

Received 14 February 2023, accepted 7 March 2023, date of publication 10 March 2023, date of current version 27 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3255403

RESEARCH ARTICLE

Interpretable Classification of Pneumonia Infection Using eXplainable AI (XAI-ICP)

RUEY-KAI SHEU^{ID1}, MAYURESH SUNIL PARDESHI^{ID2}, KAI-CHIH PAI^{ID3}, LUN-CHI CHEN^{ID3}, CHIEH-LIANG WU^{ID4,5,6}, AND WEI-CHENG CHEN¹

¹Department of Computer Science, Tunghai University, Taichung 407224, Taiwan

²Artificial Intelligence (AI) Center, Tunghai University, Taichung 407224, Taiwan

³College of Engineering, Tunghai University, Taichung 407224, Taiwan

⁴Department of Critical Care Medicine, Taichung Veterans General Hospital, Taichung 40705, Taiwan

⁵Department of Industrial Engineering and Enterprise Information, Tunghai University, Taichung 407224, Taiwan

⁶Department of Automatic Control Engineering, Feng Chia University, Taichung 40724, Taiwan

Corresponding author: Mayuresh Sunil Pardeshi (mayuresh@thu.edu.tw)

This work was supported by Ministry of Science and Technology, Taiwan (MOST 111-2321-B-075A-001).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Committee a.k.a. Institutional Review Board under TCVGH-IRB No. CE19309B.

ABSTRACT Open-box models in the medical domain have high acceptance and demand by many medical examiners. Even though the accuracy predicted by most of convolutional neural network (CNN) models is high, it is still not convincing as the detailed discussion regarding the outcome is semi-transparent in the functioning process. Pneumonia is known as one of the top contagious infections that makes most of the population affected due to low immunity. Therefore, the goal of this paper is to implement an interpretable classification of pneumonia infection using eXplainable AI (XAI-ICP). Thus, XAI-ICP is the highly efficient system designed to solve this challenge by adapting to the recent population health conditions. The aim is to design an interpretable deep classification and transfer learning based evaluation for pneumonia infection classification. The model is primarily pre-trained using the open Chest X-Ray (CXR) dataset from National Institutes of Health (NIH). Whereas, the training input and testing given to this system is Taichung Veterans General Hospital (TCVGH) for independent learning, Taiwan + VinDr open dataset for transfer learning of pneumonia affected patients with labeled CXR images possessing three features of infiltrate, cardiomegaly and effusion. The data labeling is performed by the medical examiners with the XAI human-in-the-loop approach. XAI-ICP demonstrates the XAI based reconfigurable DCNN with human-in-the-loop as a novel approach. The interpretable deep classification provides detailed transparency analysis and transfer learning for competitive accuracy. The purpose of this work, to design a re-configurable model that can continuously improve itself by using a feedback system and provide feasibility for the model deployment across multiple countries to provide an efficient system for the pneumonia infection classification. The designed model then provides detailed decisions taken at each step as transparency and features used within the algorithm for the pneumonia classification during the hospitalization. Thus, the scope can be given as explainable AI usage for the diagnosis classification using data preprocessing and interpretable deep convolutional neural network by the CXR evaluation. The accuracy achieved by using independent learning classification is 92.14% and is further improved based on successive transfer learning based evaluation is 93.29%. The XAI-ICP model adapts to the different populations by using transfer learning, while providing competitive results to the affected conditions.

INDEX TERMS Medical XAI, XAI pneumonia, transfer learning, pneumonia infection.

I. INTRODUCTION

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar^{ID}.

Recent advances in the computer based medical diagnosis (CAD) has improved the quality control process within the

TABLE 1. Recent developments within XAI pneumonia.

Reference	Study Aim	Source / Input Data	Preprocessing / Statistical Analysis	AI / XAI Algorithms	Evaluation
Zou L. et al. 2022 [23]	An ensemble XAI model for predicting pneumonia and covid-19 infections mortality	CXR of 1,475 adult patients.	SHAP and GRAD-CAM++	Ensemble XAI algorithm.	Precision, recall and Intersection over union (IOU).
Nagaoka T. et al. 2022 [24]	A deep learning system for predicting Covid-19	132 scans of CT slice images	GRAD-CAM	Inception v3 with feature map/ weighted feature map with hyper-parameters	Sensitivity, Specificity, Accuracy, AUC and ROC curve.
Leracitano C. et al. 2022 [25]	A Fuzzy based deep learning for Covid-19 pneumonia detection.	CXR images of 121 patients.	Saliency maps and Fuzzy edge detection algorithm.	A customized deep CNN (CovNNet) algorithm.	Sensitivity, specificity, positive predicted value (PPV), negative predicted value (NPV) and accuracy.
Hu B. et al. 2022 [26]	An explainable medical images and CXR evaluator for Covid-19.	1400 CXR images and ISIC 2017 skin lesion dataset consisting of 2000 images.	Learned embedded spaces.	Deep embedded network.	Similarity based saliency maps for occlusion, attention and activation mapping.
XAI-ICP	An explainable and transfer learning evaluator for pneumonia infection classification.	CXR images of 2301 patients. NIH and VinDr open dataset.	Pre-training and preprocessing methods.	Interpretable and DCNN classification algorithm with transfer learning.	SHAP, GRAD-CAM, recall, F1 score, specificity, precision, accuracy, AUROC curves and XAI Scoring.

International Skin Imaging Collaboration (ISIC) as a large-scale dataset of dermatoscopy images, Area Under the Receiver Operating Characteristic (AUROC).

hospitals [1], [2], [3], [4]. The range of applications within the medical field is enormous as most of the diagnosis deals with the symptoms, electronic health records (EHR), chest x-ray

(CXR), computed tomography (CT) scan, etc. The input data is taken into single or multiple forms, which may constitute a distinct approach to handle such an approach within machine learning (ML), deep learning (DL), explainable AI (XAI), etc. Therefore, a transparent system needs to be designed which can provide interpretability and explainability. Such systems will be impressive enough to the medical practitioners as it may convince them about the algorithm process and provide effective results [5], [6], [7], [8], [9], [10]. The problem statement for XAI-ICP is given as design and implementation of an interpretable deep classification algorithm for pneumonia prediction using explainable AI. The

pre-processing methods for improving the feature identification is an initial step, which can also be visualized by the interpretation process [11], [12], [13], [14]. The recent developments within XAI are shown in Table 1.

Several techniques can be applied for pre-processing including segmentation, reshape, augmentation, class weights, etc. [15], [16], [17]. The purpose of visualization provides details about the data distribution and analysis. Thus it can be used to decide the method selection based on its effectiveness and make the experimentation process efficient with appropriate planning [18]. The features can be specifically highlighted on the CXR image data during pre-processing by using SHapley Additive exPlanations (SHAP), heat-maps/Gradient-weighted Class Activation Mapping (GRAD-CAM), and Local interpretable model-agnostic explanation (LIME) for identifying high

scoring features, which is later prioritized by the classification algorithm for the evaluation, achieving better solutions. The selected features are crucial for the dataset explainability as well as for the input to the system model, which provides refined features that can optimize model implementation by avoiding unnecessary data processing and provide highly accurate results for the evaluation [19], [20], [21], [22]. The accuracy achieved will be higher for gaining user's trust and eliminating false alarms. The aim for this work is to design an interpretable deep classification and transfer learning based evaluation for pneumonia infection classification. Whereas, the scope can be given as explainable AI usage for the diagnosis classification using data preprocessing and interpretable deep convolutional neural network by the CXR evaluation. Even though the SHAP based top features identification is used, the XAI human-in-the-loop approach helps to accurately identify the features labeled by medical examiners, which effectively improves the accuracy and avoids false alarms in the system. The research gap is concerned about how to make the system an open-box model for high acceptability? Also the system should be feasible for testing across the countries [11], [50]. Therefore, the XAI provides complete model transparency and has higher acceptability with the detailed explanations. Whereas, the transfer learning and human-in-the-loop approach increases the feasibility for this application.

