Automatic Pneumonia Detection Using Deep Learning

Luis Vinícius de Moura
School of Technology, Medical Imaging Computing Laboratory - MEDICOM
PUCRS
Porto Alegre, Brazil
ORCID: 0000-0003-3429-3289

Caroline Machado Dartora
School of Medicine, Graduate Program in Biomedical Gerontology
PUCRS
Porto Alegre, Brazil
ORCID: 0000-0003-1357-3230

Christian Mattjie de Oliveira
School of Medicine, Graduate Program in Biomedical Gerontology
PUCRS
Porto Alegre, Brazil
ORCID: 0000-0002-3745-8686

Rodrigo Coelho Barros
School of Technology, Machine Intelligence and Robotics Research Group
PUCRS
Porto Alegre, Brazil
ORCID: 0000-0002-0782-9482

Ana Maria Marques da Silva
School of Technology, Medical Imaging Computing Laboratory - MEDICOM
PUCRS
Porto Alegre, Brazil
ORCID: 0000-0002-5924-6852

***Abstract*— Radiological chest examinations like chest X-ray play a fundamental role in the fight against the outbreak of COVID-19 pneumonia, caused by the coronavirus strain SARS-Cov-2. This study aims to investigate classification models to differentiate chest X-ray images of COVID-19-based and typical pneumonia using hand-crafted radiomic features, understanding the distinctive radiographic features of COVID-19. A total of 136 segmented chest X-rays from two public databases were used to train and evaluate the classification methods. The PyRadiomics library was used to extract first and second-order statistical texture features in the right (R), left (L), and in superior, middle and bottom lung zones for each lung side. For performing feature selection, data was split in training (80%) and test (20%) sets. Stratified K-folds (K=5) was used within the training dataset for cross-validation. The most relevant radiomic features were selected after measuring validation accuracy and relative feature importance. Support vector machines (SVM), random forest (RF), AdaBoost (AB), and logistic regression (LR) were analyzed as potential classifiers. The AB model was the best discriminant method between features related to COVID-19-based when compared to typical pneumonia, using a model of lung segmentation in six distinct lung zones (AUC = 0.98). Our study shows a predominance of radiomic feature selection in the right lung, with a tendency to the upper lung zone.**

Keywords - coronavirus, atypical pneumonia, radiography, segmentation, radiomics, feature selection, classification.

I. Introduction

The outbreak of COVID-19 pneumonia, caused by the coronavirus strain SARS-Cov-2 (severe acute respiratory syndrome coronavirus 2), has caused global turmoil and was declared a pandemic by the World Health Organization on March 13, 2020. On June 21, more than 9 million cases worldwide were confirmed. More than 1 million of these cases are located in Brazil.

The incubation period of COVID-19 is 5.2 days and can last up to 14 days [1]. Clinical features include respiratory symptoms, fever, cough, dyspnea, and viral pneumonia. COVID-19 shows to be more transmissible when people display the symptoms [2]. However, there are several cases in which subjects are asymptomatic. Radiological chest examinations, as chest X-ray or computed tomography (CT), already play a fundamental role in the fight against COVID-19 [3].

Observation of radiological lung patterns can reveal different types of pulmonary diseases. These patterns can be described based on disease and the affected tissue region. Knowledge of disease-related patterns is very important for differentiation and follow-up of pulmonary diseases. Studies have shown that COVID-19, for instance, induces abnormal pneumonia that leads to a bilateral, peripheral, ground-glass opaque pattern [4]. Apart from typical visual analysis, lung diseases pattern can be studied through feature analysis using a technique called radiomics.

Computer-based texture analysis is used to quantify numerically specific features of an image. The quantitative analysis of morphological, intensity, and texture features is helpful on diagnosis and prognosis. Texture analysis can be further categorized into structural, model-based, transformational, and statistics-based [5].

Most recent COVID-19 radiological studies are focused on CT findings, which has better sensitivity than X-ray. However, CT is more expensive and scarcer when compared to conventional X-rays, requiring a more complicated process of decontamination after COVID-19 patient scanning. The American College of Radiology (ACR) recommends CT exams to be used sparingly and reserved for hospitalized COVID-19 symptomatic patients with specific clinical indications. A portable chest X-ray is suggested as a viable option to minimize the risk of cross-infection and avoid overload and disruption of radiological departments [6].

