RMJ Ramathibodi Medical Journal

Rama Med J | Original Article

NoduleNet: A Lung Nodule Classification Using Deep Learning

Ganesh Shah[1], Ratchainant Thammasudjarit[1], Ammarin Thakkinstian[1], Thitiporn Suwatanapongched[2]

[1] Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

[2] Department of Diagnostic and Therapeutic Radiology, Faculty of Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

Background: Accurate detection and classification of lung nodules at an early stage can help physicians to improve the treatment outcomes of lung cancer. Several lung nodule classifications using deep learning have been proposed but they are lag of external validation to Thai patient data.

Objective: To propose a deep learning model called NoduleNet for lung nodule classification and perform internal and external validation of the proposed model.

Methods: Two datasets were performed; internal validation using LUNA16 (the public lung CT database), and external validation using ChestRama (37 chest CT scans retrospectively identified from the CT database of Ramathibodi Hospital between 2017 and 2019). The NoduleNet was built on top of pretrained architecture, VGG16, and VGG19 with customization.

Results: The NoduleNet showed impressive results in nodule classification. The best model achieved accuracy of 0.95 (0.94 - 0.96), sensitivity of 0.84 (0.82 - 0.86), and specificity of 0.97 (0.97 - 0.98) for internal validation, where the external validation results was accuracy of 0.95 (0.87 - 1.00), sensitivity of 0.91 (0.82 - 1.00), and specificity of 1.00 (1.00 - 1.00). There were 3 misclassified samples in external validation which are all false-negative.

Conclusions: The NoduleNet is able to generalize from non-Thai patient data to Thai patient data. It could be further improved by taking sequence of images into account, integrating with an automatic nodule detection algorithm, and adding more nodule types.

Keywords: Deep learning, Lung nodule classification, Transfer learning

Rama Med J: doi:10.33165/rmj.2020.43.4.241727

Received: May 29, 2020 **Revised:** October 7, 2020 **Accepted:** October 30, 2020

Corresponding Author:
Ratchainant Thammasudjarit
Department of Clinical Epidemiology and Biostatistics,
Faculty of Medicine
Ramathibodi Hospital,
Mahidol University,
270 Rama VI Road, Ratchathewi,
Bangkok 10400, Thailand.
Telephone: +66 2201 1269
E-mail: ratchainant.tha@mahidol.edu

Introduction

Lung cancer remains the leading cause of deaths with 18.4% death rates in both men and women worldwide. International Agency for Research on Cancer (IARC) reported in 2018 that 2.094 million new cases of lung cancer were diagnosed, of which the death accounted for about 1.761 million.[1] Early diagnosis of lung cancer is a good prognosis to receive early treatments, as for 10-year survival of patients with stage I lung cancer was longer than those with advanced stages.[2]

Computed tomography (CT) scan is commonly used for the diagnosis of lung cancer after screening by general chest X-ray. Analyzing CT scan images is required a well-trained radiologist to interpret findings that could take time approximately 9.1 ± 2.3 minutes per CT scan.[3] Several automatic nodule classifications were therefore proposed to reduce such workload, in which training data were mostly from non-Thai patient data where performances are questionable whether a model train from non-Thai patient data can be generalized to Thai patient data or not.

Nodule classification is considered as an image classification problem. Since lung nodule is a low-contrast tissue which is uneasy to distinguish, the early works employ predefined features obtained from image processing or manually hand-crafted to extract meaningful features including shapes, textures, densities, and intensity, and then classify using machine learning.[4-12] However, such feature engineering tasks were not flexible enough to generalize to other domains. The deep learning approach was introduced as an end-to-end solution for image classification without involving feature engineering because the model could learn features by itself. The state-of-the-art model for image classification problem was the convolutional neural network (CNN), which was widely proven in many domains including medical areas.[13]

Training strategy for CNN depends on the size of training data and domain-specific. If training data is sufficient, CNN is flexible enough to be trained by any strategy. Conversely, if the training data is small, transfer learning[14] is commonly used for training strategy that leverages a pretrained model. The pretrained model is another CNN that was trained from a large training dataset $\mathcal{D}_\ell$. The transfer learning takes pretrained model parameters as the initialized parameters then learn on the new dataset which is relatively small $\mathcal{D}_s$. Although domains are different between $\mathcal{D}_\ell$ and $\mathcal{D}_s$, transfer learning is still powerful enough to apply.

