RESEARCH ARTICLE

Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning

Mizuho Nishio1,2*, Osamu Sugiyama2, Masahiro Yakami1,2, Syoko Ueno3, Takeshi Kubo1, Tomohiro Kuroda4, Kaori Togashi1

1 Department of Diagnostic Imaging and Nuclear Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, 2 Preemptive Medicine and Lifestyle-related Disease Research Center, Kyoto University Hospital, Kyoto, Japan, 3 Department of Social Informatics, Kyoto University Graduate School of Informatics Yoshidakonmachi, Kyoto, Japan, 4 Division of Medical Information Technology and Administrative Planning, Kyoto University Hospital, Kyoto, Japan

* nmizuho@kuhp.kyoto-u.ac.jp, jurader@yahoo.co.jp

Abstract

We developed a computer-aided diagnosis (CADx) method for classification between benign nodule, primary lung cancer, and metastatic lung cancer and evaluated the following: (i) the usefulness of the deep convolutional neural network (DCNN) for CADx of the ternary classification, compared with a conventional method (hand-crafted imaging feature plus machine learning), (ii) the effectiveness of transfer learning, and (iii) the effect of image size as the DCNN input. Among 1240 patients of previously-built database, computed tomography images and clinical information of 1236 patients were included. For the conventional method, CADx was performed by using rotation-invariant uniform-pattern local binary pattern on three orthogonal planes with a support vector machine. For the DCNN method, CADx was evaluated using the VGG-16 convolutional neural network with and without transfer learning, and hyperparameter optimization of the DCNN method was performed by random search. The best averaged validation accuracies of CADx were 55.9%, 68.0%, and 62.4% for the conventional method, the DCNN method with transfer learning, and the DCNN method without transfer learning, respectively. For image size of 56, 112, and 224, the best averaged validation accuracy for the DCNN with transfer learning were 60.7%, 64.7%, and 68.0%, respectively. DCNN was better than the conventional method for CADx, and the accuracy of DCNN improved when using transfer learning. Also, we found that larger image sizes as inputs to DCNN improved the accuracy of lung nodule classification.
Introduction

Computer-aided diagnosis refers to software that helps clinicians to diagnose disease, and it has the potential to optimize clinicians’ workloads [1,2,3–7]. Computer-aided diagnosis can be divided into software that detects lesions (CADe, computer-aided detection) and software that classifies lesions (CADx, computer-aided diagnosis). However, for CADe or CADx to assist clinicians effectively, they must perform reliable and efficient image recognition. If a method that can better recognize an image is applied to computer-aided diagnosis, its performance can be improved.

Lung cancers are the leading cause of cancer-related death in the United States because they are frequently diagnosed at an advanced stage [8]. Results from the National Lung Screening Trial showed that lung cancer screening by computed tomography (CT) has significantly reduced lung cancer mortality among heavy smokers, but that false positives were problematic, accounting for 96.4% of positive screening results [9]. Another study has indicated that CADe might help radiologists to detect missed lung cancers on CT screening by assisting with image interpretation [7]. Experience with CADe suggests that CADx might help reduce the number of false positives identified by CT during lung cancer screening.

Deep learning is a new technique that is overtaking conventional methods of computer vision, such as hand-crafted imaging feature plus machine learning, and is increasingly being used in CAD [10]. Deep convolutional neural network (DCNN) has attracted the attention of researchers since its introduction in 2012 at the IMAGENET Large Scale Visual Recognition Challenge [11]. The DCNN method has continued to improve, and it has been shown that image recognition by DCNN was identical or superior to that by humans in general object recognition [12].

Many studies have used DCNN to improve the performance of CAD [10,13–20,21]. Several studies have also proposed the use of DCNN-based CAD for lung nodules. For example, Teramoto et al. proposed that use of DCNN in CADe could reduce the false positive rate in positron emission tomography/CT images of lung nodules [21]. The results of Ciompì et al. also show that DCNN was useful for CADx, helping to classify lung nodules into six types [19].

