

Deep Learning-Based Medical Image Classification Using Convolutional Neural Networks for Clinical Diagnosis

Dr. Bal Krishna Sharma

Professor

Department of Computer Science and Application

Mandsaur University, Mandsaur

drbks7426@gmail.com

Corresponding Author: * Bal Krishna Sharma (drbks7426@gmail.com)

Received: 28 September 2026 | Revised: 01 October 2026 | Accepted: 07 October 2026

Abstract—Computer-aided detection of diseases using machine learning mechanisms on medical images has been an interesting applied research topic in both academia and the healthcare sector. Practical studies aimed at improving clinical decision-making through accurate medical image classification can significantly assist medical professionals in disease diagnosis. This paper presents an investigation into the classification of chest X-ray images for clinical diagnosis using a proposed Convolutional Neural Network (CNN). The Chest X-Ray Images (Pneumonia) dataset containing 5,863 images was utilized, and preprocessing techniques including removal of corrupted and duplicate images, denoising, resampling, label encoding, data standardization, GLCM-based feature extraction, and SMOTE-based class balancing were applied to enhance data quality and model performance. The dataset was divided using a stratified 80:20 train-test split, and the proposed CNN model was trained to classify images into Normal and Pneumonia classes. Experimental results demonstrate that the proposed CNN achieved an accuracy of 99.5%, precision of 99.2%, recall of 99.6%, and an F1-score of 99.4%, outperforming ResNet50 (91.55%), InceptionV3 (94.70%), and VGG16 (96.75%). The findings indicate that the proposed CNN provides highly accurate and reliable medical image classification, making it a promising decision-support tool for early pneumonia detection and improved clinical diagnosis.

Keywords—Medical image classification, Computer-aided diagnosis, Artificial intelligence, Machine learning, Deep learning, Chest X-Ray.

I. INTRODUCTION

Health is a fundamental prerequisite for individual well-being and societal development, while effective healthcare systems rely on timely disease detection, accurate diagnosis [1], and appropriate clinical decision-making to improve patient outcomes. Clinical diagnosis is a fundamental component of modern healthcare [2][3], enabling physicians to identify diseases, determine their severity, and formulate appropriate treatment strategies [4]. Medical imaging techniques such as X-ray, Computed Tomography (CT),

Magnetic Resonance Imaging (MRI), ultrasound, and histopathological imaging provide detailed anatomical and functional information that supports accurate disease detection [5][6][7]. Medical image analysis has become indispensable in diagnosing conditions such as cancer, pneumonia, cardiovascular disorders, neurological diseases [8][9][10], and gastrointestinal abnormalities. However, the rapid increase in the volume and complexity of medical imaging data has made manual interpretation by radiologists increasingly time-consuming and susceptible to human error and inter-observer variability [11], highlighting the need for intelligent computer-aided diagnostic systems that can improve diagnostic accuracy, speed, and consistency [12][13].

Machine learning (ML) has emerged as an effective computational approach for assisting medical image analysis by learning meaningful patterns from imaging data [14]. Traditional ML-based medical image classification generally involves multiple stages, including image preprocessing, handcrafted feature extraction, feature selection, and classifier training [15]. Features such as texture, shape, intensity, color, and statistical characteristics are manually extracted to represent image content before classification. Although these approaches have demonstrated satisfactory performance in several clinical applications, their effectiveness depends heavily on the quality of handcrafted features and expert knowledge [16][17], limiting their ability to capture complex and high-dimensional image representations required for accurate disease diagnosis [18][19].

Recent advances in artificial intelligence have accelerated the adoption of deep learning (DL) for medical image classification. Unlike conventional ML approaches, DL automatically learns hierarchical feature representations directly from raw medical images, eliminating the need for manual feature engineering [20][21][22]. Among various DL architectures, Convolutional Neural Networks (CNNs) have become the most widely adopted models due to their ability to perform convolution, pooling, and nonlinear feature extraction across multiple layers. CNN-based models have

demonstrated superior performance in disease classification, tumor detection [23][24], organ segmentation, and lesion identification across diverse medical imaging modalities, achieving higher accuracy and robustness than conventional image analysis techniques [25][26].

A. Motivation and Contribution

The motivation for this study is the increasing need for accurate and automated medical image classification to support timely and reliable clinical diagnosis. Traditional image analysis methods often struggle with complex medical images and require extensive manual feature extraction, leading to reduced efficiency and diagnostic accuracy. Convolutional Neural Networks (CNNs) provide an effective solution by automatically learning discriminative features from medical images. Therefore, this work aims to develop a robust CNN-based classification model that improves diagnostic performance and supports intelligent clinical decision-making. This research offers several key contributions as listed below:

- Developed a Convolutional Neural Network (CNN)-based framework for accurate medical image classification to support clinical diagnosis.
- Utilized the Chest X-Ray Images (Pneumonia) dataset for automated classification of Normal and Pneumonia chest X-ray images.
- Applied Grey Level Co-occurrence Matrix (GLCM) for extracting informative texture features from chest X-ray images.
- Addressed the class imbalance problem using the SMOTE technique, improving the model's robustness and reducing prediction bias.
- Evaluated the proposed CNN using standard performance metrics, including accuracy, precision, recall, and F1-score.
- Demonstrated superior classification performance by outperforming ResNet50, InceptionV3, and VGG16, providing more accurate and reliable pneumonia detection from chest X-ray images.
- Proposed an efficient and reliable computer-aided diagnosis (CAD) system that can assist healthcare professionals in early and accurate clinical decision-making.

