IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 14 March 2024, accepted 26 March 2024, date of publication 1 April 2024, date of current version 5 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3383781

RESEARCH ARTICLE

Leveraging Radiomics and Genetic Algorithms to Improve Lung Infection Diagnosis in X-Ray Images Using Machine Learning

A. BEENA GODBIN AND S. GRACELINE JASMINE
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai 600127, India
Corresponding author: S. Graceline Jasmine (graceline.jasmine@vit.ac.in)

This work was supported by Vellore Institute of Technology, Chennai.

ABSTRACT Radiomics, an emerging discipline in medical imaging, focuses on extracting detailed quantitative features from images to unveil subtle patterns imperceptible to the naked eye. This study specifically employs radiomics and machine learning techniques to discern cases of viral pneumonia and COVID-19. By harnessing intricate radiomic features derived from medical images, the objective is to train a machine learning model capable of accurately distinguishing between patients with viral pneumonia, COVID-19, and those unaffected. To optimize the performance of machine learning models, the paper incorporates genetic algorithms for hyperparameter optimization. A comparative analysis is conducted among the genetic algorithm-based TPOT (Tree-based Pipeline Optimization Tool) settings, namely TPOT-Default, TPOT-Light, and TPOT-Sparse, to select the most effective hyperparameters. Custom modifications are introduced to the TPOT model to align it with the specific requirements of the current model, resulting in a noteworthy achievement. The proposed model attains a remarkable 94% AUC (Area Under the Curve) when employing the Random Forest algorithm. Furthermore, the study systematically evaluates the execution time taken by each TPOT model. The model's performance is comprehensively assessed through key metrics, including accuracy, precision, sensitivity, specificity, and F1-score. The techniques suggested in this article could aid radiologists in identifying anomalies in chest X-ray (CXR) images, offering a more accurate and efficient interpretation to improve medical decision-making.

INDEX TERMS Radiomic feature, chest X-ray, viral pneumonia, COVID-19, machine learning.

I. INTRODUCTION

The latest viral epidemic known as Covid-19 originated in Wuhan City, China at the beginning of this year. As a result of its rapid spread, nearly every country on earth has been affected [1]. Due to its high transmissibility, the primary mode of spreading is through inhaling droplets containing the infectious agent. An outbreak of this highly infectious virus that spread rapidly around the world has caused a worldwide health emergency [2]. A virus called Covid-19 is officially referred to as Coronavirus 2 (SARS-CoV-2) which is a member of the family of Coronaviruses. Over the last few decades, several respiratory diseases, such as pneumonia, cardiomegaly, pleural effusion, tuberculosis, and most recently, COVID-19, have evolved into epidemics among humans. Among these ailments, COVID-19 stands out as the most life-threatening. Indeed, pneumonia alone impacts approximately 7% of the global population, affecting around 450 million people and contributing to nearly 2 million deaths annually [3]. Over the past three years, the impact of COVID-19 has been staggering, with an estimated 6.27 million lives lost and approximately 522 million individuals affected by the virus [4]. Approximately 1.3 million deaths were projected to occur from tuberculosis in the year 2020; therefore, in overall, 4 million people will die every year from lung-related diseases [5]. Radiologists usually detect these types of diseases by taking CXR images of the lung to detect them. This is a cheap and easy process that only requires a few

The associate editor coordinating the review of this manuscript and approving it for publication was Hengyong Yu.

steps to diagnose these diseases. The number of computed tomography (CXR) images taken for medical purposes alone in the US is more than 35 million each year [6]. The average number of CXR images radiologists assess every day has risen to over 100, which has led to increased workload, fatigue, and incorrect diagnoses for some radiologists. Due to this reason, the conventional process tends to take a long time and costs a lot of money. In addition, due to human error, the diagnosis made by the radiologists could also differ from the diagnosis made by the patient's physician. Accurate diagnosis is of utmost importance during a pandemic, ensuring timely treatment, efficient resource allocation, effective public health measures, reduced transmission rates, and prevention of misdiagnosis. The sooner predictions can be made for conditions like COVID-19, tuberculosis, pneumonia, etc., the greater the opportunity to save millions of human lives. Rapidly spreading variants of COVID-19, such as Omicron and Delta underscore the urgency of proactive measures.

The rapid spread of newer variants of COVID-19 poses a challenge for medical specialists to meet the heightened demand for timely treatment. Consequently, the utilization of automated technologies trained to predict symptoms associated with various lung-related anomalies through specific chest X-ray (CXR) images holds the promise of enhancing confidence and accuracy for radiologists and medical doctors in the prompt and precise diagnosis of such conditions.Based on the results of the CXR images, machine learning, and deep learning techniques have been applied to develop a computer-aided intelligent system that can detect lung-related diseases automatically by using the CXR images as data. Efforts have been made in the past to automate the identification of lung-related abnormalities using CT images, driven by the recent availability of more extensive datasets for analysis. Medical imaging, especially X-ray examinations, is an important part of the early detection and diagnosis of various pulmonary 37, including lung infections, and a crucial part of early detection and diagnosis. Due to the ever-increasing volume of medical data available today, it is imperative that advanced computational techniques can be integrated into diagnostic processes to improve their accuracy and efficiency as well. The motivation behind the proposed model is to enhance accuracy in diagnosing viral pneumonia and COVID-19 from medical imaging data. It utilizes radiomics and machine learning techniques to extract subtle patterns from images, overcoming interpretability challenges. The goal is to optimize machine learning performance and assist radiologists in making more accurate clinical decisions, thereby improving patient outcomes. The purpose of this study is to examine the synergies between radiomics, machine learning, and genetic algorithm optimization to advance the state-of-the-art in lung infection detection using X-ray images to advance the state-of-the-art in lung infection detection. Radiomics is one of the most technologically advanced disciplines within medical imaging, which aims to extract quantitative information from radiographic images. Using these features, we can capture intricate details embedded within images, extending beyond the capabilities of conventional visual interpretations of images. With radiomics, this study can leverage a wealth of information to uncover subtle patterns and nuances that would otherwise be difficult to detect without radiomic analysis. It has been demonstrated that machine learning algorithms, specifically those training on large datasets, have demonstrated remarkable success when it comes to classifying images. The challenge lies in optimizing these algorithms to achieve peak performance for specific diagnostic objectives. This problem is addressed by incorporating genetic algorithm optimization as part of the approach. Using genetic algorithms, inspired by natural selection, machine learning models are iteratively refined to ensure they evolve to optimize configurations to detect lung infections. The forecasting process is typically a multi-stage process involving several steps including data preprocessing, feature extraction, and model fitting, which are all steps in the forecasting pipeline. Each of these steps requires choices to be made, and corresponding hyperparameters to be tuned, which may have an impact on the accuracy of predictions. The manual tuning of an algorithm often relies on simple assumptions, which may not fully capture the characteristics of the underlying data, which results in an unsatisfactory result. Automation of machine learning (AutoML) is a rapidly growing field of AI research that allows users to construct fully automated pipelines for classification and regression tasks that are capable of performing high-performance classification and regression. A variety of applications have shown that AutoML can be used to automatically extract high-quality models that are competitive with manual tuning, often outperforming the manual-tuned models in various applications. We used the TPOT model, and we compared the results of models with and without sparse settings, as well as those configured with light settings. In addition, tailored adjustments were made to the TPOT model so that it would align with the specific requirements of the proposed model to improve its accuracy and performance. As a whole, this paper provides a framework that integrates radiomics, machine learning, as well as genetic algorithm optimization for the detection of lung infection in X-ray images based on the use of radiomics. This paper discusses the methodology, experimental setup, and results obtained. It sheds light on the potential of this approach to significantly improve diagnostic accuracy in pulmonary infections. The primary contributions of this work are as follows:

- This paper introduces a novel approach that involves utilizing gray scale conversion and binary thresholding techniques to generate segmented lung images from chest X-ray images along with their corresponding masks
- A custom radiomics feature extraction technique was employed to categorize lung disease detection using chest X-ray images. The software provides a wide range of radiomic features that can be used to extrapolate

quantitative information from medical images using a variety of radiomics tools.

- In the study, a customized TPOT model was created. This customization involved fine-tuning the model parameters, feature selection strategies, and other aspects to better align with the unique characteristics and demands of the specific dataset and objectives. The goal was to create a TPOT model that could effectively capture the nuances of the dataset and yield improved results for the specific use case. A 94% AUC score was achieved for the TPOT model as proposed.
- To validate the effectiveness of the customized TPOT model, a thorough comparison was conducted with other TPOT- Light, Sparse, Default configuration models that employed different configurations or default settings. This comparative analysis aimed to assess whether the customizations made to the TPOT model resulted in performance improvements compared to the out-of-the-box or default versions of TPOT.
- Moreover, the systematic examination of the execution time taken by each TPOT model indicates that, beyond the initial discussion or action mentioned, the focus of the investigation extended to a systematic.

The remaining sections of the manuscript are structured as follows. An overview of relevant prior work in the field is presented in Section II. Section III provides a detailed explanation of the methodology that has been proposed. Section IV discusses the experimental results that were obtained during the study. In Section V, the paper concludes with a discussion of future research directions that could be pursued in the future.