Several pretrained models are available (VGG,[15] RESNet,[16] Inception,[17] Xception,[18] MobileNet,[19] DenseNet,[20] and NASNet[21]) where each model performance does not very different based on ImageNet dataset.[22] Selecting pretrained models was very much dependent on the number of trainable parameters and model complexity relative to the number of training data. Although existing works[4-12] proposed various CNN architectures for lung nodule classification using non-Thai patient data, generalizability to Thai patient data is still questionable. The main contribution of this study emphasizes on proposing a CNN model and externally validate with Thai patient data for lung nodule classification to either benign or malignancy.

Methods

The study design was a cross-sectional study that uses 2 data sources: 1) LUNA16 dataset,[23] and 2) ChestRama dataset. The LUNA16 was constructed by the public lung CT scan database namely LIDC/IDRI[24] which aimed to classify nodule to either benign or malignancy. The number of samples in LUNA16 is 8106 cropped nodule images; 6755 (83.33%) images of benign nodules, and 1351 (16.67%) images of malignant nodules. Each image size is the fixed size at 50×50 pixels of height and width. The LUNA16 included nodule size greater or equal to 3 mm, less than 3 mm, and non-nodule size greater or equal to 3 mm. The ChestRama dataset was constructed by the Department of Radiology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Thailand which consists of 37 full chest CT scans performed by experienced chest radiologists from 2017 to 2019. The inclusion criteria were defined as follows: 1) chest CT examination obtained with a slice thickness of less than or equal to 2.5 mm, 2) the selected

nodules were annotated by the pulmonary radiologist with nodule sizes of 3 mm or larger. Among 37 CT scans, 15 (40.5%) and 22 (59.5%) were benign and malignant images, respectively. The study was approved by the Human Research Ethics Committee of Faculty of Medicine Ramathibodi Hospital, Mahidol University, Thailand, No. MURA2018/616 on October 12, 2018.

Nodule was cropped manually from each original image, then resize to 50×50 pixels. The LUNA16 dataset was used for development and internal validation of the model by split data into 3 groups (training set, test set, and validated set) using stratify sampling with about 60% for training, and 15% to 20% for testing and internal validation, respectively (Table 1).

The best model performance from internal validation was later used for the external validation on the ChestRama data in order to evaluate the generalizability of the proposed model.

Data preprocessing was required for training the purposed model, called NoduleNet, because the input was a tensor object with a dimension of 50×50×3 where 50×50 represents image size (height, width) in pixels, and 3 represents the number of color channels. However, the LUNA16 dataset was the gray image dataset where each image could be transformed into a tensor object with a dimension of 50×50×1. Such a dimension disagreement requires Gray-to-RGB (Red-Green-Blue) method that stacks 2 dummy tensor objects with the original tensor object, where each dummy tensor object is copied from the original tensor object (Figure 1).

The proposed NoduleNet model was built from transfer learning on top of the selected pretrained models: VGG16 and VGG19. Such pretrained models were selected because their architectures could be applied with small size images. Other pretrained models were inappropriate for small size images because the nature of CNN reduced image size over each layer. Three additional dense layers were added from the pretrained model output where each layer had 64 hidden units with ReLu activation function. Batch normalization and dropout were applied to all additional layers.

Hyperparameter configurations for the NoduleNet were designed (Table 2). The fixed hyperparameters were required to configure the learning process to reduce the overfitting problem. The maximum epoch indicated to what extension of the learning duration which was set at 1000. Early stopping indicated to what extent the model should keep learning even its performance was not improved, which was set at 20 epochs. The controlled hyperparameters were required to search for the optimum model parameters. Learning rate indicates to what extent of model parameter adjustment over each epoch. The experiment set the learning rate to be 0.0001 and 0.00001. Batch size indicates the number of training samples in each mini-batch which is normally applied with data augmentation in order to increase variation in training samples. This experiment applied randomly horizontal flip and randomly zoom where the zoom ratio is set to 0.2.

Performance evaluation of NoduleNet was reported as classification accuracy, sensitivity, and specificity along with a 95% confidence interval (CI).

