Differentiating novel coronavirus pneumonia from general pneumonia based on machine learning

Chenglong Liu
University of Shanghai for Science and Technology https://orcid.org/0000-0003-1392-4600
Xiaoyang Wang
Ruian City People's Hospital
Chenbin Liu
Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Radiation Oncology, Shenzhen
Qingfeng Sun
Ruian City People's Hospital
Wenxian Peng (pengwx@sumhs.edu.cn)

Research

Keywords: Machine learning, Novel coronavirus pneumonia, Common pneumonia, Chest CT

Posted Date: May 29th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-31313/v1

Version of Record: A version of this preprint was published on August 19th, 2020. See the published version at https://doi.org/10.1186/s12938-020-00809-9.

Abstract

Background

Chest CT screening as supplementary means is crucial in diagnosing novel coronavirus pneumonia (COVID-19) with high sensitivity and popularity. Machine learning was adept in discovering intricate structures from CT images and achieved expert-level performance in medical image analysis.

Methods

To develop and validate an integrated machine learning framework on chest CT images for differentiating COVID-19 from common pneumonia (CP). Seventy-three confirmed COVID-19 cases were consecutively enrolled together with twenty-seven confirmed common pneumonia patients from Ruian People's Hospital, from January 2020 to March 2020. Statistical textual features of COVID-19 and CP images were extracted. After feature selection, the reserved features were applied to the ensemble of bagged tree (EBT) and four other machine learning classifiers with 10-fold cross-validation.

Results

The classification accuracy, precision, sensitivity, specificity and F1 score of our proposed method are 91.66%, 97.91%, 85.26%, 98.15% and 91.15% respectively. The AUC of its receiver operating characteristic is 0.98.

Conclusions

The experimental results indicate that the EBT algorithm with statistical textural features on chest CT for differentiating COVID-19 from common pneumonia achieved high transferability, efficiency, specificity, and impressive accuracy.

Background

Since the first COVID-19 case was discovered in 2019, more than 4.78 million cases of novel coronavirus pneumonia have been diagnosed worldwide, with 318,789 deaths recently according to World Health Organization Coronavirus disease (COVID-2019) situation report – 121. Currently, the detection of COVID-19 mainly relies on nucleic acid testing. However, many infected patients with obvious typical symptoms were received multiple nucleic acid tests but diagnosed positive consequently [1], resulting in delayed treatment and even aggravating the spread of the epidemic. On February 5, the Chinese National Health and Health Commission launched the "Novel Coronavirus Pneumonia Diagnosis and Treatment Program (Trial Version 5)", which updated the diagnostic criteria for novel coronavirus pneumonia with adding CT imaging examinations as one of the main basics for clinical diagnosis of COVID-19. CT screening is

considerably popular, easy to operate and sensitive to COVID-19, which is critical for both early diagnosis and epidemic control.

Nevertheless, influenza virus pneumonia and other types of pneumonia might occur as well. In some aspects, especially according to clinical features, it is troublesome to differentiate COVID-19 from common pneumonia. For instance, the main manifestations of COVID-19 in the early stage were fever, fatigue, dry cough, and expiratory dyspnea while patients with common pneumonia have similar symptoms [2]. Pneumonia places a huge burden on the health care system because of its high morbidity and mortality. Therefore, early diagnosis and isolation of CP patients and COVID-19 patients can better prevent the spread of the epidemic and optimize the allocation of medical resources. However, except for the overlapping symptoms and detection abnormalities, CT manifestations of CP and COVID-19 were similar, causing instability and uncertainty for distinguishing them [3, 4].

In the current shortage of medical resources, some countries even recruited medical interns who have not graduated yet and aged retired doctors to return to the hospital for treatment of COVID-19 patients, including radiologists. Nonetheless, the accuracy of CT diagnosis varied from houseman to experts while analyzing thousands of CT images would be no doubt an extravagant challenge for inexperienced doctors.

To tackle this problem, we developed a machine learning method, particularly focusing on differentiating COVID-19 from CP, demonstrating high efficiency in the identification of COVID-19 and CP, helping to reduce misdiagnosis and control pandemic transmission.

Results

After feature selection by T-test results ($p < 0.05$), twenty- five features showed the significant difference between COVID-19 and CP groups were included. Table 1 shows the feature names.

Table 1 Twenty-five features with significant difference between COVID-19 and CP groups ($p < 0.05$).

Statistical method	Feature name
Histogram	1.Mean; 2.Standard Deviation; 3.Smoothness; 4.Cubic moment; 5.Uniformity; 6. Entropy;
GLCM	7.Angular second moment; 8 .Entropy; 9.Contrast; 10 .Inverse difference moment; 11 .Sum average; 12 .Sum entropy; 13 .Sum variance; 14 .Variance;15.Difference average; 16 .Inertia; 17.Difference variance; 18.Difference entropy;
GLGCM	19.High gradient advantage; 20.Average gray; 21.Average gradient; 22 .Gray-scale mean -variance; 23.Gradient scale mean - variance; 24.Coherence; 25.Inertia;

Comparison of EBT and other algorithms

The classification accuracies of EBT, SVM, LR, DT and KNN are 91.66%, 85.90%, 80.98%, 88.66% and 85.99% respectively. The AUCs of EBT, SVM, LR, DT and KNN are 0.98, 0.93, 0.88, 0.91, and 0.86 respectively. Notably, EBT-1 means the EBT algorithm without feature selection. The accuracy of it is 90.58% and the AUC is 0.97. Under most circumstances, EBT achieves the best performance among others. Table 2 shows more specific results of five algorithms. Figure 1 shows the ROCs of five classifiers.

