Efficient Pneumonia Detection for Chest Radiography Using ResNet-Based SVM

Marwa M. Eid and Yasser H. Elawady

Abstract — Chest radiography has a significant clinical utility in the medical imaging diagnosis, as it is one of the most basic examination tools. Pneumonia is a common infection that rapidly affects human lung areas. So, finding an advanced automated method to detect Pneumonia is assigned to be one of the most recent issues, which is still prohibitively expensive to mass adoption, especially in the developing countries. This article presents an innovative approach for distinguishing the residence of pneumonia by embedding computational techniques to chest x-rays images which eliminating the demands for single-image investigation and significantly decrease the total costs. Recent advances in deep learning achieved remarkable results in image classification on different domains; however, its application for Pneumonia diagnosis is still restricted. Hence, the main focus is to provide an investigation that will improve the research in this area, presenting a new proposal to the applications of pre-trained convolutional neural networks (CNNs) as a stage of features extraction to detect this disease. Specifically, we propose to combine deep residual neural networks (ResNets), which extract the hierarchical features from the individual x-ray images with the boosting algorithm to select the salient features, and support vector machine for classification (AdaBoost-SVM). After conducting the performance analysis on the available dataset, we have concluded that the precision of the introduced scheme in Pneumonia classification is superior to the most concurrent approaches, resulting in a great improvement in clinical outcomes.

Index Terms — Deep-learning; convolutional Neural Network; ResNet; AdaBoost; SVM; Pneumonia.

I. INTRODUCTION

Images is a major area of interest within the field of health informatics. Especially in rural zones, where we do not have enough highly experienced radiologists, such tools can be an immense help by automatically screening people who demand critical medical care or additional diagnosis. Medical imaging plays a vital role in early diagnosis and independently tailored therapy. Especially, x-ray imaging is a key part of a noninvasive medical examination which supports physicians to diagnose and monitor treating for medical conditions. It encompasses exposing a portion of the body to a small dosage of ionizing radiation to display an image that represents the inside part of the body. Remarkably, pneumonia is one of the most common principal causes of morbidity which often demands hospitalization and significant causation of mortality. In the issues of public health concern, it is attracting considerable critical attention, as its infection fulminates rapidly in the areas of human lungs. Moreover, it is considered

one of the most common type of infection found in the world. Despite the huge improvements in treatment and therapy, over two million children under 5 years of age die from pneumonia each year worldwide, rating nearly one in five under 5 deaths [1]. Being the biggest infectious murderer of children by 2030, it is expected to kill nearly 11 million children [2]. However, the determination of this serious disease is a technically challenging task that relies on the availability and accessibility of expert radiologists. It has been noted that few studies analyzed the effect of pneumonia on a human lung for early detection and diagnosis.

Medical images are abounding with the subtle features that could be significant for investigation. Usually, there are several processes in CAD systems: preprocessing, isolating Region of Interest (RoI), extracting features, and classifying the disease according to these features. Regularly, there are several methods to highlight the ROI regions, extract salient features, and suppress the related noises. The rule-based approaches have weak performance, and they are usually combined to increase the classification accuracy. In terms of extracting features, the obtained features by conventional techniques like geometric features, shape, texture, or handcrafted features that are regularly handled to decrease the dimensionality and features redundancy, can have errors that influence the classification performance. Thus, the stage of feature extraction and the techniques of improvement and segmentation are very critical. Furthermore, deep CNNs hold distinctive characteristics that indorse them natural applicants when involving medical image classification tasks. Pre-trained networks perform rather well in classifying images that have many apparent differences [3], such as abdominal x-rays and need only a sufficient amount of training dataset. When applying the pre-trained networks to medical image classification tasks, they might need more expensive training phases with a bigger dataset to accomplish reasonably a good result. For such cases, the use of pre-trained networks as a classifier is frequently not the favored form to employ CNNs to CAD diagnosis tasks. By using pre-trained networks like DCNNs, extra difficult datasets such as the presence/absence Pneumonia, did not seem well. Hence, more training samples and data augmentation could increase efficiency [4]. While the CNN-based identification systems introduce higher accuracy in various fields, the major drawback of the CNN-based approach is that more intensive training is needed, which takes extremely processing time with much training dataset.