Several models of classification and prediction of COVID-19-based pneumonia using binary or multilabel classification techniques have been developed to aid the diagnosis using CT and X-ray images [7]–[13]. Studies have been trying to identify patterns and features related to this new disease.

This study aims to investigate classification models to differentiate chest X-ray images of COVID-19-based and typical pneumonia using hand-crafted radiomic features, understanding the distinctive radiographic features of COVID-19. Our study analyzed texture-features in two different approaches of lung segmentation, showing a predominance of radiomic feature selection in the right lung, with a tendency to the upper lung zone.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brasil (CAPES) – Finance Code 001.

2471-7819/20/$31.00 ©2020 IEEE
DOI 10.1109/BIBE50027.2020.00078

II. Materials and Methods

A. Image Dataset

A total of 136 anteroposterior (AP) and posteroanterior (PA) chest X-rays from two public databases were used to train and evaluate the classification methods. Sixty-eight COVID-19 images were provided by the COVID-19 Image Data Collection [14], composed of images and information of multiple centers, some subjects with longitudinal studies. We chose to use just those images of the first time point of the subject. Images have various matrix sizes, ranging between 156×156 pixels and $3{,}520 \times 4{,}280$ pixels, stored in JPEG and PNG. The CheXpert dataset [15] was used to get an additional 68 chest X-ray pneumonia images, which were randomly selected. Image sizes vary from 320×320 pixels to 320×394 pixels, all in JPEG file format. The entire dataset was split into training (80% , $N = 108$, with 54 COVID-19 images) and test sets (20%, $N = 28$, with 14 COVID-19 images). The testing set was never seen by the model until the very last metric evaluation.

B. Lung Segmentation

Chest X-ray images were rescaled to 256×256 pixels, and we have applied a histogram equalization procedure. Lungs were segmented using a pre-trained open code segmentation model [1] to generate lung masks. All lung segmentations were visually inspected. Two different lung masks were produced: the division of upper, middle, and bottom zones in each lung side; and the split in the left and right lungs (Fig.1).

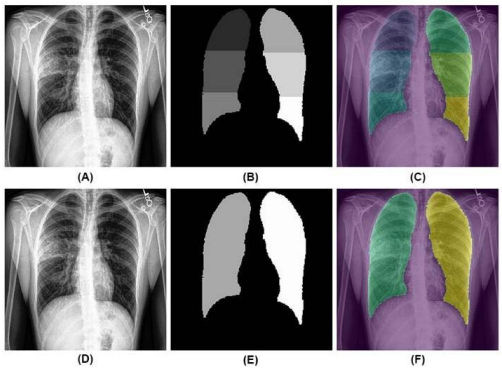

Fig. 1. Lung Segmentation. (A, D) Original chest X-ray, lung zones (B) and left-right (E) masks, and chest X-ray overlapped with lung masks (C, F). Source: The authors (2020).

Lung masks were stretched back to 512×512 pixels. An opening morphological operator with a structural element and an 8-connected neighborhood was applied to remove background clusters and fill holes of the lung mask. Clusters with less than 5 pixels were removed. A division of connected areas was made, and areas with less than 75 pixels were excluded.

The split between left and right lungs used the centroid of two areas; if the centroid was in the first half of the matrix size (from left to right), it was considered the right lung (observe the radiological image in chest X-ray is mirrored). The height of each lung was divided into upper, middle, and bottom zones, determined by the difference in the extremity points, divided in $1/3$.

Lung masks were applied in the respective chest X-ray and divided in the left and right side (L-R) and then divided into superior, middle, and bottom zones. This separation into lung zones is similar to the proposal applied in [16] regarding CT lung images.

C. Radiomic Feature Extraction and Extraction

The PyRadiomics library was used for the extraction of statistical texture features of the first and second order for each lung mask. The number of radiomic features is divided into five classes [17]:

1. First-order features (18 features): based on the first-order histogram and related to the pixel intensity distribution.

2. Gray-level co-occurrence matrix or GLCM (24 features): gives information about the gray-level spatial distribution, considering the relationship between pixel pairs and the frequency of each intensity within an 8-connected neighborhood.

3. Gray-level Run Length Matrix or GLRLM (16 features): like GLCM, it is defined as the number of contiguous pixels with the same gray level considering a 4-connected neighborhood, indicating the homogeneity of pixel intensity.