A. OBJECTIVES

1. To design an explainable AI system for the classification of pneumonia infection: The XAI building process is a crucial measure for the complete transparency and acceptance by the user. The XAI based decisions helps to gain the user's trust and accomplishes the complex tasks effectively. Subsequently, it needs to apply the XAI human in the loop approach for the training dataset and model enhancement.

2. To perform the pre-training of the XAI model and design re-configurable system: The pre-training with the globally compatible NIH CXR dataset makes the model flexible and better to classify for the test data. The model's performance is improved and makes it more effective. The re-configurable model can add/delete features based on the user specification.

3. To apply SHAP based processing for the effective DCNN evaluation: During the DCNN processing stage, high attention is given to the feature scoring process based on its current SHAP scores. Multiple case based feature scoring is analyzed and explained during the evaluation process. Therefore, as the dataset is updated, the feature scoring will adapt to that specific situation.

4. To implement an interpretable and transfer learning based evaluation system: The interpretable XAI system provides stepwise decisions with the dense neural network (DNN) layers, which helps to monitor the feature processing within every CXR section and evaluation by CNN models. The XAI-ICP classification algorithm evaluates the whole model with the open dataset and after feedback, the model

is updated to provide classification. The TCVGH hospital dataset used within this work is first evaluated using Deep Convolutional Neural Network (DCNN). Later, the data from multiple sources are combined as transfer learning to check for the system feasibility and evaluation.

One of the most prominent fields for XAI is in the medical system including pneumonia respiratory infection. Successive versions of AI had proved beneficial for evaluating input data source for the pneumonia, given as distinct vital signs, CXR, computational tomography (CT), scan, etc. For the model effectiveness, the pre-training by the large scale NIH CXR image dataset, which is afterwards fine-tuned on specific tasks having a nominal training dataset. The pre-training helps to quickly adapt to the new specific challenge and solve it. The evaluation of pneumonia detection is the popular practice observed in the recent research but it is lacking transparency. To address this issue, an XAI system is required that can provide explainability and high transparency within the system implementation. Therefore, this research work of XAI-ICP explains every decision taken by the operating algorithm in detail and supports transfer learning based re-evaluation of the system. Even though many input features are available for this system, the best pre-processing and feature highlighting will help to select the top features, selected from multi-class/binary classes, high confidence features, etc. The use of CXR images helps to completely evaluate the respiratory system of affected human body features. Whereas, the transfer learning can help to train CNN models with small data that can re-use the pre-train model on the new challenge case. Ultimately, the system will provide feature importance with classification and transfer learning based improvements to achieve a phenomenal system with explainability.

Pneumonia can be found in an affected patient either in a slight, moderate and severe, which needs to be treated in high priority. During the standard treatment period, the progress is observed and is helpful for the complete recovery, where a transparency in the system provides deep insights and feature importance. Therefore, the system model is suitable for the diagnosis analysis for the infection prediction. Several research works presented have insufficient specific human subject analysis based on the available data. To overcome such issues, XAI-ICP is designed to provide explainability and transparency of the feature importance and classification with transfer learning re-evaluation analysis.

II. LITERATURE SURVEY

The recent developments within the medical fields by using machine/deep learning has seen phenomenal achievements. Several papers are categorized based on their preprocessing, feature evaluation, system design, interpretation of the model and approach. The following survey will discuss recent design works and their implementation details by Figure 1.

The interpretability within XAI medical analysis includes GRAD-CAM, Guided GRAD-CAM, Respond CAM, Multi-layer CAM, layerwise relevance propagation (LRP), LRP Deeplight, etc. The gradient based localization for deep

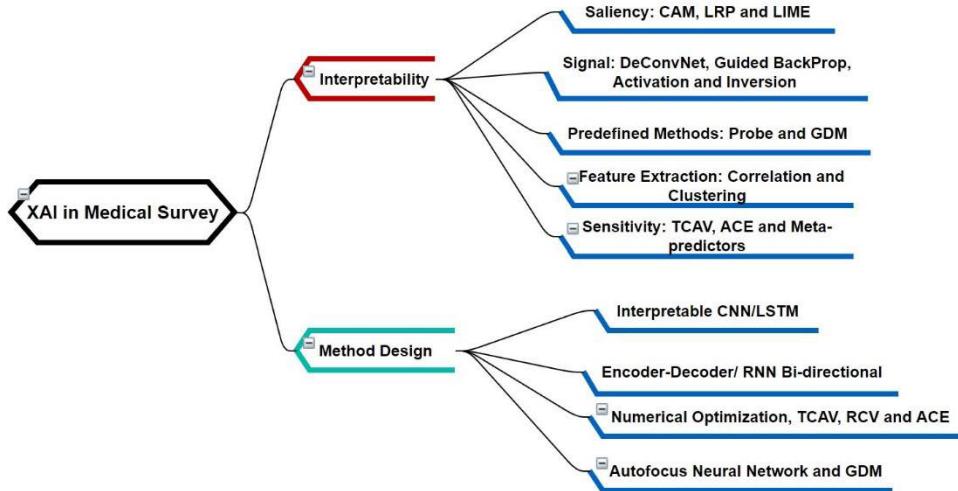


FIGURE 1. XAI based classification survey analysis.

network based visual explanations also known as GRAD-CAM is presented by Selvaraju RR et al. [20]. The data preprocessing is performed by GRAD-CAM and the evaluation is performed by Visual Geometry Group (VGG), structured CNN and reinforcement learning comparisons. The final evaluation is given by textual explanations and AUROC. The CNN pipeline for pathologies based alzheimer's disease is demonstrated by Tang et al. [27]. The data preprocessing is performed by guided GRAD-CAM and feature occlusion analysis, whereas the algorithm performs segmentation on heatmaps and CNN scoring. The evaluation is given by AUROC, precision-recall (PR) curve, t-test and p-value. The deep analyzing of 3D models by visualization is presented by Zhao et al. [28]. The respond CAM uses weighted feature maps and contours for preprocessing. The algorithm performs the sum to score property in 3D images by CNN and the evaluation is given as natural image captioning by prediction. The neural machine translator is demonstrated by Bahdanau et al. [29]. The implementation of multi-layer CAM by conditional probability for sentence pairs is preprocessed and then applied to encoder-decoder, neural machine translator and bidirectional Recurrent Neural Network (RNN). The evaluation is given by BLEU score, language translator and confusion matrix. Hans predictor unmasking and learning assessment of machines is presented by Lapuschkin et al. [30]. The data preprocessing is performed by relevance heatmaps of LRP. The algorithms are designed by class predictions with classifier, eigen based clustering, LRP and spectral relevance analysis, which is evaluated by detecting source tag, elements, orientations and Atari breakout. The LRP used for complex ML models interpretation predictions is demonstrated by Samek et al. [31]. The sensitivity analysis is performed as the part of data preprocessing. The LRP connection to the Deep Taylor Decomposition (DTD) algorithm is used and is evaluated using the qualitative and quantitative sensitivity analysis. The

use of recurrent deep learning for neuroimaging data analysis is presented by Thomas et al. [32]. The brain relevance maps are used for analysis by LRP deeplight and bi-directional long short-term memory (LSTM) for fMRI. The evaluation for brain activity is done by fine grained temporo-spatial variability, decoding accuracy and confusion matrix. In the similar way, the LRP CNN for text documents [33], LRP DNN for videos [34] and LRP Building and Interpreting Deep Similarity Models (BiLRP), for complex similarity models [35] can be applied.

Furthermore, the interpretability and method design includes local interpretable model-agnostic explanations (LIME), LIME MUSE (Model Understanding through Subspace Explanations), guideline-based additive explanation (GAX), testing with concept activation vectors (TCAV), regression concept vectors (RCV), automatic concept based explanations(ACE), etc. The LIME used to explain classifier predictions is demonstrated by Ribeiro [36]. In this work, interpretable data presentation is explained which is then sampled for local exploration and sparse linear explanations as a submodular optimization problem. It is later evaluated by recall, F1 scores, averages and prediction accuracy. A similar LIME MUSE is presented by Lakkaraju et al. [37]. The preprocessing is performed by quantifying fidelity, unambiguity, interpretability and implementing optimization procedure.

