Classification of Pneumonia Lungs Infected X-Ray Images using Statistical based Features

Saroj Agrawal[1], Yogesh Kumar Gupta[2]
[1] PhD Scholar, Banasthali Vidyapith, sarojagr708@gmail.com,
[2]Assistant Professor, Department of Computer Science, Banasthali Vidyapith, gyogesh@banasthali.in.

Abstract
Community-acquired pneumonia is classified as mild, moderate or severe by the doctors, taking into account all the risk of complications. A model is proposed for the pneumonia detection divided into two stages from chest X-Ray picture detecting & classifying the occurrence of pneumonia in different stages from images obtained from various sources such as Rajasthan hospital, Kaggle and other diagnostic centres, Jaipur. Various features are used to obtain informative and relevant data from lungs infected X-Ray images for classification of pneumonia severity in this work. It is tried to achieve a remarkable classification performance by extracting features from various infected chest X-Ray images. Classification takes place in various stages - preprocessing, feature-extraction and classification stage. Various filters are used for pre processing and statistical feature extraction algorithm for image classification. After pre-processing, feature extraction is the next step extracting the features for mild and severe stage of pneumonia applying some statistical based methods. Also the approach proposed is tested with different classifiers for classification stage in the research**.** The best results are produced when all the features are combined. The system is assessed using a dataset that was created and contains around 226 lungs infected X-ray images in which 102 mild lungs infected X Ray images and 124 severe lungs infected images.

Keywords: Pneumonia, Mild, Severe, Lungs.

1. Introduction

Pneumonia is a disease of the lung tissue that delivers oxygen to the entire body getting severely inflammated (occurring within the small air sacs called alveoli which make up the lung) cause of the infection by viruses, bacteria, fungi or as a result of damage by chemicals such as alkalis/acids.

The vital task of taking oxygen from the air and delivering it to red blood cells is carried out by the lungs. Lung diseases can be classified as: I airway disease, (ii) lung tissue disease and (iii) lung circulation disease. Lung inflammation leads to lung tissue damage, which limits lung growth. The clotting/inflammation of blood vessels that causes lung circulation disorder inhibit the lung's ability to inhale and exhale.

It is characterized by the presence of a liquid-like substance in the lungs' air sacs (alveoli). It is an infectious illness marked by coldness, shivering, and breathing difficulties. According to a World Health Organization (WHO) estimate, almost 18% of children worldwide

The primary distinction between mild and severe pneumonia is that the patient suffering from mild pneumonia does not have to go through much discomfort whereas severe pneumonia quickly deteriorates sepsis or respiratory distress if not treated on quick basis. In these pneumonia a spurt increase in the respiratory and pulse rate because of the low oxygen saturation in the body while circulating blood and cause of the shallow and short breathing. Oxygen might be needed to be supplied to the patient from an external source. Mild pneumonia can also be treated at home if not allied with high blood pressure, diabetes etc. The symptoms of mild pneumonia are fever, cough with expectoration, weakness and general body pain. On the other hand the symptoms of severe pneumonia are difficulty in breathing on slightest exertion, intense high grade fever with chills, labored breathing, congestion of the chest, hemoptysis (blood in sputum), and cyanosis and chest pain while breathing.

1.1 Objectives of the study

- To detect pneumonia stages through infected X-Ray Images.
- To find appropriate research gap.
- To design the proposed model for classification the X-Ray images
- To statistically analyzed the proposed model.
- To find best accuracy model.

