

Attention mechanism-based model for cardiomegaly recognition in chest X-Ray images

Sara El Omary¹, Souad Lahrache², Rajae El Ouazzani¹

¹Image Laboratory, Higher School of Technology, Moulay Ismail University of Meknes, Meknes, Morocco

²LabSIV, Faculty of Sciences, Ibnou Zohr University of Agadir, Agadir, Morocco

Article Info

Article history:

Received Aug 15, 2022

Revised Mar 9, 2023

Accepted Aug 2, 2023

Keywords:

Attention

Cardiomegaly

Cardiovascular diseases

Chest X-Ray

Convolutional neural networks

MobileNet

ABSTRACT

Recently, cardiovascular diseases (CVDs) have become a rapidly growing problem in the world, especially in developing countries. The latter are facing a lifestyle change that introduces new risk factors for heart disease, that requires a particular and urgent interest. Besides, cardiomegaly is a sign of cardiovascular diseases that refers to various conditions; it is associated with the heart enlargement that can be either transient or permanent depending on certain conditions. Furthermore, cardiomegaly is visible on any imaging test including Chest X-Radiation (X-Ray) images; which are one of the most common tools used by Cardiologists to detect and diagnose many diseases. In this paper, we propose an innovative deep learning (DL) model based on an attention module and MobileNet architecture to recognize Cardiomegaly patients using the popular Chest X-Ray8 dataset. Actually, the attention module captures the spatial relationship between the relevant regions in Chest X-Ray images. The experimental results show that the proposed model achieved interesting results with an accuracy rate of 81% which makes it suitable for detecting cardiomegaly disease.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Sara El Omary

Image Laboratory, Higher School of Technology, Moulay Ismail University of Meknes

Marjane, Meknes 50050, Morocco

Email: elomarysr@gmail.com

1. INTRODUCTION

Cardiovascular diseases (CVDs) harm more than 23 million people around the world, which makes heart diseases a principal health problem [1]. In the United Kingdom, CVDs are among the primary causes of sudden deaths and disability. CVDs refer to a variety of conditions that affect the heart system and blood arteries [1], [2]. In fact, there are no particular reasons behind CVDs, but there are many factors that can augment their risk of development. The following are some factors of CVDs: high blood pressure is the most critical factor, as it can cause damage to the blood vessel, cigarettes, cholesterol, diabetes, alcohol, unhealthy food, physical inactivity, obesity, and family medical history [3]. However, people may prevent CVDs by adopting a healthy lifestyle and making adjustments that reduce the risk of heart diseases. Besides, cardiomegaly is a type of CVDs that can be described as a medical condition where the heart becomes larger and often goes unnoticed until signs occur or the doctor orders imaging tests. Furthermore, the symptoms of cardiomegaly do not appear until it has reached a critical stage that is characterized by abnormal heartbeats, breathing problems, sensation of instability, fatigue, and swelling of certain parts of the body [4]. In addition,

these symptoms are caused by ventricular hypertrophy or dilatation; ventricular hypertrophy corresponds to a thickening of the ventricular wall, and ventricular dilatation to a thinning of the ventricular wall [5]. Doctors use various techniques to diagnose Cardiomegaly including Echocardiogram, Chest X-Ray, computerized tomography (CT) scan, and electrocardiogram [4]. The clinical diagnosis of chest radio-graphs can be difficult and challenging; actually, reading and interpreting a Chest X-Ray image and extracting key information is a difficult and time consuming task that requires an interesting doctors' experience and can sometimes lead to wrong results.

However, the existing approaches [6], [7] that have used Chest X-Ray images to detect Cardiomegaly have various limitations such as ignoring the spatial relationship between regions of interest (ROI) in an image. Actually, the analysis of ROI could improve the model performance. Hence, to overcome this restriction, we design an innovative attention-based deep learning (DL) model that uses an attention module with MobileNet to effectively capture critical regions, both at local and global levels, in Chest X-Ray images. Our proposed DL model presents a novel approach that seamlessly integrates the attention mechanism into the MobileNet framework, enabling more accurate identification of significant features and enhancing the model's performance for Chest X-Ray image analysis. In particular, the attention module plays a crucial role in capturing the most relevant and crucial regions within the Chest X-Ray images. Simultaneously, the MobileNet convolution module focuses on detecting the activated regions through the rectified linear unit (ReLU) function, employing a fixed kernel size. Moreover, the proposed model benefits from a reduced number of trainable parameters as it utilizes pre-trained weights from the MobileNet architecture. Additionally, its streamlined design eliminates the need for separate feature extraction and classification steps, commonly seen in traditional machine learning approaches. As a result, the model becomes more efficient and readily deployable for training and real-world applications.

