Ensemble Learner for Covid-19 from Lung X-Ray Images

Yasmin binti Mohd Yacob, Rafikha Aliana A Raof, Phak Len Eh Kan, Norsuhaida Ahmad, Sabarina Ismail.

Faculty of Electronic Engineering Technology, University Malaysia Perlis, 02600 Arau, Perlis, Malaysia

Abstract. Despite Reverse Transcription-Polymerase Chain Reaction (RT-PCR) is the gold standard of Covid-19 detection, some underdeveloped countries are lacking financially and suffer underdeveloped health system to perform fast Covid-19 detection. Both RT-PCR and Computed Tomography (CT) scan are costly diagnosis tool, thus computed diagnostic chest x-ray (CXR) is seen as fast and affordable option to perform Covid-19 diagnosis for underdeveloped countries. Despite of other works suggest to perform Local binary Pattern (LBP) and recent feature extraction methods such as Local Phase Quantization (LPQ), this works employed Gray-Level Co-Occurrence Matrix (GLCM) because it is a powerful method to extract textured features from gray-level images of chest x-ray. The learner to classify Covid-19 detection is tested via non tree-based learner such as k-Nearest Neighbour (kNN). This work also compared the performance especially in the tree-based and voting approach classifier. The experimentation shows that tree-based which uses voting and ensemble approach to detect Covid-19 from CXR images is a possible candidate learner to be improved for the underdeveloped countries.

Corresponding author: yasmin.yacob@unimap.edu.my

1. Introduction
To date (14 October 2020), Covid-19 cases has surpassed 38.4 million infections worldwide with 1,091,747 million deaths ever since World Health Organization (WHO) officially announced it as global pandemic on 11 March 2020 [1]. Currently, the D614G mutation surfaces as the dominant variant and are believed to be 10 times more infectious in Covid-19 pandemic [2]. Although Covid-19 is seen as a disease that primarily affects the lungs, it can damage other organs such as legs, liver and kidneys. This condition is because Covid-19 can make blood cells to clump up and form clots. While large clots can cause heart attacks and strokes, small clots block tiny blood vessels that may cause damages to the organs [3]. The virus can infect wide range of age; from baby to elderly. Those with underlying health problem such as high blood pressure, diabetes, cancer, respiratory problem are more likely to be badly affected which can lead to death.

It is important to diagnose Covid-19 as quick and accurate as possible for controlling the spread of the disease and treating patients. Not only Covid-19 is overwhelming the public health systems in many developed countries, but also causes dire situation in countries with underdeveloped health systems. Even the most basic public health interventions like frequent handwashing are impossible for many
people in less developed countries. Suppressing the spread of the coronavirus through lockdowns and milder forms of social distancing are far more difficult to implement in less developed countries, in particular in slums or in refugee camps. Less developed countries lack not only the necessary testing capacities due to financial capability, but the technologies and governance structures to contain the spread effectively [4]. To name a few of the underdeveloped countries are most of African countries, Afghanistan, Bangladesh, Bhutan, Cambodia, Laos, Myanmar, Nepal, Yemen, Tuvalu, Vanuatu and a few others [5].

Reverse transcription-polymerase chain reaction (RT-PCR) is the ground truth of Covid-19 diagnosis. The fact that Covid-19 affected respiratory system makes chest X-ray (CXR) and computed tomography (CT) scan as acceptable diagnosis method too. Despite the CT scan being the gold standard for the pneumonia diagnosis, this work focused only on CXR images due to its reduced cost, fast result and its general availability. This is because CT scan machines are still scarce and costly.

Less developed countries suffers financially to provide fast and effective Covid-19 detection. Due to the lack of technology and resources these countries have less capability to perform RT-PCR test. Having said that, this work is concerned on fast and inexpensive approach to detect Covid-19 via intelligent diagnosis system from CXR beneficial to the less developed countries.

2. Related Works
This section deliberates relevant works related to image characteristics, image pre-processing and machine learning of Covid-19 lung infection or disease.