In this paper, we investigate the discrimination power of the x-ray image representations derived from adopting the extracted features from the last convolutional layer of a 50-

layer deep residual network. DNNs are responsive to vanishing gradients that result in intractable learning. Especially, ResNets employ the skip connections to disseminate information over layers, which allows data to build deeper networks. Skipping the connections helps the network to understand the global features and detects the smaller objects in the chest x-ray image. So, thanks to the recommended ResNet as a feature extractor module, the stacking additional layers can build a deeper network. Thus, it can offset the vanishing gradient and dip the layers of it to be less relevant through the training phase. As the extracted feature vectors are large in dimension, we recommend using AdaBoost for reducing dimensionality. Therefore, results are presented to confirm the efficiency of ResFeats with state-of-the-art identification accuracies. Not only extracting features but also selecting features, AdaBoost gives a promising and fulfilled achievement. Moreover, there is an added gain that it is simply to be performed and fitted to a real-time automatic human-computer interaction in artificial intelligent systems. With the excellent achievement of deep learning in extracting features or classification, the proposed deep learning-based approaches have consistently become a great solution. Due to the high variance and low bias of the introduced model, it can allow the classifier to reach lower EER for the investigation process. Thus, we confirm that ResFeats accomplish a superior identification accuracy in comparison with off-the-shelf pre-trained CNN structures i.e. (VGG 19, AlexNet, and GoogLeNet) and other non-CNN based methods. This improves the capability for physicians to assess the area of treatment and give great support to the radiologist for reducing workload.

The rest of this paper is organized as follows; Section II overviews the related work. Section III describes the materials and methods used in the experiment. Section IV describes the proposed system used. Section V presents the discussion of the obtained results. Finally, section VI introduces the conclusions of the work presented in the paper.

II. RELATED WORK

Nowadays, the accessibility of huge datasets and the latest enhancements in deep-learning models, led to have power-assisted algorithms to beat medical personnel in various medical imaging tasks like classification of carcinoma [5], [6], identification of hemorrhage [7], detection of arrhythmia [8], and detection of diabetic retinopathy [9].

Automatic diagnosis of chest diseases through radiographs have received huge interests and attentions. These approaches are growingly used for detecting lung nodule [10] and classifying pulmonary tuberculosis [11]. The performance of several deep convolutional network models on different abnormalities that exist in the OpenI dataset [12] does not perform well with all abnormalities [13], ensemble models improve classification accuracy significantly when compared with single model, and finally, using deep-learning methods enhanced the classification accuracy rather than using rule-based methods.

Statistical dependency was studied between Multi-label Disease Classification (MDC) so that, the accuracy levels of the predictions are obtained [14]-[16]. The detection of diseases from x-ray images were accomplished in [17]-[19],

the image classifications from chest x-ray were fulfilled in [20], and computed tomography from chest x-ray images from body parts segmentation was carried out in [5], [17]. On the contrary, learning image selections and developing image descriptions relative to what someone's would describe have to be utilized.

An overview of the applications that use deep learning techniques for the chest x-ray images analysis are listed in Table I [21].

TABLE I: OVERVIEW OF APPLICATIONS THAT USE DEEP LEARNING TECHNIQUES FOR THE CHEST X-RAY IMAGES ANALYSIS [21]

Application	Remarks
Detection of Nodule	Classification with CNN two-layer, each with 125×5 filters.
Retrieval of Image	Combination of classical features with pre-trained CNN features for retrieval of image by using SVM.
Detection of Pathology	Take features from a pre-trained CNN and detect various diseases from a low level features.
Retrieval of Image	Continuation of Anvi adding features of age and gender
Detection of Pathology	Continuation of Bar, adding more experiments and feature selection.
Detection of Pathology	Detecting five common abnormalities, validated and trained on a big-data set using GoogLeNet CNN
Detection of Tuberculosis	Using 6 convolution layers pre-trained fine-tuned network to process entire radiographs.
Detection of Tuberculosis	Producing heat map of suspicious regions via de-convolution using MIL framework.
Detection of Pathology	Using huge data set (seven k images), frequent networks produce short captions, and CNN detects seventeen diseases.
classification of Frontal/lateral	Classification of Frontal/lateral using Pre-trained CNN.
Bone suppression	Using cascaded CNNs at high resolution to learn bone images from radiographs.
Nodule classification	Combination of CNN features and classical features from a pre-trained ImageNet CNN.

Several Recent papers have introduced an automated system to detect pneumonia from chest x-ray images. Deep learning used by Yuan Tian in 2017 [22] to train an AI algorithm to detect pneumonia through analyzing the chest x-ray images. The CNN classifier achieved an accuracy of 91%. A proposed CNN model introduced by Okeke Stephen et al. in 2019 [5], the proposed CNN model is constructed to detect pneumonia through extracting features from chest x-ray images. They deployed several data augmentation algorithms to improve the validation and classification accuracy of the CNN model. The classifier achieved an accuracy of 95.31%.