In the current study, we focused on developing CADx by DCNN for lung nodules. Our aim was to evaluate the following: (i) the usefulness of DCNN for CADx compared with conventional methodology (i.e. hand-crafted imaging feature plus machine learning), (ii) the effectiveness of transfer learning, and (iii) the effect of image size as an input to DCNN.

Methods

This retrospective study was approved by the ethical committee of Kyoto University Hospital, which waived need for informed consent. We used a database which were built for previous research of CADx [4,22]. Because the previous studies focused on CADx without DCNN, the purpose of the current study is different from those of the previous studies.

CT image database

The database contained the CT images and clinical information of 1240 patients who had at least one lung nodule. The CT images were acquired using a 320-detector-row or a 64-detector-row CT scanner (Aquilion ONE or Aquilion 64; Toshiba Medical Systems, Otawara, Japan). CT scan parameters were as follows: tube current, 109 ± 53.3 mA (range, 25–400 mA); gantry rotation time, 0.500 ± 0.0137 s (range, 0.400–1.00 s); tube potential, 120 ± 1.69 kV (range, 120–135 kV); matrix size, 512 × 512 and slice thickness, 1 or 0.5 mm. Lung nodules diagnosed as benign nodules, primary lung cancers, or metastatic lung cancers were selected.
and the CT images, final diagnosis, and nodule positions of these nodules were used for development and evaluation of CADx.

**Image pre-processing**

The CT images were loaded, and their voxel sizes converted to $1 \times 1 \times 1$ mm. In each case, because the position of the center of the lung nodule was available, the CT images including the lung nodule were cropped with a volume of interest set to $64 \times 64 \times 64$ mm (voxels). The cropped CT images were then input for CADx.

**Conventional CADx**

From the cropped CT images, feature extraction was performed by rotation-invariant uniform-pattern local binary pattern on three orthogonal planes (LBP-TOP) [23,24,25], which has been successfully used for CADx of lung nodules [3]. The results of LBP-TOP were fed to support vector machine (SVM) with kernel trick (radial basis function) [26]. LBP-TOP had two hyperparameters ($LBP_R$ and $LBP_P$), and SVM had two hyperparameters ($C$ and $\gamma$).

**CADx by DCNN with and without transfer learning**

To utilize DCNN for 2D images (2D-DCNN), the 3D cropped CT images were converted to 2D images. Three orthogonal planes (axial, coronal, and sagittal) were set on the center of the 3D images, and 2D images ($64 \times 64$) in the three orthogonal planes were extracted. At extraction, the sizes of 2D images were converted to $L \times L$, where $L$ was set to 56, 112, or 224. With this image processing, each lung nodule was represented as the three 2D images (size = $L \times L$). We referred to a pair of these 2D images and the corresponding final diagnosis as a batch. Before feeding batches to DCNN, the pixel value range of the 2D images was changed from $-1000, 1000$ to $-1, 1$ by the transformation $y = x/1000$, where $x$ and $y$ were the pixel value before and after the transformation, respectively.

The architecture of 2D-DCNN in our CADx was derived from VGG-16 convolutional neural network [27], which was modified to perform transfer learning (Fig 1). First, fully-connected (FC) layers of VGG-16 were removed, and a new FC layer was added, whose number of units was denoted by $F$. Next, an FC layer with three units, whose output would be converted to a probability of the three classes, was added as the prefinal DCNN layer. Dropout was applied between the two FC layers, with strength denoted by $D$ (0 = no dropout; 1 = full dropout and no connection between the two FC layers). We then used rectified linear units as the activation function of the FC layer with $F$ units. To convert the output of the FC layer with three units to a probability of the three classes, a softmax layer was used. For transfer learning, we used VGG-16 parameters pretrained with IMAGENET [11] and finetuned by stochastic gradient descent. The initial learning rate of stochastic gradient descent was represented as $R$. Parameter finetuning was not performed in several VGG-16 layers, and the number of layers without finetuning is represented by $V$. In CADx by DCNN without transfer learning, training was performed without VGG-16 parameters pretrained with IMAGENET. Data augmentation was performed for 2D-DCNN training. Hyperparameters of 2D-DCNN were summarized in Supporting Information.

