FEC-NET: Attention Based Featured Extraction Classifier For Anomalies Detection in Chest X-Ray Images

Jaiman Pandya
Trabuco Hills High School
Rancho Santa Margarita, USA
jaimanpandya@gmail.com

Due to their exceptional performance in classification challenges, such as the automatic analysis of X-rays, and their capacity to extract features, convolutional neural networks (CNNs) have recently come to dominate the area of computer vision. It is urgently necessary to use cutting-edge technology, such as deep learning, to detect anomalies in chest X-ray images in order to increase productivity and diagnosis precision. The over-fitting and low transfer rate of existing neural networks makes them difficult to fine-tune for medical image processing. To overcome these challenges, the proposed study designed an attention-based CNN model 'Feature Extraction Classifier Network' (FEC-NET) to automatically extract the features from customized inception blocks for efficient classification of chest X-ray images. The proposed model is small and lightweight compared to the existing pre-trained model. Moreover, we also turned the proposed model to achieve better performance of the model. According to the experimental findings, our light model FEC-NET had improved accuracy, recall, F1 score, and AUC values. It demonstrates that developing a suitable CNN is preferable to optimizing deep networks and that increasing the training data is essential to improving CNN performance.

Keywords—X-Ray Images, Attention, CNN, Radiology

I. INTRODUCTION

In recent years, the medical field has been troubled by different significant issues that impede patient health: medical staff burden and human error in interpreting the findings of medical reports [1][2]. There is no simple solution of these problems, which is particularly hazardous in medicine, where interpreting mistakes can result in severe complications of health. First, overwork in medicine, exacerbated in recent years due to pandemic of COVID-19 globally, can result in diagnostic errors and treatment delays. As mentioned previously, the exploration of certain medical tests is also subjective. The expert of medical staff interpreting these tests, such as X-rays, may make an incorrect diagnosis due to the presence of signs of multiple disorders to varying degrees [3]. Due to its low cost, rapidity of acquisition, and lack of preparation requirements, this form of imaging test is one of the most frequently used for a variety of diagnoses [4]. The diagnosis of certain thoracic disorders affecting organs including the heart, lungs, and bones can be helped by thoracic X-rays. Convolutional neural networks (CNN) are the most effective for analyzing X-rays [5] because of their properties. By expediting the diagnostic process and lessening physicians' workload, the combination of AI algorithms and medical expertise can improve the performance quality of medical staff [6] and decrease patient wait times.

The chest X-ray is one of the furthermost nearby inspections for lung disease screening and diagnosis, such as Infiltration, Atelectasis, Mass, Pneumothorax, Consolidation, Cardiomegaly, Emphysema, Pneumonia, and Hernia [7]. Every year, more than 35 million chest X-ray scans are taken only in the United States, and radiologists must interpret over 100 X-ray studies daily [8]. Due to the vast population and increased awareness of health in China, there are significantly more patients and chest X-rays. The automated methods of advanced technology can help radiologists in accurately and efficiently diagnosing disease. Regular and irregular chest X-ray images can be automatically classified into two categories by trained programs, allowing radiologists to focus more on abnormal X-ray images. In addition, trained algorithms may group chest X-ray images into additional categories based on various diseases. In addition, trained instruments can aid in disease localization and visualization.

Deep learning has recently received a great deal of attention, particularly in the field of image classification. This recognition is a result of a classification technique's success that makes use of convolutional neural network architecture. These networks have a remarkable capacity for autonomous backpropagation, which allows them to automatically collect significant image features [9]. As a result, they are able to learn from photos and recognize relevant patterns without the use of detailed manual instructions or pre-processing. Deep learning has gained significant acceptance and adoption in several image analysis applications as a result of its ability to do this task. Due to its capacity to derive visual features, CNNs have revolutionized computer vision. These architectures consist of multiple layers. The initial model's convolutional layers are inspired by the concept of cells in visual neuroscience. The animal visual brain served as inspiration for the designs. These architectures stand out because of their outstanding capacity to identify patterns in data, which improves the efficiency of earlier systems based on machine learning models. They excel in a variety of tasks, including speech recognition, computer vision, and text analysis, making them a standard in deep learning thanks to this advantage [10].

Because of the unique characteristics of chest X-rays, CNNs can analyze them successfully. The system can learn from enormous volumes of data and find patterns that are hard for people to recognize, thus there is no longer any requirement for manual extraction of visual characteristics or segmentation based on CNN. Although the focus of this article is on classification issues, other issues can also be resolved, including X-ray segmentation [11], localization, regression (such as medicine dose prediction), and others. On the basis of big data and automated feature extraction, CNNs may be used as an analytical tool for chest radiographs. However, due to concerns about patient privacy, huge training sets are often not available in the medical industry. The OpenI was the largest dataset, containing 7,470 images of chest x-rays and 3,955 radiology reports. In recent years, ChestX-ray14, an archive containing more than 30,000 patients record and 112,120 classified chest x-ray images, was released [7]. Despite being unbalanced and defective, ChestX-ray14 is sufficiently large for training, so we select it as our dataset.

