International Journal of Advances in Intelligent Informatics
Vol. 7, No. 2, July 2021, pp. 151-162
ISSN 2442-6571
151

Lung cancer medical images classification using hybrid CNN-SVM

Abdulrazak Yahya Saleh [a,1,*], Chee Ka Chin [b,2], Vanessa Penshie [a,3], Hamada Rasheed Hassan Al-Absi [c,4]

[a] Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak, Sarawak, Malaysia
[b] Faculty of Engineering, Universiti Malaysia Sarawak, Sarawak, Malaysia
[c] College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar
[1] ysahabdulrazak@unimas.my; [2] cheekachin@yahoo.com.my; [3] vanessapenshie@gmail.com, [4] hrha@outlook.com
* corresponding author

ARTICLE INFO

Article history
Received January 7, 2020
Revised May 7, 2021
Accepted May 7, 2021
Available online May 30, 2021

Keywords
Lung Cancer
Classification
Convolutional Neural Network
SVM
Computer-aided detection (CAD)

ABSTRACT

Lung cancer is one of the leading causes of death worldwide. Early detection of this disease increases the chances of survival. Computer-Aided Detection (CAD) has been used to process CT images of the lung to determine whether an image has traces of cancer. This paper presents an image classification method based on the hybrid Convolutional Neural Network (CNN) algorithm and Support Vector Machine (SVM). This algorithm is capable of automatically classifying and analyzing each lung image to check if there is any presence of cancer cells or not. CNN is easier to train and has fewer parameters compared to a fully connected network with the same number of hidden units. Moreover, SVM has been utilized to eliminate useless information that affects accuracy negatively. In recent years, Convolutional Neural Networks (CNNs) have achieved excellent performance in many computer visions tasks. In this study, the performance of this algorithm is evaluated, and the results indicated that our proposed CNN-SVM algorithm has been succeed in classifying lung images with 97.91% accuracy. This has shown the method's merit and its ability to classify lung cancer in CT images accurately.

1. Introduction

Lung cancer is one of the leading causes of death in Malaysia and around the world [1]. In 2018, lung cancer had the highest percentage of new cases and deaths worldwide, by 11.6% [2]. According to the statistics by World Health Organization, lung cancer has the second-highest rate of cases in Malaysia in 2018 after breast cancer. About 16.6% of the patients are male, while 5.4% are female. Based on The Malaysia National Cancer Registry Report 2012-2016, lung cancer stands as third-highest cancer detected among Malaysians, affecting males more than females [3]. It begins in the lungs, where cancer cells then may spread to lymph nodes or other organs in the body, such as the brain [4]. This disease is categorized into two main types called small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). These two types of cancer are treated differently. The common type which affects about 80%-85% of lung cancer patients is non-small cell cancer [5].

A small round or oval-shaped growth in lung sections is known as a pulmonary nodule. It is also called a 'coin lesion', which is usually larger than three centimeters in diameter [6]. Any nodule smaller than three centimeters is known as micro-nodules. Non-nodules such as blood vessels and bronchi walls have the same appearances as nodules and can cause false positives during the detection process [7].

Pulmonary nodules can be detected on CT scan images or X-ray films. By identifying and classifying malignant from benign nodules, the malignant nodules can be removed at an early stage to prevent further spreading of tumor cells which leads to fatality. Lung cancer classification is significant because most people diagnosed with this disease are within the age range of 55-65. Early diagnosis and treatment are crucial for their survival. An early diagnosis for lung cancer detection is by lung radiology test (lung image), called Chest X-Ray test. Lung image will show different results between normal lung and abnormal lung. The existence of nodules in the lung image indicates that the lung is not normal [8]. Fig.1 displays a standard lung image and a denoised lung image from level 1 to 10 using Haar Wavelet.