The evaluations are done by fidelity vs. interpretability tradeoffs and average time comparisons within the model. The lung nodules CAD given by GAX is demonstrated by Zhu and Ogino [38]. It shows anatomical features defined on the basis of the guideline are first generated using rule-based segmentation and anatomical regularities. The perturbation-based analysis algorithm is then used for measuring recall and global explanations by importance of anatomical features. The deep neural based features learning is given as t-SNE: Multifaceted feature visualization and is presented by Nguyen et al. [39]. The activation

maximization is used prior to the multifaceted feature visualization center-biased regularization algorithm. It is later evaluated as visualizing the different facets of a neuron, layerwise object comparison and model activation mapping comparison. The noisy brain data classification by using interpretable neural network based on node grouping is demonstrated by Yan et al. [40]. The t-SNE graph convolutional layer is used within architecture for relationships between brain subnetworks and cognitive functions. Evaluation is based on classification accuracy, runtime and F1 score of the validation set. To show that GroupINN is parsimonious, model sizes comparisons and interpretability is shown. Quantitative TCAV for high interpretability is presented by Kim et al. [41]. The pre-processing includes saliency maps with vanilla gradient, guided backprop, integrated gradient, smoothgrad, directional derivatives and conceptual sensitivity and TCAV. The evaluation is TCAV results with approximated ground truth, TCAV and accuracies of CAVs at each layer, empirical deepdream using CAVs for each layer in Googlenet.

The histopathology bidirectional explanations based on RCV is demonstrated by Graziani et al. [42]. The preprocessing is performed by Whole Slide Images (WSIs) with ground truth annotation on open tumor dataset. The algorithm consists of correlation to network prediction, regression concept vectors, sensitivity and evaluation of explanations. Additional experiments include comparison of TCAV ($\in [0, 1]$) and Br ($\in [-1, 1]$) scores. R^2 and RCV at different layers in the network, statistical significance of the scores, network accuracy and pearson's correlation. Automatic concept based explanations (ACE) is presented by Ghorbani et al. [43]. The preprocessing is similarity metric by euclidean distance in several layers of the ImageNet trained Inception-V3 architecture and chose the “mixed_8” layer. K-Means clustering is performed and outliers are removed using euclidean distance to the cluster centers. The algorithm consists of principles and desired data for concept based explanation by aggregating related local image segments across diverse data. The evaluation is prediction accuracy having (a) Intuitive correlations, (b) Unintuitive correlations and (c) Different parts of an object as separate but important concepts. Implementing semantic segmentation by autofocus layer is demonstrated by Qin et al. [44]. An autofocus convolutional layer with the number of candidate dilation rates. (a) The attention model. (b) A weighted summation of activations from parallel dilated convolutions is used in preprocessing. The design for the model is based on dilated convolution, autofocus convolutional layer and autofocus neural networks. Applying perturbations on the black box interpretations [45], use of linear classifier probes on intermediate layer explanations [46] and medical imaging multivariate inference by degenerative discriminative models [47] are also studied and useful in medical pneumonia [48] experiments. Ductal carcinoma diagnosis of 2D tissue slides using RenNet-32 and FastAI is demonstrated by Praveen et al. [49]. The ResNet-32 used within this work

is efficient, handling vanishing gradient problems supporting effective feature learning with GPU acceleration and fast callback mechanism. Evaluation is given by deep learning statistics of sensitivity, specificity, F1 Score, accuracy and Multi-fold cross-validation.

Detecting brain tumor in CT images by correlation learning mechanism (CLM) of Deep neural network is presented by Wozniak et al. [50]. The architecture consists of artificial neural network (ANN), back propagation training algorithm for ANN, multi-threading training model and adam algorithm for CLM. The model is evaluated by specificity, sensitivity, fallout, loss, accuracy, AUC and parallel processing time.

III. METHODOLOGY

The XAI-ICP dataset is taken from the Taichung Veterans General Hospital (TCVGH), Taiwan. This is a regional level hospital which provides capacity up to 1500 beds. The ethical review committee board panel from TCVGH has waived the consent of the patients for this research work purpose.

All the patient's identity related data is being de-identified. The patients considered within this dataset consist of adult patient and are admitted to the general ward within the hospital. Therefore, Intensive Care Unit (ICU) and extreme case condition patient's data is not included. The pneumonia criteria used to identify a patient is taken from the International Classification of Diseases (ICD) 10 and 11, which is at least collected for a single day CXR.

This work presents an XAI based Pneumonia infection classification, which provides XAI-ICP system transparency to the user at every decisional step. The continuous learning with human-in-the-loop approach known as user feedback makes the system re-usable for a long time. Also the system is pre-trained with the NIH dataset and transfer learning is added with TCVGH + Vindr open dataset which makes it adaptable to any international hospital. Many research works usually focus on evaluating either open datasets or local hospital datasets, so we have provided a combined work that can evaluate any dataset type. We have designed the XAI-ICP system model that consists of XAI based Deep CNN for providing a highly configurable model suitable for recent CXR classification. The XAI algorithm provides a detailed explanation for the designed system, which will be beneficial to the further developments and will serve as a new standard XAI process. The DCNN model is re-configurable and adaptable to the new dataset as well as feature set, which makes it unique for the AI field and infection classification application. Independent samples taken from the CXR for processing every feature/category are presented as above in Figure 2. Figure 2 provides the pneumonia infected patient's CXR sections. All the 7 sections are infected, and serves as the sample figure for the representational purpose from the open dataset.

The algorithm is constructed by interpretation of a neural net's internal state in terms of human-friendly concepts. For the training data, the XAI model presents training for

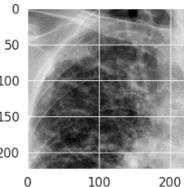
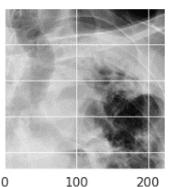
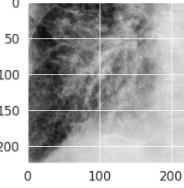
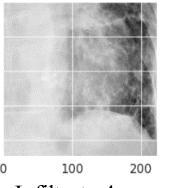
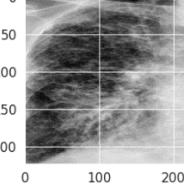
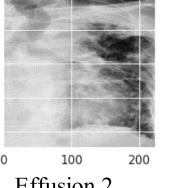
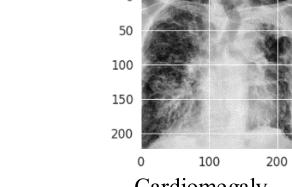
CXR Categories	Left Section	Right Section
Infiltrate (Top)		
	Infiltrate 1	Infiltrate 3
Infiltrate (Bottom)		
	Infiltrate 2	Infiltrate 4
Effusion		
	Effusion 1	Effusion 2
Cardiomegaly		
	Cardiomegaly	

FIGURE 2. Training CXR image data categories.

different features (infiltrate, effusion and cardiomegaly) in the different CXR sections. Infiltrate is observed in 4 sections, effusion in 2 sections and cardiomegaly in one section. It would be possible to divide cardiomegaly in two sections but due to less infected samples available, the accuracy is affected.

As per the medical AI regulations set by the European legislation in the General Data Protection Regulation (GDPR) standard indicates the importance of XAI [51]. It provides accurate explanation by the XAI system termed as “right to explanation”, for privacy protection and human dignity rights. Therefore, the final decision is taken by human in the loop process co-operating with automated processing and informed patient’s consent. The feedback based system is considered to be the crucial factor for the continuous improvement in the XAI system [52]. No system is perfect but it can be kept in practice and thought to be applicable until it is continuously adapting to the environment of various diseases/infections.

With the human in the loop approach, the dataset can be improved by the medical examiners by specifying the

TABLE 2. Open datasets used within XAI-ICP.