Table 1. Dataset Information

Dataset	No. of Images			
	LUNA16		ChestRama	
	Benign	Malignant	Benign	Malignant
Training set	4342	845	-	-
Test set	1340	282	-	-
Validate set	1073	224	15	22
Total	6755	1351	15	22

Figure 1. Data Preprocessing

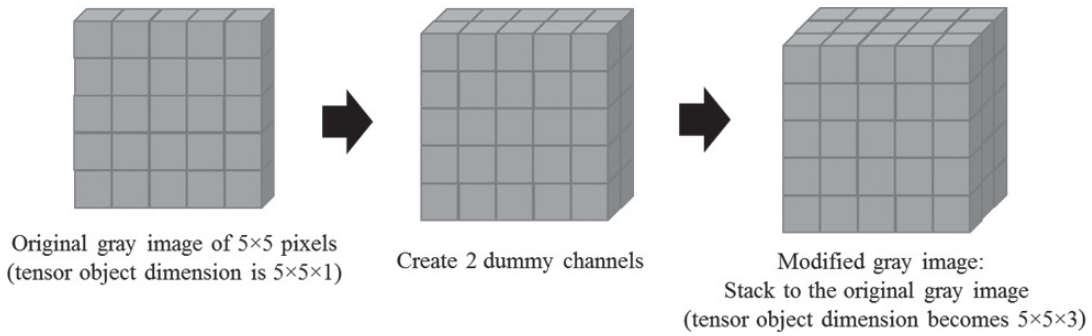


Table 2. Hyperparameter Configuration

Model	Fixed Hyperparameters		Controlled Hyperparameters	
	Max Epoch	Early Stopping	Learning Rate	Batch Size
VGG16	1000	20	0.0001	8
			0.0001	16
			0.0001	32
			0.00001	8
			0.00001	16
			0.00001	32
VGG19	1000	20	0.0001	8
			0.0001	16
			0.0001	32
			0.00001	8
			0.00001	16
			0.00001	32

Results

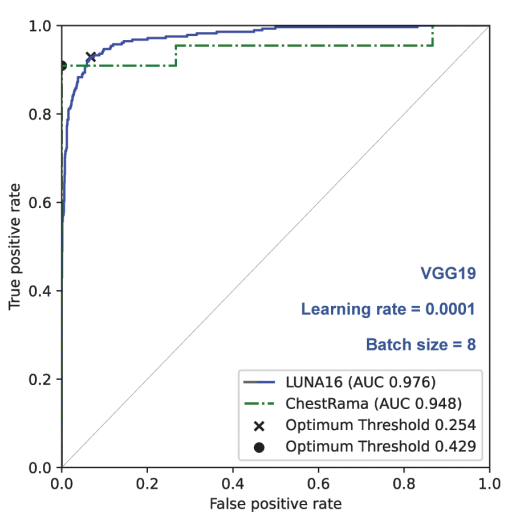
The NoduleNet achieved impressive performance where the bold characters represented the highest performance in each metric. The hyperparameter conficurations were learning rate and batch size. The best model was selected if it gained high sensitivity and specificity, or in other words, low false-positive and false-negative. Internal validation showed that VGG19 was slightly better than VGG16 for all matrices with sensitivity, specificity, and accuracy (performance [95% CI], 0.84 [0.82 - 0.86], 0.97 [0.97 - 0.98], and 0.95 [0.94 - 0.96], respectively). In addition, the external validation showed that VGG19 was very much better than VGG16 for all metrics with corresponding performance (95% CI) values of 0.87 (0.76 - 0.98), 1.00 (1.00 - 1.00), and 0.92 (0.83 - 1.00). Such results indicated that the NoduleNet with VGG19 architecture performed well with Thai patient data and may be able to generalize to other data sets (Table 3).

The receiver operating characteristic (ROC) curve of the best model from this experiments was determined The area under curve (AUC) were reported including the calibrated optimal thresholds. The AUC obtained from internal validation with LUNA16 was 0.976 where the optimal threshold was 0.254. The AUC obtained from external validation with ChestRama was 0.948 where the optimal threshold was 0.429 (Figure 2).