II. LITERATURE REVIEW

Loey et al. employed a CNN model to extract key features of COVID-19, followed by a Bayesian model for patient classification based on these features [7]. Utilizing 3,616 COVID-19 data points, they achieved a notable accuracy of 96%. Bhattacharyya et al. developed the COVIDCXNet model by using a pre-trained CheXNet method, trained with 1,326 COVID-19 images [8]. Through preprocessing techniques like CLAHE(Contrast Limited Adaptive Histogram Equalization), they achieved an accuracy of 87.88%. Ieracitano et al. proposed a fuzzy logic-based DL (deep learning) technique to identify CXR images of patients with idiopathic pneumonia from those with COVID-19 pneumonia [9]. With 121 images, their AI model reached an accuracy of about 81%, offering a diagnostic method for pneumonia. Agrawal and Choudhary designed a deep CNN for COVID-19 detection, using the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance [10]. Testing with 1,525 COVID-19 samples, they achieved 96% and 94.45% accuracy for two and three classes, respectively. Gayathri et al. utilized pretrained CNN models like Resnet101 and InceptionResnetV2 applied a sparse autoencoder for image dimension reduction, and employed a FNN(feedforward neural network) for COVID-19 detection [11]. Trained on 504 COVID-19 images, their models demonstrated an accuracy of 95.78% and an AUC of 98.21%. Kassani et al. compared different transfer learning models for detecting COVID-19 in both computed tomography (CT) and chest X-ray (CXR) images [12]. According to Angeli et al. [13], similar to clinical data and laboratory results, CT was shown to be able to provide prognostic information. As a means of selecting features and training a model, Univariate and Multivariate Logistic Regression are used to predict whether there will be improvement/recovery compared with ICU admission or death. The PC score was not significantly associated with the PI score, and an AUC (area under the curve) of 0.722 was obtained when taking into account only the PI value when considering the results. It was demonstrated that 14339 Chest Tomography images were used in [14] to predict whether a patient would survive or not (alive or dead). Radiomic features such as texture, intensities, and shapes were calculated based on COLI-Net segmentation of lungs [15] as well as ANOVA, Kruskal-Wallis, Recursive Feature Elimination, and Relief algorithms were used to select the features. To categorize the data, they used logistic regression, linear discriminant analysis, LASSO, a Random Forest, a Naive Bayes, AdaBoost and a Multilayer Perceptron. In comparison to other ANOVA feature selectors and Random Forests, the amalgamation of ANOVA feature selector and Random Forests demonstrated superior performance, achieving an AUC of 0.83 (95% CI: 0.81-0.85), a sensitivity of 0.81, and a specificity of 0.72. A study conducted in [16] used a court of 188 patients to detect whether the disease progression would worsen or improve over time. An automatic algorithm was used to generate ROI (Regions of Interest) encompassing lesions, and these ROIs were subsequently revised manually based on the results obtained from 188 CT scans. Their approach included integrating radiomic data with demographic data and laboratory data that were selected based on the results of F-tests and ICC. Linear regression, decision tree, SVM, Random Forest, and XGBoost methods were employed to explore the clinical signature, the radiomic signature alone, and their combination. There was an AUC of 0.843, 0.813, and 0.865 for each of the radiomics, clinical, and combined features in the test set when comparing the radiomics, clinical, and combined features. Using 284 CT images from [17], the advancement of COVID-19 was categorized into four groups based on how the scans within these CT images were classified: early, progressive, severe, and absorption. A total of 1688 radiomic features were extracted from the ROIs, which were segmented manually and included logarithm, wavelet, gradient, exponent, square, local binary pattern, and square root features in addition to original features. The ElasticNet algorithms and select K-best method were used to select thirty-eight radiomic features from the data. After training the SVM, the test dataset resulted in a microaudit score of 0.89 and a macroaudit score of 0.90 on the basis of a microaverage AUC. As outlined in [18], the study utilized a primary cohort consisting of 157 COVID-19

patients and a validation cohort of 105 COVID-19 patients sourced from three distinct hospitals. The objective of the study was to develop a radiomic nomogram for assessing the severity of infection in COVID-19 patients, distinguishing between Mild/Moderate and Severe/Critical cases. Using the Rad-Score, radiomic signature, comorbidity, and abnormal White Blood Cell counts, a multivariate logistic regression model has been trained obtaining an AUC of 0.988 in the validation cohort in combination with the radiomic signature, Rad-Score, and comorbidities. It has been reported that in the study by [19], 820 CXR images were used in order to predict prognosis (MILD vs. SEVERE). The researchers aimed to assess shallow learning (SL) methods like SVM, and deeper learning (DL) methods such as MLP and DL. They evaluated three approaches–handcrafted, hybrid, and end-to-end deep learning model using CXR images alone and in combination [20]. The same model was also proposed in [21] for several deep model based architectures on DenseNet-121 and ResNet-18 algorithms. Using the best classification algorithm, we computed a sensitivity of 0.79, a specificity of 0.82, and an AUC of 0.84. Grad-CAM was used as the explainability algorithm. Several convolutional models and dense network models have been proposed in [22] for the prediction of prognosis (MILD vs. SEVERE). Godbin et al. proposed utilizing machine learning models with radiomics, like GLCM features from CT scans, to achieve 99.94% accuracy in distinguishing COVID-19 from similar conditions like pneumonia [23]. In study [24], the authors introduced the Brixia score as a method to evaluate COVID-19 infections. The evaluation involved dividing each image into six zones, covering the middle, upper, and lower sections of both the right and left lungs. Using the Brixia score, an experienced thoracic radiologist was able to determine whether or not 100 CXRs were recovered or if they were dying and assigned the score manually. For comparisons between CXR scores and outcomes in selected patients, Mann-Whitney U-tests and weighted Kappa (kw) were used (kw, 0.82, 95% CI, 0.79-0.86). There was a proposal in [25], utilizing 5000 CXR images, to estimate Brixia score using the BS-net. A particular end-to-end scheme was developed to execute the Brixia score segmentation, aligning, and prediction using the BS-Net as an end-to-end method. The study utilized ResNet for feature identification, Nested U-Net for segmentation, and synthetic transformations for alignment [26]. This process included rotation, shift, scale, elastic transformation, optical distortion and grid distortion along with optional hard self-attention. The Brixia score matrix was computed using the ROI pooling method, and the Feature Pyramid Network combined features across various scales [27]. The BS-Net demonstrated high accuracy in predicting Brixia scores and other relevant scores like Toussie and GE-LO. The study proposed an approach inspired by LIME [28] to enhance Grad-CAM's explainability using super-pixels. This involved comparing probability maps from each replica with those from the model, improving the understanding of model activities in lung areas and addressing Grad-CAM's limited localization capability. F. Prinzi et.al proposed the significance of employing explainable AI algorithms in clinical settings to ensure the interpretability of predictive models for COVID-19 prognosis. This emphasis on interpretability facilitates informed clinical decision-making [19]. De et al. [29] introduced a widely practiced intervention for modifying cardiac health, highlighting the varied effects of physical activity on older adults. However, current machine learning (ML) models struggle to explain predictions at a personalized patient level, hindering their application in tailoring interventions for individual needs and optimizing patient outcomes. Papanikolaou et.al focuses on avoiding methodological errors and developing clinically meaningful Radiomic Signatures [30]. Zekun Jiang suggested that Wavelet transformation could improve computed tomography texture features. This study compares various wavelet kernels to assess their impact on predictive radiomic models [31].In their work, they undertake a thorough examination, comparing various wavelet kernels and gauging their effects on predictive radiomic models [32].

III. METHODOLOGY

This article aims to create machine learning models for predicting the prognosis of COVID-19 patients with a specific emphasis on ensuring the explainability of the models as established in Section I of this article. The overall workflow is shown in figure 1 below. Each block of the processing pipeline is described in more detail in the following subsections. The proposed architectural diagram outlines the model designed for the identification of lung boundaries within medical images, comprising four essential modules. The initial module, named lung segmentation, is dedicated to delineating the boundaries of the lungs within the image. Following this, the feature extraction module utilizes the ''pyradiomics'' tool [33] to extract pertinent features from the segmented lungs. This step is crucial in various medical applications where precise lung segmentation is a prerequisite for subsequent in-depth analysis. The feature extraction process is pivotal for discerning different types of lung abnormalities. Features such as shape, texture, and intensity play a significant role in identifying and characterizing these abnormalities. The extracted features serve as crucial inputs for subsequent stages of the model, contributing to a more nuanced and accurate analysis of medical images. The third module focuses on model selection and hyperparameter tuning. This step is essential for optimizing the outcomes of the machine learning model used for classification. Leveraging TPOT techniques, this module aids in selecting the most suitable machine learning model and fine-tuning its hyperparameters.

The goal is to achieve the highest levels of accuracy in the classification of lung diseases, enhancing the model's efficacy in providing meaningful insights from medical images. The fourth module, titled ''classification,'' involves the application of the selected machine learning model to the extracted features. This results in the creation of

a classification of lung abnormalities, with the output representing a predicted class of the specific lung disease. This classification holds immense value in the diagnostic process, guiding healthcare professionals in planning an effective treatment strategy for the patient. Collectively, this block diagram offers a comprehensive overview of how machine learning methodologies can be harnessed to aid in the diagnosis of lung diseases. The inclusion of sample images from the chest X-ray dataset in Figure 2 further enhances the visual representation, providing a tangible connection between the architectural components and the real-world application of the model.

A. LUNG SEGMENTATION

Pyradiomics played a pivotal role in extracting radiomic features from the image data. In conjunction with the medical image required by Pyradiomics, the inclusion of a mask image defining the Region of Interest (ROI) was imperative. The images are in RGB format but appear grayscale, so we converted them into grayscale images. To ensure uniformity, the mask image needed to share the same format, dimensions, and voxel size as the medical image. The mask image needed to have non-zero values in voxels within the ROI and zero values outside the ROI. Figure 3 provides an illustrative example, featuring a sample image alongside its corresponding segmented mask. To create the segmented mask, a chest X-ray scan image was employed along with a corresponding mask image for each X-ray image. To match the input size of a machine learning model (299 × 299 pixels), the mask image underwent resizing. Additionally, the mask image underwent a conversion to grayscale, simplifying subsequent processing steps. To maintain consistency, a similar grayscale conversion was applied to the original image. Subsequently, grayscale masks were transformed into binary masks through the application of a binary threshold. The binary mask, as a final step, was then applied to the original image to generate a masked image. Algorithm 1 outlines the sequential steps involved in this process. The bitwise AND operation is employed to retain only those pixels in the original image that correspond to non-zero pixels in the binary mask. The resultant masked image can be displayed or subjected to further processing, representing a crucial preparatory step for subsequent radiomic feature extraction and analysis. Figure 3 displays both the mask and the segmented mask. This meticulous approach ensures that the radiomic analysis is conducted with precision and accuracy, contributing to the overall effectiveness of the model in identifying lung abnormalities from medical images.

