Proceedings of 12th IOE Graduate Conference
Peer Reviewed
ISSN: 2350-8914 (Online), 2350-8906 (Print)
Year: 2022 Month: October Volume: 12

Explanation-Driven Deep Learning Model for Diagnosis And Prediction of Lungs Diseases Using Chest X-ray Images

Sabita Dahal [a], Sashidhar Ram Joshi [b]

[a, b] *Department of Electronics & Computer Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal*
[a] 076mscsk011.sabita@pcampus.edu.np [b] srjoshi@ioe.edu.np

Abstract

Research on COVID-19 has been an important topic as it has negative impacts on economics of the globe. Although many researchers have built classification model with large accuracy, we have observed a strong gap for Explanation based classification and diagnosis model for the interpretation of the model's prediction. So,this study purposed an Explanation based computer aided diagnosis system (CADx) based on deep learning using an Convolutional Neural Network, model for four categorical classification of lungs diseases (COVID-19, non-COVID pneumonia, tuberculosis and healthy) based on chest x-ray images. We used SHAP and LIME as different Explainable AI algorithms for the interpretation of different types of lungs diseases predicted by model. This purposed model used clinician dataset of Nepal for the validation purpose. A total of 105 Chest x-ray images belonging to different classes are collected for validation which are further augmented to increase the amount.we have got an overall validation of 85% and Test set accuracy of 91%. The f1 score of 86.75% is obtained. To show how the model performs locally, LIME builds sparse linear models around each prediction. When interpreting the results, we employed SHAP to assure consistency and local accuracy because Shapely values look at all future predictions using all potential input combinations. The shapley value of the SHAP method helps to determine the category of the x-ray image, and the output of the LIME algorithm displays the region of the lungs used for categorization.

Keywords

Covid-19, CADx, CNN, Data augmentation, Deep learning, Explainable AI, Grad CAM, LIME, ROC curve, SHAP, Transfer Learning

1. Introduction

The importance of explanation and interpretation in the development of trust and integration into therapeutic practice cannot be overstated. Deep Learning models require both explanation and interpretability, along with the high level of accuracy. As a result, we proposed a Deep Learning model that uses a convolutions neural network (CNN) for the classification of 4 category of lungs infectious diseases Covid-19 pneumonia, Non-Covid Pneumonia ,tuberculosis and normal or healthy from an chest x-rays images. In the recent years, academics have become increasingly interested in the study of explainable AI. Due to the deep learning model's lack of interpretation and justification for predictions, it is known as a "black box." We used SHAP and LIME as different Explainable AI algorithms for the interpretations of different types of lungs diseases predicted by model. Shapely values investigate all feasible predictions using all possible combinations of inputs, therefore we employed SHAP to assure qualities like consistency and local correctness for interpretation [1]. LIME, on the other hand, builds sparse linear models over each prediction to represent how the model works in the immediate vicinity. This study focuses on interpretability and high accuracy, which are critical for identifying inequalities in predicting performance, building trust, and integrating into clinical practice [1].

In a developing country like Nepal, efficient medical resources and expert knowledge are lacking during the COVID epidemic, and many people have died as a result. The inability to discriminate between COVID and respiratory infection is one of the main causes. In many situations, even though the patient did not have COVID pneumonia, they were segregated owing to Resembling symptoms, and the therapy and medication were inadequate. Taking this incidence into view, we are motivated to conduct research for effective categorization of Covid-19 from other lethal respiratory track diseases from x-rays dataset. Despite the popularity and applicability of DL models in clinical decision making, the interpretability and lack of transparency of algorithm-driven conclusions remains the most significant obstacle.

2. Related Literature

According to the WHO, nearly 408 million people have been confirmed as having COVID-19, and approximately 5.1 million people have died as a result of the virus [?]. It is a global issue that must be addressed as soon as possible. Because COVID is a respiratory track illness, it shares many symptoms with other diseases caused by respiratory infections, making identification and management of the disease more complex. Many studies have been conducted in the field to differentiate covid pneumonia from non-covid pneumonia and other lungs diseases. If a large amount of annotated dataset is available, chest X-ray images can be effective in diagnosing covid-19, but due to the restricted availability of data, categorization and diagnosis are difficult [2].

