ISSN 2347 - 3983
Volume 8. No. 1.1, 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter1281.12020.pdf
https://doi.org/10.30534/ijeter/2020/1281.12020

COVID-19 Detection for Chest X-Ray Images using Local Binary Pattern

Nurbaity Sabri[1], Raseeda Hamzah[2], Shafaf Ibrahim[3], Khyrina Airin Fariza Abu Samah[4]
[1, 3,4]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Melaka, Kampus Jasin, 77300 Merlimau, Melaka, Malaysia
[2]Faculty of Computer and Mathematical Sciences, UiTM Shah Alam, 40450 Shah Alam, Selangor, Malaysia
nurbaity_sabri@uitm.edu.my, raseeda@uitm.edu.my, shafaf2429@uitm.edu.my, khyrina783@uitm.edu.my

ABSTRACT

The increase in patients with COVID-19 is overwhelming in healthcare systems around the world. Due to the large number of people affected by this pandemic, the medical and healthcare departments are facing a delay in the detection of COVID-19. Besides, it is not an easy task to clarify the images from the radiograph on what types of infection between bacteria pneumonia and COVID-19. The automatic feature analysis can help physicians more precisely in the treatment and diagnosis of diseases. In this research, Local Binary Pattern (LBP) texture features algorithm has been proposed to automate the current manual approach. This process starts by extracting the intensity grayscale texture from the normal, bacteria pneumonia and COVID-19 chest x-ray images. To prove the accuracy of LBP, a simple classifier k-Nearest Neighbour (k-NN) has been implement to classify the chest x-ray images into normal, bacterial and pneumonia class. The 10-fold cross validation has been used to validate the chest x-ray images. From the evaluation, 96% accuracy can be achieved by using LBP as a feature extraction feature. It shows that LBP is a powerful texture features to detect COVID-19 from the x-ray images. More samples will be collected in the future and neural network approach is suggested as a classifier in the future due to its ability to imitate human respond.

Key words : COVID-19, Local Binary Pattern, Chest X-Ray Images, K-Nearest Neighbour, Pneumonia

1. INTRODUCTION

The sudden increase in the number of patients with COVID-19, a new respiratory virus, has placed unprecedented burdens on healthcare systems around the world. Health care systems have already been overwhelmed in many countries. Severe Acute Respiratory Syndrome Coronavirus 2 or SARS-Cov-23 is a virus that causes the most lethal virus that is COVID-19. A significant family of such viruses is responsible for numerous diseases such as flu, Middle East Respiratory Syndrome and Extreme Acute Respiratory Syndrome [1].

Dyspnea, fever and cough are signs of infection and, in more severe cases, can lead to pneumonia, septic shock, SARS, organ failure and death. Most of COVID cases commonly used by clinical as it is cheaper and easier to be used for diagnosis and medical imaging economically [2]. However, it is not an easy task to clarify the images from the radiograph on what types of infection it is [3]. Both have similar symptoms that often blur into each other, but there have difference appearances. Bacterial pneumonia affects the air sacs in the lungs and to identify any abnormality exists, the radiologist will look onto any existence of abnormal density in the lungs.

The existence of lobar consolidations refers as a bacterial infection where it describes the consolidation in one of the lobes of the lung, and infers as an alveolar spread of disease. Besides, the bacterial pneumonia can also detect when both diaphragm and lungs have different densities, where the lungs are known to have air density, while soft tissue density is the diaphragm. If by chance, the line between both densities not seen on x-ray, it means it is possible to have bacterial pneumonia or Right Lower Lobe Pneumonia [4]. The x-rays of the COVID patient are more different from those of the pneumonia patient, where the chest x-ray called 'ground glass' means that there is a blurred lung opacity that is not dense enough to obscure any underlying pulmonary vessels or bronchial walls. Consolidation, on the other hand, applies to thick opacities that surround vessels and bronchial walls [5].

Digital image processing has been widely used in the medical field. However, most approaches also require manual processing. Automatic features analysis can assist physicians in the treatment and diagnosis of diseases with higher accuracy, speed up the diagnosis process, and increase performance. Therefore, Chest x-ray scans can be helpful in suggesting the diagnosis for a patient and also, for monitoring patient responses. To help and improve current practices in analyze the chest x-ray, an automate detection of chest x-ray images is important for detection of COVID-19. This work discusses the classification of chest x-ray images based on the Local Binary Pattern (LBP) features and the classification is presented.