III. MATERIALS AND METHODS

In this paper, a proposal will be presented for the application of pre-trained CNNs in Pneumonia detection. Therefore, a new paradigm might have been required for diagnosing diseases through CNNs models. In this situation, we have proposed a method for the classification of Pneumonia disease images using a deep residual CNN structure in combination with a multiple instance learning (MIL) routine to get a more viable way for fulfilling this task. In the proposed proposal, a CNN architecture is used to extract features from a resized radiographic chest image. Then, the previously extracted features by a pre-trained network will be used to train a strong classifier to reach a new paradigm that

can accomplish a superior result. In terms of selecting the best classifier, the performance of support vector machine (SVM) from traditional algorithms may be more beneficial.

All experiments in this paper are performed under Windows 10 (Microsoft Corporation) installed on a DELL LATITUDE E6410 personal computer with 4 GB internal working memory equipped with a Core i5 processor M520 (Intel Corporation) and 64-bit operating system.

A. Dataset

Kaggle dataset Chest x-ray Images (Pneumonia) [23] is used in this paper. The dataset consists of 5,863 x-ray images in (JPEG) format. It is categorized in two cases *Pneumonia* or *Normal*. The Kaggle data set selected to be used in this paper because it is used globally in many researches and able to make a good comparison enriching scientific research.

Fig. 1 shows normal image samples without Pneumonia while Fig. 2 shows image samples with Pneumonia.

Fig. 1. A sample for normal images without Pneumonia.

Fig. 2. A sample for images with Pneumonia.

B. Convolutional Neural Network (CNN)

CNN is a type of deep neural networks (DNN) which is specialized in image analysis and achieving higher accuracy than prior methods in the detection of diseases. Thus, it is used widely in computer vision applications such as clustering, object detection, and image classification.

The following part outlines several CNN models such as AlexNet [24], [25], VGG-Net [26], GoogLeNet [27] and ResNet [28]. These models have different numbers of convolution layers used. In CNN models, if the number of convolution layers increased the higher classification accuracy achieved.

AlexNet [25]: Krizhevsky proposed AlexNet for the first time in 2012 and won the (ILSVRC) [28] in this year which is ImageNet Large-Scale Visual Recognition Challenge. It is composed of three fully connected layers and five convolutional layers.

VGGNet [26]: In 2014, VGGNet won the ILSVRC competition for localization and classification tracks. It has

two popular architectures: VGGNet-19 and VGGNet-16. It has simple architecture, which makes it used in wide range. It has three fully connected layers, five pooling layers, and 13 convolutional layers.

GoogLeNet [27]: It is another model of CNN architecture. It has two main advantages. The first, GoogLeNet used different sizes of filter kernels at the same layer. The second; it uses a reduced number of network parameters, which allows it to be deeper and makes it less sensitive to over fitting. Finally, GoogLeNet has twelve times fewer parameters than AlexNet.

ResNet (Residual Network) [28]: It is one of the most efficient models of CNNs. In the Conference of Computer Vision and Pattern Recognition (2016), ResNet was announced to be the best paper. The ResNet idea is that each layer should only learn a residual correction of the previous layer not the whole feature space transformation, which gives a deeper networks training that works efficiently. Fig. 3 shows the residual blocks that ResNets built on. The ResNets main idea is to add the activation of layer (1) to the output of layer (1+2) which is called identity shortcut connection [22]. ResNet allows us to train much deeper neural networks without suffering from gradient vanishing and exploding issues. In this Paper, we will use ResNet50 pre-trained on the ImageNet classification dataset.

Fig. 3. Residual block in ResNet.

C. Adaptive Boosting (AdaBoost) Algorithm

Yoav Freund and Robert Shapire proposed the AdaBoost algorithm in 1995 [29]. The algorithm maintains a collection of weights over training data to create a set of poor learners and adaptively adjusts them after each weak learning cycle. The weights of the training samples which are correctly classified will be decreased while the weights of misclassified by current weak learner will be increased [29].

The advantages of AdaBoost algorithm:

- Fast converging.
- Easy to be implemented by machine learning algorithm.
- Does not require any previous knowledge about the weak learner.
- Can be easily combined with other algorithms such as support vector machine to find weak hypothesis.

In many applications, AdaBoost algorithm can be used as a classifier and feature selection or extraction. In this paper AdaBoost algorithm is used as feature selection.