Deep learning models outperform conventional machine learning algorithms in terms of accuracy, while deep learning-based CNN models outperform classical machine learning algorithms in the medical imaging domain [3]. By employing the backpropagation technique to suggest changes to a machine's internal parameters that are used to compute the model in each layer from the representation in the previous layer, deep learning can discover detailed structure in massive data sets [3]. Different performance metrics such as accuracy, responsiveness, specificity, precision, F1 value, and Deep Learning were introduced by the authors [4].

Wang et al. developed a deep learning-based model that detects COVID-19 using X-ray images and released the dataset of 13975 x-ray images as an open dataset [5]. In the interpretation of radiographic images, a number of artificial intelligence systems based on deep learning have been utilized, including CXR and CT, which have shown to be promising in terms of accuracy and sensitivity.The author discusses the value or benefits of employing chest x-ray radiographic pictures to distinguish Covid-19 from other non-Covid diseases during the global COVID-19 pandemic, particularly in resource-limited locations and areas that have been severely hit. Chest x-ray scans provide a quick way to diagnose problems, are widely accessible in hospitals and imaging facilities, are portable, and are reasonably priced.[5] Yeh, Chun-Fu et al. developed a COVID-19 pneumonia screening platform that is also AI-based and uses chest X-ray pictures to identify whether a patient has COVID-19 disease or not. It employs a cascade learning-based model to improve the sensitivity and specificity of a CNN-based model while employing a smaller dataset of publicly available X-rays [6].

Rajpurkar, Pranav, and colleagues created a 121-layer convolutional neural network trained on x-ray images to diagnose pneumonia from these pictures. This model uses x-ray pictures as input and outputs the likelihood of pneumonia as well as the most damaged body part [7]. For a number of reasons, the suggested architecture is a CNN-based DL architecture. First, we found that CNN-based deep learning is highly adept at lowering parameter thresholds while maintaining model quality. It can automatically extract features from an image, negating the need for human feature engineering.Chest X-rays (CXR) and computed tomography or computerized X-ray imaging (CT) scans are the two most used imaging tests for pneumonia and subsequently COVID-19. Since it produces incredibly detailed images, the CT scan is the gold standard for diagnosing lung diseases. Despite this, CXR is still very effective in some situations due to its affordability, speedy image generation, low radiation exposure to patients, and greater use in emergency rooms [8].

For a number of medical diagnostic tasks, deep learning techniques have proven to be quite successful, sometimes even outperforming human experts. The algorithms' black-box nature has, however, limited their therapeutic application. Recent studies on explainability seek to identify the factors most responsible for a model's choice [9]. A possible technique to improve the effectiveness and accessibility of the diagnosis process is through computer-aided diagnostics (CAD), which uses AI. The most effective artificial intelligence (AI) technique for a variety of challenges, including those involving medical imaging, is deep learning [9].

Interpretability, a vital aspect of model understanding, is one of the issues with model prediction. Knowing how to create a classification model that is accurate could provide us more assurance that the model is accurately capturing the patterns in the target area [10]. The implementation of an algorithm called gradient-weighted class activation mapping (Grad-CAM) to create a heat map to visually verify where in the image the CNN model is looking at, to ensure the model is functioning correctly, and to show the effectiveness and utility of the suggested method is one of the study's main contributions [10].

3. Methodology

Figure below depicts the proposed framework for deep learning model with Explainable AI. For the suggested method, we will have 3 different phases the first phase will be data pre-processing, second phase will be deep learning classification model and third phase will be explanation generating phase.