For processing x-rays of the chest and other medical pictures, transfer learning is considered as the preferred method in numerous publications for the classification task [12]. Researchers train ChestX-ray14 by improving present ImageNet-trained deep neural networks. Four standard CNN architectures (GoogleNet, AlexNet, ResNet and VGGNet) were fine-tuned in the only peer-reviewed publication, and ResNet obtained the best results [13][14]. The variance between medical and natural images can result in low transmission performance, although fine-tuning already-existing deep neural networks is not necessarily the best answer.

Furthermore, too many factors could result in overfitting and wasteful space use. ChestX-ray14 is sufficiently large for the research community and necessitates the training of a customized lightweight model without excessive time or memory consumption. In this study, we suggest a brand-new CNN architecture called FEC-NET and train it using ChestX-ray14 from the beginning.

II. RELATED WORK

Since the 1980s, when AI approaches were first adopted in medicine, the use of these algorithms has significantly increased, especially more recently. Diverse sorts of clinical data analysis have effectively used deep learning algorithms [15]: bio signals, involving electrical [16], mechanical [17], [18] and heat signals [19]; biological medicine, examining the molecules of biological function [20], [21]; electronic health records (EHR), aimed at improving diagnosis [22], [23]; and clinical scans, widely used in the identification of multiple diseases [24], [25] similar to how our issue is. Evidence-based medicine has replaced observation-based medicine in the delivery of healthcare. Due to their ability to recognize some radiological signals that medical personnel are unable to notice, methods for large data and deep learning are particularly helpful in this sector [35]. Although the focus of this work is related to classification, there are papers that use these algorithms for regression issues, including estimating drug dosage [26], producing medical reports from clinical trials [27], assisting healthcare administration [28], or image processing, including segmentation [29] and rebuilding of images [30].

The study into the use of deep learning and machine learning in medical image processing has been significantly influenced by the COVID-19 epidemic. As might be predicted, a large portion of the analyzed categorization schemes has concentrated on identifying symptoms of bilateral COVID-19-associated pneumonia. In Ahmed et al. [4], they categorize chest X-rays using VGG16 and ResNet, two distinct pre-trained structures and optimize the hyperparameters. In Pham [31], they train three alternative pre-trained architectures—AlexNet, GoogleNet, and SqueezNet—with six datasets separately, evaluating various train set sample sizes (50 and 80%), and SqueezeNet achieves an accuracy of 99.85%. The ensemble system developed by Ahmad et al. [32] obtains 96.49% accuracy and is based on MobileNet and InceptionV3. The use of binary classification to multiclass circumstances quickly made it possible to determine if pneumonia is caused by COVID-19 or another virus/bacteria or whether the patient is healthy. Similar to binary classification issues, several studies, including that by Avola et al. [33], search through the most effective topologies, including ResNet, GoogleNet, AlexNet, and ShuffleNet, among others. On a dataset of 6330 samples, MobilNet_v3 produces the highest result with an accuracy of 84.92%. Zebin and Rezvy [34] train a pre-trained model in addition to generating a heatmap-based visualization that displays two pictures for each sample. The first is for the original X-ray image and the second for the class activation areas which is more important for the classification model. However, because the pictures do not overlap, interpretation is challenging. Other studies, like Teixeira et al. [35], use a segmentation approach to eliminate any system components that are not necessary, which should enhance performance and visualization. Three separate datasets make up their dataset labelled as COVID-19, normal, and lung opacity.

Although samples of a specific ailment and healthy samples make up the majority of medical datasets, chest X-rays frequently show symptoms of many pathologies. Due to this, several scientists have released multilabel radiological datasets during the past five years. With the extra difficulty of class imbalance, these datasets are more realistic than binary ones. The prevalence of disease in society should determine the proportion of each class in a real dataset, meaning that certain classes should be greater in number than others. These datasets' intriguing qualities must be carefully examined in order to properly solve the issue. Since its release in 2019 [36], the ChestX-ray 14 dataset has become one of the most frequently used datasets in the area of chest X-ray classification. By using DenseNet-121 to optimize its hyperparameters, for instance, Wang et al. [56] were able to achieve an average AUC of 0.82. The pneumonia class's AUC was 0.662, the lowest possible, while the hernia class' AUC was 0.923, the best possible. Other researchers have achieved an AUC for pneumonia of 0.73 using other designs like InceptionResNet_v2 and ResNet152_v2 [37]. A lot of the work on this dataset involves retraining cutting-edge architectures, however, there are various methods for enhancing classification performance. For instance, Almezhghwi et al. [38] switched the classification algorithm from AlexNet and VGG16 to SVM in an effort to enhance the findings of earlier manuscripts, and both architectures achieved an AUC for pneumonia of 0.98. The performance of the classification model can be increased with different types of X-ray images or with customized deep-learning models.

