SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Computed Tomography Image Analysis on COVID-19 Cases using Machine Learning Approaches

Jasmine Wang Thye Wei[1], Hang See Pheng[1,*], Ong Kok Haur[2]

[1] Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
[2] Bioinformatics Institute, A*STAR, Singapore

ARTICLE INFO

Article history:
Received 10 April 2023
Received in revised form 2 September 2023
Accepted 10 September 2023
Available online 27 September 2023

Keywords:
COVID-19; Computer Tomography image; image segmentation; GLCM feature extraction; medical image classification; Machine Learning; Deep Learning

ABSTRACT

Coronavirus disease (COVID-19) has become a serious worldwide health concern affecting the respiratory system since December 2019. Computed Tomography (CT) image analysis and identification are powerful tools for diagnosing COVID-19. However, due to the disparity of the distribution and form of COVID-19 infection and the diverse degrees of infection severity, the classification of the CT images is challenging, especially manually. Therefore, this study aims to employ artificial intelligence techniques, including machine learning and deep learning algorithms, to classify COVID-19 from CT images. The grey-level co-occurrence matrix (GLCM) features were computed and fed into machine learning classifiers, namely Support Vector Machine, Random Forest, K-Nearest Neighbour, Logistic Regression, and Naïve Bayes model for training purposes. Deep learning models, including ResNet50, Densenet121, Inception, and VGG16, were trained using raw data scaled and transformed to greyscale mode. The performances of machine learning and deep learning models were assessed on the testing data. Random Forests, ResNet50, and DenseNet121 outperform all the other models by achieving 100% accuracy, precision, sensitivity, and specificity on the dataset applied. The performance of machine learning models can be further improved by obtaining the optimised parameters in future research.

1. Introduction

According to the World Health Organization (WHO), Coronavirus disease (COVID-19) is a viral infection caused by the SARS-CoV-2 virus [1]. Symptoms of infection include dry cough, fever, and a loss of taste. People with underlying medical conditions are more likely to develop serious illnesses, including severe respiratory issues, organ failure, and death. Since COVID-19 is extremely infectious and spreading worldwide aggressively, early detection is critical [2]. Computer-Aided Diagnosis (CAD) systems that use medical image processing techniques, as well as artificial intelligence (AI), assist clinicians by improving their diagnostic efficiency since the clinical diagnosis time can be reduced, potentially lowering death rates [3]. CAD is based on highly complex pattern recognition, which is a branch of mathematical principles and methods for classifying and identifying objects, events,

* *Corresponding author.*
E-mail address: sphang@utm.my

https://doi.org/10.37934/araset.32.2.404416

processes, signals, and situations. The development of high-performance CAD is a way to alleviate the rising pressures on radiologists by aiding in the classification of COVID-19 cases [4].

Medical image processing is the use and investigation of 3D image collections of the human body, most typically collected from an X-ray, Computed Tomography (CT), Positron Emission Tomography (PET), and Magnetic Resonance Imaging (MRI) scanner. These images are employed to diagnose diseases, assist in surgical planning, and facilitate medical research endeavors. The experimental findings obtained using CT images were superior to those obtained using X-ray images [5]. The Hounsfield unit (HU) is a relative quantitative measurement of radiodensity used to analyses computed tomography (CT) images [6].

The segmentation method is used to portray the image in a form that makes it easier to analyses and more meaningful [7]. In the medical field, image segmentation aims to identify specific areas of interest, such as locating the tumor, abnormalities, and degree of infection [8]. Medical image segmentation is vital for further clinical analysis, diagnoses, therapy planning, and illness progression monitoring [9]. Threshold-based, edge-based, and clustering-based segmentation are some image segmentation techniques [10]. The edge detection approach is the most common method that works based on detecting discontinuities in grey level and colour. These edges frequently indicate boundaries between objects [11]. However, the limitation of the edge-based segmentation method is that the presence of noise may impact its performance [12]. On the other hand, the marching squares method that creates scalar field contours in two dimensions is another popular approach to finding the contour [13].

Meanwhile, the image features summaries the data in the pixels that make up the image [14]. Feature extraction is a step in the dimensionality reduction process, in which an initial collection of raw data is separated and reduced to more manageable categories. Haralick *et al.,* [15] introduced the Grey Level Co-occurrence Matrix (GLCM) feature extraction technique often used in biomedical image processing. As a function of the grey level, GLCM indicates the occurrence frequency of each grey level in a pixel located at a specific geometric position compared to that of other pixels [16].

Medical image classification is one of the most significant problems in image recognition, intending to classify medical images to aid clinicians in illness diagnosis or further study [17]. There are various medical image classification techniques, such as the decision tree method, neural network classifier, Bayers classification methods, and rule-based classification [18]. The structures and concepts of machine learning models such as Support Vector Machine, Random Forest, and K-nearest Neighbour can be illustrated in Figure 1, 2, and 3, respectively.