Table.2 Precision, recall and F1-score of body parts based on different classifiers

Classifier	Accuracy	Precision	Sensitivity	Specificity	F1-score	AUC
LR	80.98%	83.33%	77.81%	84.20%	80.48%	0.88
SVM	85.90%	88.50%	82.78%	89.08%	85.54%	0.93
KNN	85.99%	91.13%	79.97%	92.10%	85.19%	0.86
DT	88.66%	90.63%	86.42%	90.92%	88.47%	0.91
EBT-1	90.58%	97.49%	83.44%	97.82%	89.92%	0.97
EBT	91.66%	97.91%	85.26%	98.15%	91.15%	0.98

Comparison of COVID-19 and CP

As is shown in Table 3, no matter what algorithm it is, the number of correctly classified CP images is slightly larger than that of COVID-19. The results demonstrated that approximately from 15.12% (93 out of 615) to 21.46% (132 out of 615) COVID-19 samples are wrongly classified while the maximum error rate of CP is 14.21% (83 out of 584).

Table 3
TP & TN of five classifiers

	TP	TN
LR	467	501
SVM	490	528
KNN	475	554
DT	519	555
EBT-1	515	585
EBT	526	585

Discussion

Because of the absence of CP images, we conducted data augmentation on the ROIs of CP images to balance with the amount of COVID-19 images. Compared with the diverse features of COVID-19 CT images, it is much uncomplicated to discriminate CP samples. It caused superior classification result than that of COVID-19. Whereas without data augmentation, the total accuracy rate of the algorithms would decline. More CP cases should be accumulated for further study.

In CT images of COVID-19, there were multiple small patches, obvious stromal changes in the lung exudate, and the damage to the alveolar epithelium was very obvious [5–8]. To be more specific, many patients presented with ground-glass opacity (GGO), consolidation, GGO plus a reticular pattern, vacuolar sign, microvascular dilation sign, fibrotic streak, pleural thickening, pleural retraction sign, and pleural effusion [9–12]. Patients with common pneumonia generally have flaky high-density shadows in the lung, with relatively few ground-glass changes and alveolitis manifestations which are rather grievous in COIVD-19 CT images [13]. Multiple ground-glass shadows and consolidation in double lungs with a small amount of pleural effusion appears in common pneumonia CT images as well [14–16]. The differences between GGO change, consolidation, density, and other distinct characteristics can be extremely well distributed in statistical textural features. The pronouncedly superior performance of the proposed method benefits overwhelmingly from this advantage.

At present, many scholars have been working on the prediction and diagnosis COVID-19, and most of the algorithms they developed are based on deep learning, which obtained impressive results [17, 18]. For example, the deep learning model (COVNet) developed by Li et al.[19] can accurately detect COVID-19 and distinguish it from lung diseases such as community-acquired pneumonia (CAP). The sensitivity and specificity for detecting COVID-19 in the independent test group were 114 (90% [95% CI: 93%, 98%]) out of 127 (90% [95% CI: 83%, 94%])]) and 307 (96% [95% CI: 93%, 98%]) of 294 (90% [95% CI: 93%, 98%]), the AUC was 0.96 (p-value < 0.001). The sensitivity and specificity for detecting the CAP in the independent

test set were 87% (152 (175)) and 92% ((239 (259)), and the AUC was 0.95 (95% CI: 0.93, 0.97). However, development of the practical deep learning diagnostic systems for epidemic response is rather different from the development of traditional deep learning diagnostic systems due to deep learning-based COVID-19 diagnostic systems is time-consuming to some extent, while initially the COIVD-19 samples are in shortage, thus not enough to train deep learning algorithms that often require a large scale of training data. Furthermore, deep neural networks require high-end Graphics Processing Units (GPUs), which are extremely expensive, to be trained within a reasonable amount of time. It is not practical to train deep neural networks to achieve high performance without GPUs. To effectively leverage such high-end GPUs, fast-access Central Processing Units (CPUs), Solid State Disk (SSD) and large-capacity random-access memory (RAM) are also required. In contrast, machine learning is a relatively fabulous choice. Machine learning can achieve relatively considerable results in a small data set. At the same time, machine learning does not need high-performance hardware, and can try a variety of different machine learning algorithms in a short time to choose the best one. Traditional machine learning algorithms are based on feature engineering which is easier to explain and understand than deep networks that lack transparency and interpretability (e.g., it is impossible to determine what imaging features are being used to determine the output). Our proposed machine learning method in combination with statistical textural features accomplished the accuracy of 91.66% for distinguish COVID-19 from CP and AUC of 0.98. It also has a high sensitivity and specificity of 85.26% and 98.15% respectively.

Limitations

1. The ROI is manually delineated which is time-consuming especially when doctors are racing against time to save lives. 2. The CT images, particularly CP images, are not adequate. Despite that the proposed model attained state-of-the-art performance, more clinical images are required to test the generalizability of the machine learning model to other patients. 3. Instead of distinguishing CP from COVID-19, the result of our established model didn’t determine which specific pneumonia it was, such as viral or bacterial, mainly due to insufficient data and quiet over-lapping CT manifestations. Prognosis of CPs will be considered in our future study.

