RMSE	Accuracy	Loss	RMSE
Covid19	EfficientNetB0	89.09	0.399	0.633	82.76	0.856	0.925
	NASNetLarge	50.34	1.660	1.288	50.35	1.375	1.172
	DenseNet169	81.06	0.558	0.746	80.46	0.695	0.833
	ResNet152V2	81.89	0.851	0.922	81.84	0.523	0.723
	InceptionResNetV2	83.61	0.435	0.659	86.00	0.300	0.547
Pneumonia	EfficientNetB0	82.37	0.853	0.923	90.09	0.299	0.546
	NASNetLarge	50.35	1.344	1.159	53.34	1.460	1.208
	DenseNet169	80.42	0.628	0.792	82.06	0.489	0.699
	ResNet152V2	81.05	0.574	0.757	82.89	0.442	0.664
	InceptionResNetV2	86.59	0.399	0.631	86.61	0.341	0.583
SARS	EfficientNetB0	92.37	0.753	0.867	99.09	0.199	0.446
	NASNetLarge	60.35	1.244	1.115	63.34	1.360	1.166
	DenseNet169	90.42	0.528	0.726	92.06	0.389	0.623
	ResNet152V2	91.05	0.474	0.688	92.89	0.342	0.584
	InceptionResNetV2	96.59	0.299	0.546	96.61	0.241	0.490
Tuberculosis	EfficientNetB0	92.37	0.753	0.867	96.00	0.853	0.923
	NASNetLarge	60.35	1.244	1.115	59.35	1.854	1.361
	DenseNet169	90.42	0.528	0.726	90.42	0.758	0.870
	ResNet152V2	91.05	0.474	0.688	84.05	0.764	0.874
	InceptionResNetV2	96.95	0.284	0.546	87.59	0.859	0.926

Table 2 displays the results of the model’s classification accuracy, loss, & RMSE measurements in training and testing for each illness class in the dataset. In particular, EfficientNetB0 showed the highest accuracy (88.09%) in training and predicting Covid disease. EfficientNetB0 has the best testing accuracies (89.09%, 99.09%, and 96%, respectively) for diseases including pneumonia, severe acute respiratory syndrome (SARS), and tuberculosis. Training accuracies of 86.59%, 96.59%, and 96.95% were attained using InceptionResNetV2 for the same disorders. DenseNet169 achieved 98.42% accuracy in both training and testing for MERS. The illness dataset was split into its component classes and then concatenated into two sub-datasets, one for training and one for testing, with four categories in each. The results of evaluating the models’ functionality on both datasets are summarized in Table 3.

Table 3. Model analysis during training and testing phase

Models	Training			Testing		
	Acc	Loss	RMSE	Acc	Loss	RMSE
EfficientNetB0	90.22	0.279	0.578	82.37	0.853	0.923
NASNetLarge	53.34	1.460	1.208	53.66	1.367	1.132
DenseNet169	82.06	0.489	0.699	80.42	0.628	0.792
ResNet152V2	82.89	0.442	0.664	81.05	0.574	0.757
InceptionResNetV2	86.61	0.341	0.583	88	0.399	0.631

EfficientNetB0 was found to be the most effective model as it achieved an accuracy of 90.22%, a loss of 0.279, and an RMSE value of 0.578 during training dataset. During model testing, InceptionResNetV2 achieved the highest accuracy, lowest loss, and lowest root mean squared error (RMSE) values. Specifically, the accuracy, loss, and RMSE values were 88%, 0.399, and 0.631, respectively. The NASNetLarge model has achieved a testing accuracy of 53.66%, a loss of 1.367, and an RMSE value of 1.132. Furthermore, the performance of the models was assessed using recall, accuracy, and F1 scores for the complete dataset, as shown in Table 4.

Table 4. Model performance evaluation

Models	Precision	Recall	F1 score
EfficientNetB0	75	87	90
NASNetLarge	89	66	65.87
DenseNet169	83.6	79.5	79
ResNet152V2	85.71	84	82.62
InceptionResNetV2	67.58	92	87.34

NASNetLarge had the greatest precision score (89%), whereas InceptionResNetV2 had the lowest score (67.58%). InceptionResNetV2 earned the greatest recall score of 92%, while NASNetLarge had the lowest score of 66%. EfficientNetB0 scored 90% on the F1 score, whereas NASNetLarge scored the lowest again, at 65.87%. Besides this, the paper also validated the model’s effectiveness and ensures their reliability through rigorous evaluation and analysis, as shown in Table 5, by addressing the challenges faced by the researchers in Sec 1.