Algorithm 1 Proposed Lung Segmentation Algorithm:

1: Input:Read the chest X-ray scan image and its corresponding mask image.
OriginalImage,Mask = ReadImages [XRayImagePath, MaskImagePath]
2: Resize the mask image to 299 × 299 pixels:
ResizedMask = ResizeImage [Mask, 299, 299]
3: Convert the resized mask image to grayscale:
GrayscaleMask [i,j] = ConvertToGrayscale [ResizedMask(i,j)]
4: Convert the original chest X-ray scan image to grayscale:
GrayscaleImage [i,j] = ConvertToGrayscale [OriginalImage(i,j)]
5: Convert the grayscale mask to a binary mask using a binary threshold:
BinaryMask [i,j] = 1 if GrayScaleMask [i,j] > Threshold; otherwise 0
6: Apply the binary mask to the original grayscale image using bitwise AND operation:
FUNCTION ApplyBinaryMask(GrayscaleImage, BinaryMask):
rows = number of rows in GrayscaleImage
columns = number of columns in GrayscaleImage
MaskedImage = create a new matrix of size rows x columns, initialized with zeros
7: **for** $i = l$ to rows: **do**
8: **for** $j = l$ to columns: **do**
pixelValueG = GrayscaleImage[i, j]
pixelValueB = BinaryMask[i, j]
resultBinary = pixelValueG BITWISE AND pixelValueB
MaskedImage[i,j] =ConvertToInteger(resultBinary)
9: **end for**
10: **end for**
11: **return** MaskedImage [where i,j represents pixel coordinates]
12: Output: Display or process the masked image as needed:
DisplayOrProcess [MaskedImage]

B. DATASET

This study uses Kaggle's COVID-19 Radiography open data set [37] to build the dataset for this study. It contains a chest X-ray that is normal (i.e. healthy) and shows COVID-19, viral pneumonia, and normal (i.e. healthy) chest X-rays. There are X-ray images included in the dataset as well as masks that correspond to the X-ray images. The image distribution consists of 500 COVID-19 images, 500 images of viral pneumonia, and 500 images of normal X-rays. Figure 1 illustrates a sample image from each class in terms of its composition and content. The stratified selection approach was then used to divide the data set into two parts at random, and 80% of the images were destined for the training stage, while twenty percent were destined for the validation stage.

C. RADIOMIC FEATURE EXTRACTION

Radiomic features constitute a collection of quantitative descriptors extracted from medical images, including X-rays, CT scans, or MRIs, employing the field of radiomics. These features serve to encapsulate a diverse array of information pertaining to the spatial distribution of pixel intensities,

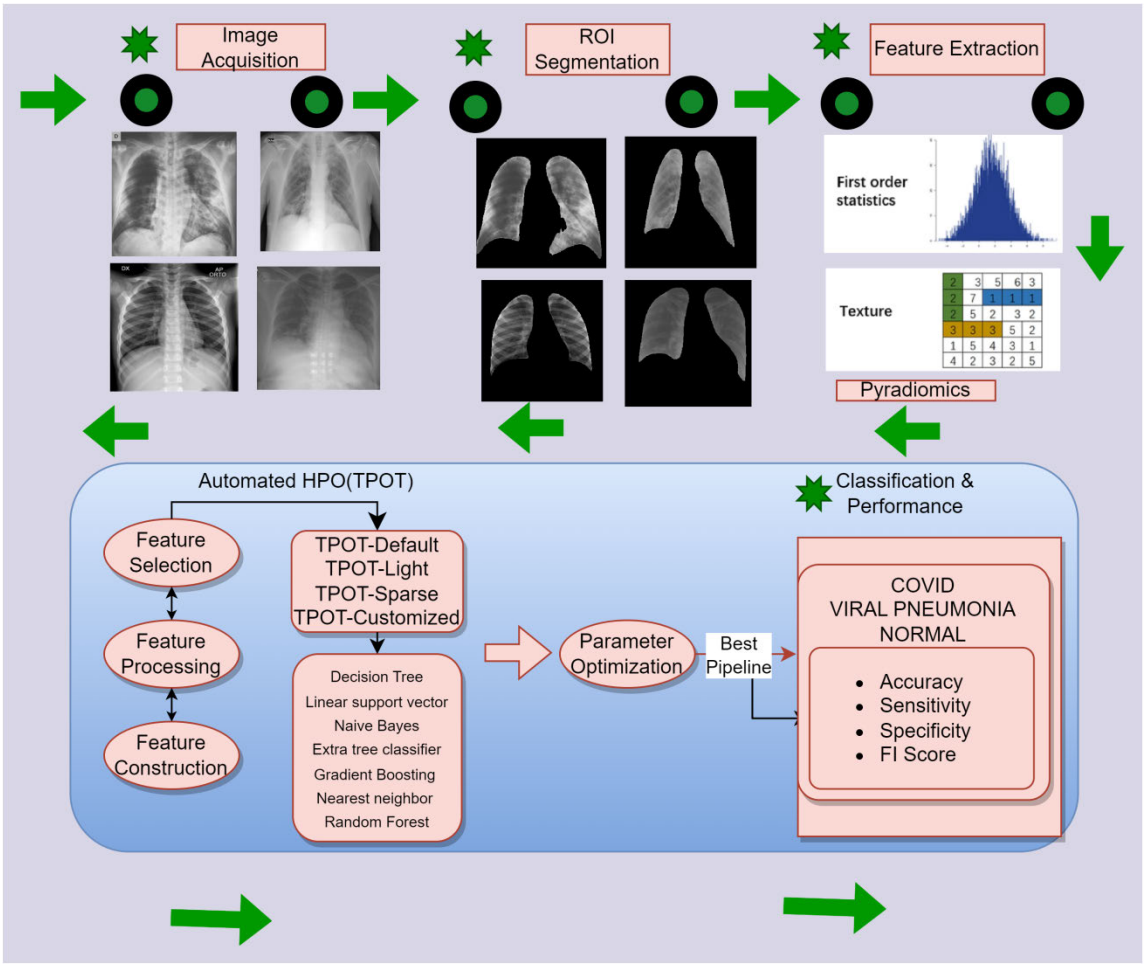

FIGURE 1. Proposed model.

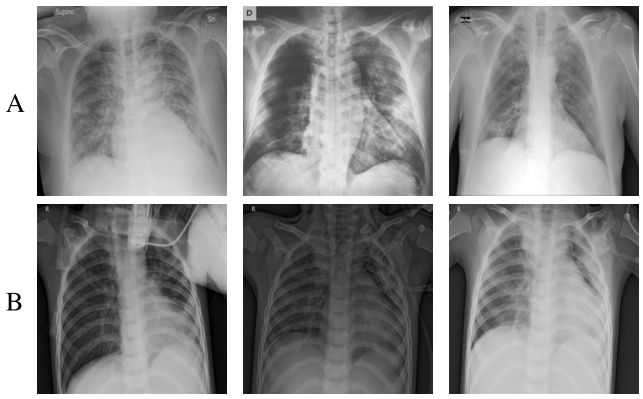

FIGURE 2. Sample images from the data set (A) COVID images (B) viral pneumonia images.

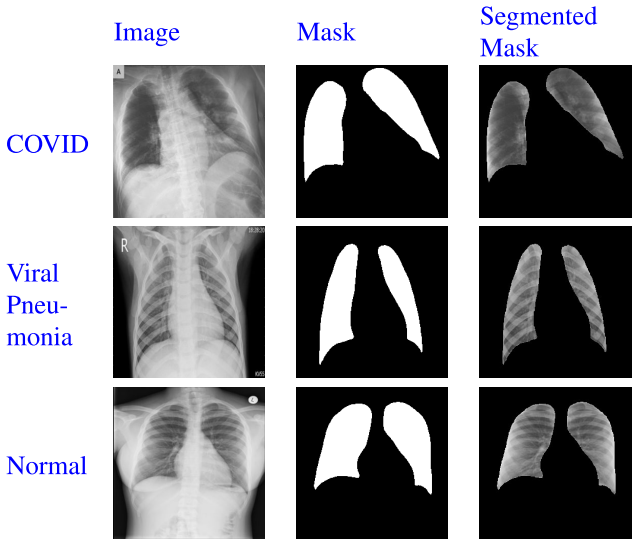

FIGURE 3. Chest X-rays, Lung mask and segmented mask for COVID-19, viral pneumonia, and healthy lung image.

textures, shapes, and other discernible patterns within the images. The overarching goal of radiomics is to transform medical images into mineable data, facilitating a more exhaustive and objective analysis for diagnostic, prognostic, or predictive purposes. In the extraction of these features, the Pyradiomics library proved instrumental shown in figure 4. This sophisticated tool offers a broad spectrum of radiomics features, providing a means to extract quantitative details

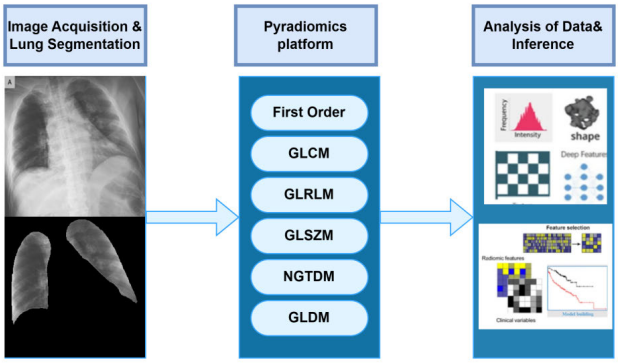

FIGURE 4. Pyradiomics feature extraction.

TABLE 1. PyRadiomics features.

Class	Description
First-Order	Retrieves the voxel-intensities with in the ROI in an image
GLCM	Quantifies pixel intensity patterns in an image for texture analysis.
GLRLM	Captures the distribution of consecutive pixel values in an image, offering insights into texture and structure.
GLSZM	Characterizes the size and intensity of connected regions with the same pixel value in an image, providing information about texture and spatial patterns.
NGTDM	quantifies variations in pixel intensity between a pixel and its neighbors in an image, offering information about texture and local contrast.
GLDM	characterizes the spatial dependence between pixel pairs with specific intensity values in an image, providing insights into texture and structure.

from medical images through a versatile set of radiomics functions.We have utilized the parameters binWidth and resampledPixelSpacing in the feature extraction settings and conducted the testing accordingly. These functions are broadly categorized into three main groups:

First-order features: First-order features, also known as statistical features, are a category of radiomic features that describe the basic statistical properties of pixel intensities in a medical image. These features provide a fundamental understanding of the distribution of pixel intensities, offering insights into the overall characteristics of the image.This feature describes how voxel intensities vary within a ROI. There are mean, median, standard deviation, skewness, kurtosis, minimum, and maximum intensities. It is essential to note that these features gives basic knowledge about the intensity distributions within the ROI and may be useful in analyzing the overall distribution of intensity within the ROI [34].

