Int. J. Nonlinear Anal. Appl.
Volume 12, Special Issue, Winter and Spring 2021, 1047-1058
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/IJNAA.2021.5551

Classification of lung nodules in CT images using conditional generative adversarial – convolutional neural network

Nur Nabila Mohd Isham[a], Siti Salasiah Mokri[a,*], Ashrani Aizuddin Abd Rahni[a], Nurul Fatihah Ali[a]

[a]*Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan, Malaysia.*

(Communicated by Madjid Eshaghi Gordji)

Abstract

Based on Global Cancer 2015 statistics, the lung cancer of all types constitutes 27% of overall cancers while 19.5% of cancer deaths are due to lung cancer. In lieu of this, an effective lung cancer screening test using Computed Tomography (CT) scan is crucial to detect cancer at the early stage. The interpretation of the CT images requires an advanced CAD system of high accuracy for instance, in classifying the lung nodules. Recently, Deep Learning method that is Convolution Neural Network (CNN) shows an outstanding success in lung nodules classification. However, the training of CNN requires a great number of images. Such a requirement is an issue in the case of medical images. Generative adversarial network (GAN) has been introduced to generate new image datasets for CNN training. Thus, the main objective of this study is to compare the performance of CNN architectures with and without the implementation of GAN for lung nodules classification in CT images. Here, the study used Conditional GAN (cGAN) to generate benign nodules images. The classification accuracy of the combined cGAN-CNN architecture was compared among CNN pre-training networks namely GoogleNet, ShuffleNet, DenseNet, and MobileNet based on classification accuracy, specificity, sensitivity, and AUC-ROC values. The experiment was tested on LIDC-IDRI database. The results showed cGAN-CNN architecture improves the overall classification accuracy as compared to CNN alone with the cGAN-ShuffleNet architecture performed the best, achieving 98.38% accuracy, 98.13% specificity, 100% sensitivity and AUC-ROC at 99.90%. Overall, the classification performance of CNN can be improved by integrating GAN architecture to mitigate the constraint of having a large medical image dataset, in this case, CT lung nodules images.

Keywords: Computed yomography, Convolution neural network, Generative adversarial network, Lung nodules, Classification

*Corresponding author
Email addresses: a165617@siswa.ukm.edu.my (Nur Nabila Mohd Isham), siti1950@ukm.edu.my (Siti Salasiah Mokri), ashrani@ukm.edu.my (Ashrani Aizuddin Abd Rahni), p111397@siswa.ukm.edu.my (Nurul Fatihah Ali)

Received: June 2021 *Accepted:* September 2021

1. Introduction

According to World Health Organization (WHO) in 2014, the ratio of deaths due to lung cancer in Malaysia is 19.1 deaths per 100,000 and this accounts for 3.22% of all deaths [6]. Correspondingly significant, there are 131,880 deaths due to lung cancer in the United States in this year alone (2021) that consume 28.99 % of all cancer deaths [3]. One of the biggest risk factors of lung cancer is smoking; either the primary or secondary smokers [10]. Lung cancer or lung carcinoma occurs when there is out-of-control cell growth in the lung tissues. Normally, the body has its own inspection and balance system to monitor the process of cell division to produce new cells. However, when the system is disrupted, the grow of out-of-control cells happens leading to tumors. The tumors are categorized into benign (non-cancerous) and malignant (cancerous) [23].

Early screening is very important to detect lung cancer so that proper intervention can be planned. Computed tomography (CT) screening is the commonly used screening method to detect lung cancer due to its high resolution capability [19]. In CT machine, computer aided diagnostic (CAD) system is utilized to assist the physicians and radiologists to detect and diagnosis the nodules or tumors. The CAD systems consist of two types; computer aided detection (CADe) system and computer aided diagnosis (CADx) system [16]. The CADe system is to detect nodular from non - nodular parts and CADx system is used to classify lung nodules into benign or malignant nodules. However, most of the available CAD systems that are based on non- artificial intelligent based algorithms have the disadvantage to generate fast diagnostic results [4].

Currently, deep learning has become the most advanced machine learning method and is commonly used in many applications in conjunction with the readily available supercomputing technologies and large databases [8]. There are various architectures that have been introduced in deep learning. Significantly, convolution neural network (CNN) architecture is considered as the most established deep learning network especially in image recognition application [18]. To further improve CNN performance, various types of CNN extended networks have been proposed. However, CNN training requires large datasets that becomes an issue for medical images. This problem is alleviated through the application of GAN architecture that can generate additional images [12].

There are several published methods that work on lung cancer classification that integrate GAN architecture. Yan Kuang et al. [9] compared Multi-Discriminators GAN and Encoder (MDGAN and Encoder), GAN and Encoder, and GAN architecture for lung nodule classification. They have showed that MDGAN and Encoder showed the best performance with 95.32% accuracy, 90.79% specificity, and 94.15% sensitivity.

