Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia
Vol. 8, 2022, page 40-48

Performance Evaluation of Lung Segmentation of Ground Glass Opacity COVID-19 Based on X-Ray Images

Nur Hazreena Azmil, Norma Alias
Department of Mathematical Sciences, Faculty of Science
Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia

Abstract
The current COVID-19 pandemic directly threatens the health and well-being of individuals all over the world and vulnerable to be affected by the COVID-19 virus. The automatic detection of such a virus is an important concern. According to (Ai *et al.,*2020), the role for deep neural network approaches such as convolutional neural network and other features classifier for the detection or characterization the accuracy, precision, and sensitivity of COVID-19 X-Ray images. The purpose of this study is to investigate the classification of X-ray images affected by COVID-19. The application of machine learning in data analytics in term of the accuracy, sensitivity, and precision of data images. Based on this research drawback, the accuracy can be improved in diagnosing this disease using machine learning. The goal of this research is to govern and modify a machine learning and unsupervised clustering-based lung segmentation algorithm. Data source is based on lesion region in two-dimensional enhanced X-Ray images. Image processing of structured data can be analysed by Convolutional Neural Network (CNN), X-Ray training model, confusion matrix, and prediction parameters. The classification addresses the appearance of the ground glass opacities (GGOs) nodules on lung of Covid-19 patient. A modified K-means algorithm and Fuzzy C-Means will be used to detect and segment of these regions. Support Vector Machine (SVM) able to be detected, the infected areas by obtaining in the solid false positives of the GGOs. This is inputs and targets for the supervised machine learning classifiers. The classification of the clinical specimens obtained from the corona-infected patients with the help of some machine learning techniques like Convolutional Neural Network, Support Vector Machine, K means clustering, and Fuzzy C-Means. The result shows that the RT-PCR test is the best options for diagnosing Covid-19 patients, as well as being less expensive than the PCR test. The use of advanced artificial intelligence (AI) techniques in conjunction with radiological imaging can aid in the accurate detection of this disease, as well as in overcoming the problem of a lack of specialized physicians in remote villages. Th implement of Phyton and MATLAB support the computation. As conclusion, the result Fuzzy C-Means achieves high (AUC) compared to K-means clustering consists of VGG16 90.04%, ResNet152V2 95.13%, and InceptionV3 91.20%. The accuracy between fuzzy compared to without fuzzy, achieves less accuracy but with fuzzy, recall reduced 19.44% and improved the precision. This research reviews the method of analysing X-Ray images that being affected by COVID19 and the techniques of machine learning to extract the features of lung accurately.

Keywords COVID-19, Ground Glass Opacity, Support Vector Machine, CNN

1. Introduction
Coronavirus disease 2019 (COVID-19) has rapidly spread worldwide, and the World Health Organization (WHO) declared it a pandemic on March 11, 2020.The present of COVID-19 has impacted every element of life and has drastically altered our behaviour patterns. Most countries including Malaysia have reported some COVID-19 cases which have progressively increased over a time, which had an impact on many sectors as well as the economy, industry, and health-related that had impact on people. According to

(Sohrabi,2020), the rapid rate of spread virus has strained healthcare systems worldwide due to inadequate medical equipment, key protective equipment, and qualified providers .Since the outbreak began, the WHO has reported 29,892 deaths were caused by the SARS-CoV-2 .As a result, COVID-19's effects will persist and have a long-term impact on numerous elements of our life as well as the long-term surgery of COVID-19 survivors, especially those who were infected by this virus that led to the appearance of ground lesion.

Thus, it's critical to research and analysed the appearance of ground glass opacity that affects lung segmentation based on chest X-ray. Patients infected with the COVID-19 virus develop severe pneumonia, which generally leads to death it come to the higher stage. Radiological evidence has demonstrated that the disease causes interstitial involvement in the lungs and lung opacities, as well as bilateral ground glass opacities and patchy opacities.

Previous researcher (Long & Ehrenfeld,2020) stated that the use of artificial intelligence is an inevitable requirement for mitigating the effects of the COVID-19 pandemic crisis and increasing and accelerating disease diagnosis and diagnostic success [2]. Lung and chest radiological images, whether COVID-19 or non-COVID-19, are important clinical data for diagnosis in this context. As a result, according to the research of [3] taking insufficient specimens could result in a false result.

