A Deep Learning-based Radiomics Approach for COVID-19 Detection from CXR Images using Ensemble Learning Model

Márcus V. L. Costa*, Erikson J. de Aguiar*, Lucas S. Rodrigues*,
Jonathan S. Ramos*†, Caetano Traina Jr.*, and Agma J. M. Traina*
*Institute of Mathematics and Computer Science (ICMC), University of São Paulo (USP), São Carlos, SP, BR 13566-590
{marcusvlc, erjulioaguiar, lucas_rodrigues, jonathan}@usp.br, {caetano, agma}@icmc.usp.br
†Federal University of Rondônia (UNIR), Porto Velho, RO, BR 76801-059
jonathan@unir.br

***Abstract*—Medical image analysis plays a major role in aiding physicians in decision-making. Specifically in detecting COVID-19, Deep Learning (DL) and radiomic approaches have achieved promising results separately. However, DL results are hard to interpret/visualize, and the radiomic approach encompasses successive steps, such as image acquisition, image processing, segmentation, feature extraction, and analysis. In this paper, we integrate DL with radiomic approaches, aiding in detecting COVID-19. We use DL models to extract 128 relevant deep radiomic features to assess COVID-19 from several image sources of 392 representative chest X-ray (CXR) exams. We avoid successive radiomic steps by employing DL (transfer learning) from Imagenet's VGG-16, ResNet50V2, and DenseNet201 networks. We considered a set of Machine Learning (ML) algorithms to further validate our results, providing an ensemble model to detect COVID-19. Our experimental results show that our approach achieved 95% AUC using 128 relevant features from DenseNet201. Conversely, our ensemble model presented 91% AUC, indicating that deep learning-based radiomics could increase binary classification performance in a real scenario. In addition, we highlight that our approach can be adapted to create other DL-based radiomics tools. For reproducibility, we made our code available at https://github.com/usmarcv/CBMS-DL-based-radiomics.**

***Index Terms*—Deep Learning, Deep Radiomics, Deep Features, Radiomics, Medical Images**

I. Introduction

In recent years, humanity has faced outbreaks of coronaviruses, commonly named COVID-19, leading to a pandemic that changed our way of life. Several methods can be employed to detect COVID-19, such as molecular analysis (RT-PCR), chest computed tomography (CT), and chest X-Ray (CXR) [1].

From a Computer-Aided Diagnosis (CAD) perspective, image-based techniques can offer a feasible way to identify COVID-19. For example, Deep Learning (DL) techniques have recently been widely used to classify medical images, such as CXR or CT, and achieved promising results [2]–[4]. However, DL models are seen as black boxes since learned patterns are challenging to interpret/analyze.

In a clinical routine, a black-box method introduces complexity in the widespread application/implementation of DL methods [3]. However, explainable methods with visual explanations can help the specialist understand DL decision-making methods, such as the Gradient-weighted Class Activation Mapping (Grad-CAM) technique [5]. Thus, the classical radiomics approach usually works by selecting a Region of Interest (ROI) and extracting a large number of features from it, employing these features as the basis for the analyses. Despite the limitations of DL models, the radiomic approach can be helpful in COVID-19 identification, complementing DL models and leading to a paradigm shift. We applied a fine-tuning technique to DL models to avoid quantitative imaging features, using the 128 most relevant ones.

Recent studies reveal methods of deep learning-based radiomics regarding its applicability and effectiveness in the automated identification and segmentation of brain metastases and gliomas [6], [7]. It exemplifies that high-throughput features from many medical images have the potential for clinical prediction [8]. We highlighted that the radiomics features are not predefined but can be learned from DL models, since deep learning-based radiomics automatically identifies and extracts high-dimensional features from the input images [9], [10]. However, DL models must improve their way of convincing the users of their procedures and analysis.