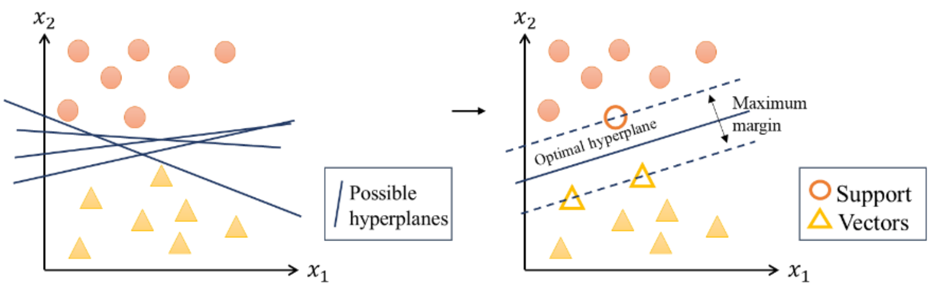

Fig. 1. Structure of Support Vector Machine Classifier

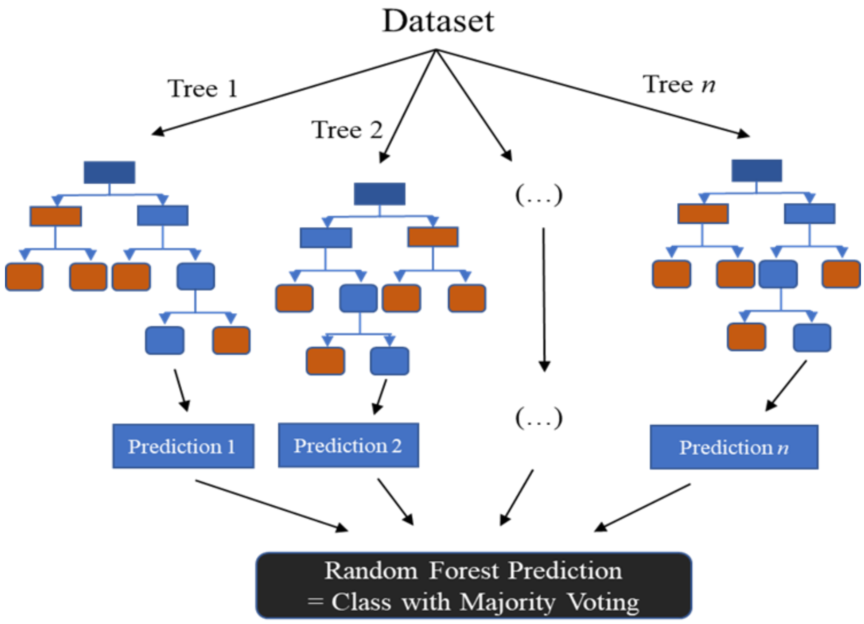

Fig. 2. Structure of Random Forest Classifier

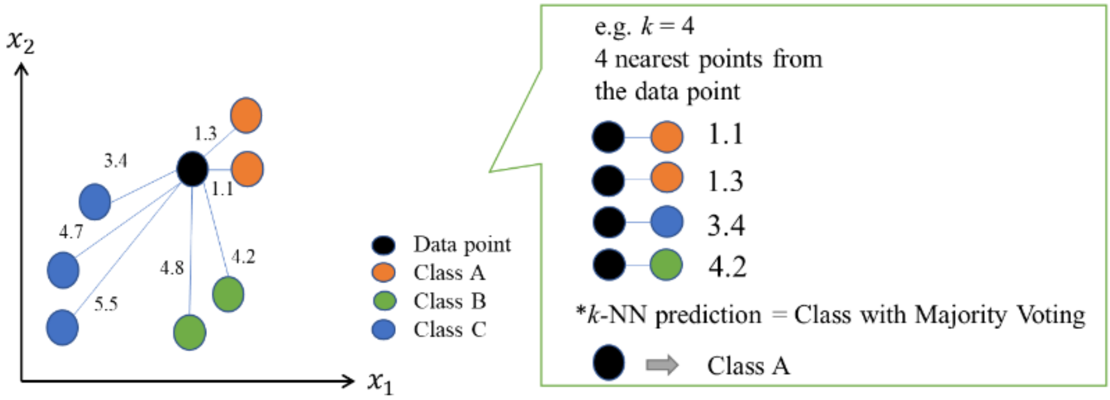

Fig. 3. Structure of K-nearest Neighbour Classifier

RF classifier exhibits strong generalisation since it combines individual tree judgments for the final decision [19]. The COVID-19 lung image classification based on the Support Vector Machine has good performance in predicting COVID-19 and normal patients using chest X-ray images, achieving an accuracy of 96% [20]. Yoo *et al.,* [21] suggested that the decision tree classifier could be utilised to pre-screen patients for triage and decision-making before RT-PCR data are available.

Furthermore, deep learning has achieved significant success in numerous fields of medical imaging as it has increased the capacity to recognize and classify characteristics in pictures using deep models [22]. DenseNet, Xception, and VGGNet are some of the feature extractors that have been proposed that work for medical image classification. Wang *et al.,* [23] redesigned COVID-Net to enhance the COVID-19 CT images' prediction accuracy and computation efficiency. Tabik *et al.,* [24] proposed a COVID-SDNet methodology based on a convolutional neural network (CNN) for the classification of chest X-rays (CXR). Although CXR is a time-effective assessment tool, it is inferior to CT scans which provide additional viewpoints and much more detail.

A chest CT scan is useful even before COVID-19 symptoms develop, and it can precisely detect the aberrant characteristics identified in images [25]. It is widely recognised that bilateral ground-glass opacity (GGO) with associated consolidation is the most prevalent computed tomography (CT) finding in COVID-19 infection [26]. However, these symptoms might occur in irregular distribution and shape. Also, the severity of each patient might differ from others. Additionally, the relatively large number of variable objects, particularly the scanned areas beyond the lungs that are irrelevant to the diagnosis of COVID-19, makes the classification of the CT images challenging, especially manually [27].