D. Support Vector Machine (SVM) Classifier

Vapnik developed SVM [29] from Structural Risk Minimization theory. SVM is considered as a supervised machine-learning algorithm and it is used in solving many classification problems. Moreover, it is known as a binary classification algorithm, which transforms the data and according to these transformations it finds the optimal boundary between the possible outputs, this technique is called the kernel trick. Non-linear SVM means that the boundary that the algorithm calculates does not have to be a straight line. It is beneficial because, you do not have to carry out complex transformations by yourself and you can capture many relationships that are more complex between your data points.

SVMs are important for the following reasons [29]:

- Powerful in the case of having very large number of variables and small samples.
- Both simple and highly complex classification models can be learned.
- avoiding over-fitting by utilizing advanced mathematical principles.

Many researchers attempt to employ ensemble methods, such as conventional Bagging and AdaBoost [29] to improve the classification performance of the real SVM. Finally, yet importantly, SVM is known as a stable and strong classifier.

IV. PROPOSED SYSTEM

An overview of the procedure implemented in the proposed system is shown in Fig. 4.

A. The Deep Residual CNNs Structure

ResNet-50 model [30]-[32] has been adopted to extract the salient features from chest x-rays images by altering the nodes in the fully connected layer and performing the fine-tuning process using the experimental dataset. Moreover, to apply the ResNet-50 model throughout the fine-tuning; the input should be resized to 224×224 pixels for each image [33]. Hence, the size of the original images has been changed in the proposed system to be suitable for the recommended structure as introduced in Fig. 5 and Table II.

Moreover, after these pre-stages, (i.e., resizing and zero-center normalization), the generated image is used as the final input for the proposed CNN. The structure of the adopted CNN regarding the number of layers and filters and their specifications are selected identical to the model of ResNet-50 CNN as in [33] excepting the number of nodes in the output FC layer. This research focuses on classifying two categories, thus the output nodes in the FC layer is set to two (Pneumonia or Normal).

Fig. 4. Overview of the proposed system.

Fig. 5. The CNN structures.

Table II presented the main features of convolutional layers.

In the feature extraction step, the pooling or subsampling layers (Max pool and AVG pool) select the values of maximum and average pixels to get the salient features [33], [34]. From Table II, the convolutional layers are (Conv1 ~ Conv5).

As exhibited in Fig. 5, Conv3 ~ Conv5 are holding the structure of a bottleneck where the size of channels before and after it is 256 and 512 respectively. Then, for practical considerations and more efficient in computations during the convolution operation, they are diminished to 128 by the filters of size $(3 \times 3 \times 128)$. Depending on the number of iterations in Table II, the layers (Conv2 ~ Conv5) are iterated

to operate consecutively. Furthermore, the holding data in the features map preceding to Conv2-1, Conv3-1, Conv4-1, and Conv5-1 is individually summed elementwise over the alternative layer into the output maps (i.e., Conv2-3 ~ Conv5-3) as shown in Fig. 5. Table II shows the deep details of the ResNet-50 structure. In Table II, "3*" denotes to 3 pixels are included in the padding process in these positions (left, right, up, and down) of the input while "1*" is indicated that 1-pixel

is involved in left, right, up, and down positions of feature map. where "2/1**" implies that 2 at the first iteration and 1 from the second iteration. Hence, the ResNet-50 structure can hold more accurate filters and classifiers through applying the batch normalization [35] process after every convolutional layer then employing the ReLU [36] as an activation function.

TABLE II: CONVOLUTION LAYERS DATA OF THE IMAGE

No. of iterations	Padding Amount	No. of strides	Filter Size	Filters No.	The feature map size	Layer
1	3*	2	7 x 7 x 3	64	224 x 224	input layer
1	0	2	3 x 3	1	112 x 112	Conv 1
3	0	1	1 x 1 x 64	64	56 x 56	Max pool
	1*	1	3 x 3 x 64	64	56 x 56	Conv2-1
	0	1	1 x 1 x 64	256	56 x 56	Conv2-2
						Conv2-3
	0	1	1 x 1 x 64	256	56 x 56	Conv2-4 (shortcut)
4	0	2/1**	1 x 1 x 256	128	28 x 28	Conv3-1
	1*	1	3 x 3 x 128	128	28 x 28	Conv3-2 (Bottleneck)
	0	1	1 x 1 x 128	512	28 x 28	Conv3-3
	0	2	1 x 1 x 256	512	28 x 28	Conv3-4 (shortcut)
6	0	2/1**	1 x 1 x 512	256	14 x 14	Conv4-1
	1*	1	3 x 3 x 256	256	14 x 14	Conv4-2 (Bottleneck)
	0	1	1 x 1 x 256	1024	14 x 14	Conv4-3
	0	2	1 x 1 x 512	1024	14 x 14	Conv4-4 (shortcut)
3	0	2/1**	1 x 1 x 1024	512	7 x 7	Conv5-1
	1*	1	3 x 3 x 512	512	7 x 7	Conv5-2 (Bottleneck)
	0	1	1 x 1 x 512	2048	7 x 7	Conv5-3
	0	2	1 x 1 x 1024	2048	7 x 7	Conv5-4 (shortcut)
1	0	1	7 x 7	1	1 x 1 x 2048	AVG pool
1					2	FC layer
1					2	softmax