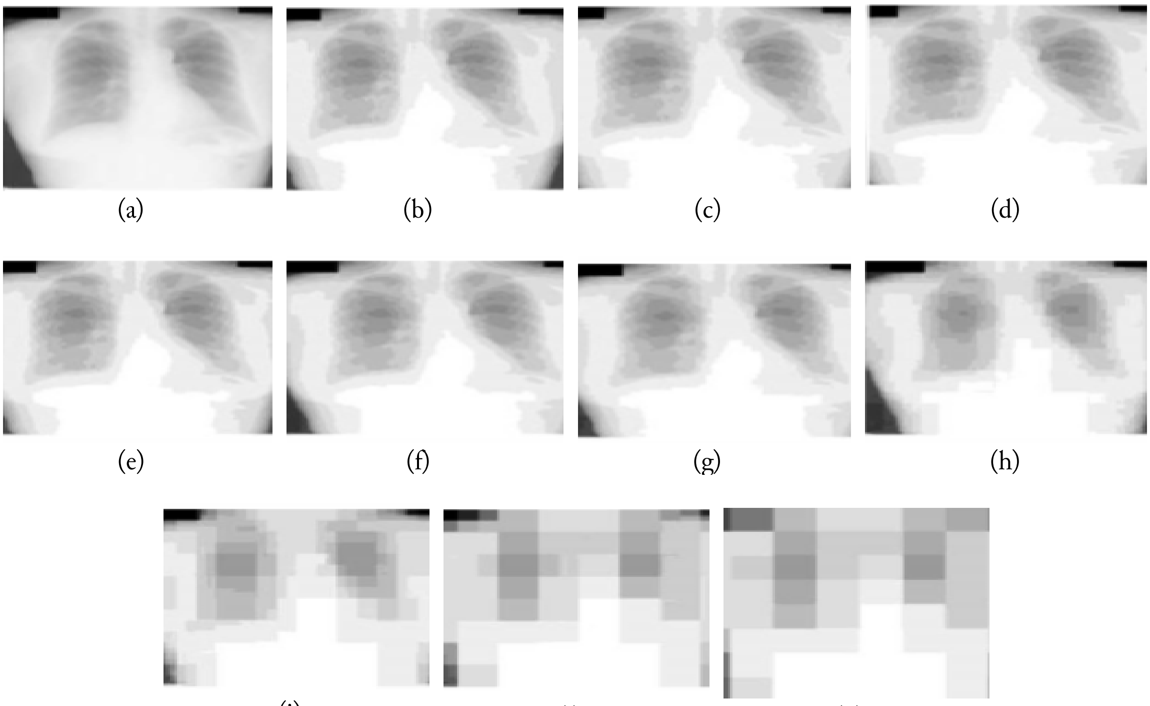

Fig. 1. (a) Normal Lung Image, (b) Denoised Lung Image with Haar Wavelet Levell, (c) Denoised Lung Image with Haar Level 2, (d) Denoised Lung Image with Haar Level 3, (e) Denoised Lung Image with Haar Level 4, (f) Denoised Lung Image with Haar Level 5, (g) Denoised Lung Image with Haar Level 6, (h) Denoised Lung Image with Haar Level 7, (i) Denoised Lung Image with Haar Level 8, (j) Denoised Lung Image with Haar Level 9, (k) Denoised Lung Image with Haar Level 10 [8].

Detection of lung cancer in its early stages could reduce the fatality rate [9][10]. Early detection of cancerous cells is crucial to prevent them from growing and remove them early before they start to grow rapidly. Also, accurate location and size detection of benign cells is another crucial point to be taken as in CT scans. It is difficult to differentiate cancerous nodules from other pulmonary nodules that are not cancerous by only depending on visual estimation, which is a traditional way for healthcare workers to examine CT scans. Computer-Aided Detection (CAD) systems have been utilized to process medical images such as CT scan images to assist the medical personnel in the early detection of lung cancer [11]–[14]. It has been reported that CAD systems have an enormous positive impact on the job of radiologists in detecting cancer, and the systems could be used as a second reader [15]. CAD systems are based on computer algorithms that process images such as CT scan images of the lung to determine whether a given image has any sign of cancer; this process is called image classification. Several machine learning algorithms have been proposed to classify medical images and determine the existence of lung cancer, such as those of [16]–[18]. Singh & Gupta [19] have researched various machine learning approaches for the classification and detection of lung cancers using CT scan images. Different classifiers are put to the test to determine the best machine learning algorithm from the classification process. The results show that multilayer perceptron or neural networks can be applied to classify lung cancer CT scan images because it obtained a high percentage of accuracy precision value. Over the years, deep learning has been used increasingly in various fields, and research on deep learning are growing rapidly. Various techniques are proven to give accurate and precise results, which help humans detect errors and mistakes at early stages and help to predict the outcomes.

Machine and deep learning methods are used widely in the medical and healthcare field for monitoring, detecting, classifying, and predicting diseases [20]–[24]. For instance, an established Artificial Intelligent system for detecting and classifying hemoglobin to monitor blood loss, the Triton™ system by Gauss Surgical™, a medical technology company, has been used in a few hospitals such as at Mount Sinai Hospital to monitor blood loss and prevent hemorrhage by early detection [25]. Various studies proved that computational methods applied in Triton™ have high prediction accuracy towards early detection of hemorrhage [26]. The results produced by deep learning are convincing and deep learning should have focused more on helping in the medical sector. Convolutional Neural Networks (CNNs) have recently become a popular diagnostic tool due to the high performance in segmentation, pattern detection, recognition, and classification [27]. Moreover, CNN is known to have the best performance when using large amounts of data lacking in medical imaging due to several factors such as ethics and lack of well-labeled data. Many researchers have tried to overcome this problem by proposing advanced off-the-shelf CNNs, advanced implementation techniques, and many more to propose an algorithm that leverage performance even with small dataset [7].

Deep learning is utilized in this study for lung cancer classification due to its merit as a popular and powerful method of pattern recognition and classification [28][29]. Deep learning, specifically CNN, is known to have a high success rate if only a huge amount of data is implied. A large number of well-labeled training data such as ImageNet is required for CNN to have a good performance which medical images lack. Large datasets are not always available due to several factors such as the costly expert explanatory notes, ethical issues, and shortage of disease images [30]. The model with a large number of parameters will fail to learn the patterns if supplied with small datasets and can easily cause overfitting [31]. Most traditional CNN architectures' performance depends heavily on the size of data. They initially have numerous parameters and state of art CNN models trained with large datasets such as ImageNet is unsuitable for datasets with hundreds or thousands of instances [32]. This poses a problem that researchers must address in order to improve model output despite having a large amount of labelled data. As a result, conventional CNN is unsuitable for medical imaging processes involving small datasets (hundreds to thousands number of data). It is important to investigate a new algorithm in order to obtain a more accurate result [17][33][34].

Although many deep learning-based algorithms for lung cancer have been proposed [35]–[40], there is much room for their accuracy to be improved. Therefore, a CNN-SVM architecture has been proposed in this paper, which removes and eliminates useless information that negatively impacts accuracy. This is accomplished in the CNN architecture's pooling step, which is used to classify lung cancer in the fully connected layer using a modified SVM architecture.

2. Method

Deep learning is no longer a stranger to the medical imaging analysis field. It is a growing trend, and the demand for deep learning in achieving accurate and precise results is growing [41]. Deep learning involves imitating how the human brain works in dealing with data and recognizing patterns for decision-making [42]–[45]. With the emerging of technology and better algorithms, more machinery gives high accuracy and reliability for medical analysis. Detection of cancerous or malignant cells is crucial for the treatment of lung cancer. Application of image analysis deep learning on computed tomography (CT) scan images helps to detect malignant cells early before they develop and become lethal [5]. Deep learning has been proven to have significant performance in image processing, especially in object detection and localization [19]. Deep learning uses multiple layers to represent the abstractions of data to build computational models [46]. A deep learning algorithm with Convolutional Neural Networks

(CNNs) is applied to a dataset to perform a classification task. Fig. 2 shows an illustration of a Convolutional Neural Network in conjunction with a support vector machine.