Defang Zao et al. [22] proposed a method of increasing the dataset to improve the classification performance of lung lesions via Forward and Backward GAN (F&BGAN). There are two levels of F&BGAN that is Forward GAN (FGAN) where it can produce various images while Backward GAN (BGAN) is used to increase image quality. They compared five architectures namely M-VGG16, VGG-16, Resnet, GoogleNet, and CNN. It was shown that the M-VGG16 architecture obtained an excellent performance compared to all other architectures with 95.24% accuracy, 92.47% specificity, and 98.67% sensitivity.

Yuya Onishi et al. [13] proposed three-dimensional convolution neural network (3DCNN) using lung nodules synthesized by 3DGAN. It is a multi-scale 3DCNN (M-scale 3DCNN) architecture where nodule areas of different sizes are inserted into the 3DResnet and combined into the last convolution layer to produce a comprehensive classification result. The convolution layer uses a 3x3x3 pixel kernel and produces a feature map with the same shape. Then, the convolution layers from each 3DResnet are combined. After that, the feature map is expanded to one fully connected dimension and the classification probabilities are extracted through the softmax function. The proposed method

obtained a sensitivity of 90.9%, a specificity of 74.1% and an accuracy of 82.5% . In the next study, they proposed a classification model based on the CNN and GAN called as WGAN [14].

On a different note, this paper presents the comparison between CNN and GAN-CNN specifically Conditional GAN (cGAN) to classify 2D CT scans of lung nodules into benign and malignant. Four pre-training CNN networks are chosen and the percentage of improvement of each network when cGAN is utilized is investigated. These CNN pretraining networks are GoogleNet, ShuffleNet, DenseNet, and MobileNet. Correspondingly, the ideal ratio between benign and malignant images used for CNN training is also studied so that a significant classification improvement can be achieved as compared to sole CNN classification.

2. Methodology

In this study, The Lung Image Database Consortium image collection (LIDC-IDRI) datasets was used [1]. Figure 1 shows the information about the dataset. This dataset consists of 1018 cases collected from seven academic institutions and eight medical imaging companies. There are 1646 data containing benign and malignant nodules. There are 1423 malignant cases and 223 benign cases. Then, these data were resized to 244×244 and saved in PNG file.

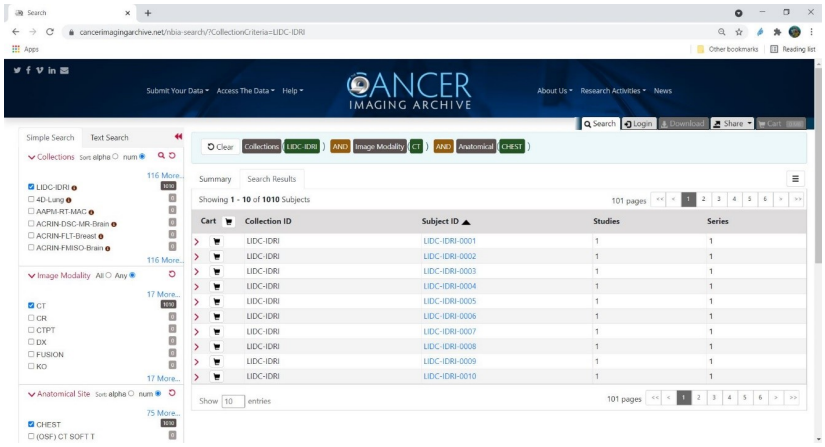

Figure 1: LIDC-IDRI datasets

The software used in this study was MATLABR2020a Deep Learning Toolbox. It is used to design the deep learning architecture, set the appropriate parameters, and conduct the training and testing for each of the selected architecture. The computer environment used in this study was Intel (R) Core (TM) i5-7200U CPU@ 2.50 Hz, NVIDIA GeForce 920MX graphics card. The effectiveness of the GAN-CNN classifier to classify CT lung nodules was studied. There were four CNN pretraining networks selected, namely GoogleNet, ShuffleNet, DenseNet and MobileNet.

Generative Adversarial Network (GAN) [20] consists of two deep learning networks, which are the generator and the discriminator. The generator network generates new fake image where it receives random input or noise. On the other hand, the discriminator network receives inputs from two images; the real image and the fake image generated by the generator. The discriminator will then discriminate each data as fake or not, that is it belongs to the actual training dataset or not repeatedly. This is explained in Figure 2. As there are 1423 malignant cases and 223 benign cases in LIDC-IDRI dataset, Conditional GAN (cGAN) [11] that is an extension of GAN was utilized to generate more benign images to match the number of malignant images. In cGAN, the label of the generated images can be determined beforehand (Figure 3). The generator of this scheme receives

two inputs: random (noise) input as in GAN scheme as well as label input. This is to ensure that the generated images of cGAN follows the predetermined label from the label input. In our case, cGAN was used to generate benign images so that they match the number of malignant images in LIDC-IDRI dataset.