Training a deep convolutional neural network (CNN) from scratch is challenging due to the fact that requires a big amount of labelled training data and a high level of expertise to achieve convergence [43]. Fine-tunned networks that have already been trained on huge datasets are a viable option. For semantic segmentation, the study [39] fine-tuned all layers using backpropagation over the whole FCN-AlexNet, FCN-VGG16, and FCN-GoogLeNet. The study [40] used a self-paced fine-tuning network (SPFTN)-based architecture to separate objects into categories in movies with sparse labelling. A fine-tuning approach to update the network that was previously trained on photos of urban environments was provided in a study [41], and it successfully transfers semantic properties to a different context. Regarding medical pictures that diverge greatly from natural ones, several studies have varied opinions. The study [42] showed that in the multi-class classification evaluation of knee osteoarthritis, fine-tuning considerably outperformed feature extraction from scratch. However, [43] showed that in the categorization of cytopathology images, a novel CNN that extracts features surpassed fine-tuning.

III. PROPOSED MODEL

The conventional method employed to enhance the performance of image classification entails the stacking of convolutional layers within convolutional neural networks (CNNs). Although this technique has proven effective to some extent, it brings about certain drawbacks that limit its overall efficacy. Layer-stacking in traditional image classification methods significantly increases parameters and computational complexity, posing challenges for scalability, real-time processing, and resource utilization. Alternative approaches are needed to mitigate these drawbacks while maintaining or improving performance. The elevated parameter count and computational demand can hinder real-time processing capabilities and impede the scalability of the network architecture. Moreover, the traditional CNNs such as VGG [44], ResNet [45], and AlexNet [46], use the single kernel size which results in the loss of fine-grained features important for analyzing abnormal spots found in chest X-ray images. Consequently, the compromised retention of significant feature details adversely affects the accuracy and robustness of disease recognition algorithms utilized within the context of chest X-ray classification.

To enhance chest X-ray image classification, we introduce FE-unit, a feature extraction module comprising three convolutional units designed to address the identified challenges. RIBFE-unit emphasizes network width to overcome limitations in existing methods. The first convolutional block employs two 1D 3 Conv receptive fields, achieving an equivalent 5×5 receptive field. This configuration enables the detection of larger disease spots, capturing global feature information vital for disease recognition. The second unit utilizes 1D 3 Conv convolutions to detect smaller disease spots and gather local feature information. These units together reduce information loss and enhance recognition accuracy. To address

issues of gradient vanishing and accelerate optimization, the third unit adopts a residual structure [47]. It overcomes depth-related degradation by using a mapping relationship expressed as H(x) = F(x) + x, allowing small weight adjustments for optimal network performance. The 1D convolutions are employed in all units to enhance compatibility with input data, facilitating down sampling, up sampling, and flexible adjustment of output size. Figure 1 (inception modified) shows the customized Inception unit or modified inception unit called FE-unit.

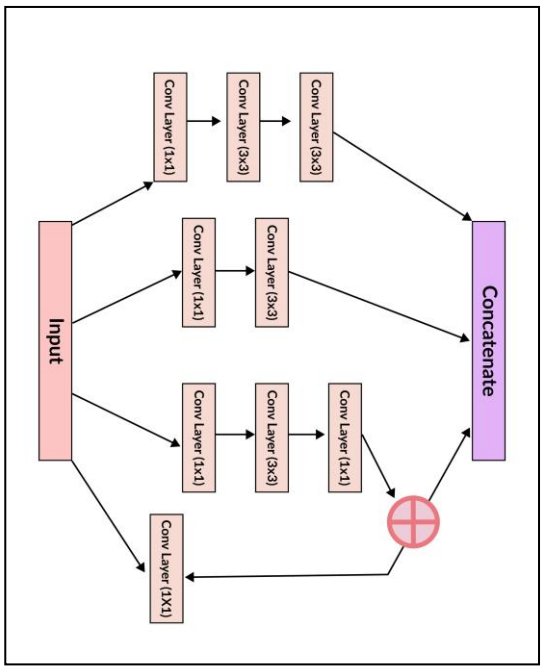

Fig 1. Customized Inception Unit Architecture

Novel Inception structure tries to reduce the model parameters but still suffers from overfitting in complex image scenarios like medical image classification where data is always limited. To overcome these challenges associated with the Inception network, we have devised an alternative Chest diseases classification network called FENet. FENet consists of three simplified FE-unit1, each designed to address specific shortcomings of the Inception network. Specifically, the FE-unit1 produces 512 feature maps, the FE-unit2 generates 1024 feature maps, and the FE-unit3 yields 2048 feature maps. This hierarchical design enables the feature extraction module, consisting of the three FE units, to perform a localized perception of Chest disease present in the input images. By leveraging the varying channel outputs and increasing feature maps, the network becomes capable of capturing a diverse range of local feature details pertaining to Chest diseases. Following the extraction of local features, the FENet incorporates higher-level synthesis operations to integrate the acquired local information and derive comprehensive feature representations. This synthesis process aggregates and synthesizes the local feature information from multiple FE-Blocks, allowing the network to capture more holistic and discriminative representations of Chest diseases. By combining localized perception with higher-level synthesis, the FENet effectively captures both detailed local features and their contextual relationships, enhancing the overall capability of the network for accurate Chest disease classification. By incorporating the FE-Block module, the resulting FE-Net network possesses a broader network structure, enabling the selection of fine-grained information while reducing the model complexity. Figure 2 illustrates the structure of FE-Net (arch without attention).