According to (Lai *et al.,*2020) indicate that the virus is harboured most with little or no symptoms but can also lead to a rapidly progressive and often fatal pneumonia in 2–8% of those infected [4]. Due to the rapid increase in number of new and suspected COVID-19 cases, there may be a role for deep neural network approaches such as convolutional neural network and other features classifier for the detection or characterization the accuracy, precision, and sensitivity of COVID-19 X-Ray images .As a result,(Ai *et al.,* 2020) mentioned the appearance of ground glass opacity that formed by pneumonia can be detected on Ct Scan and X-ray images according to the stage of disease and disease severity and associated comorbidities.[5]

Thus, to reduce of accuracy in taking PCR test or RTK test , the development of a CNN model consists of convolutional layers is used in validation the accuracy of the train and test data images. Data analytics can be used in predict the appearance of consolidation or ground glass opacity (GGOs) is Fuzzy C-means (FCM) and K-Means. To classify the stages of COVID-19, Support Vector Machine (SVM) is being used to categorize the diseases based on severity.

This research is to govern mathematical modelling for classifying the MRI images. The modification unsupervised clustering integrated with lung segmentation algorithm is used to detect lesion region in two-dimensional enhanced X-Ray images. The implementation is based on various combinations of feature reduction, and classifier of multi -class segmentation of lung infections types and multi-class independent component analysis (ICA) dimensionality reduction method. Some indicator of performance evaluations in feature extraction, localization, and characterization of opacity /lesion region in X-Ray image are under investigation.

2. Literature Review

2.1. Convolutional Neural Network

CNN model has been used as feature extraction in analysing lung segmentation of COVID-19. Filter-based feature extraction is used in the CNN model, which can be effective for classification. CNNs can classify images with complex identities. Many weight parameters can be reduced using the CNN architecture. According to [6], has mentioned that CNN model was developed to solve COVID-19 detection problem using the chest radiography data. In this study, CNN was used to develop an automatic diagnostic system that used chest X-ray analysis results to determine whether a person is COVID-19-affected or normal.

2.*2 Support Vector Machine*

In recent years, the support vector machine algorithm has been used extensively in bioinformatics, face detection, text categorization, and other fields. have explained how the SVM is being used in the early detection of heart failure, which can help cardiologists to improve the diagnosis process. Thus, SVM also can be used in detection severity of COVID-19. This system occur by uses two separate models first to eliminate irrelevant features, and the second model is used as a predictive model. [7] applied the SVM classification method based on phenological metrics to identify the changes and the main agricultural classes in the concerned area.

2.*3 Stages of COVID-19*

Staging systems provide valuable frameworks and benchmarks for clinical decision making in patient management and as evidence-based treatment selection .[8] stated that, evidence increased indicates that endothelial cells in many organs as well as other sever acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

- **COVID-19 Stage 1:** Viral Entry and Replication (Asymptomatic)
- **COVID-19 Stage 2:** Viral Dissemination (Mild or Moderate)
- **COVID-19 Stage 3:** Multi-system Inflammation (Severe)
- **COVID-19 Stage 4:** Endothelial Damage, Thrombosis, and Multi-organ Dysfunction (Critical)

2.*3 Ground Glass Opacity (GGOs)*

The segmentation and classification texture features in GGOs and lesion region will be classified into stage 1,2,3,4.The stages have been defined as below [9]:

Stage 1: Predominantly basilar perivascular GGOs and develop fibrotic changes.
Stage 2: Subtle to more diffuse GGOs in the acute phase, which may mimic pulmonary enema. There are distinct tiny centrilobular pulmonary nodules, measuring less than 6mm.
Stage 3: Undergone chemoradiation followed by immunotherapy and complicated by pneumonitis. More than 6mm size of centrilobular pulmonary nodules.
Stage 4: GGOs typically recede and there may be residual centrilobular nodules due to passage of blood into the alveoli. The distinct feature of hypersensitivity pneumonitis or fibrotic stage

3. Methodology

3.1. Data Accusation

The data in this study is data Driven and a machine learning solution for image processing is the important methodology to be focused. Radiological imaging, 624 Covid-19 and 20 normal chests X-Ray images is a repository of images established by a team of board-certified radiologists