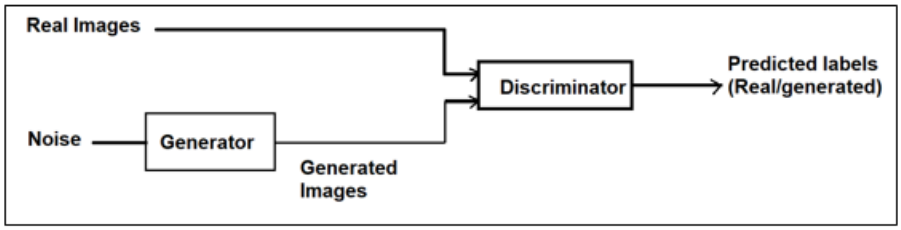

Figure 2: Generative Adversarial network (GAN) architecture

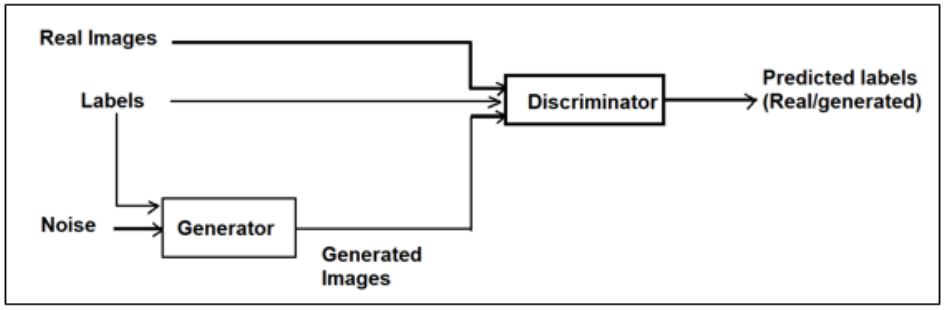

Figure 3: Conditional Generative Adversarial network (cGAN) architecture

In our implementation, malignant (1423) and benign (223) images of LIDC-IDRI database were fed to the cGAN architecture in which the benign images were arranged first in the sequence of 1 to 223 while the malignant images were arranged from 224 to 1646. The datasets were also randomly augmented through horizontal flipping and were reshaped to 64×64. The generator network was fed with two inputs; labels and noise (random vectors). The random vectors were of size 100 and it will be converted to $4 \times 4 \times 1024$ arrays. On the other hand, the labels were also changed to vectors and reshaped to 4×4 array. These two types of inputs were then concatenated to produce $4 \times 4 \times 1025$ array. This array was upscaled to $64 \times 64 \times 3$ via transposed convolution layers of 5×5 filter size and followed by ReLU tanh layers. The embedding dimension for the labels input was 50.

The discriminator network inputs were coming from $64 \times 64 \times 1$ images and labels. The images (input) were added with noise with a dropout probability of 0.7. The output was the probability score obtained via series of convolutional layers (5×5 filter size) and ReLU layers (scale 0.3). This network was completed with a convolutional layer (4×4). The cGAN network was trained according to Table 1 using Adam optimization method. To improve the learning of the generator and the discriminator, the labels of some of the real images were flipped with 0.5 flipping factor.

Four pre-training CNN architectures (GoogleNet, ShuffleNet, DenseNet and MobileNet) [2] to investigate the percentage of improvement when integrating cGAN into the CNN framework were compared. In basis, the CNN network consists of three layers (Figure 4). They are convolutional, pooling and fully-connected layers. The input layer holds the input values. The convolutional layer describes the output of neurons connected to local regions of the input through the calculation of the scalar product between their weights and the region connected to the input volume. The pooling layer downsamples along the spatial dimensionality of the given input, further reducing the number

Table 1: Training Parameters for cGAN Network

Parameter	Values
Mini-batch size	128
Number of epochs	400
Learning rate	0.0001
Gradient decay factor	0.25
Squared gradient decay factor	0.9999
Flipping factor	0.5

of parameters within that activation. Finally, the fully connected layers produce class scores from the activations, to be used for classification.