Table 3. Model Performance

Model	Validation	Hyperparameters		Performance (95% CI)		
		Learning Rate	Batch Size	Accuracy	Sensitivity	Specificity
VGG16	Internal	0.0001	8	0.92 (0.9 - 0.93)	0.83 (0.82 - 0.85)	0.94 (0.92 - 0.95)
		0.0001	16	**0.93 (0.92 - 0.94)**	0.86 (0.84 - 0.88)	**0.95 (0.94 - 0.96)**
		0.0001	32	0.91 (0.90 - 0.92)	**0.87 (0.85 - 0.88)**	0.92 (0.91 - 0.93)
		0.00001	8	0.91 (0.89 - 0.92)	0.73 (0.71 - 0.76)	0.94 (0.93 - 0.95)
		0.00001	16	0.92 (0.90 - 0.93)	0.79 (0.77 - 0.81)	0.94 (0.93 - 0.95)
		0.00001	32	0.91 (0.90 - 0.93)	0.83 (0.82 - 0.85)	0.93 (0.92 - 0.94)
	External	0.0001	8	0.73 (0.59 - 0.87)	0.86 (0.75 - 0.97)	0.53 (0.37 - 0.69)
		0.0001	16	0.68 (0.52 - 0.83)	**0.86 (0.75 - 0.97)**	0.40 (0.24 - 0.56)
		0.0001	32	0.59 (0.44 - 0.75)	**0.86 (0.75 - 0.97)**	0.20 (0.07 - 0.33)
		0.00001	8	0.57 (0.41 - 0.73)	0.82 (0.69 - 0.94)	0.20 (0.07 - 0.33)
		0.00001	16	0.65 (0.49 - 0.80)	**0.86 (0.75 - 0.97)**	0.33 (0.18 - 0.49)
		0.00001	32	0.65 (0.49, 0.80)	0.82 (0.69 - 0.94)	0.40 (0.24 - 0.56)
VGG19	Internal	0.0001	8	0.95 (0.94 - 0.96)	0.84 (0.82 - 0.86)	0.97 (0.97 - 0.98)
		0.0001	16	0.94 (0.93 - 0.95)	**0.84 (0.83 - 0.86)**	0.96 (0.95 - 0.97)
		0.0001	32	0.93 (0.92 - 0.95)	0.83 (0.81 - 0.85)	0.96 (0.95 - 0.97)
		0.00001	8	0.91 (0.90 - 0.92)	0.67 (0.64 - 0.69)	0.96 (0.95 - 0.97)
		0.00001	16	0.91 (0.90 - 0.93)	0.79 (0.77 - 0.81)	0.94 (0.93 - 0.95)
		0.00001	32	0.91 (0.90 - 0.93)	0.83 (0.81 - 0.85)	0.93 (0.92 - 0.94)
	External	0.0001	8	0.92 (0.83 - 1.00)	0.87 (0.76 - 0.98)	**1.00 (1.00 - 1.00)**
		0.0001	16	0.95 (0.87 - 1.00)	0.91 (0.82 - 1.00)	1.00 (1.00 - 1.00)
		0.0001	32	0.92 (0.83 - 1.00)	**0.91 (0.82 - 1.00)**	0.93 (0.85 - 1.00)
		0.00001	8	0.84 (0.72 - 0.96)	0.77 (0.64 - 0.91)	0.93 (0.85 - 1.00)
		0.00001	16	0.92 (0.83 - 1.00)	0.86 (0.75 - 0.97)	**1.00 (1.00 - 1.00)**
		0.00001	32	0.89 (0.79 - 0.99)	0.86 (0.75 - 0.97)	0.93 (0.85 - 1.00)

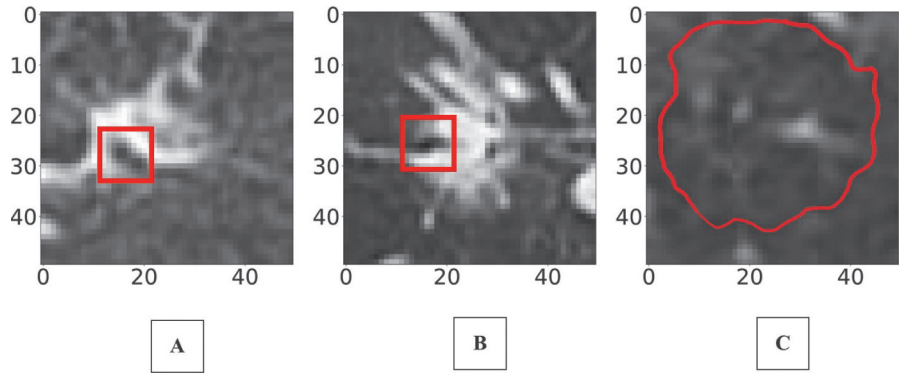
Figure 2. Receiver Operating Characteristics