**Statistical analysis**

We used 1113 training cases for learning and 123 validation cases for performance evaluation, which did not overlap. Validation loss and validation accuracy were calculated 10 times with the same CADx hyperparameters [19]; splitting of the training and validation sets was random.
each time. The averaged values for validation loss and validation accuracy were obtained for each set of hyperparameters and were used to evaluate the performance. For the conventional method, we selected the best LBP-TOP and SVM hyperparameters by grid search [28]. For the DCNN method, we performed random search to optimize the hyperparameters [29]. The detail of random search was described in Supporting Information.

Results

For benign nodules, primary lung cancers, and metastatic lung cancers, the following number of lung nodules were selected from the database for development and evaluation of CADx: benign nodules, n = 412; primary lung cancers, n = 571; and metastatic lung cancers, n = 253. Four lung nodules were excluded because they did not fit one of these three types (for example, carcinoid). All diagnoses of primary lung cancer were confirmed pathologically. Benign nodules were primarily confirmed by stability or shrinkage on repeat CT scans over a 2-year follow-up period, but 57 were also diagnosed pathologically. Most of the metastatic lung cancers were diagnosed radiologically and clinically, and the diagnosis of 90 metastatic lung cancers was confirmed pathologically. As shown in Table 1, mean and standard deviation of size of these lung nodules were 20.52 ± 10.22 mm.

The current study included 709 men and 527 women, and the patient demographics of these 1236 patients are shown in Table 1. Mean and standard deviation of patient age and smoking history (Brinkman Index) was 65.76 ± 12.65 and 605.1 ± 774.2, respectively. Their smoking status was as follows: current smoker, n = 266; ex-smoker, n = 456; and never smoker, n = 514. Previous history of malignant tumor was confirmed in 545 patients. Contrast-enhanced CT was performed in 531 patients.

Fig 2 shows representative CT images of a benign nodule, a primary lung cancer, and a metastatic lung cancer. Fig 3 shows three representative CT images of a lung nodule obtained from the three orthogonal planes and used as the input to 2D-DCNN.
Table 1. Summary of patient demographics.

| Variables                      | All       | Benign nodule | Primary lung cancer | Metastatic lung cancer |
|--------------------------------|-----------|---------------|---------------------|------------------------|
|                                | Mean      | SD            | Mean                | SD                     |
| N                              | 1236      | 412           | 571                 | 253                    |
| Age (y)                        | 65.76     | 12.65         | 64.81               | 13.80                  |
| Sex (number of men)            | 709       | 237           | 331                 | 141                    |
| Smoking history (Brinkman Index)| 605.1     | 774.2         | 543.8               | 747.7                  |
| Smoking status                 |           |               |                     |                        |
| Current smoker                 | 266       | 70            | 151                 | 45                     |
| Ex-smoker                      | 456       | 161           | 219                 | 76                     |
| Never smoker                   | 514       | 181           | 201                 | 132                    |
| Previous history of malignant tumor | 545      | 148           | 144                 | 253                    |
| Nodule size (mm)               | 20.52     | 10.22         | 18.28               | 8.54                   |
| Contrast-enhanced CT           | 531       | 113           | 287                 | 131                    |

Because Brinkman index was not clearly described in 20 patients, Mean and SD of brinkman index were calculated without these 20 patients. Abbreviation: SD, standard deviation; CT, computed tomography.

https://doi.org/10.1371/journal.pone.0200721.t001

Fig 2. Representative CT images of lung nodules. (A) benign nodule, (B) primary lung cancer and (C) metastatic lung cancer.

https://doi.org/10.1371/journal.pone.0200721.g002

Fig 3. Three CT images obtained from three orthogonal planes used for input to 2D-DCNN. Fig 2(B) is identical to Fig 3(A). (A) axial image, (B) coronal image and (C) sagittal image. Abbreviations: DCNN, deep convolutional neural network.