Hence, this study aims to:

i. propose artificial intelligence methods for COVID-19 cases detection on CT images
ii. compute Grey-Level Co-occurrence Matrix features from automatic lung segmentation results in the training of artificial intelligence classification models
iii. assess the performance of artificial intelligence models that consist of machine learning and deep learning approaches in COVID-19 CT image classification. This study will compare the classification performances of machine learning models using handcrafted GLCM features and deep learning models with automatic feature extraction.

2. Methodology

2.1 Materials and Methods

Medical Image Processing, Analysis, and Visualization (MIPAV) and Python are used to run and process the data in this study. The overview of the research methodology is illustrated in Figure 4.

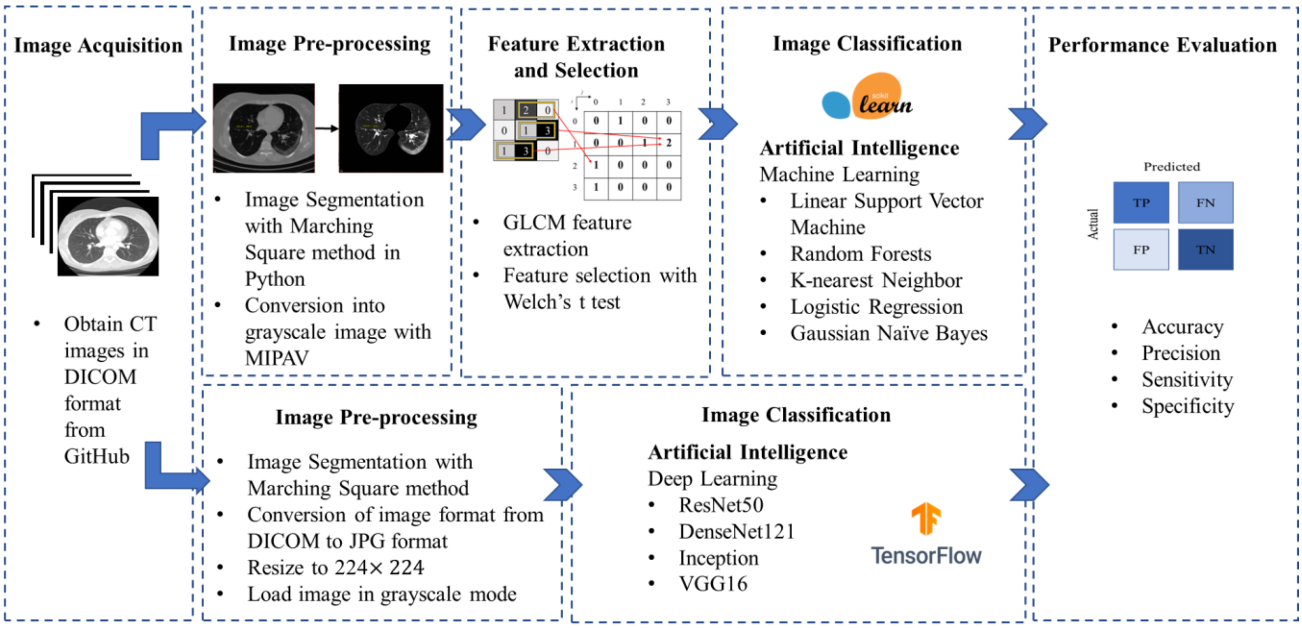

Fig. 4. Overview of the Research Methodology

2.2 Data Acquisition

In this study, a collection of a publicly available dataset, namely 'COVID-CT-MD' obtained from GitHub, is used as the raw image data and to be processed as the input data. This dataset contains volumetric chest CT scans of 169 patients with positive COVID-19 and 76 normal lung cases stored in the Digital Imaging and Communication in Medicine (DICOM) format, collected from Babak Imaging Center, Tehran, Iran [28]. COVID-CT-MD is intended to support the development of advanced machine learning and deep neural network-based solutions.

2.3 CT Image Segmentation

The lung regions were segmented using a threshold based on the CT Hounsfield Unit (HU) and Marching Squares method. The marching squares method that creates scalar field contours in two dimensions is a popular approach to finding the contour, which is the line connecting all the spots along an image's boundaries with the same intensity. Constraints are applied in the segmentation of lung regions. Firstly, the contours must be a closed set. Since a contour is made up of a set of points, the distance between the first and last points of the contour must be less than 1 pixel to ensure that the particular contour is a closed set. In addition, the contour must have a minimum volume of 1500 pixels and a maximum volume of 100000 pixels.

2.4 Feature Extraction

This study used the GLCM-based feature extraction technique to generate a compact set of discriminative texture features. The GLCM determines how frequently a pixel with a grey-level value *i* appears horizontally, vertically, or diagonally to neighbouring pixels with value *j*. Figure 5 shows the formation of the GLCM.