The FE-Net (Feature Extraction Network) encounters challenges in accurately detecting Chest disease due to the distinctive characteristics of Chest disease spots. These characteristics manifest as greater inter-disease and minimal intra-disease differences. The presence of such variations in disease patterns complicates the accurate identification and classification of Chest diseases within the FE-Net network. Overcoming these challenges requires sophisticated feature extraction mechanisms capable of capturing the subtle differences between different Chest disease types. We propose integrating a modified Convolution Block Attention Module (CBAM) [48] into the FE-Net network to reduce the redundancy in the information.

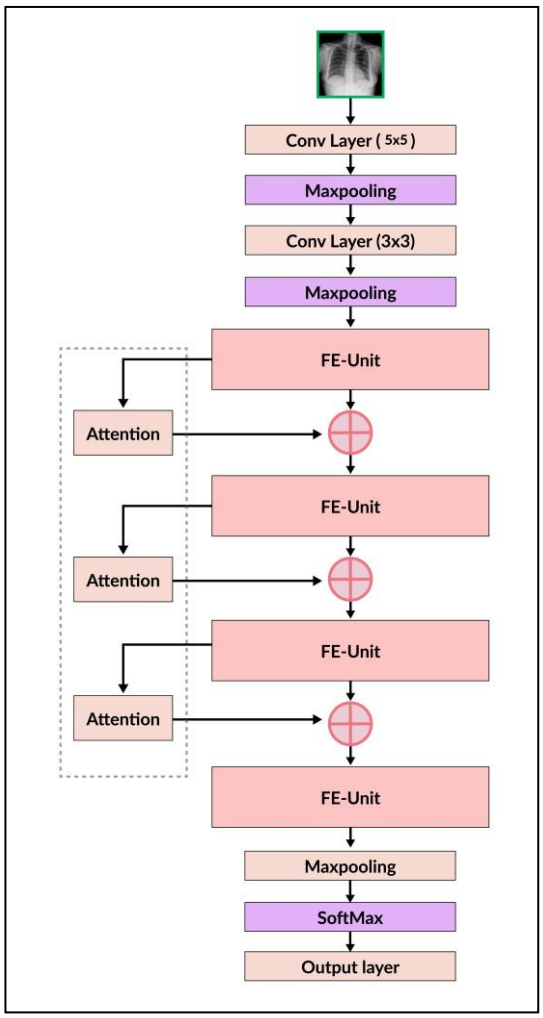

Fig 2. Complete Architecture of the proposed model. Gray Box add Attention Blocks.

The creation of a channel attention map is facilitated through the utilization of the inter-channel associations within features. Given that each channel within a feature map is construed as a feature detector, the focus of channel attention is on discerning the significant elements of the input image. The reduction of the spatial dimension of the input feature map is instrumental in computing channel attention efficiently. It has been generally recognized that average-pooling is an effective method for the consolidation of spatial information." This module enhances disease spot localization, leading to improved recognition accuracy. The modified CBAM module seamlessly integrates with the FE-Net network as a channel and spatial attention mechanism. The channel attention module operates on an input feature map encompassing height (H), width (W), and channel (C) dimensions. Through average pooling and maximum pooling, the feature map (F) is transformed into distinct feature representations, namely $Z^c{}_{avg}$ and $Z^c{}_{max}$. To preserve disease region features, the original Dense layers are replaced with two 1D Conv layers, which serve as feature information detectors rather than dimensionality reducers.

Qualitative analysis of kernel size ensures effective detection of disease spot information across various sizes in $Z^c{}_{avg}$ and $Z^c{}_{max}$. The feature representations are undergoing a join operation and generates the channel attention map $A_c(Z) \in R^{1x1xc}$. This map is fed to spatial attention to produce respective output map$A_S(Z) \in R^{HxW}$ Experimental observations indicate that $A_S(Z)$ reflects scores associated with the diseased region and non-diseased region in the original representation Z, reflecting the proportion of such region within the image. To preserve unaffected convolution calculations in subsequent modules, a weighting operation is introduced involving a channel-wise weighting of $A_S(Z)$ output by multiplying it by the original image, to focus on the crucial of lesions. This magnifies disease region information while suppressing non-disease features and eventually enhancing robustness. The modified attention module is depicted in Figure 3.

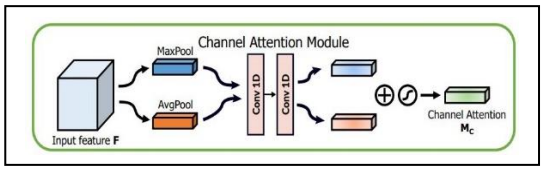

Fig 3. Modified Attention Unit

To further enhance the performance of our FE-Net for Chest regions with varying shapes and sizes and considering the limited adaptability of the CBAM [48] to FE-Net, we propose the FEC-Net. FE-Net given the limitation of original CBAM in generalization. FEC-Net is built upon the foundation of FEC-Net, incorporating modified CBAM. The middle layer of FEC-Net employs four FE-Blocks, and the output features from each FE-Block undergo further feature extraction using modified CBAM, ensuring a more fine-grained representation of the input X-ray image. The SoftMax activation is employed in the final stage to compute the recognition probabilities for each disease category. FEC-Net operates in both the channel-wise and spatial domains to get finer identification of Chest diseases and effectively filters non-disease feature information, thereby enhancing the accuracy of disease identification. Figure 2 provides an illustration of the FEC-Net network architecture (Gray dotted box showing the attention blocks).