The increasing demand for accurate and automated medical image analysis justifies the development of an efficient deep learning framework for clinical diagnosis. The novelty of this work is the integration of the three types of image preprocessing: comprehensive image pre-processing, texture feature extraction using GLCM, and class balancing using SMOTE into a single classification framework along with a CNN model. This combination increases the feature representation, mitigates class imbalance effect and better classifies chest X-ray images. The proposed framework provides an effective computer-aided diagnosis system for distinguishing Normal and Pneumonia cases with high reliability.

B. Organization of the Paper

The paper is organized into five sections: Section II reviews related literature, Section III describes the proposed methodology, Section IV discusses the experimental results and comparative analysis, and Section V concludes the study with future research directions.

II. LITERATURE REVIEW

A comprehensive review of recent studies on Medical Image Classification for Clinical Diagnosis was conducted to support this research. Table I summarizes the proposed models, datasets, key findings, and research challenges.

Liu et al. (2026) propose the Cross-Supervision Similarity Network (CSSN) for medical image classification on imbalanced small datasets with limited categories. CSSN transforms classification into a similarity comparison task using cross-supervision between class and patch features. Soft logarithmic supervision balances the two similarity branches during training. Experimental results on the PCR-ISD dataset demonstrate improvements of 15% in F1-score, 6% in accuracy, and 9% in balanced accuracy, highlighting its effectiveness in identifying minority classes and enhancing classification performance [27].

Wu et al. (2025) examine three scenarios: a novel CLIP variant with CNN and ViT image encoders for brain and skin cancer classification, federated learning with 12 deep models for privacy preservation, and traditional machine learning to improve generalization. Experimental results show MaxViT achieved the highest average accuracy (87.03%) on the HAM10000 dataset, ConvNeXt_1 attained an F1-score of 83.98% in federated learning, and SVM improved the average test performance of the Swin Transformer series by approximately 2% on the ISIC2018 dataset [28].

Xiang (2025) approach employs adaptive feature tokenization, B-spline parameterized learnable activation functions, and hybrid loss formulation. Experiments on Diverticulosis, Neoplasm, Peritonitis, and Ureters datasets achieve 95%+ accuracy. Grad-CAM visualizations confirm clinically meaningful attention on diagnostically relevant regions, validating interpretability for clinical deployment. Their architecture demonstrates superior performance while maintaining computational efficiency suitable for real-world medical applications [29].

Mamasiddikovich et al. (2024) focus on the area of interest to apply machine learning algorithms in the diagnosis of UTIs using medical images. Using various ultrasound, CT, and MRI scans, they arranged a large dataset and applied preprocessing with CNN, which enabled them to accurately detect UTIs. The models got to approve the inherent precision of 92. show that the sensitivity of this method is 5%, which is higher than conventional methods of diagnosis [30].

Korkmaz and Soygazi (2024) investigate gastrointestinal image classification using transfer learning-based deep learning models on the Kvasir dataset. Eight CNN architectures, including VGG19, ResNet, EfficientNet, InceptionV3, DenseNet201, and Xception, are evaluated using accuracy, precision, recall, and F1-score. Experimental results show that VGG19 achieved the highest prediction accuracy of 88.6% [31].

Phine (2023) develops a deep learning-based system for pneumonia detection and classification from chest X-ray images using the VGG-19 model. The Kaggle chest X-ray dataset is evaluated under both balanced and unbalanced conditions. Experimental results achieve 86% accuracy on the unbalanced dataset and 94% accuracy on the balanced dataset [32].

TABLE I. RECENT STUDIES ON MEDICAL IMAGE CLASSIFICATION FOR CLINICAL DIAGNOSIS USING MACHINE LEARNING TECHNIQUES

Author	Proposed Work	Dataset	Results	Limitations & Future Work
Liu et al. (2026)	Cross-Supervision Similarity Network (CSSN) for imbalanced medical image classification	PCR-ISD	Improved F1-score (15%), accuracy (6%), and balanced accuracy (9%)	Requires validation on larger, multi-center datasets and diverse disease categories.
Wu et al. (2025)	CLIP-based multimodal learning with federated learning and ML integration	HAM10000, ISIC2018	MaxViT achieved 87.03% accuracy; ConvNeXt_1 F1-score 83.98%	Increased model complexity; further optimization and broader clinical validation are needed.
Xiang (2025)	Adaptive feature tokenization with hybrid deep learning architecture	Diverticulosis, Neoplasm, Peritonitis, Ureters	Achieved over 95% accuracy with improved interpretability	Performance should be evaluated on larger and more heterogeneous clinical datasets.
Mamasiddikovich et al. (2024)	CNN-based UTI diagnosis using multimodal medical images	Ultrasound, CT, MRI	Achieved 92% accuracy with improved sensitivity	Requires larger datasets and enhanced model generalization across imaging modalities.
Korkmaz & Soygazi (2024)	Transfer learning-based gastrointestinal image classification	Kvasir	VGG19 achieved 88.6% accuracy	Future work should improve robustness and evaluate advanced architectures on diverse datasets.
Phine (2023)	VGG19-based pneumonia detection from chest X-ray images	Kaggle Chest X-ray	Achieved 94% accuracy on balanced dataset	Requires validation on real-world clinical datasets and improved performance on imbalanced data.