B. Feature Extraction Via Convolutional Layers

By enforcing a typical 2D convolution procedure to each image, the salient features can be extracted. Furthermore, the area to be considered in an image changes depending on the size of each filter which had been involved. Moreover, depending on the padding options, the filter investigation and both the height and the width of resulting images change [36]. Hence, the basic factors to be considered are the number of filters, filter sizes, strides, and padding options. As exhibited in Table II, Conv1 consists of 64 filters of (7×7×3) filter size and investigates in the both directions while striding by a unit which is composed of two-pixels. Moreover, its padding process is employed in a unit of three-pixels to each direction. While the Max pool consists of a single (3×3) filter to investigate the directions of vertical and horizontal pixels and stride by a 2-pixels unit. As presented in Table II, Conv2-1 gives the convolution process with a number of 64 filters whose sizes are (1×1×64) and investigates in each direction and strides by a single pixel unit. Furthermore, Conv2-2 makes the process of convolution with (3×3×64) filters which explore in every direction and stride by a 1-pixel unit with the padding of a single pixel also. Within (Conv2 ~ Conv5), the obtained features are in two categories as shown in Fig. 5. The first is the sequentially operating convolutional layers (Conv2 ~ Conv5). While the second category is the residual information in the features map ahead of (Conv2-1 ~ Conv5-

1). In addition, it is element-wise features which embedded in the shortcut layer within the output map (Conv2-3 ~ Conv5-3) as presented in Fig. 5 and Table II. Moreover, through applying smaller filter size as in the case of (1×1), the filter parameters demanding in the training operation are remarkably decreased.

As illustrated, the fine-tuning had been performed on each layer from Conv-1 to (FC layer) as shown in Table II. Moreover, following to separate convolutional layer, the batch normalizing is accomplished depended on the computed average value and the result of the standard deviation to the whole data. In addition, the activation function (ReLU) is subsequent to every batch normalizing procedure. The process, which is employed to get the result for a certain input in a ReLU layer as in (1):

$$z = \max(0; y) \quad (1)$$

where y is the ReLU input function and z is the value of the output. By applying equation 1, the z ranges can be demoted to zero or small positive values, which can gain easier training to the CNN model. Furthermore, in the case of positive value to the variable y, the output z is likewise to y, and the first order derivative is one, which marks more manageable for the training mathematical equation. Simply, this equation can restrict the problems of vanishing gradient [37] which is occurred in employing a sigmoid function or hyperbolic

tangent in learning within the back propagation [38]. Moreover, it exceeds the non-linear activation function in the computation time. Thus, the learning efficiency totally raises because of decreasing the elapsed time in learning. Furthermore, regarding Conv5, the AVG pool layer is applied, where a 7×7 filter is used by a stride of single pixel. Therefore, the extracted feature map will be passed as the input to the following step is (1×1×2048).

The classification techniques of different types of Deep Convolutional Neural Networks (DCNNs) used in this paper are visualized in Fig. 6.

Fig. 6. Different types of DCNNs used in the classification.

V. RESULTS AND DISCUSSION

For the datasets, the images had been separated randomly into (60%) images for training, (20%) in validation, and (20%) images to apply the testing phase. The training data were employed to teach the network while the validation data were used to model selection, and the testing data to evaluate model efficiency on unforeseen cases. Several performance metrics have been computed to assess the performance such as, elapsed time, accuracy, sensitivity, false-positive rate, precision, and F-score. This metrics can be defined as follows:

Accuracy [39]: it measures the ability of a predictor to identify the whole samples correctly, regardless being positive or not. It is calculated by (2):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (2)$$

where, TP is the value of the true positive, FN is false negative value, TN is true negative, and FP is the rate of the false positive.