Discussion

The appearance of each misclassified samples were further investigated and had similar issues of misclassification. They are malignant nodules because the white pixels indicate nodule's solidity. However, the black pixels in the red rectangle could confuse human or machine inference because such a region looks like air bronchogram where the surrounding white region was likely to be parenchymal disease such as infection or either malignancy (Figure 3A and Figure 3B). Another was also malignant nodule but such an image was harder to identify the nodule and hypodensity may be caused by emphysema and less likely to be a malignancy (Figure 3C).

Figure 3. Misclassified Samples



A, 1st misclassified malignant nodule with air bronchogram in the red rectangle.

B, 2nd misclassified malignant nodule with air bronchogram in the red rectangle.

C, 3rd misclassified malignant nodule with hypodensity in the red rectangle.

The NoduleNet assumes the input image as a cropped nodule image that can be either cropped manually by human or automatically cropped by nodule detection algorithm. Such an assumption can be advantage in term of flexibility to integrate with legacy system that does not have nodule detection algorithm yet. However, cropping manually could be prone to human error and eventually, effect to the classification performance. Another limitation is the sequence of images was not addressed. The NoduleNet just considers a snapshot of cropped nodule image while radiologists, in common practice, consider the sequence of images to make conclusions. One possible future direction is to integrate the NoduleNet to existing nodule detection algorithm for fully automated classification thathelps reducing radiologist workload including taking the sequence of cropped nodule images into account to improve classification performance. Another possible future direction is to add more nodule type for generalizing classification tasks.

Although some existing work[25] achieved higher classification accuracy, a comparison could not be made directly because the training data were different. Such existing works used training data from the original datasets (LIDC and IDRI) and extract nodule image but the image resolution did not mention clearly did not mention clearly. Since the NoduleNet learns from the image resolution of 50×50 pixels, classification can be improved if the cropped nodule images have higher resolution.

Conclusions

The NoduleNet, a nodule classifier using deep learning to classify a given cropped nodule image to be either benign or malignant nodules, showed impressive classification accuracy in both internal validation using LUNA16 dataset and external validation using ChestRama dataset. The main contribution was addressed by externally validate the model trained from non-Thai patient data and generalize to Thai patient data. The NoduleNet is flexible enough to integrate with the legacy system even it does not have a nodule detection algorithm.

Acknowledgments

The authors would like to thank the information technology section, Department of Diagnostic and Therapeutic Radiology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, for deidentification and transferring of CT data images; and also acknowledge Chaiyawat Suppasilp, MD, for the interpretation of misclassified samples in an analysis of errors.