The subsequent sections of the paper are structured as follows; the second section consists of related work. Then, the third section presents the methodology details, in particular data preprocessing and the convolutional neural network (CNN) architecture. Subsequently, the fourth section provides the results obtained from the implemented CNN architecture in addition to the results discussion. Finally, the fifth section encompasses the conclusion and outlines potential areas for future work.

2. RELATED WORK

Recently, machine and DL studies are widely investigated, as they have shown sophisticated results in many problems such as CVDs. For example, El Omary *et al.* [8] employed several CNNs for the purpose of detecting cardiac arrhythmia based on electrocardiogram (ECG) two-dimensional (2D) images; in addition, they [9] utilized a variety of pre-trained CNN models to diagnose heart failure in Radiograph images. Next, Yang *et al.* [10] introduced a model aiming at early heart failure diagnosis using a combination of Bayesian principal component analysis (BPCA) and support vector machine (SVM) resulting in an accuracy rate of 74.4%. Afterwards, Miao *et al.* [11] employed DL to devise a system that enhances the dependability and efficiency of CVDs diagnosis. Their approach involved a multi-layer model, leading to a recall rate of 72.86% and a sensitivity rate of 93.51%. Furthermore, Son *et al.* [12] presented a model specifically designed for early-stage diagnosis of heart failure in emergency rooms. They harnessed the potential of rough sets (RS) and decision trees (DT) techniques, and attained an accuracy value of 97.5%. Further, Bar *et al.* [13] utilized an image Net-based CNN architecture to identify various pathologies in Chest X-Ray images and obtained an accuracy rate of 89%. Then, Acharya *et al.* [14] suggested a CNN model using electrocardiogram (ECG) signals, and this model acquired an accuracy rate of 98.97%. Finally, Rubin *et al.* [15] proposed a new network called DualNet that analyzes frontal and lateral Chest X-Ray images from the MIMIC Chest X-Ray (MIMIC-CXR) dataset, and they procured an accuracy rate of 91%.

Actually, a single DL model may not be able to provide enough discriminative information for Chest X-Ray images classification [16]. Due to this major issue, many researchers utilized ensemble learning methods to train a set of algorithms to form robust models. Ensemble methods can be defined as a technique that uses a collection of models rather than a single model to significantly improve experimental results [17]. There are three primary types of ensemble methods including bagging, boosting, and stacking [17]. Besides, several studies have used ensemble methods, for example, Zhou *et al.* [18] employed a combination of various artificial neural networks (ANNs) to recognize lung cancer cells. Next, Sasaki *et al.* [19] utilized an ensemble model that can detect abnormalities in Chest X-Ray images. Meanwhile, Li *et al.* [20] used a variety of CNNs with Chest X-Ray images of lung nodules to reduce the rate of false positive. Additionally, Islam *et al.* [21] presented an ensemble model created by combining several different pre-trained DL models to detect lung nodules as well. Finally, Chouhan *et al.* [22] suggested a model that combines ResNet-18, DenseNet-121, AlexNet, GoogleNet, and Inception-V3 to diagnose pneumonia. However, ensemble learning methods still have some weaknesses such as overfitting due to the small amount of medical data. Moreover, ensemble

methods can be time and memory consuming, as they use a large number of parameters to extract key patterns from input images.

3. METHODOLOGY

3.1. Data description

Actually, Chest X-Ray images are among the most common and economical medical imaging procedures. National Institutes of Health (NIH) ChestX-ray8 is a public dataset containing various Chest X-Ray images [23]. The NIH Chest X-Ray8 includes various images of 14 diseases, in particular 112,120 Chest X-Ray images, including Atelectasis, Consolidation, Infiltration, Pneumothorax, Edema, Emphysema, Fibrosis, Effusion, Pneumonia, Pleural-thickening, Cardiomegaly, Chouhan *at al.* [22]. These diseases are labeled from 30,805 patients, and the authors utilized natural language processing (NLP) tools to extract and classify diseases using the associated radiology reports [24]. To classify patients with Cardiomegaly, we create two classes, the first one represents Cardiomegaly, and the second one groups the other diseases under the healthy class. Figure 1 exposes Chest X-Ray images of two distinct cases: one displaying a healthy patient and the other showing a patient diagnosed with Cardiomegaly.