We propose a new approach using deep learning-based radiomics, which focuses on extracting core deep radiomics features from DL models to aid physicians' decision-making in COVID-19 cases. In summary, our work gives the following contributions: (i)

- a method to collect image datasets from different sources, processing and enhancing the quality of images using Contrast Limited Adaptive Histogram Equalization (CLAHE);
- an approach using deep learning-based radiomics to extract the most relevant features from the images;
- the design and evaluation of an ensemble learning model to improve the measures for classification tasks.

Paper outline. First, Section II describes the fundamental concepts concerning radiomics, Deep Learning and the relevant works in the literature related to our proposal. Section III formally presents our proposed approach using Deep Learning-based radiomics. Section IV offers and discusses the materials used to carry out experiments and the results achieved in this work. Finally, Section V summarizes the contributions and results of our work.

II. Background and Related Work

Medical imaging is a very attractive and promising medical science field since hospitals and clinical facilities faced digitalization. Recently, this scenario has become widespread due to the high demand for imaging exams and telemedicine. Considering, for example, human tissue/organ analysis with non-invasive procedures, each patient's information assessment capability makes a feasible and highly recommended routine for diagnosing, treating, and monitoring pathologies.

The radiomics approach aims at extracting quantitative features from medical imaging data [11]. This approach requires successive steps, such as image acquisition, image processing, segmentation, feature extraction, and analysis. However, human intervention is needed, usually by medical specialists in the area [12]. It is essential to highlight that the radiomics approach deals well with the inherent variation among the data patients, but can also improve the accuracy and precision of models from Artificial Intelligence (AI).

In contrast, deep learning-based radiomics automatically identifies and extracts high-dimensional values from the input data. On the other hand, the image segmentation step, which is usually costly and depends on specialist intervention, is optional for deep learning-based radiomics approaches. More specifically, different scaling and abstraction levels result in feasible, robust, and reliable models for pattern recognition or classification tasks [9]. The emphasis on learning is on successive layers, contributing to the depth and exploration of specific models [13]. However, the radiomics approach with DL can aid physicians' decision-making of procedures with more confidence [14].

There are several DL Model proposals inside approaches for medical image analysis based on Convolutional Neural Networks (CNN). These approaches have a grid-based format to process images, comprising spatial hierarchies of low and high-level features [15]. DL models based on CNN can assist physicians in diagnosing. Therefore, explainable methods can be helpful, such as Grad-CAM. The Grad-CAM uses salient maps to localize essential regions on the image classification [5]. Our work aims at getting the best of DL and the radiomics approach to support decision-making with visual assessment.

In recent years, several advancements were proposed for DL-based approaches to radiomics issues, applying quantitative features from medical images to build prediction models to assist decision-making on medical images. Therefore, we present and discuss deep learning-based radiomics approaches.

In the work of [8], the authors proposed an approach to encode the features learned in CNN (e.g., DarkNet-19 and ResNet50) layers using the Gaussian Mixture Methods (GMM), called CNN-GMM. The features are fed into an ML algorithm random forest to classify COVID-19 from pneumonia types, achieving up to 97% Accuracy and 99% AUC on a controlled image's dataset. In [16], the authors developed a DL model that integrates radiomic features for diagnosing COVID-19 in CXR images. Its distinguishing feature incorporates the extracted radiomic feature map in the CNN convolution process, VGG-16, VGG-19, and DenseNet121. The authors reported up to 97% Accuracy in differentiating between COVID-19 and Non-COVID-19, considering images from the same source.

Even though these works exploit deep learning-based radiomics, they work on images from the same source and do not provide clues about the data behavior during feature extraction and use. Our paper fills these gaps and evaluates the most relevant deep radiomics features in different ML algorithms and Ensemble Model learning.

III. Proposed Approach

In this work, we propose an approach using deep learning-based radiomics to classify COVID-19 cases. We present and discuss our proposal belonging to this section.