2. Literature Survey

The respiratory discomfort is accompanied by symptoms such as, high fever, breathing problem, chest pain and mucus accumulation [1].After research it is found that Community-acquired pneumonia in elderly patients, they conclude that main causes for pneumonia and what are these treatments [2]. Pneumonia is a disease in which lungs are severely contaminated. Every year around 40000-70000 people are affected. The author conducted research on the kids under the age of 12 in the areas of Malaysi, Tawau etc. The objective of this exploration is to investigate the preponderant factors of lung infected disease pneumonia in Tawau hospice, finally conclude that which part of the lungs is affected by pneumonia. The outcomes will be helping the government in avoiding the risk and creating measures to overcome with it [3]. Pneumonia develops speedily posing a threat to human health and survival. Presently, binary classification algorithms are used for computer aided diagnosis (CAD) of pneumonia which does not provide medical practitioners with location information. The current research puts forth a highly effective end to end algorithm for identifying pneumonia infection grounded on a CNN - Pneumonia Yolo (PYolo). It is a refined variation of the Yolov3 algorithm. A total of 6000 chest X-Ray’s were considered for the present study. The Radiological Society of North America (RSNA) delivered these X-Rays. An anchor box is created using dilated convolution and K-means to enhance localization accuracy. 46.84 mean averages precision (MAP) was obtained by the algorithm. Therefore, pneumonia could be determined using an algorithm that can provide location information to doctors of bruises. Thus, the research examined the issue associated with small textures of pneumonia bruises in X-ray. It also proposed an improved pneumonia detection algorithm based on Yolov3.

A total of 6000 chest X-Ray’s were considered for the present study. These X-rays’s were rendered by the Radiological Society of North America [4]. 5,856 X-rays of front back chests were selected from pediatric clients ranging between the ages of 1 to 5 years. These scans were subjected to a CNN model prepared from the start to classify and detect pneumonia. The algorithm aided in enhancing the CNN models validation and classification. Through data augmentation, learning rate variation, and reinforcing is achieved significant validation accuracy which was used to help small data sets into deep neural network architectures. Subsequently the model can be expanded to identify and classify X-rays comprising of lung cancer and pneumonia. It can also be used to distinguish lung cancer from pneumonia in the X-rays. [5]

One of the researches aimed at discerning the etiology, by understanding the relationship among factors related with bacterial pneumonia and the relationship among co-infections and disease brutality in hospitalized kids due to serious pneumonia. Children hospitalized due to severe pneumonia between 1 to 5 years of age were considered for the purpose of this study. Bacteria alone were concluded to be the significant cause of severe pneumonia in developing countries. Bacterial etiology was notably linked with male gender and existence of crepitations. Higher CRP was linked to co-infection but not to other parameters of serious clinical illness [6].

The current research stresses upon pneumothorax, pneumoconiosis and emphysema. Chest X-rays were used as the dataset. Balance contrast enhancement techniques (BCET) were adopted for data analysis. The technology helps extract lungs features as the lung regions, lungs sizes and shape irregularities. Two CXR datasets were adopted for the experimental research: The first data set contained 138 frontal chests X-ray in Montgomery County (MC) set. The other set contained 662 frontal chests X-rays Shenzhen Hospital (SH) set. An automated system is used for lung boundary detection and categorization of CXR images using PNN binary classifier with 94.98% to 95.77%, [7]. The study proposed an efficacious model DCNN (deep convolution neural networks) for big data with deep layers. The model helps classify if an individual is affected with pneumonia or not. The model achieved the predicted accuracy of 84% by extracting the features of X-rays of superior quality. Different metrics were adopted to compare the results of this model with other ordinary classifiers like SVM, random forest etc [8].

The formation of images is affected by machine-/deep-learning. Van Slounet investigated the advantages of data-driven deep learning in different facets of ultrasound imaging. From views located at the intersection of raw signal acquisition and image generation, to learning compression codes for color Doppler acquisition, to learning clutter suppression strategies using positional information.

This offers an impressive vision of ultrasounds future formulated upon portable and intelligent imagining enables intelligent wireless probes for many unique applications. The researcher places of interest that high volume of data processing limits sonographic instruments for implementing the accessible methods. [9]

The present research used Contrast limited adaptive histogram equalization (CLAHE) to increase the visibility level of foggy pictures & videos in real time. Normal histogram equalization varies from adaptive histogram equalization (AHE) as AHE adopts various methods with each being in accordance with different parts of an image. ‘Distribution’ Constraints are used to describe the shape of the histogram for CLAHE, which gives better quality results compared to AHE.