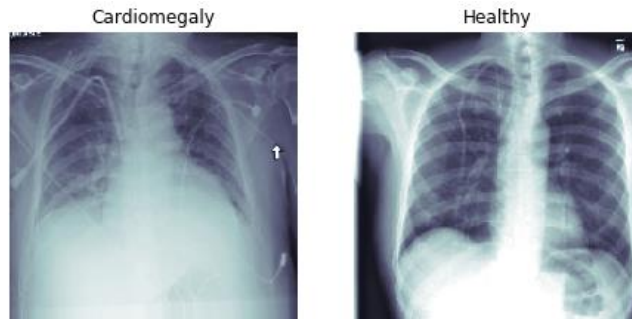


Figure 1. Healthy and Cardiomegaly patients' Chest X-Ray images [23]

3.2. Proposed architecture

We have proposed a model that combines the MobileNet and attention modules. MobileNet can be defined as a simplified DL architecture that creates lightweight deep CNN using depthwise separable convolutions and offers efficient models suitable for mobile and embedded vision applications [25]. MobileNet has many advantages, including reduced network size, fewer parameters, speed, and applicability to real-time applications [25]. In fact, the MobileNet model was chosen because it is among the five most accurate models and has a small kernel size that allows extraction of low-level features, which is suitable for Chest X-Ray images with fewer layers [26]. Moreover, the MobileNet provides an excellent feature extraction capability of Chest X-Ray image classification. Figure 2 shows the construction of MobileNet using depthwise separable filters [27].

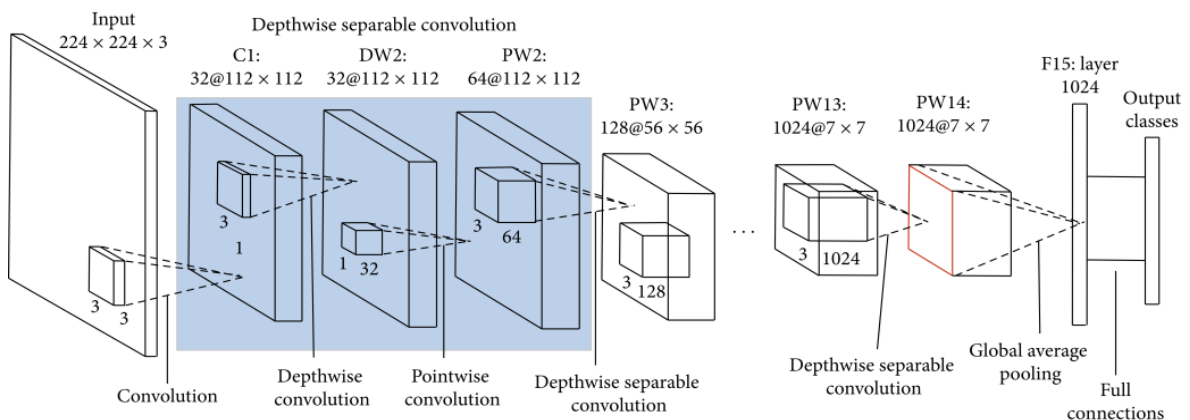


Figure 2. The MobileNet architecture [27]

Actually, depthwise separable convolution filters combine depthwise and pointwise convolution filters [27], [28]. There are two primary types of separable convolutions exist: spatial separable convolutions and depthwise separable convolutions [27]. First, the spatial separable convolution works mainly on the height and width of the images and the kernel is divided into smaller elements [27]. For example, a kernel of 3×3 might be divided into 3×1 and 1×3 [27]. Then, the depth separable convolution name is derived from its consideration of both the depth dimension (the number of channels) and the spatial dimensions [27]. An RGB image has 3 channels which are red, blue, and green [27]. Moreover, a depth-separable convolution divides a kernel into two distinct kernels that perform two convolutions which are the depth convolution and the pointwise convolution, as for the spatial separable convolution [27]. The pointwise convolution is a convolutional operation that utilizes a 1×1 kernel: a kernel that involves iterating through each point, and its depth is equivalent to the number of channels in the input image [27]. As illustrated in Figure 3, that contains subfigures 3(a) Standard convolutional filters, Figure 3(b) depthwise filters, and Figure 3(c) point filters. First, standard convolutional filters, which can be defined as small matrices of numerical values, that slide over an input image, performing a convolution operation at each location, to extract specific features from the image such. Second, depthwise filters refer to a type of filter that performs convolution independently on each channel of the input image, that can be defined as the depthwise convolution filter applies one convolution per input channel, while the point convolution filter linearly mixes the depthwise convolution result with 1×1 convolutions.

In the following, we will highlight the modules that mainly compose the proposed architecture. The architecture is composed of four building blocks which are the convolutional module, the attention module, the fully connected (FC), and classification layers. First, the attention module is employed to retain the spatial relationship of the visual information contained in the Chest X-Ray images. Next, the convolutional module is used to extract the main features figured in our input data using the convolutional layers of the MobileNet model, and then its output is given to the attention module. Further, we have the FC layers to concatenate the features produced by the convolutional and attention blocks into a 1D representation. Finally, the last dense layer is used to classify the input image as healthy or patient with Cardiomegaly disease using the sigmoid function. However, in Figure 2, the global average pooling layer is oversimplified and the input images have some regions that are more important than the others. Thus, we designed an attention mechanism to turn pixels on and off, and then rescale the results using the Lambda layer based on the pixels' amount. Furthermore, the attention layer is used to weight the processing of specific regions in the average pooling layer. Figure 4 and Figure 5 expose more details about the proposed architecture. Then, Figure 6 illustrates the layers of the entire model and each layer with its name, input vector, output vector, and how the components of the proposed architecture are related to each other.

