Diagnosing Heart Disease Types From Chest X-Rays Using A Deep Learning Approach

1st Roozbeh Sadeghian [1]
Data Analytics
Harrisburg University of Sceince and Technology
Harrisburg, US
RSadeghian@harrisburgu.edu

1st Siamak Aram
Information System Engineering and Managemnet
Harrisburg University of Sceince and Technology
Harrisburg, US
SAram@harrisburgu.edu

2nd Iheb Abdellatif
Information System Engineering and Managemnet
Harrisburg University of Sceince and Technology
Harrisburg, US
IAbdellatif@harrisburgu.edu

3rd Stanley Nwoji
Information System Engineering and Managemnet
Harrisburg University of Sceince and Technology
Harrisburg, US
SNwoji@harrisburgu.edu

Abstract— **A chest x-ray is the most commonly performed diagnostic examination of heart disease. The interpretation of chest radiographs is crucial to detect a variety of conditions that affect millions all around the world every year. These diseases include cancer, heart and lung diseases, tuberculosis, fibrosis, and others. Trained radiologists perform these difficult time-consuming and challenging interpretation tasks. Misdiagnosis can occur. Computer-aided techniques could lead to more accurate and accessible diagnoses. The present paper used the VGG16 architecture to develop a deep convolutional neural network, which addressed the issue of ensuring accurate diagnoses of various diseases in chest radiography. The model performed well for chest radiography, resulting in high levels of diagnostic accuracy and sensitivity and the ability to classify 14 different diseases.**

Keywords— Chest X-Ray, Diagnostic Radiography, Deep learning

I. Introduction

Chest x-rays aid radiologists and other doctors in classifying diseases [1, 2]. The different diseases that can be classified using chest x-rays include the following: cardiomegaly [3], emphysema [4], edema [5], hernia [6], pneumothorax [7], effusion [8], mass [9], fibrosis [10], atelectasis [11], consolidation [12], pleural thickening [13], nodule [14], pneumonia [15], and infiltration [16]. Chest x-rays, though the most common and accessible diagnostic tools available to radiologists [17], take time to evaluate and can lead to misdiagnosis [3]. There is, therefore, the need to train computers to aid doctors in making faster and more accurate diagnoses. This paper, therefore, seeks to classify these diseases detected from chest x-rays using Machine Learning-based approaches. Chest x-ray data used in this work were extracted from ChestX-ray14 [17], a dataset compiled by the National Institutes of Health from over 30,000 patients.

Our team developed a model that helped computers detect and classify these diseases automatically. This model can also aid doctors and other medical practitioners in several ways. First, the model will serve as adjuncts to the diagnosis made by doctors through either confirming the diagnosis or helping doctors avoid misdiagnosis. With the availability of huge numbers of chest x-ray datasets, analytic techniques will help interpret data to aid the doctor in making better decisions that will prevent, cure, and manage diseases more efficiently. Second, this research will help clinicians interpret chest x-rays rapidly and accurately [18]. Third, it has the potential of empowering patients to make medical decisions for themselves or, at least, in partnership with medical practitioners. Fourth, it will lead to a reduction of medical errors

The use of deep learning in the analysis of chest x-rays is on the increase [19, 20]. There is now a paradigm shift in medical practice towards the analysis of chest x-ray datasets using the application of artificial intelligence and machine learning. However, existing literature on analysis of chest x-rays using deep learning focuses on identification of one or a few diseases. Few articles concentrate on the analysis of chest x-rays using deep learning to classify and identify all the 14 diseases for which chest x-rays are carried out [19, 20]. This paper addressed this gap.

II. Related Works

Chest x-rays aid the diagnosis of 14 different diseases [21, 22]. A misdiagnosis, therefore, would affect the lives of a great many patients as well as the reputation of the medical institution. There are numerous articles on how deep learning aids the diagnosis of diseases using chest x-rays. Most are still in the development stage and need further clinical evaluation [23].

A. Chest X-Rays in Medical Diagnosis

A chest x-ray is inexpensive, fast, painless, and has few risks, yet it aids doctors in the diagnosis of heart and lung diseases, cancer, tuberculosis, fibrosis, etc. [24, 25]. The Mayo Clinic shows that chest x-rays can detect the condition of the lungs, heart-related lung conditions, the size and outline of the heart, blood vessels, calcium deposits, fractures, postoperative changes, pacemakers, defibrillators or catheters, and even help determine if treatment is progressing as it should [25]. Many studies have confirmed the importance of chest x-rays as a diagnostic tool [26].