In this research a variety of unique and varied methods for the concept of using computer-aided diagnosis in healthcare to find lung issues. The authors analyzed some of the newest methods for identifying various normal and pathological tissues as well as lung nodules. They concentrated on the creation of CAD systems for the processing of medical images. In their subsequent research, they focused on CT image reconstruction techniques and imaging protocols for the potential prediction of lung tumors. [10]

In this paper focus on the general and comparative study of parallel and non-parallel computing feature extraction methods used to extract important features from medical images without altering their dependable performance metrics. The use of intellectual analytics techniques in the healthcare area has demonstrated the difficulty of this endeavor. They examined different scalable analytical algorithms and models to address the problem with large-scale image processing in order to gain high performance utilizations for effective diagnosis. [11]

A technique for extracting wavelet features from medical images that come from different imaging modalities has been proposed. For the deconstruction of medical pictures, they employed Discrete Wavelet Transform. After calculating the extracted features' mean and standard deviation, the researchers entered the findings into the K-Nearest Classifier for classification. As a consequence, they successfully discovered that DWT is more accurate than DCT at 99.96 percent. They will develop another cutting-edge strategy to lessen the dimensionality and temporal complexity as future work.[12]

The author explains in detail the steps involved in picture processing. It classifies medical images using the association rule mining a priory method. The primary goals of this study are to improve diagnostic precision and shorten decision-making times. [13]

This paper emphasizes the importance of strategies for classifying medical image data. They discuss supervised, unsupervised, and fundamental data mining techniques, wavelet and Fourier transforms for texture classification, and Bayesian decision theory for neural networks. [14]

In this paper to achieve computational accuracy and speed, we describe a proposed framework consisting of six phases, including data acquisition, data preprocessing, data splitting, Soft set classification, data analysis, and performance development. It explains why classification is necessary and how related technologies and methods such as texture classification, k-nearest neighbors, neural networks, and SVM can be used to classify medical images. [15]

2.1 Research Gap

The requirement for an accurate and effective method for diagnosing pneumonia stages (mild and severe) from infected chest X-ray pictures is the research gap that this study fills. This is a serious problem because pneumonia is a dangerous lung condition that puts people's lives at risk. Although numerous deep learning algorithms have been utilised in the past to detect pneumonia, a trustworthy and accurate method that can aid in the early detection and treatment of the condition is still required. Random forest Analysis (RFA), the suggested algorithm in this work, achieves a high level of accuracy, which may close the research gap in creating a reliable and effective approach for detecting severity of pneumonia from infected chest X-ray pictures.

3. Materials and Methods

3.1 Dataset

A. Dataset Description

Dataset utilized contains 226 lungs X-beam pictures with classifications: Mild and severe infected. Those pictures are taken from Rajasthan Hospital (Jaipur) and Kaggel.com. All are physically examined to vacate inferior quality pictures, as characterized by doctors to affirm miss-characterization and ground truth. 226 lungs infected X-Ray images in which 102 mild lungs infected X-Ray images and 124 severe lungs infected images.

- Training Dataset – This Dataset comprises of 149 images along with 82 mild infected and 67 severe X-Ray images.
- Test Dataset – The validation dataset comprises of 77 images with 42 mild infected and 35 severe X-beam images.
- Labels – Mild : 0, Severe : 1

B. Mild, Severe Pneumonia

A wide range of highlights are utilised in a x-beam picture by the radiologists to identify the Pneumonia. In later phases of infection GGO examples are different when they are contrasted with the beginning phases. In radiography to speak to aspiratory combination murkiness is used which helps in comparing to the white territory, as exposed in the figure 1, mild infected patient X-beam picture and Figure 2 severe patient X-beam picture.

3.2 Approaches to Our Model

Description of the applied methodology is discussed in this section. The proposed model consists of the pre-processing, feature extraction and classification stage.

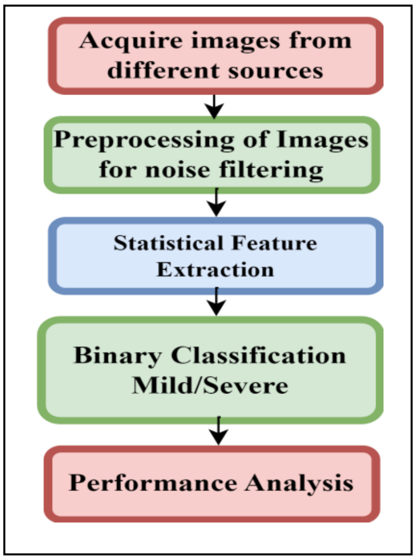

Figure 1 Proposed Model

3.2.1 The Pre-Processing Stage

Homomorphic and the Clahe are the two commonly used filtering techniques employed for digital image enhancement. These are also used in various imaging applications including robotic, medical and biometric vision.