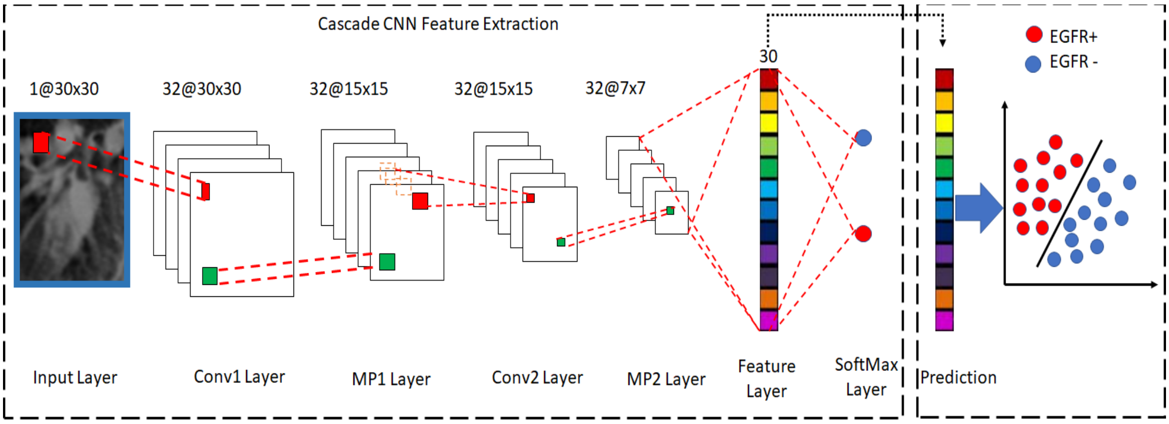

Fig. 2. Illustration of a Convolutional Neural Network in conjunction with a support vector machine (Adopted from [47]).

In Fig. 2, Yu *et al.* [47] proposed a system for predicting epidermal growth factor receptor (EGFR) mutations that have been known to be associated with lung cancer. In the system, six layers of CNN are constructed to learn deep image features, and then the Support Vector Machine (SVM) classifier is used for prediction. This system was tested on two datasets, and it achieved accuracy levels of 76.16 % and 67.55% on Dataset1 and Dataset2, respectively. By learning from the deep-layered and hierarchical data models, the deep learning algorithms can outperform the traditional machine learning models [33]. With deep learning, the data are processed easily. According to [48], deep learning has been found to identify the underlying structure of data using auto-encoders and other techniques. A CNN is the most popular choice when it comes to image classification. This is because CNN architecture implicitly combines the benefits of the standard neural network training with the convolution operation to classify images efficiently. da Silva *et al.* [33] tested the functionality of the deep learning algorithms for lung cancer diagnosis with cases from Lung Image Database Consortium (LIDC) database. The results showed the promising performance of deep learning algorithms, and the accuracy achieved was an impressive rate of 79.40%. The researchers intended to test the deeper structured schemes for lung cancer diagnosis and seek more efficient ways to minimize the downsample effect. They compared their method with Deep Belief Networks (DBNs) and Stacked Denoising Autoencoder (SDAE) using the same dataset. They achieved accuracies of 81.19% and 79.29% on DBN and SDAE, respectively. The comparison results demonstrated the great potential of deep structured algorithms and computer learned features in the domain of medical imaging. In the article by Ali *et al.* [49], the authors reported developing a reinforcement learning model based on deep artificial neural networks for the early detection of lung cancer. The authors used the LUNA dataset, a subset of the LIDC/IDRI and achieved an accuracy of 99.1% on the training set and 64.4% on the testing set. Another system was proposed by [50], where the authors developed a CAD system for lung nodule classification based on multi-view multi-scale CNN. The system was tested on both the LIDC/IDRI and ELCAP datasets and achieved 92.3% and 90.3% accuracy on both datasets, respectively. The aforementioned studies show that CNN performance in many instances could be improved to achieve higher accuracy and assist in the early detection of lung cancer with minimum errors.

2.1. Dataset

This study utilized the CT images from the Chest CT-Scan Images Dataset [16]. This dataset consists of CT images of normal lung images and lung cancer images. Since the data was cleaned and of high resolution, the database was perfect for deep learning. We created a subset of the dataset with 1000 images comprising cancer images (adenocarcinoma, large cell carcinoma, squamous cell carcinoma) and normal images. Examples of the images are shown in Fig. 3 and Fig. 4. Each training, testing, and validation set has four classes of images (adenocarcinoma, large cell carcinoma, normal and squamous cell carcinoma). The data was separated into training, testing, and validation sets to evaluate the classification model.

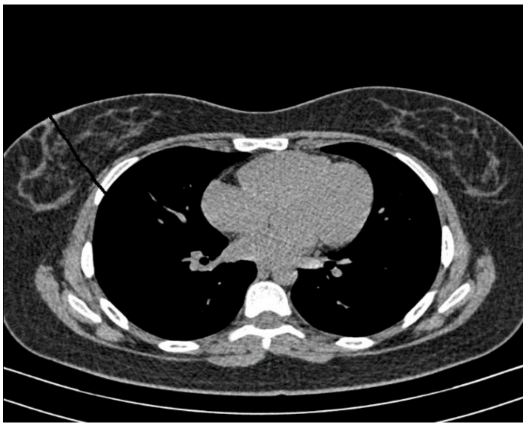

Fig. 3. Lung image with no cancer.