[1] **Co-first Authors (Siamak Aram, Roozbeh Sadeghian) and Corresponding Author (Stanley Nwoji)**

978-1-7281-5584-5/19/$31.00 ©2019 IEEE
DOI 10.1109/CSCI49370.2019.00173

B. Diagnosing Heart Disease Types Using Chest X-rays

Ischemic heart disease kills more people in the world today than other diseases [27]. This fact makes faultless diagnosis of diseases associated with the heart essential. Existing literature suggests that 14 different heart diseases can be diagnosed using chest x-rays [28] [29] including atelectasis, cardiomegaly, effusion, infiltration, mass, nodule, pneumonia, pneumothorax, consolidation, edema, emphysema, fibrosis, pleural thickening, and hernia.

C. Deep Learning

Deep Learning, a subset of Machine Learning, is a branch of artificial intelligence [30]. In artificial intelligence, machines are made to act as intelligent humans, and in Machine Learning, algorithms learn from data without programming. Deep Learning uses Artificial Neural Networks to mimic the human brain [30]. Deep Learning led to a paradigm shift from computerized processing of images, audios, speeches, and videos, which, until now, were human-centric activities [31]. The "deep" in "Deep Learning" refers to the multiple hidden layers between inputs and outputs in its architecture. These hidden layers are mathematical operations on inputs. The output of each layer serves as an input for the following layer.

D. Deep Learning in Medical Diagnosis

Medicine and technology are compatible; Deep Learning is proving that. Deep learning can enhance diagnosis and treatment and can relieve medical practitioners from tasks that hinder practitioner-patient relationships [32]. Convolutional Neural Network, a type of deep neural network, is designed explicitly for mostly image analytics. Its accuracy and diagnostic capabilities are now approaching or, sometimes, surpassing those of humans [33]. Deep Learning is now crucial to clinicians in diagnosing and treating heart and lung diseases, cancer and diabetes management, and in managing neurological conditions [33].

E. Diagnosing Heart Disease Types from Chest X-Rays Using Deep Learning

Deep Learning is essential in diagnosing heart diseases using chest x-rays [34]. The number of articles published just in 2019 alone on this topic (e.g. Google Scholar has 584 results) stresses the importance of the method. As previously noted, scientists have been able to classify 14 different heart conditions using deep neural networks. The accuracy needs to be high so the outcomes can help medical practitioners avoid misdiagnosis. Deep Learning could also help underdeveloped countries access better treatments for their patients.

III. Methodology

Our team developed a model to detect 14 severe pathologies found in chest radiographs within an acceptable range of accuracy. This research could ~~also~~ be used to provide further access to quick, high-quality interpretations of chest radiographs.

A. Data

The dataset used is the ChestX-ray14, which was created by Wang et al. [31] and contains 112,120 x-ray images (frontal-view) from 30,805 unique patients. The dataset has been extracted from the PACS database at the National Institutes of Health Clinical Center, consisting of ~60% of all frontal-view chest x-rays in the entire hospital. Each image is annotated by Wang et al. [31] "with up to 14 different thoracic pathology labels using automatic extraction methods on radiology reports."

B. Deep Learning Approach

The dataset was split randomly in the disease task into validation, tests, and training. We reduced the dimension of the images before they were fed to the network by a windowing technique. They were normalized based on the standard deviation of images, as well as the mean in the ImageNet training set. In the study, the chosen framework was a convolutional neural network (CNN). We also used several architectures of CNN to achieve the best performance. This CNN is a specific type of neural network designed for the handling of image data.

The neural network is trained to simultaneously detect each of the 14 pathologies in frontal-view chest radiographs. These 14 pathologies are not unique to each patient, which means we have patients who are suffering from more than one disease. Convolutional neural networks can scan over an image to acquire knowledge regarding the features from local structures thereby making a prediction on the entire image. By using sample data, we predicted previously concealed frontal-view chest radiographs. The final CNN model uses VGG16 architecture [40].