4. Gray-level Size Zone Matrix or GLSZM (16 features): used for texture characterization, it provides statistical representation by the estimation of a bivariate conditional probability density function of the image distribution values. It is invariant to image rotation.

5. Gray-level Dependence Matrix or GLDM (14 features): quantifies the dependence of image gray-level by calculating the connectivity at a certain distance when their difference in pixel intensity is < 1.

For performing feature selection, data was split into training (80%) and test (20%) sets. Stratified K-folds (K=5) was used within the training set for cross-validation. Data were scaled between 0 and 1, and a Random Forest (RF) model was trained in 4 folds and validated in the 5th folder. Validation accuracy and the relative feature importance were measured. After stratifying K-folds five times, importance feature mean was measured. For statistical accuracy, this method was run 100 times. Relative feature importance average and confidence interval of the 100 runs were calculated. The most relevant features were selected.

D. Machine Learning Models

Four classification models were trained with the selected radiomic features. The classification methods used in this paper are Support Vector Machines (SVMs, more specifically its SVC version [18]), Random Forest (RF) [19], Adaptative Boosting or AdaBoost (AB) [20], and Logistic regression (LR) [21]. All models were tuned by grid search with cross-validation, looking for the best sensitivity. The tuned parameters were kernel, C, and degree for SVC; the number of estimators, criterion and maximum depth for RF; the number of estimators and learning rate for AB and penalty, C and maximum iteration for LR. All models were implemented using the scikit-learn library on Python version 3.6.5 [22].

[1] https://github.com/imlab-uiip/lung-segmentation-2D

For evaluation, sensitivity (Eq. 1), accuracy (Eq. 2), and the area under the curve (AUC) of the receiver operating characteristic (ROC) graph were used. The final model was selected based on the best sensitivity achieved in the test dataset. Each metric was calculated as following [23]:

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN}) \quad (1)$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (2)$$

where: TP = true positive, TN = true negative, FP = false positive and FN = false negative.

III. Results

A. Selected Features

For each approach (lung zones and L-R), 88 features were extracted. The relative importance of each feature and each approach was measured with 100 runs of the RFC model with Stratified K-Folds (K=5).

The relative importance of each radiomic feature in each approach is shown in Fig. 2.

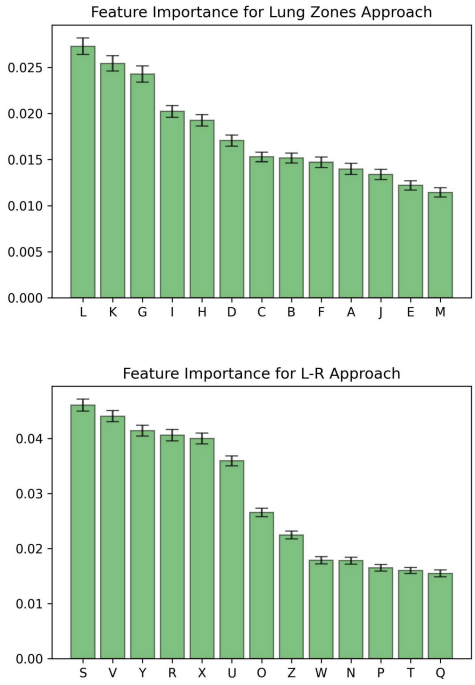

Fig.2 Feature importance for Lung Zones approach (top) and L-R approach (bottom).

B. Classification

Each model used was tuned by a Grid Search with Cross-Validation [24] in the training dataset, seeking for the best sensitivity. After the best hyper-parameters selection, all models were trained in the entire training dataset. The performance was evaluated in the test dataset. Table I shows the best metrics for each approach.

TABLE I. Model Metrics

Approach	Model	Sensitivity	Accuracy	AUC
Lung Zones	**LR**	0.8947	0.6786	0.75
	SVC	0.8947	0.8571	0.92
	RF	0.8947	0.8571	0.95
	AB	**0.9471**	**0.9286**	**0.98**
L-R	**LR**	0.8421	**0.8571**	**0.94**
	SVC	**0.8947**	**0.8571**	0.91
	RF	0.8421	**0.8571**	0.91
	AB	0.8421	0.8214	0.92

The performance of each model was measured using the test dataset. Table 1 shows the AUC, sensitivity, and accuracy for each model. Figure 3 shows the ROC graph for each model approach.