Dataset Name	Category	Contents	Source
National Institutes of Health (NIH)	Open	112,120 CXR images	U.S. Department of Health and Human Services.
VinDr	Open	18,000 CXR images	Vietnam
Taichung Veteran General Hospital (TCVGH)	Private	2301 CXR images	Taiwan

identification of particular symptoms within the CXR, EHR, clinical tests, etc. It helps to ensure the system is up to date with training and is capable of classifying millions of affected cases.

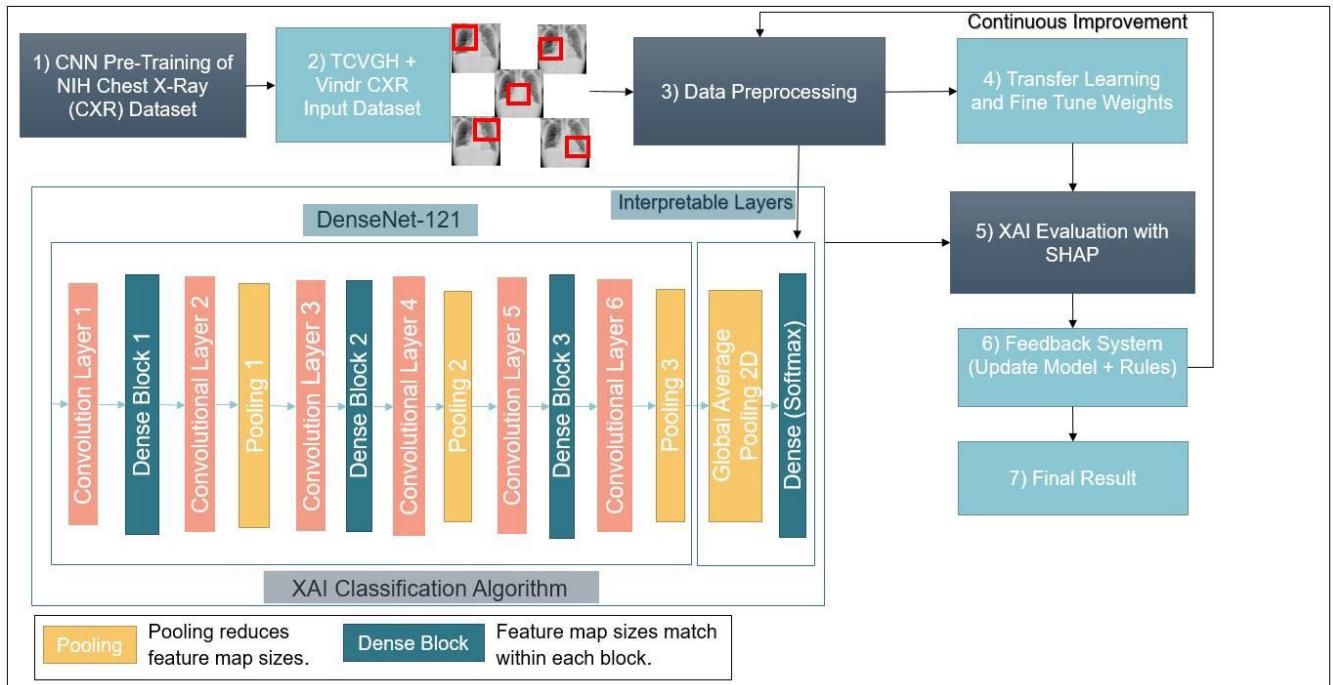
A. MEDICAL DATASETS AND TRAINING

The above Table 2 presents the open datasets that are used within XAI-ICP for pre-training weights, transfer learning and evaluation as NIH, VinDr and TCVGH datasets respectively. NIH dataset is selected for pre-training because it consists of numerous samples from diverse countries that are found to be best for the model. The TCVGH used for training and VinDr open datasets for transfer learning are used to provide the models efficiency in handling various test evaluations.

B. SYSTEM MODEL AND FLOWCHART DESIGN

The system model as shown above in Figure 3, can be explained into 7 major parts. At the beginning, the XAI-ICP model is pre-trained with the NIH CXR dataset to quickly adapt to the new CXR vision problem. The further training input given to the system is TCVGH + VinDr open dataset [53] of CXR images for pneumonia respiratory infection with the XAI human in the loop approach by three labeled features of infiltrates, cardiomegaly and effusion. Later, the data preprocessing performed on this model is to segmentation, standard scalar, data augmentation, etc. Next, the transfer learning is performed by utilizing a pre-trained network with input dataset and fine tune network weights. Successively, the XAI-ICP classification algorithm consists of a Dense-121 network with global average pooling 2D and dense layer having softmax for multiclass output. The detailed XAI based explanation is given in algorithm 1.

Exclusive DenseNet-121 Deep neural networks (DNNs) are models only used in model training. Therefore, in this process, an interpretable DNN model is used for the detailed decision transparency for the classification purpose. Further



on the feedback based evaluation collects data from the users to improve the system with rules for classification, which ultimately results in the XAI post-evaluation survey. The model is then made recursive to update the training data regularly, such that the resulting evaluation will be improved and also the features are suggested by the domain experts/users choice from their experiences.

Ultimately, the system will keep on improving continuously with the data and suggestions for providing the best accuracy. The best accuracy can be described as the highest scores achieved by the machine and human improvements. The XAI evaluation provides the transparency, interpretability and human in the loop approach as the final evaluation results. Thus, the final result decision is then given for the pneumonia diagnosis for the purpose of infection or not.

The XAI-ICP flowchart as shown in Figure 4 presents the system working in detail with the control flow and its respective operations. At the start, the model performs pre-training with an open dataset then training data of CXR images of the patient are taken. The dataset is then filtered with the pneumonia hospitalization criteria and is checked whether the patients are adults and are admitted in the general ward with the exception of ICU ward. The data preprocessing of the input dataset is performed to utilize a consistent and high quality data. These features are then given as input to the XAI DCNN classification algorithm with tuned hyper-parameters for obtaining better classification results. In every step of XAI process, the detailed explanation of the decisions taken by the algorithms are made available to the user and also data transparency is present. The disease severity analysis of the

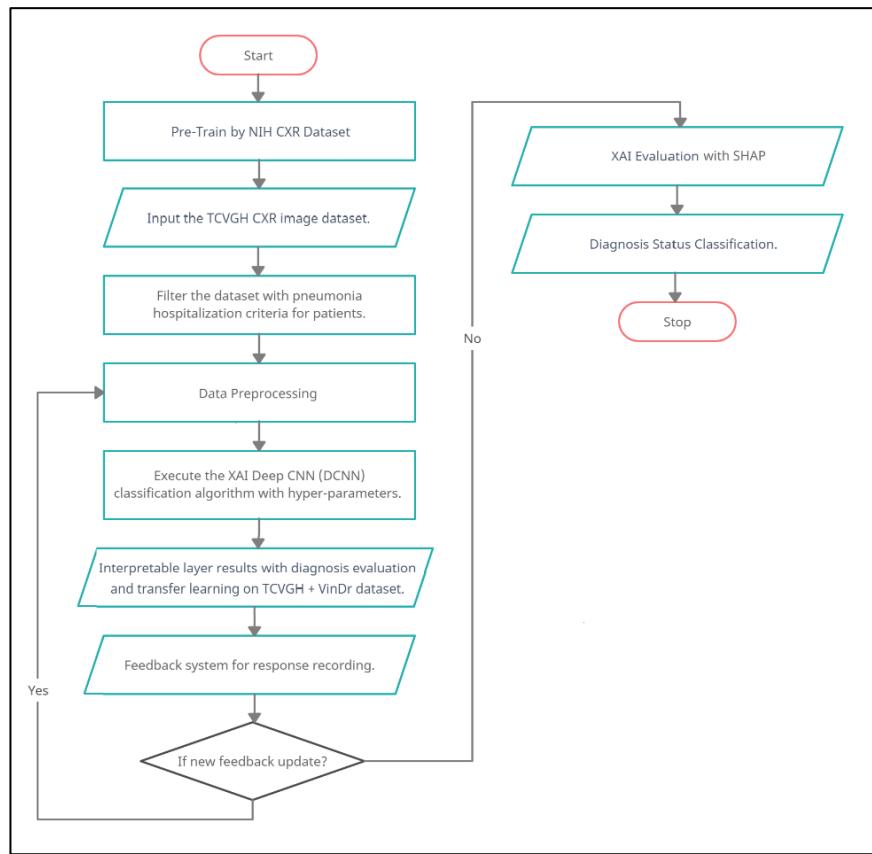
different sections within the CXR is given by different interpretable layers for the CXR section wise analysis. Therefore, a detail implementation of the patient's health status is made available. Successively, the transfer learning is made on a new external dataset for validation. A feedback system is then provided after the classification, so that the domain expert can provide their suggestions and diagnosis for the patient's condition better than the XAI classification results.