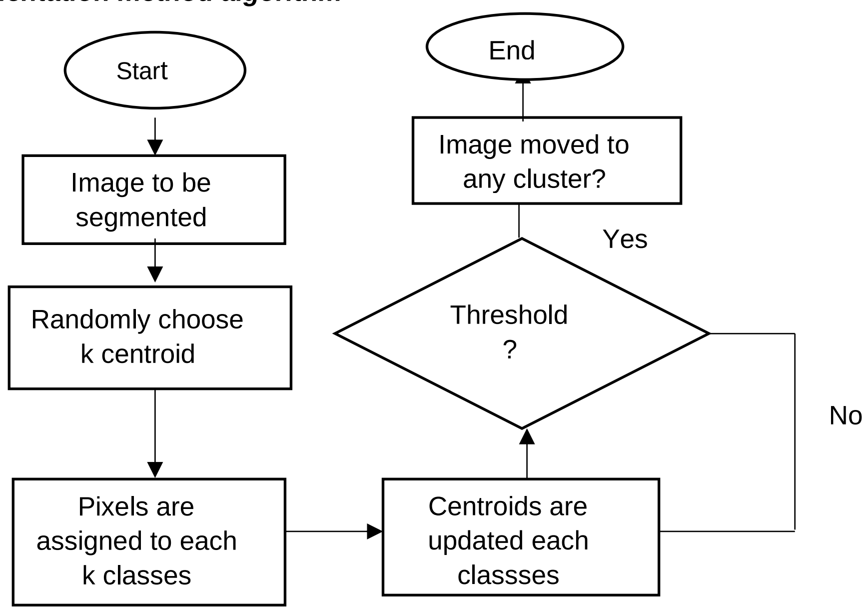
3.2. Area of ground glass opacity

The purpose of determining the area of grounded glass opacity is to determine the index on the level of grounded in the X-ray and the classification stage of the COVID-19 that is affected by the patient by using image segmentation in MATLAB. The ground truth is from lung opacity Kaggle.

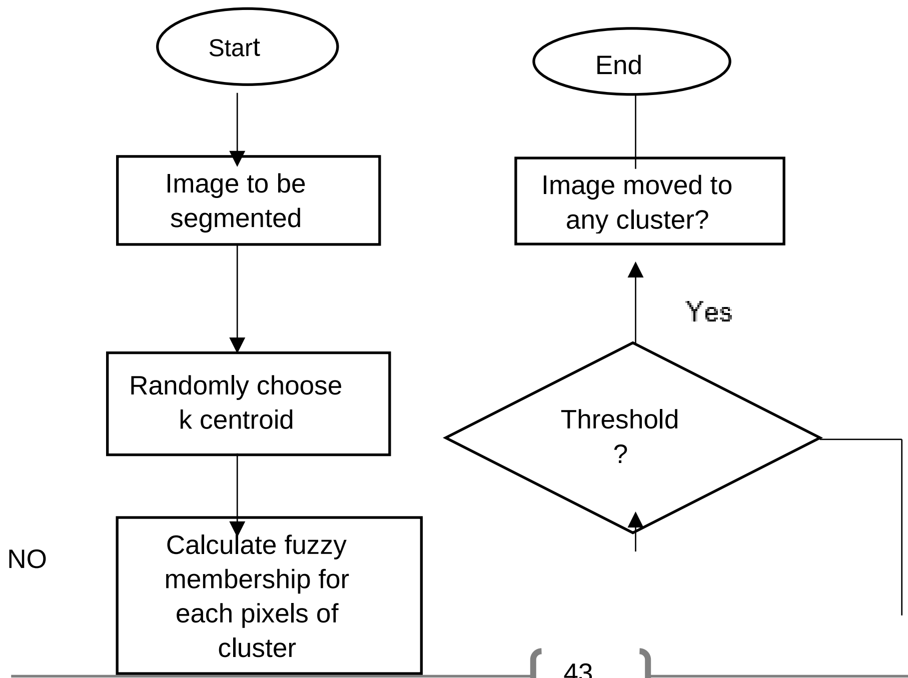
3.3. Segmentation approaches

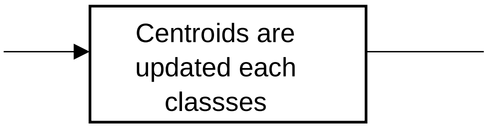
(Yang *et al.,*2020) decompose and partition the 2 D segmentation problem will reduce the model complexity by an order of magnitude, size and at the same time, significantly improve the segmentation accuracy. Modify and proposed several clustering-based GGOs segmentation approaches using K -Mean and Fuzzy-C-Means will improve the accuracy of detection