https://doi.org/10.1371/journal.pone.0200721.g003
The best averaged validation accuracy for the conventional method was 55.9%, and the following optimal hyperparameters were used: $LBP_R = 4$, $LBP_P = 40$, $C = 1024$, and $\gamma = 4$. Table 2 shows validation loss, validation accuracy, and the optimal hyperparameters for $L$ values of 56, 112, and 224 for CADx by DCNN with transfer learning. The best averaged validation loss and validation accuracy for DCNN with transfer learning were, respectively, as follows: 0.822 and 60.7% when $L = 56$; 0.783 and 64.7% when $L = 112$; and 0.774 and 68.0% when $L = 224$. Table 2 also shows validation loss, validation accuracy, and the optimal hyperparameters for $L$ values of 56, 112, and 224 for DCNN without transfer learning. The best averaged validation loss and validation accuracy for DCNN without transfer learning were, respectively, as follows: 0.843 and 60.2% when $L = 56$; 0.824 and 62.4% when $L = 112$; and 0.860 and 58.9% when $L = 224$. The raw results for optimal CADx with DCNN are shown in Supporting Information, as are the averaged validation loss and validation accuracy data in all trials of random search.

Figs 4 and 5 show representative results for loss and accuracy during DCNN training with and without transfer learning, respectively. Tables 3 and 4 show the corresponding confusion matrices between true labels and predicted labels obtained from CADx by DCNN with and without transfer learning, respectively. In addition, averaged confusion matrix was shown in Table 5, where the best averaged validation accuracy (68.0%) was obtained.

### Discussion

The current results show that CADx of the ternary classification (benign nodule, primary lung cancer, and metastatic lung cancer) was better when using DCNN than when using the conventional method, and that transfer learning improved image recognition with the DCNN method. In addition, larger image sizes as inputs to DCNN improved the accuracy of lung nodule classification.

The averaged validation accuracies of CADx were 68.0% and 55.9% by the DCNN and conventional methods, respectively. These results confirm that DCNN was more useful for the CADx of lung nodules. While a major advantage of DCNN is that its performance for image recognition is superior to the conventional method, disadvantages are (i) that it is difficult to train because it frequently leads to overfitting and (ii) that large-scale data are needed for effective training. To prevent overfitting, we therefore used transfer learning to provide better diagnostic accuracy for lung nodules. We speculated that transfer learning was effective because our database was medium-scale (>1000 lung nodules).

The previous study [4] evaluated the performance of CADx without DCNN using the data for 1000 lung nodules obtained from our database. The study produced classification validation loss and validation accuracy were calculated 10 times with the same CADx hyperparameters, and their averaged values were shown. Abbreviations: CADx, computer-aided diagnosis; DCNN, deep convolutional neural network; TF, transfer learning.

Table 2. Optimal hyperparameters and classification results for CADx by DCNN with and without transfer learning.

| Type         | $L$ | $E$  | $R$  | $V$   | $F$  | $D$   | Validation Accuracy (%) | Validation Loss |
|--------------|-----|------|------|-------|------|-------|--------------------------|----------------|
| DCNN with TF |     |      |      |       |      |       |                          |                |
| 56           | 20  | 0.000002 | 4    | 384   | 0.6  | 60.7  | 0.822                    |                |
| 112          | 20  | 0.000002 | 11   | 384   | 0.4  | 64.7  | 0.783                    |                |
| 224          | 20  | 0.000002 | 11   | 384   | 0.4  | 68.0  | 0.774                    |                |
| DCNN without TF |     |      |      |       |      |       |                          |                |
| 56           | 30  | 0.000007 | 0    | 384   | 0.6  | 60.2  | 0.843                    |                |
| 112          | 25  | 0.0001   | 0    | 384   | 0.4  | 62.4  | 0.824                    |                |
| 224          | 15  | 0.0001   | 0    | 384   | 0.4  | 58.9  | 0.860                    |                |

https://doi.org/10.1371/journal.pone.0200721.t002
accuracies of 57.7% and 61.3% based on the conventional method and their proposed method (feature vectors calculated based on radiological findings), respectively. Because we used different methods for evaluating CADx performance, it was difficult to directly compare the performance with that of the previous study. However, according to both studies, the accuracy of CADx with the conventional method was nearly 60% for our database.