III. RESEARCH METHODOLOGY

The proposed methodology utilizes the Chest X-Ray Images (Pneumonia) dataset, where the images undergo preprocessing steps including removal of corrupted and duplicate images, denoising, resampling, label encoding, data standardization, GLCM-based feature extraction, and SMOTE-based class balancing before being divided into training and testing sets using a stratified 80:20 split. A Convolutional Neural Network (CNN) is then trained to automatically learn discriminative features for classifying chest X-ray images into Normal and Pneumonia categories. Finally, the model performance is evaluated using accuracy, precision, recall, and F1-score to assess its effectiveness for reliable clinical diagnosis. Fig. 1 illustrates the proposed flowchart for Medical Image Classification for Clinical Diagnosis using deep learning. The following section presents a detailed explanation of each step involved in the proposed methodology:

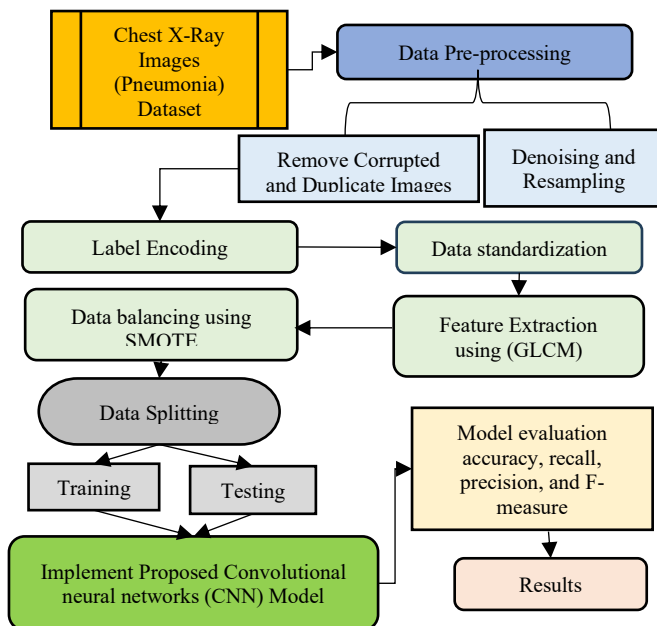


Fig. 1. Proposed Flowchart for Medical Image Classification for Clinical Diagnosis

A. Data Gathering and Analysis

This study utilizes the Chest X-Ray Images (Pneumonia) Dataset from Kaggle. The dataset comprises 5,863 chest X-ray images in JPEG format. The images belong to two categories: Pneumonia and Normal. Data visualizations such as bar plots and heatmaps were used to examine image distribution etc., are given below:



Fig. 2. Class Distribution Bar Plot

The comparative distribution of chest X-ray images between two classes, Normal and Pneumonia, is shown in Fig. 2. It displays there are significantly more images in the Pneumonia class than in the Normal class, so there is an imbalance in data and class balancing techniques can be used when training the model to help it make more accurate predictions.

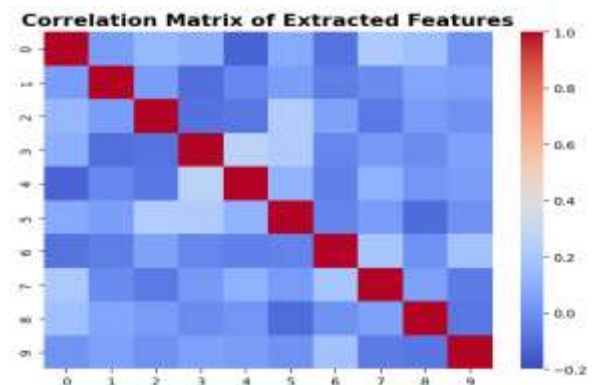


Fig. 3. Heatmap of Correlation Matrix

The heatmap in Fig 3 indicates the correlation between the extracted features, with values on the diagonal corresponding to a self-correlation of 1.0. The correlations between most of the off-diagonal elements are weak to moderate, suggesting that the features are not strongly correlated and provide distinct information. This low inter-feature correlation helps reduce redundancy and enhances the effectiveness of the classification model.

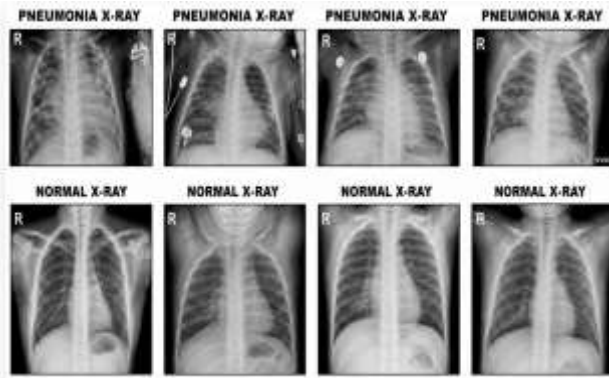


Fig. 4. Representative Chest X-Ray Images of Pneumonia and Normal Cases

Fig. 4 shows sample chest X-ray images from both the Pneumonia and Normal classes used in the dataset. Medical image classification is demonstrated by the visible lung opacities and abnormal infiltrates in the pneumonia images, and the clear lung fields in the normal images.