Contrast limited adaqptive histogram equalization (CLAHE) improves the visibility of foggy videos and images. This is why it is used in this document to improve the quality of X-ray images. After that Homomorphic filtering is used which works in frequency domain to reduce the low frequency components and It works using the illumination reflection model. The image is characterized by two main components of this model. The first is the illumination of the incident light source, denoted as i(x,y). The second is the reflection component of the scene r(x,y).

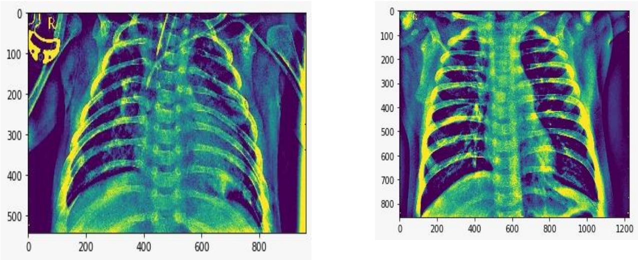

Figure 2 (A) Mild Filter Image (B) Severe Filter Image

3.2.2 Feature Extraction-

For numerically quantifying and characterizing the contents of an image an abstraction of an image is used. Usually real, integer, or digital valued. List of numbers used to represent an image is called a feature vector. Once pre-processing is done feature extraction is the next step. The classification of images is done according to mild and severe infectious images being the first stage of image is preprocessing. In this research, First-order histogram, GLCM is used to extract features, through First-order histogram there are the following five features are derived: variance, mean, kurtosis, skewness, and standard deviation of lungs infected X-Ray image. Through GLCM, these eight features are extracted such as dissimilarity, ASM, IDM, homogeneity, correlation, contrast, entropy, and energy. A total of 13 features were derived using all statistical based features.

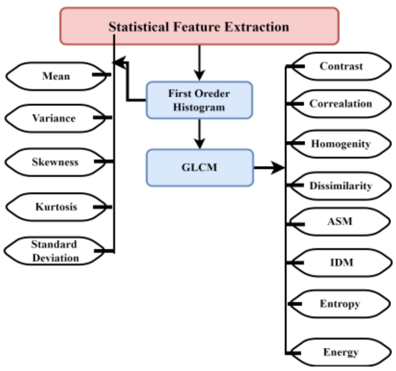

Figure 3 Feature Extraction Model

A. The first-order histogram

The first-order histogram is used in order to extract first-order statistical values. This is also used to calculate different statistical measures. A picture is represented in numerical form as a matrix for two space variable x and y, through this method below mentioned six features are extracted: skewness, mean, kurtosis, variance, and standard deviation of grey pictures.

Values are selected randomly from the vector database as shown in Table (1). The mean value of mild is lesser than infected as compared to white pixel because of white cloud in severe infected X-Ray images. The shape of probability distribution is represented by kurtosis. Value greater than zero represents peak is low around mean value and is also in shape of a rounded rectangle. On the other hand if the value is lower than zero or negative, it means that peak is high around mean value.

B. GLCM

One technique for obtaining second-order statistical values (i.e. texture feature) is the GLCM approach. It basically consists of a matrix with as many (i) rows and (j) columns as there are gray levels in the image. Eight features are retrieved in this study: energy, homogeneity, correlation, IDM, ASM, entropy, contrast, and dissimilarity. The degree of contrast among a pixel and its neighbour over the entire x-ray image is measured by contrast, which also determines the dissimilarity among the greatest and smallest values of a surrounding pixels's group.

Correlation is a measurement of the joint probability of grey scale pixel pairs. Using the Homogeneity feature of GLCM, the proximity of the distribution of grey scale pixels is calculated. In image processing, discrete entropy represents the number of bits. Random sample values of GLCM features are represented in table (1) below.