3.4 K-Means Segmentation method algorithm

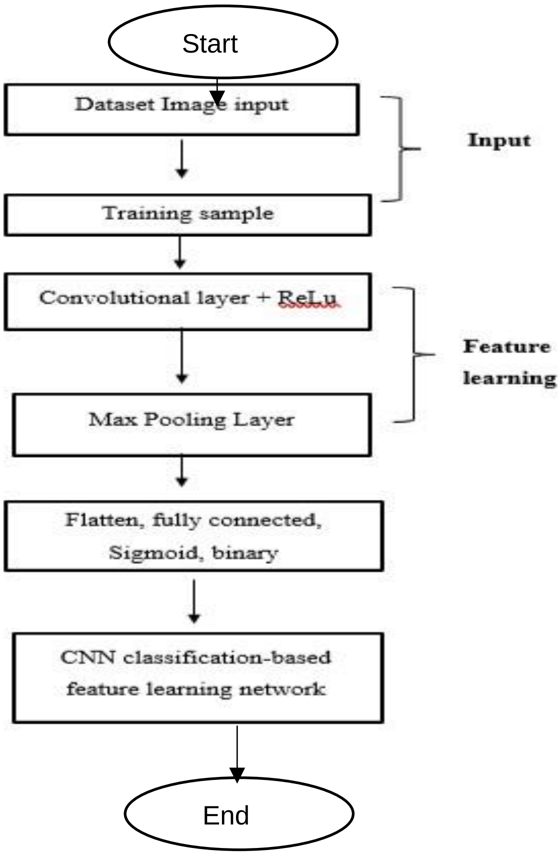


3.5 Fuzzy C-Means Algorithm





3.6. Convolutional Neural Network algorithm



3.7 Parameter Evaluations Metrics

Parameter of CNN Model ,K-means and Fuzzy C-means are calculated by using evaluation metrics as below

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3.1)$$

$$Percision = \frac{TP}{TP + FP} \qquad (3.2)$$

$$Sensitivity = \frac{TP}{TP + FN} \qquad (3.3)$$

$$Recall = \frac{TP}{TP + FN} \quad (3.4)$$

$$F1 - Score = 2.\frac{Percision \times Recall}{Percision + Recall} \quad (3.5)$$

4. Results and discussion

4.1 CNN model

The dataset contains two class which is training and testing of the CNN algorithm. The training dataset contained 624 COVID-19 X-ray images and 20 normal X-ray images in the training dataset and for a total of 644 X-ray images. The Figure 4.1 depicts model accuracy during training and validation as a graph, with the blue curve representing CNN training accuracy and the orange curve representing validation accuracy. According to Figure 4.1, the CNN's training accuracy remains consistent after the 10 epochs and the CNN also shows a consistent validation accuracy after 25 epochs. The loss during the CNN training and validation is depicted in Figure 4.2 CNN's training loss of is minimal and consistent from the first epoch, while validation decreases till last epochs. The following results demonstrate the efficiency of the CNN model proposed in this study

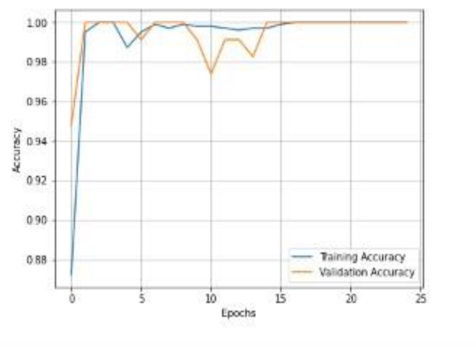


Figure 4.1 Model Accuracy

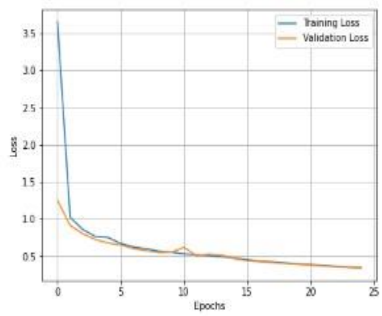


Figure 4.2 Model Loss

Figure 4.3 Model performance on test data

Based on Figure 4.3 above, the average of testing accuracy, sensitivity, precision and FI-Score is characterized. We accomplished 100% testing accuracy, 89.49% of testing sensitivity and 90% of testing precision as well as 91.01% testing FI-Score. Furthermore, we depicted the confusion matrix of worst and best cases in figure 4.5. In the worst of confusion matrix 20 and 216 testing images correctly classified as 1 is Normal and 0 is Pneumonia respectively.

```
[ ]  print(confusion_matrix(test_labels, pred))

     [[ 20 370]
      [ 18 216]]
```

Figure 4.4 Value of confusion matrix

```
print("True Positive:", TP)
print("False Negative:", FN)
print("False Positive:", FP)
print("True Negative:", TN)

True Positive: [ 20 216]
False Negative: [370  18]
False Positive: [ 18 370]
True Negative: [216  20]
```

Figure 4.5 Value of TP,FN,FP,TN

4.2 K-Means clustering & Fuzzy C-Means

Table 1 : K-Means Performance Metrics

Model	**AUC (%)**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
VGG16	86.15	76.39	71.11	88.89
ResNet152V2	70.02	52.77	51.72	83.33
InceptionV3	87.03	50.00	50.00	100