2.1. Image Characteristics of Infected Lung
Covid-19 infection to the lung is categorized to mild and moderate, severe and critical cases. Around 80% of people who have Covid-19 get mild to moderate symptoms. They may suffer dry cough or sore throat. Some people have pneumonia; a lung infection whereby the alveoli suffers inflammation. In CXR, the inflammation appears as “ground-glass opacity” since it looks like the frosted glass on a shower door. Around 14% of Covid-19 cases are severe. When the swelling becomes worse, the lungs are filled with fluids and debris [6]. In severe cases, they might have serious pneumonia where the air sacs are filled with mucus, fluid and other cells that are trying to fight the infection. In critical cases of Covid-19 infection, about 5% of the total cases, the infection can damage the walls and linings of the air sacs in lung. Your body tries to fight it, the lungs become more inflamed and filled with fluids which makes the patients’ lung even harder to swap oxygen and carbon dioxide. In CXR image, infected lungs appear as white patches. The white patches are the lungs’ air sacs filled with pus or water [6]. The Covid-19 infected lungs CXR images are shown in Figure 1 as a comparison to a normal lung in Figure 2. However, these white patches can also be confused with tuberculosis or bronchitis. This work aims to explore the detection of pneumonia variations caused by multiple pathogens inclusive of Covid-19 infections using only CXR images.

2.2. Image Processing
There are many approaches used by researchers to perform pre-processing for medical images. Parveen and Sthik (2010) employed Discrete Wavelet Transform (DWT), Wavelet Frame Transform (WFT), or Wavelet Packet Transform (WPT) for pneumonia detection in CXR images [7]. Nanni et al. (2010) employed Local binary patterns variants as texture descriptors for medical image analysis [8]. Scalco and Rizzo (2017) applied Grey-level histogram, Gray-Level Co-occurrence Matrix (GLCM),
Neighbourhood Gray Tone Difference Matrix (NGTDM), Gray Level Run-Length Matrix (GLRLM) and Gray Level Size Zone Matrix (GLSZM) to describe the grey-level patterns of medical images for radiotherapy applications [9]. Pereira et al. (2020) employed Local Binary Pattern (LBP), Elongated Quinary Pattern (EQP) which is a variation of LBP and LTP descriptors. The significant difference from EQP to LBP and LTP descriptors is that the EQP employed a quinary pattern. Local directional number (LDN), Locally Encoded Transform Feature Histogram (LETRIST) descriptors, Binarized statistical image features (BSIF) is based both on LBP and LPQ descriptors. Pereira et al. (2020) highlighted that BSIF uses a schema based on statistics of natural images and not on heuristics, such as the descriptors LBP and LPQ. Local phase quantization (LPQ) was initially proposed to provide a good texture description for noised images, affected by blur. However, LPQ emerged to be quite effective also to describe the textural content even for images not affected by blur [6]. Oriented basic image features (oBIFs) automatically learned features with inception-V3. Pereira et al. extracted 59 features from LBP, 256 features from EQP, 56 features from LDN, 413 features from LETRIST, 256 features from BSIF, 256 features from LPQ, 484 features from oBIFs and 2048 features from Inception-V3 [6]. Regardless of varieties of approaches to extract features, this works decided to employ GLCM which consists of contrast, correlation, energy and homogeneity; each of 4 directions of 0, 45, 90 and 135 degrees which
totaled up to 16 feature descriptors. This is because the method is a powerful entry-point texture feature descriptors due to the fact that the x-ray images is a gray-level image and carry textured-features.

2.3. Machine Learner

There are varieties of machine learners or classifiers commonly employed and proposed by studies related to lung disease which also includes Covid-19 works. Considering lung as the focus area, it can be deduced that the machine learner should be able to detect abnormalities from gray-level image of chest x-ray (CXR) or CT scan images. On top of that, this works is aware of difficulties to differentiate between abnormalities caused by pneumonia from virus such as SARS and MERS, bacterial and even fungus [10].