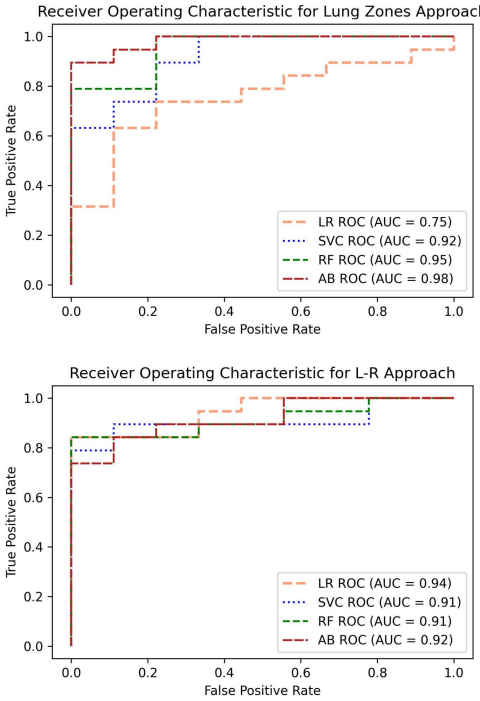

Fig.3 ROC curve for each model for Lung Zones approach (top) and L-R approach (bottom).

Based on the sensitivity values, the best model was AdaBoost (AB). The parameters used in this model were the learning rate equal to 1 and 50 estimators.

Figure 4 shows the time to run each step of the proposed pipeline to make a prediction of a new case/subject using the AB model.

The thirteen most relevant radiomic features for each approach are shown in Table II.

IV. Discussion

A wide variety of computer-aided methods are being applied to aid the diagnosis and prognosis of COVID-19 [7], [8], [25], [26]. Clinical evaluation of symptomatic COVID-19 patients exhibits pulmonary problems and atypical pneumonia. Deep learning (DL) “black box” techniques are

the most common strategy. Our approach used hand-crafted radiomics features and machine learning models to differentiate pneumonia patterns from chest X-rays.

Our analysis aimed to find a group of meaningful radiomic features and the best classification model to differentiate between COVID-19-based pneumonia and typical pneumonia. Two lung segmentation approaches were performed to assess the influence in pneumonia types differentiation: left and right lung sides (L-R), and in zones in each lung (upper, lower, bottom). The segmentation and use of masks in specific regions avoid the features unrelated to the lung disease pattern, such as lung borders, presence of heart, muscles, and bones, restricting the evaluation in the lung tissues.

The use of the lung zones approach, separating the upper, middle, and bottom regions, achieved better performance classification metrics than L-R. It might be associated with the use of smaller regions for the feature analysis, making it more representative of small structures, suppressed in L-R group due to the prevalence of bigger homogeneous regions. Our best result reached 94% sensitivity in differentiating COVID-19-based pneumonia from typical pneumonia using the AdaBoost model with the separated lung zones.

GLSZM radiomic features were selected in five of six lung zones, composing eleven of thirteen significant selected features. These features are based on a gray level zone, defined as the number of 8-connected pixels sharing the same gray level intensity. These features are invariant to rotation, with the initial matrix calculated in all directions at once [27].

When the importance of each feature in the classification (Fig.2) is correlated with the features (Table II), it can be seen that two GLSZM are between the most relevant in both segmentation approaches: the size zone non-uniformity normalized feature and the small area emphasis. The first quantifies the variability on the size zones, with a lower value representing a higher homogeneity among the size of zones. The second feature, on the other hand, quantifies the distribution of small size zones in the image, having higher values when fine textures are present [17].

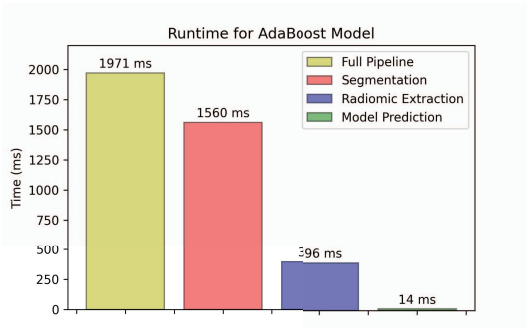

Fig.4 Time to run the pipeline with AdaBoost model for each image.

In the GLDM class, which quantifies gray level dependency on an image, defined by the difference between neighboring pixels connected with the central pixel, based on the difference between these gray levels, the dependence non-uniformity feature is between the most relevant features. It quantifies the similarity between pixels, where small values represent a higher homogeneity [28].