Our approach collects interoperable medical images from different sources, processing this data to improve the quality when extracting features. We employ DL models based on CNNs when extracting the most relevant deep radiomics features. Besides, we employ an evaluation of the core deep radiomics features by using a set of traditional ML classifiers to validate our approach. We introduce an Ensemble Model composed of ML classifiers to create a robust model, enhancing the classification tasks, especially for COVID-19 scenarios, bringing more assurance and confidence results to the physicians. We provide an end-to-end and visual assessment, easing COVID-19 analysis for making decisions. Therefore, our paper brings contributions concerning ways to evaluate deep learning-based radiomics and visual behavior from deep radiomics features assessment to specialists.

Figure 1 shows the proposed workflow. We organized our proposal into four steps: (i) data pre-processing; (ii) feature extraction based on deep radiomics; (iii) ensemble model classifier; and (iv) deep radiomics analysis for performance evaluation measurements and *visual assessment*.

A. Dataset Pre-processing

We collected COVID-19 images from diverse sources that comprise an interoperable medical dataset with balanced classes. The collected images come with different resolutions and contrast. Afterward, we pre-processed the exams to achieve their best quality: we employed the Contrast Limited Adaptive Histogram Equalization (CLAHE) since this technique enhances the contrast of the input image in operation on the local histogram information, especially from medical images [17]. Finally, we reshaped all images to 224x224 pixels with three color channels and normalized them to [0, 255] range for the best input data into the CNN architecture.

B. Deep Radiomics Features Extraction

We extracted core deep radiomic features from three pre-trained DL architectures based on CNNs, such as VGG-16, ResNet50V2, and DenseNet201, both pre-trained on ImageNet[1]. These DL models perform better than others when

[1]https://keras.io/api/applications/

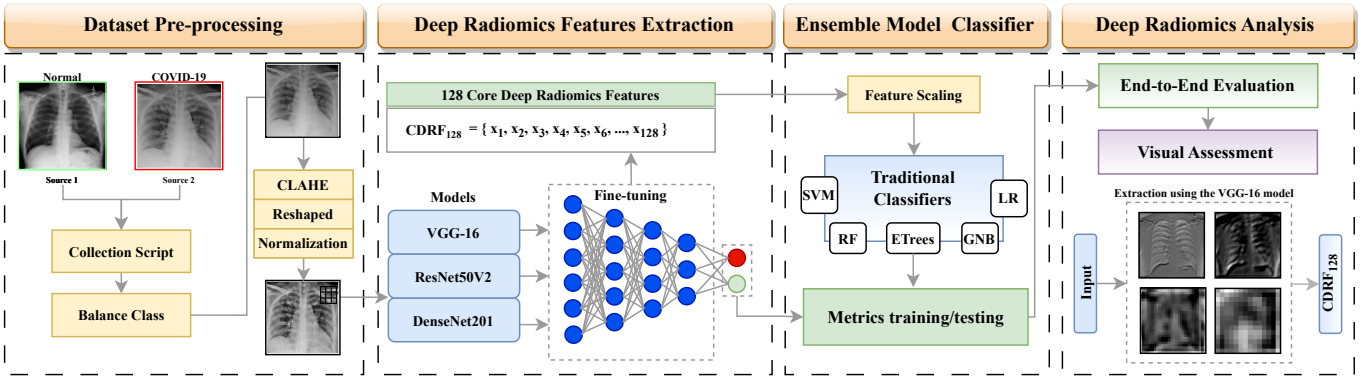

Fig. 1: Workflow employed in this study.

addressing the diagnosis of the classification task of COVID-19 using CXRs images [18]. We applied fine-tuning techniques using the transfer learning approach and the global average pooling layer to minimize overfitting, which reduces the total number of parameters. A sequence of four dense layers with 2048, 1024, 512, and 128 neurons employs the ReLU activation functions, and we added dropout layers. Further, we consider Adam[2] optimizer to update the network parameters with the batch size, learning rate, and epochs 32, $1e^{-4}$, and 100, respectively. We defined a 50% dropout rate randomly chosen and used early stopping[3] to avoid overfitting. The hyperparameters of our fine-tuning were chosen empirically after experimentation. Table I shows the top-down structure for each CNN architecture. Note that we employed the same fine-tuning structure for DL models.