Conclusions

This study explored an ensemble of bagged tree (EBT) algorithm with statistical textural features for differentiating novel coronavirus pneumonia from common pneumonia. The experimental results show that, as compared to the other four classifiers and EBT without feature selection, the proposed method achieved pronouncedly superior performance with a small amount of CT images.

Methods

Figure 2 illustrates the scheme of this study. Machine learning algorithms integrated with statistical textural features are leveraged to differentiate COVID-19 from CP. Firstly, manually delineation of the ROIs of both COVID-19 and CP would be performed. Then we wrote a Matlab algorithm to crop the ROIs from

DICOM images. Secondly, input the ROIs to feature extraction algorithms. Thirdly, select features with significant difference for time-saving and reducing overfitting. Finally, input the selected features with labels to five classifiers and then the result is the prediction of the classifiers.

Patient selection

From January 2020 to March 2020, there were 73 COVID-19 cases confirmed by nucleic acid test positive and 27 common pneumonia cases enrolled in this study (age ranges from 14 to 72 years). Both COVID-19 and CP patients who had undergone chest CT scans were retrospectively reviewed by two senior radiologists. Of the COVID-19 cases, twelve patients without obvious characteristics on CT images were excluded (negative rate 16.4%, 12/73). Finally, sixty-one confirmed COIVD-19 cases, and 27 common pneumonia cases were enrolled in this study.

Delineation of ROIs

The images were independently assessed by two radiologists. If the radiologists disagreed with each other, a senior radiologist would be invited to review the pulmonary CT images and make the final examination. All the CT images were generated from the Siemens Sensation 16-layer spiral CT (Siemens, Erlangen, Germany). The image format was Digital Imaging and Communications in Medicine (DICOM). The scan parameters were: tube voltage 120 kV; tube current automatic regulation; 1–2 millimeters cross-sectional thickness; 1–2 millimeters cross-sectional distance; scan pitch 1.3; and 16 × 0.625 millimeters collimation.

The software of MRIcro 1.4 was used to extract the rectangle ROI of COVID-19 and CP. ROIs were sketched from CT images based on features of COVID-19 and CP, such as multiple ground-glass opacity in both lungs in the progression stage for COVID-19 and individual ground glass nodule for CP. The main processes of ROI delineation: 1) A rectangular region as large as possible, which is the ROI, was delineated within the features and export the whole image with delineation to a PNG image; 2) Binarize the PNG image to get the ROI boundary and fill the rectangular region to get the ROI template; 3) Using the ROI template to extract the ROI in the original DICOM image; 4) Convert the gray level of the ROI image to 256 gray levels and resize the image to 32 × 32 pixels.. Figure 3 demonstrates the flowchart of ROI delineation. Consequently, 615 COVID-19 CT images and 146 CP ROIs were cropped. Because COVID-19 images were four times larger than CP images which could reduce the classification accuracy, we rotated the CP images by 90, 180, 270 degrees. Ultimately, the number of CP images were augmented to 584. Typical ROIs cropped from COVID-19 and CP cases were shown in Fig. 4.

Sub-image from number 1 to number 8 are the ROIs of COVID-19, sub-image from a to h are the ROIs of CP.

Feature extraction

The statistical texture feature is based on the gray-scale feature parameters extracted by a certain image processing algorithm based on pixels and their neighborhood. The gray level co-occurrence matrix (GLCM) [20] and the gray level-gradient co-occurrence matrix (GLGCM) [21] are the second-order statistical texture features of the gray-scale variation of the image and the basic functions to describe the characteristics of texture structure[22], which have been widely used in medical image processing[23, 24]. In this paper, 13-dimensional GLCM features and 15-dimensional GLGCM features are applied. In the gray level co-occurrence matrix, supposing that the distance d of the pixel is 1, firstly, calculate the gray level co-occurrence matrix in four directions of 0, 45, 90 and 135 degrees, and secondly, calculate the texture features based on the matrix in each direction. Eventually, the average of the texture features in 4 directions was taken as the final feature. Studies have shown that the statistical texture features extracted by GLCM and GLGCM can achieve effective identification of medical images. Also, the histological characteristics of COVID-19 and CP can be well reflected in the gray mode, and the gray histogram is an intuitive statistical method [25]. It is a one-dimensional function of the gray level and belongs to the first-order statistical method. A total of 34-dimensional features including 6-dimensional first-order histogram features, 13-dimensional GLCM features, and 15-dimensional GLGCM features of second-order statistical texture features are utilized in this paper. After obtaining all texture feature data, due to the different calculation methods of each feature, the numerical value changes in a wide range. Therefore, to facilitate calculation, all data are normalized to [0, 1] based on their respective dimensions, the normalized Eq. (1) is as follows:

$$X^* = (X - MIN) / (MAX - MIN)$$

Where X is the original data of the N_{th} dimension, MIN is the minimum value in the N_{th} dimension, MAX is the maximum value in the N_{th} dimension, X^* is the normalized feature.