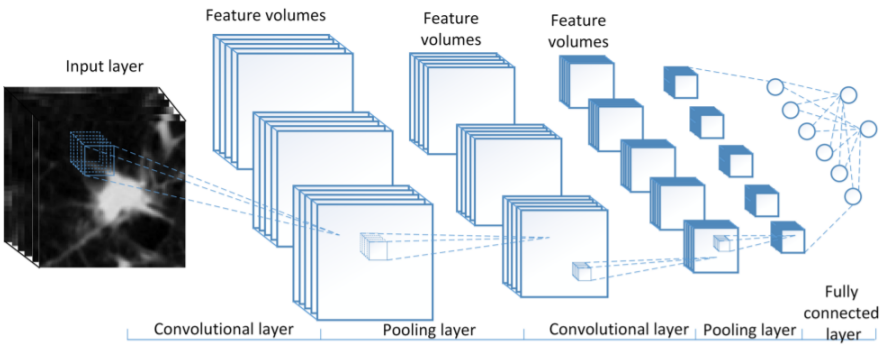

Figure 4: Convolutional Neural Network (CNN) architecture [7]

Google introduces GoogleNet [17] that contains an inception module. It is a 1×1 convolution layer that is capable to minimize the dimensionality of the array. This enables less computational burden. Exhaustive computational requirement when processing deep learning is also alleviated with the introduction of ShuffleNet [21]. In this network, shuffle channel is used to perform shuffle operation before feeding the data to the next convolution layer. Point wise group convolution is also performed before the shuffle operation under ShuffleNet scheme. These features increase the computation efficiency of ShuffleNet.

On the other hand, DenseNet [5] has a closed packed network that requires less parameters but able to achieve good accuracy. The computational constraint and memory are used efficiently as well. MobileNetV2 [15] is a memory efficient network created by Google via the use of an inverted residual structure module. Table 2 shows the values of training parameters for these CNN based pretraining networks used in the experiment.

The performance of each of the four networks with and without cGAN was assessed based on accuracy, specificity, sensitivity, and area under the curve of the receiver operating characteristic curve (AUC-ROC). Accuracy is defined as the ratio of correctly predicted data with respect to the total data as represented in Equation 1. TP is True Positive, TN is True Negative, FP is False Positive and FN is False Negative.

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

Specificity is the ratio of correctly predicted data as negative with respect to the whole negative

Table 2: Training Parameters for CNN Based Networks

Parameter	Values
Mini-batch size	64
Number of epochs	20
Learning rate	0.011
Learn rate drop factor	0.2
Learn rate drop period	5
Validation frequency	50
Momentum	0.9

data. It is defined as:

$$Specificity = \frac{TN}{TN + FP} \tag{2}$$

Sensitivity is the ratio of correctly predicted data as positive with respect to the whole positive data. It is calculated as:

$$Sensitivity = \frac{TN}{TN + FP} \tag{3}$$

Area under the curve (AUC) is measured from the Receiver Operating Characteristics (ROC) curve. It indicates the accuracy of the classifier to classify the data into the correct class. It is in the range of [0,1] whereby 1 implies perfect classification while 0 implies zero classification accuracy.

3. Results and discussion

3.1. Image generation using cGAN

Figure 5 shows some of the generated benign images using cGAN at 500, 4000 and 6500 training iteration respectively. As the iteration increases, the generated images mimic closer to the actual images. This is because the discriminator has learned the strong features efficiently to differentiate fake images from the truth images while the generator network has learned the powerful features to produce realistic images.

3.2. Generated images at iteration 500 (top left), 4000 (top right), and 6500 (buttom)

The performance of the cGAN-CNN when different proportion of malignant and benign dataset is used was also investigated. As previously mentioned, LIDC-IDRI dataset has an imbalance proportion of benign images (223) in comparison to malignant images (1423). The performance of the cGAN-CNN when 500, 1000 and 1500 benign images versus 1423 malignant images were compared. As the original benign images is 223, cGAN was used to provide the additional 277 images so that the total amount of benign images is 500. The same approach was carried out to prepare the 1000 and 1500 benign images. Figure 6 shows some of the benign and malignant datasets used in PNG format.

Table 3 shows the performance of different cGAN-CNN pretrained networks when different ratio of benign and malignant was used. Here, first stage implies that 500 benign images were used, second stage is when 1000 benign images were used and lastly third stage is when 1500 benign images were used. In all stages, 1423 malignant images were used. In overall, second stage and third stage obtain better performance than the first stage in terms of the accuracy. Moreover, both