Zhou et al. [11] discussed work related to deep learning model to differentiate novel coronavirus pneumonia and influenza pneumonia in chest computed tomography (CT). On top of that, He at al. [12] proposed a 3D deep learning model called as COVNet which was fed by series of CT images to identify 3 classes; Covid-19, other viral pneumonia, and non-pneumonia. Narin et al. [13] identified Covid-19 via deep convolution neural network (CNN) using CXR images. Still more works via deep learning model by Gozes et al. to detect Covid-19 from CT images [14]. Wang and Wong detected Covid-19 via x-ray images by development of deep neural network [15]. Khan et al. [16] also identified Covid-19 via Convolution neural network from x-ray images. Still, Ozturk et al. proposed a deep model for early detection of COVID-19 cases using X-ray images [17]. [18] proposed prior-attention residual learning mechanism in CNN using CT images.

[9,11-14,18] works mainly extracted features from CT scan images and employed convolution neural network (CNN). Unlike other works, [19] employed Boosted Random Forest. Basically many works revolves around CNN, however other classifier such as Random Forest is also a potential learner and can be further improved with good techniques of voting or also called ensemble method.

3. Methodology

3.1. Dataset

The data set were acquired from Kaggle.com [20] whereby they are chest CXR images and has been labelled by the contributor as normal and Covid-19 images. Referring to Table 1, there were a total of 980 images whereby 750 used as training and 230 as testing set. For normal-labelled images, there are 105 whereby 75 were used as training and 30 as testing. On the other hand, for Covid-19-labelled images, there are 875 whereby 675 were used as training and 200 as testing.

Table 1. Distribution of Normal and Covid-19 X-ray Images Dataset.

|       | All   | Normal | Covid-19 |
|-------|-------|--------|----------|
| Train | 750   | 75     | 675      |
| Test  | 230   | 30     | 200      |
| Total | 980   | 105    | 875      |

3.2. Feature Extraction from Images

All of the images are set to grayscale images. This work decided to employ the texture-based descriptors which are the Gray-Level Co-Occurrence Matrix (GLCM) because it is one of the powerful method to extract textured features. The feature descriptors are contrast, correlation, energy and homogeneity whereby each of them were extracted from 0, 45, 90 and 135 degree of direction. Thus, there are 16 features extracted from each instances.

3.3. Machine Learners and Evaluation Measures

Regardless of most of the researchers opted to employ convolution neural network (CNN) as the learner in Covid-19 studies, this works sees ensemble tree-based learner is a possible candidate. This is because
ensemble learners is unbiased as it consider classification from tree-growing and voting approach. The experiment were run using MATLAB Machine Learning toolbox. Among the non-tree-based learners employed in the experiment are k-Nearest Neighbour (kNN) [21]. The tree-based learners compared in the experiment are AdaBoost [22], Decision Tree (DT) [23], Tree Bagger [24] and Bagged Ensemble [25]. Tree Bagger is also called as Random Forest in some literatures. Due to characteristics of tree-bagging characteristics that grows tree, this work run the experiment by considering incremention by 25 in the number of tree from 50 to 150 trees. The performance of the learners were evaluated based on standard error and Area Under Curve (AUC). This is done to compare and contrast or have a better picture of the learners performance. On top of that, out-of-bag classification error were employed for the ensemble-based learners.

4. Results and Discussion
Classification error rate from Decision Tree is 8.13% and standard error is 8.50% during training. Figure 3(a) and (b) shows that the out-of-bag error for Tree Bagger is between 0.05 and 0.06 twice when the number of trees are between 20 to 30 and also 60 to 70. As the number of trees exceeds 80, the out-of-
Figure 4. Ensemble tree learner with varying size of tree.