The extraction of features is a critical step, as it is a significant factor in the efficiency of the model classifier and the accuracy of the image classification counts on the numerical properties of the image function [6]. There are three types of features that can be derived that are color, form and texture. Image can be represented by a boundary representation and a region representation. Boundary representation shows the intrinsic characteristics of an object such as shapes. Meanwhile, the representation of the region focuses on internal properties such as color and texture [7].

In this research, region representation method is used to represent the different type's chest x-ray images. Local Binary Patterns (LBP) technique for features extraction is proposed to extract texture by applying thresholding the neighbourhood of each pixel to label the image's pixels as mentioned [8]. This method delivers a good performance based on combination of structural statistical methods for texture analysis [9]. To clarify, structural method represents texture by patterns that has regular appearance and located at the surface consistently. On the contrary, statistical method represents texture by non-absolute properties and random distributed elements. In other words, statistical method represents texture that has irregular appearance and located at the surface inconsistently.

Others texture features has also been implements in medical imaging such as GLCM and Gabor filter. Gray level co-occurrence matrix (GLCM) texture analysis is a captive of image texture proposed [10]. GLCM is a method that can operate co-occurrence greyscale matrices through a statistical method to extract any attribute from the image [11]. However, this algorithm is lacks of relationship between texture patterns and sensitivity to image noise [12]. Gabor filter is used to improve the lung images [13]. However, algorithm performance of Gabor filter is not mutually orthogonal between the texture characteristics and the high cost of the Gabor filter [12].

For classification purpose, the simple machine learning algorithm which is k-Nearest Neighbour (kNN) will be used. This algorithm is based on the shortest distance from the test sample to the training sample. KNN is a supervised machine learning algorithm that able to group an unknown data into the class where majority of its k nearest neighbours belong [14]. This technique is commonly used for pattern recognition where Euclidean distance is usually used for calculation between two points in order to determine its nearest neighbours [15]. Benefits of KNN are this technique effective with large training data and strong against noisy data [16].

2. METHODOLOGY

In this research, a model for chest x-ray image detection has been proposed and tested using LBP features algorithm as shown in Figure 1. It starts with input chest x-ray image, feature extraction, and classification. MATLAB software has been used to extract features from the 101 chest images using LBP algorithm. For testing purpose, Weka software has been used to verify the accuracy of the proposed features.

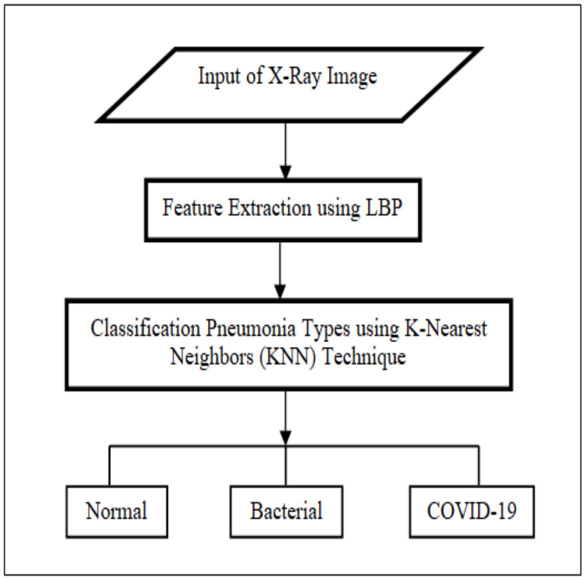

Figure 1: The framework of overall process for COVID-19 detection

2.1 Input X-ray image

Image acquisition under the input image is the early stage where the data are collected. The images are collected from the Kaggle dataset consist of three types of chest x-ray images. Total of 101 of chest x-ray images collected where 33 x-ray images of COVID, 28 of normal x-ray images and 40 of pneumonia bacterial x-ray images. Sample images for normal, bacterial and COVID x-ray images are as Figure 2 below.