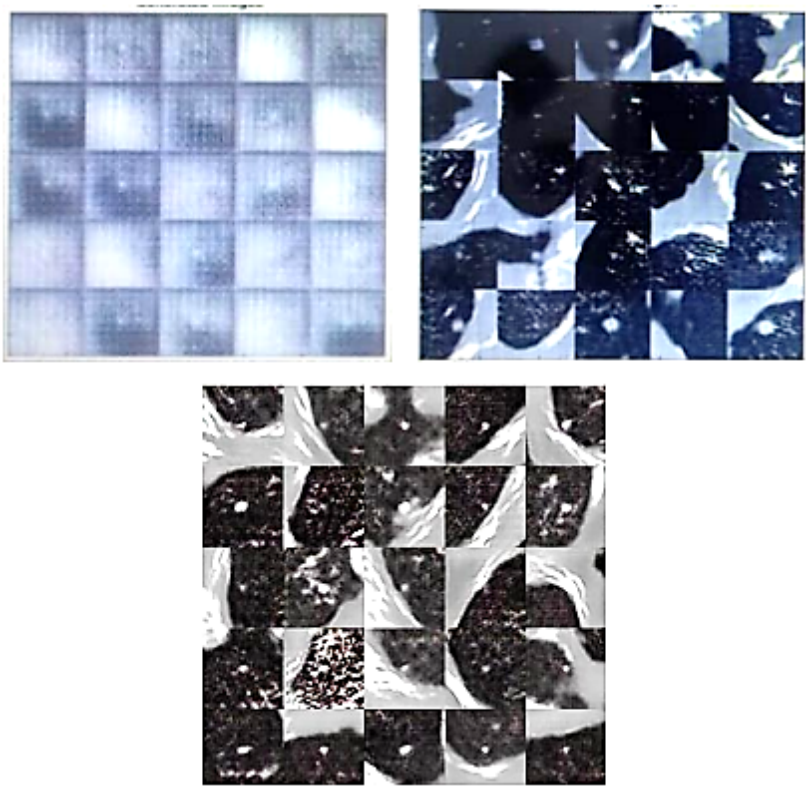

Figure 5: Generated images at iteration 500 (top left), 4000 (top right), and 6500 (buttom)

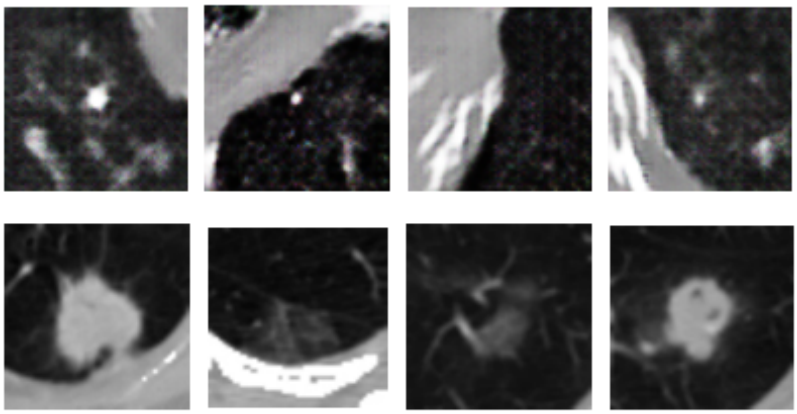

Figure 6: Benign (above) and Malignant (below) nodules

second and third stage show comparable performance with the third stage is slightly better in cGAN-GoogleNet and cGAN-DenseNet than the second stage. However, the second stage is slightly better in case of cGAN-ShuffleNet and cGAN-MobileNet. This experiment implies that the accuracy of lung lesions classification using cGAN-CNN can be improved if the ratio of benign images with respect to malignant datasets is more than 0.7.

3.3. Performance of cGAN-CNN vs CNN

The performance results for cGAN-CNN vs CNN are shown in Table 4. Here, only the third stage of cGAN-CNN implementation is presented. In general, the introduction of cGAN to CNN in increasing benign images to be comparable to the number of malignant images has greatly improved

Table 3: Comparison of cGAN-CNN of Different Ratio of Benign-Maligant Images.

Network	Accuracy	Specificity
cGAN-GoogleNet		
First stage	95.34%	98.83%
Second stage	95.55%	97.89%
Third Stage	95.95%	99.06%
cGAN-ShuffleNet		
First stage	97.77%	98.13%
Second stage	98.79%	98.59%
Third Stage	98.38%	98.13%
cGAN-DenseNet		
First stage	96.96%	96.72%
Second stage	97.17%	96.96%
Third Stage	97.98%	98.59%
cGAN-MobileNet		
First stage	91.90%	91.10%
Second stage	95.55%	98.83%
Third Stage	97.94%	95.78%

the original CNN classifier. Considering CNN classifier alone, in overall, ShuffleNet obtains the best classification performance, followed by DenseNet, GoogleNet and lastly MobileNet. Although the accuracy achieved by ShuffleNet is equal to DenseNet, this network performs better in terms of sensitivity and AUC-ROC values. On the other hand, MobileNet stands last with 92.91% accuracy, 62.69% sensitivity and AUC-ROC of 82.11%.