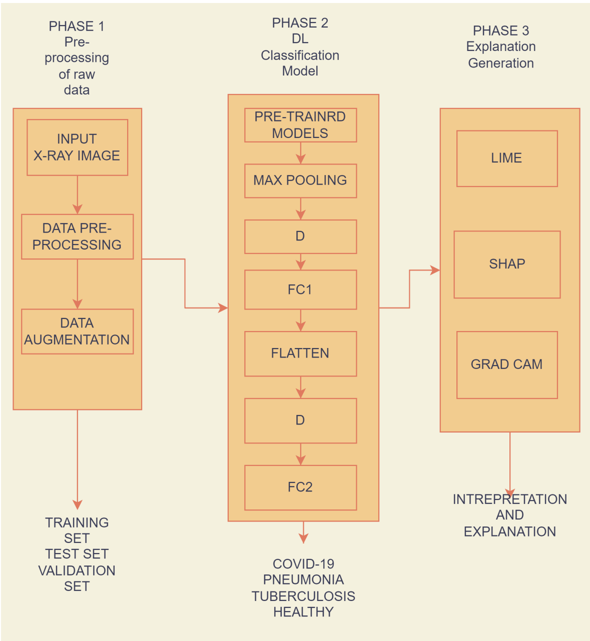

Figure 1: Framework of proposed Model. MAX is maximum pooling layer, FC fully connected layer and D is dropout layer, F is flattened layer.

3.1 Data Pre-processing

This study make use of two datasets, clinician dataset for validation and public dataset for the training and testing of the model. The dataset of four category of lungs diseases including Covid-19, Pneumonia, Tuberculosis and Normal chest x-ray images are collected. A total of 7135 dataset were utilized, of which 713 were used as testsets, 6317 were used for training, and 105 were used for model validation.

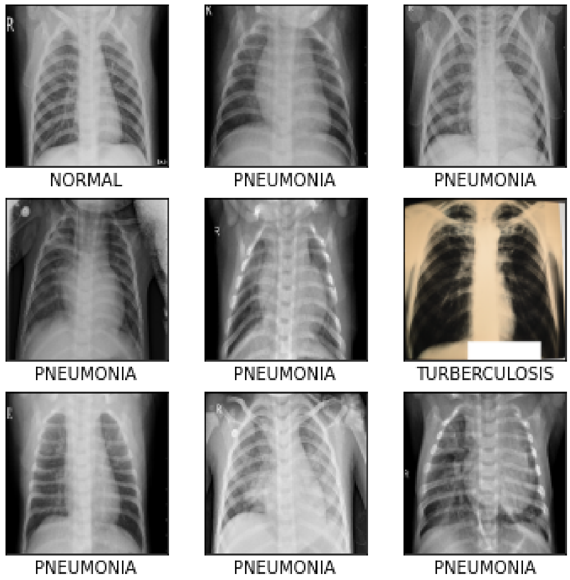

Figure 2: Sample CXR Image representatives

The images size has been resized to 180*180 pixel. Generally, the medical image has higher dimension and large resolution. The use of higher resolution image in Neural Network will result in large computational complexity hence the dataset are resized to smaller pixel size which became appropriate for the model. Normalization helps to keep data inside a range and lowers skewness, allowing you to learn more quickly and effectively. Chest x-ray pictures are divided by 255 for image normalization, with pixel values ranging from 0 to 1.Normalizing images entails changing them into values such that the image's mean and standard deviation are 0.0 and 1.0, respectively.

3.1.1 Data Augmentation

This technique is used as the part of image pre-processing. As the dataset available for COVID-19 related X-rays images are limited for the training of the model and deep learning model are resource constraint and required a large amount of data which may be expensive and have large computational complexity, so to achieve larger dataset from available small dataset augmentation is used. Rotation,Horizontal flip,Vertical flip,brightness and contrast enhancement are augmentation technique used here.

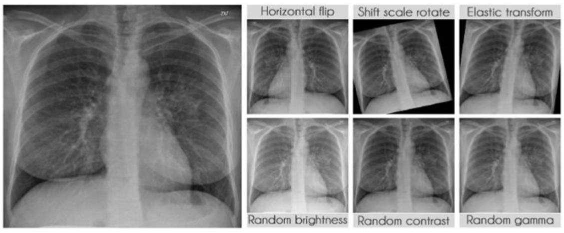

Figure 3: Data Augmentation

3.2 Classification Phase

The second phase of the framework is the classification model, where a sequential CNN model is used. The convolution layer will produce the features maps with the filter of size of 3*3 and stride of 1. The convolution layer of the CNN model is followed by a maximum pooling layer, which reduces the quantity of dataset by taking the Maximum of each feature map from the preceding layer. This prevents the model from overfitting and increases classification accuracy. GAP takes the average of all values to lower, It lowers the spatial dimension and reduce each feature map to a single integer. The features produced by global average pooling are given to the fully connected layer, which contains the weights and biases of each neuron. The features are then subjected to mathematical functions, for weight updating back propagation algorithm is used. resulting in image categorization.