TABLE I: Layer types used in our fine-tuning technique.

Layer (type)
VGG-16, ResNet50V2, DenseNet201 (Models)
global_average_pooling2d
dense (Dense)
dropout (Dropout)
dense_1 (Dense)
dropout_1 (Dropout)
dense_2 (Dense)
dropout_2 (Dropout)
deep_radiomic_feature (Dense)
dropout_3 (Dropout)

Finally, we extracted 128 core deep radiomics features of every image from the last dense layer before the classification layer (e.g., softmax). Also, the number of features was defined empirically after experimentation. The output data are structured as a vector of radiomic features with 128 neurons of the dense layer.

C. Ensemble Model Classifier

We propose an ensemble method to classify COVID-19 and Normal (healthy) CXR images, combining five ML classifiers as follows:

- **Support Vector Machine (SVM)**: finds the best hyperplane that separates the classification groups;
- **Gaussian Naive Bayes (GNB)**: a probabilistic model for binary and multi-class classification;
- **Logistic Regression (LR)**: determines the probability of an event occurring;
- **Random Forest (RF)**: creates randomized decision trees using the "bagging" method and increases the prediction performance;
- **Extra Tress (ETrees)**: uses several randomized decision trees on subsamples of the data set. With the combination of the results from each tree, find the final answer. ETress uses the mean to improve predictive accuracy and control over-fitting.

As aforementioned, the chosen ML classifiers allow the creation of a promising and robust model. We proposed the Ensemble Model, a mixture of learning algorithms to improve quality, classification and regression problems. Besides, we implemented a soft voting ensemble learning to make a final prediction with a better quality considering our model. Thus, we aim at providing the best classification or even regression results over COVID-19 data collection. We highlight that our Ensemble Model improves the classification and regression results over interoperable image collections, as we will show in the experiments. We employed the parameters of ML classifiers using standards from the *scikit-learn* library.

D. Deep Radiomics Analysis

One main task in deep radiomics is how we can provide insights regarding the deep learning approaches and their results. Our approach can present the decision-making behavior in extracting core deep radiomics features from DL models for the exams regarding COVID-19. Therefore, this behavior can be helpful in the clinical routine. The specialist can visualize the model's decisions and support the diagnosis using a method from deep learning-based radiomics. Besides, when considering related models, we can evaluate our ensemble model in classification and regression scenarios, particularly COVID-19.

[2]https://keras.io/api/optimizers/adam/
[3]https://keras.io/api/callbacks/early_stopping/

IV. Experimental Evaluation

This section presents materials, methods, and our experimental evaluation regarding the proposed approach. We first describe the representative database employed, collected from different sources and computational resources, in Section IV-A. We divide our experimental analysis into four main parts. First, in Section IV-B, we compared CNN end-to-end architectures on the dataset. In Section IV-C, we applied ML algorithms using the core deep radiomics features from the CNN architectures. Section IV-D presents a perspective of interpretability based on our approach. Finally, we discuss the results in Section IV-E.

A. Dataset and Setup Specifications

We collected and combined images from two public repositories with balanced classes, such as *covid-chestxray-dataset* [19] and *Chest X-Ray Images (Pneumonia)* [20]. The dataset comprised 392 CXR representative scan images, with 196 COVID-19-infected patients and 196 Normal (Healthy) patients. Table II shows the exact number of posterior-anterior (PA) view CXR images. In addition, the dataset split was randomized into 80% for the training set and 20% for the test set.

We performed the experiments on an RYZEN(R) 9 CPU, 32GB RAM, NVIDIA RTX3090 24GB RAM GPU. We implemented the deep learning and machine learning models using Keras with Tensorflow as a backend and Miniconda with CUDA toolkit[4] V11.2.67 on a Linux Pop!OS 5.11.