Texture features: A texture feature, also known as a texture analysis feature or a second-order feature, is a class of radiomic features that detect patterns as well as spatial arrangements of pixel intensities in medical images. This type of information provides information about the texture and fine details that can be seen in an image that may be difficult or impossible to see with the naked eye. Table 1 enumerates all the texture features. A texture analysis is particularly valuable in medical imaging since it is able to reveal subtle variations in tissue properties, with the purpose of better identifying different pathologies based on these subtle changes. These features are able to describe a number of spatial patterns and relationships between the voxels intensities within the ROI that are able to be described as a consequence of the structural features [35]. In radiomics, the following are some of the most common textural features that are used:

Gray Level Co-occurrence Matrix (GLCM): GLCM, or Gray-Level Co-occurrence Matrix, is a statistical technique employed in image processing and computer vision to analyze texture patterns within images. This method focuses on representing the distribution of co-occurring pixel intensity values at a specified offset or distance in an image. Figure 5 illustrates the calculation process of GLCM.Essentially, the GLCM provides valuable insights into the spatial relationships between pixel values, capturing patterns that contribute to the overall texture of the image. The elements of the GLCM are computed based on the occurrences of pixel pairs with specific intensity values at a given distance and angle. The general formula for computing the GLCM element P $(i, j; d, \theta)$ is shown in equation 1.

$$P(i,j;d,\theta) = \frac{(Number\ of\ Occurences\ of pixel\ pair\ (i,j))}{(Total\ number\ of\ pixel\ pairs\ at\ distance)} \tag{1}$$

Gray Level Run Length Matrix (GLRLM): GLRLM stands for Gray-Level Run Length Matrix, a texture analysis technique utilized in image processing, similar to GLCM (Gray-Level Co-occurrence Matrix). It assesses the distribution of pixel runs with a defined gray level and length in an image. The GLRLM is a matrix that describes how runs of consecutive pixels with the same gray level are distributed in a particular direction within an image, which can be horizontal, vertical, or diagonal. The matrix P(i, j) indicates the count of runs with gray level i and length j. The GLRLM features are shown in equations 2 - 11.

$$ShortRun = \sum_i \sum_j (Q(i,j)/j^2)/S \tag{2}$$

$$LongRun = \sum_i \sum_j O^2 Q(i,j)/S \tag{3}$$

$$Graylevelnon - uniformity = \sum_i (\sum_j, Q(i,,,J))^2/S \tag{4}$$

$$Runlengthnon - uniformity = \sum_j (\sum_i, Q(i,,,J))^2/S \tag{5}$$

$$Runratio = \sum_i \sum_j S/jQ(i,,,j) \tag{6}$$

$$LowGraylevelRun = \sum_i \sum_j Q(i,j)/Si^2 \tag{7}$$

$$HighGraylevelRun = \sum_i \sum_j i^2 Q(i,j)/S \tag{8}$$

$$ShortRunlowgraylevel = \sum_i \sum_j Q(i,j)/j^2 i^2 \tag{9}$$

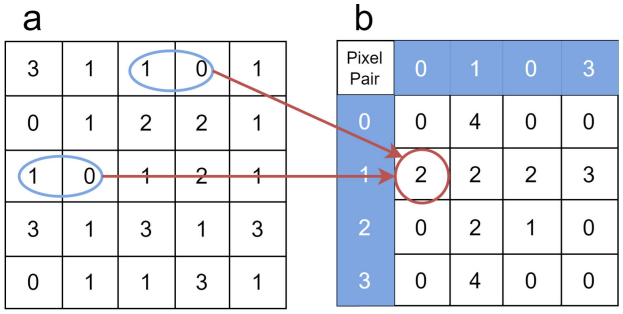

FIGURE 5. GLCM matrix.

$$ShortRunHighgraylevel = \sum_i \sum_j i^2 Q(i,j)/Sj^2 \quad (10)$$

$$LongRunlowgraylevel = \sum_i \sum_j j^2 Q(i,j)/Si^2 \quad (11)$$

Gray Level Size Zone Matrix (GLSZM): GLSZM or Gray-Level Size Zone Matrix is an acronym which stands for Grey-Level Size Zone Matrix. It is a texture analysis technique which is used in medical imaging and other fields to quantify the distribution of connected regions (zones) within an image based on their gray-level values at different points in the image [36]. The GLSZM is a method of analyzing the spatial distribution of regions of different sizes and gray levels within an image, thereby offering insight into the texture and heterogeneity of the underlying structures by analyzing the spatial distribution of these regions.The GLSZM equations are displayed in 12-16.

$$ProbabilityofOccurrence = \frac{\text{GLSZM}(i,j)}{N} \quad (12)$$

$$GrayLevelNon - Uniformity = \sum_{i=1}^{N_g} (P(i,\cdot))^2 \quad (13)$$

$$ZoneSizeNon - Uniformity = \sum_{j=1}^{N_s} (P(\cdot,j))^2 \quad (14)$$

$$GrayLevel\text{Non-Uniformity}Normalized = \frac{GLN}{N^2} \quad (15)$$

$$ZoneSize\text{Non-Uniformity}Normalized = \frac{ZSN}{N^2} \quad (16)$$

Gray Level Dependence Matrix (GLDM): The GLDM features describe how the spatial distribution of voxel intensity values is dependent on the intensity values of voxels. In medical imaging, another texture analysis method commonly used in a quantitative analysis of pixel intensities is the Gray Level Dependence Matrix, or GLDM, which quantifies the spatial dependencies of pixel intensity values within an image. As a result, it is capable of characterizing the relationships between pixels based on their gray-level values. As a result of analyzing the distribution of gray level pairs within the image at a given distance and angle, the GLDM matrix can be calculated. These features can be seen in equations 17 to 20 which shows the characteristics of the system.

$$Contrast = \sum_i P(ij).(i-j)^2 \quad (17)$$

$$Dissimilarity = \sum_{i,j} P(i,j).|i-j| \quad (18)$$

$$Homogeneity = \sum_{i,j} \frac{P(i,j)}{1+|i-j|} \quad (19)$$

$$Energy = \sum_{i,j} P(i,j)^2 + \epsilon \quad (20)$$

Neighborhood Gray-Tone Difference Matrix (NGTDM): The Neighborhood Gray-Tone Difference Matrix, or NGTDM for short, is another procedure for analyzing texture that is commonly used in medical imaging to estimate the differences in pixel intensities within a certain neighborhood. NGTDM aims to capture the spatial distribution of gray levels in an image, just like other techniques for texture analysis. There are several applications for this tool, but it can be particularly useful when describing texture homogeneity and coarseness. This is a technique that measures the difference between the intensity values of adjacent pixels in a matrix and those of a particular gray level intensity value based on the spacing between the pixels. Several of these features can be used to provide more detailed information regarding the texture and spatial patterns of the tissues within the ROI, as well as to help identify subtle changes in the characteristics of certain tissues within the ROI on a more detailed level.NGTDM equations are shown in 21-23. First-order analysis, GLCM, GLRLM, GLSZM, GLDM, and NGTDM were used to extract 19 features, 24, 16, 16, 14, and 5 features using GLCM, GLRLM, GLSZM, GLDM, and NGTDM, respectively, in this study.

Figure 6. highlights the importance of radiomic features in descending order.Fig.7 provides a visual representation of the correlations between the top features in the data set, aiding in exploratory data analysis and feature selection, where darker colors represent stronger correlations.The number of radiomic features used for this study was 94.

Shape Features: images ROI that can be described by these features. It consists of volumes, surface areas, sphericity, compactness, and elongation of the objects. In addition to giving information about the shape and size of the ROI, these features can help to Several geometrical properties describeidentify the morphology of tumors or other anatomical structures in order to characterize them.

$$Busyness = \frac{1}{N_g - 1} \sum_{i=1}^{N_g} i \cdot p_i \quad (21)$$

$$Complexity = \frac{1}{N_g} \sum_{i=1}^{N_g} \left(\frac{p_i}{1+(i-\mu)^2} \right) \quad (22)$$

$$Strength = \sum_{i=1}^{N_g} p_i^2 \quad (23)$$

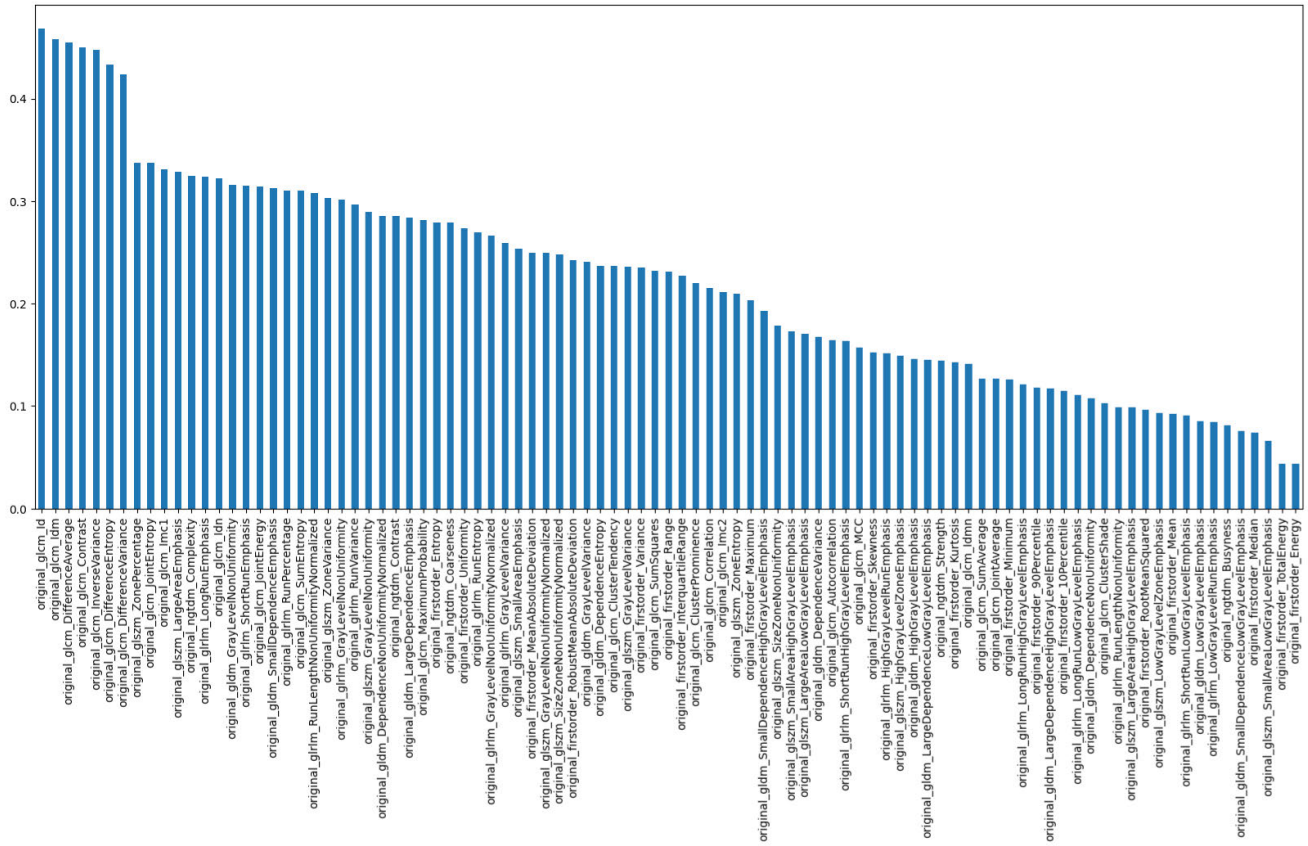

FIGURE 6. Bar plot showing the feature importance of radiomic features in descending order.