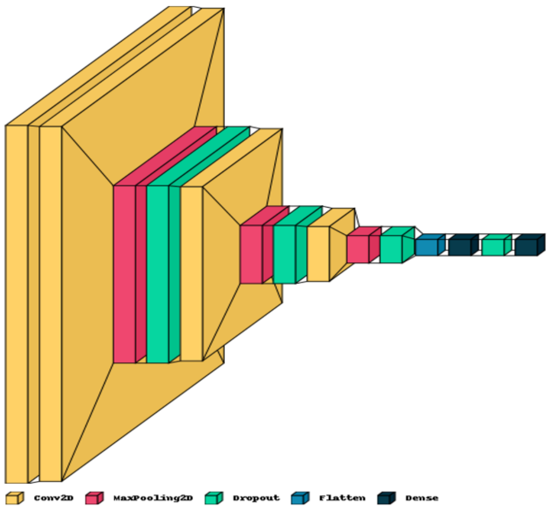

Figure 4: Convolutional Neural Network Model

The dropout layer, which prevents overfitting by dropping or nullifying the function of part of the features in the next layer. Dropout probability of 0.1 has been used here. Dropout layer is used for regulations and flatten layer is used to convert the 2D array output to single linear vector After the dropout layer, the final Fully connected layer is used for the four-category categorization. To find the best hyperparameter for the model and the data augmentation strategy, a random search was conducted. Hyperparameter optimization is a technique for improving performance of model in a reasonable amount of time,Because these parameters govern how the model learns

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 178, 178, 3)	84
conv2d_5 (Conv2D)	(None, 176, 176, 32)	896
max_pooling2d_3 (MaxPooling) 2D)	(None, 88, 88, 32)	0
dropout_4 (Dropout	(None, 88, 88, 32)	0
conv2d_6 (Conv2D)	(None, 86, 86, 96)	27744
max_pooling2d_4 (MaxPooling 2D)	(None, 28, 28, 96)	0
dropout_5 (Dropout	(None, 28, 28, 96)	0
conv2d_7 (Conv2D)	(None, 26, 26, 128)	110720
max_pooling2d_5 (MaxPooling 2D)	(None, 13, 13, 128)	0
dropout_6 (Dropout)	(None, 13, 13, 128)	0
flatten_1 (Flatten)	None, 21632)	0
dense_2 (Dense)	None, 512)	11076096
dropout_7 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 4)	2052

Total params: 11,217,592
Trainable params: 11,217,592
Non-trainable params: 0

Figure 5: CNN Model Summary

A patience level of 3 has been applied throughout the 50 iterations of this model. early stopping was set up to account for validation loss. The cross-entropy loss function and the Adam optimizer are used to control the learning process.

3.3 Explainable AI

Deep Convolutions neural networks are called as ”black boxes” since they are not automatically Interpretable. ”Transparency, interpretability, and Explainability are important to create patient and provider trust,” according to an international statement on the ethics of artificial intelligence in radiology [11].

Medical and health professionals are unlikely to place all their faith in an algorithm's predictions,By explaining why model predicts COVID-19 or other pulmonary infection disorders, we may improve the system and acquire a better understanding of why the model makes certain decisions.

3.3.1 LIME (Local Interpretable Model-Agnostic Explanation)

To interpret the predictions of the deep learning model that is used for prediction, we will employ Local Interpretable Model-Agnostic Explanations (LIME). LIME perturbs the features in an x-ray image and fits a linear model to approximate the neural network in the local feature space surrounding the image for explanation. The linear model is then used to discover which attributes contributed the most to the model's forecast for that image [11].