References

1. International Agency for Research on Cancer, World Health Organization. *Cancer fact sheets: all cancers*. https://gco.iarc.fr/today/data/factsheets/cancers/39-All-cancers-fact-sheet.pdf. Accessed May 13, 2020.
2. Al Mohammad B, Hillis SL, Reed W, Alakhras M, Brennan PC. Radiologist performance in the detection of lung cancer using CT. *Clin Radiol*. 2019;74(1):67-75. doi:10.1016/j.crad.2018.10.008.
3. Nair A, Screaton NJ, Holemans JA, et al. The impact of trained radiographers as concurrent readers on performance and reading time of experienced radiologists in the UK Lung Cancer Screening (UKLS) trial. *Eur Radiol*. 2018;28(1):226-234. doi:10.1007/s00330-017-4903-z.
4. Jiang H, Ma H, Qian W, Wei G, Zhao X, Gao M. A novel pixel value space statistics map of the pulmonary nodule for classification in computerized tomography images. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2017:556-559. doi:10.1109/EMBC.2017.8036885.
5. El-Baz A, Nitzken M, Vanbogaert E, Gimel'Farb G, Falk R, Abo El-Ghar M. A novel shape-based diagnostic approach for early diagnosis of lung nodules. *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 2011:137-140. doi:10.1109/ISBI.2011.5872373.
6. McNitt-Gray MF, Hart EM, Wyckoff N, Sayre JW, Goldin JG, Aberle DR. A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: preliminary results. *Med Phys*. 1999;26(6):880-888. doi:10.1118/1.598603.
7. Mukherjee J, Chakrabarti A, Shaikh SH, Kar M. Automatic detection and classification of solitary pulmonary nodules from lung CT images. *Proceedings of the International Conference on Emerging Applications of Information Technology*. 2014:294-299. doi:10.1109/EAIT.2014.64.
8. Iwano S, Nakamura T, Kamioka Y, Ishigaki T. Computer-aided diagnosis: A shape classification of pulmonary nodules imaged by high-resolution CT. *Comput Med Imaging Graph*. 2005;29(7):565-570. doi:10.1016/j.compmedimag.2005.04.009.
9. Fernandes VPM, Kanehisa RFA, Braz G, Silva AC, de Paiva AC. Lung nodule classification based on shape distributions. *Proceedings of the ACM Symposium on Applied Computing*. 2016:84-86. doi:10.1145/2851613.2851877.
10. Xie Y, Zhang J, Liu S, Cai W, Xia Y. Lung nodule classification by jointly using visual descriptors and deep features. In: Muller H, Kelm BM, Arbel T, eds. *Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging*. Athens, Greece: Springer-VDI-Verlag GmbH & Co. KG; 2017:116-125. doi:10.1007/978-3-319-61188-4_11.
11. Song J, Liu H, Geng F, Zhang C. Weakly-supervised classification of pulmonary nodules based on shape characters. *Proceedings of the 2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress*. 2016:228-232. doi:10.1109/DASC-PICom-DataCom-CyberSciTec.2016.58.
12. Al-Saffar AAM, Tao H, Talab MA. Review of deep convolution neural network in image classification. *Proceeding ot the International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications*. 2017:26-31. doi:10.1109/ICRAMET.2017.8253139.
13. He T, Zhang Z, Zhang H, Zhang Z, Xie J, Li M. Bag of tricks for image classification with convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019:558-567. doi:10.1109/CVPR.2019.00065.
14. Simonyan K, Zisserman A. Very deep convolutional networks

for large-scale image recognition. *Proceedings of the International Conference on Learning Representations;* 2015. https://arxiv.org/pdf/1409.1556.pdf. Accessed May 19, 2020.

15. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:770-778. doi:10.1109/CVPR.2016.90.
16. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015:1-9. doi:10.1109/CVPR.2015.7298594.
17. Chollet F. Xception: deep learning with depthwise separable convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017:1800-1807. doi:10.1109/CVPR.2017.195.
18. Howard AG, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017. https://arxiv.org/abs/1704.04861v1. Accessed May 19, 2020.
19. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017:2261-2269. doi:10.1109/CVPR.2017.243.
20. Zoph B, Vasudevan V, Shlens J, Le QV. Learning transferable architectures for scalable image recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018:8697-8710. doi:10.1109/CVPR.2018.00907.
21. Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge. *Int J Comput Vis*. 2015;115:211-252. doi:10.1007/s11263-015-0816-y.
22. Setio AAA, Traverso A, de Bel T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Med Image Anal*. 2017;42:1-13. doi:10.1016/j.media.2017.06.015.
23. Armato SG 3rd, McLennan G, Bidaut L, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. *Med Phys*. 2011;38(2):915-931. doi:10.1118/1.3528204.
24. Sajja TK, Devarapalli RM, Kalluri HK. Lung cancer detection based on CT scan images by using deep transfer learning. *Trait Signal*. 2019;36(4):339-344. doi:10.18280/ts.360406.