Table 1 Statistical Features Values

	Path	Mean	Std Dev	Variance	Skewness	Kurtosis	Contrast	Dissimilarity	Homogeneity	Energy	Correlation	ASM	Entropy	IDM	Category	Label
1	newMMS/mild/mild5.jpeg	82.13	69.06	4769.35	0.55	-0.51	1.33	0.62	0.74	0.26	0.96	0.06	3.98	0.66	mild	0
2	newMMS/mild/mild96.jpeg	88.16	86.08	7410.03	0.49	-1.11	1.92	0.76	0.71	0.35	0.97	0.12	3.73	0.66	mild	0
3	newMMS/mild/mild58.jpeg	86.42	77.73	6042.55	0.59	-0.79	1.99	0.78	0.71	0.26	0.96	0.06	4.10	0.63	mild	0
4	newMMS/mild/mild74.jpeg	88.41	75.56	5709.84	0.45	-0.85	2.23	0.88	0.67	0.27	0.95	0.07	4.10	0.60	mild	0
5	newMMS/severe/severe104.jpeg	103.96	80.51	6482.11	0.18	-1.12	2.56	0.97	0.64	0.23	0.95	0.05	4.29	0.57	severe	1
6	newMMS/severe/severe90.jpeg	114.18	90.57	8203.20	-0.04	-1.44	1.51	0.59	0.77	0.31	0.98	0.09	3.70	0.71	severe	1
7	newMMS/severe/severe69.jpeg	118.42	85.27	7271.22	-0.06	-1.28	1.66	0.66	0.74	0.24	0.97	0.05	4.03	0.67	severe	1
8	newMMS/severe/severe86.jpeg	102.21	87.41	7639.89	0.24	-1.30	1.33	0.58	0.76	0.31	0.98	0.09	3.78	0.69	severe	1

3.2.3 Classification:

The proposed method in this paper classifies lungs infected X-Ray images into two classifications: the Using features with a statistical, mild and severe infected pneumonia. If the image is determined to be anomalous, it will be reviewed and tested further. Eight classifiers, such as Logistic Regression, Quadratic Discriminant Analysis, Decision Tree, Support Vector (SV) Machine, Linear Discriminant Analysis, Random Forest, K-Nearest Neighbors, and Naive Bayes, are used to test the suggested methodology, as shown in the given figure.

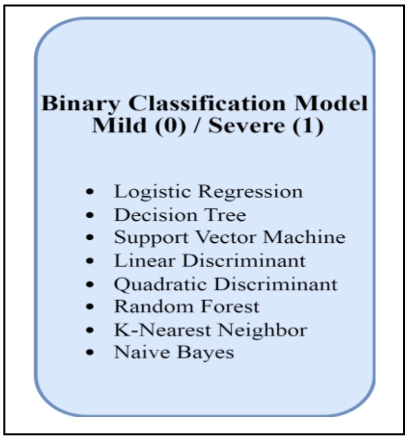

Figure 4 Classification Model

- Logistic Regression: predictive modeling concept is used as regression in this but it is used to classify samples therefore, it falls under the classification algorithm.
- Decision Tree: As a trained learning process, it is employed for the purpose of regression and classification task, but it is primarily favored, in order to address categorization issues. It functions like a tree-structured classifier.
- Support Vector Machine: Using maximum margin, SVM builds the hyper-plane to categorize the data into the relevant groups. Kernel functions transform decreased-dimension data into increased-dimension data for better performance.
- Linear Discriminant Analysis: A dimensionality lessening method is commonly used for supervised categorization issues in LD analysis, It is also known as Discriminant function analysis(DFA), used to project the features in an increased measurement space into a decreased measurement space, and also it is used to separate two or more classes. Unlike linear

discriminant analysis, quadratic discriminant analysis assumes that each class has the same mean and covariance. We must calculate it independently because of this.

- Random Forest: This method is applied for categorization and regression issues in machine learning, which is also used for integrating different classifiers to handle a complicated task and boost the performance. It has many decision trees on dissimilar subsets of the given dataset, and it takes the average to raise the dataset's predictive exactness.
- K-Nearest Neighbor: One of the supervised knowledgable method, is K-Nearest Neighbor. Based on similarity, this algorithm saves and categorizes the entire accessible data and any new data. Using the KNN method, newly emerging data may be promptly categorized into a suitable category. It is a non-parametric algorithm that is primarily used for categorized problems, while it can also be utilized for regression problems. K-NN is non-parametric; therefore, no assumptions are made about the underlying data. Naive Bayes is a trained knowledgeable approach for categorizing issues is depends on the Bayes theorem. Being a probabilistic categorizer, it predicts assumptions based on the probability of an entity.