Four different machine learning models for classification (SVC, RF, AB, and LR) between COVID-19-based and typical pneumonia were used with the previous thirteen selected features. We decided to don't use DL models, like convolutional neural networks, because we aimed to find the

TABLE II. SELECTED FEATURES FOR EACH APPROACH

Approach	Abbreviation	Side	Zone Location	Class	Feature
Lung Zones	A	Left	Middle	GLSZM	Size Zone Non-Uniformity Normalized
	B	Left	Middle	GLSZM	Small Area Emphasis
	C	Left	Upper	GLSZM	Size Zone Non-Uniformity Normalized
	D	Left	Upper	GLSZM	Small Area Emphasis
	E	Right	Bottom	GLSZM	Size Zone Non-Uniformity Normalized
	F	Right	Bottom	GLSZM	Small Area Emphasis
	G*	Right	Middle	GLDM	Dependence Non-Uniformity
	H	Right	Middle	GLSZM	Size Zone Non-Uniformity Normalized
	I	Right	Middle	GLSZM	Small Area Emphasis
	J	Right	Upper	GLDM	Dependence Non-Uniformity
	K*	Right	Upper	GLSZM	Size Zone Non-Uniformity Normalized
	L*	Right	Upper	GLSZM	Small Area Emphasis
	M	Right	Upper	GLSZM	Zone Entropy
L-R	N	Left	-	First Order	Minimum
	O	Left	-	GLDM	Dependence Non-Uniformity
	P	Left	-	GLDM	Large Dependence Low Gray Level Emphasis
	Q	Left	-	GLRLM	Long Run Low Gray Level Emphasis
	R	Left	-	GLSZM	Size Zone Non-Uniformity Normalized
	S*	Left	-	GLSZM	Small Area Emphasis
	T	Left	-	GLSZM	Zone Entropy
	U	Right	-	GLDM	Dependence Non-Uniformity
	V*	Right	-	GLDM	Gray Level Non-Uniformity
	W	Right	-	GLRLM	Long Run Low Gray Level Emphasis
	X	Right	-	GLSZM	Size Zone Non-Uniformity Normalized
	Y*	Right	-	GLSZM	Small Area Emphasis
	Z	Right	-	GLSZM	Zone Entropy

Asterisk (*) next to the letter in the abbreviation is marking the three most important features in the classification model.
GLSZM is Gray-level Size Zone Matrix; GLDM is Gray-level Dependence Matrix.

best model keeping feature's information and its importance in the classification.

The highest sensitivity was obtained using the AB model with the lung zones approach. AdaBoost is referred, in other studies, as the best option to boost the performance of decision trees on binary classification problems. AdaBoost creates a collection of weak learners by maintaining a set of weights over training samples and adjusting after each weak learning cycle adaptively. The samples weights which are misclassified by the current weak learner are increased, while the weights of the samples which are correctly classified are decreased [20].

When we used L-R lung segmentation, all models perform similarly. This might be due to the accuracy classification limit using the whole lung segmentation, where a vast area is used to the extraction of the radiomic attributes.

In our study, we achieved 93% and 95% of accuracy and sensibility. Even though it is a binary classification problem, our study used lung segmentation, with features related only to the pulmonary tissue patterns.

Other classification studies classifying COVID-19-based pneumonia from other pulmonary diseases have emerged in the last few months. Accuracies varying between 90% and 98% in pneumonia classification have been achieved with DL models [9]–[12]. Apostolopoulos *et al.* (2020) [7] extracted biomarkers from chest X-ray images to differentiate seven classes: COVID-19-based pneumonia, edema, pleural effusion, emphysema, fibrosis, pneumonia, and normal. DL was used for the extraction of high order features and a MobileNet v2 for multilabel classification. Accuracy was 87.7%, with 99.18% and 97.36% of accuracy and sensitivity just for COVID-19 pneumonia.

Asnaoui and Chawki (2020) [8] have used X-ray and CT images with DL models to classify between atypical COVID-19-based pneumonia, typical pneumonia, and normal subjects. DL models were: VGG16, VGG19, DenseNet201, Inception_ResNet_V2, Inception_V3 and Resnet50 and MobileNet_V2. They obtained 82.8% of sensitivity in COVID-19 classification. However, it is important to note that typical pneumonia and normal subjects were taken from a pediatric database, while the COVID-19 dataset has only images of adult subjects. Thereby, this model might be differentiating the relation of adult versus pediatric X-ray, and not the pulmonary diseases.