According to Litjens et al. [10], few studies have performed a thorough investigation of whether transfer learning gives better results for medical image analysis. Indeed, the results of two studies have left controversy about the efficacy of transfer learning [30,31]. By contrast, another two studies have shown that transfer learning with Google's Inception v3 architecture can achieve diagnostic accuracy to expert human level in dermatology and ophthalmology [32,33]. In conjunction with the results of our study, CADx with transfer learning should improve diagnostic accuracy provided sufficient training data are used.

It was notable that image size ($L$) affected the accuracy of CADx by DCNN. Although image size is a simple factor, its effect on the accuracy of CADx was large in our study. Similar results were obtained in the previous study, where slice thickness of CT images could affect the detectability of CADe [34]. We speculated that, because VGG-16 was originally pretrained with an image size of $224 \times 224$, the best accuracy was obtained by finetuning VGG-16 with 2D CT images of the same size in our study. In the review of CAD by Litjens et al. [10], it was suggested that the exact architecture of deep learning was not the most important determinant.
Fig 5. Representative results of loss and accuracy during DCNN training without transfer learning. Abbreviations: DCNN, deep convolutional neural network.

Table 3. Representative result of confusion matrix between true labels and predicted labels by DCNN with transfer learning.

| True label          | Predicted label                  |
|---------------------|----------------------------------|
|                     | Benign nodule | Primary lung cancer | Metastatic lung cancer |
| Benign nodule       | 22            | 8                   | 5                     |
| Primary lung cancer | 6             | 46                  | 8                     |
| Metastatic lung cancer | 5         | 4                   | 19                    |

Because splitting of the training and validation sets was random each time, the ratio between the 3 classes was variable. This confusion matrix corresponds to the results of Fig 4. Abbreviations: DCNN, deep convolutional neural network.

Table 4. Representative result of confusion matrix between true labels and predicted labels by DCNN without transfer learning.

| True label          | Predicted label                  |
|---------------------|----------------------------------|
|                     | Benign nodule | Primary lung cancer | Metastatic lung cancer |
| Benign nodule       | 15            | 19                  | 6                     |
| Primary lung cancer | 10            | 43                  | 6                     |
| Metastatic lung cancer | 6           | 3                   | 15                    |

Because splitting of the training and validation sets was random each time, the ratio between the 3 classes was variable. This confusion matrix corresponds to the results of Fig 5. Abbreviations: DCNN, deep convolutional neural network.
of a good solution, and that data pre-processing or augmentation based on expert knowledge about the task could provide advantages beyond simply adding more layers to DCNN. Our results also show that a pre-processing step, such as adjusting the image size, should be performed carefully to obtain accurate results from CADx.

We developed a CADx method which classifies lung nodules into benign nodules, primary lung cancer, or metastatic lung cancer. A Lung CT Reporting and Data System (Lung-RADS) has been proposed for estimating lung cancer risk and the optimal follow-up strategy based on nodule-specific characteristics (i.e. nodule type, nodule size) [35]. Ciompi et al. developed CADx with DCNN for classifying the nodule type based on Lung-RADS [19]. However, although the nodule type is an important factor when evaluating lung cancer risk, it is not directly associated with pathological or clinical diagnosis. In contrast to this, our CADx method using DCNN can directly output the probabilities of the three classifications and would be more useful for clinicians than CADx which classifies nodule type.

Both our database and that of The Lung Image Database Consortium and Image Database Resource Initiative (LIDC/IDRI) [36] contain in excess of 1000 cases and CT images. However, clinical diagnostic results are only partially available in the LIDC/IDRI database. Few studies exist in which CADx was performed by DCNN with directly outputted probabilities of disease classification. We built our database to include both clinical diagnosis and radiological image findings [22].