IV. Results and Discussion

A. Dataset and Preprocessing

ChestX-ray14 is the most useful database that is currently available for training a classification algorithm for chest X-ray images. The collection contains information on almost 30,000 individuals, 297,541 labelled chest X-rays, and 14 distinct categories of abnormal images. The statistical information on the dataset can be found in Table I. On the other hand, as can be shown in Table I, there is a significant disparity in the number of chest X-ray images that belong to each of the categories in Chest X-ray 14. Due to the fact that some of the images in the dataset have multi-label data, there are more than 15 categories present in those images. The work that was proposed chose images that had any anomaly, and the number of samples that had an anomaly label and belonged to each category is shown in Table I. In addition to this, each and every image was resampled to the dimensions of 224 by 224.

TABLE I. Statistics of the chest X-ray14 dataset.

Class	Samples	Class	Samples
Infiltration	19894	Pleural Thickening	3385
Effusion	13317	Cardiomegaly	2778
Atelectasis	11559	Emphysema	2516
Nodule	6331	Edema	2305
Mass	5782	Fibrosis	1686
Pneumothorax	5296	Pneumonia	1431
Consolidation	4643	Hernia	227

B. Training and Testing

We split the Chest ChestX-ray14 dataset into the training, testing and validation set with the ratio of 70%, 20% and 10% respectively. The 14,084 samples of the training set and 2,012 samples of the validation set were used for the training of the model. Two different architectures of the proposed model (proposed model without attention and proposed model with attention blocks) were used to classify the chest X-ray images. Each architecture was trained multiple times on different hyperparameters. Each time the X-ray images of shape 224*224 were used as input. The model showed the best results with Adam optimizer with 0.003 learning rate, 64 batch size and categorical-cross-entropy as loss function. All the experiments were performed in the virtual environment of Python with the TensorFlow framework. Moreover, the GeForce GTX 1080 of 11GB on the Ubuntu operating system was used to perform the experiments. By following every epoch, validation samples were used to estimate the performance of the model. The decay was also set by a factor of 2, 5 and 10 after every 25 epochs. In each experiment, the model was trained on 100 epochs with a training and validation set

C. Classification Results

The proposed model was trained without an attention mechanism and with an attention mechanism on multiple hyperparameter values. Both models showed 0.95% and 0.88% training accuracy during the training of the proposed model. By following the training of the model, 4,024 test samples of the chestXray14 dataset were used to evaluate the performance of the model. The evaluation process showed that the proposed model with attention mechanism outperforms the other variants of the proposed model or the available pre-trained model. The trained model showed the 0.90% highest detection score for the hernia class and 0.27% lowest detection score for the pneumonia class (Table II). The overall accuracy of the proposed model for the ChestX-ray14 dataset was significantly considerable compared to the pre-trained model on the ImageNet dataset. The complete overview of the achieved results is shown in Table II.

TABLE II. Evaluation Scores of each Class in the dataset.

Classes	Accuracy	Precision	Recall	F1-Score
Infiltration	0.9673	0.9367	0.9296	0.9331
Effusion	0.9607	0.8846	0.8500	0.8669
Atelectasis	0.8965	0.8620	0.8771	0.8695
Nodule	0.9674	0.7523	0.8688	0.8064
Mass	0.9613	0.6793	0.8650	0.7610
Pneumothorax	0.9550	0.6542	0.6603	0.6572
Consolidation	0.9551	0.5833	0.7543	0.6578
Pleural Thickening	0.9428	0.3993	0.7396	0.5186
Cardiomegaly	0.9362	0.2555	0.6607	0.3685
Emphysema	0.9613	0.4319	0.7986	0.5602
Edema	0.9736	0.5208	0.8695	0.6514
Fibrosis	0.9773	0.4710	0.7738	0.5855
Pneumonia	0.9415	0.0973	0.2797	0.1444
Hernia	0.9884	0.1785	0.9090	0.2985
Average	0.9560	0.5504	0.7740	0.6199

The overall accuracy of the proposed model for ChestX-ray14 dataset was significantly considerable compared to pretrained model on the ImageNet dataset. The complete comparison of the achieved results is shown in Table III.

TABLE III. OVERVIEW OF EVALUATION SCORE ON DIFFERENT MODELS.