(a)

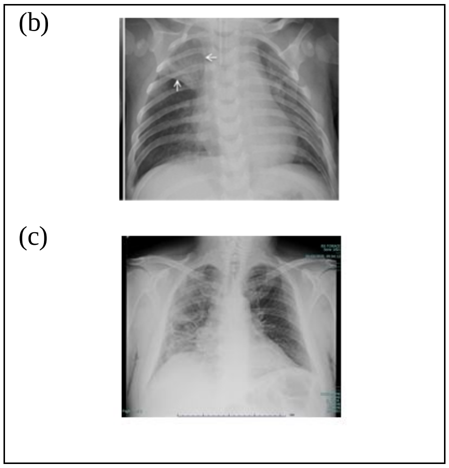

Figure 2: The x-ray images of (a) normal, (b) bacteria and (c) COVID-19

2.2 Feature Extraction using LBP

Extraction of features in the form of numerical values is conducted to represent an image for classification purpose. Texture is chest x-ray images due to the presence of repetitive pattern that depict the chest images have been generated with the absorption of different spectrums based on the tissue density. The suggested steps below are used to extract the desired features [17]:

1. Firstly, a clear image is divided to form cells such that for each cell, image is divided into 3 x 3 pixel.
2. Secondly, pixels comparison is made according to its neighbour such as top-left and top-right. Later, form a circle that contains the comparison of the pixels.
3. Value 1 is replaced if value of center pixels than the value of neighbourhood pixels apart from that value 0 is replaced in the cells.
4. Next, histogram is calculated and normalized.
5. Lastly, histogram of each cells is collected.

Equation (1) shows the LBP algorithm:

$$LBP_{R,N}\,(p,q) = \sum_{i=0}^{N-1} s(n_i - n_c)2^i, s(p) = \begin{cases} 1, p > 0 \\ 0, \ otherwise \end{cases} \qquad (1)$$

Differences in neighbourhood are portrayed as N-bit binary number, which leads to distinct values of 2^n, for the binary pattern. Based on the (1), n_c refers to grey level of local neighbourhood center pixel and refers to equally spaced pixel of *Radius, R* and *S* represent sample images.

2.3 Classification using K-Nearest Neighbour (k-NN)

Generally, k-NN is type of lazy learning as it is learning method in testing phase instead of in training phase. Classification is achieved by identifying the closest neighbors and by using those neighbors to ensure the class of inquiry [18]. Conceptually, the minimum distance of the certain dataset decided which class the given point is belongs is based on the calculation using Euclidean Distance. This algorithm are used to calculate the minimum value between the test data and training the x ray images data of normal, COVID and bacteria.

The technique raises comparative estimations as per class of test data and classifies the data dependent on the most elevated number of votes from the k-neighbors [19]. The data is set to the class most common among its k-nearest neighbors. K is a small positive integer, if k = 1, the sample is set to the class of its nearest neighbor. Neighbors are excluded from a group of items for which the correct classification is established. It can be thought of as a training set for the algorithm, although no specific training phase is needed. Equation (2) below shows the Euclidean distance where u = (x1, y1) and v = (x2, y2) are two points.

$$(u, v) = \sqrt{(x_1 - x_2)^2} + (y_1 - y_2)^2 \qquad (2)$$

3. RESULT AND ANALYSIS

A total of 101 data consist of 33 COVID-19, 28 normal and 40 bacteria of chest X-Ray images are tested. The extracted features are tested using Weka software where it able to analyze the accuracy of k-NN classifier using 10-folds cross-validation. The result represents with true positive (TP) and false positive (FP) for all tested images. In this research, 3776 attributes from x-ray images have been used for classification purpose. Matrices used to measure the efficiency of the classifier precision and recall. Precision represent the percentage of x-ray images that are classified as true. Meanwhile, recall is the percentage of relevant x-ray images that labeled as "true" by the classifier Result of classification shown in Table 1.

Table 1: Table result of k-NN classification using LBP features

Class	TP	FP	Precision	Recall
COVID	**0.939**	**0.029**	**0.939**	**0.939**
Normal	**1.000**	**0.014**	**0.966**	**1.000**
Bacteria	**0.950**	**0.016**	**0.974**	**0.950**
Average	**0.960**	**0.020**	**0.960**	**0.960**

Based on Table 1, it is clearly observed that LBP able to produce a good classification accuracy with average of 0.960. 96%. The analysis also reveals that the maximum precision and recall are obtained for the LBP algorithm, with both values at 0.96.

4. CONCLUSION

This paper presents a research on detection of chest x-ray images using Local Binary Pattern (LBP) algorithm. The objective of this research is to classify an x-ray chest images into three classes which is COVID-19, Normal and bacteria. The performance of proposed algorithm has been reported. It proves that LBP a powerful feature where it able to classify the chest X-ray images for normal, bacterial and COVID with accuracy of 96%. In this research, k-NN classifier has been used to measure the accuracy of the propose features. This research is ongoing work involving only a limited sample of images below 50 data per class. Future work will increase the sample size and build a much larger training database. Future research will also involve the comparison of clustering, classifying and neural network algorithms for chest x-ray images.