B. Data Pre-Processing

The Chest X-Ray Images (Pneumonia) Dataset was used for data preparation, which included image collection, cleansing, and feature extraction. The preprocessing pipeline consisted of removing corrupted and duplicate images, image denoising, resampling, label encoding, data labeling, and data standardization to improve data quality and ensure consistency for proposed model training. The main preprocessing procedures are summarized as follows:

- **Remove Corrupted and Duplicate Images:** Corrupted, unreadable, and duplicate medical images are identified and removed to eliminate data redundancy, prevent training errors, and ensure a clean, high-quality dataset for model development.
- **Denoising:** Noise can have an impact on medical imaging modalities, leading to unpredictable variations in image intensity. In order to diminish noise, pictures might undergo filtering in both the spatial and frequency domains.
- **Resampling:** This technique is used to change the size of specific image pixels or voxels without changing their location in the patient's coordinate system.
- **Label Encoding:** Categorical variables, including image categories and class labels, were transformed into numerical representations using label encoding to ensure compatibility with the deep learning model.

C. Data Standardization

Data standardization, sometimes referred to as data scaling, is an important data pre-processing method used in a variety of data analysis and deep learning activities. In the proposed work, data standardization aims to convert a

dataset's features into a common scale (between 0 and 1) to improve the accuracy of the model. The relationship used to achieve standardization is given as Equation (1):

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Here, x represents the original feature value, μ denotes the mean of all feature values, σ is the standard deviation that measures the spread of the data, and z is the standardized value obtained after scaling. This transformation ensures that the data have a mean of zero and a standard deviation of one, improving the stability and performance of the deep learning model during training.

D. Feature Extraction Using Grey Level Co-Occurrence Matrix (GLCM)

The goal of feature extraction in image processing is to lessen computational complexity by removing pointless and unneeded properties. The feature extraction technique determines the characteristics of the images using microanalysis and macroanalysis. Feature extraction is performed using the Grey Level Co-occurrence Matrix (GLCM) to obtain meaningful texture information from chest X-ray images. GLCM analyzes the spatial relationship between neighboring pixels and extracts discriminative texture features such as contrast, correlation, energy, and homogeneity. These features effectively represent image patterns while reducing data dimensionality and computational complexity.

E. Data Balancing Using SMOTE

Data balancing is the process of addressing an imbalanced dataset, where one class has significantly fewer samples than the others, by adjusting the distribution of classes to improve the performance of machine learning and deep learning models. In the proposed work, the Synthetic Minority Over-sampling Technique (SMOTE) was employed to balance the class distribution by generating synthetic samples for the minority class instead of simply duplicating existing instances. This approach enhances the diversity of minority class samples, reduces classification bias toward the majority class, and enables the CNN model to learn representative features from all classes. Consequently, SMOTE improves the robustness, generalization capability, and overall classification accuracy of the proposed medical image classification system.

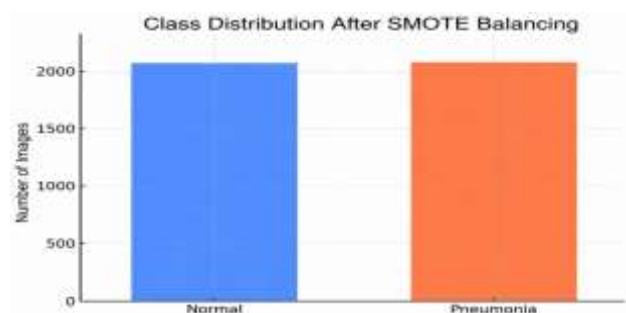


Fig. 5. Bar Graph of Class Distribution After SMOTE

Fig. 5 the bar chart of the class distribution after applying the SMOTE technique, where the Normal and Pneumonia classes contain an equal number of samples. This balanced distribution minimizes class imbalance, reduces prediction bias, and improves the reliability and accuracy of the classification model during training.

F. Data Splitting

The dataset was divided into training and testing sets using a stratified 80:20 split. This approach ensured that the class distribution in both the training and testing sets remained proportional to that of the original dataset, thereby preserving class balance and improving the reliability of model evaluation.

G. Proposed Convolutional Neural Networks (CNN) Model

Convolutional neural networks (CNN) perform feature learning via non-linear transformations implemented as a series of nested layers. The input data is a multidimensional data array, called tensor, that is presented at the visible layer. This is typically a grid-like topological structure, e.g., time-series data, which can be seen as a 1D grid taking samples at regular time intervals, pixels in images with a 2D layout, a 3D structure of videos, etc. Then a series of hidden layers extracts several abstract features. For instance, given a two-dimensional input x , a two-dimensional kernel h computes the 2D convolution by Equation (2)

$$(x * h)_{i,j} = x[i,j] * h[i,j] = \sum_n \sum_m x[n,m] \cdot h[i-n][j-m] \quad (2)$$

the dot product between their weights and a small region they are connected to in the input.