D. GA BASED HYPER-PARAMETERS TUNING

Genetic algorithms (GAs) represent optimization techniques inspired by the principles of natural selection, offering a robust approach to finding optimal solutions in scenarios where problem complexities abound, and variable relationships are not clearly defined within the vast search space. In the realm of machine learning, genetic algorithms prove particularly valuable for hyperparameter tuning, seeking to identify the most effective set of hyperparameters for a given model. In this context, the TPOT model, or Tree-based Pipeline Optimization Tool, comes into play. TPOT, an acronym for Tree-based Pipeline Optimization Tool, is an automated machine learning (AutoML) tool meticulously crafted to streamline the tasks of model selection, feature engineering, and hyperparameter tuning for data scientists and machine learning practitioners. TPOT can perform feature selection as part of its automated machine learning pipeline. Leveraging genetic programming, TPOT systematically explores diverse combinations of preprocessing steps and machine learning models. The overarching goal is to identify the optimal pipeline tailored to a specific dataset, thereby automating and enhancing the efficiency of the machine learning workflow. Moore et al [38] developed a framework based on Tree-Based Pipeline Optimization Tool (TPOT).

TPOT Configurations: The default version of TPOT serves as a well-balanced starting point, offering solid performance across a diverse range of tasks without requiring extensive manual tuning. It provides an efficient option for users seeking automated machine learning capabilities without the need for intricate customization, allowing for a quick and effective approach. While TPOT's default setting provides a strong out-of-the-box solution, users retain the flexibility to fine-tune its behavior. Adjusting various parameters and constraints enables users to tailor TPOT to specific requirements or preferences, accommodating a more personalized and targeted optimization process. TPOT-Light, a variant of the Tree-based Pipeline Optimization Tool, is designed for scenarios with limited computational resources or time constraints during model training. Operating on a reduced search space compared to the full TPOT, TPOT-Light focuses on lighter, quicker-to-train models. Despite this reduced complexity, TPOT-Light performs automated feature engineering, selection, and hyperparameter tuning to find an effective machine learning pipeline. Users can leverage TPOT-Light by providing a dataset and specifying the target variable. The tool autonomously explores various combinations of preprocessing steps, models, and hyperparameters to maximize performance based on a specified evaluation metric. TPOT-Light streamlines the machine

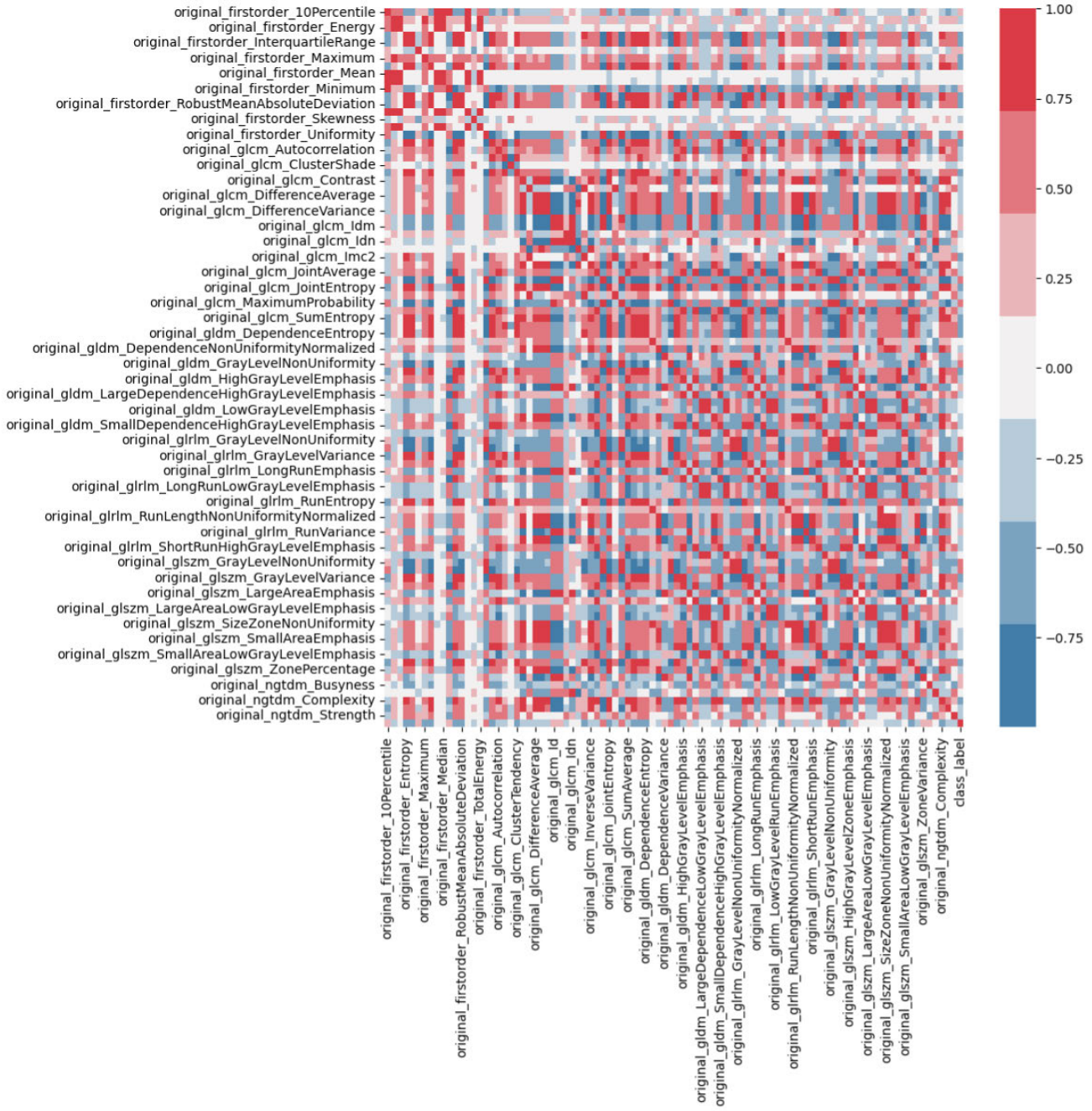

FIGURE 7. Heat map shows the effect of extracted radiomic features using the Pyradiomics package.

learning pipeline optimization process, particularly beneficial in resource-constrained environments. The ''Sparse'' option in TPOT specifically addresses datasets with a substantial number of features, commonly encountered in real-world applications. This feature enhances TPOT's adaptability to handle complex datasets, contributing to its effectiveness in diverse and challenging machine learning scenarios. Table 2 outlines the TPOT configuration types.

Customized TPOT: The process of customizing TPOT entails the procedure of tailoring the Tree-based Pipeline Optimization Tool to meet the specific preferences, requirements, or constraints of users or to fit a particular use case. There are several aspects of the optimization process that can be customized as part of this process. Depending on the preferences of the users, the search space can be defined or restricted in terms of the preprocessing techniques, feature selectors, and machine learning models that are considered. According to the priority of the problem, TPOT can use a variety of evaluation metrics to assess the pipeline performance, depending on the choice made by the user.

TABLE 2. TPOT configurations.

Name	Description
Default TPOT	TPOT conducts an extensive search across various preprocessors, feature selectors, feature constructors models, and parameters to identify a combination that minimizes the error in model predictions. Certain operators within this search space may be intricate and time-consuming, particularly when dealing with large datasets
TPOT Light	The tool aims to discover a combination of operators that reduces the model prediction error within a limited search space of preprocessors, feature selectors, feature constructors, models, and parameters. TPOT light is beneficial for swiftly identifying straightforward pipelines in classification and regression problems, as it focuses on simpler and faster operators, making it suitable for scenarios where simplicity is preferred.
TPOT Sparse	TPOT utilizes a configuration dictionary that incorporates a one-hot encoder, along with operators typically included in TPOT that also support sparse matrices.

Furthermore, the number of generations or the time allotted for optimizing can also be fine-tuned to reduce the amount of computational resources required. Several advanced features can be incorporated into the pipeline by advanced users, including custom preprocessing steps, feature engineering techniques, and even unique machine learning models. Through this customization, users will be able to better align TPOT with their specific needs, which will enhance its efficiency and versatility across a wide range of scenarios. Making sure that a balance is struck between customization and generalization will ensure that the tool will be effective and adaptable to a variety of situations.

To tailor the TPOT model to suit our study's needs, we implemented distinct modifications.The chosen settings were selected based on several considerations to optimize model performance and align with the specific requirements of the study. These adjustments ensure that the automated optimization process aligns effectively with our research objectives and dataset characteristics. These modifications directly contributed to the achievement by optimizing the model's relevance, feature representation, configuration, alignment with domain knowledge, and interpretability.

IV. RESULTS AND DISCUSSION

A. EVALUATION PARAMETERS

To assess the discrimination performance, several indicators were employed, including accuracy, sensitivity, specificity, positive predictive value, and negative predictive value (NPV). These metrics were crucial in evaluating the effectiveness of discrimination. Four performance indexes were calculated: false positive index (FP), true positive (TP), true negative index (TN), and false negative index (FN). A TP (True Positive) case signifies accurate identification of COVID-19, while an FP (False Positive) case indicates misidentification of COVID-19 as viral pneumonia. Conversely, an FN (False Negative) case occurs when COVID-19 is correctly identified as viral pneumonia, and a TN (True Negative) case denotes correct identification of viral pneumonia as COVID-19.