This architecture consists of 16 convolutional layers, and due to its monotone frame is often used in many image-recognition approaches. The input image dimension is degraded 80x80 RGB to make the computations faster. An image augmentation algorithm is applied to raw images to generate a higher number of training data. This data augmentation rotates the images by 20 degrees and shifts the width and height of images by 0.2 and flips the images horizontally. These images are passed through four stacks of convolutional layers, applying a small filter of 3x3 to them. Each stack contains four convnet architecture.

A normalization layer is applied ~~on~~ to each stack independently, and at the end of each set of layers, a max-pooling layer is applied to down-sample the images, which will help the features contained in the sub-region binned. The activation functions are considered to be ReLU. The number of neurons in each stack is 32, 64, 128, and 256 respectively. A dropout methodology is added to each of the stacks to prevent any possible overfitting. Figure 1 shows a sample of images. This image belongs to a male 58-year-old patient diagnosed with "cardiomegaly."

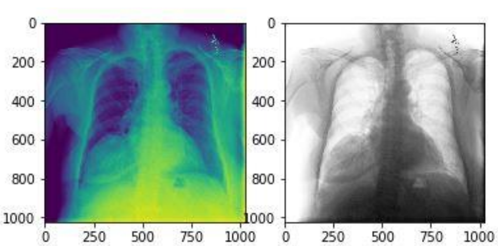

Figure 1- A sample of chest x-rays in the dataset.

Two fully connected layers, in the output layer, are considered; the first one had 1024 layers, and the last one had

14 outputs. Another dropout algorithm is applied to handle the problem of overfitting in the output. The learning rate for training is 10^{-3}.

IV. Result and Discussion

During training, data were split randomly to 80% for training and 20% for testing. Validation was done using 20% of the training data. Figure 2 shows the learning curve of the model after 180 epochs.

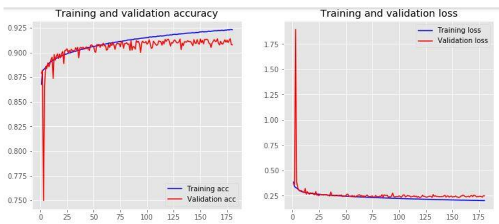

Figure 2- Training and validation learning curves for accuracy and loss.

As shown, the training and validation followed a similar pattern, which showed that using drop out helped prevent overfitting. The accuracy of the test model is 92.6%. Table 1 listing indicates the accuracy of our model in identifying each type of heart disease. Note that some of the patients (samples) are suffering from more than one condition. This analysis shows the sensitivity of 92.9%.

Table 1- The accuracy of different types of heart diseases and the number of samples for each case

Heart disease	Accuracy	Number of samples
Cardiomegaly	94.47	1963
Emphysema	95.37	1119
Effusion	90.13	4062
Hernia	95.01	1300
Infiltration	82.96	5009
Mass	92.00	2386
Nodule	89.06	2749
Atelectasis	87.21	3725
Pneumothorax	92.33	2122
Pleural-Thickening	91.52	2514
Pneumonia	93.03	21.07
Fibrosis	90.24	2622
Edema	93.87	2299
Consolidation	91.07	3151

V. Conclusion

In this work, through a collection of different x-ray images, we demonstrated how to classify chest x-ray data accurately. The results showed us that CNNs could aid in the detection of diseases in chest x-rays. We constructed our model using various techniques, such as data augmentation and windowing. These techniques separate our model from other methods for all ~~the~~ 14 different diseases. This study analyzed all 14 conditions accurately while combining a variety of techniques, which enhanced the model's performance and reduced the rate of errors. The implications of these performances, clinically, are of the utmost importance, especially regarding settings where radiology is scarce. In these places, many x-rays are not given to an expert radiologist to read and analyze. The high value of sensitivity shows how dependable this technology can be for implementation in real-world scenarios.

VI. Acknowledgement

The authors would like to thank the supercomputer center of Harrisburg University for providing access to the GPU enabled cloud system for Deep Learning analysis. Also we would like to thank Christine Altieri for the proofreading.

References

[1] X. Baur, S. Sanyal, and J. L. Abraham, "Mixed-dust pneumoconiosis: Review of diagnostic and classification problems with the presentation of a work-related case," *Science of the Total Environment,* vol. 652, pp. 413-421, 2019.