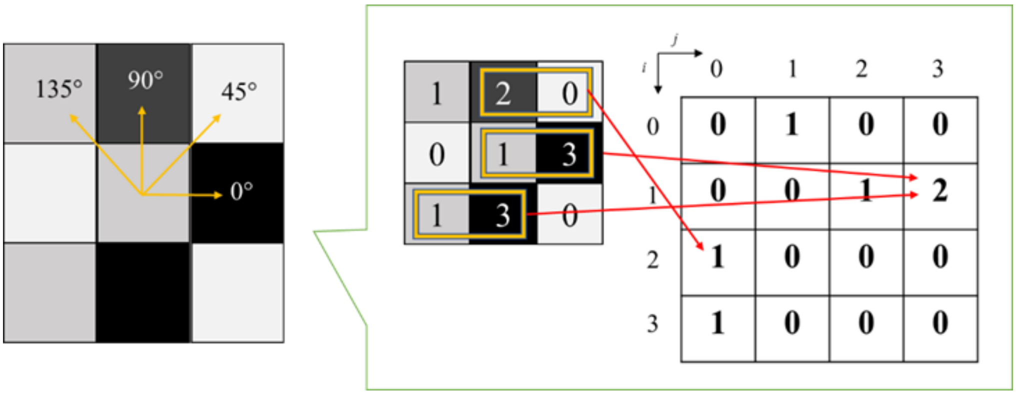

Fig. 5. Formation of Grey Level Co-occurrence Matrix (GLCM)

Mathematically, GLCM, ***P*** may be built as a frequency matrix by counting the number of times each pair of quantised grey levels appears as neighbours in the quantised image, *I*, and each P_{ij} in GLCM ***P*** can be computed as shown in Eq. (1), where *m* and *k* indicate the image's row and column number of pixels.

$$P_{ij} = \sum_{m=1}^{M} \sum_{k=1}^{K} \begin{cases} 1, & if\ I(m,k) = i, I(m + \Delta x, k + \Delta y) = j, \\ 0, & otherwise \end{cases} \tag{1}$$

A normalised GLCM, $\widetilde{\boldsymbol{P}}$ can be computed as expressed in Eq. (2).

$$\widetilde{\boldsymbol{P}} = \frac{\boldsymbol{P}}{\sum_i \sum_j P_{ij}} \tag{2}$$

Contrast, which measures the difference in grey level between the reference pixel and its neighbour pixel, can be computed as in Eq. (3).

$$Contrast = \sum_{i}\sum_{j}(i-j)^2\tilde{P}_{ij} \tag{3}$$

Angular second moment (ASM) represents the degree of grey distribution homogeneity, and texture thickness can be computed as in Eq. (4).

$$ASM = \sum_{i}\sum_{j}\left(\tilde{P}_{ij}\right)^2 \tag{4}$$

Energy gives the sum of square elements in GLCM can be expressed as in Eq. (5).

$$Energy = \sqrt{ASM} \tag{5}$$

The homogeneity of an image represents the similarity between pixels can be computed as in Eq. (6).

$$Homogeneity = \sum_{i=0}^{N}\sum_{j=0}^{N}\frac{\tilde{P}_{ij}}{1+(i-j)^2} \tag{6}$$

The dissimilarity is a distance measured between two pixels in the region of interest and can be computed as shown in Eq. (7).

$$Dissimilarity = \sum_{i}\sum_{j}\tilde{P}_{ij}|i-j| \tag{7}$$

The GLCM mean, and standard deviation for rows and columns can be computed using Eq. (8), (9), (10), and (11).

$$\mu_x = \sum_{i}\sum_{j}i\tilde{P}_{ij} \tag{8}$$

$$\mu_y = \sum_{i}\sum_{j}j\tilde{P}_{ij} \tag{9}$$

$$\sigma_x = \sqrt{\sum_{i}\sum_{j}(i-\mu_x)^2\tilde{P}_{ij}} \tag{10}$$

$$\sigma_y = \sqrt{\sum_{i}\sum_{j}\left(j-\mu_y\right)^2\tilde{P}_{ij}} \tag{11}$$

The linear dependency of grey level values in the co-occurrence matrix is represented by correlation computed as shown in Eq. (12).

$$Correlation = \sum_i \sum_j \tilde{P}_{ij} \frac{(i-\mu_x)(j-\mu_y)}{\sigma_x \sigma_y} \quad (12)$$

2.5 Feature Selection

Feature selection is the process of minimising the number of input variables when creating a predictive model. Welch's t-test is performed to determine if the means of these two data sets are significantly different. The null hypothesis, H_0 and alternative hypothesis, H_1 are shown in following

i. H_0: *There is no significant difference between the features of the two sets of data*
ii. H_1: *There is a significant difference between the features of the two sets of data*

At a 95% confidence interval, if the *p*-value obtained from Welch's t-test is less than 0.05, the null hypothesis will be rejected, indicating that features extracted from COVID-19 and normal lung CT images significantly differ at a 0.05 significance level. The result of Welch's t-test is adequate if the data fulfil the normality assumption. The Anderson-Darling test tests whether the data follows normal distribution [29]. The modified AD statistics, A*, can be expressed in Eq. (13)

$$A^* = A^2\left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right) \quad (13)$$

The normality is rejected if the value of A* exceeds the critical value of 0.752 and *the p*-value is less than 0.05 at a 0.05 significance level [30]. However, according to the central limit theorem (CLT), when the sample size is equal to or greater than 30, the distribution of the sample approaches a normal distribution [31].