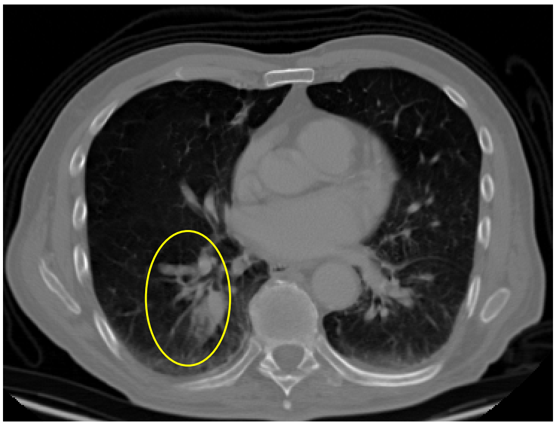

Fig. 4. Lung image with cancer (large cell carcinoma) – yellow circle.

2.2. Data Augmentation

Data augmentation methods were added to the workflow of these experiments to enlarge the data set and avoid overfitting issues [51]–[53]. This technique aims to increase the number of deep neural model training datasets, balance the size of the datasets, and boost their efficiency, but it is still the subject of research [54]. The dataset used in this paper has been increased from 1000 images to 5103 images via the color transformation technique by intensity adjustment. There was no specific requirement for deep learning, but the training data set needs to be high-quality in medical imaging [55]. Based on [56], at least 500 to more than 1000 images per class is good enough for classification task in Deep Neural Network (DNN). In this experiment, 20% of the total dataset was utilized as a testing set, and the remaining data was divided into the train-validation split of 80-20.

2.3. Convolution Neural Network (CNN)

The deep learning algorithm used in this study is Convolutional Neural Network. A convolutional neural network (CNN) is a type of multilayer feedforward biologically influenced neural network [57]–[60]. A CNN has several layers, which can be classified into three types: convolutional (compute the output of the connected local input region neurons), max-pooling (sub-sampling the inputs), and fully connected layers (used in computing each class' activation). The input to a CNN is an n x n x m image, where n is the height and width and m is the number of channels, and there will be k convolutional filters of size a x a in the convolutional layer, where a < n. The CNN is built and then fitted into the dataset. In CT scan images, a slice has a width and height of 512x512 which are multiplied by the number of images, in our case 5103 images. In this study, the CNN model consists of 4 steps: 1) the convolution layer; 2) the pooling layer; 3) the flattening layer; and 4) the full connection layer. Fig. 5 shows the architecture of the model.

In the Convolution Layer, feature maps are created by multiplying the input images with feature detectors. Many feature maps are created, and the size of the image becomes smaller for easier processing. Due to the use of several feature detectors, different feature maps are created. Within this step, the ReLU activation function is used to increase the non-linearity in the CNN (since images are generally non-linear). In the Pooling layer, the feature maps' resolution is reduced; in this step, the max-pooling is applied where a window function is applied to the input feature map and computes the maximum in the neighborhood. In this step, a large amount of useless information is eliminated, and this has a great impact on producing better results since unrelated/useless information will not be fed into the fully connected layer. In the flattening step, the pooled feature map is flattened into a column. This is done to further process it in the Support Vector Machine (SVM). Lastly, all features are processed in a fully connected layer that consists of SVM, which produces one of four options: adenocarcinoma, large cell carcinoma, normal or squamous cell carcinoma. In this study, the above architecture was implemented with Python 3.7 using Jupyter Notebook, which was installed in Intel Core i5 10th Gen processor with 16 GB of memory and NVIDIA GeForce RTX 2060 GPU support with 6 GB of memory.

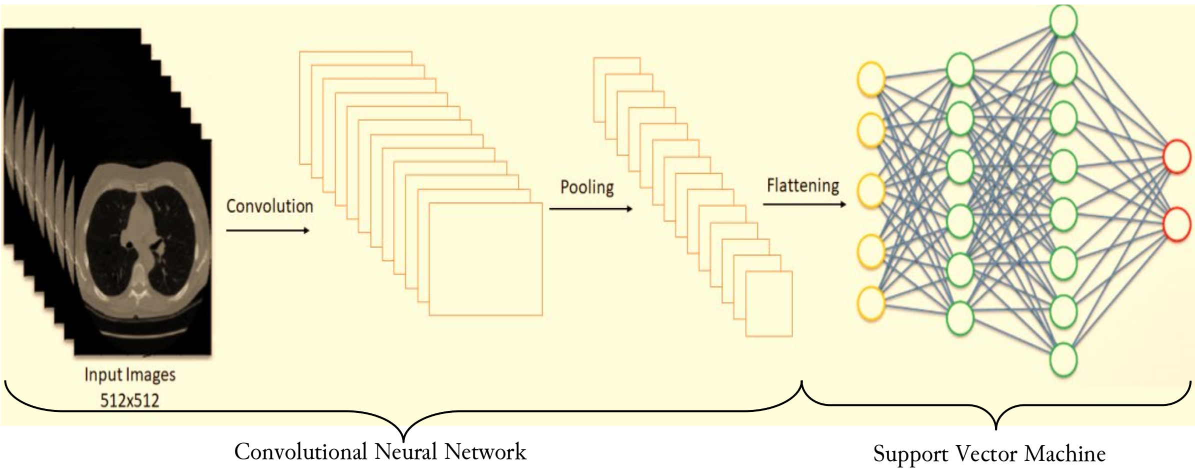

Fig. 5. A flow chart of the architecture of the CNN-SVM model.

2.4. Model Training

In the model training, Adam was used as a neural work optimizer and categorical cross entropy was used as a loss function. The important parameters such as batch size, epoch values, and learning rate were the same as those utilized in the training of the proposed CNN-SVM model. These values were 64 for batch size, 60 epoch values with 2000 steps per epoch, and a learning rate of 0.0001.

3. Result and Discussion

The last part is evaluating whether the proposed method has succeeded in enhancing the accuracy of prediction. Accuracy refers to the closeness of a measured value to a standard or known value. The accuracy of the result is very important in determining the best algorithm to use in the future. The higher the accuracy, the more excellent the results of the research are. The accuracy can be calculated based on (1).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

where TP refers to True Positives, TN refers to True Negative, FP refers to the False Positive, and FN refers to False Negative. The experiment runs 60 times using the datasets for training and validation. At the beginning of the training, the accuracy was low because the training was executed only a few times. After the number of epochs increased, the accuracy improved. Table 1 shows the final accuracy obtained in the Chest CT-Scan images dataset.

