Clinical Decision Support for Early Diagnosis of Cardiomegaly by Using Deep Learning Techniques on Chest X-rays

Erdem Yanar
Aselsan Inc /Gazi University
Ankara, Turkey
eyanar.bme@gmail.com

Fırat Hardalaç
Gazi University
Ankara, Turkey
firat@gazi.edu.tr

***Abstract*— Computer-aided diagnosis (CAD) systems have been widely researched and used in the medical field since their introduction in the 1960s. The system acts as an aid to radiologists in medical examinations using imaging technology such as X-rays, MRI, and CT scans to diagnose diseases and treat injuries. As the technology has evolved over the years, concepts such as neural networks have emerged and become increasingly popular in the field. In the field of neural networks, deep learning methods such as convolutional neural networks (CNNs) have shown promising and impressive results in many areas, including image recognition. Increased development has led to the use of CNNs in CAD systems as an enhancement to existing systems. A well-performing CAD system using a CNN relies on a well-performing CNN model. Obtaining a well-performing CNN model is not an easy task, as it depends heavily on finding the right hyperparameters and a sufficient data set. There are several approaches to finding good hyperparameters today, including hyperparameter tuning algorithms and manual trial and error. The aim of this thesis was to evaluate the effect of different configurations of hyperparameters on the performance of a specific CNN model. A promising and popular CNN model called VGGNet-16 was used in the study to diagnose a condition called cardiomegaly, where a patient suffers from an enlarged heart, using chest X-rays. To further improve these procedures, three different model was composed to create an ensemble model and we got a unique detection model. A dataset of X-rays labeled 'Cardiomegaly' and 'No findings' was used to train the model with different values of hyperparameters for each training session. The results showed that the learning rate and number of epochs had the greatest impact on the performance of the model and can therefore be considered as the most important hyperparameters. It was also found that a lower value of the learning rate generally resulted in higher performance compared to higher values and that the combination of a low learning rate and a low batch size is preferable to achieve a higher performance of the model. In conclusion, our new ensemble model got a comparable performance with respect to the literature.**

Keywords—Cardiomegaly; Deep Learning; Chest x-ray; Convolution Neural Network; CTR; VGG16; Resnet50; InceptionV3

I. INTRODUCTION

The heart is a vital organ in the human body, and its failure can affect other vital organs such as the brain and kidneys. The heart acts as a pump, circulating blood throughout the body, and if it stops beating, death can occur within minutes [1]. Figure 1 shows an image of a person with a normal heart and an image of a person with an enlarged heart.

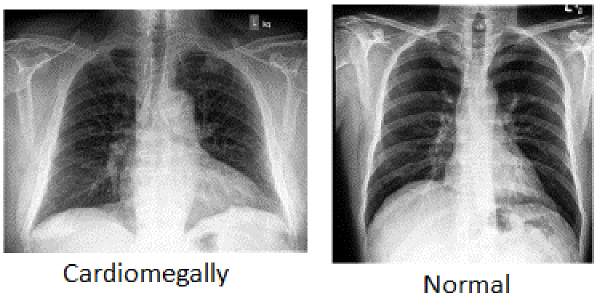

Figure 1 Chest X-rays of Cardiomegaly and Normal Cases[1]

Diagnosing heart abnormalities often involves a series of tests. Chest X-rays are a common imaging modality for the assessment of cardiomegaly due to their availability and low cost; radiologists typically use cardiothoracic ratio (CTR) values on anteroposterior chest X-rays. The cardiothoracic ratio (CTR) is a simple indicator used in clinical applications to assess heart size during manual examination, including typical conditions (CTR between 0.42 and 0.50), mild to moderate cardiomegaly (CTR between 0.50 and 0.60) and severe cardiomegaly (CTR above 0.60) [2].

CTR can be estimated using three different methods: [3].

