Automated Chest X-Ray Projection Classification

Mustafa Melik Ayanoğlu, Emir Çil, Emir Ersel Bilgiç, Batu Arıbakır
Department of Electrical and Electronics Engineering,
Izmir Katip Celebi University, Izmir, Turkey

Abstract

Chest radiographs are among the most widely used imaging techniques in the evaluation of pneumonia conditions. In recent years, the interpretation of these images has increasingly involved the use of computer-aided diagnosis (CAD) systems, which aim to improve diagnostic accuracy and minimize inter-observer variability. Identification of the projection plane is essential for the performance of these systems, as variations in projection views—typically frontal affect the appearance of anatomical structures and pathological features. This study proposes a deep learning-based approach for the automatic classification of chest X-ray projections. Therefore, various CNN architectures and ViT models have been trained to detect anatomical differences and categorize the images into frontal views. Among the fine-tuned models using transfer learning, ShuffleNet_V2_x1_0 and EfficientNet-B0 yielded the best results on the Chest X-Ray Images (Pneumonia) dataset. The training process was conducted using the Adam optimization algorithm and the cross-entropy loss function. Evaluation metrics such as accuracy, sensitivity, specificity, and F1-score demonstrated the effectiveness of the proposed model in projection classification tasks.

Keywords

Projection, Image Classification, Chest X-ray, Convolutional Neural Networks, Vision Transformer.

I. INTRODUCTION

Contemporary medical imaging technologies play a significant role in early diagnosis and disease detection processes [1], [2]. In particular, chest X-ray images are widely used for the assessment of lung diseases, cardiac abnormalities, and breast tissue pathologies [3]–[9]. However, the interpretation of these images depends largely on the experience of radiologists, and variations in expertise among healthcare professionals can lead to time-consuming evaluation processes. In this context, this study aims to develop a deep learning-based artificial intelligence model for the automatic classification of chest X-ray images the projection direction of the X-ray.

Among chest X-ray imaging techniques, frontal projections allow the evaluation of anatomical structures from different angles. Frontal X-ray images provide a general view of the heart, lung fields, ribs, and diaphragm domes, while lateral X-ray images help distinguish overlapping structures in frontal projections, such as the posterior border of the heart, retrosternal, and retrocardiac spaces. Therefore, correct classification of the projection direction directly affects the clinical interpretability of the images.

The model developed in this study integrates the lightweight and efficient ShuffleNet_V2_x1_0 and EfficientNet-B0 architecture with transformer-based Vision Transformer (ViT) components to learn anatomical differences and perform projection classification tasks. ShuffleNet_V2_x1_0 and EfficientNet-B0, known for its depthwise separable convolutions and inverted residuals, achieved the highest performance scores among the evaluated architectures, making it the core of the proposed method. The model was trained on the Chest X-Ray Images (Pneumonia) dataset, utilizing pre-processing techniques such as contrast enhancement to emphasize anatomical details like organs and bone structures. This approach not only facilitates the identification of biological differences but also enhances the model's generalization by analyzing features obtained from different projection angles.

II. METHOD

In this section, the proposed CNN-based X-ray projection classification approach is presented as shown in 1 and 2. The methodology involves pre-processing the Chest X-Ray Images (Pneumonia) dataset, correcting misclassified projection labels, and training multiple deep learning models for automatic classification.

To enhance contrast, the Contrast Limited Adaptive Histogram Equalization method was applied. This technique improves local contrast, making anatomical structures more distinguishable and allowing the model to extract more effective features. The images processed with CLAHE were directly fed into CNN models, ensuring that the learning process was based on sharpened visual representations.