The evaluation metric for predictive performance, based on the ROC(receiver operating characteristic) curve, involved calculating the AUC (area under the curve) for each machine learning model. The AUC served as a comprehensive measure to assess predictive accuracy across different models. Python, along with the scikit-learn library, was employed for performance verification and statistical analysis, providing a robust platform for conducting the analysis efficiently and effectively.The formulas for performance measures are shown from 24 to 28.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{24}$$

$$Precision = \frac{TP}{TP + FP} \tag{25}$$

$$F_1 - score = 2 * \frac{Precision * sensitivity}{Precision + sensitivity} \tag{26}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{27}$$

$$Specificity = \frac{TN}{TN + FP} \tag{28}$$

B. EXPERIMENTAL RESULTS

The efficacy of the proposed technique is assessed utilizing a publicly available dataset known as the Covid-19 CXR dataset, which is categorized into three groups: Covid-19, viral pneumonia, and normal cases. This dataset is partitioned into 80% for training purposes and 20% for testing. Following the extraction of radiomics features, the genetic algorithm-based TPOT technique is employed for hyperparameter tuning and model optimization. In the pursuit of model development, a comparative analysis is conducted among different configurations of TPOT, namely TPOT-Default, TPOT-Light, and TPOT-Sparse. This comparative approach delves into the influence of hyperparameter variations on the overall performance of the model. Notably, this study goes a step further by introducing a customized TPOT model for optimal pipeline optimization, resulting in an impressive 94% AUC (Area Under the Curve). This achievement underscores the robustness and effectiveness of the proposed technique in leveraging radiomics features and machine learning for accurate classification within the given dataset. In the context of Class COVID, the model demonstrated accurate predictions for 74 instances (True Positives), but it misclassified 10 instances as Class viral pneumonia (False Positives) and overlooked 8 instances (False Negatives). Moving on to Class normal, the model accurately predicted 88 instances (True Positives) but made errors by misclassifying 7 instances as Class COVID (False Positives) and misclassifying 8 instances as Class normal (False Negatives). Furthermore, for Class viral pneumonia, the model achieved 88 correct predictions (True Positives), misclassified 12 instances as Class COVID (False Positives), and misclassified 5 instances as Class viral pneumonia (False Negatives). These results pertain specifically to the TPOT-Default configuration, as illustrated in Figure 8.a.

TPOT-Light is specifically designed for scenarios where computational resources are limited or where there are constraints on the time available for model training. It operates on a reduced search space compared to the full TPOT, focusing on lighter, quicker-to-train models. Confusion matrix for TPOT-light shown in figure.10 reflects 75 correct positive predictions, 81 correct negative predictions, along with 12 instances of incorrectly predicting positive and 10 instances of incorrectly predicting negative, providing insights into the model's classification accuracy.The confusion matrix, depicted in Fig. 8.c, summarizes the classification model's performance. It indicates 74 true positive, 87 true negative, 10 false positive, and 8 false negative predictions. The confusion matrix, displayed in Fig. 8.d, highlights the classification model's performance with 85 true positives, 94 true negatives, 2 false positives, and 5 false negatives.

The Receiver Operating Characteristic (ROC) curve serves as a graphical tool for assessing the performance of a classification model across various classification thresholds. Figure 9.a displays the ROC curve for TPOT-Default, illustrating the true positive rate (sensitivity or recall) against the false positive rate at different threshold values for all three classes. This graphical representation provides a comprehensive view of the model's discrimination ability and its performance across a range of classification scenarios.The ROC curve for TPOT Light is depicted in Fig. 9.b. The Receiver Operating Characteristic (ROC) curve presented in Figure 9.c for TPOT-Sparse is a visual representation commonly employed in classification to showcase the model's performance across various discrimination thresholds. This curve plots the True Positive Rate (Sensitivity) against the False Positive Rate (1 - Specificity) for different threshold values.

The Area Under the Curve (AUC), as depicted in Figure 9.d, provides valuable insights into the discriminatory performance of TPOT with customized values for different classes in its ROC curve. Specifically, for the COVID class, the 94% AUC indicates a robust ability to differentiate between COVID and non-COVID cases, with a 94% probability of correctly ranking a randomly chosen COVID case higher than a non-COVID case. Likewise, the 97% AUC values for both the Viral Pneumonia and Normal classes signify high discriminatory accuracy. In these instances, the model has a 97% probability of correctly ranking cases within each class higher than those outside the class.

Collectively, these AUC values suggest that the model effectively distinguishes between the specified classes, with higher AUC values indicating superior discriminatory performance. It is crucial to consider these metrics alongside other evaluation measures for a comprehensive assessment of the model's overall performance

Table 3 presents the performance measures for each TPOT model, detailing the default settings utilized by TPOT, specifically for the XGBClassifier–a implementation of the XGBoost algorithm. In this configuration, the learning rate

TABLE 3. TPOT performance measures.

Model	Accuracy	Precision	Sensitivity	Specificity	F_1-score
TPOT-Default	0.8333	0.7957	0.8043	0.8744	0.7989
TPOT-Light	0.8207	0.7353	0.8152	0.8232	0.7739
TPOT-Sparse	0.83	0.796	0.804	0.875	0.80
TPOT-Customized	**0.91**	**0.922**	**0.913**	**0.904**	**0.998**

is set at 1.0, influencing the step size in each boosting round. The maximum depth of the trees is limited to 9, with a minimum child weight of 10 implemented to control overfitting. The model incorporates 100 boosting rounds, employs a single thread for training ($n_jobs = 1$), and utilizes a subsample fraction of 0.95 to prevent overfitting. Furthermore, the verbosity level is set to 0, minimizing the printed output during the training process. These specified parameter values establish a baseline configuration for the XGBClassifier within TPOT, serving as the initial point for automated hyperparameter optimization. This optimization aims to enhance the model's performance on the provided dataset by systematically tuning these parameters based on their impact on the model's accuracy and generalization capabilities.

TPOT-Light defines the pipeline used for the KNeighborsClassifier algorithm used for classification purposes in the TPOT-Light configuration. We initially scale the input matrix using the MaxAbsScaler, which ensures uniformity across the features by scaling each feature with its maximum absolute value, which ensures uniformity across the feature sets. A KNeighborsClassifier is then employed utilizing the following hyperparameters: it takes into account nine neighbors in making predictions, the Manhattan distance (L1 norm) is the distance metric used (p=1), and the weights are assigned based on the distance, with the closer neighbors having a bigger impact ('distance weighting'). This configuration is customized for TPOT-Light, a tool that enables you to explore different pipelines and hyperparameter settings iteratively, with the objective of optimizing the performance of the KNeighborsClassifier on a given dataset by iteratively experimenting with them. It is shown in Table 4 that the best hyperparameters for each TPOT model can be found. The TPOT-Sparse configuration explains how to set up the XGBClassifier within the Tree-based Pipeline Optimization Tool (TPOT) using the Tree-based Pipeline Optimization Tool (TPOT). The learning process involves the use of a dataset which may have sparse representations of the features in it. The XGBClassifier is configured with a learning rate of 0.1, which facilitates a balance between the number of boosting rounds and the convergence of the classifier. There is a maximum depth of nine for each tree to ensure the level of complexity that is required to capture patterns without exposing the system to the risk of overfitting. A minimum child weight of four is specified, as well as

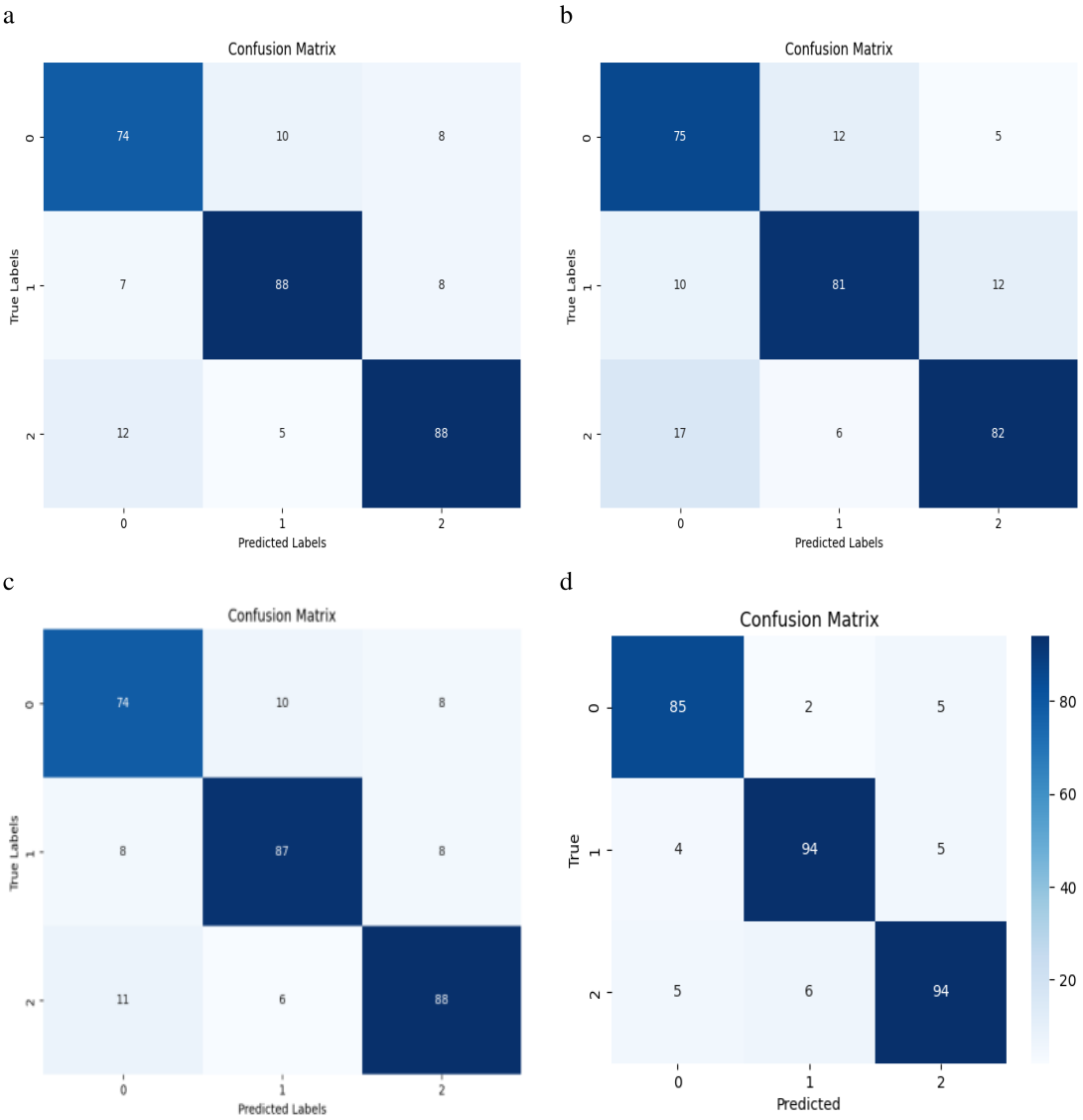

FIGURE 8. Confusion matrix. a) TPOT-Default b)TPOT-Light c)TPOT-sparse d)TPOT-customized.