To assess trust—which is essential if one intends to act on a prediction—or to decide whether to use a new model, it is crucial to understand the motivations behind predictions. Insights into the model can be gained from such an understanding and used to change an unreliable model or prediction into a reliable one. a new method of explanation that builds a local interpretable model around the prediction to accurately and elucidate the predictions of any classifier [12].

LIME considers the image's features to be super pixels. These super pixels will be created using the Quick Shift mode-seeking segmentation technique. Then, by turning on and off the super pixels, many samples similar to the input image are created. When a super pixel is 1, it is on, and when it is zero, it is off. The prediction of each modified image is then generated in the next step. After that, the cosine distance between each perturbed image and the input image is calculated. The greater the similarity between a modified image and the input image, the greater its weight and contribution to prediction will be.

3.3.2 Time and Space Complexity measure of LIME

Let's write the model being explained as $f : R^d \rightarrow RF(x)$ is the likelihood (or a binary indicator) that x belongs to a particular class in classification [12]. In order to determine locality around x, we further use $pix(z)$ as a closeness measure between an instance z and x. Let $L(f,g,\pi_x)$ be a measure of how inaccurately g approximates f in the locale denoted by x. We need to keep omega(g) down enough to be interpretable by humans while minimizing $L(f,g,\pi_x)$ to ensure both interpretability and local faithfulness [12]. Complexity is measured by the constant

$$\Omega(g)$$

. The explanation produced by LIME is given as

$$\xi(x) = \operatorname*{argmin}_{g \in G} \mathscr{L}(f,g,\pi_x) + \Omega(g) \tag{1}$$

3.3.3 SHAP (SHapley Additive exPlanation)

SHAP is another explainable algorithm that we are going to implement in this study. For SHAP the score for each pixel on the predicted image is used as contribution for the classification of Pulmonary abnormalities. The SHapley value represents each feature for subtypes of lungs diseases. The algorithm used to calculate SHapley values is as shown.

Algorithm 1 Algorithm to calculate the Shapley Values

INPUT: Number of iterations M,
instance of interest x, feature index j, data matrix X,
and CNN Model

1: for Every Iteration 1......M do
2: Draw random instance Z from the data matrix X
3: Choose a random permutation of the feature values
4: Order instance X (X0...., Xj......, Xp)
5: Order instance Z (Z0......, Zi......, Zp)
6: Construct two new instances:
7: With j: X_{-j} = (X0...., Xj, Z_{j+1},.....Zp)
8: Without j: X_{+j} = (x0.... Xj, Zj+1.... Zp)
9: Compute the Marginal Distribution:

$$\phi_j^M = f(X+j) - f(X-j) \tag{1}$$

10: end for
11: Compute Shapley Values:

$$\phi_j(X) = \frac{1}{M}\sum_{M=1}^{M} \phi_j^M \tag{2}$$

Figure 6: Algorithm to calculate SHapley additive values

4. Result

Following the classification phase, the accuracy and quantity of incorrect predictions are used to assess how well CNN models performed. The curves for CNN's results are shown in the figure below.The model has been trained for 50 epoch, the patience level was set to 3 and the early stopping was set up to account for validation loss. The cross-entropy loss function and the Adam optimizer are used to control the learning process

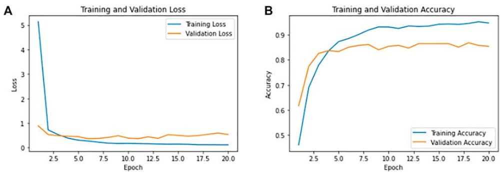

Figure 7: Training VS Validation Result of Classification model. A) Shows the training and validation loss with respect to number epochs. B) Shows the training vs the validation accuracy with respect to number of epochs

4.1 Classification Report

	precision	recall	f1-score	support
COVID19	0.84	0.73	0.78	63
NORMAL	0.90	0.91	0.91	165
PNEUMONIA	0.94	0.95	0.95	413
TURBERCULOSIS	0.81	0.86	0.83	72
accuracy			0.91	713

Figure 8: Classification Report

From the Classification report we can say that the overall accuracy in the test dataset is 91 percent. The overall f1 score of 0.867 and has been obtained for the study. The recall is 0.8625 and precision of 0.8725 i.e., 87 percentage is obtained for the classification of four category of the lungs diseases consisting of Covid-19, Pneumonia, Tuberculosis and Normal. The accuracy for the test dataset is obtained as 91%.