The purpose of this system is to be open for the supervision and to achieve continuous improvement for the system adaptation to the future classification improvements by XAI human in the loop approach as data labeling. Thus, if a new update is made in the form of suggestion by the domain expert then the new data recorded from the last update and hyper-parameter tuning is performed within the system. Grid search for the optimal hyper-parameters can be used to check for such cases optionally.

The XAI post survey evaluation presents the results after the improvement of the hyper-parameter, if any. The SHAP interpretation provides the highlighting of features with transparency to the medical examiners about the system process. Finally, the diagnosis decision is given by the XAI system from the patient's current health condition immediately during the hospitalization. A negative decision may help medical examiners to change the treatment process and improve the possibility of the patient's early diagnosis in the future.

C. ALGORITHM DESCRIPTION

The algorithm 1 for XAI-ICP prediction and classification accepts details for the CXR of the patient. The prediction is

**FIGURE 4.** XAI-ICP flowchart.

used for the shape detection in the CXR, whereas the classification is performed on the pneumonia detection. In step 1, the *IRaw* is the input given as the CXR image of the patient. In step 2, the *ILabelled* is the labelling performed by the medical examiner to improve the pneumonia features classification accuracy. In step 3, the *RInference* takes reference to specific CXR features necessary for fitness analysis by segmentation. In step 4, the threshold is determined by the medical examiner's experience for special cases. In step 5, the *XAIAnalysis* provides explainability to the XAI system for gaining the user trust as output. In step 6, the variables *IProcessed_{CG}*, *IProcessed_{IP}*, and *IProcessed* is initialized to NULL. In step 7, the DenseNet121 is pre-trained with NIH dataset and is termed as CNN. In step 8, the input image *IRaw* is imputed and categorized for features and stored in *IProcessed_{CG}*. In step 9, the *IProcessed_{CG}* is then resized to a standard scale and grey-scale is used for processing, which is then labelled and stored in *IProcessed_{IP}*. In step 10, the *IProcessed_{IP}* is segmented, standard scalar is applied and data augmentation makes it ready to be processed and is stored in *IProcessed*. In step 11, transfer learning is applied on DCNN, so that it can learn from VinDr dataset knowledge and apply in current classification and for performance scaling the interpretable layers are added as shown in the Figure 3 system model. In step 12, the do while loop is initiated, which

has the default first entry. In step 13, If *IProcessed* given as input to DCNN is available and valid for processing, then it is printed in step 14. In step 15, If the *RInference* specified by the medical examiner for image quality is greater than or equal to threshold then it is found to be valid and printed in step 16. In step 17, If the confidence of no diagnosis is classified as higher by the system than diagnosis confidence, then it is printed with percentage difference in step 18. In step 19, else for higher diagnosis confidence is printed in step 20.

The user in step 21, can suggest a feedback by providing valid or invalid CXR features for better performance in the future, which is taken as input by step 22. Else in step 23, if the CXR accuracy is not better than it is printed in step 24. In step 25, if the suggestions provided by the user/medical examiner is valid then do while loop is utilized. In step 26, if the accuracy of cardiomegaly is greater than that of CXR 4 section infiltrates then it is printed with the percentage in step 27, else in step 28, the higher accuracy percentage of CXR 4 section infiltrates is printed in step 29. Finally, in step 30, if the classification recall of the class weights performs better than without class weights then it is printed in step 31 with percentage. Else if the recall of not using class weights is better in step 32, then it is printed with percentage in step 33. Therefore, in step 34 all the explainable steps

Algorithm 1 XAI-ICP classification

```

1. Input:  $I_{Raw}$ , Input CXR image of patient.
2.  $I_{Labelled}$ , Labelling on CXR image by
   medical examiner.
3.  $R_{Inference}$ , Inference rules with threshold
   specified by the medical
   examiner's.
4.  $Threshold$ , Determined by the medical
   examiner's experience.
5. Output:  $XAI_{Analysis}$ , Explainable AI Analysis
6. Initialize ( $I_{ProcessedCG}$ ,  $I_{ProcessedIP}$ )
 $I_{Processed} = \emptyset$ 
7.  $CNN = \text{Pre-Train}_{NIH}(\text{DenseNet-121})$ 
8.  $I_{ProcessedCG} = \text{Categorization}(\text{Imputation}(I_{Raw}))$ 
9.  $I_{ProcessedIP} = \text{Grey-Scale}(\text{Resize}(I_{ProcessedCG}))$ 
 $\cup I_{Labelled}$ 
10.  $I_{Processed} =$ 
    Data-Augmentation(StandardScalar(Segmentation
        ( $I_{ProcessedIP}$ )))
11.  $DCNN = \text{TransferLearning}_{Vindr}(\text{CNN}(I_{ProcessedIP}))$ 
 $\cup \text{Interpretable-Layers}$ 
12. Do {
13.   If ( $DCNN(I_{Processed})$  is Valid)
14.     FPrint("CXR shape detection and classification
           is successful")
15.   If ( $R_{Inference} \geq Threshold$ )
16.     FPrint("All CXR features are valid as per
           inference rules")
17.   If (Confidence(No Diagnosis)
        > Confidence(Diagnosis))
18.     FPrint("No Diagnosis confidence is higher by
           (No Diagnosis - Diagnosis) × 100 percent")
19.   Else
20.     FPrint("Diagnosis confidence is higher by
           (Diagnosis - No Diagnosis) × 100 percent")
21.   If (Can you suggest some valid/invalid features in
        the CXR to be altered for better performance?)
22.     Input("User comments for valid/invalid features")
23.   Else
24.     FPrint("CXR performs better due to more
           symptoms features")
25. }While(Valid Suggestions)
26. If (Cardiomegaly.accuracy > 4infiltrates.accuracy)
27.   FPrint("The CXR 4-section infiltrates
           including cardiomegaly has the accuracy
           (Cardiomegaly.accuracy) × 100 %")
28. Else
29.   FPrint("The CXR 4-section infiltrates performs
           better with the accuracy
           (4infiltrates.accuracy) × 100 %")
30. If (ClassWeights.recall [1] >
      NoClassWeights.recall [1])
31.   FPrint("Class weights performs better with
           (ClassWeights.accuracy) × 100 %")
32. Else
33.   FPrint("No Class weights performs better with
           (NoClassWeights.accuracy) × 100 %")
34.  $XAI_{Analysis} = \text{FPprint}$ 
35. Return  $XAI_{Analysis}$ 

```

are collectively stored in the $XAI_{Analysis}$ by this algorithm. Finally, in step 35, the $XAI_{Analysis}$ is returned to the calling algorithm.

D. MATHEMATICAL MODEL

In preprocessing, to make the present image appear more natural, the pixel intensity is adjusted by using image normalization. The training model's computational complexity is reduced using Z-score, which is given in equation 1 as

$$Z\text{-score} = \frac{X - \mu}{\alpha} \quad (1)$$

where x is image pixel, μ belongs to the per pixel mean and α belongs to the standard deviation.