Feature selection & Classification

Feature selection can usually improve classification performance and reduce overfitting, especially for problems of small sample size. T-test, also known as Student's T-test, is mainly used for normal distribution data with a small sample size and unknown population standard deviation σ. It uses the T-distribution theory to infer the probability of the difference [26], to determine whether the difference between the two samples is significant. Feature with $p < 0.05$ indicates that it has a significant difference, which matches the inclusion criteria.

Ensemble of bagged tree (EBT) algorithm adopts the idea of self-help method while it forms multiple training samples and generates numerous tree models [27]. Besides, the result was predicted by majority voting. EBT uses the same algorithm to train the samples multiple times with various independent classifiers. As for classification, the final output is the voting result of each classifier [28, 29]. It improves the generalization ability, and each prediction function of EBT can be generated in parallel, thereby saving a lot of time.

Different feature classification algorithms have their advantages. For comparison with the performance of the EBT algorithm, support vector machine (SVM), logistic regression (LR), decision tree (DT), K-nearest neighbor with Minkowski distance equal weight (KNN) are implemented with the same texture feature extraction methods and the same feature selection method. To superiorly identify the differences of the results, a 10-fold cross-validation method is adopted. All image data are divided into ten equal parts. Nine of them are employed as the training set, and the remaining one is applied as the validation set to calculate the corresponding accuracy. Take turns until every part is utilized as the validation set. The average value of 10 accuracies is taken as the final accuracy of the classification algorithm.

Statistics

The classification metrics used included area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy, precision, and F1 score. Let TP (true positive) denote the number of samples belonging to class positive and correctly classified; FN (false negative) denote the number of samples belonging to class negative but misclassified; FP (false positive) denote the number of samples not belonging to class positive but misclassified as class positive; FN (false negative) denote the number of samples not belonging to class negative but misclassified as class negative [30]. Classification accuracies are reported in terms of accuracy, sensitivity, specificity, precision and F1 score as

Accuracy = (TP+TN) ⁄ (TP + TN + FP + FN)
Sensitivity = Recall = TP ⁄ (TP+FN)
Precision=TP ⁄ (TP+FP)
Specificity = TN ⁄ (TN+FP)
F1score = 2 × Precision × Recall / (Precision + Recall)

Abbreviations

CT: Computed Tomography; COVID-19: novel coronavirus pneumonia; CP: common pneumonia; EBT: ensemble of bagged tree; DICOM: digital imaging and communications in medicine; kV: kilovolt; GLCM: gray level co-occurrence matrix; GLGCM: gray level-gradient co-occurrence matrix; SVM: support vector machine; LR: logistic regression; DT: decision tree; KNN: K-nearest neighbor with Minkowski distance equal weight; ROC: receiver operating characteristic curve; AUC: the receiver operating characteristic curve; GPU: Graphics Processing Unit; CPU: central processing units; SSD: solid state disk; RAM: random-access memory.

Declarations

Availability of data and materials

The dataset analyzed during the current study was derived from the following public domain resources: https://pan.baidu.com/s/1Ux9dpa1wtquNee4hEh1OWQ, code: k23c.

Ethics approval and consent to participate

The Institutional Review Board of Ruian People's Hospital, Ruian city, Zhejiang province of China approved the retrospective study (YJ202014), and the requirement for written informed consent was waived.

Consent for publication

All authors consent for the publication of this manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the funding of Ruian Science and Technology Bureau (MS2020023, MS2020025).

Authors' contributions

CL and XW are equally contributed as the first authors. Corresponding Author: WP. CL, XW, and WP accomplished the manuscript writing, data analysis, and machine learning model developing. CL and QS accomplished data collecting, clinical expertise providing, and manuscript revising. All authors read and approved the final manuscript.

Acknowledgements

We would like to acknowledge the funding agencies for the support of the work. The content is solely the responsibility of the authors and does not necessarily represent the official views of Ruian Science and Technology Bureau.

Author details

[1] School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2] College of Medical Imaging, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China. [3] Department of Radiology, Ruian People's Hospital, Zhejiang, 325200, China. [4] Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, 518117, China. [5] Infectious Disease Department, Ruian People's Hospital, Zhejiang, 325200, China.