Table 1. Results of the dataset at the beginning and at the end of the training and testing.

Epoch No	Loss	Accuracy	Validation loss	Validation accuracy
1	1.2785	0.3582	1.5085	0.3635
			....	
60	0.0992	0.9671	0.0732	0.9791

As shown in Table 1, the performance of the CNN-SVM method reached a high rate of 96.71% on the training set and 97.91% on the validation set. In the reported result, there are 4 components: accuracy, validation accuracy, loss, and validation loss. Fig. 6 and Fig. 7 show training loss versus validation loss and training accuracy versus validation accuracy, respectively. As shown in Fig. 7, the training accuracy increases as the number of epochs increases. This means that the higher the number of epochs, the more the CNN-SVM algorithm adapts to the dataset in classifying the images.

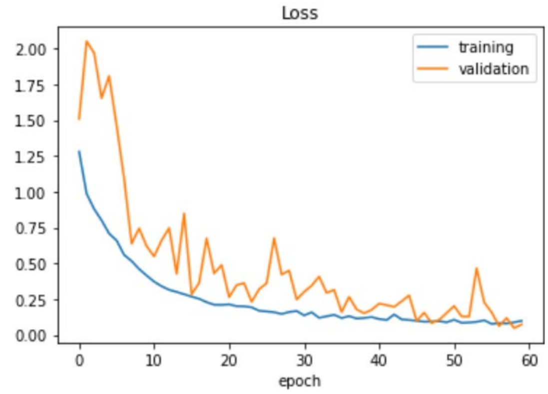

Fig. 6. The training loss and validation loss.

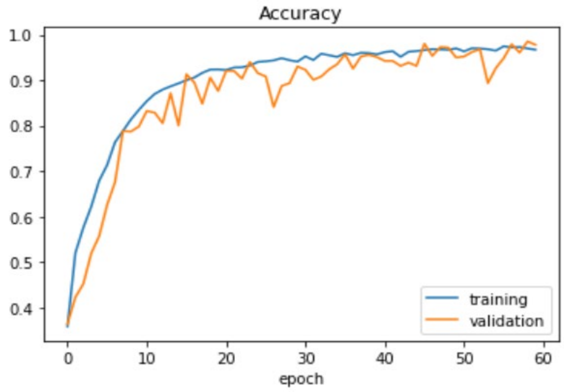

Fig. 7. The training accuracy and validation accuracy.

Consequently, using the confusion matrix (shown in Fig. 8), the sensitivity, specificity and precision were computed. These performance measures are determined using (2) to (4).

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} . 100\% \quad (2)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} . 100\% \quad (3)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} . 100\% \quad (4)$$

Fig. 8 and Fig. 9 show the confusion matrix and Receiver Operating Characteristics (ROC) curve performance, respectively. Table 2 illustrates the performance measure in this experiment.

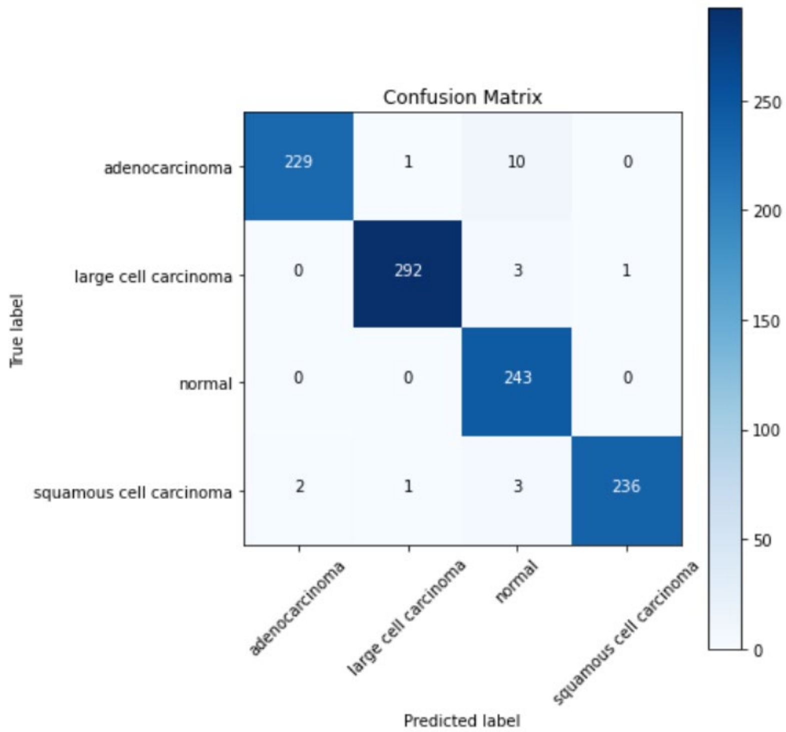

Fig. 8. Confusion Matrix Performance.

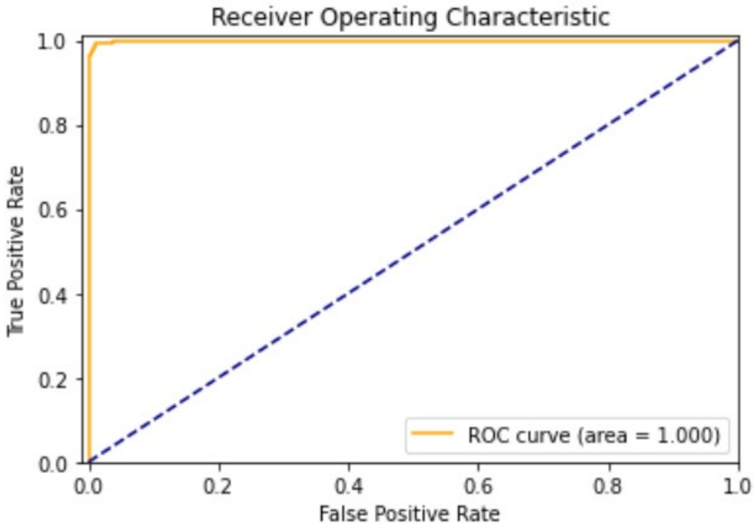

Fig. 9. ROC Curve Performance.