The conventional CNN use non-linear transformation H_l as the output layers (l^{th}) by using the output of previous layers X_{l-1} given as equation 2

$$X_l = H_l(X_{l-1}) \quad (2)$$

In case of DenseNet [54], the input and output layer function maps are concatenated instead of summing them up. The layers consisting of information flow as an easy communication model is provided by DenseNet. The previous layer feature input's is received from l^{th} layer. The equation 3 is then represented as

$$X_l = H_l[(X_0, X_1, X_2, \dots, X_{l-1})] \quad (3)$$

The output maps achieved from past layers are concatenated to obtain single tensor as $[X_0, X_1, X_2, \dots, X_{l-1}]$. The non-linear transformations function is given as H_l with three major operations as convolution (CONV), batch normalization (BN), and activation (ReLU). The l^{th} layer is generated by the growth rate K as $K^{[l]} = (K^{[0]} + K(l-1))$, where the number of channels is $K^{[0]}$.

Whereas, the training samples are improved for the actual model without collecting additional data is achieved by data augmentation. The source image is transformed while holding the semantic information with an increased number of samples as bias free is given by data augmentation. In model training, the techniques of random cropping, color space augmentation and horizontal flipping are employed to avoid bias in image because if similarity in the underlying image part. Ultimately, data augmentation prevents overfitting by learning diverse feature collections leading to increase dataset size.

To minimize regression or classification model error domain divergence, good features needs to be found using inductive transfer learning [55]. In case of supervised feature construction methods that are similar to multi-task learning, which shares low-dimensional representation within the related tasks. Therefore, the regression or classification model error is reduced using learned representation. The optimization problem is solved by learning common features of inductive transfer learning.

$$\arg \min_{A, U} \sum_{t \in \{T, S\}} \sum_{i=1}^{n_t} L(y_t, \langle a_t, U^T x_{t_i} \rangle) + \gamma \|A\|_{z, 1}^2 \quad (4)$$

Here, the tasks of source domain S and target domain T are given. The parameter matrix is $A = \{a_s, a_T\} \in R^{d \times 2}$ is used

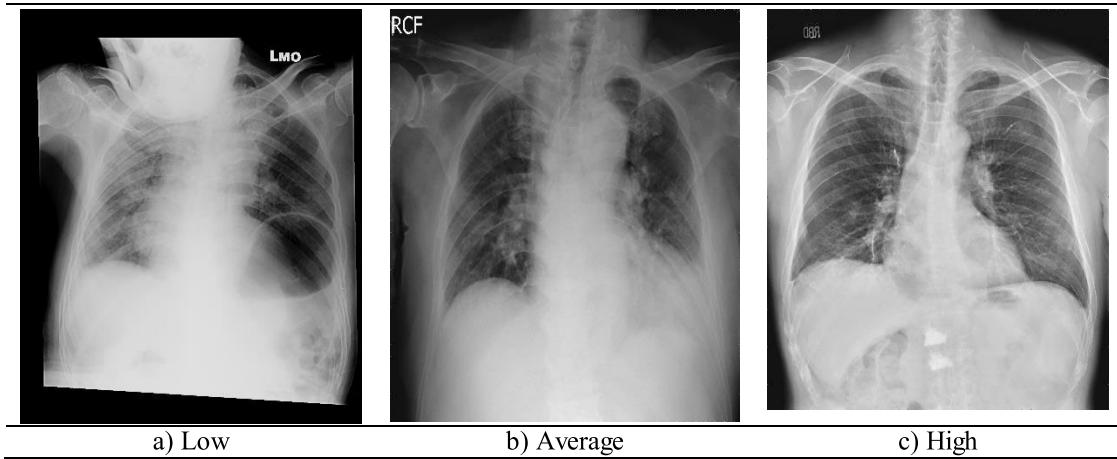


FIGURE 5. CXR image quality of the patient's data.

to map high dimensional to low dimensional data representations. $\|A\|_{r,p} := (\sum_{i=1}^d \|a^i\|_r^p)^{\frac{1}{p}}$ is given by (r, p) -norm of A. The equation 4, presents the optimization problem for $U^T X_T, U^T X_S$ as estimates for low-dimensional representations and model having parameter A simultaneously. Nevertheless, this optimization problem can be further transformed into convex optimization problem for solving.

In case of unsupervised feature construction, sparse coding is applied for gaining higher level features by transfer learning. In this approach, $b = \{b_1, b_2, b_3, \dots, b_S\}$ basis vectors of higher level features are captured for solving optimization problem by source domain data is given as

$$\min_{a,b} \sum_i \left\| x_{S_i} - \sum_j a_{S_i}^j b_j \right\|_2^2 + \beta \| \alpha_{S_i} \|_1$$

Such that, $\|b_j\|_2 \leq 1, \quad \forall j \in 1, \dots, S$

(5)

In the above equation 5, basis b_j has a new representation $a_{S_i}^j$ for input x_{S_i} and feature construction as well as regularization. Successively basis vector b learning, higher level features are learned by using equation 6 optimization function applying on target domain data.

$$a_{T_i}^* = \arg \min_{a_{T_i}} \left\| x_{T_i} - \sum_j a_{T_i}^j b_j \right\|_2^2 + \beta \| \alpha_{T_i} \|_1$$
(6)

The regression or classification models are trained by applying discriminative algorithm with labels to $\{a_{T_i}^*\}$'s for use in target domain.

IV. EXPERIMENTS

In the experiment section, details about the experiment's setup, XAI process implementation and final results are demonstrated. To implement a system consisting of multiple DNNs, a parallel architecture is required that can execute the process within acceptable time limits. Thus a new system is assigned for this work and utilized to perform experiments.

TABLE 3. System configuration.

System	Workstation (Ubuntu 18.04 LTS)
Processor	Intel Core i7 – 8700K
Memory	32G, DDR4
Graphics Card (GPU)	TitanV – CUDA 9
Python Library	Numpy, Pandas, Matplotlib, Seaborn, OpenCV and Keras.

The system configuration for the experiment hardware is given in Table 3.

A. MEDICAL DATASET LABELLING AND PREPROCESSING

In the first phase of the experiments, the open dataset included is in the form of CXR. The CXR is labeled by the medical examiners to know whether a patient has symptoms for the pneumonia respiratory infection. The readings are recorded for the 2031 patients within the VinDr dataset on the diagnosis day. The corrupted CXR data is discarded of a total 331 images, so the remaining TCVGH 601 original data is combined with open data of 1700 high quality images.

The CXR image quality analysis is crucial as the low image quality may lead to misclassification and will not be acceptable by the medical examiner as shown in Figure 4. Thus, algorithm 1 checks for the input of CXR image quality and notify the user regarding the input data quality.

In Figure 5, CXR images of low quality are not recommended to be used for training as it affects the model learning. Whereas, the average quality having light flashes or other blind spots are not acceptable.

Therefore, only the high quality CXR images are recommended for the model training and others are rejected from the filtering. Later, the CXR is divided in 4 parts as given in Figure 6 SHAP based highlighting of different CXR sections of 3 patients CXR images that are affected with pneumonia. The SHAP based feature highlighting provides the graphical view of the pneumonia infection on the different CXR