Table 5. Comparison Analysis of the current study with the existing techniques

Ref	Diseases	Dataset	Techniques	Remarks
(5)	Covid	CXR images	CNN Model	Limited sample size
(6)	Covid	5 different types of dataset	Logistic Regression, Random Forest, Support Vector Machine	Low sample of 777 patients
(9)	Cholera	.csv data from Jan 2015 to Dec 2017	ADASYN, PCA, XgBoost	Class imbalance
(11)	Chikunguniya	Data of 250 patients	Expert System	Couldn’t handle the uncertainties
(12)	Tuberculosis	Electronic Health Record Data	CNN, MIDDM	Limited generalizability

Continued on next page

Table 5 continued

Our Study	COVID-19, MERS, Pneumonia, SARS, and tuberculosis	29,252 images of different infectious diseases	EfficientNetB0, DenseNet169, ResNet152V2, and InceptionResNetV2	NASNetLarge,	Our proposed system demonstrates improved classification rates and enhanced generalizability by incorporating diverse infectious diseases
------------------	---	--	---	--------------	---

4 Conclusion

Infectious disease prediction and prognosis methods based on deep learning have become more effective. These methods have been revolutionary developments in epidemiology, allowing researchers to better comprehend disease patterns and devise efficient intervention plans. Accurate forecasts and predictions are made possible by deep learning models' ability to uncover complex patterns and correlations that may not be immediately evident to human observers. This study used infectious illness data to train eight pre-trained models, including EfficientNetB0, NASNetLarge, DenseNet169, ResNet152V2, and InceptionResNetV2. The models performed well in both stages of the 75:25 training/testing dataset. The study revealed shortcomings. NASNetLarge performed poorly on the testing data set, and preprocessing and ROI extraction took a long time. Specific models are overfitted, underlining the need to resolve this problem in future work. Despite training and testing on 29,252 images, deep learning models might be improved. To enhance model accuracy, optimize hyperparameters. This study shows the benefits and drawbacks of developing deep-learning models to identify infectious diseases. The future of study in deep learning-based methods for the prognosis and optimization of infectious illnesses involves overcoming restrictions, optimizing model hyperparameters, and creating a unified platform for thorough disease identification. In the future, ways to improve accuracy and reliability in illness prediction and prognosis will be refined as techniques using deep learning for predicting & prognosis of infectious diseases are explored.

References

- Santangelo OE, Gentile V, Pizzo S, Giordano D, Cedrone F. Machine learning and prediction of infectious diseases: a systematic review. 2023. Available from: <https://doi.org/10.3390/make5010013>.
- Agrebi S, Larbi A. Use of artificial intelligence in infectious diseases. Academic Press. 2020. Available from: <https://doi.org/10.1016/B978-0-12-817133-2.00018-5>.
- Thakur K, Kaur M, Kumar Y. A Comprehensive Analysis of Deep Learning-Based Approaches for Prediction and Prognosis of Infectious Diseases. *Archives of Computational Methods in Engineering*. 2023;30(7):4477–4497. Available from: <https://doi.org/10.1007/s11831-023-09952-7>.
- Çinar A, Yildirim M. Classification of Malaria Cell Images with Deep Learning Architectures. *International Information and Engineering Technology Association*. 2020;25(1):35–39. Available from: <https://doi.org/10.18280/isi.250105>.
- Irmak E. COVID-19 disease severity assessment using CNN model. *IET Image Processing*. 2021;15(8):1814–1824. Available from: <https://doi.org/10.1049/ipr2.12153>.
- Callejon-Leblic MA, Moreno-Luna R, Cuvillo AD, Reyes-Tejero IM, Garcia-Villaran MA, Santos-Peña M, et al. Loss of Smell and Taste Can Accurately Predict COVID-19 Infection: A Machine-Learning Approach. *Journal of Clinical Medicine*. 2021;10(4):570. Available from: <https://doi.org/10.3390/jcm10040570>.
- Feng K, He F, Steinmann J, Demirkiran I. Deep-learning Based Approach to Identify Covid-19. *SoutheastCon 2021*. 2021;p. 1–4. Available from: <https://doi.org/10.1109/SoutheastCon45413.2021.9401826>.
- Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mazhar R, et al. Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization. *IEEE Access*. 2020;8:191586–191601. Available from: <https://doi.org/10.1109/ACCESS.2020.3031384>.
- Leo J, Luhanga E, Michael K. Machine Learning Model for Imbalanced Cholera Dataset in Tanzania. *The Scientific World Journal*. 2019;2019:1–12. Available from: <https://doi.org/10.1155/2019/9397578>.
- Tran NK, Albahra S, May L, Waldman S, Crabtree S, Bainbridge S, et al. Evolving Applications of Artificial Intelligence and Machine Learning in Infectious Diseases Testing. *Clinical Chemistry*. 2021;68(1):125–133. Available from: <https://doi.org/10.1093/clinchem/hvab239>.
- Hossain MS, Sultana Z, Nahar L, Andersson K. An intelligent system to diagnose chikungunya under uncertainty. *Journal of Wireless Mobile Networks*. 2019. Available from: <https://doi.org/10.22667/JJOWUA.2019.06.30.037>.
- Wang M, Wei Z, Jia M, Chen L, Ji H. Deep learning model for multi-classification of infectious diseases from unstructured electronic medical records. *BMC Medical Informatics and Decision Making*. 2022;22(1):1–3. Available from: <https://doi.org/10.1186/s12911-022-01776-y>.
- Taresh MM, Zhu N, Ali TAA, Hameed AS, Mutar ML. Transfer Learning to Detect COVID-19 Automatically from X-Ray Images Using Convolutional Neural Networks. *International Journal of Biomedical Imaging*. 2021;2021:1–9. Available from: <https://doi.org/10.1155/2021/8828404>.
- Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, et al. Can AI Help in Screening Viral and COVID-19 Pneumonia? *IEEE Access*. 2020;8:132665–132676. Available from: <https://doi.org/10.1109/ACCESS.2020.3010287>.
- Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*. 2018;172(5):1122–1131. Available from: <https://doi.org/10.1016/j.cell.2018.02.010>.
- Munadi K, Muchtar K, Maulina N, Pradhan B. Image Enhancement for Tuberculosis Detection Using Deep Learning. *IEEE Access*. 2020;8:217897–217907. Available from: <https://doi.org/10.1109/ACCESS.2020.3041867>.