Low-level features, such as edges, textures, and fundamental geometric structures learned in the early convolutional layers, remain valid for medical imaging, enabling these models to achieve higher accuracy with less data. However, training deep learning models from scratch requires a large amount of labeled data, significantly increasing computational costs and the risk of

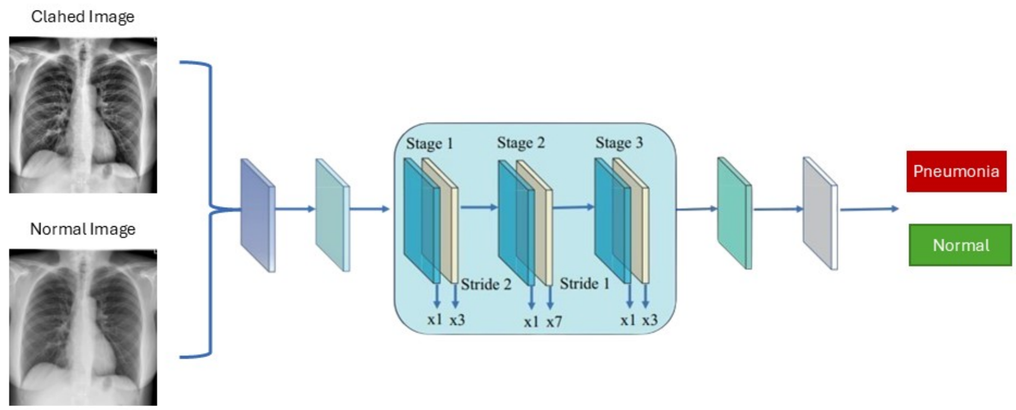

Figure 1: The Proposed ShuffleNetV2 Approach

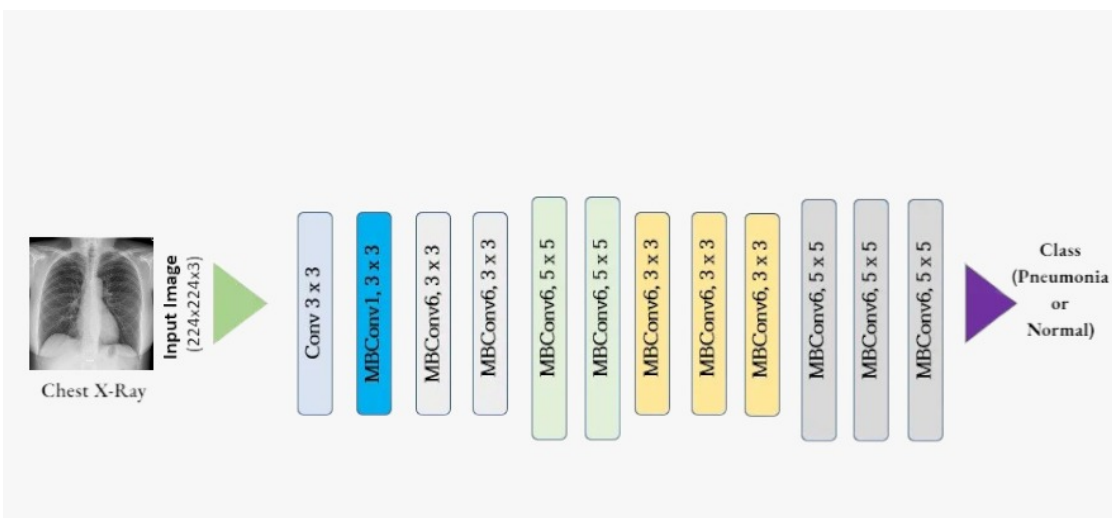

Figure 2: The Proposed EfficientnetB0 Approach

overfitting [10], [11]. Therefore, a transfer learning approach was adopted, utilizing models pre-trained on large-scale datasets to enable the network to learn general visual representations beforehand. Transfer learning facilitates an optimal initialization in the parameter space, thereby accelerating and stabilizing the gradient descent process [12], [13]. Various pretrained CNN models, including ShuffleNet_V2_x1_0 and EfficientNet-B0 were employed to classify the projection direction.