Model	Accuracy	Precision	Recall	F1-Score
VGG19	0.4078	0.3654	0.4014	0.3825
ResNet50	0.4421	0.4325	0.4255	0.4289
Efficient Net	0.5488	0.5410	0.5324	0.5366
FEC-NET (without attention)	0.8886	0.5360	0.7224	0.6153
FEC-NET (with-attention)	**0.9560**	**0.5504**	**0.7740**	**0.6199**

The results of Table 3 showed the performance of the proposed model over the pre-trained model with parameter tunning options. The results reveal that the utilization of an already trained deeper model with parameter tunning is less significant than the newly designed CNN model. The chest X-ray images for anomaly classification are requiring the lightweight CNN model with timely attention to integrating the convolutional layers in the model. The proposed model also contributed in this regard to achieving a good accuracy score compared to the available models. We also achieve the best accuracy score (Table 3) with our proposed attention-based FEC-NET. One thing that needs to be considered is that the selected dataset is a high-class imbalance dataset and accuracy is not a good evaluation measure in this scenario. To evaluate the actual performance of the proposed model, recall should be considered as it will describe the percentage of accurately classified samples of each class. Hence, the proposed model also showed a 0.77% recall which is significant compared to the other listed models. However, there is a need to investigate the weights of the model to tunned and update the model in future for better experimental results.

V. CONCLUSION

In this study, we used deep learning approaches, concentrating on convolutional neural networks (CNNs), to meet the pressing demand for improved technology to detect abnormalities in chest X-ray pictures. CNNs have gained popularity in the field of computer vision because to their outstanding performance in a variety of classification tasks, including the automated analysis of X-rays. When used for medical image processing, the current neural networks suffer from overfitting and poor transferability. To effectively classify chest X-ray pictures, we presented an attention-based CNN model named "Feature Extraction Classifier Network" (FEC-NET). Customized inception blocks were inserted into our model to extract pertinent characteristics, and attention mechanisms were added to improve the model's capacity to concentrate on prominent areas of the images. Notably, compared to current pre-trained models, FEC-NET showed greater performance despite being smaller and lighter. We conducted considerable experimentation, and the findings were compelling, validating the efficacy of our suggested FEC-NET paradigm. When compared to earlier methods, the model showed considerable gains in accuracy, recall, F1 score, and AUC values. Our model has the capacity to accurately detect anomalies in chest X-ray pictures as seen by the recall score of 0.77% that was acquired. This result represents a significant improvement in classification performance. We stress that creating an appropriate CNN architecture, such as FEC-NET, is superior to focusing only on deep network optimization. Personalized feature extraction blocks and attention methods may help the model successfully extract pertinent data from chest X-ray pictures, according to our findings. Additionally, our work emphasizes the need of expanding the training data set to improve CNN performance in medical image classification tasks. To capture the intrinsic complexity and diversity contained in medical pictures and enable the model to build robust representations and enhance its accuracy, sufficient training data is crucial.

REFERENCES

[1] E. Moustaka and T. C. Constantinidis, "Sources and effects of Work-related stress in nursing," *SCIENCE JOURNAL ® VOLUME*, vol. 4, no. 4, 2010, Accessed: Jun. 24, 2023. [Online]. Available: www.hsj.gr

[2] S. Domínguez-Rodríguez *et al.*, "Testing the Performance, Adequacy, and Applicability of an Artificial Intelligent Model for Pediatric Pneumonia Diagnosis," *SSRN Electronic Journal*, May 2022, doi: 10.2139/SSRN.4095071.

[3] N. J. Shaw, M. Hendry, and O. B. Eden, "Inter-Observer Variation in Interpretation of Chest X-Rays," *http://dx.doi.org/10.1177/003693309003500505*, vol. 35, no. 5, pp. 140–141, Oct. 1990, doi: 10.1177/003693309003500505.

[4] K. Ben Ahmed, G. M. Goldgof, R. Paul, D. B. Goldgof, and L. O. Hall, "Discovery of a Generalization Gap of Convolutional Neural Networks on COVID-19 X-Rays Classification," *IEEE Access*, vol. 9, pp. 72970–72979, 2021, doi: 10.1109/ACCESS.2021.3079716.

[5] Y. Lecun, Y. Bengio, and R. 4g332, "Convolutional Networks for Images, Speech, and Time-Series".

[6] "Deep learning drops error rate for breast cancer... - Google Scholar." https://scholar.google.com/scholar?q=Deep%20learning%20drops%20error%20rate%20for%20breast%20cancer%20diagnoses%20by%2085 (accessed Jun. 24, 2023).

[7] 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." pp. 2097–2106, 2017. Accessed: Jun. 24, 2023. [Online]. Available: https://uts.nlm.nih.gov/metathesaurus.html

[8] S. I. Kamel, D. C. Levin, L. Parker, and V. M. Rao, "Utilization Trends in Noncardiac Thoracic Imaging, 2002-2014," *Journal of the American College of Radiology*, vol. 14, no. 3, pp. 337–342, Mar. 2017, doi: 10.1016/J.JACR.2016.09.039.

[9] S. W. E. Baalman *et al.*, "A morphology based deep learning model for atrial fibrillation detection using single cycle electrocardiographic samples," *Int J Cardiol*, vol. 316, pp. 130–136, Oct. 2020, doi: 10.1016/J.IJCARD.2020.04.046.