When comparing cGAN-CNN and CNN networks, cGAN-GoogleNet improves the GoogleNet performance with an increment of 0.62% in accuracy, 11.94% in sensitivity and 2.30% in AUC-ROC. cGAN-ShuffleNet improves the accuracy of ShuffleNet classifier by 4.86% while the sensitivity and AUC-ROC are increased by 40.3% and 17.48% respectively. The third architecture is cGAN-DenseNet that shows an increase for accuracy value by 4.46%, sensitivity by 29.85% and AUC by 13.80% in comparison to DenseNet. Lastly, in terms of cGAN-MobileNet, an improvement of accuracy value by 2.03% is seen while the sensitivity and AUC-ROC are improved by 26.86% and 16.54% respectively as compared to MobileNet.

Here, the best architecture is cGAN-ShuffleNet that achieves 100% of sensitivity and 99.99% of AUC-ROC. The minimum percentage of improvement is seen in cGAN-GoogleNet & GoogleNet as compared to other models. In obvious, cGAN-CNNs improve the CNN classifiers with an average improvement of 2.99 % in accuracy, an average improvement of 27.24% in sensitivity and an average improvement of 12.53% in AUC-ROC. Figure 7 shows the percentage of improvement of cGAN-CNN on all pretrained networks.

Figure 8 shows the receiver operating curve (ROC) for each cGAN-CNN architecture. cGAN-ShuffleNet achieves an AUC of 99.90% followed by cGAN-DenseNet (AUC 99.98%), then cGAN-MobileNet (AUC 98.65%) and lastly cGAN-GoogleNet (AUC 93.04%).

The next simulation is the prediction of each cGAN-CNN architecture (Third Stage) to correctly classify a randomly selected image. In this case, it is an image with a malignant lesion. Figure 9 shows the prediction results. All cGAN-CNN have successfully classify this image as malignant with cGAN-ShuffleNet obtains the highest prediction of 0.96, followed by cGAN-DenseNet, cGAN-GoogleNet and finally cGAN-MobileNet with a prediction accuracy of 0.91.

Table 4: Comparison of CGAN-CNN of Different Ratio of Benign-Maligant Images.

Architecture	Acc	Sens	AUC-ROC
GoogleNet	93.93%	61.19%	82.41%
cGAN-GoogleNet (Third Stage)	95.55%	73.13%	94.71%
ShuffleNet	93.52%	59.70%	82.42%
cGAN-ShufleNet (Third Stage)	98.38%	100%	99.90%
DenseNet	93.52%	64.18%	86.00%
cGAN-DenseNet (Third Stage)	97.89%	94.03%	99.80%
MobileNet	92.91%	62.69%	82.11%
cGAN-MobileNet (Third Stage)	94.94%	89.55%	98.65%

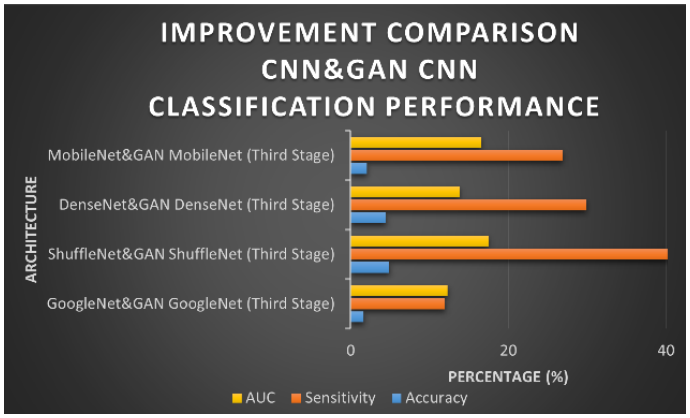

Figure 7: Improvement of Classification Performance by cGAN-CNN on four pretrained CNN networks