RMJ Ramathibodi Medical Journal

Rama Med J | Original Article

NoduleNet: การจำแนกก้อนเนื้อในปอดด้วยวิธีการเรียนรู้เชิงลึก

กาเนซ ชาห์[1], รัตน์ชัยนันท์ ธรรมสุจริต[1], อัมรินทร์ ทักขิญเสถียร[1], ฐิติพร สุวัฒนะพงศ์เชฏ[2]

[1] ภาควิชาระบาดวิทยาคลินิกและชีวสถิติ คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล กรุงเทพฯ ประเทศไทย

[2] ภาควิชารังสีวิทยา คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล กรุงเทพฯ ประเทศไทย

บทนำ: การจำแนกก้อนเนื้อ (Nodule) ในปอดได้อย่างแม่นยำตั้งแต่ระยะเริ่มต้นของโรคมะเร็งปอด สามารถช่วยให้แพทย์ตัดสินใจด้านการวางแผนการรักษาได้ดีขึ้น ที่ผ่านมามีการใช้การเรียนรู้เชิงลึกมาช่วยในการตรวจจับและจำแนกก้อนเนื้อในปอด แต่ยังขาดการนำมาประเมินความถูกต้องกับข้อมูลของคนไข้ไทย

วัตถุประสงค์: เพื่อสร้างตัวแบบชื่อ NoduleNet ด้วยวิธีการเรียนรู้เชิงลึกสำหรับการจำแนกก้อนเนื้อในปอดและนำมาประเมินความถูกต้องกับข้อมูลคนไข้ไทย

วิธีการศึกษา: ข้อมูลชุดแรก (LUNA16) ซึ่งเป็นข้อมูลสาธารณะถูกนำมาใช้สอนตัวแบบด้วยการเรียนรู้เชิงลึก ส่วนข้อมูลชุดที่สอง (ChestRama) ถูกนำมาใช้เพื่อการประเมินความถูกต้องของตัวแบบที่สร้างจากตัวแบบที่เคยผ่านการเรียนรู้มาก่อน (VGG16 และ VGG19) ผ่านวิธีการเรียนรู้แบบถ่ายทอด (Transfer learning) ซึ่งเป็นรูปแบบหนึ่งของการเรียนรู้เชิงลึก

ผลการศึกษา: NoduleNet สามารถทำงานได้ดีโดยสมรรถนะจากความตรงภายใน (Internal validation) มีค่าความแม่นยำเท่ากับ 0.95 (0.94 - 0.96) ความไวเท่ากับ 0.84 (0.82 - 0.86) ความจำเพาะเท่ากับ 0.97 (0.97 - 0.98) และความตรงภายนอก (External validation) มีค่าความแม่นยำเท่ากับ 0.95 (0.87 - 1.00) ความไวเท่ากับ 0.91 (0.82 - 1.00) และความจำเพาะเท่ากับ 1.00 (1.00 - 1.00) โดยมี 3 ข้อมูลตัวอย่างจาก ChestRama เท่านั้นที่ทำนายคลาดเคลื่อนเป็นผลลบปลอม (False-negative)

สรุป: NoduleNet ที่สร้างขึ้นโดยการเรียนรู้จากข้อมูลผู้ป่วยที่ไม่ใช่คนไทยสามารถขยายผลไปสู่ข้อมูลผู้ป่วยที่เป็นคนไทยได้ดี นอกจากนี้ยังสามารถพัฒนาต่อยอดในอนาคตด้วยการควบรวมเข้ากับอัลกอริธึมการตรวจจับแบบอัตโนมัติเพื่อค้นหาตำแหน่งก้อนเนื้อในปอดสำหรับการจำแนกชนิดของก้อนเนื้อในปอด

คำสำคัญ: การเรียนรู้เชิงลึก การเรียนรู้แบบถ่ายทอด การจำแนกก้อนเนื้อในปอด

Rama Med J: doi:10.33165/rmj.2020.43.4.241727

Received: May 29, 2020 **Revised:** October 7, 2020 **Accepted:** October 30, 2020

Corresponding Author:
รัตน์ชัยนันท์ ธรรมสุจริต
ภาควิชาระบาดวิทยาคลินิก
และชีวสถิติ
คณะแพทยศาสตร์
โรงพยาบาลรามาธิบดี
มหาวิทยาลัยมหิดล
270 ถนนพระรามที่ 6
แขวงทุ่งพญาไท เขตราชเทวี
กรุงเทพฯ 10400 ประเทศไทย
โทรศัพท์ +66 2201 1269
อีเมล ratchainant.tha@mahidol.edu