TABLE II: Dataset for class distributions for training and test.

Class type	Training	Testing
COVID-19	157	39
Normal	157	39
Total	**314**	**78**

B. Deep Learning Models Performance

We evaluated the classification performance of DL models and traditional ML classifiers with well-established measures, such as Accuracy, Area Under the Curve (AUC), F1-Score, Recall, and Precision. Table III shows the performance results of the end-to-end DL models that extract 128 core deep radiomic features for each model. The VGG-16 presented the best Accuracy, F1-Score, Recall, and Precision compared with the other models. DenseNet201 presented the best AUC, and ResNet50V2 performed worse than the other two models.

C. Traditional Classifiers Performance

We analyzed the performance of each ML classifier using 128 core deep features. Table IV shows the performance results with the traditional ML classifiers using the core deep radiomics features from deep learning-based radiomics.

- **SVM:** DenseNet201 achieves the best performance for accuracy, AUC, F1-Score, and Recall, as well as ResNet50V2 has the best Precision. Furthermore, the VGG-16 achieves lower results for these metrics.
- **GNB**: DenseNet201 reaches the best performance for all metrics, and the VGG16 is the worst.
- **LR**: DenseNet201 presents the best performance for all metrics. VGG-16 and ResNet50V2 have the worst.
- **RF**: DenseNet201 presents the best results for all metrics, and the VGG-16 and ResNet50V2 have the lowest.
- **ETrees**: DenseNet201 presents the best results for all metrics, and the VGG-16 and ResNet50V2 have the lowest.
- **Ensemble Learning**: DenseNet201 achieves the best performance for all metrics, and the VGG-16 and ResNet50V2 have the lowest ones.

[4]https://anaconda.org/anaconda/cudatoolkit

TABLE III: Performance measures for DL models end-to-end used to extract core deep radiomics features. Best values in bold.

Model	Accuracy	AUC	F1-Score	Recall	Precision
VGG-16	**89.9%**	93.7%	**89.9%**	**89.9%**	**89.9%**
ResNet50V2	82.3%	86.6%	82.3%	82.3%	82.3%
DenseNet201	88.6%	**95.0%**	88.6%	88.6%	88.6%

TABLE IV: Results for traditional ML classifiers.

Classifier	DL Model	Accuracy (%)	AUC (%)	F1-Score (%)	Recall (%)	Precision (%)
SVM	VGG-16	81.0	81.2	80.8	81.2	82.7
	ResNet50V2	83.5	83.7	83.3	83.7	86.3
	DenseNet201	86.1	86.1	86.1	86.1	86.1
GNB	VGG-16	70.5	78.7	78.0	78.7	81.9
	ResNet50V2	82.3	82.4	82.0	82.4	84.4
	DenseNet201	88.6	88.7	88.5	88.7	89.8
LR	VGG-16	82.3	82.5	81.8	82.5	86.8
	ResNet50V2	86.1	86.2	86.0	86.2	86.6
	DenseNet201	91.1	91.2	91.1	91.2	91.3
RF	VGG-16	82.3	82.5	81.9	82.5	85.5
	ResNet50V2	83.5	83.7	83.4	83.7	85.3
	DenseNet201	86.6	86.6	86.6	86.6	86.6
ETrees	VGG-16	83.5	83.7	83.3	83.7	86.3
	ResNet50V2	82.3	82.4	82.1	82.4	83.6
	DenseNet201	91.1	91.2	91.1	91.2	91.2
Ensemble Learning	VGG-16	83.5	83.7	83.3	83.7	86.3
	ResNet50V2	84.8	85.0	84.6	85.0	87.1
	DenseNet201	**91.1**	**91.2**	**91.1**	**91.2**	**91.3**

Best value in **bold**.