2.6 Machine Learning Classification

Machine learning models are used to classify normal lung and COVID-19 CT images with those features selected as input information. The dataset is separated into three sections, with 60% of the data utilised to train machine learning models, 30% for cross-validation, and 10% for model testing. CT image classification employs five machine learning classifiers: Linear Support Vector Machine, Random Forest, *K*-Nearest Neighbour, Logistic Regression, and Gaussian Naïve Bayes models from the scikit-learn library of Python. A cross-validation is a resampling approach that tests and trains a model on multiple iterations using different partitions or folds of the data. This approach is used to evaluate the effectiveness of the models, especially to avoid overfitting in the prediction models.

2.7 Deep Learning Classification

CT image classification employs four pre-trained deep learning classifiers: DenseNet-121, ResNet-50, Inception, and VGG-16 models from the TensorFlow library. The loss function is a fundamental component of neural networks. The computed output of the loss function is the loss or error, which is the difference between the model's prediction using a set of parameter values and the target values, as shown in Eq. (14).

$$L = -\sum_{j=1}^{M} y_j \log(\hat{y}_j) \tag{14}$$

2.8 Performance and Evaluation Metrics

Parameters that measure the classification performance for the deep learning model are accuracy, precision, sensitivity, and specificity, which can be computed as shown in Eq. (15), (16), (17), and (18), respectively. TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively [32].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{17}$$

$$Specificity = \frac{TN}{TN + FP} \tag{18}$$

3. Results

3.1 CT Image Segmentation

The raw dataset was subjected to medical image processing to segment the region of interest, which is the lung area. The lung regions were segmented using a threshold based on the CT Hounsfield Unit (HU) and Marching Squares method to find the contours of the lung. Figure 6 illustrates the procedures of lung segmentation.

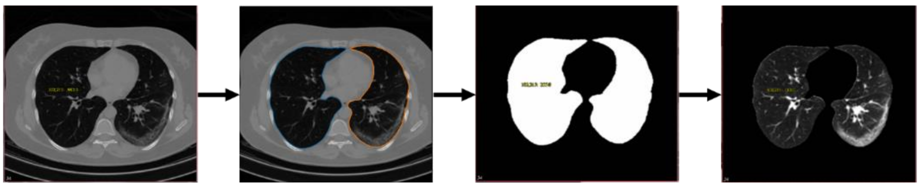

Fig. 6. Procedures of the Lung Segmentation

Mathematical morphological operations are performed on the incomplete lung segmentation as post-processing, as shown in Figure 7.

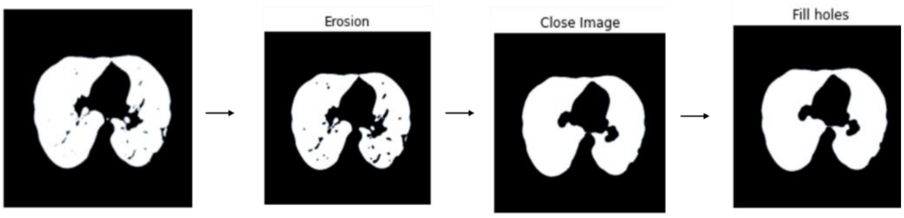

Fig. 7. Morphological Operations on Incomplete Lung Segmentation

3.2 Feature Extraction and Selection

The GLCM features of contrast, energy, dissimilarity, homogeneity, ASM, and correlation are extracted with a combination of distance and angle of the pixel of interest and its neighbouring, as displayed in Table 1.

Table 1
Combination of distance and angle of the pixel of interest and its neighbouring

Combination	$d =$ distance, $\theta =$ angle
[1]	$d = 1, \theta = 0°$
[2]	$d = 1, \theta = 45°$
[3]	$d = 1, \theta = 90°$
[4]	$d = 1, \theta = 135°$
[5]	$d = 2, \theta = 0°$
[6]	$d = 2, \theta = 45°$
[7]	$d = 2, \theta = 90°$
[8]	$d = 2, \theta = 135°$

The results of the GLCM features extracted from each normal lung (N) CT image and COVID-19 (A) CT image are visualised using a boxplot, as shown in Figure 8.

From the boxplots, the distributions of normal lungs seem to be subparts of abnormal lungs for features of contrast, energy, dissimilarity, and homogeneity, as these features do not show obvious differences between the two classes of CT images. Meanwhile, for features of ASM and correlation, the boxplot distributions are distinguishable between the lung with and without COVID-19 infection. The p-values obtained from the t-test for COVID-19 and normal lung CT images for those two features in each of the combinations of distance and angle are less than 0.001, implying that the null hypothesis is rejected. Thus, there is a highly statistically significant difference in ASM and correlation features between the COVID-19 and normal lung CT images.