The ensemble tree learner reached its one-time best training performance based on out-of-bag error approaching 0.05 when the number of tree is between 30 to 35 as indicated in Figure 4(a) and (b). As the tree size exceed 40, the out-of-bag error rate does not gets better as shown in Figure 4(c) and (d). As for ensemble testing scheme, as shown in Figure 4 (a), error rate is between 0.09 to 0.1 when the tree size is between 10 to 15 trees. Then, as the number of trees increased to 75 during testing, the error rate is between 0.09 until 0.1 when the number of trees is between 5 to 10 trees. The testing error rate stayed between 0.09 to 0.1 for number of trees between 10 to 20 although the number of trees increased until 125 during testing.

During AdaBoost training session as shown in Figure 5, the error rate is between 0.07 to 0.08 regardless of increased number of trees. Such error is possible when the number of tree is between 15 to 30 as shown in Figure 5(a), number of trees is between 15 to 40 trees in Figure 5(b), number of trees between 20 to 60 in Figure 5(c) and number of trees more than 70 in Figure 5(d). The training tends to get error less than 0.07 as trees exceed 100 trees. For AdaBoost testing, the error is less than 0.1 when the number of trees is more than 10. As shown in Figure 5(b), the error is between 0.09 to 0.1 when the number of trees is between 15 to 40 trees. As the number of trees is expanded to 100 in Figure 5(c), the error is between 0.09 to 0.1 when the trees is between 20 to 48. Surprisingly, the the error almost exceeded 0.12 when the number of trees is more than 60. The test error is at 0.1 when the number of trees is between 5 to 30. It seems that AdaBoost need less number of trees of at least 15 to be stabil with low error rate.
during training. On the other hand, if the number of trees is too big it may worsen the error rate. This pattern also holds during testing process. It was observed in Figure 3, Figure 4 and Figure 5, the out-of-bag error for the ensemble-based learners were reducing as the number of tree increases during training. However during testing, as the number of trees become too big the out-of-bag error gets worse.

To compare learner’s performance via Area Under Curve (AUC), the left line must be approaching the left side of the axis and the right line should be approaching the top horizontal axis. The area should be more than 0.7 and if it at 1.0 means the classifier is a very good discriminant or classifier. As seen in AUC for DT in Figure 6(a), the left line is the nearest to the y-axis, thus Decision Tree is the best discriminant based on AUC measure. The worst is KNN as the left line is the furthest to the right of y-axis as shown in Figure 6(b). Nevertheless, the Tree Bagger is potentially a good classifier or discriminant as the left line approaching the y-axis. Out of curiosity, this work plotted the AUC for Tree Bagger increased by 25 number of tree starting from 50 until 150. It turned out the AUC are almost the same as shown in Figure 6(c) regardless of increased tree size.

![Figure 5. AdaBoost learner with varying size of tree.](image)
Based on both error rate and AUC measure, it is highly likely that a tree-based learner that is fused with voting scheme and ensemble approach is good as learner or classifier for Covid-19 detection from lung x-ray images. Thus, further work need to be done to develop a robust and effective learner in order to benefit underdeveloped country to perform cost-saving Covid-19 detection.

References

[1] Worldometer Coronavirus, [Online] . Available: https://www.worldometers.info/coronavirus, Last Updated : 14 October 2020.
[2] D614G mutation now the dominant variant in the global COVID-19 pandemic,[Online] Available: https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/update-36-long-term-symptoms.pdf?sfvrsn=5d3789a6_2
[3] Covid (coronavirus): Long-term effects, [Online] Available: https://www.mayoclinic.org/diseases-conditions/coronavirus/in-depth/coronavirus-long-term-effects/art-20490351
[4] M. Bruckner and R. Mollerus, UN/DESA Policy Brief #66 : Covid19 and the Least Developed Countries, [Online] Available: https://www.un.org/development/desa/dpad/publication/un-desa-policy-brief-66-covid-19-and-the-least-developed-countries/.
[5] Underdeveloped Countries 2020, World Population Review, [Online] Available: https://worldpopulationreview.com/country-rankings/underdeveloped-countries, Updated :15 October 2020.
[6] R.M. Pereira, D. Bertoloni, L.O. Teixeira, C.N. Silla Jr. and Costa, Y.M.G, “Covid-19 Identification in Chest X-ray Images on Flat and Hierarchical Classification Scenarios”, Computer Methods and Programs in Biomedicine, 194(2020), 105532.
[7] N. Parveen and M.M. Sathik (2010), “Detection of pneumonia in chest x-ray images”, J. Xray Sci. Technol., 19(4), pp. 423-428, 2011.
[8] L. Nanni, A. Lumini, S. Brahanam, “Local binary patterns variants as texture descriptors for medical image analysis”, Artif. Intell. Med. 49(2), pp. 117–125, 2010.