References

1. Li D, Wang D, Dong J, Wang N, Huang H, Xu H, Xia C. "False-Negative Results of Real-Time Reverse-Transcriptase Polymerase Chain Reaction for Severe Acute Respiratory Syndrome Coronavirus 2: Role of Deep-Learning-Based CT Diagnosis and Insights from Two Cases". Korean J Radiol. 2020;21(4):505–8.
2. Cheng Z, Lu Y, Cao Q, Qin L, Pan Z, Yan F, Yang W. "Clinical Features and Chest CT Manifestations of Coronavirus Disease 2019 (COVID-19) in a Single-Center Study in Shanghai, China" *AJR. American journal of roentgenology*, pp. 1–6, 2020.
3. Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, Cui J, Xu W, Yang Y, Fayad ZA, et al., "CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, 2020.
4. Li CX, Wu B, Luo F, Zhang N, "[Clinical Study and CT Findings of a Familial Cluster of Pneumonia with Coronavirus Disease 2019 (COVID-19)]," *Sichuan da xue xue bao. Yi xue ban = Journal of Sichuan University. Medical science edition*, vol. 51, no. 2, pp. 155–158, 2020.
5. Dai H, Zhang X, Xia J, Zhang T, Shang Y, Huang R, Liu R, Wang D, Li M, Wu J, et al., "High-resolution Chest CT Features and Clinical Characteristics of Patients Infected with COVID-19 in Jiangsu, China" *International journal of infectious diseases: IJID : official publication of the International Society for Infectious Diseases*, 2020.
6. Mungmungpuntipantip R, Wiwanitkit V. "Clinical Features and Chest CT Manifestations of Coronavirus Disease (COVID-19)" *AJR. American journal of roentgenology*, pp. W1, 2020.
7. Li X, Zeng W, Li X, Chen H, Shi L, Li X, Xiang H, Cao Y, Chen H, Liu C, et al. "CT imaging changes of corona virus disease 2019(COVID-19): a multi-center study in Southwest China". Journal of translational medicine. 2020;18(1):154.
8. Zhu T, Wang Y, Zhou S, Zhang N, Xia L. "A Comparative Study of Chest Computed Tomography Features in Young and Older Adults With Corona Virus Disease (COVID-19)" *Journal of thoracic imaging*, 2020.
9. Liu M, Song Z, Xiao K. "High-Resolution Computed Tomography Manifestations of 5 Pediatric Patients With 2019 Novel Coronavirus" *Journal of computer assisted tomography*, 2020.
10. Zhong Q, Li Z, Shen X, Xu K, Shen Y, Fang Q, Chen F, Liang T, "[CT imaging features of patients with different clinical types of coronavirus disease 2019 (COVID-19)]," *Zhejiang da xue xue bao. Yi xue ban = Journal of Zhejiang University. Medical sciences*, vol. 49, no. 1, pp. 0, 2020.
11. Guan CS, Lv ZB, Yan S, Du YN, Chen H, Wei LG, Xie RM, Chen BD, "Imaging Features of Coronavirus disease 2019 (COVID-19): Evaluation on Thin-Section CT," *Academic radiology*, 2020.
12. Salehi S, Abedi A, Balakrishnan S, Gholamrezanezhad A. "Coronavirus Disease 2019 (COVID-19): A Systematic Review of Imaging Findings in 919 Patients" *AJR. American journal of roentgenology*, pp. 1–7, 2020.

13. Cao MS, Sheng J, Wang TZ, Qiu XH, Wang DM, Wang Y, Xiao YL, Cai HR. "Acute exacerbation of idiopathic pulmonary fibrosis: usual interstitial pneumonitis vs. possible usual interstitial pneumonitis pattern". Chin Med J. 2019;132(18):2177–84.
14. Ryu AJ, Navin PJ, Hu X, Yi ES, Hartman TE, Ryu JH. "Clinico-radiologic Features of Lung Disease Associated With Aspiration Identified on Lung Biopsy," *Chest*, vol. 156, no. 6, pp. 1160–1166, 2019.
15. Kloth C, Thaiss WM, Beck R, Haap M, Fritz J, Beer M, Horger M. "Potential role of CT-textural features for differentiation between viral interstitial pneumonias, pneumocystis jirovecii pneumonia and diffuse alveolar hemorrhage in early stages of disease: a proof of principle". BMC Med Imaging. 2019;19(1):39.
16. Koo HJ, Lim S, Choe J, Choi SH, Sung H, Do KH, "Radiographic and CT Features of Viral Pneumonia," *Radiographics: a review publication of the Radiological Society of North America, Inc*, vol. 38, no. 3, pp. 719–739, 2018.
17. Alimadadi A, Aryal S, Manandhar I, Munroe PB, Joe B, Cheng X. "Artificial intelligence and machine learning to fight COVID-19" *Physiological genomics*, vol. 52, no. 4, pp. 200–202, 2020.
18. Santosh KC. "AI-Driven Tools for Coronavirus Outbreak: Need of Active Learning and Cross-Population Train/Test Models on Multitudinal/Multimodal Data". Journal of medical systems. 2020;44(5):93.
19. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, et al., "Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT," *Radiology*, pp. 200905, 2020.
20. Xu X, Zhang X, Tian Q, Zhang G, Liu Y, Cui G, Meng J, Wu Y, Liu T, Yang Z, et al. "Three-dimensional texture features from intensity and high-order derivative maps for the discrimination between bladder tumors and wall tissues via MRI". Int J Comput Assist Radiol Surg. 2017;12(4):645–56.
21. Jiang S, Mao H, Ding Z, Fu Y. "Deep Decision Tree Transfer Boosting" *IEEE transactions on neural networks and learning systems*, 2019.
22. Haralick RM, Shanmugam K, Dinstein I, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.
23. Lam SW. Texture feature extraction using gray level gradient based co-occurence matrices. Presented at IEEE International Conference on Systems,1996,1:267–271.
24. Yang Q, Gao F, Nie Q. "Analysis of rotation invariance in texture image recognition". Computer Engineering Applications. 2010;46:205–7.
25. Langner T, Wikstrom J, Bjerner T, Ahlstrom H, Kullberg J. "Identifying morphological indicators of aging with neural networks on large-scale whole-body MRI" *IEEE transactions on medical imaging*, 2019.
26. Motz L, Weaver JH. "The story of mathematics" *Princeton University Press Princeton Nj*, vol. 9, 2005.
27. Al-Barazanchi KK, Al-Neami AQ, Al-Timemy AH, "Ensemble of bagged tree classifier for the diagnosis of neuromuscular disorders," Presented at 2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME), Beirut, 2017, pp. 1–4.