3.2.4 Performance Analysis

This part of the paper explains parameter evaluation metrics for performance of various machine learning algorithms. The results are explained in detail and displayed graphically.

Performance Assessment: Using several metrics, the effectiveness of the suggested scaffold is assessed. The diverse evaluation indicators were derived using data from Table 2. Various measures showed that Random Forest outperformed other models.

As shown in table below. Although Random Forest produces outstanding results when compared to other measures, we cannot entirely rely on accuracy due to the unbalanced nature of the data. Quadratic Discriminant Analysis & K-Nearest Neighbors are equally effective and formidable rivals. Random forest retains its stand when FP rate and TP rate are compared. Random Forest produces effective results on the data with an accuracy of 83%.

Table 2 Performance Evaluation

Model	Accuracy	Sensitivity	Specificity
Logistic Regression	64.94%	54.29%	73.81%
Decision Tree	64.94%	54.29%	73.81%
Support Vector Machine	54.55%	0.00%	100.00%
Linear Discriminant Analysis	64.94%	54.29%	73.81%
Quadratic Discriminant Analysis	70.13%	34.29%	100.00%
Random Forest	83.12%	82.86%	83.33%
K-Nearest Neighbors	67.53%	51.43%	80.95%
Bayes	63.64%	22.86%	97.62%

Table 3 Confusion Matrices

	True Reference	
Predicted Condition	Condition Positive	Condition Positive
Pneumonia Positive	TP(W) =29	FP(Y) = 6
Pneumonia Negative	FN(Z) = 7	TN(X) =35

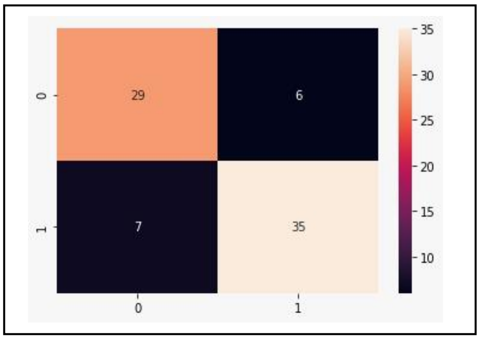

Figure 5 Confusion Matrix

4. Result and discussion

Different metric results obtained after experimenting and testing different models. It is easy to see that Random Forest Analysis has the best precision and other parameters. The conclusion is depicted in the table 2 above. In Random Forest Analysis high true positive rate and low false positive rate. Patients with pneumonia are more likely to be recognized. Low false positive rate, which shows that the probability of cases of pneumonia patients classified. When comparing the models one of the important metric to be considered is TP rate.

5. Conclusion

Pneumonia in people can become life-threatening if improperly diagnosed. The World Health Organization believes that the third-largest population in India does not have access to diagnostic radiology. In this work, the formation of efficient machine learning relies on predictive models. Chest X-Ray image reports are used to predict pneumonia in patients. Machine learning is efficient for image data processing making this task more effective. It is easy to observe that the accuracy and other parameters of the Random Forest Analysis are the best among all other models. The proposed framework is found to be more efficient compared other methods of machine learning with an average accuracy of 83.11%, resulting as a better option than all other classified method. In healthcare science, radio- logistic medical images are used. In this study, X-Ray pictures were employed.

6. Future Scope

In future may employ CT, USG, and MRI to retrieve the hidden data that is stored as pixels in these pictures. Medical image analysis contains so much important data that it has become a crucial component. Leaving the information undetected and hidden can cause dilemma in near future. In medical science, radiologists may have to make tough choices as the information is not feasibly extractable and the images have very less number of pixels.

More research will also be conducted by assessing additional models and pre-processing methods that could enhance the detection of input artifacts using deep learning approaches. Ultimately, we will create general deep learning tools that can help clinicians make diagnoses of both known and previously unknown diseases.

References

[1] Zafar, M. Z. (2016). A case study: pneumonia. *Occup Med Health Aff*, *4*(242), 2.

[2] Sufahani, S. F., Razali, S. N., Mormin, M. F., & Khamis, A. (2012). An analysis of the prevalence of pneumonia for children under 12 year old in Tawau general hospital, Malaysia. *arXiv preprint arXiv:1205.2109*.