[9] E. Scalco and G. Rizzo, “Texture analysis of medical images for radiotherapy applications”, Br. J. Radiol. 90(1070), 20160642, 2017

[10] Hwang EJ, Park S, Jin K-N, Kim JI, Choi SY, Lee JH, Goo JM, Aum J, Yim J-J, Cohen JG, Ferretti GR, Park CM, “Development and Validation of a Deep Learning– Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs”, JAMA Netw Open 2019;2:e191095.

[11] M. Zhou, Y. Chen, D. Wang, Y. Xu, W. Yao, J. Huang, X. Jin, Z. Pan, J. Tan, L. Wang, Y. Xia, L. Zou, X. Xu, J. Wei, M. Guan, J. Feng, H. Zhang, J. Qu, “Improved deep learning model for differentiating novel coronavirus pneumonia and influenza pneumonia”, medRxiv 2020.

[12] K. He, X. Zhang, S. Ren, J. Sun, “Deep residual learning for image recognition”, in the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[13] A. Narin, C. Kaya, Z. Pamuk, “Automatic detection of coronavirus disease (COVID-19) using x-ray images and deep convolutional neural networks”, arXiv: 2003.10849, 2020.

[14] O. Gozes, M. Frid-Adar, H. Greenspan, P.D. Browning, H. Zhang, W. Ji, A. Bernheim, E. Siegel, “Rapid AI development cycle for the coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis”, arXiv:2003.05037, 2020.

[15] L. Wang, A. Wong, “COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images”, arXiv: 2003.09871, 2020.

[16] A.I. Khan, J.L. Shah, M. Bhat, “CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest x-ray images”, arXiv:2004.04931, 2020.

[17] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, “Automated detection of covid-19 cases using deep neural networks with x-ray images”, Comput. Biol. Med., 103792, 2020.

[18] Wang, J., Bao, Y., Wen, Y., Lu, H., Luo, H., Xiang, Y., Li, X., Liu, C. and Qian, D., “Prior-Attention Residual Learning for More Discriminative Covid-19 Screening in CT Images”, IEEE Transaction on Medical Imaging, 39(8), pp. 2572-2583, August 2020.

[19] Iwendi C, Bashir AK, Peshkar A, Sujatha R, Chatterjee JM, Pasupuleti S, Mishra R, Pillai S and Jo O, “COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm”, Front. Public Health 8:357. doi: 10.3389/fpubh.2020.00357, 2020.

[20] Asraf, A., “Covid19 Penumonia Normal Chest Xray PA Dataset”, [Online]. Available: https://www.kaggle.com/asraf047/covid19-pneumonia-normal-chest-xray-pa-dataset/activity, Version 1 updated. 4 months.

[21] Altman, Naomi S., "An introduction to kernel and nearest-neighbor nonparametric regression", The American Statistician. 46(3), pp. 175–185, 1992.

[22] Freund, Yoav and Schapire, Robert E, "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of Computer and System Sciences, 55, pp. 119–139, 1997.

[23] Quinlan, J. R., "Simplifying decision trees", International Journal of Man-Machine Studies, 27 (3), pp. 221–234, 1987.

[24] Ho, Tin Kam, Random Decision Forests, in Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.

[25] Polikar, R., "Ensemble based systems in decision making", IEEE Circuits and Systems Magazine, 6 (3), pp. 21–45, 2006.