A. MRD: represents the width of the space between the right heart and midline.

B. MLD: stands for the left midline diameter of the heart.

C. ID: represents the chest's internal diameter.

On the other hand, this manual technique requires a considerable amount of time and the results of the diagnosis depend on the reader's opinion and previous experience. However, manually marking organ boundaries and calculating CTR are time-consuming and labor-intensive processes that are prone to error. Recent advances in computer vision techniques and machine learning, specifically deep learning, use artificial neural networks to learn patterns and features in large amounts of data [4]. These algorithms have shown potential for automating the identification of cardiac abnormalities in radiological images, as they can be trained on labeled medical images to automatically detect features indicative of cardiomegaly and classify images based on the underlying cause [5].

This paper aims to utilize deep learning concepts such as convolutional neural networks (pre-processing the chest X-ray, creating a mask, and diagnosing the disease) and they were able to achieve successful predictions of this condition. The potential benefits and limitations of this approach are discussed, demonstrating the potential of deep learning algorithms in medical image analysis for accurate and efficient detection of cardiomegaly and its underlying causes.

II. METHODS

A. Datasets

"ChestX-ray8" is in fact an outpatient database of 108,948 frontal CXR images from 32,717 different individuals, with images classified for eight different diseases. Teams from the National Library of Medicine's Department of Medicine and Imaging Sciences and the National Institutes of Health's Bethesda headquarters have released "ChestX-ray8." an eliminator [6]. This dataset is open to the public for research and testing of different computer-based detection approaches.

We extract all cardiomegaly images from the Excel spreadsheets provided by the ChestX-ray8 dataset using a primitive Python script that separates files into multiple folders based on their label. As a result, we were able to extract 1010 cardiomegaly-related images. However, these 1024x1024 resolution images are provided without any masks or annotations to indicate the location of the disease. As a result, they need to be pre-processed before using the deep learning approach.

B. Data Preprocessing

The first step is to prepare the input image by resizing, enhancing, and removing noise. Medical imaging devices often introduce noise that can hinder accurate diagnosis of disease. It is therefore essential to clean the image by addressing issues such as noise, uneven lighting, blur, and other distortions. Many techniques are available to improve the quality of the input image, such as spatial or frequency filters.

CXRs always suffer from the limited contrast that can degrade the system performance, so the pre-processing step is very important for better performance of the designed system [7]. The raw images were first sent to the pre-processing step because some factors affect the image quality. In this step, the image quality is improved by eliminating the unwanted features and noise of the image and generating a more appropriate image for better segmentation and classification.

Here we used one of the most popular histogram equalization techniques called contrast limited adaptive histogram equalization (CLAHE). It is the most popular method of histogram equalization and an advanced version of AHE, which is beneficial to adjusting the image contrast and brightness and finally increasing the visibility by enhancing the hidden information that may contain meaningful details about the image.