28. Tan AC, Gilbert D. "Ensemble machine learning on gene expression data for cancer classification. Applied bioinformatics. 2003;2(3 Suppl):S75–83.
29. Pei C, Sun Y, Zhu J, Wang X, Zhang Y, Zhang S, Yao Z, Lu Q. "Ensemble Learning for Early-Response Prediction of Antidepressant Treatment in Major Depressive Disorder" *Journal of magnetic resonance imaging: JMRI*, 2019.
30. Yan Z, Zhan Y, Peng Z, Liao S, Shinagawa Y, Zhang S, Metaxas DN, Zhou XS. "Multi-Instance Deep Learning: Discover Discriminative Local Anatomies for Bodypart Recognition". IEEE Trans Med Imaging. 2016;35(5):1332–43.

Figures

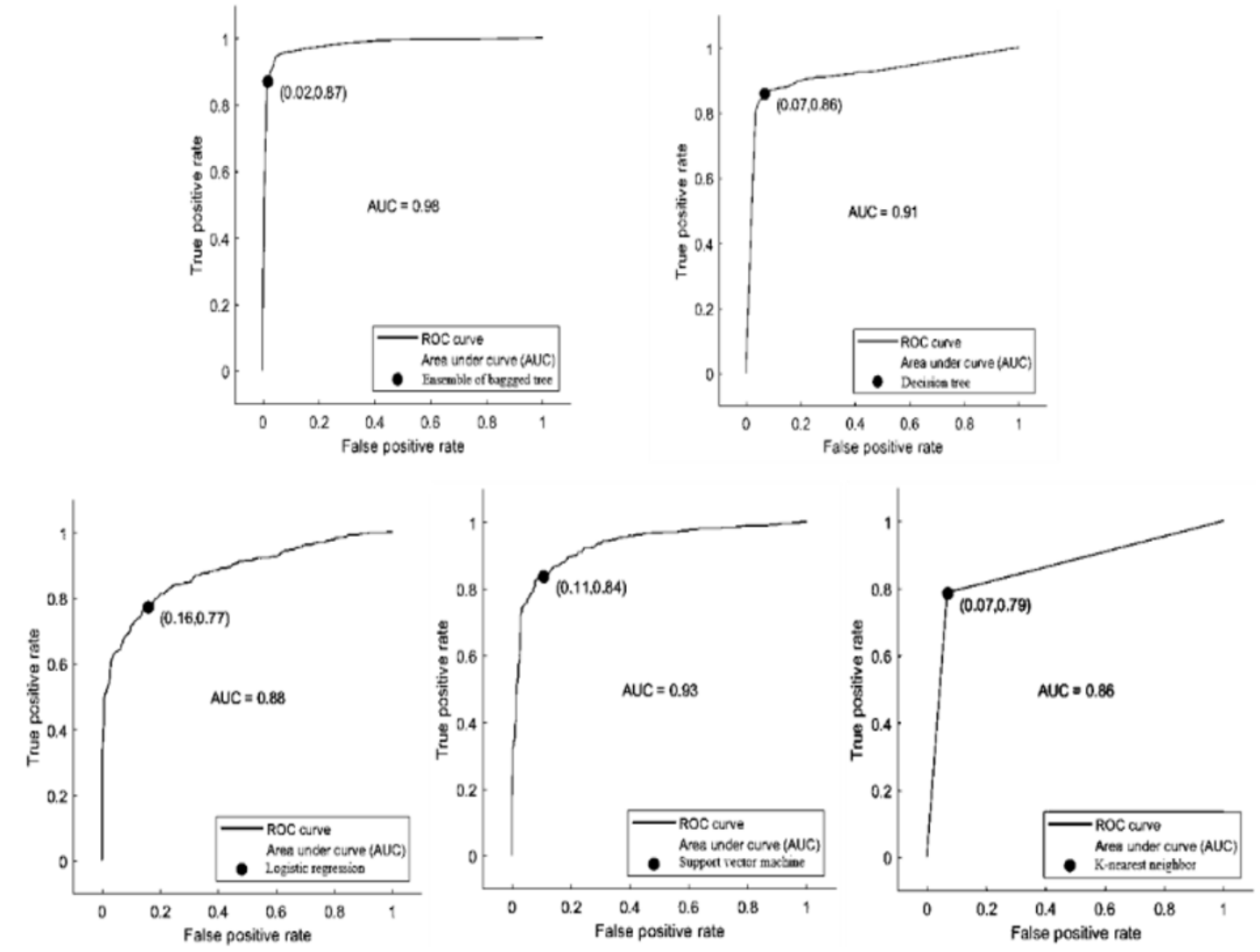

Figure 1

ROCs of five classifiers.