There were several limitations to our study. First, we ignored all nodule-specific features, such as nodule size and type. The results of a previous study [4] show that CADx using radiological findings provided better results; given this, utilizing radiological findings may improve DCNN-based CADx. We hope that our study could serve as a basis for further exploration of CADx based on lung nodule characteristics. Second, we used 2D-DCNN for the CADx of lung nodules. Through image pre-processing, the 3D CT images of the lung nodules were converted to 2D CT images in three orthogonal planes, which greatly reduced the computational burden for DCNN training and testing. We focused on 2D-DCNN in the present study because it was difficult to perform transfer learning with 3D-DCNN on medical image analysis. We will attempt 3D-DCNN for CADx of lung nodules in a future study. Third, we only investigated the effect of smaller image sizes ($L \leq 224$) because the computational cost precluded the evaluation of larger images. Given that the performance of graphic processing units has increased since the study inception, we expect to be able to evaluate the effect of larger image sizes in a future study.

In conclusion, the 2D-DCNN method was more useful for ternary classification of lung nodule than the conventional method for CADx, and transfer learning enhanced the image recognition for CADx by DCNN when using medium-scale training data. In addition, our

| True label          | Predicted label          |
|---------------------|--------------------------|
| Benign nodule       | Benign nodule 19.9        |
|                     | Primary lung cancer 12.6  |
|                     | Metastatic lung cancer 7.2|
| Primary lung caner  | Benign nodule 8.4         |
|                     | Primary lung cancer 43.5  |
|                     | Metastatic lung cancer 4.1|
| Metastatic lung caner| Benign nodule 4.6         |
|                     | Primary lung cancer 2.5   |
|                     | Metastatic lung cancer 20.2|

Because splitting of the training and validation sets was random each time, the ratio between the 3 classes was variable. This averaged confusion matrix was calculated from the 10 sets of classification results of 123 validation cases at the optimal hyperparameters. The validation accuracy of this confusion matrix was 68.0%.

Abbreviations: DCNN, deep convolutional neural network.

https://doi.org/10.1371/journal.pone.0200721.t005
results show that larger image sizes as inputs to DCNN improved the accuracy of lung nodule classification.

**Supporting information**

S1 Table. Raw results of CADx by DCNN with transfer learning in optimal hyperparameters. Raw results of CADx by DCNN with transfer learning in optimal hyperparameters. (XLSX)

S2 Table. Raw results of CADx by DCNN without transfer learning in optimal hyperparameters. Raw results of CADx by DCNN without transfer learning in optimal hyperparameters. (XLSX)

S3 Table. Averaged validation loss and validation accuracy of CADx by DCNN with transfer learning in all trials of random search. Averaged validation loss and validation accuracy of CADx by DCNN with transfer learning in all trials of random search. (XLSX)

S4 Table. Averaged validation loss and validation accuracy of CADx by DCNN without transfer learning in all trials of random search. Averaged validation loss and validation accuracy of CADx by DCNN without transfer learning in all trials of random search. (XLSX)

S1 File. Detail of conventional CADx and CADx by DCNN. Detail of conventional CADx and CADx by DCNN. (DOCX)

**Acknowledgments**

This study was supported by JSPS KAKENHI (Grant Number JP16K19883). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Author Contributions**

**Conceptualization:** Mizuho Nishio.
**Data curation:** Masahiro Yakami, Takeshi Kubo.
**Formal analysis:** Mizuho Nishio.
**Funding acquisition:** Mizuho Nishio.
**Investigation:** Mizuho Nishio.
**Methodology:** Mizuho Nishio, Osamu Sugiyama, Syoko Ueno.
**Project administration:** Mizuho Nishio.
**Resources:** Mizuho Nishio.
**Software:** Mizuho Nishio.
**Supervision:** Tomohiro Kuroda, Kaori Togashi.
**Validation:** Mizuho Nishio.
**Visualization:** Mizuho Nishio.
Writing – original draft: Mizuho Nishio.
Writing – review & editing: Mizuho Nishio, Osamu Sugiyama, Masahiro Yakami, Syoko Ueno, Takeshi Kubo, Tomohiro Kuroda, Kaori Togashi.