C. Deep Learning Models

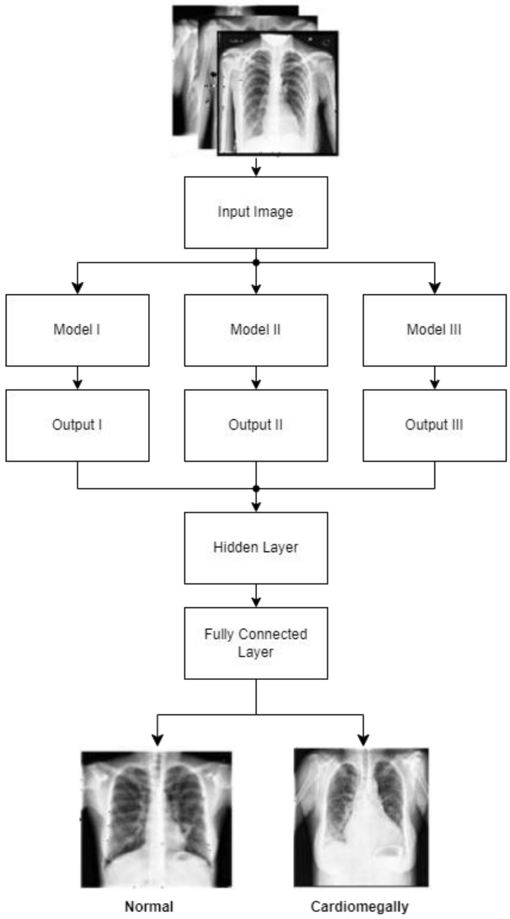

Figure 2 General Process of The Ensemble Model

Once a mask of the cardiomegaly chest X-ray has been created and the affected area has been localized, the next step is to diagnose the cardiomegaly. This is achieved by training different data samples on chest X-rays and developing a model capable of automatically diagnosing test data. Figure 2 shows the general flow of our ensemble model.

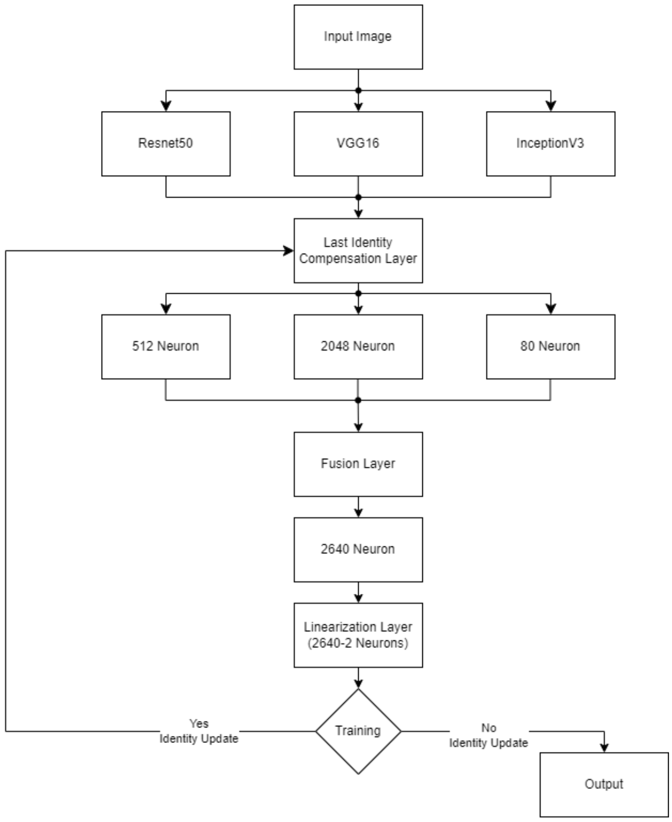

Figure 3 Details of The Ensemble Model

In Figure 3 we summarize the details of our ensemble model:

Step 1: The suggested ensemble model is formed as the final decision of the 3 pre-trained sub-models. These models' last layers are changed to the Identity layer.

Step 2: After this process, 512 for Resnet50 with Spinal Fully Connected Layer 512 for the final layer with 2048 output for VGG16, and 80 output for Inception with Spinal Fully Connected Layer is provided.

Step 3: These outputs are then combined to form a single linear layer and a hidden layer with 2640 outputs is obtained.

Step 4: Finally, this hidden layer is given a sigmoid activation function with an output of 2 connected classifier layers.

Step 5: The network created as a result of these operations is retrained and the results are obtained.

D. Metrics for Evaluation Performance

In this section, an overview of several evaluation techniques is provided, highlighting the importance of utilizing diverse approaches to assess model performance. Specifically, accuracy, recall, precision, and the F1 measure are discussed as effective methods for evaluating classification algorithms [8]. Additionally, the following terms are defined:

- True positive (TP), which pertains to instances of positive occurrences that are correctly classified.
- False negative (FN), which refers to instances of positive occurrences that are mistakenly classified.
- False positive (FP), which pertains to instances of negative occurrences that are erroneously classified.
- True negative (TN), which denotes instances of negative occurrences that are correctly classified.