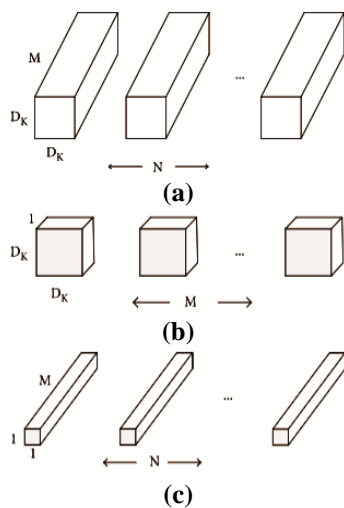


Figure 3. Illustration of different types of convolutional filters including: (a) standard convolutional filters, (b) depthwise filters, and (c) point filters [27]

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	(None, 224, 224, 3)	0
mobilenet_1.00_224 (Model)	(None, 7, 7, 1024)	3228864
pure_attention (Model)	(None, 7, 7, 1)	139969
Total params: 3,368,833		
Trainable params: 137,921		
Non-trainable params: 3,230,912		

Figure 4. Number of parameters in the proposed architecture

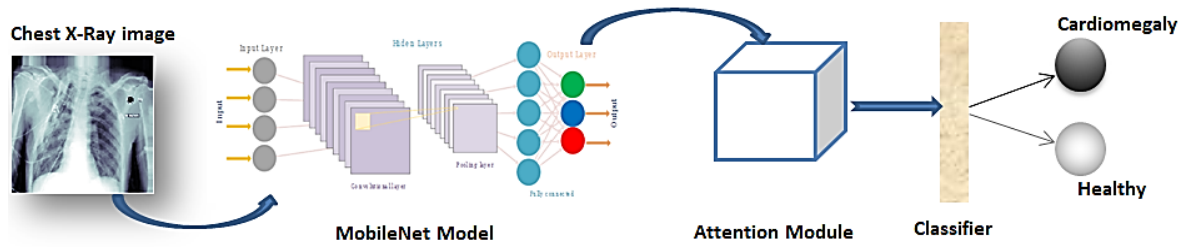


Figure 5. Summary of the proposed architecture

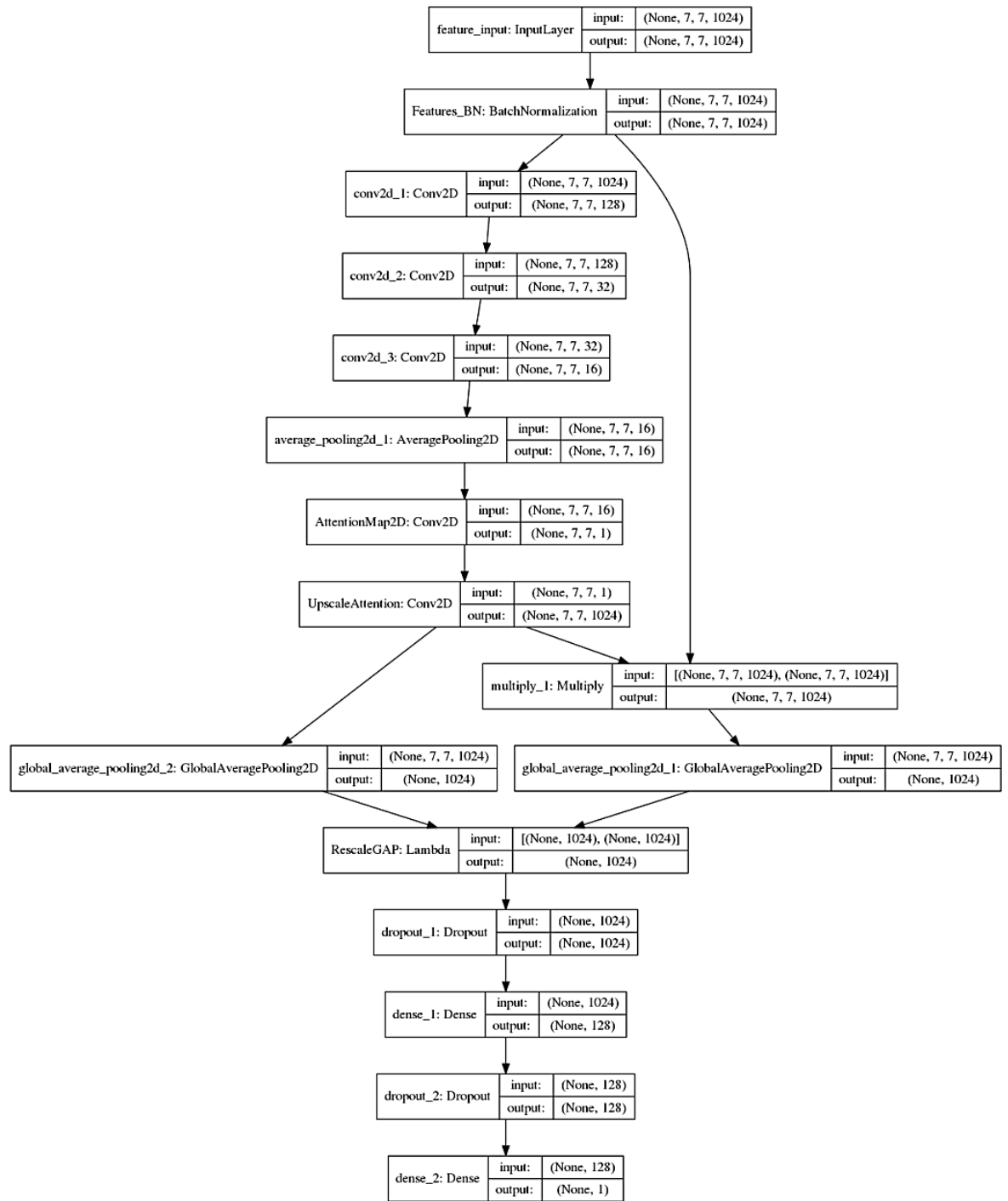


Figure 6. The detailed architecture of the proposed CNN

4. RESULTS AND DISCUSSION

In this section, we will explore the results achieved by our proposed architecture, along with a discussion of these results. Actually, the proposed CNN model and all experiments were implemented in Python language using the Tensorflow and Keras libraries, which are open-source machine learning libraries developed by Google for DL. In addition, due to resource limitations, we trained our model on the Kaggle graphics processing unit (GPU) simulator. Furthermore, to evaluate the performance of classification models, different metrics are required to differentiate between well-performing and non-performing models. Thus, we utilize accuracy, precision, recall, F1-score, and ROC curve metrics as performance metrics. In the following, we present the equations to calculate these metrics. We assume that TP are true positives, FP are false positives, TN are true negatives, FN are false negatives, i is the class index, and S is the total number of classes. The accuracy, the precision, and the recall can be calculated as (1)-(3) [8]:

$$Accuracy = \frac{1}{S} \sum_{i=1}^S \left(\frac{TP^i + TN^i}{TP^i + TN^i + FN^i + FP^i} \right) \quad (1)$$

$$Precision = \frac{1}{S} \sum_{i=1}^S \left(\frac{TP^i}{TP^i + FP^i} \right) \quad (2)$$

$$Recall = \frac{1}{S} \sum_{i=1}^S \left(\frac{TP^i}{TP^i + FN^i} \right) \quad (3)$$

The F1-score can be calculated by considering both recall and precision [8]:

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

The proposed model obtained an accuracy rate of 80%, a precision value of 66%, a recall rate of 62.5%, and an F1-score value of 64%. The achieved precision value indicates that the model correctly predicted 66% of the data, and the recall value acquired refers to 62.5% of the given results that are basically false and classified well. Besides, the receiver operating characteristic (ROC) curve illustrates the trade-off between specificity and sensitivity; where the specificity is the false positive rate (FPR), which corresponds to correctly classified data, while the sensitivity is the true positive rate (TPR). The ROC curve is constructed by plotting the TPR against the FPR. A classifier is considered effective when the curve approaches the upper left corner, and the area under the ROC curve (AUC) value is close to one (0.74), indicating a strong classifier. Figure 7 exposes the ROC curve of the proposed model.