4.2 Confusion Matrix

A confusion matrix is a N*N model used to assess the performance of a model.This Research involves a four-category classification of covid-19, non-covid pneumonia, tuberculosis and healthy person chest x-rays images hence a 4*4 confusion matrix is employed here. The matrix compares the actual value to the model's predictions. It provides us with a comprehensive picture of how our model is functioning and what mistakes it is making[13].

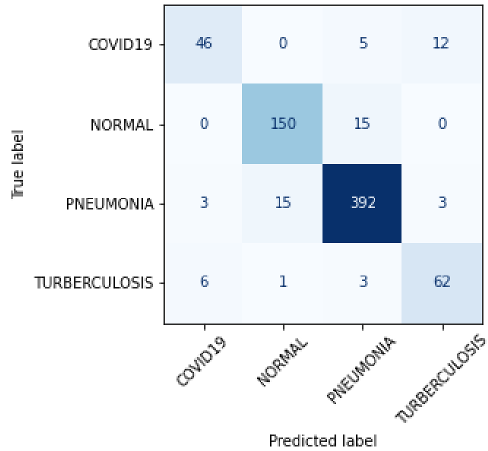

Figure 9: Confusion Matrix

The confusion matrix for the 713 test images is shown in the above figure. The model incorrectly classifies 17 Covid-19 images out of 63, 15 Normal images out of 165, 21 Pneumonia images out of 413, and 10 Tuberculosis images out of 72.

4.3 Interpretations of output of LIME

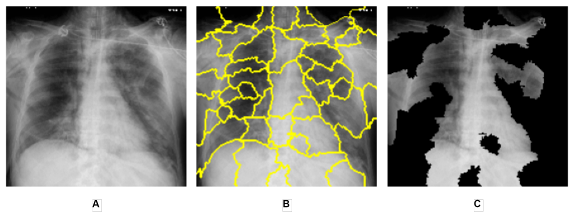

Figure 10: : LIME developed interpretations for a Covid-19 image. (A) An illustration of a Covid-19 images taken from a test image. (B) Super pixels produced via quick-shift segmentation of a sample of the Covid-19 image to produce perturbations. (C) The final perturbed image Covid-19 impact in CXR image.

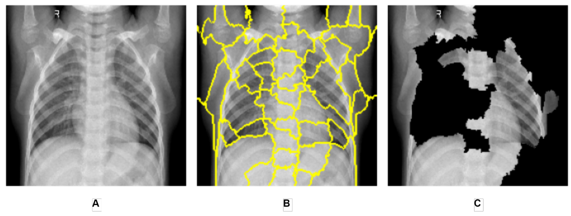

Figure 11: LIME developed interpretations for a typical image. (A) An illustration of a normal image taken from a test image. (B) Super pixels produced via quick-shift segmentation of a sample of the normal image to produce perturbations. (C) The final distorted image for no defect in CXR image.

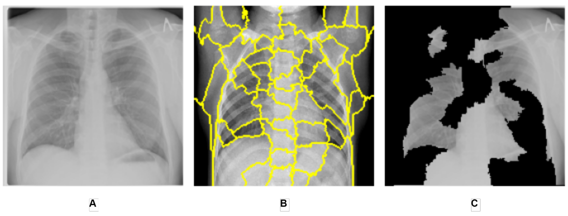

Figure 12: LIME developed interpretations for a typical image. (A) An illustration of a Tuberculosis image taken from a test image. (B) Super pixels produced via quick-shift segmentation of a sample of the normal image to produce perturbations. (C) The final distorted image for no defect in CXR image.

4.4 Interpretations of Output of SHAP algorithm

SHapley values generated after executing the above algorithm 1 represent the contribution of each feature to the model's prediction.Positive SHAP values indicate that the model has a positive impact on prediction, causing the model to predict the specific class. A negative SHAP score indicates a negative influence, which causes the model to miss to predict the class. In this research we used blue color pixel for the negative and red color for indicating the positive SHAP value.