Equation 1 Evaluation Metrics [8]

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

$$precision = \frac{tp}{tp + fp}$$

$$recall = \frac{tp}{tp + fn}$$

$$f1_{score} = \frac{2 \times precision \times recall}{precision + recall}$$

III. RESULTS

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar: Table 1 shows the comparison of our results with others in the literature:

Table 1Comparison of Model Performances

Author	Techniques	Dataset	Performance		
			Acc. (%)	Sens. (%)	Spec. (%)
Isarun Chamveha, et al. [9]	U-Net with VGG16 encoder	ChestX-ray8	69.8	0.69	0.70
Abdelilah Bouslama, et al. [10]	U-Net	ChestX-ray8	94	-	-
Qiwen Que1.et al[11]	CardioXNet	ChestX-ray8	93.75	-	-
Mohammed Innat et al. [8]	CXA_Dense121	ChestX-ray8	87	85	-
Muhammad Arsalan, et al.[13]	X-ray Net_2	JSRT	98.1	-	-
Chia-Hung Lin, et al.[14]	Multilayer 1D Convolutional Neural Network	NIH	98	-	-
Proposed Model	ELM (Ensemble Learning Model)	ChestX-ray8	94.12	96.2	92.5

Our trained model was able to classify the test data with an accuracy of 94.12%. The sensitivity and specificity for detection of cardiomegaly on chest X-ray were 96.2% and 92.5%.

IV. DISCUSSION AND CONCLUSION

This paper presents a comprehensive overview of cardiomegaly disease, including its classification algorithms and techniques employed in this field. The utilization of deep learning algorithms in detecting and diagnosing cardiomegaly through chest X-rays has exhibited significant promise. Employing sophisticated techniques and algorithms, such as convolutional neural networks. The aforementioned algorithms have yielded high efficacy and accuracy compared to traditional methods, with the potential to automate the detection process and significantly enhance patient outcomes via early identification and intervention. Similar studies were conducted in the literature, mostly binary classification.

In this study; basically, two classes (normal/abnormal). However, unlike the literature, instead of binary classification, multiclass classification was made. The reason for this is that the initial 26 classifications model's output for each class is evaluated and the ensemble is to determine the most suitable models to be used in the models. In this way, the best results in the classification study were obtained with ensemble models.

In this framework, we have designed CNN-based models for the automatic diagnosis of cardiomegaly from CXR images, which can be beneficial in healthcare planning and management strategies. CXR images always suffer from the limited contrast that can degrade the system performance, so the pre-processing step is essential for better performance of the designed system. The developed system follows 3 important steps, i.e. image quality enhancement by applying CLAHE to normalize the images, then data augmentation was applied to create the artificial samples from the actual dataset to avoid the overfitting, also generate more learning features to train the model and the last step was dedicated to data processing via our ensemble model form of custom CNN based models. We accessed a well-known publicly available dataset. Firstly, we extracted the cardiomegaly reports from the chest x-ray-8 and categorized them into two classes, i.e. cardiomegaly and normal. Out of the total CXR images, 80% were used for training and validation purposes and the remaining 20% were used for testing purposes. However, we trained the model with different epochs like 100, 120, 150 and 200. However, the model trained with 120 epochs shows a revolutionary result. Various performance parameters like precision, accuracy, specificity, sensitivity, recall, F1 score, confusion matrix, and loss graph were used to validate the accuracy of the system. The obtained accuracies of 100, 120, 150, and 200 epochs are 88.69%, 94.12%, 88.09%, and 86.64% respectively. We have comparable results with the literature.

In the future, more evolutionary algorithms can be used as feature selection techniques by combining with other classification models and deep learning approaches to achieve more accurate results with less computation time. This work can be extended by considering more diseases other than chest diseases with the large dataset.

ACKNOWLEDGMENT

The author would thank ASELSAN Inc. for supporting the academic research activities.

REFERENCES

[1] Cristina Barsanti, Francesca Lenzarini, and Claudia Kusmic, "Diagnostic and prognostic utility of non-invasive imaging in diabetes management," World Journal of Diabetes, vol. 6, pp.792 806, 2015. DOI: http://doi.org/10.4239/wjd.v6.i6.792

[2] https://www.mayoclinic.org/tests-procedures/x-ray/about/pac-20395303. last accessed 5 March 2022.