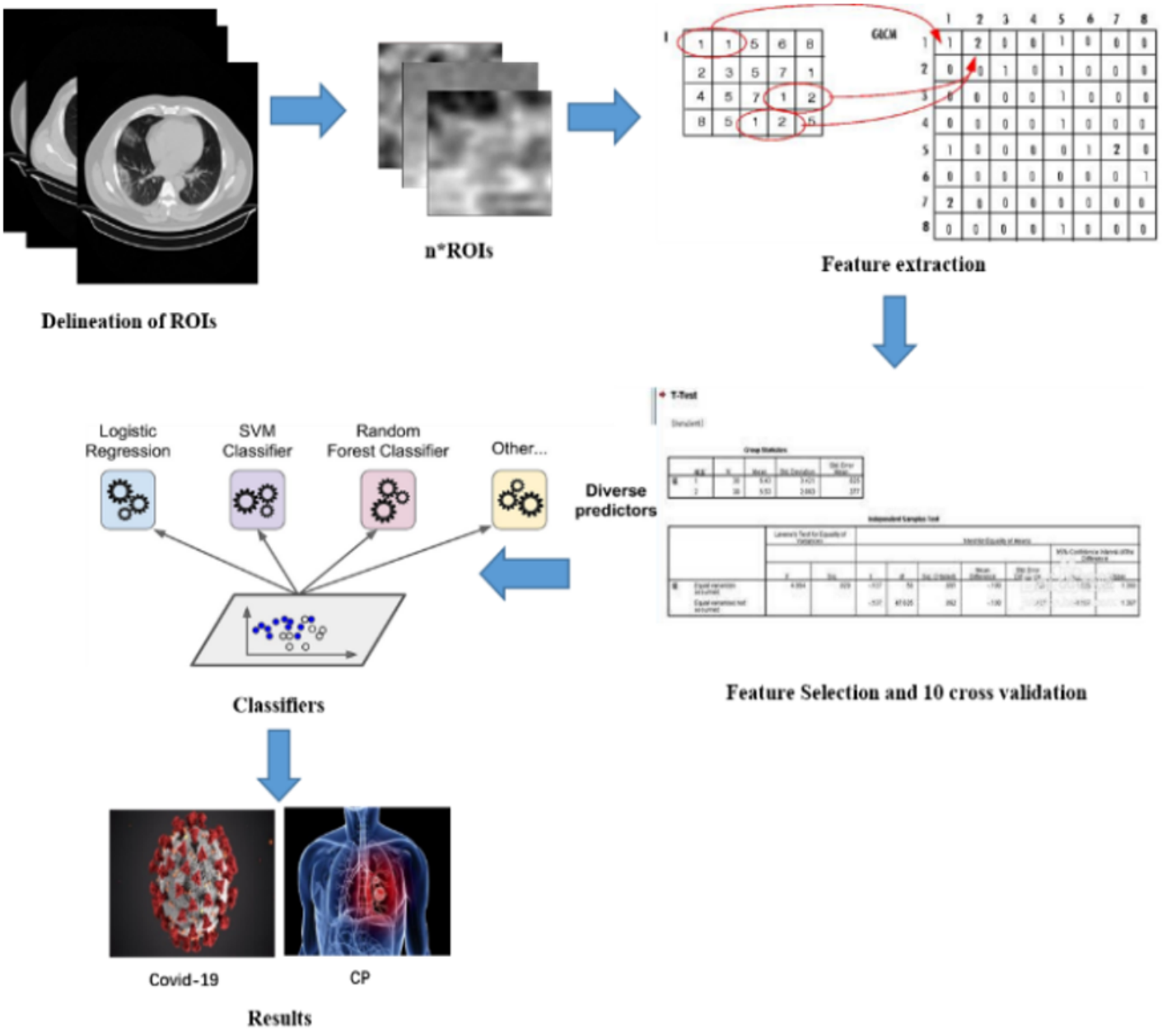

Figure 2

The flowchart of this study

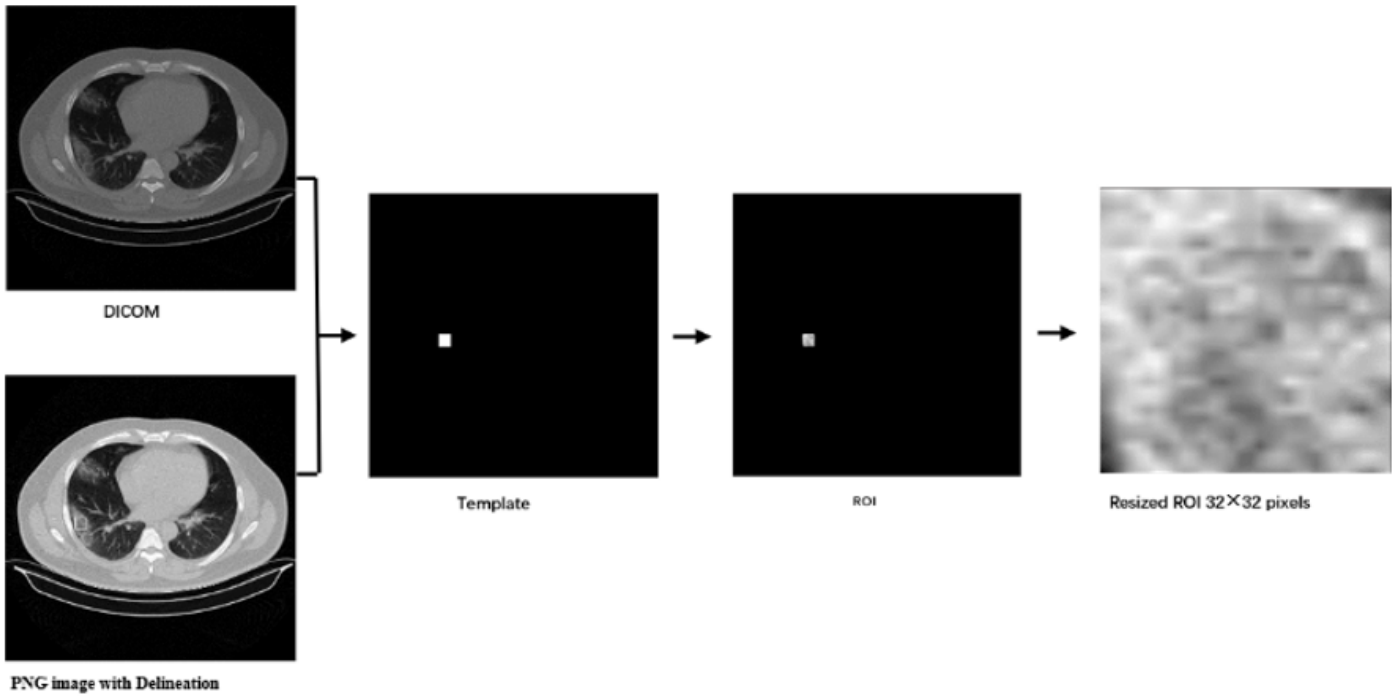