Sensitivity (recall) [39]: is the true positive rate (TPR). It computes the capability of recognizing the positive samples which is calculated by (3):

$$TPR = \frac{TP}{TP+FN} \times 100 \quad (3)$$

Specificity [39]: is the true negative rate (TNR).It gets the capability in finding the negative samples which is calculated by (4):

$$TNR = \frac{TN}{TN+FP} \times 100 \quad (4)$$

False positive rate (FPR) [39]: is the fall-out value and measures incorrect portions. It is calculated by (5):

$$FPR = \frac{FP}{FP+TN} \times 100 \quad (5)$$

Precision [39]: refers to positive predictive value (PPV) and it directs the rate of true positives amongst all positive values. It is calculated by (6):

$$PPV = \frac{TP}{TP+FP} \times 100 \quad (6)$$

F-score [39]: measures a harmonic mean of precision and sensitivity (recall). It is calculated by (7):

$$F_{Score} = 2 * \frac{PPV * TPR}{PPV + TPR} \times 100 \quad (7)$$

The results of True Positive Rate (TPR), True Negative Rate (TNR), False Predictive Rate (FPR), and Positive Predictive Value (PPV), achieved from the experiments are listed in Table III.

TANLE III: TPR, TNR, FPR, AND PPV FOR THE DIFFERENT USED CLASSIFICATION METHODS

Methodology	TPR	TNR	FPR	PPV
AlexNet	90.5473	97.8495	1.9900	97.8495
GoogLeNet	90.1615	91.4394	8.5606	91.5735
Vgg19	74.1294	75.12	0.2488	99.6656
Resnet50 +PCA+SVM	86.6431	89.03	9.0766	91.4243
Resnet50 + (AdaBoost_SVM) (proposed)	96.4029	99.8	0	100

The ROC curve of the proposed method is shown in Fig. 7. The proposed method shows that, the area under the curve (AUC) is equal to maximum and equal 100%.

Fig. 7. ROC curve of proposed method.

The results of Accuracy, F_score and elapsed time achieved from the experiments are listed in Table IV.

From Table IV, we notice that the accuracy of the proposed technique is 98.13 % which is the better compared with the other techniques listed.

TABLE IV: ACCURACY , F_SCORE AND ELAPSED TIME FOR THE DIFFERENT USED CLASSIFICATION TECHNIQUES

Methodology	Accuracy	F-score	Elapsed Time
<i>AlexNet</i>	94.2786	94.0568	176 mins.
<i>GoogLeNet</i>	90.7905	90.8620	340 mins.
<i>Vgg19</i>	86.9403	85.0214	14881.8 mins.
<i>Resnet50 +PCA+SVM</i>	88.6652	88.9695	447.82 sec
<i>Resnet50 + (AdaBoost_SVM)</i> (proposed)	98.1343	98.1685	208.13 sec

Table V represents a comparison between the proposed technique and other researcher's techniques for detecting pneumonia. **Yuan Tian** used the same dataset that we are used and give an accuracy of 91% for detecting the pneumonia with ResNet_34 technique. **Okeke Stephen** used another dataset and give an accuracy of 95.31% for detecting the pneumonia with their own technique.

TABLE V: A COMPARISON BETWEEN THE PROPOSED TECHNIQUE AND OTHER RESEARCHER'S TECHNIQUES

Method	Data set	Technique	Accuracy
[22]	5,863 chest-x ray images, Kaggle	CNN (ResNet34)	91%
[5]	5,856 chest x-ray images, retrospective pediatric patients, from 1 to 5 years old.	CNN (Own algorithm)	95.31%
Proposed Method	5,863 chest-x ray images from Kaggle	CNN (Resnet50 + Quadratic SVM)	98.13%

VI. CONCLUSION

Pneumonia is considered as one of the essential causes of morbidity, which often demands hospitalization. Finding the most effective scheme to detect the Pneumonia from chest x-ray images has taken critical attention, as its infection fulminates quickly the human lungs. To overcome the time-consuming scheme of training CNN and the low accuracy of MIL based methods, the proposed scheme for classifying Pneumonia disease images using a deep residual CNN structure in combination with an (AdaBoost -SVM) classifier introduced higher performance and lower error rate when compared with the common related works and traditional algorithms. Other than the diverse CNN structures, ResNets are more accessible to train. However, residual deep learning and the identity mappings in ResNets have been shown to achieve outstanding results in feature detection. Very deep networks are renowned to induce overfitting and saturation inaccuracy. However, residual learning and identity mappings in ResNets have been proved to defeat these problems. Moreover, ResNet features hold distinctive characteristics that make them natural candidates when applied to the classification step. Not only extracting features is critical but also selecting features, the recommended AdaBoost technique gives a fulfilled accomplishment to fit a real-time and fast CAD system. From the simulation results, we confirm that our recommended features are powerful and have a classification accuracy that is higher than the CNN off-the-shelf features. Finally, this approach effectively investigates Pneumonia diseases of the area in the chest x-ray image and increases the potential diagnosis of the missing obstacles of the pneumothorax area. Generally, our future work might be involved in improving the detection of multi-grade disease diagnosis.