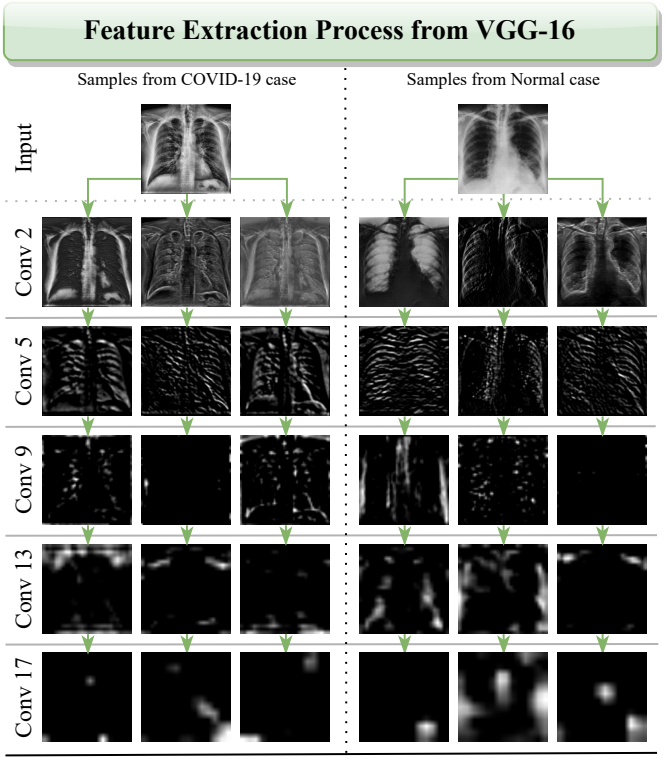

Fig. 2: Visualization of the core deep radiomic feature extraction process in the VGG-16 model. Each convolution block of the network shows different textures, gray levels, and shapes for a COVID-19 and a Normal case.

D. Visual Assessment from Deep Radiomics Analysis

We analyzed the visual interpretability of the feature extraction process for the VGG-16 model, presenting three samples from COVID-19 and Normal cases in Figure 2. The convolutional (Conv) process in DL models tries to understand the basic patterns of our images from input data. Note that the feature propagation is throughout the extraction process since DL models learn directly from the CXR images and keep the information regarding image textures, colors, and shapes. Regarding the DL convolutions operations, in Conv 2, we can highlight that there is still a shape of lung radiography. Note that essential features are propagated through the network during Conv 5, 9, and 13, resulting in Conv 17 only activation maps. Therefore, we can visualize the features process before extracting core deep radiomics features.

We applied the Grad-CAM[5] to investigate our proposed approach. The Grad-CAM provides a visual explanation algorithm, and we chose a COVID-19 case in our dataset to analyze it. Figure 3 shows the behavior from the feature extraction process in the VGG-16 model and when using Grad-CAM. Note that we tested an image from the convolutional process in each convolutional layer and compared it with Grad-CAM. We can observe in Conv 2 that our prediction is COVID-19 (0), and the heatmaps from Grad-CAM identify the lung. Therefore, Conv 5, 9, 13, and 17 show to predict Normal (Healthy - 1), and the Grad-CAM identifies the borders from each image. Finally, our output with Grad-CAM classifies correctly and shows the heatmap of the lung. Thus, our approach allows us to understand and visualize the interpretability extracted from the core deep radiomics features from the VGG-16 model.

[5]https://keras.io/examples/vision/grad_cam/

E. Discussion

Firstly, our results revealed that VGG-16 provided the best DL end-to-end model classification, and ResNet50V2 gave the worst performance. VGG-16 presented fewer trainable parameters than other models when we applied transfer learning and fine-tuning technique. DenseNet201 achieved the best AUC because it has shorter connections between layers and strengthens feature propagation. Afterward, we extract the 128 core deep features from each DL model. We tested these features into traditional ML classifiers. We highlight these tests to provide more confidence in our proposed approach.