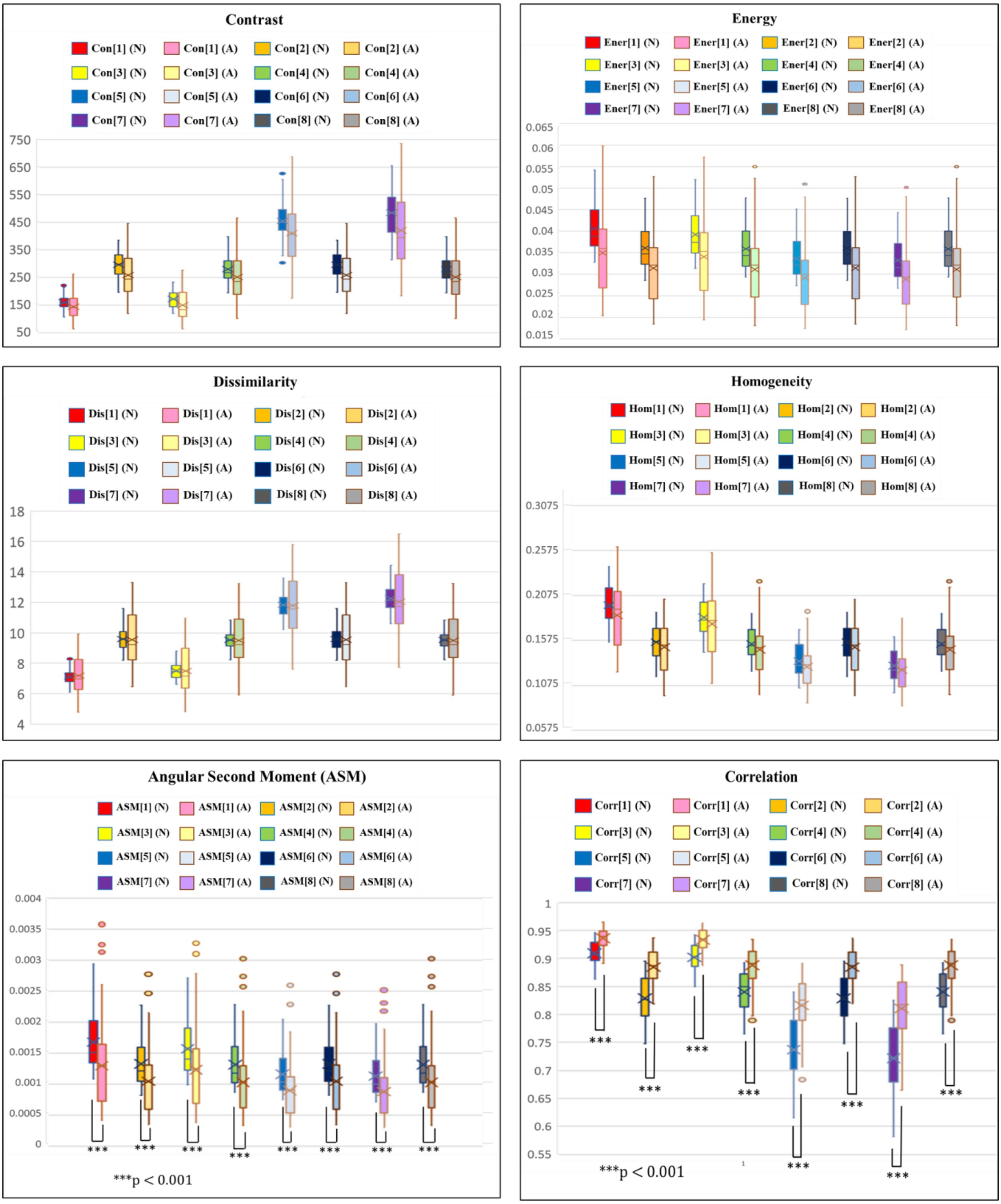

Fig. 8. GLCM Features Extracted for COVID-19 and Normal Cases

3.3 Classification Performance and Evaluation Metrics

The performance of machine learning and deep learning models in terms of accuracy, precision, sensitivity, and specificity are illustrated in Table 2. ResNet-50 and DenseNet-121 outperform all the other models by achieving 100% performance in those metrics, while Random Forest achieves 100%

accuracy for testing data, precision, sensitivity, and specificity of both normal lungs and Covid-19 cases.

Table 2
Comparison of Classification Performance over Machine Learning and Deep Learning Models

Classifier	Accuracy		Precision		Sensitivity		Specificity	
	Testing	Validation	Normal	Covid-19	Normal	Covid-19	Normal	Covid-19
Linear Support Vector Machine	86.36%	76.67%	80.00%	88.00%	67.00%	94.00%	94.00%	67.00%
Random Forest	**100.00%**	88.33%	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**
K-Nearest Neighbour	95.45%	83.33%	86.00%	100.00%	100.00%	94.00%	94.00%	100.00%
Logistic Regression	77.27%	83.33%	67.00%	79.00%	33.00%	94.00%	94.00%	33.00%
Gaussian Naïve Bayes	77.27%	88.33%	56.00%	92.00%	83.00%	75.00%	75.00%	83.00%
ResNet-50	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**
DenseNet-121	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**
Inception	95.00%	100.00%	100.00%	91.67%	88.89%	100.00%	94.12%	95.65%
VGG-16	95.00%	96.67%	100.00%	91.67%	88.89%	100.00%	94.12%	95.65%

3.4 Discussion

The results of this study suggest that deep learning models achieve better performance compared to machine learning models in the classification of COVID-19 CT images. Deep learning models, namely ResNet-50 and DenseNet-121, attain 100% accuracy in testing and data, precision, sensitivity, and specificity. The perfect accuracy is ascribable, most probably to the feature extraction and modelling done automatically after data training. A large number of significant features that can discriminate COVID-19 cases from normal lungs are extracted automatically by deep learning algorithms.