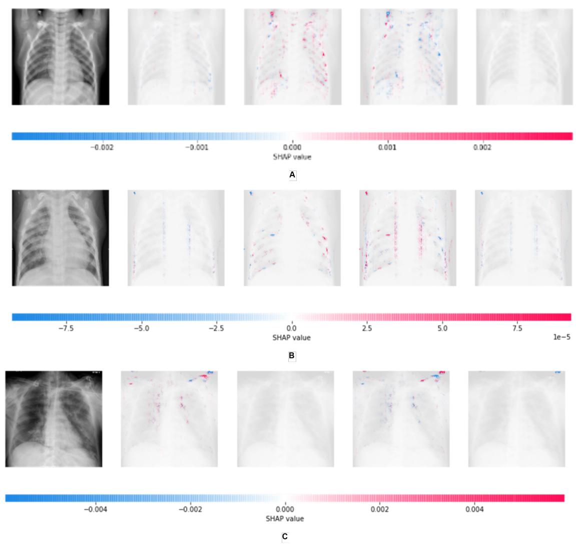

Figure 13: A) XAI SHAP implementation shows Normal image in terms of SHapley value B) From the SHapley value we can say that the image is of Pneumonia patient. C) on the basis of Shapley value, we can say that the CXR image holds Covid-19.

The first graphic displays a random image from the test dataset; the remaining figures display Covid-19, Normal, Pneumonia, and Tuberculosis, in that succession. The most red pixels in a figure indicate that it belongs to the specified category.

5. Comparison with Existing Methods

The classification accuracy for three classes—pneumonia, covid-19, and healthy—was found to be 83 percent by nishio2020automatic using 1248 datasets. DeTraC of [2] successfully identified COVID-19 X-ray images from patients of mild acute respiratory syndrome and severe acute respiratory disorder with an accuracy of 93.1% (and a sensitivity of 100%). The radiological characteristics of CT scans taken from patients with COVID-19 pneumonia and non-COVID-19 pneumonia were evaluated by Ardakani et al. (15). To identify the computer-aided diagnosis system that produced an accuracy of 91.94% utilizing an ensemble (COVIDiag) classifier, they used decision trees, K-nearest neighbor, naive Bayes, support vector machines, and ensemble classifiers. A deep CNN-based artificial intelligence technique has been developed by Alazab et al. to analyze chest X-ray pictures and identify COVID-19 patients. The prediction models' average accuracy was 94.80 and 88.43%. When compared to cutting-edge techniques, it is discovered that we have large test accuracy of 91% and validation accuracy of 85%. Additionally, we employed SHAP and LIME, two explainable AI algorithms, to explain the result or prediction generated by the model, which is a significant advancement in comparison to previous state-of-the-art methods.

6. Discussion

Using the CNN architecture and data augmentation techniques, we can develop a computer-aided diagnosis model, according to the research's findings, the algorithm successfully distinguishes between the four categories of Covid Pneumonia, viral/bacterial Pneumonia, Tuberculosis, and healthy from CXR images with a 91 percent accuracy rate.In a developing nation like Nepal, there are few effective medical resources and specialists, and during the COVID pandemic, many people did not receive the proper care and perished as a result. One of the primary causes is the inability to distinguish between pneumonia brought on by COVID and respiratory infection. Many times, despite the patient not having COVID pneumonia, they were isolated due to symptoms that were similar, and the treatment and medicines were ineffective. We are encouraged to perform research for an accurate categorization of these respiratory infection disorders in light of this incidence.

Since deep Convolutional neural networks cannot be automatically interpreted, they are referred described as ”black boxes.” An international statement on the ethics of artificial intelligence in radiology states that ”Transparency, interpretability, and Explainability are vital to foster patient and provider confidence”[11]. Health is a delicate area where any inaccuracies can end in a person’s death, therefore medical and health professionals are unlikely to put all their faith in an algorithm’s forecasts, nor should they. We may enhance the system and learn more about the model’s decision-making by describing how it predicts COVID-19 or other lung infection problems. For the interpretation and transparency of model predictions, we have employed two algorithms LIME and SHAP algorithms. Despite sharing the same primary objective they discover the critical locations in different ways.