Table 2. The performance measure in this experiment.

Type of Lung Cancer	Sensitivity	Specificity	Precision
Adenocarcinoma	95.42%	99.74%	99.13%
Large Cell Carcinoma	98.65%	99.72%	99.32%
Normal	100.00%	97.93%	93.82%
Squamous Cell Carcinoma	97.52%	99.87%	99.58%
Average	**97.90%**	**99.32%**	**97.96%**

Table 3 compares the implementation of the CNN and other recent works that are also based on deep learning. As shown in Table 3, our work achieves better accuracy if compared with most of the other recent works (which have been reported in this paper) except that of [49], which achieved 99.1% using the training set. However, the accuracy of the testing set is 64.4%, which is low compared to the test accuracy of 97.91% obtained from our work. In addition, the proposed work also shows promising performance in terms of sensitivity, specificity, precision, and AUC when compared with existing works. On top of that, the number of epochs used in [49] was 120 compared with 60 epochs in our training.

Table 3. Comparison with other published works.

Published work	Dataset	Accuracy	Sensitivity	Specificity	Precision	AUC
CNN+SVM [16]	Private	76.16%	73.80%	78.24%	-	0.828
CNN [14]	LIDC	79.40%.	-	-	-	-
DBNs [14]	LIDC	81.19%	-	-	-	-
SDAE [14]	LIDC	79.29%	-	-	-	-
Reinforcement learning deep ANN [17]	LUNA	Training: **99.1%** Testing: 64.4%	58.90%	55.30%	54.26%	-
Multi-view multi-scale CNN [18]	LIDC/IDRI	92.3%	-	-	-	-
	ELCAP	90.3%	-	-	-	-
Proposed CNN-SVM	Chest CT-Scan images Dataset	Training: 96.71% Testing: **97.91%**	**97.90%**	**99.32%**	**97.96%**	**1.000**

4. Conclusion

This study introduces a hybrid CNN-SVM method to classify lung CT images into adenocarcinoma, large cell carcinoma, normal or squamous cell carcinoma. The aim was to achieve a higher level of accuracy, which is the goal of any computer-aided detection system. The method was applied to the Chest CT-Scan images dataset, a standard and publicly available cluster of CT images. A total of 5103 images were used to test the method, and a classification accuracy rate of 97.91% was achieved. It shows the superiority of the proposed hybrid CNN-SVM and its capabilities in the applications. Furthermore, the proposed CNN-SVM method also provides the sensitivity, specificity, precision, and AUC of 97.90%, 99.32%, 97.96%, and 1.000, respectively. For future work, the researchers aim to test the method with different datasets. The level of accuracy can be further improved by increasing the number of images utilized for the procedure. Additionally, X-ray, X-beam, and PET images can be interpreted by utilizing this method. An examination should be possible for all these images. By studying and analyzing the prediction results of different types of images, the medical personnel will be able to use the most suitable images to discover lung malignancy.

Acknowledgment

The authors acknowledge the RIEC of University Malaysia Sarawak (UNIMAS) Malaysia that supported and funded the research under the Cross-Disciplinary Research (F04/CDRG/1839/2019).

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. This research is supported and funded by RIEC of University Malaysia Sarawak (UNIMAS), under the Cross-Disciplinary Research (F04/CDRG/1839/2019).
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] M. Kan Chan Siang and C. K. M. John, "A review of lung cancer research in Malaysia," *Med J Malaysia,* vol. 71, p. 71, 2016. Available at: Google Scholar.

[2] I. A. f. R. o. C. WHO, "New Global Cancer Data: GLOBOCAN 2018. World Health Organization," 2018. Available at: https://www.uicc.org/

[3] M. o. H. Malaysia., "Malaysia National Cancer Registry Report (MNCR) 2012-2016," ed. Putrajaya, 2019. Available at: Google Scholar.

[4] C. f. D. C. a. Prevention, "What is Lung Cancer?," 22 September 2020. Available at: https://www.cdc.gov/

[5] A. C. Society, "About Lung Cancer," 1 October 2019. [Online]. Available at: https://www.cancer.org/.

[6] E. J. Olson, "Lung nodules: Can they be cancerous?," 2020. Available at: https://www.mayoclinic.org/

[7] P. Monkam, S. Qi, H. Ma, W. Gao, Y. Yao, and W. Qian, "Detection and classification of pulmonary nodules using convolutional neural networks: a survey," *IEEE Access,* vol. 7, pp. 78075-78091, 2019. doi: 10.1109/ACCESS.2019.2920980

[8] D. Nurtiyasari and D. Rosadi, "The application of Wavelet Recurrent Neural Network for lung cancer classification," in *2017 3rd International Conference on Science and Technology-Computer (ICST)*, 2017: IEEE, pp. 127-130. doi: 10.1109/ICSTC.2017.8011865

[9] F. M. Sullivan *et al.*, "Earlier diagnosis of lung cancer in a randomised trial of an autoantibody blood test followed by imaging," *European Respiratory Journal,* vol. 57, no. 1, 2021. Available at: Google Scholar.