[10] Ş. Öztürk, U. Özkaya, and M. Barstuğan, "Classification of Coronavirus (COVID-19) from X-ray and CT images using shrunken features," *Int J Imaging Syst Technol*, vol. 31, no. 1, pp. 5–15, Mar. 2021, doi: 10.1002/IMA.22469.

[11] T. Agrawal and P. Choudhary, "EfficientUNet: Modified encoder-decoder architecture for the lung segmentation in chest x-ray images," *Expert Syst*, vol. 39, no. 8, p. e13012, Sep. 2022, doi: 10.1111/EXSY.13012.

[12] R. K. Sevakula, V. Singh, N. K. Verma, C. Kumar, and Y. Cui, "Transfer Learning for Molecular Cancer Classification Using Deep Neural Networks," *IEEE/ACM Trans Comput Biol Bioinform*, vol. 16, no. 6, pp. 2089–2100, Nov. 2019, doi: 10.1109/TCBB.2018.2822803.

[13] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Sep. 2014, Accessed: Jun. 24, 2023. [Online]. Available: https://arxiv.org/abs/1409.1556v6

[14] C. Szegedy *et al.*, "Going Deeper With Convolutions." pp. 1–9, 2015.

[15] F. Piccialli, V. Di Somma, F. Giampaolo, S. Cuomo, and G. Fortino, "A survey on deep learning in medicine: Why, how and when?," *Information Fusion*, vol. 66, pp. 111–137, Feb. 2021, doi: 10.1016/J.INFFUS.2020.09.006.

[16] S. W. E. Baalman *et al.*, "A morphology based deep learning model for atrial fibrillation detection using single cycle electrocardiographic samples," *Int J Cardiol*, vol. 316, pp. 130–136, Oct. 2020, doi: 10.1016/J.IJCARD.2020.04.046.

[17] H. Li *et al.*, "A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection," *Comput Biol Med*, vol. 120, p. 103733, May 2020, doi: 10.1016/J.COMPBIOMED.2020.103733.

[18] J. Pan, Y. Zi, J. Chen, Z. Zhou, and B. Wang, "LiftingNet: A Novel Deep Learning Network with Layerwise Feature Learning from Noisy Mechanical Data for Fault Classification," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 6, pp. 4973–4982, Jun. 2018, doi: 10.1109/TIE.2017.2767540.

[19] Y. S. Su, T. J. Ding, and M. Y. Chen, "Deep Learning Methods in Internet of Medical Things for Valvular Heart Disease Screening System," *IEEE Internet Things J*, vol. 8, no. 23, pp. 16921–16932, Dec. 2021, doi: 10.1109/JIOT.2021.3053420.

[20] M. Jost *et al.*, "Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs," *Nature Biotechnology 2020 38:3*, vol. 38, no. 3, pp. 355–364, Jan. 2020, doi: 10.1038/s41587-019-0387-5.

[21] G. Zampieri, S. Vijayakumar, E. Yaneske, and C. Angione, "Machine and deep learning meet genome-scale metabolic modeling," *PLoS Comput Biol*, vol. 15, no. 7, p. e1007084, Jul. 2019, doi: 10.1371/JOURNAL.PCBI.1007084.

[22] M. L. Welch *et al.*, "User-controlled pipelines for feature integration and head and neck radiation therapy outcome predictions," *Physica Medica*, vol. 70, pp. 145–152, Feb. 2020, doi: 10.1016/J.EJMP.2020.01.027.

[23] Z. Xu *et al.*, "Identifying sub-phenotypes of acute kidney injury using structured and unstructured electronic health record data with memory networks," *J Biomed Inform*, vol. 102, p. 103361, Feb. 2020, doi: 10.1016/J.JBI.2019.103361.

[24] D. Arefan, A. A. Mohamed, W. A. Berg, M. L. Zuley, J. H. Sumkin, and S. Wu, "Deep learning modeling using normal mammograms for predicting breast cancer risk," *Med Phys*, vol. 47, no. 1, pp. 110–118, Jan. 2020, doi: 10.1002/MP.13886.

[25] M. Byra *et al.*, "Knee menisci segmentation and relaxometry of 3D ultrashort echo time cones MR imaging using attention U-Net with transfer learning," *Magn Reson Med*, vol. 83, no. 3, pp. 1109–1122, Mar. 2020, doi: 10.1002/MRM.27969.

[26] S. Kazemifar *et al.*, "Dosimetric evaluation of synthetic CT generated with GANs for MRI-only proton therapy treatment planning of brain tumors," *J Appl Clin Med Phys*, vol. 21, no. 5, pp. 76–86, May 2020, doi: 10.1002/ACM2.12856.

[27] G. Liu *et al.*, "Clinically Accurate Chest X-Ray Report Generation." PMLR, pp. 249–269, Oct. 28, 2019. Accessed: Jun. 24, 2023. [Online]. Available: https://proceedings.mlr.press/v106/liu19a.html

[28] F. Piccialli, F. Giampaolo, E. Prezioso, D. Camacho, and G. Acampora, "Artificial intelligence and healthcare: Forecasting of medical bookings through multi-source time-series fusion," *Information Fusion*, vol. 74, pp. 1–16, Oct. 2021, doi: 10.1016/J.INFFUS.2021.03.004.