100 rounds of boosting. During training, we reduce the verbosity by setting the subsample fraction to 0.5, introducing randomness to ensure that there is no overfitting, and we minimize the verbosity by setting it to zero. TPOT-Sparse configuration has been fine-tuned to cater to the needs of TPOT-Sparse. The aim is to automate the exploration of the optimal settings of hyperparameters for the XGBClassifier, especially on datasets that have sparse features.

Figure.10 presents a comparative analysis of different TPOT configurations, including TPOT-Default, TPOT-Light, TPOT-Sparse, and a Customized version, based on their computational efficiency and model accuracy on a specific dataset. TPOT-Default, despite consuming a considerable amount of time (746.2 seconds), achieved an accuracy of 83%. TPOT-Light, a more computationally efficient variant, took only 45.71 seconds but sacrificed a bit of accuracy, achieving 79%. TPOT-Sparse struck a balance, taking 161.34 seconds with an accuracy of 82%. Notably, the Customized configuration outperformed others, taking

TABLE 4. TPOT configurations.

Model	Best Hyperparameters
TPOT-Default	XGBClassifier(input_matrix,learning_rate=1.0, max_depth=9,min_child_weight=10, n_estimators=100,n_jobs=1,subsample =0.9500000000000001, verbosity=0)
TPOT-Light	KNeighborsClassifier(MaxAbsScaler(input_matrix), n_neighbors=9, p=1, weights=distance)
TPOT-Sparse	XGBClassifier(input_matrix,learning_rate=0.1, max_depth=9,min_child_weight=4, n_estimators=100,n_jobs=1,subsample=0.5, verbosity=0)
TPOT-Customized	RandomForestClassifier(bootstrap=True, criterion='gini',max_features=0.55, min_samples_leaf=10,min_samples_split=19, n_estimators=100)

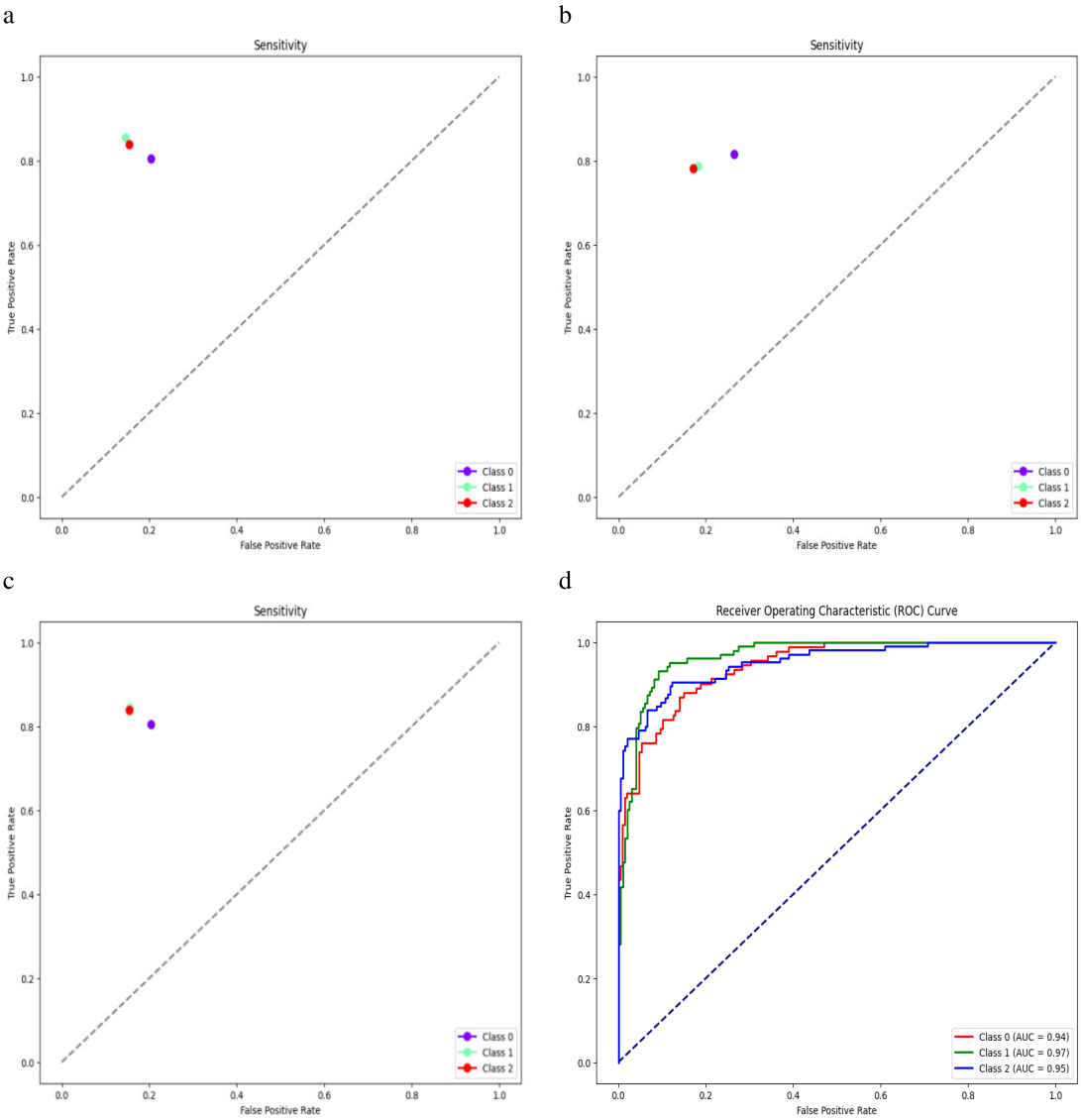

FIGURE 9. ROC AUC curve a) TPOT-Default b) TPOT-Light c) TPOT-sparse d) TPOT-customized.

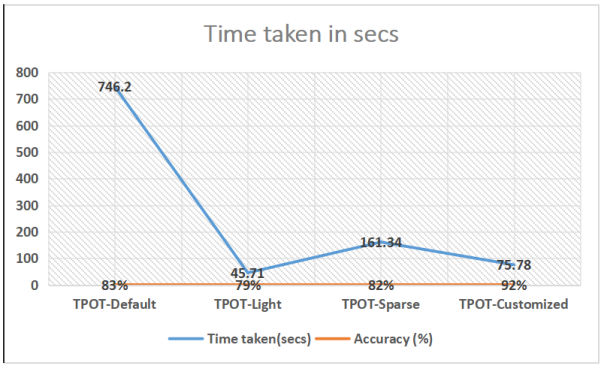

FIGURE 10. Time taken for each TPOT model.

75.78 seconds and achieving a remarkable accuracy of 92%. This indicates that the customization efforts in the TPOT pipeline led to a highly efficient and accurate model, tailored specifically to the characteristics of the dataset. The results emphasize the importance of customization in optimizing the trade-off between computational resources and model performance, highlighting the potential for tailored approaches to significantly enhance predictive capabilities. Real-time computation efficiency checking for the proposed model involves optimizing algorithm selection, leveraging hardware acceleration, and streamlining data preprocessing. By carefully balancing computational complexity, utilizing hardware acceleration techniques, and implementing efficient data pipelines, the model aims to achieve rapid and accurate diagnoses of viral pneumonia and COVID-19 from medical imaging data in real-time scenarios.

V. CONCLUSION AND FUTURE DIRECTION

The results of this study demonstrate that radiomics is rapidly becoming an important part of medical imaging

and that it can identify respiratory conditions, such as pneumonia and COVID-19, with a high degree of accuracy. This paper specifically focuses on the use of radiomic features together with machine learning in order to enhance diagnostic capabilities. There is a systematic approach to fine-tuning the machine learning model by the use of genetic algorithms when optimizing hyperparameters, in particular in the context of TPOT (Tree-based Pipeline Optimization Tool), which reflects a systematic approach to fine-tuning the machine learning model. This study makes an effort to explore the impact of hyperparameter variations on model performance by comparing different settings of the TPOT, specifically TPOT-Default, TPOT-Light, and TPOT-Sparse. This method of comparing hyperparameters adds depth to previous studies. To enhance the efficiency of the Random Forest algorithm when used with a fresh batch of data, custom modifications tailored to the specific needs of the model demonstrate a proactive approach to optimizing performance, resulting in a remarkable achievement of 94% AUC (Area Under the Curve) when utilizing the Random Forest algorithm. In medical settings where time is of the essence, focus is crucial on executing procedures in a timely manner, especially in the context of timely diagnosis. A systematic analysis of the execution times for each of the proposed TPOT models provides insights into the efficiency of the proposed method in terms of computation. We conclude with the statement that this paper contributes to the evolving field of radiomics through its presentation of the potential of combing intricate radiomic features with machine learning for the accurate identification of pneumonia and COVID-19 cases in chest X-rays and thus contributes to the field of radiomics. As a result of the use of genetic algorithms for hyperparameter optimization, together with custom modifications to the TPOT model, we have successfully developed a model that performs extremely well with an impressive AUC of 94%. An evaluation of the model's performance should include not only accuracy, but also metrics such as precision, sensitivity, and F1-score which demonstrate that a comprehensive evaluation of the model has been undertaken. As a result, the proposed methods will have practical implications, providing radiologists with the potential to diagnose abnormalities more precisely and efficiently by applying these methods.This paper contributes valuable insights into optimizing machine learning models for the analysis of medical images as the field of radiomics continues to advance, and the methodology presented here contributes valuable insights to the advancement of radiomics. It will be important to conduct further research and validate the proposed approach with a variety of data sets in order to establish its robustness and generalizability in real-world clinical settings in order to demonstrate its robustness and generalizability. The limitation of our paper is that we did not include a dedicated test dataset. However, we intend to address this limitation in future research endeavors and also plan to enhance accuracy by incorporating features like wavelets.

ACKNOWLEDGMENT

The authors would like to thank Vellore Institute of Technology, Chennai, India, for their support.

REFERENCES

[1] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, and W. Tan, "Detection of SARS-CoV-2 in different types of clinical specimens," *JAMA*, vol. 323, no. 18, pp. 1843–1844, Mar. 2020.

[2] C. I. Paules, H. D. Marston, and A. S. Fauci, "Coronavirus infections–more than just the common cold," *Jama*, vol. 323, no. 8, pp. 707–708, 2020.

[3] O. Ruuskanen, E. Lahti, L. C. Jennings, and D. R. Murdoch, "Viral pneumonia," *Lancet*, vol. 377, no. 9773, pp. 1264–1275, 2011.