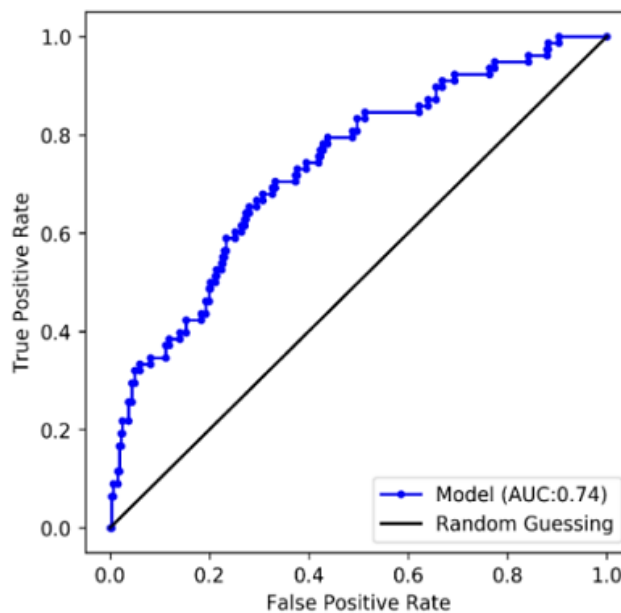


Figure 7. ROC curve of the proposed model

In fact, comparing the proposed CNN model with other available algorithms and methods may be inadequate due to various conditions including variations in the size of the used X-Ray images, the number of classes employed for classification, and the used dataset. Additionally, the models may handle different data characteristics that make the comparison unfair. However, the proposed model reached an accuracy rate of 81%, which outperformed the published results obtained by Bougias *et al.* [29], which are 71% and 81% achieved by the Inception V3 and SqueezeNet models respectively. In addition, in terms of AUC value, our proposed model achieved a value of 0.75 and surpassed the results obtained by Candemir *et al.* [30] that achieved 0.61 using Inception V3. Additionally, Son *et al.* [12] used a large dataset named MIMIC-CXR and achieved an accuracy of 89%, while it should be higher since the used dataset is huge.

Finally, the last part consists of presenting the predicted results using the unseen data of the test set. After training and evaluating the proposed model, we have generated the performance scores to assess its effectiveness. Then, we have used matplotlib functions to visualize the produced model's predictions. Figure 8 shows some examples of Cardiomegaly (True) or healthy (False), and the attention map that provides the prediction. These visualizations demonstrate the model's capability and ability to discern between Cardiomegaly and healthy cases and contribute to gaining better understanding and deeper insights into its performance and interpretability.

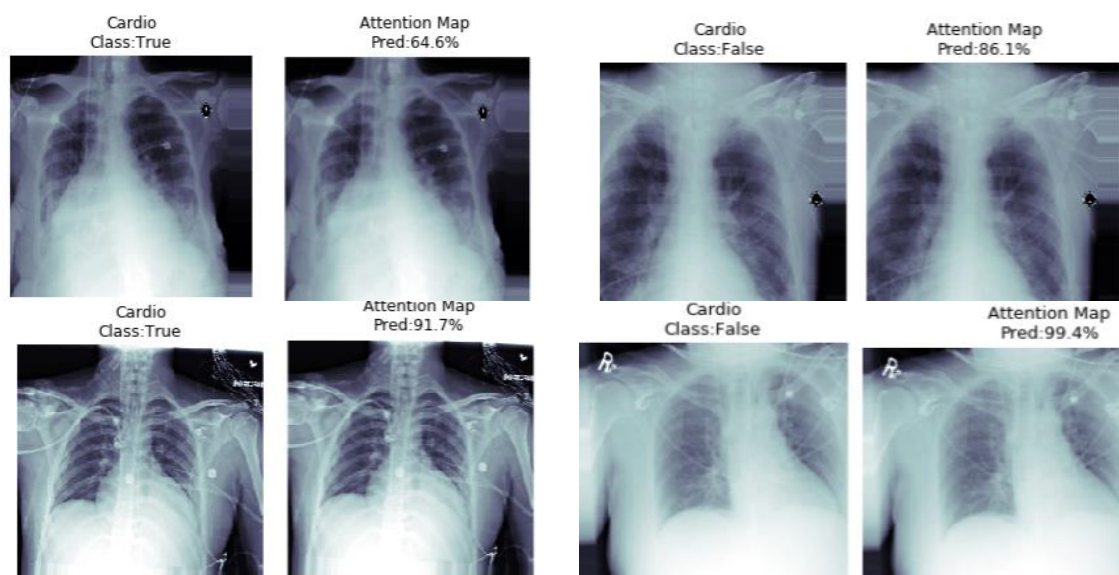


Figure 8. Examples of Cardiomegaly detection predicted using the proposed CNN model

5. CONCLUSION




Deep learning is a branch of artificial intelligence that empowers machines to acquire the ability to learn on their own. Deep learning models imitate the learning process of the human brain, and have many applications in the medical domain. In this paper, we introduced an innovative Deep Learning model that can classify Chest X-Ray images of Cardiomegaly patients using the attention module with MobileNet. In addition, the proposed model is composed of four main blocks, an attention module, a convolutional module, a fully connected, and classifier layers. According to the results, we reached a classification accuracy rate of 81%, a precision rate of 66%, a recall rate of 62.5%, and an F1-score value of 64%. In the future, we plan to use the existing techniques of data augmentation including generative adversarial networks (GANs) and convolutional autoencoder to enhance the efficiency of the classification model. In fact, data augmentation techniques are applied to increase the number of images used in the model learning phase and reduce the overfitting risk. Subsequently, we can use models that have a small filter size to extract the relevant part of the Chest X-Ray images. Furthermore, this approach has been tested on Cardiomegaly disease, but it can also be applied to detect the other diseases in the Chest X-Ray8 dataset.