The population bias was avoided by the work published by Rahimzadeh and Attar (2020) [29], where they use the same database and groups for classification, but using only adult chest X-rays. In this study, Rahimzadeh and Attar (2020) [29] concatenated two DL networks (Xception and ResNet50V2) for high order feature extraction and classification. The average performance accuracy was 80.5%, with a sensitivity of 99.5% for COVID-19 cases.

Despite showing encouraging results of classification accuracy and sensitivity, the use of DL to aid the diagnosis of diseases has shown some concerns about the model explainability. DL are “black boxes”, with fundamental issues in explaining the decision-making process for classification, and which aspects of the input data drive the decisions of the network. Furthermore, Maguolo and Nanni (2020) [30] showed that their datasets might influence many proposed DL models for COVID-19 identification. Images from the same dataset usually have similar characteristics, since most come from similar equipment and medical center. Because DL models use all available information and characteristics on the image, the models might be learning to discriminate the datasets instead of diseases, leading to biased COVID-19 identification.

Studies showed that the radiological findings in CT images of COVID-19 patients are evenly distributed between the left and right lungs [31]–[33]. However, findings are more present in the lower right lobe, followed by the upper and lower left lobes [4], [34], [35].

It is interesting to note that among the most important features selected by our method, we have two related to the bottom right lung. For the left lung side, we only have features related to the medial and upper lung zones. We hypothesize that for our method, the lack of selection of features in the lower part of the left lung might be due to the heart penumbra, which makes it harder to segment this lung region. In CT images, however, this tissue overlay does not occur. Thus, there may be a significant variation in features in this anatomical region, causing it not to be selected to perform the classification.

It is important to emphasize that the computational cost of ML models should not be too high so they could be potentially applied in clinical use. The analysis of computational time to perform the classification using our best model, including all steps, took less than two seconds using a 2.3 GHz Intel Xeon processor with a single-core, making it suitable for clinical use.

Some limitations of this study are related to the low spatial resolution of some images from the COVID-19 dataset, and a small number of chest X-ray images of subjects with COVID-19-based pneumonia in public datasets. Another limitation of our method is the segmentation step that needs further work to improve its reliability. Currently, chest X-ray segmentation models are trained in images of subjects that have pulmonary diseases without severe lung obstruction or lesions, leading to a non-generalization for more aggressive pulmonary diseases.

V. Conclusion

This paper presents the investigation of classification models to differentiate chest X-ray images between COVID-19-based and typical pneumonia. Our analysis showed that the AdaBoost ML model is the best discriminant method between features related to COVID-19-based pneumonia when compared to typical pneumonia, using a model of lung segmentation in six distinct zones. Our study showed a predominance of features being selected in the right lung, with a tendency to the upper zone.

Further studies are required to increase the number of chest X-ray images of COVID-19-based pneumonia to investigate features related to radiological findings. A more in-depth evaluation of the radiomic features related to COVID-19 in chest X-ray and CT images will be required to analyze if there is a radiomic signature of COVID-19.

Acknowledgment

We would like to thank Google Collaboratory for the free availability of computational resources for the development of this study.

REFERENCES

[1] Q. Li *et al.*, "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia," *N. Engl. J. Med.*, vol. 382, no. 13, pp. 1199–1207, Mar. 2020.

[2] N. Chen *et al.*, "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," *Lancet*, vol. 395, no. 10223, pp. 507–513, 2020.

[3] L. Hu and C. Wang, "Radiological role in the detection, diagnosis, and monitoring for the coronavirus disease 2019 (COVID-19)," *Eur. Rev. Med. Pharmacol. Sci.*, vol. 24, no. 8, pp. 4523–4528, 2020.

[4] H. Shi *et al.*, "Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study," *Lancet Infect. Dis.*, vol. 20, no. 4, pp. 425–434, Apr. 2020.

[5] Y. Yoon, T. Hwang, H. Choi, and H. Lee, "Classification of radiographic lung pattern based on texture analysis and machine learning," *J. Vet. Sci.*, vol. 20, no. 4, 2019.