[3] Yao, S., Chen, Y., Tian, X., Jiang, R., & Ma, S. (2020). An improved algorithm for detecting pneumonia based on YOLOv3. *Applied Sciences*, *10*(5), 1818.

[4] Stephen, O., Sain, M., Maduh, U. J., & Jeong, D. U. (2019). An efficient deep learning approach to pneumonia classification in healthcare. *Journal of healthcare engineering*, *2019*.

[5] Nathan, A. M., Teh, C. S. J., Jabar, K. A., Teoh, B. T., Tangaperumal, A., Westerhout,

C., ... & de Bruyne, J. A. (2020). Bacterial pneumonia and its associated factors in children from a developing country: A prospective cohort study. *PloS one*, *15*(2), e0228056.

[6] Zotin, A., Hamad, Y., Simonov, K., & Kurako, M. (2019). Lung boundary detection for chest X-ray images classification based on GLCM and probabilistic neural networks. *Procedia Computer Science*, *159*, 1439-1448.

[7] Jakhar, K., & Hooda, N. (2018, December). Big data deep learning framework using keras: A case study of pneumonia prediction. In *2018 4th International Conference on computing communication and automation (ICCCA)* (pp. 1-5), IEEE.

[8] Duncan, J. S., Insana, M. F., & Ayache, N. (2019). Biomedical imaging and analysis in the age of big data and deep learning [scanning the issue]. *Proceedings of the IEEE*, *108*(1), 3-10.

[9] Yadav, G., Maheshwari, S., & Agarwal, A. (2014, September). Contrast limited adaptive histogram equalization based enhancement for real time video system. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2392-2397). IEEE.

[10] Lowitz, G.E.: ‘Can a local histogram really map texture information?, *Pattern Recognit.*, 1983, **16**, (2), pp. 141–147.

[11] Haralick, R.M., Shanmugam, K.: ‘Textural features for image classification’, *IEEE Trans. Syst. Man Cybern.*, 1973, **3**, (6), pp. 610–621.

[12] Reed, T.R., Dubuf, J.M.H.: ‘A review of recent texture segmentation and feature extraction techniques’, *CVGIP, Image Underst.*, 1993, **57**, (3), pp. 359–372 . https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/iet-cvi.2018.5349

[13] National Center for Health Statistics (NCHS); Centers for Disease Control and Prevention (CDC) FastStats: Pneumonia. Last Updated February 2017. Available online: http://www.cdc.gov/nchs/fastats/pneumonia.htm (accessed on 21 November 2019).Heron, M. Deaths: Leading causes for 2010. Natl. Vital. Stat. Rep. **2013**, 62, 1–96.

[14] World Health Organization. The Top 10 Causes of Death;World Health Organization: Geneva, Switzerland, 2017; Available online: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed on 10 November 2019).

[15] Nathan, A. M., Teh, C. S. J., Jabar, K. A., Teoh, B. T., Tangaperumal, A., Westerhout, C., ... & de Bruyne, J. A. (2020). Bacterial pneumonia and its associated factors in children from a developing country: A prospective cohort study. *PloS one*, *15*(2), e0228056.

[16] Ayan, E., & Ünver, H. M. (2019, April). Diagnosis of pneumonia from chest X-ray images using deep learning. In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)* (pp. 1-5). Ieee..

[17] Mithlesh Arya, Namita Mittal, Girdhari Singh."Texture-based feature extraction of smear

[18] images for the detection of cervical cancer", *IET Computer Vision*, 2018.

[19] Yogesh Kumar Gupta, Saroj Agrawal, Aman Mittal. "Classification of lung infected Corona

[20] Virus X Ray Images using Deep Learning CNN Model", *IOP Conference Series: Materials Science and Engineering*, 2021.

[21] Haralick, R.M., Shanmugam, K.: ‘Textural features for image classification’, IEEE Trans. Syst. Man Cybern., 1973, 3, (6), pp. 610–621.

[22] Reed, T.R., Dubuf, J.M.H.: ‘A review of recent texture segmentation and feature extraction techniques’, CVGIP, Image Underst., 1993, 57, (3), pp. 359–372. https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/iet-cvi.2018.5349.