REFERENCES




- [1] Nawsherwan, W. Bin, Z. Le, S. Mubarak, G. Fu, and Y. Wang, "Prediction of cardiovascular diseases mortality- and disability-adjusted life-years attributed to modifiable dietary risk factors from 1990 to 2030 among East Asian countries and the world,"

- Frontiers in Nutrition*, vol. 9, Oct. 2022, doi: 10.3389/fnut.2022.898978.
- [2] L.-A. Bocancia-Mateescu, D. Stan, A.-C. Mirica, M. G. Ghita, D. Stan, and L. L. Ruta, "Nanobodies as Diagnostic and Therapeutic Tools for Cardiovascular Diseases (CVDs)," *Pharmaceuticals*, vol. 16, no. 6, p. 863, Jun. 2023, doi: 10.3390/ph16060863.
 - [3] D. Adhikary, S. Barman, R. Ranjan, and H. Stone, "A Systematic Review of Major Cardiovascular Risk Factors: A Growing Global Health Concern," *Cureus*, Oct. 2022, doi: 10.7759/cureus.30119.
 - [4] M.-P. S. T. S. Bhadauria, "Cardiomegaly: A brief review with basic and physiotherapeutic approach," *Indian Journal of Physical Rehabilitation*, vol. 2022, Available: https://www.researchgate.net/publication/363087892_Cardiomegaly_A_brief_review_with_basic_and_physiotherapeutic_approach (accessed Nov. 9, 2022).
 - [5] S. Baudet, "Hypertrophy and dilation: a TOTally new story?," *Cardiovascular Research*, vol. 46, no. 1, pp. 17–19, Apr. 2000, doi: 10.1016/S0008-6363(00)00015-8.
 - [6] A. Bouslama, Y. Laaziz, and A. Tali, "Diagnosis and precise localization of cardiomegaly disease using U-NET," *Informatics in Medicine Unlocked*, vol. 19, p. 100306, 2020, doi: 10.1016/j.imu.2020.100306.
 - [7] K. Almezghwi, S. Serte, and F. Al-Turjman, "Convolutional neural networks for the classification of chest X-rays in the IoT era," *Multimedia Tools and Applications*, vol. 80, no. 19, pp. 29051–29065, Aug. 2021, doi: 10.1007/s11042-021-10907-y.
 - [8] S. El Omary, S. Lahrahe, and R. El Ouazzani, "A Lightweight CNN to Identify Cardiac Arrhythmia Using 2D ECG Images," 2022, pp. 122–160. doi: 10.4018/978-1-6684-2304-2.ch005.
 - [9] S. El Omary, S. Lahrahe, and R. El Ouazzani, "Detecting Heart Failure from Chest X-Ray Images Using Deep Learning Algorithms," in *2021 3rd IEEE Middle East and North Africa COMMUNICATIONS Conference (MENACOMM)*, Dec. 2021, pp. 13–18. doi: 10.1109/MENACOMM50742.2021.9678291.
 - [10] Guiqiu Yang *et al.*, "A heart failure diagnosis model based on support vector machine," in *2010 3rd International Conference on Biomedical Engineering and Informatics*, Oct. 2010, pp. 1105–1108. doi: 10.1109/BMEI.2010.5639619.
 - [11] K. H. Miao and J. H., "Coronary Heart Disease Diagnosis using Deep Neural Networks," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, 2018, doi: 10.14569/IJACSA.2018.091001.
 - [12] C.-S. Son, W.-S. Kang, J.-H. Lee, and K. J. Moon, "Machine Learning to Identify Psychomotor Behaviors of Delirium for Patients in Long-Term Care Facility," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 4, pp. 1802–1814, Apr. 2022, doi: 10.1109/JBHI.2021.3116967.
 - [13] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, Apr. 2015, pp. 294–297. doi: 10.1109/ISBI.2015.7163871.
 - [14] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals," *Information Sciences*, vol. 415–416, pp. 190–198, Nov. 2017, doi: 10.1016/j.ins.2017.06.027.
 - [15] J. Rubin, D. Sanghavi, C. Zhao, K. Lee, A. Qadir, and M. Xu-Wilson, "Large Scale Automated Reading of Frontal and Lateral Chest X-Rays using Dual Convolutional Neural Networks," 2018, [Online]. Available: <http://arxiv.org/abs/1804.07839>
 - [16] U. Srinivas, "Discriminative models for robust image classification," 2016, [Online]. Available: <http://arxiv.org/abs/1603.02736>
 - [17] T. G. Dietterich, "Ensemble Methods in Machine Learning," 2000, pp. 1–15. doi: 10.1007/3-540-45014-9_1.
 - [18] C. Zhou *et al.*, "Final overall survival results from a randomised, phase III study of erlotinib versus chemotherapy as first-line treatment of EGFR mutation-positive advanced non-small-cell lung cancer (OPTIMAL, CTONG-0802)," *Annals of Oncology*, vol. 26, no. 9, pp. 1877–1883, Sep. 2015, doi: 10.1093/annonc/mdv276.
 - [19] Y. Sasaki, K. Abe, M. Tabei, S. Katsuragawa, A. Kurosaki, and S. Matsuoka, "Clinical usefulness of temporal subtraction method in screening digital chest radiography with a mobile computed radiography system," *Radiological Physics and Technology*, vol. 4, no. 1, pp. 84–90, Jan. 2011, doi: 10.1007/s12194-010-0109-7.
 - [20] C. Li, G. Zhu, X. Wu, and Y. Wang, "False-Positive Reduction on Lung Nodules Detection in Chest Radiographs by Ensemble of Convolutional Neural Networks," *IEEE Access*, vol. 6, pp. 16060–16067, 2018, doi: 10.1109/ACCESS.2018.2817023.
 - [21] S. R. Islam, S. P. Maity, A. K. Ray, and M. Mandal, "Automatic Detection of Pneumonia on Compressed Sensing Images using Deep Learning," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, May 2019, pp. 1–4. doi: 10.1109/CCECE.2019.8861969.
 - [22] V. Chouhan *et al.*, "A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images," *Applied Sciences*, vol. 10, no. 2, p. 559, Jan. 2020, doi: 10.3390/app10020559.
 - [23] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 3462–3471, 2017, doi: 10.1109/CVPR.2017.369.
 - [24] N. I. of H. C. X.-R. Dataset, "NIH Chest X-rays," *NIH Chest X-rays*, 2018.
 - [25] H. A. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *Computer Vision and Pattern Recognition*, vol. 14, no. 2, pp. 53–57, 2009, doi: 10.48550/arXiv.1704.04861.
 - [26] W. Wang, Y. Li, T. Zou, X. Wang, J. You, and Y. Luo, "A Novel Image Classification Approach via Dense-MobileNet Models," *Mobile Information Systems*, vol. 2020, pp. 1–8, Jan. 2020, doi: 10.1155/2020/7602384.
 - [27] W. Wang, Y. Hu, T. Zou, H. Liu, J. Wang, and X. Wang, "A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers," *Computational Intelligence and Neuroscience*, vol. 2020, pp. 1–10, Aug. 2020, doi: 10.1155/2020/8817849.
 - [28] G. Wang, G. Yuan, T. Li, and M. Lv, "An multi-scale learning network with depthwise separable convolutions," *IPSI Transactions on Computer Vision and Applications*, vol. 10, no. 1, p. 11, Dec. 2018, doi: 10.1186/s41074-018-0047-6.
 - [29] H. Bougias, E. Georgiadou, C. Malamateniou, and N. Stogiannos, "Identifying cardiomegaly in chest X-rays: a cross-sectional study of evaluation and comparison between different transfer learning methods," *Acta Radiologica*, vol. 62, no. 12, pp. 1601–1609, Dec. 2021, doi: 10.1177/0284185120973630.
 - [30] S. Candemir, S. Rajaraman, G. Thoma, and S. Antani, "Deep Learning for Grading Cardiomegaly Severity in Chest X-Rays: An Investigation," in *2018 IEEE Life Sciences Conference (LSC)*, Oct. 2018, pp. 109–113. doi: 10.1109/LSC.2018.8572113.