[29] T. Nemoto *et al.*, "Efficacy evaluation of 2D, 3D U-Net semantic segmentation and atlas-based segmentation of normal lungs excluding the trachea and main bronchi," *J Radiat Res*, vol. 61, no. 2, pp. 257–264, Mar. 2020, doi: 10.1093/JRR/RRZ086.

[30] D. C. Benz *et al.*, "Validation of deep-learning image reconstruction for coronary computed tomography angiography: Impact on noise, image quality and diagnostic accuracy," *J Cardiovasc Comput Tomogr*, vol. 14, no. 5, pp. 444–451, Sep. 2020, doi: 10.1016/J.JCCT.2020.01.002.

[31] T. D. Pham, "Classification of COVID-19 chest X-rays with deep learning: new models or fine tuning?," *Health Inf Sci Syst*, vol. 9, no. 1, pp. 1–11, Dec. 2021, doi: 10.1007/S13755-020-00135-3/TABLES/8.

[32] F. Ahmad, A. Farooq, and M. U. Ghani, "Deep Ensemble Model for Classification of Novel Coronavirus in Chest X-Ray Images," *Comput Intell Neurosci*, vol. 2021, 2021, doi: 10.1155/2021/8890226.

[33] D. Avola, A. Bacciu, L. Cinque, A. Fagioli, M. R. Marini, and R. Taiello, "Study on transfer learning capabilities for pneumonia classification in chest-x-rays images," *Comput Methods Programs Biomed*, vol. 221, p. 106833, Jun. 2022, doi: 10.1016/J.CMPB.2022.106833.

[34] T. Zebin and S. Rezvy, "COVID-19 detection and disease progression visualization: Deep learning on chest X-rays for classification and coarse localization," *Applied Intelligence*, vol. 51, no. 2, pp. 1010–1021, Feb. 2021, doi: 10.1007/S10489-020-01867-1/FIGURES/8.

[35] L. O. Teixeira *et al.*, "Impact of Lung Segmentation on the Diagnosis and Explanation of COVID-19 in Chest X-ray Images," *Sensors 2021, Vol. 21, Page 7116*, vol. 21, no. 21, p. 7116, Oct. 2021, doi: 10.3390/S21217116.

[36] 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." pp. 2097–2106, 2017. Accessed: Jun. 24, 2023. [Online]. Available: https://uts.nlm.nih.gov/metathesaurus.html

[37] S. Albahli, H. T. Rauf, A. Algosaibi, and V. E. Balas, "AI-driven deep CNN approach for multi-label pathology classification using chest X-Rays," *PeerJ Comput Sci*, vol. 7, p. e495, Apr. 2021, doi: 10.7717/PEERJ-CS.495.

[38] K. Almezhghwi, S. Serte, and F. Al-Turjman, "Convolutional neural networks for the classification of chest X-rays in the IoT era," *Multimed Tools Appl*, vol. 80, no. 19, pp. 29051–29065, Aug. 2021, doi: 10.1007/S11042-021-10907-Y/TABLES/5.

[39] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation." pp. 3431–3440, 2015.

[40] D. Zhang, L. Yang, D. Meng, D. Xu, and J. Han, "SPFTN: A Self-Paced Fine-Tuning Network for Segmenting Objects in Weakly Labelled Videos." pp. 4429–4437, 2017.

[41] W. Zhou *et al.*, "Transferring visual knowledge for a robust road environment perception in intelligent vehicles," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, vol. 2018-March, pp. 1–6, Mar. 2018, doi: 10.1109/ITSC.2017.8317747.

[42] J. Antony, K. McGuinness, N. E. O'Connor, and K. Moran, "Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks," *Proceedings - International Conference on Pattern Recognition*, vol. 0, pp. 1195–1200, Jan. 2016, doi: 10.1109/ICPR.2016.7899799.

[43] E. Kim, M. Corte-Real, and Z. Baloch, "A deep semantic mobile application for thyroid cytopathology," *https://doi.org/10.1117/12.2216468*, vol. 9789, pp. 36–44, Apr. 2016, doi: 10.1117/12.2216468.

[44] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," *International Journal of Scientific and Research Publications*, vol. 9, no. 10, p. 143, 2019, doi: 10.29322/IJSRP.9.10.2019.p9420.

[45] S. Targ, D. Almeida, and K. L. Enlitic, "Resnet in Resnet: Generalizing Residual Architectures," Mar. 2016, Accessed: Jun. 24, 2023. [Online]. Available: https://arxiv.org/abs/1603.08029v1

[46] M. Z. Alom *et al.*, "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches," Mar. 2018, Accessed: Jun. 24, 2023. [Online]. Available: https://arxiv.org/abs/1803.01164v2

[47] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9908 LNCS, pp. 630–645, 2016, doi: 10.1007/978-3-319-46493-0_38/TABLES/5.

[48] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module." pp. 3–19, 2018.