REFERENCES

- [1] Lakhani, P., & Sundaram, B. "Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks". *Radiology*, 284(2), pp.574–582. 2017.
- [2] Sharma, Raju, A., D., & Ranjan, S. "Detection of pneumonia clouds in chest x-ray using image processing approach". *International Conference on Engineering, Nirma University ,NUICONE* 2017.
- [3] Lakhani, P. "Deep Convolutional Neural Networks for Endotracheal Tube Position and x-ray Image Classification : Challenges and Opportunities". *Digital Imaging Journal*, 30(4), pp.460–468. 2017.
- [4] City, Q., & Hutchison, D. "Deep Learning in Medical and Multimodal Learning", (Vol. 3). 2017.
- [5] Stephen, O., Maduh, Sain, M., U. J., & Jeong, D. U. "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare". *healthcare engineering Journal*, 2019.
- [6] Esteva, A., Novoa, Kuprel, B., R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. " Dermatologist-level classification of skin cancer with deep neural networks". *Nature*, 542(7639), 115. 2017.
- [7] Grewal, M., Srivastava, M. M., Kumar, P., & Varadarajan, S. Radnet: "Radiologist level accuracy using deep learning for hemorrhage detection in ct scans". *IEEE International Symposium on Biomedical Imaging (ISBI)*. pp. 281-284. April 2018.
- [8] Rajpurkar, P., A. Y., Hannun, Haghanahi, M., Bourn, C., & Ng, A. Y. "Cardiologist-level arrhythmia detection with convolutional neural networks". *arxiv :1707.01836*, 2017.
- [9] Ibrahim, A., El-kenawy, E. S. M. . Image Segmentation Methods Based on Superpixel Techniques: A Survey. *Journal of Computer Science and Information Systems*, 15 (3 October 2020). (2020).
- [10] Ibrahim, A., El-kenawy, E. S. M. . Applications and Datasets for Superpixel Techniques: A Survey. *Journal of Computer Science and Information Systems*, 15 (3 October 2020). (2020).
- [11] Lakhani, P., & Sundaram, B. "Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks". *Radiology*, 284(2), pp.574–582. 2017.
- [12] El-Kenawy, El-Sayed M., Abdelhameed Ibrahim, Seyedali Mirjalili, Marwa Metwally Eid, and Sherif E. Hussein. "Novel feature selection and voting classifier algorithms for COVID-19 classification in CT images." *IEEE Access* 8 (2020): 179317-179335
- [13] Islam, M. T., Aowal, M. A., Minhaz, A. T., & Ashraf, K. "Abnormality detection and localization in chest x-rays using deep convolutional neural networks". *arxiv :1705.09850*. 2017.
- [14] Yao, L., Poblentz, E., Dagunts, D., Covington, B., Bernard, D., & Lyman, K. " Learning to diagnose from scratch by exploiting dependencies among labels ". *arxiv :1710.10501*. 2017.
- [15] Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases". *IEEE conference on computer vision and pattern recognition*, pp. 2097-2106. 2017.
- [16] Shin, H. C., Lu, L., Kim, L., Seff, A., Yao, J., & Summers, R. M. "Interleaved Text/Image Deep Mining on a Large-Scale Radiology Database for Automated Image Interpretation". *Journal of Machine Learning Research*, 17, pp.1-31, 2. 2016.
- [17] Fouad, Mohamad M., Ali Ibrahim El-Desouky, Rami Al-Hajj, and El-Sayed M. El-Kenawy. "Dynamic group-based cooperative optimization algorithm." *IEEE Access* 8 (2020): 148378-148403.
- [18] Melendez, J., van Ginneken, B., Maduskar, P., Philipsen, R. H., Reither, K., Breuninger, M., & Sánchez, C. I. "A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest x-rays". *IEEE transactions on medical imaging*, 34(1), pp.179-192. 2014.
- [19] Jaeger, S., Karagyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., & Thoma, G. " Automatic tuberculosis screening using chest radiographs ". *IEEE transactions on medical imaging*, 33(2), pp.233-245. 2013.
- [20] Xue, Z., You, D., Candemir, S., Jaeger, S., Antani, S., Long, L. R., & Thoma, G. R. " Chest x-ray image view classification". *IEEE International Symposium on Computer-Based Medical Systems*. pp. 66-71. June ,2015.
- [21] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. " A survey on deep learning in medical image analysis". *Medical image analysis*, 42, pp.60-88. 2017.
- [22] Hassib, Eslam M., Ali I. El-Desouky, Labib M. Labib, and El-Sayed M. El-kenawy. "WOA+ BRNN: An imbalanced big data classification framework using Whale optimization and deep neural network." *soft computing* 24, no. 8 (2020): 5573-5592.
- [23] <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>,