After the convolution, a bias term is added and a point-wise nonlinearity g is applied, forming a feature map at the filter output. Let denote the l -th feature map at a given convolutional layer as h^l , whose filters are determined by the coefficients or weights W^l , the input x and the bias b_l , then the feature map h^l is obtained as follows

$$h^l_{i,j} = g(W^l * x)_{ij} + b_l \quad (3)$$

where $*$ is the 2D convolution defined by Equation (3), while $g(\cdot)$ is the activation function.

Common activation functions encountered in deep neural networks are the rectifier, which is defined in Equation (4)

$$g(x) = x^+ = \max(0, x), \quad (4)$$

The proposed Convolutional Neural Network (CNN) was trained using 10 epochs with a batch size of 32, a learning rate of 0.001, and the Adam optimizer. The model employed the Binary Cross-Entropy loss function, ReLU activation in the hidden layers, and a Sigmoid activation function in the output layer for binary classification. Chest X-ray images were resized to 224×224 pixels, a dropout rate of 0.5 was applied to reduce overfitting, and a 20% validation split was used during training to evaluate model performance.

H. Evaluation Metrics

In this study, the trained models were evaluated using the following performance metrics: accuracy, recall, precision, and F-measure score. These evaluation metrics are calculated based on the following parameters and equations:

- **True Positive (TP):** If the input image has findings and Classified as finding a class.
- **True Negative (TN):** If the input image has no findings and classified as no finding class.
- **False Positive (FP):** In this case, the input image has no findings but is classified as a finding.
- **False Negative (FN):** In this case, the input image has findings but is classified as no finding.

With the help of these four parameters, can calculate

recall, precision, and F-measure score:

Accuracy: The ratio of the number of instances correctly predicted by the trained model to the total number of instances in the dataset (input samples). It is given as Equation (5)

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

Precision: Precision is the proportion of positive instances successfully predicted to all positive instances predicted by the model. Precision indicates. How good the classifier is in predicting the positive classes is expressed as Equation (6)

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

Recall: This metric, the ratio of events that were accurately predicted as positive to all instances that should have proved positive. In mathematical form it is given as Equation (7).

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

F1 score: It is a combination of the harmonic mean of precision and recall, that is, it helps to balance recall and precision. Its range is $[0, 1]$. Mathematically, it is given as Equation (8).

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

IV. RESULTS AND DISCUSSION

This section describes the experimental setup and highlights the performance of the proposed model across both training and testing phases, demonstrating its evaluation and efficient processing. The experiments were conducted on a high-performance PC with an NVIDIA RTX 3090 GPU, an Intel Core i9-12900K CPU, and 64GB of RAM. The models were implemented using PyTorch, with CUDA 11.7 for GPU acceleration, using Python libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Scikit-learn for system design and analysis. Table II presents the classification results of the proposed CNN model for medical image classification. The proposed model achieved an accuracy of 99.5%, a precision of 99.2%, a recall of 99.6%, and an F1-score of 99.4%, demonstrating excellent classification capability and balanced predictive performance. These results indicate that the proposed CNN model is highly accurate, reliable, and effective for medical image classification, making it a suitable approach for supporting automated clinical diagnosis.

TABLE II. CLASSIFICATION RESULTS OF PROPOSED MODEL FOR MEDICAL IMAGE CLASSIFICATION

Matrix	Testing	Training
Accuracy	99.5	100
Precision	99.2	100
Recall	99.6	100
F1-score	99.4	100

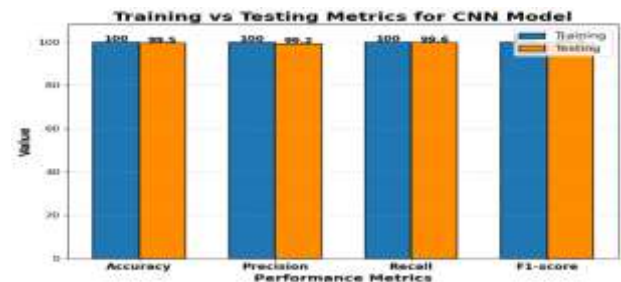


Fig. 6. Train and Test Performance Comparison of Proposed Model for Medical Image Classification for Clinical Diagnosis

Fig. 6 shows the accuracy, precision, recall and F1-score results evaluated the training and testing performance of the proposed CNN model for medical image classification and clinical diagnosis. During the training phase, a perfect score of 100% was obtained for all measures, whereas during the testing phase an accuracy of 99.5%, a precision of 99.2%, a recall of 99.6% and an F1-score of 99.4% were obtained. The minimal difference between the training and testing results indicates excellent generalization capability with negligible overfitting, demonstrating that the proposed model provides highly accurate and reliable performance for clinical diagnosis using medical images.

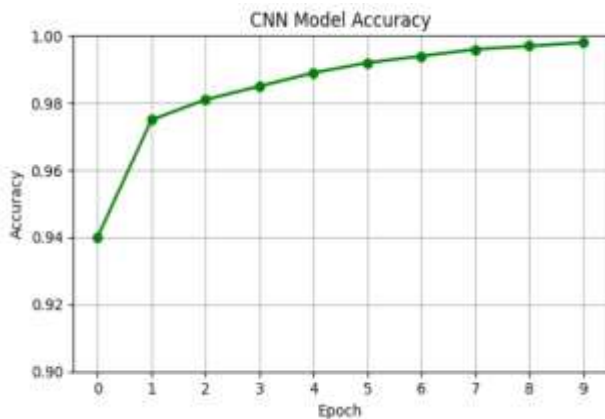


Fig. 7. Accuracy Curve of the Proposed CNN Model

Fig. 7 illustrates the accuracy curve of the proposed CNN model over 10 training epochs. The model accuracy increases steadily from approximately 94% in the first epoch to nearly 99.8% in the final epoch, demonstrating rapid learning and effective feature extraction. The stable convergence at the end of training indicates that the CNN model achieves high classification performance with excellent predictive capability.