Figure 3

The flowchart of ROI Delineation

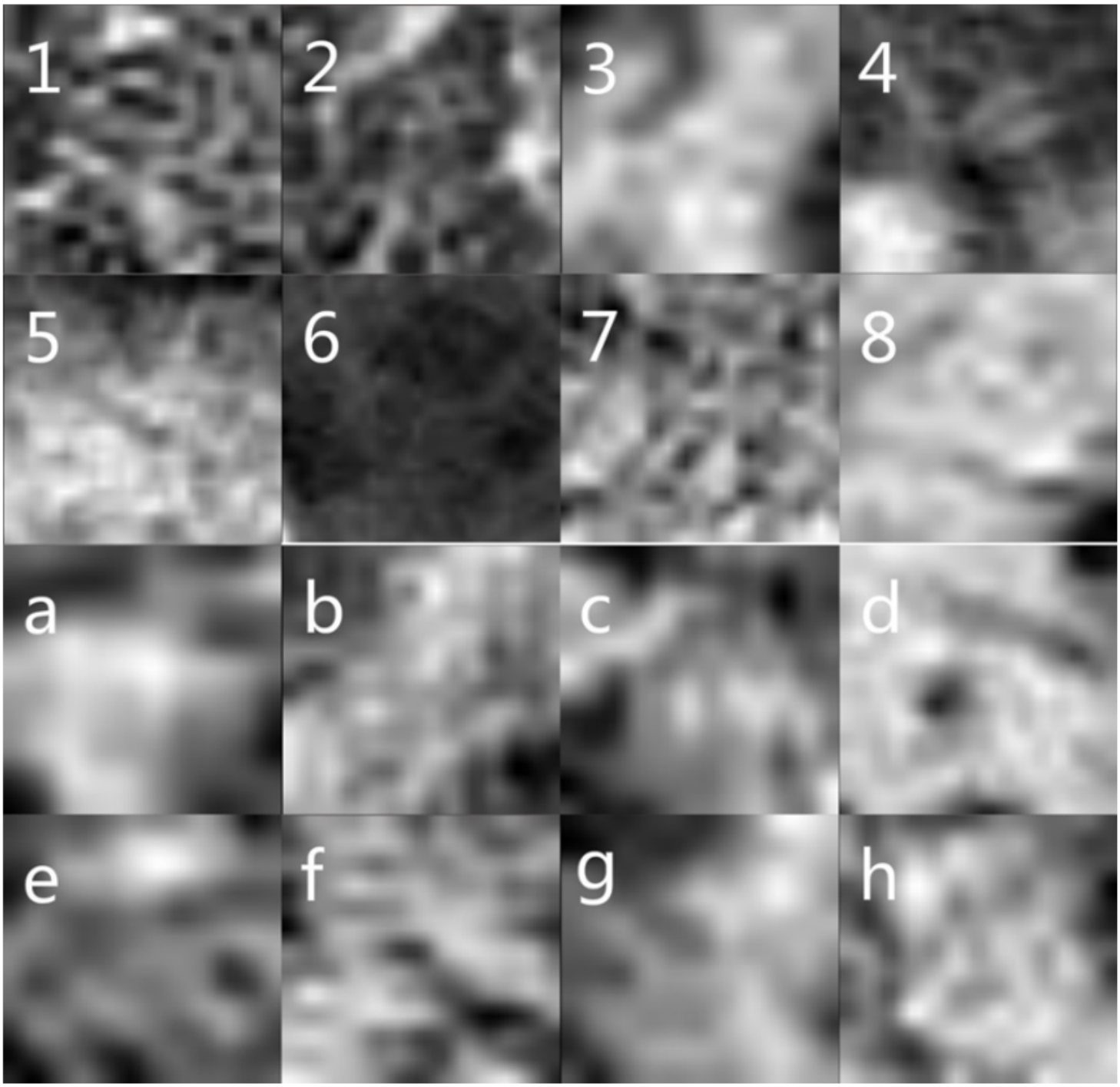

Figure 4

The ROI of COVID-19 and CP.