[2] World Health Organization, Chest Radiography in Tuberculosis Detection, Geneva, Switzerland: World Health Organization, 2012.

[3] R. Chavan, J. Pillai, S. Holkar, P. Salgaonkar, and P. Bhise, "Thoracic Diseases Prediction Algorithm From Chest X-Ray Images Using Machine Learning Techniques," in *Proceedings 2019: Conference on Technologies for Future Cities*, Navi Mumbai, India, 2019.

[4] E. Calli, K. Murphy, E. Sogancioglu and B. Van Ginneken, "FRODO: Free rejection of out-of-distribution samples: application to chest x-ray analysis," *Medical Imaging with Deep Learning,* pp. 1-4, 2019.

[5] R. Liao, J. Rubin, G. Lam, S. Berkowitz, S. Dalal, W. Wells, S. Horng, and P. Golland, "Semi-supervised Learning for Quantification of Pulmonary Edema in Chest X-Ray Images," Massachusettes Institute of Technology, Cambridge, MA, 2019.

[6] B. Eshetu, T. Mekonnen, and M. Berhane, "Lumbo-Costo-Vertebral Syndrome with Congenital Lumbar Hernia: Case Report," *Ethiop J Health Sci. ,* vol. 29, no. 3, pp. 413-416, 2019.

[7] Y. Wang, L. L. Sun and Q. Jin, "Enhanced Diagnosis of Pneumothorax with an Improved Real-time Augmentation for Imbalanced Chest X-rays Data Based on DCNN," *IEEE/ACM Transactions on Computational Biology and Bioinformatics,* pp. 1-1, 2019.

[8] S. Reuter, D. Lindgaard, C. Laursen, B. M. Fischer, P. F. Clementsen, and U. Bodtger, "Computed tomography of the chest in unilateral pleural effusions: outcome of the British Thoracic Society guideline," *J Thorac Dis.,* vol. 11, no. 4, pp. 1336-1346, 2019.

[9] I. Wong and S. Shahzad, "A Bulging Chest Wall Mass after Pneumococcal Empyema. A Case of Pneumothorax Necessitans," *American Journal of Respiratory and Critical Care Medicine,* vol. 199, no. 9, pp. e37-e38, 2019.

[10] F. Kurosaki, T. Takemura, M. Bando, T. Kuroki, T. Numao, H. Moriyama, and K. Hagiwara, "Progressive plasterer's pneumoconiosis complicated by fibrotic interstitial pneumonia: a case report," *BMC Pulmonary Medicine,* vol. 19, no. 6, pp. 1-5, 2019.

[11] S. Rakshit, I. Saha, M. Wlasnowolski, U. Maulik, and D. Plewczynski, "Deep Learning for Detection and Localization of Thoracic Diseases Using Chest X-Ray Imagery," in *International*

Conference on Artificial Intelligence and Soft Computing, Zakopane, Poland, 2019.

[12] S. Fukuhara, Y. Yamaguchi, Y. Uetani, and Y. Akasaka, "Lung Ultrasound in Children with Acute Respiratory Failure: Comparison between Chest X-ray, Chest Computed Tomography, and Lung Ultrasound: A Case Series," *Indian J Crit Care Med,* vol. 23, no. 2, pp. 95-98, 2019.

[13] A. Saito, Y. Hakamata, Y. Yamada, M. Sunohara, M. Tarui, Y. Murano, A. Mitani, K. Tanaka, T. Nagase, and S. Yanagimoto, "Pleural thickening on screening chest X-rays: a single institutional study," *Respiratory Research,* vol. 20, no. 38, pp. 1-7, 2019.

[14] L. J. M. Kroft, L. van der Velden, I. H. Giron, J. J. H. Roelofs, A. de Ross, and J. Geleijns, "Added Value of Ultra–low-dose Computed Tomography, Dose Equivalent to Chest X-Ray Radiography, for Diagnosing Chest Pathology," *Journal of Thoracic Imaging,* vol. 34, no. 3, pp. 179-186, 2019.

[15] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. P. C. Rodrigues, "Identifying pneumonia in chest X-rays: A deep learning approach," *Measurement,* vol. 145, pp. 511-518, 2019.