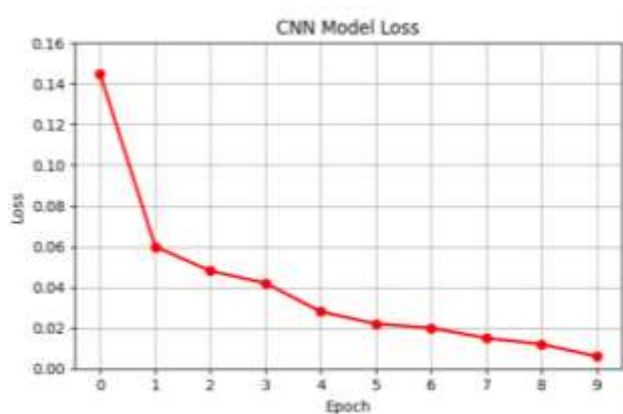


Fig. 8. Loss Curve of the Proposed CNN Model

Fig. 8 presents the loss curve of the proposed CNN model across 10 training epochs. The loss decreases consistently from approximately 0.145 to 0.006, indicating that the model progressively minimizes prediction errors during training. The smooth downward trend and near-zero final loss confirm efficient optimization, fast convergence, and improved generalization performance of the CNN model.

A. Comparative Analysis

To assess the effectiveness of the proposed model, a comparative accuracy evaluation against other existing

models is provided in Table III. The comparison includes ResNet50, InceptionV3, VGG16, and the proposed CNN using accuracy, precision, recall, and F1-score as performance metrics. ResNet50 achieved 91.55% accuracy, 92.08% precision, 92.10% recall, and 91.11% F1-score, while InceptionV3 obtained 94.70% accuracy, 85.67% precision, 96.52% recall, and 90.77% F1-score. VGG16 demonstrated improved performance with 96.75% accuracy, 96.00% precision, 97.00% recall, and 96.45% F1-score. In comparison, the proposed CNN outperformed all existing models, achieving 99.5% accuracy, 99.2% precision, 99.6% recall, and 99.4% F1-score. These results demonstrate that the proposed CNN provides superior classification performance, excellent feature learning capability, and enhanced reliability for medical image classification and clinical diagnosis.

TABLE III. COMPARISON OF DIFFERENT MACHINE LEARNING AND DEEP LEARNING MODELS FOR MEDICAL IMAGE CLASSIFICATION FOR CLINICAL DIAGNOSIS

Model	Accuracy	Precision	Recall	F1-score
ResNet50 [33]	91.55	92.08	92.1	91.11
Inceptionv3 [34]	94.70	85.67	96.52	90.77
VGG16 [35]	96.75	96	97	96.45
Proposed CNN	99.5	99.2	99.6	99.4

The proposed CNN model achieved an excellent 99.5% accuracy, outperforming the existing models and demonstrating its effectiveness in medical image classification for clinical diagnosis. The classification accuracy is high, suggesting that the model can effectively learn high-level image features and correctly classify the disease into different categories. This enhanced performance makes computer-aided diagnosis more reliable and decreases the misclassification errors. Therefore, the proposed CNN model provides a highly accurate and efficient solution for supporting clinical decision-making.

V. CONCLUSION AND FUTURE STUDY

Recent years have provided extremely interesting and exciting developments and applications of Machine Learning (ML) and Deep Learning (DL) to automate healthcare delivery. ML techniques are surpassing human performance in detecting, predicting, diagnosing diseases, and even recommending appropriate treatments. ML, together with Deep Learning, has revolutionized medical imaging technology. Based on the experimental results, the comparative analysis clearly shows that the proposed CNN model outperforms existing deep learning models, achieving an accuracy of 99.5%, compared to 96.75% for VGG16, 94.70% for InceptionV3, and 91.55% for ResNet50. The superior performance of the proposed model demonstrates its ability to accurately classify chest X-ray images into Normal and Pneumonia categories, making it a reliable and effective approach for clinical diagnosis and supporting early disease detection in real-world healthcare applications.

The proposed study is limited to the Chest X-Ray Images (Pneumonia) dataset and binary classification of Normal and Pneumonia cases, which may affect its generalizability to other medical imaging tasks. Future studies will be conducted on larger multi-center datasets, with the aim of extending the model to multi-class disease classification and implementing more sophisticated deep learning architectures to further enhance the robustness and applicability of the model in clinical practice.

- **Funding:** This research received no external funding.
- **Institutional Review Board Statement:** Not applicable.
- **Informed Consent Statement:** Not applicable.
- **Data Availability Statement:** The data supporting the findings of this study are available from the corresponding author upon reasonable request.
- **Conflicts of Interest:** The author declares no conflict of interest.