[4] World Health Org. (2021). *Who Coronavirus (COVID-19) Dashboard*. Accessed: May 22, 2022. [Online]. Available: https://COVID19.who.int/table

[5] World Health Org. (2021). *Global Tuberculosis Report*. Accessed: May 22, 2022. [Online]. Available: https://www.who.int/teams/global-tuberculosis-programme/

[6] S. I. Kamel, D. C. Levin, L. Parker, and V. M. Rao, "Utilization trends in noncardiac thoracic imaging, 2002-2014," *J. Amer. College Radiol.*, vol. 14, no. 3, pp. 337–342, Mar. 2017.

[7] M. Loey, S. El-Sappagh, and S. Mirjalili, "Bayesian-based optimized deep learning model to detect COVID-19 patients using chest X-ray image data," *Comput. Biol. Med.*, vol. 142, Mar. 2022, Art. no. 105213.

[8] A. Haghanifar, M. M. Majdabadi, Y. Choi, S. Deivalakshmi, and S. Ko, "COVID-CXNet: Detecting COVID-19 in frontal chest X-ray images using deep learning," *Multimedia Tools Appl.*, vol. 81, no. 21, pp. 30615–30645, Sep. 2022.

[9] C. Ieracitano, N. Mammone, M. Versaci, G. Varone, A.-R. Ali, A. Armentano, G. Calabrese, A. Ferrarelli, L. Turano, C. Tebala, Z. Hussain, Z. Sheikh, A. Sheikh, G. Sceni, A. Hussain, and F. C. Morabito, "A fuzzy-enhanced deep learning approach for early detection of COVID-19 pneumonia from portable chest X-ray images," *Neurocomputing*, vol. 481, pp. 202–215, Apr. 2022.

[10] T. Agrawal and P. Choudhary, "FocusCovid: Automated COVID-19 detection using deep learning with chest X-ray images," *Evolving Systems*, vol. 13, no. 4, pp. 519–533, 2022.

[11] J. L. Gayathri, B. Abraham, M. S. Sujarani, and M. S. Nair, "A computer-aided diagnosis system for the classification of COVID-19 and non-COVID-19 pneumonia on chest X-ray images by integrating CNN with sparse autoencoder and feed forward neural network," *Comput. Biol. Med.*, vol. 141, Feb. 2022, Art. no. 105134.

[12] S. H. Kassania, P. H. Kassanib, M. J. Wesolowskic, K. A. Schneidera, and R. Detersa, "Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: A machine learning based approach," *Biocybernetics Biomed. Eng.*, vol. 41, no. 3, pp. 867–879, Jul. 2021.

[13] E. Angeli, S. Dalto, S. Marchese, L. Setti, M. Bonacina, F. Galli, E. Rulli, V. Torri, C. Monti, R. Meroni, G. D. Beretta, M. Castoldi, and E. Bombardieri, "Prognostic value of CT integrated with clinical and laboratory data during the first peak of the COVID-19 pandemic in northern Italy: A nomogram to predict unfavorable outcome," *Eur. J. Radiol.*, vol. 137, Apr. 2021, Art. no. 109612.

[14] I. Shiri et al., "COVID-19 prognostic modeling using CT radiomic features and machine learning algorithms: Analysis of a multi-institutional dataset of 14,339 patients," *Computers in biology medicine*, vol. 145, Sep. 2022, Art. no. 105467.

[15] I. Shiri, H. Arabi, Y. Salimi, A. Sanaat, A. Akhavanallaf, G. Hajianfar, D. Askari, S. Moradi, Z. Mansouri, M. Pakbin, S. Sandoughdaran, H. Abdollahi, A. R. Radmard, K. Rezaei-Kalantari, M. Ghelich Oghli, and H. Zaidi, "COLI-Net: Deep learning-assisted fully automated COVID-19 lung and infection pneumonia lesion detection and segmentation from chest computed tomography images," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 1, pp. 12–25, Jan. 2022.

[16] D. Wang, C. Huang, S. Bao, T. Fan, Z. Sun, Y. Wang, H. Jiang, and S. Wang, "Study on the prognosis predictive model of COVID-19 patients based on CT radiomics," *Sci. Rep.*, vol. 11, no. 1, p. 11591, Jun. 2021.

[17] Z. Xu, L. Zhao, G. Yang, Y. Ren, J. Wu, Y. Xia, X. Yang, M. Cao, G. Zhang, T. Peng, J. Zhao, H. Yang, J. Hu, and J. Du, "Severity assessment of COVID-19 using a CT-based radiomics model," *Stem Cells Int.*, vol. 2021, pp. 1–10, Sep. 2021.

[18] H. Shi, Z. Xu, G. Cheng, H. Ji, L. He, J. Zhu, H. Hu, Z. Xie, W. Ao, and J. Wang, "CT-based radiomic nomogram for predicting the severity of patients with COVID-19," *Eur. J. Med. Res.*, vol. 27, no. 1, p. 13, Dec. 2022.
[19] F. Prinzi, C. Militello, N. Scichilone, S. Gaglio, and S. Vitabile, "Explainable machine-learning models for COVID-19 prognosis prediction using clinical, laboratory and radiomic features," *IEEE Access*, vol. 11, pp. 121492–121510, 2023.
[20] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Amsterdam, The Netherlands: Elsevier, 2011.
[21] C. A. Barbano, E. Tartaglione, C. Berzovini, M. Calandri, and M. Grangetto, "A two-step radiologist-like approach for COVID-19 computer-aided diagnosis from chest X-ray images," in *Proc. Int. Conf. Image Anal. Process.*, 2022, pp. 173–184.
[22] V. Guarrasi and P. Soda, "Multi-objective optimization determines when, which and how to fuse deep networks: An application to predict COVID-19 outcomes," *Comput. Biol. Med.*, vol. 154, Mar. 2023, Art. no. 106625.
[23] A. B. Godbin and S. G. Jasmine, "Screening of COVID-19 based on GLCM features from CT images using machine learning classifiers," *Social Netw. Comput. Sci.*, vol. 4, no. 2, p. 133, Dec. 2022.
[24] A. Borghesi and R. Maroldi, "COVID-19 outbreak in italy: Experimental chest X-ray scoring system for quantifying and monitoring disease progression," *La Radiologia Medica*, vol. 125, no. 5, pp. 509–513, May 2020.
[25] A. Signoroni, M. Savardi, S. Benini, N. Adami, R. Leonardi, P. Gibellini, F. Vaccher, M. Ravanelli, A. Borghesi, R. Maroldi, and D. Farina, "BS-Net: Learning COVID-19 pneumonia severity on a large chest X-ray dataset," *Med. Image Anal.*, vol. 71, Jul. 2021, Art. no. 102046.
[26] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Proc. Int. Workshop Deep Learn. Med. Image Anal.*, 2018, pp. 3–11.
[27] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 936–944.
[28] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?" in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 1135–1144.
[29] D. R. Loh, S. Y. Yeo, R. S. Tan, F. Gao, and A. S. Koh, "Explainable machine learning predictions to support personalized cardiology strategies," *Eur. Heart J. Digit. Health*, vol. 3, no. 1, pp. 49–55, Mar. 2022.
[30] N. Papanikolaou, C. Matos, and D. M. Koh, "How to develop a meaningful radiomic signature for clinical use in oncologic patients," *Cancer Imag.*, vol. 20, no. 1, p. 33, Dec. 2020.
[31] Z. Jiang, J. Yin, P. Han, N. Chen, Q. Kang, Y. Qiu, Y. Li, Q. Lao, M. Sun, D. Yang, S. Huang, J. Qiu, and K. Li, "Wavelet transformation can enhance computed tomography texture features: A multicenter radiomics study for grade assessment of COVID-19 pulmonary lesions," *Quant. Imag. Med. Surgery*, vol. 12, no. 10, pp. 4758–4770, Oct. 2022.
[32] F. Prinzi, C. Militello, V. Conti, and S. Vitabile, "Impact of wavelet kernels on predictive capability of radiomic features: A case study on COVID-19 chest X-ray images," *J. Imag.*, vol. 9, no. 2, p. 32, Jan. 2023.
[33] J. J. M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. H. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, and H. J. W. L. Aerts, "Computational radiomics system to decode the radiographic phenotype," *Cancer Res.*, vol. 77, no. 21, pp. e104–e107, Nov. 2017.
[34] N. Aggarwal and R. K. Agrawal, "First and second order statistics features for classification of magnetic resonance brain images," *J. Signal Inf. Process.*, vol. 3, no. 2, pp. 146–153, 2012.
[35] M. Avanzo, J. Stancanello, and I. El Naqa, "Beyond imaging: The promise of radiomics," *Phys. Medica*, vol. 38, pp. 122–139, Jun. 2017.
[36] L. J. Jensen, D. Kim, T. Elgeti, I. G. Steffen, B. Hamm, and S. N. Nagel, "Stability of radiomic features across different region of interest sizes—A CT and MR phantom study," *Tomography*, vol. 7, no. 2, pp. 238–252, Jun. 2021.
[37] *COVID-19 Radiography Database*. Accessed: Jul. 2, 2020. [Online]. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database
[38] J. H. Moore, P. H. Ribeiro, N. Matsumoto, and A. K. Saini, "Genetic programming as an innovation engine for automated machine learning: The tree-based pipeline optimization tool (TPOT)," in *Handbook of Evolutionary Machine Learning* (TPOT). Cham, Switzerland: Springer, 2024, pp. 439–455.

A. BEENA GODBIN received the B.Tech. degree in information technology and engineering from Anna University, Chennai, Tamil Nadu, in 2005, and the M.E. degree in computer science and engineering from Anna University, Coimbatore, Tamil Nadu, in 2019. She is currently pursuing the Ph.D. degree with Vellore Institute of Technology, Chennai. Her research interests include medical image processing, machine learning, and deep learning.

S. GRACELINE JASMINE received the master's degree in computer application and the Ph.D. degree in remotely sensed image processing from VIT. She is currently an Associate Professor with the School of Computer Science and Engineering, VIT, Chennai Campus. With 15 years of extensive experience in teaching and research, she has published around 50 papers in international journals and conferences. She is also the Research Group Chair of the Imaging and Computer Vision Research Group, VIT Chennai, and is engaged in diverse roles, including industrial consultancy and overseeing projects related to image processing, computer vision, and remote sensing. She is actively involved in numerous professional bodies and holds certifications, such as NASSCOM Certified Master Trainer of Associate Analytics and Remote Sensing and Digital Image Analyst certified by ISRO. Her recent projects include AI-based smart agriculture, computer vision for ST microelectronics, and 3D weather reconstruction for Indian Meteorological Department, showcasing her expertise in cutting-edge technologies and applications.