[16] H. Tanaka, M. Yano, C. Kuwabara, A. Kume, Y. Tamura, M. Murakami, R. Shimizu, H. Saito, and Y. Suzuki, "Massive Hemoptysis Due to the Rupture of Thoracic Aortic Aneurysm Caused by Leukemic Cell Infiltration in a Patient With Chronic Myelomonocytic Leukemia," *J Clin Med Res.,* vol. 11, no. 2, pp. 145-150, 2019.

[17] 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," National Institutes of Health, Bethesda, MD, 2017.

[18] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine,* vol. 25, pp. 44-56, 2019.

[19] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest Pathology Detection Using Deep Learning with Non-Medical Training," IEEE, Tel Aviv University, Tel Aviv, Israel, 2015.

[20] H. T. Huynh and V. N. N. Anh, "A deep learning method for lung segmentation on large size chest X-Ray image," IEEE, Binh Duong New City, Vietnam, 2019.

[21] R. B. Lanfredi, J. D. Schroeder, C. Vachet and T. Tasdizen, "Adversarial regression training for visualizing the progression of chronic obstructive pulmonary disease with chest x-rays," University of Utah, Salt Lake City, UT, 2019.

[22] S. Purkayastha, S. B. Buddi, S. Nuthakki, B. Yadav, and J. W. Gichoya, "Evaluating the Implementation of Deep Learning in LibreHealth Radiology on Chest X-Rays," Indiana University & Oregon University, Indianapolis, IN & Portland, OR, 2019.

[23] M. Harris, A. Qi, L. Jeagal, N. Torabi, D. Menzies, A. Korobitsyn, M. Pai, R. R. Nathavithrana and F. A. Khan, "A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis," *PLoS One,* vol. 14, no. 9, pp. 1-19, 2019.

[24] National Institutes of Health, "Chest X-Ray," National Institutes of Health, 2019. [Online]. Available: https://www.nhlbi.nih.gov/health-topics/chest-x-ray. [Accessed 09/18/2019 September 2019].

[25] Mayo Clinic, "Chest X-Rays," Mayo Clinic, 26 April 2018. [Online]. Available: https://www.mayoclinic.org/tests-procedures/chest-x-rays/about/pac-20393494. [Accessed 18 September 2019].

[26] D. C. Moncada, Z. V. Rueda, A. Macias, T. Suarez, H. Ortega, and L. A. Velez, "Reading and interpretation of chest X-ray in adults with community-acquired pneumonia," Colombian Administrative Department of Science, Technology, and Innovation, Medellin, Colombia, 2011.

[27] World Health Organization, "Global Health Observatory (GHO) Data," World Health Organization, 2018. [Online]. Available: https://www.who.int/gho/mortality_burden_disease/causes_death/top_10/en/. [Accessed 18 September, 2019].

[28] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, R. L. Ball, C. Langlotz, K. Shpanskaya, M. P. Lungren and A. Y. Ng, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," Stanford University, Stanford, CA, 2017.

[29] G. Liu, T. H. Hsu, M. McDermott, W. Boag, W. Weng, P. Szolovits, and M. Ghassemi, "Clinically Accurate Chest X-Ray Report Generation," Massachusetts Institute of Technology, Cambridge, MA, 2019.

[30] M. Hodnett, J. F. Wiley, Y. Liu and P. Maldonado, Deep Learning with R for Beginners, Birmingham: UK, 2019.

[31] S. Yao, Y. Zhao, A. Zhang, S. Hu, H. Shao, C. Zhang, L. Su, and T. Abdelzaher, "Deep Learning for the Internet of Things," *Computer,* no. May 2018, pp. 32-41, 2018.

[32] A. O'Connor, "How Artificial Intelligence Could Transform Medicine," New York Times, 12 March 2019. [Online]. Available: https://www.nytimes.com/2019/03/11/well/live/how-artificial-intelligence-could-transform-medicine.html. [Accessed 18 September 2019].

[33] J. Bresnick, "What Is Deep Learning and How Will It Change Healthcare?," Health IT Analytics, 30 November 2018. [Online]. Available: https://healthitanalytics.com/features/what-is-deep-learning-and-how-will-it-change-healthcare. [Accessed 18 September, 2019].

[34] I. M. Baltruschat, H. Nickisch, M. Grass, T. Knopp and A. Saalbach, "Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification," *Scientific Reports,* vol. 9, pp. 6381-6391, 2019.