The performance of the trained models was evaluated using metrics such as accuracy, sensitivity, specificity, and F1-score, providing a comprehensive analysis of classification effectiveness. The developed model aims to automatically classify X-ray projections and reduce the workload of radiologists by accelerating the evaluation process. By comparing various CNN architectures, the most effective model for projection classification was determined, ensuring a more reliable classification in medical imaging workflows.

III. Experimental Evaluations

In this section, we evaluate the performance of our proposed approach for each radiology report, using the Chest X-Ray Images (Pneumonia) dataset for our experimental evaluations.

A. Dataset and Performance Metrics

Profile (projection) detection in medical imaging aims to accurately classify projection orientations in chest X-ray images. The dataset used in this study was obtained from the Chest X-Ray Images (Pneumonia) dataset, which contains chest X-ray images labeled with projection information. In addition to applying CLAHE preprocessing was also performed to enhance contrast and improve the model's feature extraction capability [14].

The performance of the proposed model was evaluated using various assessment metrics. Accuracy represents the proportion of correctly classified instances among all samples. Sensitivity (recall) measures the model's ability to correctly identify a specific projection type, while specificity indicates the extent to which incorrect classifications are avoided. The F1-score is calculated as the harmonic mean of precision and recall, providing a balanced evaluation of the model's performance [15].

Additionally, the area under the receiver operating characteristic curve (AUC-ROC) was computed to evaluate the model's ability to distinguish between different projection orientations [16]. A confusion matrix was also used to analyze misclassification patterns and investigate class-specific errors [17]. The final classification performance of the model was determined by averaging all evaluation metrics. Through this comprehensive evaluation approach, the reliability of the CNN-based profile detection model was systematically validated.

B. Results and Discussion

This section presents a comparative analysis of the classification performance of various Convolutional Neural Network (CNN)-based architectures and the transformer-based Vision Transformer model in identifying chest X-ray projection views. The models were evaluated under two preprocessing conditions: original images and contrast-enhanced images using CLAHE. The performance metrics considered include accuracy, sensitivity, specificity, precision, F1-score, and loss, and are summarized in Tables I.

The results demonstrate that CNN-based models consistently outperform the ViT architecture in all preprocessing scenarios. For original images, modern architectures such as ShuffleNetV2 and EfficientNetB0 achieved accuracies above 0.98*** and maintained a strong balance between sensitivity and specificity, as shown in Table I. In contrast, classical architectures like AlexNet and VGG16 yielded relatively lower sensitivity and precision, indicating limited ability to distinguish certain projection views.

As shown in , nearly all CNN-based models achieved near-perfect or perfect scores, with ShuffleNetV2 and EfficientNetB0 reaching 1.000 in accuracy, sensitivity, specificity, precision, and F1-score. This indicates that CNNs are highly effective in utilizing the improved contrast. However, the ViT model struggled under the same conditions, with substantially lower scores across all metrics. This outcome supports existing findings that transformer-based models often require larger and more balanced datasets to generalize effectively in medical imaging tasks.

While ViT achieved an accuracy of XXXXXX and an F1-score of XXXXXX, models such as ShuffleNetV2 and EfficientNetB0 maintained high classification performance with F1-scores exceeding 0.97***. In particular, classical CNNs showed inconsistent performance across preprocessing types, further emphasizing the importance of using up-to-date and optimized architectures for projection view classification.

Overall, the findings highlight the robustness and adaptability of CNN-based models, particularly lightweight and optimized architectures, across different preprocessing techniques. Although ViT presents potential, its current performance suggests that further optimization and significantly larger, more diverse datasets are necessary to match the reliability and efficiency demonstrated by CNNs in this domain.