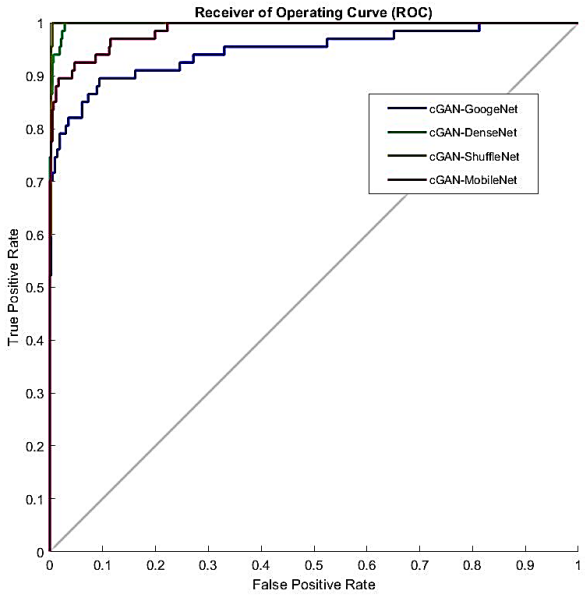

Figure 8: Comparison of ROC curve for each cGAN-CNN network

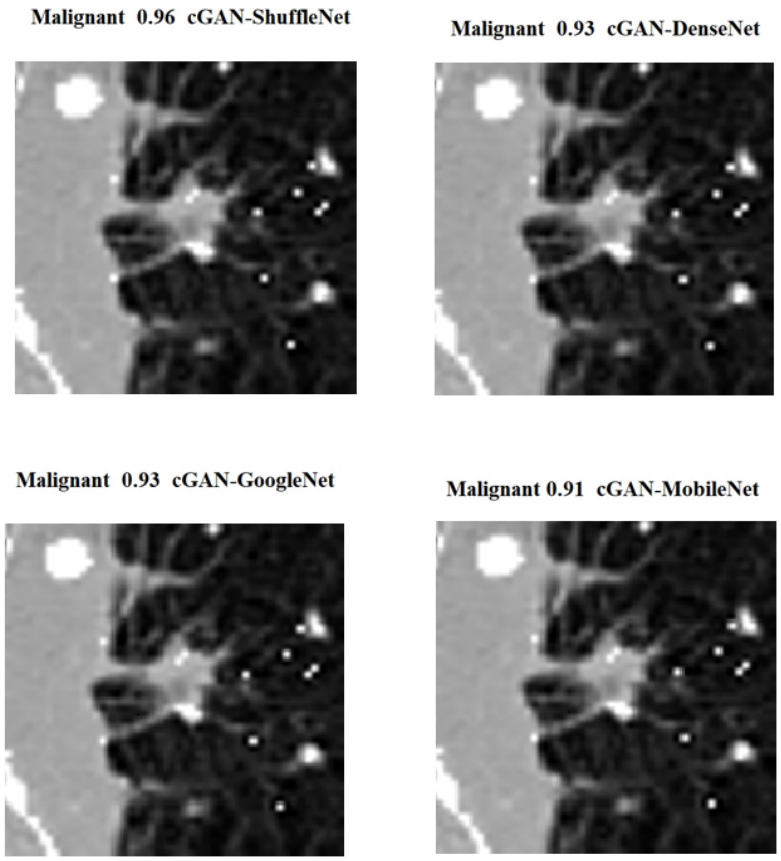

Figure 9: Prediction data for each GAN CNN architectures (Third Stage)

4. Conclusion

This paper has presented the comparison of 4 CNN based pretraining networks with and without cGAN for lung lesions classification in LIDC-IDRI CT images. In overall, the classification accuracy of CNN classifier could be improved via cGAN as it provides large number of datasets used for CNN training. By comparison, cGAN-ShuffleNet achieves the best classification performance of 98.38% accuracy, 98.13% specificity, and 100% sensitivity than cGAN-DenseNet, cGAN-GoogleNet and cGAN-MobileNet. Furthermore, it is also shown that the classification accuracy of cGAN-CNN is significantly improved if the ratio of benign images to malignant images is more than 0.7:1.

Acknowledgement

The authors would like to acknowledge Universiti Kebangsaan Malaysia and the Ministry of Education, Malaysia (MOE) for the Fundamental Research Grant Scheme (FRGS) with code: FRGS/1/2019/TK0
to support this project.

References

[1] S. Armato III, G. Mclennan, L. Bidaut, M. McNitt-Gray, C. Meyer, A. Reeves, B. Zhao, B., et al, *The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans*, Medical Physics, 38(2011) 915–93, doi: 10.1118/1.3528204.

[2] S.M. Ashhar, S.S. Mokri, A.A. Abd Rahni, A.B. Huddin, N. Zulkarnain, N.A. Azmi, and T. Mahaletchumy, *Comparison of deep learning convolutional neural network (CNN) architectures for CT lung cancer classification*, International Journal of Advanced Technology and Engineering Exploration, 8 (74)(2021)126, doi: 10.19101/IJATEE.2020.S1762126.

[3] M. Bittoni, J.C. Yang, J.Y. Shih, N. Peled, E.F. Smit, D.R. Camidge, and P.K. Paik, *Real-world insights into patients with advanced NSCLC and MET alterations*. Lung Cancer, 159 (2021) 96-106, doi: 10.1016/j.lungcan.2021.06.015.

[4] S.A. El-Regaily, M.A. Salem, M.H. Abdel Aziz, and M.I. Roushdy, *Survey of Computer Aided Detection Systems for Lung Cancer in Computed Tomography*, Current Medical Imaging Reviews, 14 (1)(2017) 3–18, doi: 10.2174/1573405613666170602123329.

[5] G. Huang, and K.Q. Weinberger, *Densely Connected Convolutional Networks*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2018) 2261-2269, doi: 10.1109/CVPR.2017.243.