On the other hand, machine learning models require features extracted from the GLCM as input data. However, in this study, GLCM features that can distinguish the two classes of CT images are features of correlation and ASM. With only these two features, the training of the models might not be as strong as deep learning models. Moreover, the parameters used in machine learning models from the Scikit-learn library in Python are the default parameters, indicating that the parameters of machine learning models are not optimal for the classification of Covid-19 CT images.

When comparing machine learning models, the Random Forest model outperforms all the other models in terms of its accuracy for testing and validation data, precision, sensitivity, and specificity. The 10-fold cross-validation has been performed to avoid biased results from testing data by ensuring that all validation data is used for testing. The results show that the Random Forest model has good consistency in the classification task, achieving high accuracy for both testing and validation data. Since the Random Forest model aggregates individual tree judgments for the conclusion, it has a high degree of generalization. Although it is obvious that the deep learning models, ResNet-50 and DenseNet-121 have outstanding performance in the COVID-19 CT image classification, the machine learning model, the Random Forest model continues to hold significance in this classification of Covid-19 CT images. Since the data size used in this study is small, the results obtained are only valid for the dataset used and cannot represent all the other data.

Limitations of this study include applying a limited amount of CT images, and the parameters of machine learning models used in this study are default parameters that are not optimised yet. The performance of machine learning models can be further improved by obtaining the optimised parameters in future research. More datasets can be used in future studies to improve the training of machine learning and deep learning models. Also, mathematical optimisation methods can be applied to determine the optimal parameters of machine learning models.

4. Conclusions

In this paper, artificial intelligence techniques, including machine learning and deep learning algorithms, have been employed to classify COVID-19 CT images. A comparison has been made between deep learning and machine learning classifiers. The result of this study shows that ResNet-50 and DenseNet-121 outperform all the other models by achieving 100% accuracy, precision, sensitivity, and specificity on the dataset applied.

Acknowledgement
This work was supported by the Malaysian Ministry of Higher Education (MOHE) and Universiti Teknologi Malaysia (UTM) under UTM Tier 2 Grant (Cost centre no: R.J130000.2654.17J37).

References

[1] World Health Organization. "Coronavirus disease (COVID-19), 12 October 2020." (2020).
[2] Saygılı, Ahmet. "A new approach for computer-aided detection of coronavirus (COVID-19) from CT and X-ray images using machine learning methods." *Applied Soft Computing* 105 (2021): 107323. https://doi.org/10.1016/j.asoc.2021.107323
[3] Kassania, Sara Hosseinzadeh, Peyman Hosseinzadeh Kassanib, Michal J. Wesolowskic, Kevin A. Schneidera, and Ralph Detersa. "Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach." *Biocybernetics and Biomedical Engineering* 41, no. 3 (2021): 867-879. https://doi.org/10.1016/j.bbe.2021.05.013
[4] Jadhav, Shreeraj, Gaofeng Deng, Marlene Zawin, and Arie E. Kaufman. "COVID-view: Diagnosis of COVID-19 using Chest CT." *IEEE transactions on visualization and computer graphics* 28, no. 1 (2021): 227-237. https://doi.org/10.1109/TVCG.2021.3114851
[5] Benmalek, Elmehdi, Jamal Elmhamdi, and Abdelilah Jilbab. "Comparing CT scan and chest X-ray imaging for COVID-19 diagnosis." *Biomedical Engineering Advances* 1 (2021): 100003. https://doi.org/10.1016/j.bea.2021.100003
[6] DenOtter, Tami D., and Johanna Schubert. "Hounsfield unit." (2019).
[7] Jaglan, Poonam, Rajeshwar Dass, and Manoj Duhan. "A comparative analysis of various image segmentation techniques." In *Proceedings of 2nd International Conference on Communication, Computing and Networking: ICCCN 2018, NITTTR Chandigarh, India*, pp. 359-374. Springer Singapore, 2019. https://doi.org/10.1007/978-981-13-1217-5_36
[8] Arunkumar, N., Mazin Abed Mohammed, Mohd Khanapi Abd Ghani, Dheyaa Ahmed Ibrahim, Enas Abdulhay, Gustavo Ramirez-Gonzalez, and Victor Hugo C. de Albuquerque. "K-means clustering and neural network for object detecting and identifying abnormality of brain tumor." *Soft Computing* 23 (2019): 9083-9096. https://doi.org/10.1007/s00500-018-3618-7
[9] Chen, Xu, Bryan M. Williams, Srinivasa R. Vallabhaneni, Gabriela Czanner, Rachel Williams, and Yalin Zheng. "Learning active contour models for medical image segmentation." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11632-11640. 2019. https://doi.org/10.1109/CVPR.2019.01190
[10] Chowdhary, Chiranji Lal, and D. Prasanna Acharjya. "Segmentation and feature extraction in medical imaging: a systematic review." *Procedia Computer Science* 167 (2020): 26-36. https://doi.org/10.1016/j.procs.2020.03.179
[11] Sharma, Neeraj, and Lalit M. Aggarwal. "Automated medical image segmentation techniques." *Journal of medical physics/Association of Medical Physicists of India* 35, no. 1 (2010): 3. https://doi.org/10.4103/0971-6203.58777
[12] Mittal, Mamta, Amit Verma, Iqbaldeep Kaur, Bhavneet Kaur, Meenakshi Sharma, Lalit Mohan Goyal, Sudipta Roy, and Tai-Hoon Kim. "An efficient edge detection approach to provide better edge connectivity for image analysis." *IEEE access* 7 (2019): 33240-33255. https://doi.org/10.1109/ACCESS.2019.2902579
[13] Gong, Sui, and Timothy S. Newman. "Dual marching squares: description and analysis." In *2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pp. 53-56. IEEE, 2016. https://doi.org/10.1109/SSIAI.2016.7459173