[3] Tarun Agrawal, and Prakash Choudhary, "Segmentation and classification on chest radiography: a systematic survey," The Visual Computer, Springer, 2022. DOI: https://doi.org/10.1007/s00371 021 02352 7

[4] Abdelilah Bouslama, Yassin Laaziz, and Abdelhak Tali, "Diagnosis and precise localization of cardiomegaly disease using U-NET," Informatics in Medicine Unlocked, Elsevier, 2020. DOI: https://doi.org/10.1016/j.imu.2020.100306

[5] Mu Sook Lee, Yong Soo Kim, Minki Kim, Muhammad Usman, Shi Sub Byon, Sung Hyun Kim, Byoung Il Lee, and Byoung-Dai Lee, "Evaluation of the feasibility of explainable computer-aided detection of cardiomegaly on chest radiographs using deep learning," Scientific Reports, Nature, 2021. DOI: https://doi.org/10.1038/s41598 021 96433 1

[6] Mohammad Tariqul Islam, Md Abdul Aowal, Ahmed Tahseen Minhaz, Khalid Ashraf, "Abnormality Detection and Localization in Chest X-Rays using Deep Convolutional Neural Networks", arXiv:1705.09850v3 [cs.CV, 2017.

[7] Fabián Torres-Robles, Alberto Jorge Rosales-Silva, Francisco Javier Gallegos-Funes and Ivonne Bazán-Trujillo, "A robust neuro-fuzzy classifier for the detection of cardiomegaly in digital chest radiographs," Dyna (Medellin, Colombia), pp. 35-41, 2014. DOI: http://dx.doi.org/10.15446/dyna.v81n186.37797

[8] Innat, M., Hossain, M.F., Mader, K. et al. A convolutional attention mapping deep neural network for classification and localization of cardiomegaly on chest X-rays. Sci Rep 13, 6247 (2023). https://doi.org/10.1038/s41598-023-32611-7

[9] I. Chamveha,T. Promwiset, T. Tongdee, P. Saiviroonporn, and W. Chaisangmongkon, "Automated Cardiothoracic Ratio Calculation and Cardiomegaly Detection using Deep Learning Approach," pp. 1–11, 2020, [Online]. Available:http://arxiv.org/abs/2002.07468. :doi.org/10.48550/arXiv.2002.07468

[10] Isarun Chamveha, Treethep Promwiset, Trongtum Tongdee, Pairash Saiviroonporn, and Warasinee Chaisangmongkon, "Automated Cardiothoracic Ratio Calculation And Cardiomegaly Detection Using Deep Learning Approach," arXiv:2002.07468v1 [eess.IV] 2020. DOI: https://doi.org/10.48550/arXiv.2002.07468

[11] A. Bouslama, Y. Laaziz, and A. Tali, "Diagnosis and precise localization of cardiomegaly disease using U-NET," Informatics Med. Unlocked, vol. 19, no. January 2021, 2020, doi: 10.1016/j.imu.2020.100306.

[12] Tawsifur Rahman, Amith Khandakar, Yazan Qiblawey....and Muhammad E.H. Chowdhury, "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images", Computers in Biology and Medicine, Elsevier, 2021. DOI: https://doi.org/10.1016/j.compbiomed.2021.104319

[13] C. H. Lin et al., "Posteroanterior Chest X-ray Image Classification with a Multilayer 1D Convolutional Neural Network-Based Classifier for Cardiomegaly Level Screening," Electron., vol. 11, no. 9, 2022, doi: 10.3390/electronics11091364.

[14] M. Arsalan, M. Owais, T. Mahmood, J. Choi, and K. R. Park, "Artificial intelligence-based diagnosis of cardiac and related diseases," J. Clin. Med., vol. 9, no. 3, 2020, doi: 10.3390/jcm9030871.