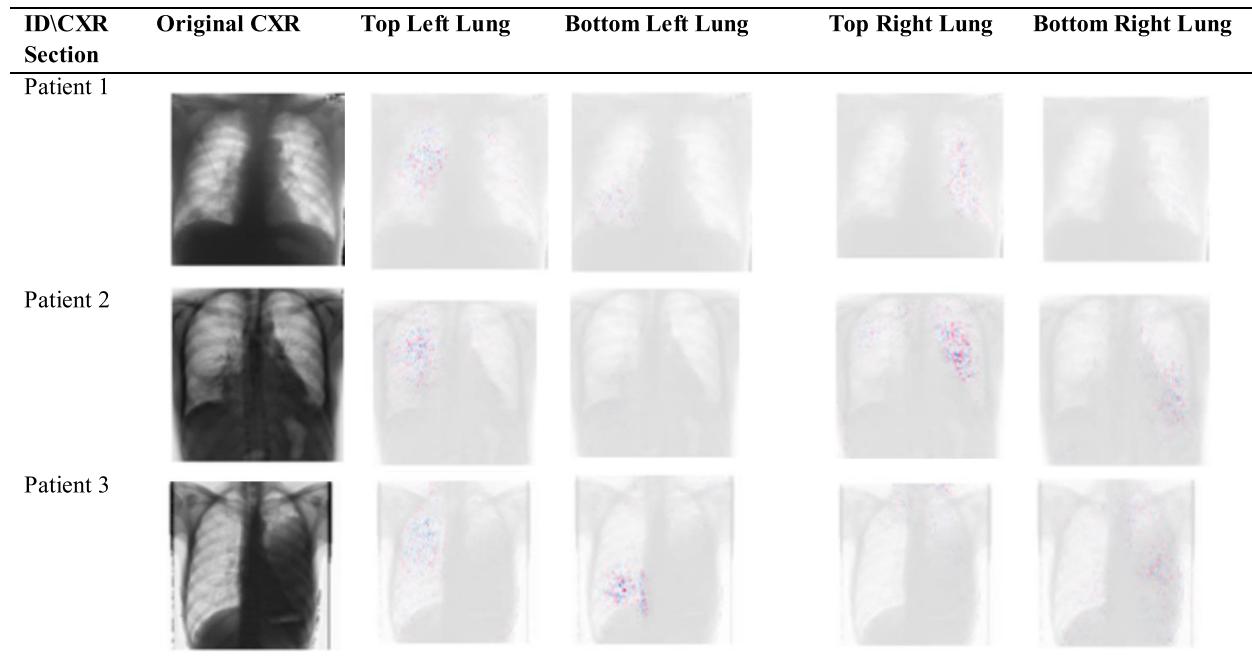


FIGURE 6. SHAP highlighting of different CXR sections.

ID\Context	CXR Image	Image Quality	Infiltrate (TL)	Infiltrate (TR)	Infiltrate (BL)	Infiltrate (BR)	Cardio--megaly	Effusion (Left)	Effusion (Right)
Patient 2		2 (High)	3	3	3	2	3	0	0
Patient 3		1 (Moderate)	1	3	0	1	3	0	0
Patient 4		2 (High)	2	3	3	3	0	3	0

Labelling: 0: Normal, 1: Slight, 2: Moderate, 3: Severe

FIGURE 7. Labelling by the medical domain expert.

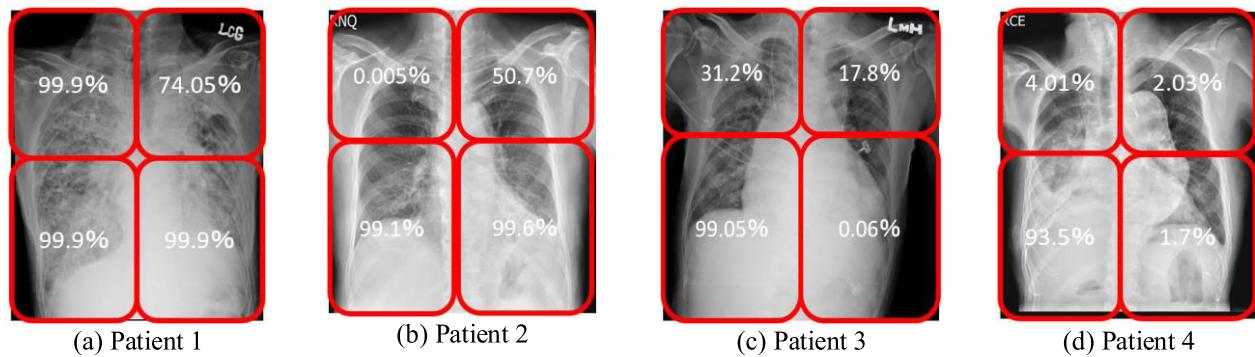
sections. The SHAP displays intensity of highlighting within the different blue and red colors for the low to high infection value.

The exact position of SHAP colors is presented as a part of XAI evaluation to the medical examiners for ease of locating the infection, understanding the severity and providing the correct treatment.

Successively, the domain experts labeling for the 500 CXR images is performed for training the model to identify the level of image quality, infiltrate, cardiomegaly and effusion as given in Figure 7. This labeling is necessary for the system

to later evaluate the test set and provide accurate results about the current condition of the patient's health status. In case the training images are not correctly classified, the medical experts add labeling to such images which provides better model training and evaluation by classification. Therefore, it will become possible during the interpretation of DNN layers for the different categories of patients affected with the severity level. Thus a highly transparent and explainable system is made available that can help to understand the differences and learn about the exact affected CXR image parts of the patient's, which are utilized for this work. The

CXR Location \ Status	Severity Level 0 (Normal)	Severity Level 1 (Slight)	Severity Level 2 (Moderate)	Severity Level 3 (Severe)
Infiltration Top Right Lung				
Infiltration Bottom Right Lung				
Infiltration Top Left Lung				
Infiltration Bottom Left Lung				
Cardiomegaly				
Effusion Right Lung				
Effusion Left Lung				

FIGURE 8. CXR infiltrate severity levels during CNN interpretations.**FIGURE 9.** CXR analysis for the section-wise classification.

data augmentation performed during preprocessing is given as rotation range = 30, wide shift range = 0.30, height shift range = 0.30 and zoom range = 0.30. The patients CXR images are also stored and processed for data analysis for the interpretable classification of pneumonia.

B. SEVERITY ANALYSIS BY 4-CLASS MODEL INTERPRETATION

For presenting the CNN interpretation results, different severity levels are being given as slight, moderate and high in Figure 8. As the interpretations are given in higher detail, the

TABLE 4. DenseNet121 based evaluation.

DenseNet121 \ Evaluation	Precision	Recall	F1-Score	Accuracy
Infiltration (Top Right Lung)	0.83	0.82	0.82	0.82
Infiltration (Bottom Right Lung)	0.81	0.79	0.78	0.79
Infiltration (Top Left Lung)	0.83	0.82	0.83	0.82
Infiltration (Bottom Left Lung)	0.78	0.78	0.78	0.78

classification will provide better accuracy with status prediction. The XAI based severity categorization provides higher classification of the CXR infection level. Proper categorization can help medical examiners to provide correct treatment and avoid high dosage for recovery. Different patients are shown with their respective CXR sections affected with severity levels. The severity levels are given in detail for the infiltrate, cardiomegaly and effusion with the affected regions.

C. EVALUATION OF THE 4-CLASS CXR DATASET USING CONVOLUTIONAL NEURAL NETWORK (CNN)

In this subsection the CNN based evaluation is provided as by Dense Convolutional Network (DenseNet121) and Residual Neural Network (ResNet152 v2) neural networks in Table 4 and 5 respectively. The highest values within that category are highlighted with bold text. It can be seen that DenseNet121 evaluates better on infiltrate category for the top section with top right and top left lungs, which are basically the highly affected section within the CXR. Whereas, the ResNet152 v2 has higher classification of the infiltrate category within the top left CXR section.

Figure 9 shows the classification provided by the DCNN model for the pneumonia affected patient. The infection classification is shown in the form of percentage in Figure 9 as a visual representation of Figure 6. It records the affected patient's infection scores based on the CXR section. For patient 1, the higher accuracy can be noticed by the 4 sections as the pneumonia classification presented, which is very high. For patient 2, the lower sections are showing the higher values,

Which is usually rare to be seen but the ground truth here is false. So patient 2 does not have pneumonia. Similarly, in case of patient 3 and 4, only the bottom left section is showing affected. Therefore, it is concluded that the patient does not have pneumonia as shown by the low classification results of the bottom left CXR section by the DenseNet121 and ResNet152 v2. The medical examiner's team needs to be trained for such cases to have proper understanding of the model prediction and final classification. This challenge is overcome by the XAI that explains the system transparency for the patient 1 as:

TABLE 5. ResNet152 V2 based evaluation.

ResNet152 \ Evaluation	Precision	Recall	F1-Score	Accuracy
Infiltration (Top Right Lung)	0.83	0.82	0.82	0.82
Infiltration (Bottom Right Lung)	0.81	0.79	0.78	0.79
Infiltration (Top Left Lung)	0.83	0.82	0.83	0.82
Infiltration (Bottom Left Lung)	0.78	0.78	0.78	0.78

The CXR left section indicates severe infection.