How to cite this article:

B. K. Sharma, "Deep Learning Based Medical Image Classification Using Convolutional Neural Networks for Clinical Diagnosis," *Journal of Global Research in Multidisciplinary Studies*, vol. 2, no. 10, pp. 01–08, Oct. 2026, Doi: XXXX.

REFERENCES

- [1] S. Mukherjee, "An Effective System for Medical Image Diagnosis Using Deep Convolutional Networks (CNNs) in Healthcare Sector," in *2026 14th International Symposium on Digital Forensics and Security (ISDFS)*, Boston, MA, USA: IEEE, Mar. 2026, pp. 01–06. doi: 10.1109/ISDFS69419.2026.11459010.
- [2] A. Warriar, "RESTful API Design Patterns for HIPAA-Compliant Healthcare Data Exchange," Zenodo, 2019. doi: doi.org/10.5281/zenodo.17251112.
- [3] R. Lingam, "Integrating Trustworthiness Into the AI Lifecycle (Trustworthiness)," in *AI Safety and Preventing Harm in AI Systems*, IGI Global Scientific Publishing, 2026, pp. 213–248. doi: 10.4018/979-8-3373-6935-8.ch008.
- [4] K. K. Mohammed, "A Survey on Digital Health Care Data Analysis Techniques for Developing Machine Learning Models," *Int. J. Softw. Eng. Technol.*, vol. 13, no. 5, pp. 158, October, 2025, doi: 10.5281/zenodo.17339813.
- [5] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, and P. Kumar, "Artificial intelligence to deep learning: machine intelligence approach for drug discovery," *Mol. Divers.*, vol. 25, no. 3, pp. 1315–1360, Aug. 2021, doi: 10.1007/s11030-021-10217-3.
- [6] Ramineni Snehamrutha, "Integration of Analytical Techniques (HPLC & FTIR) for Enhanced Quality Assurance in Drug Formulation," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 3, no. 1, pp. 953–964, Aug. 2023, doi: 10.48175/IJARST-12750C.
- [7] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. Guevara Lopez, "Representation learning for mammography mass lesion classification with convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 127, pp. 248–57, Apr. 2016, doi: 10.1016/j.cmpb.2015.12.014.
- [8] S. Mahmud, K. K. Mohammed, V. R. Jain, and S. A. Shah, "Artificial Intelligence-Driven Predictive Models for Identifying Risk Factors of Chronic Diseases," in *2026 14th International Symposium on Digital Forensics and Security (ISDFS)*, Boston, MA, USA: IEEE, Mar. 2026, pp. 01–06. doi: 10.1109/ISDFS69419.2026.11459003.
- [9] B. Sahu, J. Mishra, S. Parveen, B. Das, A. Panigrahi, and A. Pati, "STIR-MO: An Enhanced Hybrid Cancer Predictive Model," in *2025 International Conference on Intelligent and Cloud Computing (ICoICC)*, Bhubaneswar, India: IEEE, 2025, pp. 1–6, May. doi: 10.1109/ICoICC64033.2025.11052206.
- [10] S. Gupta and D. K. Jayswal, "Enhanced Artificial Intelligence (AI)-Based Risk Prediction Framework for Cardiovascular Health Monitoring," in *2026 2nd International Conference on Artificial Intelligence Systems (AIS)*, San Antonio, TX, USA: IEEE, 2026, pp. 282–288, May. doi: 10.1109/AIS69919.2026.11621536.
- [11] A. Srivastava, "Securing PHI in Cloud-Based Healthcare Systems: Challenges, Frameworks, and Future Directions," *J. Comput. Anal. Appl.*, vol. 33, no. 2, pp. 2126–2133, Jan, May 2025, doi: 10.48047/jocaaa.2024.33.02.33.
- [12] S. V J and J. F. D, "Deep Learning Algorithm for COVID-19 Classification Using Chest X-Ray Images," *Comput. Math. Methods Med.*, vol. 2021, p. 9269173, Nov. 2021, doi: 10.1155/2021/9269173.
- [13] M. D. K. Hasan *et al.*, "Deep Learning Approaches for Detecting Pneumonia in COVID-19 Patients by Analyzing Chest X-Ray Images," *Math. Probl. Eng.*, vol. 2021, pp. 1–8, May 2021, doi: 10.1155/2021/9929274.
- [14] P. Ramaraj and M. Mittal, "Deep Learning Models For Automated Medical Imaging: Accuracy, Ethics, And Deployment Challenges," *Int. J. Artif. Intell. Deep Learn.*, vol. 2, no. 1, pp. 11–31, Jun. 2025, doi: 10.34218/IJAIDL_02_01_002.
- [15] L. Nwosu, X. Li, L. Qian, S. Kim, and X. Dong, "Calibrated Bagging Deep Learning for Image Semantic Segmentation: A Case Study on COVID-19 Chest X-ray Image," *PLoS One*, vol. 17, no. 11, p. e0276250, May 2022, doi: 10.1371/journal.pone.0276250.
- [16] R. Snehamrutha, "Role of In-Process Quality Control (IPQC) in Reducing Deviations during ANDA-Oriented Pharmaceutical Manufacturing," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 11, no. 6, pp. 554–568, Nov. 2025, doi: 10.32628/CSEIT251117150.
- [17] A. Warriar, "Federated Learning Architectures for Multi-Hospital Research Data Collaboration," *Int. J. Emerg. Res. Eng. Technol.*, vol. 1, no. 1, 2020, doi: 10.63282/3050-922X.IJERET-V1I1P110.
- [18] K. Usui, T. Yoshimura, S. Ichikawa, and H. Sugimori, "Development of Chest X-ray Image Evaluation Software Using the Deep Learning Techniques," *Appl. Sci.*, vol. 13, no. 11, p. 6695, May 2023, doi: 10.3390/app13116695.
- [19] A. Victor Ikechukwu and M. S, "CX-Net: an efficient ensemble semantic deep neural network for ROI identification from chest-x-ray images for COPD diagnosis," *Mach. Learn. Sci. Technol.*, vol. 4, no. 2, p. 025021, Jun. 2023, doi: 10.1088/2632-2153/acd2a5.
- [20] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, May 2016, doi: 10.1109/TMI.2016.2538465.
- [21] M. Aamir, Z. Rahman, U. A. Bhatti, W. A. Abro, J. A. Bhutto, and Z. He, "An automated deep learning framework for brain tumor classification using MRI imagery," *Sci. Rep.*, vol. 15, no. 1, p. 17593, May 2025, doi: 10.1038/s41598-025-02209-2.
- [22] R. Lingam, "Agentic AI and Federated Learning for Privacy-Preserving Dental Imaging," in *2026 Innovations in Machine, Engineering, and Digital Conference (IMED)*, Kota Kinabalu, Malaysia: IEEE, Mar. 2026, pp. 1–6. doi: 10.1109/IMED68921.2026.11484887.
- [23] S. Mukherjee, "Performance Evaluation of Deep Learning Models for Early Detection of Infectious Diseases in Healthcare Systems," in *SoutheastCon 2026*, Huntsville, AL, USA: IEEE, 2026, pp. 1–6, March. doi: 10.1109/SoutheastCon63549.2026.11475977.
- [24] S. A. D. Poovaiah and P. K. Pareek, "AI-Optimized Multimodal Neuroimaging Pipelines for Early Alzheimer's Disease Detection: A System-Level Perspective," in *Alzheimer's Disease - Diagnostics, Pathogenesis, and Treatment*, IntechOpen, 2026. doi: 10.5772/intechopen.1015774.
- [25] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.
- [26] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic Segmentation of MR Brain Images With a Convolutional Neural Network," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1252–1261, May 2016, doi: 10.1109/TMI.2016.2548501.
- [27] X. Liu, Y. Luo, Y. Yi, J. Zhang, X. Gao, and X. Zhuang, "Cross-Supervision Similarity Network for Medical Image Classification on Imbalanced Small Datasets," *IEEE Trans. Med. Imaging*, vol. 45, no. 5, pp. 2429–2442, May 2026, doi: 10.1109/TMI.2026.3656364.
- [28] Y. Wu, M. Owais, R. Kateb, and A. Chaddad, "Deep Modeling and Optimization of Medical Image Classification," in *2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*, IEEE, Apr. 2025, pp. 1–4. doi: 10.1109/ISBI60581.2025.10981184.
- [29] Y. Xiang, "Medical Image Classification Using U-KAN Architecture for Gastrointestinal Pathology Detection," in *2025 4th International Conference on Image Processing, Computer Vision and Machine Learning (ICICML)*, IEEE, Nov. 2025, pp. 1442–1447. doi: 10.1109/ICICML67980.2025.11333778.
- [30] S. R. Mamasiddikovich, M. Al-Farouni, N. Giri, T. Vignesh, F. Khan, and P. K. Rao, "Detection of Urinary Tract Infections in