References
1. Suzuki K. Computer-Aided Detection of Lung Cancer. In: Image-Based Computer-Assisted Radiation Therapy. Singapore: Springer Singapore; 2017:9–40. https://doi.org/10.1007/978-981-10-2945-5_2
2. El-Baz A, Beache GM, Gimel'farb G, Suzuki K, Okada K, Elnakib A, et al. Computer-aided diagnosis systems for lung cancer: challenges and methodologies. Int J Biomed Imaging. 2013; 2013:942353. https://doi.org/10.1155/2013/942353 PMID: 23431282
3. Nishio M, Nagashima C. Computer-aided Diagnosis for Lung Cancer: Usefulness of Nodule Heterogeneity. Acad Radiol. 2017; 24(3):328–336. https://doi.org/10.1016/j.acra.2016.11.007 PMID: 28110797
4. Kawagishi M, Chen B, Furukawa D, Sekiguchi H, Sakai K, Kubo T, et al. A study of computer-aided diagnosis for pulmonary nodule: comparison between classification accuracies using calculated image features and imaging findings annotated by radiologists. Int J Comput Assist Radiol Surg. 2017; 12(5):767–776. https://doi.org/10.1016/j.ijacrs.2016.09.023 PMID: 28526968
5. de Oseas Carvalho Filho A, Corrêa Silva A, de Cardoso Paiva A, Acatauas Nunes R, Gattass M, Acatauassu Nunes R. Computer-Aided Diagnosis of Lung Nodules in Computed Tomography by Using Phylogenetic Diversity, Genetic Algorithm, and SVM. J Digit Imaging. 2017. https://doi.org/10.1007/s10278-017-9973-6 PMID: 28526968
6. Nomura Y, Higaki T, Fujita M, Miki S, Awaysa Y, Nakanishi T, et al. Effects of Iterative Reconstruction Algorithms on Computer-aided Detection (CAD) Software for Nodules in Ultra-low-dose CT for Lung Cancer Screening. Acad Radiol. 2017; 24(2):124–130. https://doi.org/10.1016/j.acra.2016.09.023 PMID: 27986507
7. Liang M, Tang W, Xu DM, Jirapatnakul AC, Reeves AP, Henschke CI, et al. Low-Dose CT Screening for Lung Cancer: Computer-aided Detection of Missed Lung Cancers. Radiology. 2016; 281(1):279–288. https://doi.org/10.1148/radiol.2016150063 PMID: 27019363
8. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016; 66(1):7–30. https://doi.org/10.3322/caac.21332 PMID: 26742998
9. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. N Engl J Med. 2011. https://doi.org/10.1056/NEJMoa1102873 PMID: 21714641
10. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A Survey on Deep Learning in Medical Image Analysis. arXiv. 2017; 1702.05747(1995):1–34. https://doi.org/10.1016/j.media.2017.07.005 PMID: 28778026
11. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis. 2015; 115(3):211–252. https://doi.org/10.1007/s11263-015-0816-y
12. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing Human-Level Performance on ImageNet Classification. 2015. 10.1109/ICCV.2015.123.
13. Liu S, Zheng H, Feng Y, Li W. Prostate Cancer Diagnosis using Deep Learning with 3D Multiparametric MRI. SPIE Med Imaging. 2017; 10134:1–4. https://doi.org/10.1117/12.2277121
14. Roth HR, Lu L, Liu J, Yao J, Self A, Cherry K, et al. Improving Computer-aided Detection using Convolutional Neural Networks and Random View Aggregation. 2015:1–12. https://doi.org/10.1109/TMI.2015.2482920 PMID: 26441412
15. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging. 2016; 35(5):1285–1298. https://doi.org/10.1109/TMI.2015.2528162 PMID: 26886976
16. Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks. Int J Comput Assist Radiol Surg. 2017. https://doi.org/10.1007/s11548-017-1605-6 PMID: 28501942
17. Hussein S, Gillies R, Cao K, Song Q, Bagci U. TumorNet: Lung Nodule Characterization Using Multi-View Convolutional Neural Network with Gaussian Process. 2017. http://arxiv.org/abs/1703.00645.
18. Yu-Jen Chen Y-J, Hua K-L, Hsu C-H, Cheng W-H, Hidayati SC. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. Onco Targets Ther. 2015:2015. https://doi.org/10.2147/OTT.S80733 PMID: 26346558
19. Ciompi F, Chung K, van Riel SJ, Setio AAA, Gerke PK, Jacobs C, et al. Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Sci Rep. 2017; 7(April):46479. https://doi.org/10.1038/srep46479 PMID: 28422152
20. Anirudh R, Thiagarajan JJ, Bremer T, Kim H. Lung Nodule Detection using 3D Convolutional Neural Networks Trained on Weakly Labeled Data. *SPIE Med Imaging*. 2015. https://doi.org/10.1117/12.2214876