In experiments of traditional classifiers, we highlight the best result provided by LR, RF, ETrees, and Ensemble Model generated from core deep radiomics features from DenseNet201. We observe this behavior since the RF and ETrees are based on decision trees. These algorithms provide the users with more interpretability in decision-making based on binary classification. The same happens with LR, but it employs a statistical method. Our Ensemble Model performs the best Accuracy, F1-Score, Recall, and Precision when comparing DL end-to-end and traditional ML classifiers, up to 2% based on heterogeneous datasets.

Additionally, the proposed approach provides a deep learning-based radiomics perspective of interpretability. We can observe from the VGG-16 model the extraction process behavior from core deep radiomics features. Thus, we employ the Grad-CAM as a support tool to provide Visualization to specialists. The hyperparameters in the CNN are trained without human knowledge intervention and are usually hard to interpret. Our results revealed that ML classifiers perform better using core deep radiomics features. Therefore, ML algorithms working with DL make the last ones more interpretable in decision-making classification tasks.

V. Conclusion

Our work proposed and evaluated deep learning-based radiomics to classify COVID-19 infected and Normal (Healthy) patients from CXR images. Our approach allows the use of interoperable datasets, and we enhanced the image quality of our dataset using CLAHE. We fine-tuned the DL models by adding new layers to extract core deep radiomics features. Besides, we compared the performance of DL models end-to-end and ML traditional algorithms using core deep radiomics features.

The results show that VGG-16 achieves the best accuracy and DenseNet201 the best AUC for end-to-end performance. Hence, we extracted the core radiomic features from each DL model and evaluated them using several well-known ML

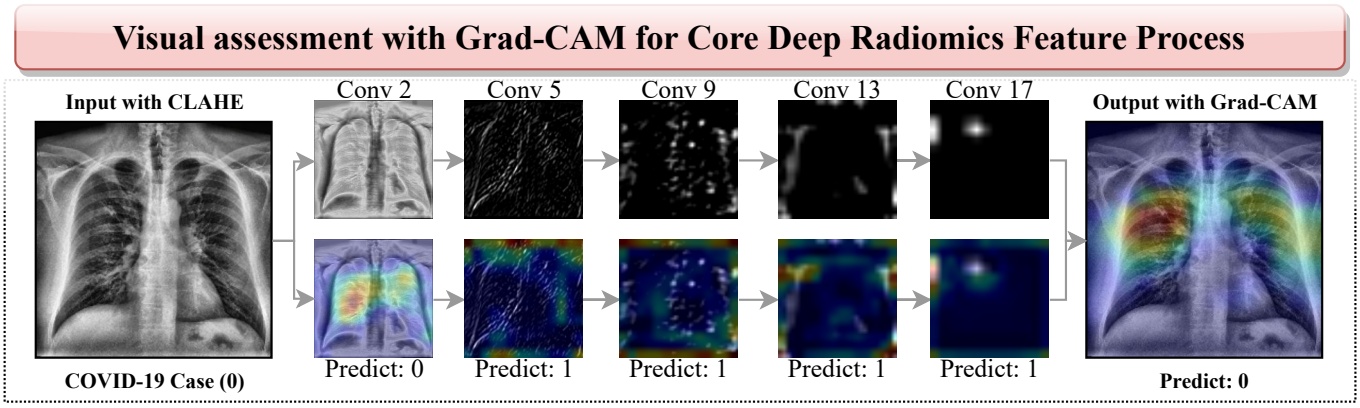

Fig. 3: Visual assessment with Grad-CAM from a COVID-19 case, showing the behavior from Grad-CAM about one example extracted from each convolution process. Remark that COVID-19 case is labeled 0 and Normal (Healthy) is 1.

traditional classifiers. Our Ensemble Model learning mixed the ML classifiers to improve classification tasks and achieves measures of Accuracy, AUC, F1-Score, Recall, and Precision exceeding 91% for binary classification in heterogeneous datasets. Therefore, the visual assessment allows us to observe and understand the feature extraction process in this core deep radiomics on DL models.

This study is the first step in evaluating deep learning-based radiomics with visual assessment and performance evaluation. We intend to extend our study in future work by applying the same proposed pipeline to more extensive multiclass datasets and employing other DL models and traditional ML classifiers.