BIOGRAPHIES OF AUTHORS

Sara El Omary    received a B.Sc. degree in computer science from the Higher School of Technology of Oujda, Morocco in 2016, then an M.Sc. degree in data science from the Faculty of Science Semlalia of Marrakech, Morocco in 2020. She is currently a Ph.D. candidate at the Moulay Ismail University of Meknes (Morocco). Her main areas of interest include machine learning, deep learning, image preprocessing, and computer vision. She can be contacted at email: elomarysr@gmail.com.



Souad Lahrache    Professor at the Faculty of Science, University Ibnou Zohr, Agadir, Morocco. She obtained a Ph.D. from the Faculty of Sciences of the University Moulay Ismail of Meknes, Morocco. She has published several papers in peer-reviewed journals and international conferences. Her research interests include image processing, pattern recognition, computer vision, and machine learning. She can be contacted at email: souadlahrache@gmail.com.



Rajae El Ouazzani    Received her master's degree in computer science and telecommunication by the Mohammed V University of Rabat (Morocco) in 2006 and the Ph.D. in image and video processing by the High National School of Computer Science and Systems Analysis (Morocco) in 2010. From 2011, she is a Professor in the High School of Technology of Meknes, Moulay Ismail University in Morocco. Since 2007, she is an author of several papers in international journals and conferences. Her domains of interest include multimedia data processing and telecommunications. She can be contacted at email: elouazzanirajae@gmail.com.