[10] H. Shin *et al.*, "Early-stage lung cancer diagnosis by deep learning-based spectroscopic analysis of circulating exosomes," *ACS nano,* vol. 14, no. 5, pp. 5435-5444, 2020. doi: 10.1021/acsnano.9b09119

[11] H. Fujita, "AI-based computer-aided diagnosis (AI-CAD): the latest review to read first," *Radiological physics and technology,* vol. 13, no. 1, pp. 6-19, 2020. doi: 10.1007/s12194-019-00552-4

[12] A. Saygılı, "A new approach for computer-aided detection of coronavirus (COVID-19) from CT and X-ray images using machine learning methods," *Applied Soft Computing,* vol. 105, p. 107323, 2021. doi: 10.1016/j.asoc.2021.107323

[13] W. K. Moon *et al.*, "Computer-aided tumor detection in automated breast ultrasound using a 3-D convolutional neural network," *Computer methods and programs in biomedicine,* vol. 190, p. 105360, 2020. doi: 10.1016/j.cmpb.2020.105360

[14] X.-F. Cao, Y. Li, H.-N. Xin, H.-R. Zhang, M. Pai, and L. Gao, "Application of artificial intelligence in digital chest radiography reading for pulmonary tuberculosis screening," *Chronic Diseases and Translational Medicine,* vol. 7, no. 1, pp. 35-40, 2021. doi: 10.1016/j.cdtm.2021.02.001

[15] M. Liang *et al.*, "Low-dose CT screening for lung cancer: computer-aided detection of missed lung cancers," *Radiology,* vol. 281, no. 1, pp. 279-288, 2016. doi: 10.1148/radiol.2016150063

[16] M. Hany, "Chest CT-Scan images Dataset CT-Scan images with different types of chest cancer," 2020. Available at: https://www.kaggle.com/

[17] H. Wang *et al.*, "Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18 F-FDG PET/CT images," *EJNMMI research,* vol. 7, no. 1, pp. 1-11, 2017. doi: 10.1186/s13550-017-0260-9

[18] K. Liu and G. Kang, "Multiview convolutional neural networks for lung nodule classification," *International Journal of Imaging Systems and Technology,* vol. 27, no. 1, pp. 12-22, 2017. doi: 10.1002/ima.22206

[19] G. A. P. Singh and P. Gupta, "Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans," *Neural Computing and Applications,* vol. 31, no. 10, pp. 6863-6877, 2019. doi: 10.1007/s00521-018-3518-x

[20] F. Ali *et al.*, "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion," *Information Fusion,* vol. 63, pp. 208-222, 2020. doi: 10.1016/j.inffus.2020.06.008

[21] C. Bhatt, I. Kumar, V. Vijayakumar, K. U. Singh, and A. Kumar, "The state of the art of deep learning models in medical science and their challenges," *Multimedia Systems,* pp. 1-15, 2020. Available at: Google Scholar.

[22] K. Muhammad, S. Khan, J. Del Ser, and V. H. C. de Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," *IEEE Transactions on Neural Networks and Learning Systems,* 2020. doi: 10.1109/TNNLS.2020.2995800

[23] S. Mittal and Y. Hasija, "Applications of deep learning in healthcare and biomedicine," in *Deep Learning Techniques for Biomedical and Health Informatics*: Springer, 2020, pp. 57-77. doi: 10.1007/978-3-030-33966-1_4

[24] M. Masud, N. Sikder, A. A. Nahid, A. K. Bairagi, and M. A. AlZain, "A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework," *Sensors,* vol. 21, no. 3, p. 748, 2021. doi: 10.3390/s21030748

[25] D. Katz *et al.*, "The association between the introduction of quantitative assessment of postpartum blood loss and institutional changes in clinical practice: an observational study," *International journal of obstetric anesthesia,* vol. 42, pp. 4-10, 2020. doi: 10.1016/j.ijoa.2019.05.006

[26] N. Bernal, J. Muniz Castro, K. Burton, and R. Thurer, "Accurate measurement of intraoperative blood loss during wound excision leads to more appropriate transfusion and reduced blood utilization," *Journal of Anesthesia & Clinical Research,* vol. 8, no. 11, pp. 1-6, 2017. doi: 10.4172/2155-6148.1000783

[27] T. Kadir and F. Gleeson, "Lung cancer prediction using machine learning and advanced imaging techniques," *Translational lung cancer research,* vol. 7, no. 3, p. 304, 2018. doi: 10.21037/tlcr.2018.05.15

[28] W. Sun, B. Zheng, and W. Qian, "Computer aided lung cancer diagnosis with deep learning algorithms," in *Medical imaging 2016: computer-aided diagnosis*, 2016, vol. 9785: International Society for Optics and Photonics, p. 97850Z. doi: 10.1117/12.2216307

[29] O. Echaniz and M. Graña, "Ongoing work on deep learning for lung cancer prediction," in *International Work-Conference on the Interplay Between Natural and Artificial Computation*, 2017: Springer, pp. 42-48. doi: 10.1007/978-3-319-59773-7_5

[30] X. Zhao, L. Liu, S. Qi, Y. Teng, J. Li, and W. Qian, "Agile convolutional neural network for pulmonary nodule classification using CT images," *International journal of computer assisted radiology and surgery,* vol. 13, no. 4, pp. 585-595, 2018. doi: 10.1007/s11548-017-1696-0

[31] S. J. Pawan, "Learning from small data," 9 March 2019. Available at: https://blog.goodaudience.com/

[32] R. Keshari, M. Vatsa, R. Singh, and A. Noore, "Learning structure and strength of CNN filters for small sample size training," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9349-9358. doi: 10.1109/CVPR.2018.00974

[33] G. da Silva, A. Silva, A. de Paiva, and M. Gattass, "Classification of malignancy of lung nodules in CT images using convolutional neural network," in *Anais do XVI Workshop de Informática Médica*, 2016: SBC, pp. 21-29. Available at: Google Scholar.

[34] W. Alakwaa, M. Nassef, and A. Badr, "Lung cancer detection and classification with 3D convolutional neural network (3D-CNN)," *Lung Cancer,* vol. 8, no. 8, p. 409, 2017. doi: 10.14569/IJACSA.2017.080853.