Model	Loss	Accuracy	TP	FP	TN	FN	Sensitivity	Specificity	Precision	F1-Score
ShuffleNetV2_x1_0	0.047	0.981	779	8	1486	37	0.955	0.995	0.990	0.972
EfficientNet-B0	0.048	0.981	791	19	1475	25	0.969	0.987	0.977	0.973

Table I: Chest X-ray Projection Classification Performance Metric Results (Normal)

Model	Loss	Accuracy	TP	FP	TN	FN	Sensitivity	Specificity	Precision	F1-Score
ShuffleNetV2_x1_0	0.047	0.981	779	8	1486	37	0.955	0.995	0.990	0.972
EfficientNet-B0	0.048	0.981	791	19	1475	25	0.969	0.987	0.977	0.973

Table II: Chest X-ray Projection Classification Performance Metric Results (CLAHE)

IV. CONCLUSION

In this study, we evaluated the performance of various CNN architectures and a Vision Transformer model for classifying projection views of chest X-ray images using the Chest X-Ray Images (Pneumonia) dataset. Each model was trained and tested on two image variations: the original images and contrast-enhanced images using CLAHE. Our findings indicate that both preprocessing methods and model architecture significantly impact classification accuracy. Notably, the ViT model trained on normal images achieved the highest overall performance, suggesting that contrast enhancement can improve anatomical feature recognition. The results demonstrate that preprocessing algorithms have a decisive impact on medical image analysis and that convolutional and transformer-based architectures offer effective approaches in radiological classification tasks.

REFERENCES

[1] Y. Alaca, "Machine learning via darts-optimized mobilevit models for pancreatic cancer diagnosis with graph-based deep learning," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, pp. 1–21, 2025.

[2] S. Bharati, P. Podder, and M. R. H. Mondal, "Hybrid deep learning for detecting lung diseases from x-ray images," *Informatics in Medicine Unlocked*, vol. 20, p. 100391, 2020.

[3] N. R. S. Parveen and M. M. Sathik, "Detection of pneumonia in chest x-ray images," *Journal of X-ray Science and Technology*, vol. 19, no. 4, pp. 423–428, 2011.

[4] J.-I. Toriwaki, Y. Suenaga, T. Negoro, and T. Fukumura, "Pattern recognition of chest x-ray images," *Computer Graphics and Image Processing*, vol. 2, no. 3-4, pp. 252–271, 1973.

[5] D. A. Moses, "Deep learning applied to automatic disease detection using chest x-rays," *Journal of Medical Imaging and Radiation Oncology*, vol. 65, no. 5, pp. 498–517, 2021.

[6] V. Chang, V. R. Bhavani, A. Q. Xu, and M. Hossain, "An artificial intelligence model for heart disease detection using machine learning algorithms," *Healthcare Analytics*, vol. 2, p. 100016, 2022.

[7] N. Ghaffar Nia, E. Kaplanoglu, and A. Nasab, "Evaluation of artificial intelligence techniques in disease diagnosis and prediction," *Discover Artificial Intelligence*, vol. 3, no. 1, p. 5, 2023.

[8] A. Alqahtani, S. Alsubai, M. Sha, L. Vilcekova, and T. Javed, "Cardiovascular disease detection using ensemble learning," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 5267498, 2022.

[9] F. Yasmin, S. M. I. Shah, A. Naeem, S. M. Shujauddin, A. Jabeen, S. Kazmi, S. A. Siddiqui, P. Kumar, S. Salman, S. A. Hassan *et al.*, "Artificial intelligence in the diagnosis and detection of heart failure: the past, present, and future," *Reviews in cardiovascular medicine*, vol. 22, no. 4, pp. 1095–1113, 2021.

[10] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, pp. 1–40, 2016.

[11] S. Panigrahi, A. Nanda, and T. Swarnkar, "A survey on transfer learning," in *Intelligent and Cloud Computing: Proceedings of ICICC 2019, Volume 1*. Springer, 2021, pp. 781–789.

[12] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1, no. 2.

[13] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[14] S. M. Pizer, M. Amburn, J. Austin *et al.*, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.

[15] C. J. Van Rijsbergen, "Precision and recall," *Journal of Information Retrieval*, vol. 7, pp. 313–323, 2013.

[16] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[17] R. Kohavi, "Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid," *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 202–207, 1996.