- [24] Han, J., Zhang, D., Cheng, G., Liu, N., & Xu, D. "Advanced deep-learning techniques for salient and category-specific object detection: a survey". *IEEE Signal Processing Magazine*, 35(1), pp.84-100. 2018.
- [25] El-kenawy, E. S. M. T. "A Machine Learning Model for Hemoglobin Estimation and Anemia Classification." *International Journal of Computer Science and Information Security (IJCSIS)* 17, no. 2 (2019).
- [26] Simonyan, K., & Zisserman, A. "Very deep convolutional networks for large-scale image recognition". *arxiv arxiv:1409.1556*. 2014.
- [27] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. "Going deeper with convolutions". *IEEE conference on computer vision and pattern recognition*. pp. 1-9., 2015.
- [28] He, K., Zhang, X., Ren, S., & Sun, J. "Deep residual learning for image recognition". *IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.
- [29] Wang, R." AdaBoost for feature selection, classification and its relation with SVM, a review". *Physics Procedia*, 25, pp.800-807. 2012.
- [30] Jung, H., Choi, M. K., Jung, J., Lee, J. H., Kwon, S., & Jung, W. Y. "ResNet-Based Vehicle Classification and Localization in Traffic Surveillance Systems". *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. pp. 934-940. July ,2017.
- [31] Zagoruyko, S., & Komodakis, N. "Wide Residual Networks". Retrieved from *arxiv.org*, 2016.
- [32] Mahmood, A., M., Bannamoun, An, S., & Soheli, F. "Resfeats: Residual network based features for image classification". *International Conference on Image Processing, ICIP*. pp.1597-1601. 2018.
- [33] He, K.; Zhang, X.; Ren, S.; Sun, J. "Deep residual learning for image recognition". *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp. 770-778. June, 2016.
- [34] Lecun, Y.; Bottou, Bengio, L.; Y.; Haffner, P." Gradient-based learning applied to document recognition". *Proc. IEEE*. 86, pp.2278-2324. 1998.
- [35] Ioffe, S.; Szegedy, C. "Batch normalization: Accelerating deep network training by reducing internal covariate shift". In *Proceedings of the International Conference on Machine Learning*, Lille, France, pp. 448-456. July, 2015.
- [36] E.-S. El-Kenawy and M. Eid, "Hybrid gray wolf and particle swarm optimization for feature selection," *International Journal of Innovative Computing Information and Control*, vol. 16, no. 3, pp. 831-844, 2020.
- [37] Glorot, X.; Bordes, A.; Bengio, Y. "Deep sparse rectifier neural networks". *14th International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, FL, USA. pp. 315-323. April, 2011.
- [38] El-Kenawy, El-Sayed M., Marwa Metwally Eid, Mohamed Saber, and Abdelhameed Ibrahim. "MbGWO-SFS: Modified binary grey wolf optimizer based on stochastic fractal search for feature selection." *IEEE Access* 8 (2020): 107635-107649.
- [39] Jiao, Y., & Du, P. "Performance measures in evaluating machine learning based bioinformatics predictors for classifications". *Quantitative Biology*, 4(4), pp.320-330. 2016.

Marwa M. Eid received the Ph. D. degree in electronics and communications engineering from the Faculty of Engineering, Mansoura University, Egypt, in 2015. She has been an Assistant Professor with the Delta Higher Institute for Engineering and Technology, since 2011. Her current research interests include image processing, encryption, wireless communication systems, and field programmable gate array (FPGA) applications.

Yasser H. Elawady is Assistant Professor in the computer engineering Department, MISR higher institute of engineering and Technology, he received the PHD from faculty of computers and information, Mansoura University, Egypt. He received his M.Sc. from the Department of Computer Engineering, Faculty of Engineering, Mansoura university, Mansoura, Egypt, in 2003. His subject of interest includes Remote Access, FPGA programming, Hardware Design, computer architecture and organization and Networking.