- 17) Agarwal V. Complete architectural details of all EfficientNet models. 2020. Available from: <https://towardsdatascience.com/complete-architectural-details-of-all-efficientnet-models-5fd5b736142>.
- 18) Punn NS, Agarwal S. Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks. *Applied Intelligence*. 2021;51(5):2689–2702. Available from: <https://doi.org/10.1007/s10489-020-01900-3>.
- 19) Nair K, Deshpande A, Guntuka R, Patil A. Analysing X-Ray Images to Detect Lung Diseases Using DenseNet-169 technique. *SSRN Electronic Journal*. 2022. Available from: <http://dx.doi.org/10.2139/ssrn.4111864>.
- 20) Kakkar B, Goyal M, Johri P, Kumar Y. Artificial Intelligence-Based Approaches for Detection and Classification of Different Classes of Malaria Parasites Using Microscopic Images: A Systematic Review. *Archives of Computational Methods in Engineering*. 2023;23:1–20. Available from: <https://doi.org/10.1007/s11831-023-09959-0>.
- 21) Ibrahim DM, Elshennawy NM, Sarhan AM. Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. *Computers in Biology and Medicine*. 2021;132:104348–104348. Available from: <https://doi.org/10.1016/j.compbiomed.2021.104348>.
- 22) Kaur S, Kumar Y, Koul A, Kamboj SK. A Systematic Review on Metaheuristic Optimization Techniques for Feature Selections in Disease Diagnosis: Open Issues and Challenges. *Archives of Computational Methods in Engineering*. 2023;30(3):1863–1895. Available from: <https://doi.org/10.1007/s11831-022-09853-1>.
- 23) Kaur I, Kumar Y, Sandhu AK, Ijaz MF. Predictive Modeling of Epidemic Diseases Based on Vector-Borne Diseases Using Artificial Intelligence Techniques. *Computational Intelligence in Medical Decision Making and Diagnosis*. 2023;p. 81–100. Available from: <https://doi.org/10.1201/9781003309451-5>.
- 24) Modi K, Singh I, Kumar Y. A Comprehensive Analysis of Artificial Intelligence Techniques for the Prediction and Prognosis of Lifestyle Diseases. *Archives of Computational Methods in Engineering*. 2023;24:1–24. Available from: <https://doi.org/10.1007/s11831-023-09957-2>.
- 25) Daid R, Kumar Y, Gupta A, Kaur I. An effective mechanism for early chronic illness detection using multilayer convolution deep learning predictive modelling. *2021 International Conference on Technological Advancements and Innovations (ICTAI)*. 2021;p. 649–652. Available from: <https://doi.org/10.1109/ICTAI53825.2021.9673393>.
- 26) Koul A, Bawa RK, Kumar Y. Artificial intelligence techniques to predict the airway disorders illness: a systematic review. 2023. Available from: <https://doi.org/10.1007/s11831-022-09818-4>.