ACKNOWLEDGMENT

The authors would like to thank Research Management Centre under the grant of 600-IRMI/REI 5/3 (010/2018) of University of Technology MARA.

REFERENCES

1. A. Mangal, S. Kalia, H. Rajgopal, K. Rangarajan, V. Namboodiri, S. Banerjee and C. Arora. **CovidAID: COVID-19 detection using chest x-ray**, *arXiv preprint arXiv:2004.09803*, 2020.
2. C. Qin, D. Yao, Y. Shi and Z. Song. **Computer-aided detection in chest radiography based on artificial intelligence: a survey**, *Biomedical Engineering Online*, vol. 17, pp. 113, 2018.
3. P. Rajpurkar, J. Irvin, R.L. Ball, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C.P. Langlotz, and B.N. Patel. **Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists**, *PLoS medicine*, vol. 15, 2018.
4. S. Rajaraman, S. Candemir, I. Kim, G. Thoma, and S. Antani. **Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatrics chest radiographs**, *Applied Sciences*, vol. 8, 2018.
5. S. H. Yoon, K. H. Lee, J. Y. Kim, Y. K. Lee, H. Ko, K. H. Kim, and Y. H. Kim. **Chest radiographic and CT findings of the 2019 novel coronavirus disease (COVID-19): analysis of nine patients treated in Korea**, *Korean Journal of Radiology*, vol. 21, pp. 494-500, 2020.
6. S. A. Medjahed. **A comparative study of feature extraction methods in images classification. International journal of image**, *Graphics and Signal Processing*, vol. 7, 2015.
7. K. Prasad, S. Rani, and S. Basha. **An effective texture features based mammogram mass detection system**, *International Journal of Engineering &Technology*, vol. 7, pp. 601-605, 2018.
8. Y. H. Chan, Y. Z. Zeng, H. C. Wu, M. C. Wu, and H. M. Sun. **Effective pneumothorax detection for chest X-ray images using local binary pattern and support vector machine**, *Journal of healthcare engineering*, 2018.
9. A. Humeau-Heurtier. **Texture feature extraction methods: A survey**, *IEEE Access*, vol. 7, pp. 8975-9000, 2019.
10. R. M. Haralick, K. Shanmugam, and I. H. Dinstein. **Textural features for image classification**, *IEEE Transactions on systems, man, and cybernetics*, vol. 6, pp. 610-621, 1973.
11. V. Patel, S. Shah, H. Trivedi, and U. Naik. **An analysis of lung tumor classification using SVM and ANN with GLCM features**, in *Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019)*, pp. 273-284, Springer, Singapore, 2020.
12. R. A. A. Al-falluji. **Color, shape and texture-based fruit recognition system**, *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 5, 2016.
13. M. S. Al-Tarawneh, S. Al-Habashneh, N. Shaker, W. Tarawneh, and S. Tarawneh. **Lung cancer detection using morphological segmentation and Gabor filtration approaches**, *International Journal of Engineering Research*, vol. 3, 2014.
14. S. Pudaruth, K. M. S. Soyjaudah, and R. P. Gunputh. **An innovative multi-segment strategy for the classification of legal judgments using the k-nearest neighbour classifier**, *Complex & Intelligent Systems*, vol. 4, pp. 1-10, 2018.
15. P. Kalakuntla. **Performance Analysis of kNN query processing on large datasets using CUDA & Pthreads**, *Telecommunication Systems*, 2017.
16. O. Ardhapure, G. Patil, D. Udani and K. Jetha. **Comparative study of classification algorithm for text-based categorization**, *Int. J. Res. Eng. Technol*, vol. 5, pp. 217-220, 2016.
17. T. Vigneshl and K. K. Thyagharajan. **Local binary pattern texture feature for satellite imagery classification,** *International Conference on Science Engineering and Management Research (ICSEMR) IEEE*, pp. 1-6, 2014.
18. B. Ren, and Y. Shi. **Research on spam filter based on improved Naive Bayes and KNN algorithm**, *4th International Conference on Machinery, Materials and Computing Technology*, *Atlantis Press*, 2016.

A. Al-Taei. **A smartphone-based model for human activity recognition**, *Ibn Al-Haitham Journal for Pure and Applied Science*, vol. 30, pp. 243-250, 2017..