[14] Corke, Peter, and Peter Corke. "Image Feature Extraction." *Robotics, Vision and Control: Fundamental Algorithms In MATLAB® Second, Completely Revised, Extended And Updated Edition* (2017): 413-458. https://doi.org/10.1007/978-3-319-54413-7_13
[15] Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. "Textural features for image classification." *IEEE Transactions on systems, man, and cybernetics* 6 (1973): 610-621. https://doi.org/10.1109/TSMC.1973.4309314
[16] Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar. "Image texture feature extraction using GLCM approach." *International journal of scientific and research publications* 3, no. 5 (2013): 1-5.
[17] Lai, ZhiFei, and HuiFang Deng. "Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron." *Computational intelligence and neuroscience* 2018 (2018). https://doi.org/10.1155/2018/2061516
[18] Kotadiya, Hiral, and Darshana Patel. "Review of medical image classification techniques." In *Third International Congress on Information and Communication Technology: ICICT 2018, London*, pp. 361-369. Springer Singapore, 2019. https://doi.org/10.1007/978-981-13-1165-9_33
[19] Misra, Siddharth, Hao Li, and J. He. "Noninvasive fracture characterization based on the classification of sonic wave travel times." *Machine learning for subsurface characterization* 4 (2020): 243-287. https://doi.org/10.1016/B978-0-12-817736-5.00009-0
[20] Nazish, Syed Irfan Ullah, Abdus Salam, Wajid Ullah, and Muhammad Imad. "COVID-19 lung image classification based on logistic regression and support vector machine." In *European, Asian, Middle Eastern, North African Conference on Management & Information Systems*, pp. 13-23. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-77246-8_2
[21] Yoo, Seung Hoon, Hui Geng, Tin Lok Chiu, Siu Ki Yu, Dae Chul Cho, Jin Heo, Min Sung Choi *et al.,* "Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging." *Frontiers in medicine* 7 (2020): 427. https://doi.org/10.3389/fmed.2020.00427
[22] Lahsaini, Ilyas, Mostafa El Habib Daho, and Mohamed Amine Chikh. "Deep transfer learning based classification model for covid-19 using chest CT-scans." *Pattern recognition letters* 152 (2021): 122-128. https://doi.org/10.1016/j.patrec.2021.08.035
[23] Wang, Zhao, Quande Liu, and Qi Dou. "Contrastive cross-site learning with redesigned net for COVID-19 CT classification." *IEEE Journal of Biomedical and Health Informatics* 24, no. 10 (2020): 2806-2813. https://doi.org/10.1109/JBHI.2020.3023246
[24] Tabik, Siham, Anabel Gómez-Ríos, José Luis Martín-Rodríguez, Iván Sevillano-García, Manuel Rey-Area, David Charte, Emilio Guirado *et al.,* "COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on chest X-ray images." *IEEE journal of biomedical and health informatics* 24, no. 12 (2020): 3595-3605. https://doi.org/10.1109/JBHI.2020.3037127
[25] Ibrahim, Dina M., Nada M. Elshennawy, and Amany M. Sarhan. "Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases." *Computers in biology and medicine* 132 (2021): 104348. https://doi.org/10.1016/j.compbiomed.2021.104348
[26] Arslan, Sevtap, and Emre Ünal. "One of the many faces of COVID-19 infection: an irregularly shaped pulmonary nodule." *Insights Into Imaging* 12, no. 1 (2021): 1-2. https://doi.org/10.1186/s13244-021-00987-7
[27] Wang, Shuai, Bo Kang, Jinlu Ma, Xianjun Zeng, Mingming Xiao, Jia Guo, Mengjiao Cai *et al.,* "A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)." *European radiology* 31 (2021): 6096-6104. https://doi.org/10.1007/s00330-021-07715-1
[28] Afshar, Parnian, Shahin Heidarian, Nastaran Enshaei, Farnoosh Naderkhani, Moezedin Javad Rafiee, Anastasia Oikonomou, Faranak Babaki Fard, Kaveh Samimi, Konstantinos N. Plataniotis, and Arash Mohammadi. "COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning." *Scientific Data* 8, no. 1 (2021): 121. https://doi.org/10.1038/s41597-021-00900-3
[29] Razali, Nornadiah Mohd, and Yap Bee Wah. "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests." *Journal of statistical modeling and analytics* 2, no. 1 (2011): 21-33.
[30] Stephens, Michael A. "The anderson-darling statistic." *No. TR-39). Stanford Univ Ca Dept. of Statistics* (1979).
[31] Islam, Mohammad Rafiqul. "Sample size and its role in Central Limit Theorem (CLT)." *Computational and Applied Mathematics Journal* 4, no. 1 (2018): 1-7. https://doi.org/10.31295/pm.v1n1.42
[32] Lascu, Mihaela-Ruxandra. "Deep learning in classification of Covid-19 coronavirus, pneumonia and healthy lungs on CXR and CT images." *Journal of Medical and Biological Engineering* 41, no. 4 (2021): 514-522. https://doi.org/10.1007/s40846-021-00630-2