[6] A. C. of Radiology, "ACR Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection," 2020. [Online]. Available: https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection. [Accessed: 22-Jun-2020].

[7] I. D. Apostolopoulos, S. I. Aznaouridis, and M. A. Tzani, "Extracting Possibly Representative COVID-19 Biomarkers from X-ray Images with Deep Learning Approach and Image Data Related to Pulmonary Diseases," *J. Med. Biol. Eng.*, vol. 40, no. 3, pp. 462–469, Jun. 2020.

[8] K. El Asnaoui and Y. Chawki, "Using X-ray images and deep learning for automated detection of coronavirus disease," *J. Biomol. Struct. Dyn.*, pp. 1–12, May 2020.

[9] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, p. 103792, Jun. 2020.

[10] M. E. H. Chowdhury *et al.*, "Can AI help in screening Viral and COVID-19 pneumonia?," Mar. 2020.

[11] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," Mar. 2020.

[12] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," Mar. 2020.

[13] J. Pu *et al.*, "Any unique image biomarkers associated with COVID-19?," *Eur. Radiol.*, May 2020.

[14] J. P. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," Mar. 2020.

[15] J. Irvin *et al.*, "CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison," Jan. 2019.

[16] D. Colombi *et al.*, "Well-aerated Lung on Admitting Chest CT to Predict Adverse Outcome in COVID-19 Pneumonia," *Radiology*, p. 201433, Apr. 2020.

[17] J. J. M. Van Griethuysen *et al.*, "Computational radiomics system to decode the radiographic phenotype," *Cancer Res.*, vol. 77, no. 21, pp. e104–e107, 2017.

[18] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New Support Vector Algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, May 2000.

[19] L. Breiman, "Random forests" *Mach. Learn.*, pp. 5–32, 2001.

[20] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.

[21] J. I. E. Hoffman, "Logistic Regression," in *Basic Biostatistics for Medical and Biomedical Practitioners*, Elsevier, 2019, pp. 581–589.

[22] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[23] R. Parikh, A. Mathai, S. Parikh, G. Chandra Sekhar, and R. Thomas, "Understanding and using sensitivity, specificity and predictive values," *Indian J. Ophthalmol.*, vol. 56, no. 1, p. 45, 2008.

[24] P. M. Lerman, "Fitting Segmented Regression Models by Grid Search," *Appl. Stat.*, vol. 29, no. 1, p. 77, 1980.

[25] S. Wang *et al.*, "A deep learning algorithm using CT images to screen for coronavirus disease (COVID-19)," *medRxiv*, p. 2020.02.14.20023028, 2020.

[26] S. F. Ardabili *et al.*, "COVID-19 Outbreak Prediction with Machine Learning," *SSRN Electron. J.*, 2020.

[27] G. Thibault *et al.*, "Texture indexes and gray level size zone matrix. Application to cell nuclei classification", 2009.

[28] C. Sun and W. G. Wee, "Neighboring gray level dependence matrix for texture classification," *Comput. Vision, Graph. Image Process.*, vol. 23, no. 3, pp. 341–352, Sep. 1983.

[29] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2," *Informatics Med. Unlocked*, vol. 19, p. 100360, 2020.

[30] G. Maguolo and L. Nanni, "A Critic Evaluation of Methods for COVID-19 Automatic Detection from X-Ray Images," Apr. 2020.

[31] Y. Wang *et al.*, "Temporal Changes of CT Findings in 90 Patients with COVID-19 Pneumonia: A Longitudinal Study," *Radiology*, p. 200843, Mar. 2020.

[32] M. Chung *et al.*, "CT imaging features of 2019 novel coronavirus (2019-NCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, 2020.

[33] R. C. Chate *et al.*, "Presentation of pulmonary infection on CT in COVID-19: initial experience in Brazil," 2020.

[34] W. Zhao, Z. Zhong, X. Xie, Q. Yu, and J. Liu, "Relation Between Chest CT Findings and Clinical Conditions of Coronavirus Disease (COVID-19) Pneumonia: A Multicenter Study," *Am. J. Roentgenol.*, vol. 214, no. 5, pp. 1072–1077, May 2020.

[35] S. H. Yoon *et al.*, "Chest Radiographic and CT Findings of the 2019 Novel Coronavirus Disease (COVID-19): Analysis of Nine Patients Treated in Korea," *Korean J. Radiol.*, vol. 21, no. 4, p. 494, 2020.