Acknowledgment

This study was financed in part by the São Paulo Research Foundation (FAPESP – grants 2016/17078-0, 2020/07200-9, and 2021/08982-3), National Council for Scientific and Technological Development (CNPq – grants 152760/2021-0, 308544/2021-8, 308738/2021-7), and Coordination for Higher Education Personnel Improvement (CAPES – grant 001).

References

[1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu *et al.*, "A novel coronavirus from patients with pneumonia in china, 2019," *New England journal of medicine*, 2020.

[2] V. Guarrasi, N. C. D'Amico, R. Sicilia, E. Cordelli, and P. Soda, "A multi-expert system to detect COVID-19 cases in X-ray images," in *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2021, pp. 395–400.

[3] Q. Ye, J. Xia, and G. Yang, "Explainable AI for COVID-19 CT classifiers: an initial comparison study," in *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2021, pp. 521–526.

[4] P. Ghose, M. A. Uddin, U. K. Acharjee, and S. Sharmin, "Deep viewing for the identification of Covid-19 infection status from chest X-Ray image using CNN based architecture," *Intelligent Systems with Applications*, vol. 16, p. 200130, 2022.

[5] B. H. Van der Velden, H. J. Kuijf, K. G. Gilhuijs, and M. A. Viergever, "Explainable artificial intelligence (XAI) in deep learning-based medical image analysis," *Medical Image Analysis*, p. 102470, 2022.

[6] P. Prasanna, A. Karnawat, M. Ismail, A. Madabhushi, and P. Tiwari, "Radiomics-based convolutional neural network for brain tumor segmentation on multiparametric magnetic resonance imaging," *Journal of Medical Imaging*, vol. 6, no. 2, pp. 024 005–024 005, 2019.

[7] E. Grøvik, D. Yi, M. Iv, E. Tong, D. Rubin, and G. Zaharchuk, "Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI," *Journal of Magnetic Resonance Imaging*, vol. 51, no. 1, pp. 175–182, 2020.

[8] A. Chaddad, L. Hassan, and C. Desrosiers, "Deep radiomic analysis for predicting coronavirus disease 2019 in computerized tomography and X-ray images," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 1, pp. 3–11, 2021.

[9] P. Lohmann, K. Bousabarah, M. Hoevels, and H. Treuer, "Radiomics in radiation oncology—basics, methods, and limitations," *Strahlentherapie und Onkologie*, vol. 196, pp. 848–855, 2020.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[11] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, 2016.

[12] B. Koçak, E. Ş. Durmaz, E. Ateş, and Ö. Kılıçkesmez, "Radiomics with artificial intelligence: a practical guide for beginners," *Diagnostic and interventional radiology*, vol. 25, no. 6, p. 485, 2019.

[13] F. Chollet, *Deep learning with Python*. Simon and Schuster, 2021.

[14] J. E. Van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, and B. Baessler, "Radiomics in medical imaging—"how-to" guide and critical reflection," *Insights into imaging*, vol. 11, no. 1, pp. 1–16, 2020.

[15] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[16] Z. Hu, Z. Yang, K. J. Lafata, F.-F. Yin, and C. Wang, "A radiomics-boosted deep-learning model for COVID-19 and non-COVID-19 pneumonia classification using chest x-ray images," *Medical physics*, vol. 49, no. 5, pp. 3213–3222, 2022.

[17] E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, and S. M. Pizer, "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," *Journal of Digital imaging*, vol. 11, pp. 193–200, 1998.

[18] M. Roberts, D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, A. I. Aviles-Rivero, C. Etmann, C. McCague, L. Beer *et al.*, "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans," *Nature Machine Intelligence*, vol. 3, no. 3, pp. 199–217, 2021.

[19] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "COVID-19 image data collection: Prospective predictions are the future," *arXiv 2006.11988*, 2020. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset

[20] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.