21. Teramoto A, Fujita H, Yamamuro O, Tamaki T. Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique. *Med Phys*. 2016; 43(6):2821–2827. https://doi.org/10.1118/1.4948498 PMID: 27277030

22. AOYAMA G, KUBO T, SAKAMOTO R, YAKAMI M, FUJIMOTO K, EMOTO Y, et al. Integrated Lung Nodule Database Consisting of CT Images, Structured Imaging Findings and Clinical Information. *Med Imaging Technol*. 2016; 34(5):267–278.

23. Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit*. 1996; 29(1):51–59. https://doi.org/10.1016/0031-3203(95)00067-4

24. Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell*. 2002; 24(7):971–987. https://doi.org/10.1109/TPAMI.2002.1017623

25. Zhao G, Pietikainen M. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions. *IEEE Trans Pattern Anal Mach Intell*. 2007; 29(6):915–928. https://doi.org/10.1109/TPAMI.2007.1110 PMID: 17431293

26. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol*. 2011; 2(3):1–27. https://doi.org/10.1145/1961189.1961199

27. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv Prepr*. 2014:1–10.

28. Hsu C-W, Chang C-C, Lin C-J. A Practical Guide to Support Vector Classification. http://www.csie.ntu.edu.tw/. Accessed February 4, 2018.

29. Bergstra J, Bengio Y. Random Search for Hyper-Parameter Optimization. *J Mach Learn Res*. 2012; 13(Feb):281–305. http://www.jmlr.org/papers/v13/bergstra12a.html. Accessed June 29, 2017.

30. Antony J, McGuinness K, O’Connor NE, Moran K. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In: 2016 23rd International Conference on Pattern Recognition (ICPR). IEEE; 2016. 10.1109/ICPR.2016.7899799.

31. Kim E, Corte-Real M, Baloch Z. A deep semantic mobile application for thyroid cytopathology. In: Zhang J, Cook TS, eds. SPIE Medical Imaging 2016. Vol 9789. International Society for Optics and Photonics; 2016:97890A. 10.1117/12.2216468.

32. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*. 2016; 316(22):2402. https://doi.org/10.1001/jama.2016.17216 PMID: 27898976

33. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017; 542(7639):115–118. https://doi.org/10.1038/nature21056 PMID: 28117445

34. Narayanan BN, Hardie RC, Kebede TM. Performance analysis of a computer-aided detection system for lung nodules in CT at different slice thicknesses. *J Med Imaging*. 2018; 5(1):1. https://doi.org/10.1117/1.JMI.5.1.014504 PMID: 29487880

35. Lung CT Screening Reporting and Data System (Lung-RADS). https://www.acr.org/Quality-Safety/Resources/LungRADS. Accessed August 20, 2017.

36. Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, 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. https://doi.org/10.1118/1.3528204 PMID: 21452728