-
- Medical Images Using Machine Learning Algorithms,” in *2024 International Conference on Advances in Computing, Communication and Materials (ICACCM)*, IEEE, Nov. 2024, pp. 1–4. doi: 10.1109/ICACCM61117.2024.11059039.
- [31] İ. Korkmaz and F. Soygazi, “Gastrointestinal Image Classification Using Deep Learning Architectures via Transfer Learning,” in *2024 Medical Technologies Congress (TIPTEKNO)*, IEEE, Oct. 2024, pp. 1–4. doi: 10.1109/TIPTEKNO63488.2024.10755310.
- [32] S. Phine, “Pneumonia Classification Using Deep Learning VGG19 Model,” in *2023 IEEE Conference on Computer Applications (ICCA)*, IEEE, Feb. 2023, pp. 67–71. doi: 10.1109/ICCA51723.2023.10181954.
- [33] H. Slimi, A. Balti, S. Abid, and M. Sayadi, “Trustworthy pneumonia detection in chest X-ray imaging through attention-guided deep learning,” *Sci. Rep.*, vol. 15, no. 1, p. 40029, Nov. 2025, doi: 10.1038/s41598-025-23664-x.
- [34] D. Zhang, F. Ren, Y. Li, L. Na, and Y. Ma, “Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network,” *Electronics*, vol. 10, no. 13, p. 1512, Jun. 2021, doi: 10.3390/electronics10131512.
- [35] Nisha, “Deep learning-based medical image classification using CNN and attention mechanisms for disease detection and diagnosis,” *Int. J. Comput. Artif. Intell.*, vol. 6, no. 1, pp. 187–194, Jan. 2025, doi: 10.33545/27076571.2025.v6.i1c.153.