The SHAP (SHapley Additive Explanation) value is a two-dimensional value made up of predicted class labels in rows and columns, with the value of each feature influencing the prediction of the image’s specific class. In this study, the absence of any contribution is represented by the color blue, and the high probability of class prediction is represented by the color red.

The predictions of the trained neural network classifier have been explained using local interpretable model-agnostic explanations (LIME). In order to approximate the neural network at the local region in the feature space around the example, LIME perturbs the features in the example and fits a linear model. Then, it makes use of the linear model to identify the characteristics that contributed most to the example’s prediction by the model [11].

7. Conclusion and Future Work

The suggested study used an explanation-driven CNN model to successfully classify photos, achieving an accuracy of 91 percent even for low-quality images. The CNN model was applied to the SHAP and LIME algorithms as well as the dataset of chest x-ray images used to train and evaluate the CNN model. With deep features as inputs from the dataset, the proposed CNN model is inherently interpretable and offers a reasonable explanation. These characteristics are used to describe the pixels on an image that contain characteristics that allow researchers to differentiate between the various categories of lung ailments that have been employed in the study. Even if the qualitative descriptions of the results are enough for the average person, image analysts and medical professionals can more fully comprehend the results.

Generative Adversarial Networks (GAN) can be utilized to produce more images by training the model using the available dataset for the study’s future advancement. The model may be tested and trained using more clinical datasets rather than just validation, which increases the model’s predictability and accuracy. On the other hand, explainability techniques like Grad-Cam can be used to better depict model predictions by highlighting the characteristics that do so. To avoid starting from scratch with a model and to make the most efficient use of resources, pre-trained CNN architectures can be employed.

References

[1] Loveleen Gaur, Mohan Bhandari, Tanvi Razdan, Saurav Mallik, and Zhongming Zhao. Explanation-driven deep learning model for prediction of brain tumour status using mri image data. *Frontiers in Genetics*, page 448, 2022.

[2] Asmaa Abbas, Mohammed M Abdelsamea, and Mohamed Medhat Gaber. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. *Applied Intelligence*, 51(2):854–864, 2021.

[3] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[4] Dimitrios Kollias, Athanasios Tagaris, Andreas Stafylopatis, Stefanos Kollias, and Georgios Tagaris. Deep neural architectures for prediction in healthcare. *Complex & Intelligent Systems*, 4(2):119–131, 2018.

[5] Linda Wang, Zhong Qiu Lin, and Alexander Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports*, 10(1):1–12, 2020.

[6] Chun-Fu Yeh, Hsien-Tzu Cheng, Andy Wei, Hsin-Ming Chen, Po-Chen Kuo, Keng-Chi Liu, Mong-Chi Ko, Ray-Jade Chen, Po-Chang Lee, Jen-Hsiang Chuang, et al. A cascaded learning strategy for robust covid-19 pneumonia chest x-ray screening. *arXiv preprint arXiv:2004.12786*, 2020.

[7] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet: Radiologist-level pneumonia detection on

chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

[8] Wesley H Self, D Mark Courtney, Candace D McNaughton, Richard G Wunderink, and Jeffrey A Kline. High discordance of chest x-ray and computed tomography for detection of pulmonary opacities in ed patients: implications for diagnosing pneumonia. *The American journal of emergency medicine*, 31(2):401–405, 2013.

[9] Amitojdeep Singh, Sourya Sengupta, and Vasudevan Lakshminarayanan. Explainable deep learning models in medical image analysis. *Journal of Imaging*, 6(6):52, 2020.

[10] Eri Matsuyama et al. A deep learning interpretable model for novel coronavirus disease (covid-19) screening with chest ct images. *Journal of Biomedical Science and Engineering*, 13(07):140, 2020.

[11] Blake VanBerlo. Investigation of explainable predictions of covid-19 infection from chest x-rays with machine learning. 2020.

[12] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.

[13] Igor Puleko, Oleksandra Svintsytska, Victor Chumakevych, Vadym Ptashnyk, and Yuliia Polishchuk. The scalar metric of classification algorithm choice in machine learning problems based on the scheme of nonlinear compromises. 2022.