[6] C.S. Kan, and K. Chan, *A Review of Lung Cancer Research in Malaysia*, The Medical Journal of Malaysia, (71) (2016)70–78, PMID: 27801389.

[7] G. Kang, K. Liu, B. Hou, and N. Zhang N, *3D multi-view convolutional neural networks for lung nodule classification*, PLoS ONE, 12(11)(2017) e0188290, doi: 10.1371/journal.pone.0188290.

[8] K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, and D.I. Fotiadis, *Machine learning applications in cancer prognosis and prediction*, Computational and Structural Biotechnology Journal, 13(2015) 8–17, doi: 10.1016/j.csbj.2014.11.005.

[9] Y.A.N. Kuang, T. Lan, and X. Peng, *Unsupervised Multi-Discriminator Generative Adversarial Network for Lung Nodule Malignancy Classification*, IEEE Access, 8 (2020) 77725–77734, doi: 10.1109/ACCESS.2020.2987961.

[10] J.H.B. Masud, *Lung diseases and smoking: A systematic analysis of big data in the era of artificial intelligence*, Tobacco Induced Diseases, 19(1)(2021) , doi:10.18332/tid/141067.

[11] M. Mirza, and S. Osindero, *Conditional generative adversarial nets*, arXiv preprint arXiv:1411.1784, 2014.

[12] Y. Onishi, A. Teramoto, M. Tsujimoto, T. Tsukamoto, K. Saito, H. Toyama, K. Imaizumi, and et al, *Automated Pulmonary Nodule Classification in Computed Tomography Images Using a Deep Convolutional Neural Network Trained by Generative Adversarial Networks*, Biomed Res Int., 2019 (6051939) 2019, doi: 10.1155/2019/6051939.

[13] Y. Onishi, A. Teramoto, M. Tsujimoto, T. Tsukamoto, K. Saito, and H. Toyama, *Investigation of pulmonary nodule classification using multi - scale residual network enhanced with 3DGAN - synthesized volumes*, Radiol Phys Technol, 13(2)(2020) 160-169, doi: 10.1007/s12194-020-00564-5.

[14] Y. Onishi, A. Teramoto, M. Tsujimoto, T. Tsukamoto, K. Saito, and H. Toyama, *Multiplanar analysis for pulmonary nodule classification in CT images using deep convolutional neural network and generative adversarial networks*, Int J Comput Assist Radiol Surg, 15(1)(2020) 173-178, doi: 10.1007/s11548-019-02092-z.

[15] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.C. Chen, *MobileNetV2: Inverted Residuals and Linear Bottlenecks*, IEEE Conference on Computer Vision and Pattern Recognition, (2018) 4510-4520, doi: 10.1109/CVPR.2017.243.

[16] Q.Z. Song, L. Zhao, X.K. Luo, and X.C. Dou, *Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images*, Journal of Healthcare Engineering, 2017 (Article ID 8314740) (2017), doi: 10.1155/2017/8314740.

[17] C. Szegedy, S. Reed, P. Sermanet, V. Vanhoucke, and A. Rabinovich, *Going deeper with convolutions*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2015) 1-9, doi: 10.1109/CVPR.2015.7298594.

[18] A. Teramoto, H. Fujita, O. Yamamuro, and T. Tamaki, *Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique*, Medical Physics, 43 (6)(2016) 2821–2827, doi: 10.1118/1.4948498.

[19] G.S. Tran, T.P. Nghiem, V.T. Nguyen, C.M. Luong, J.C. Burie, and Y. Levin-Schwartz, *Improving Accuracy of Lung Nodule Classification Using Deep Learning with Focal Loss*, Journal of Healthcare Engineering, 2019 (2019), doi: 10.1155/2019/5156416.

[20] X. Yi, E. Walia, and P. Babyn, *Generative adversarial network in medical imaging: A review*, Medical image analysis, 58 (2018) 101552, doi: 10.1016/j.media.2019.101552.

[21] X. Zhang, X. Zhou, M. Lin, and J. Sun, *Shufflenet: An extremely efficient convolutional neural network for mobile devices*, In Proceedings of the IEEE conference on computer vision and pattern recognition, (2018) 6848-6856, doi: 10.1109/CVPR.2018.00716.

[22] D. Zhao, D. Zhu, J. Lu, and Y. Luo, *Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16*. Symmetry, 10 (519) (2018), doi: 10.3390/sym10100519.

[23] M. Zhang, N. Zhuo, Z. Guo, X. Zhang, W. Liang, S. Zhao, and J. He, *Establishment of a mathematic model for predicting malignancy in solitary pulmonary nodules*, Journal of thoracic disease, vol. 7(10) (2015) 1833-1841, doi: 10.3978/j.issn.2072-1439.2015.10.56.