[35] D. Riquelme and M. A. Akhloufi, "Deep learning for lung cancer nodules detection and classification in CT scans," *AI,* vol. 1, no. 1, pp. 28-67, 2020. doi: 10.3390/ai1010003

[36] A. Asuntha and A. Srinivasan, "Deep learning for lung Cancer detection and classification," *Multimedia Tools and Applications,* vol. 79, no. 11, pp. 7731-7762, 2020. doi: 10.1007/s11042-019-08394-3

[37] J. H. Lee *et al.*, "Performance of a Deep Learning Algorithm Compared with Radiologic Interpretation for Lung Cancer Detection on Chest Radiographs in a Health Screening Population," *Radiology,* vol. 297, no. 3, pp. 687-696, 2020. doi: 10.1148/radiol.2020201240

[38] S. Takahashi *et al.*, "Predicting Deep Learning Based Multi-Omics Parallel Integration Survival Subtypes in Lung Cancer Using Reverse Phase Protein Array Data," *Biomolecules,* vol. 10, no. 10, p. 1460, 2020. doi: 10.3390/biom10101460

[39] A. Bhandary *et al.*, "Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images," *Pattern Recognition Letters,* vol. 129, pp. 271-278, 2020. doi: 10.1016/j.patrec.2019.11.013

[40] L. Cong, W. Feng, Z. Yao, X. Zhou, and W. Xiao, "Deep Learning Model as a New Trend in Computer-aided Diagnosis of Tumor Pathology for Lung Cancer," *Journal of Cancer,* vol. 11, no. 12, p. 3615, 2020. doi: 10.7150/jca.43268

[41] B. Marr, "The top 10 artificial intelligence trends everyone should be watching in 2020.," 6 January 2020. Available at: https://www.forbes.com/sit.

[42] M. Mahmud, M. S. Kaiser, T. M. McGinnity, and A. Hussain, "Deep learning in mining biological data," *Cognitive Computation,* vol. 13, no. 1, pp. 1-33, 2021. doi: 10.1007/s12559-020-09773-x

[43] F. Piccialli, V. Di Somma, F. Giampaolo, S. Cuomo, and G. Fortino, "A survey on deep learning in medicine: Why, how and when?," *Information Fusion,* vol. 66, pp. 111-137, 2021. doi: 10.1016/j.inffus.2020.09.006

[44] T. Nakaura, T. Higaki, K. Awai, O. Ikeda, and Y. Yamashita, "A primer for understanding radiology articles about machine learning and deep learning," *Diagnostic and Interventional Imaging,* 2020. doi: 10.1016/j.diii.2020.10.001

[45] S. Cui, H. H. Tseng, J. Pakela, R. K. Ten Haken, and I. El Naqa, "Introduction to machine and deep learning for medical physicists," *Medical physics,* vol. 47, no. 5, pp. e127-e147, 2020. doi: 10.1002/mp.14140

[46] S. Pouyanfar *et al.*, "A survey on deep learning: Algorithms, techniques, and applications," *ACM Computing Surveys (CSUR),* vol. 51, no. 5, pp. 1-36, 2018. doi: 10.1145/3234150

[47] D. Yu *et al.*, "Convolutional neural networks for predicting molecular profiles of non-small cell lung cancer," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017: IEEE, pp. 569-572. doi: 10.1109/ISBI.2017.7950585

[48] S. M. Salaken, A. Khosravi, A. Khatami, S. Nahavandi, and M. A. Hosen, "Lung cancer classification using deep learned features on low population dataset," in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2017: IEEE, pp. 1-5. doi: 10.1109/CCECE.2017.7946700

[49] I. Ali *et al.*, "Lung nodule detection via deep reinforcement learning," *Frontiers in oncology,* vol. 8, p. 108, 2018. doi: 10.3389/fonc.2018.00108

[50] X. Liu, F. Hou, H. Qin, and A. Hao, "Multi-view multi-scale CNNs for lung nodule type classification from CT images," *Pattern Recognition,* vol. 77, pp. 262-275, 2018. doi: 10.1016/j.patcog.2017.12.022

[51] H. Cao, S. Pu, W. Tan, and J. Tong, "Breast mass detection in digital mammography based on anchor-free architecture," *Computer Methods and Programs in Biomedicine,* vol. 205, p. 106033, 2021. doi: 10.1016/j.cmpb.2021.106033

[52] S. Kaur, H. Aggarwal, and R. Rani, "Diagnosis of Parkinson's disease using deep CNN with transfer learning and data augmentation," *Multimedia Tools and Applications,* vol. 80, no. 7, pp. 10113-10139, 2021. doi: 10.1007/s11042-020-10114-1

[53] S. Dutta, P. Prakash, and C. G. Matthews, "Impact of data augmentation techniques on a deep learning based medical imaging task," in *Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications*, 2020, vol. 11318: International Society for Optics and Photonics, p. 113180M. doi: 10.1117/12.2549806

[54] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 international interdisciplinary PhD workshop (IIPhDW)*, 2018: IEEE, pp. 117-122. doi: 10.1109/IIPHDW.2018.8388338

[55] J. Cho, K. Lee, E. Shin, G. Choy, and S. Do, "How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?," *arXiv preprint arXiv:1511.06348,* 2015. Available at: Google Scholar.

[56] A. Sarkis, "How to create AI ready data for object detection," 2018. Available at: https://medium.com.

[57] H.-y. Lee, "Deep learning tutorial," *Open Course, Online Available,* 2020. Available at: Google Scholar.

[58] E. MalekHosseini, M. Hajabdollahi, N. Karimi, S. Samavi, and S. Shirani, "Splitting Convolutional Neural Network Structures for Efficient Inference," *arXiv preprint arXiv:2002.03302,* 2020. Available at: Google Scholar.

[59] E. C. Putro, R. M. Awangga, and R. Andarsyah, *Tutorial Object Detection People With Faster region-Based Convolutional Neural Network (Faster R-CNN).* Kreatif, 2020. Available at: Google Books.

[60] H. Dao, "Image Classification Using Convolutional Neural Networks," 2020. Available at: Google Scholar.