Table 2 : Fuzzy C-Means Performance Metrics

Model	**AUC (%)**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
VGG16	90.04	59.72	100	19.44
ResNet152V2	95.13	77.77	69.23	100
InceptionV3	91.20	87.50	82.92	94.44

Table 3: Without Fuzzy C-Means

Model	**AUC (%)**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
VGG16	98.91	97.22	97.22	97.22
ResNet152V2	61.11	61.11	56.25	100
InceptionV3	85.26	84.72	76.59	100

The first result in Table 4.1 shows that VGG16 in k-means has an accuracy of 76.38% while fuzzy c-mean only has 59.72. In fuzzy, the accuracy of ResNet152v2, and InceptionV3 in fuzzy is greater than accuracy of K-means. However, the AUC of fuzzy is greater than that of K-means for all types of models that result in a higher rate of true positive in the image analysis. InceptionV3 k-means, fuzzy and without fuzzy achieves better recall. According to the results, the fuzzy filter reduced the recall in Table 4.3 over the without fuzzy filter. As a result, we can see that fuzzy improved the performance of ResNet15V2, and InceptionV3. To summarise, because of the high AUC of fuzzy c-means in correctly identifying the lesion region, this method can be used to assist the radiologist in making clinical decisions.

4.3 Support Vector Machine

Class	Precision	Recall	f1-score	Support
Not infected	0.75	0.75	0.75	12
Mildly infected	0.82	0.74	0.78	19
Severely infected	0.94	1.00	0.97	29

Figure 4.6 Classification of MRI Images

Figure 4.6 shows, high success rate of predicting severely infected cases, which is very crucial for COVID-19 prediction. However, the score for the rest two classes in on the lower side because of the variable nature of COVID-19 because people with no symptoms are also getting infected with COVID-19. We cannot guarantee if the person showing mild symptoms will not be affected by COVID-19 . Because of these uncertainties in the dataset, the accuracy is on the lower side.

5. Conclusion

In this study, show that by using transfer learning and leveraging pre-trained models, we can achieve very high accuracy in detecting COVID-19. Also, with the fuzzy filter, it is possible to achieve a recall of 1.0 with more than one pre-trained model. The best model was a combination of VGG16 and ResNet152V2. Furthermore, CNN model has been governed. The result of this analysis is to provide radiologists, data scientists, and the research community with a multi-input CNN model and SVM model that may be help to diagnose the stage of COVID-19 early.

Acknowledgement

The researcher would like to express my thanks and gratitude to research supervisor Prof. Madya Dr. Norma Binti Alias for her supervisions, supports, advice, guidance, suggestions, and comments in project from beginning to the last. Thanks to all people that share their knowledge and lend their times to assist writing this project.

References

[1] Sohrabi, C., Alsafi, Z., O'neill, N., Khan, M., Kerwan, A., Al-Jabir, A., ... & Agha, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID19). International journal of surgery, 76, 71-76.

[2] Long, J. B., & Ehrenfeld, J. M. (2020). The role of augmented intelligence (AI) in detecting and preventing the spread of novel coronavirus. Journal of Medical Systems, 44(3), 1-2.

[3] Yang, Y., Yang, M., Shen, C., Wang, F., Yuan, J., Li, J., ... & Liu, Y. (2020). Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. MedRxiv.

[4] Lai, C. C., Liu, Y. H., Wang, C. Y., Wang, Y. H., Hsueh, S. C., Yen, M. Y., ... & Hsueh, P. R. (2020). Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): Facts and myths. Journal of Microbiology, Immunology and Infection, 53(3), 404-412.

[5] Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., ... & Xia, L. (2020). Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology.

[6] Reshi, A. A., Rustam, F., Mehmood, A., Alhossan, A., Alrabiah, Z., Ahmad, A., ... & Choi, G. S. (2021). An efficient CNN model for COVID-19 disease detection based on X-ray image classification. Complexity, 2021.

[7] Lebrini, Y., Boudhar, A., Hadria, R., Lionboui, H., Elmansouri, L., Arrach, R., ... & Benabdelouahab, T. (2019). Identifying agricultural systems using SVM classification approach based on phenological metrics in a semi-arid region of Morocco. Earth Systems and Environment, 3(2), 277-288.

[8] Libby, P., & Lüscher, T. (2020). COVID-19 is, in the end, an endothelial disease. European heart journal, 41(32), 3038-3044.

[9] Parekh, M., Donuru, A., Balasubramanya, R., & Kapur, S. (2020). Review of the chest CT differential diagnosis of ground-glass opacities in the COVID era. Radiology.