The CXR right section indicates severe infection.

Therefore, the complete CXR on an average indicates severe infection.

D. EVALUATION OF THE 2-CLASS TCVGH CXR DATASET USING DEEP CONVOLUTIONAL NEURAL NETWORK (DCNN)

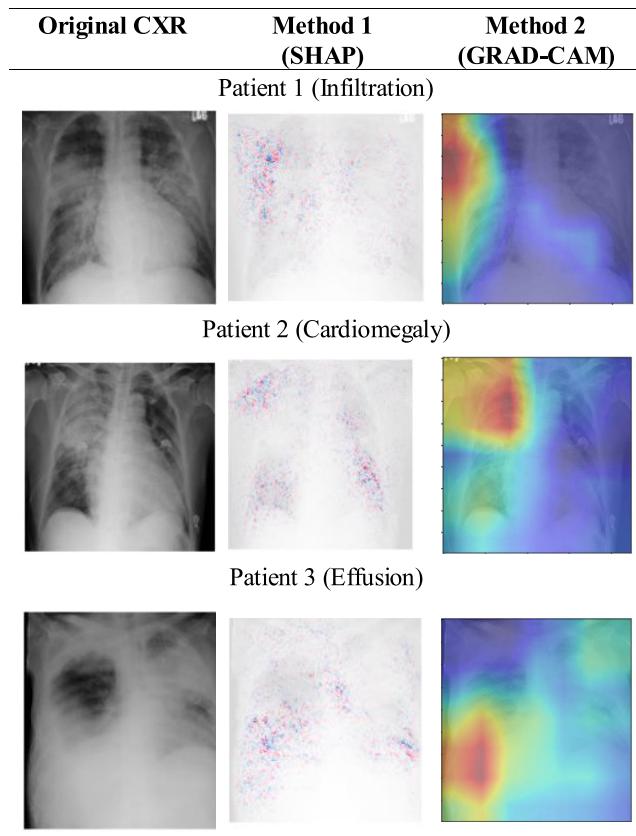
The Chest radiographs are referred from the TCVGH CXR Dataset, which is classified by the XAI-ICP model as shown with results in Table 6. The pneumonia based different features of infiltrate, cardiomegaly and effusion are separately classified for the better understanding. For the data adaptability, as shown in Table 6 the model is at its best to classify the TCVGH dataset with better accuracy and Area Under the Curve (AUC).

The performance for 2-class was performing better, where classification is given as 0 → Normal or 1 → Slight /Medium/Severely infected. In this case, the precision, recall and F1-score for normal are higher for class 1 infected due to the input strength of normal patients CXR data is higher than symptomatic patients count. The Figure 10 shows the SHAP diagram which is crucial to highlight top features within the CXR image dataset. The top features help to provide better classification and the false positive rate is controlled. The XAI based evaluation for the CXR by the model interpretation for different sections is given in Figure 10. The CXR after preprocessing shows that SHAP evaluation is more precise in locating the features as compared to the GRAD-CAM. Whereas, the comparison of XAI-ICP with the DCNN model having multiple features is presented in Table 6 with details. The dominance of XAI-ICP with different factors and results can be noticed clearly to achieve high classification.

Even though the accuracy and AUC are found to be remarkable in the TCVGH dataset classification by DCNN, the precision, recall and F1-score are found to be lagging in the scores. Therefore, to avoid false positive prediction in the future for the sensitive medical cases, transfer learning is implemented in the next subsection that overcomes the limitations of the local dataset and local optima results.

TABLE 6. TCVGH dataset test by DCNN.

DCNNEvaluation	Normal Patient	Symp-tomatic Patient	Precision		Recall		F1-score		Accuracy	AUC
			0	1	0	1	0	1		
Infiltration (Top Right Lung)	623	137	0.97	0.76	0.94	0.86	0.95	0.81	0.93	0.97
Infiltration (Bottom Right Lung)	567	193	0.92	0.72	0.9	0.76	0.91	0.74	0.86	0.93
Infiltration (Top Left Lung)	646	114	0.92	0.81	0.98	0.49	0.95	0.61	0.91	0.95
Infiltration (Bottom Left Lung)	581	179	0.96	0.87	0.96	0.86	0.96	0.86	0.94	0.98
Cardiomegaly	683	77	0.96	0.7	0.97	0.64	0.96	0.67	0.94	0.95
Effusion (Left)	705	55	0.94	0.67	0.99	0.22	0.97	0.33	0.94	0.9
Effusion (Right)	717	43	0.96	0.37	0.97	0.33	0.96	0.35	0.93	0.87

**FIGURE 10.** XAI saliency maps.

E. TRANSFER LEARNING OF TCVGH + VINDR CXR DATASET USING DEEP CONVOLUTIONAL NEURAL NETWORK (DCNN) FOR 2-CLASS CLASSIFICATION

The CXR training and tests are referred from the TCVGH + VinDr Chest X-ray Dataset [56] is classified by the XAI-ICP model as shown with results in Table 7. In case of recall,

the transfer learning in Table 7 shows much better results due to good model fitting. Successively, the TCVGH and VinDr open dataset is combined to check for transfer learning and the accuracy and AUC is found to be performing better than before due to feasibility of the dataset. Nevertheless, the combined effusion for 2-class results show improvements over all the independent effusion (left) and effusion (right) scores. Figure 11 provides the XAI scoring system a total of score $\geq 90\%$ for the validity and effectiveness of the XAI-ICP [57]. As pneumonia is not a chronic disease, the recommendation score can be relaxed. Therefore, the XAI-ICP model is evaluated to be in Class I grade.

Figure 15 shows the confusion matrix (CM) for the TCVGH + Vindr dataset classification. The Vindr dataset is distinct and is finely adapted to be classified by the current model.

In correspondence of the confusion matrix, the Receiver Operating Characteristic (ROC) curves for the TCVGH classification is given in Figure 12 (a to h). The ROC curves are performing quite better for the TCVGH 2-class than previous 4-class classification. From Figure 14, the confusion matrix shows the less quantity of false positive and false negative cases, whereas it improves the ROC curve well in Figure 12 (i to p). The ROC curve presents the good score for the different sections within CXR of 0.98 as the highest for the infiltrate category and 0.87 as minimum recorded score for the effusion due to less occurrences for training samples. This work serves as an effective case study that will be useful to reference in the future works, presenting several opportunities to overcome such issues and provide an adaptable model.

The combined effusion shows improvements over all the independent effusion (left) and effusion (right) scores. Figure 14 shows the confusion matrix (CM) for the TCVGH dataset classification. Due to the use of pre-trained models of the NIH dataset, the current classification of TCVGH

TABLE 7. TCVGH + vindr dataset test by DCNN.

DCNNEvaluation	Normal Patient	Symp-tomatic Patient	Precision		Recall		F1-score		Accuracy	AUC
			0	1	0	1	0	1		
Infiltration (Top Right Lung)	623	137	0.97	0.8	0.95	0.85	0.96	0.82	0.93	0.97
Infiltration (Bottom Right Lung)	567	193	0.95	0.91	0.97	0.84	0.96	0.87	0.94	0.98
Infiltration (Top Left Lung)	646	114	0.98	0.67	0.92	0.88	0.95	0.76	0.92	0.96
Infiltration (Bottom Left Lung)	581	179	0.97	0.85	0.95	0.89	0.96	0.87	0.94	0.98
Cardiomegaly	683	77	0.96	0.68	0.96	0.68	0.96	0.68	0.93	0.96
Effusion (Left)	705	55	0.96	0.53	0.96	0.53	0.96	0.53	0.93	0.94
Effusion (Right)	717	43	0.96	0.47	0.98	0.35	0.97	0.4	0.94	0.89
Effusion (Combined)	684	76	0.96	0.57	0.95	0.63	0.95	0.6	0.92	0.93

TABLE 8. Comparison with recent models.

Reference / Features	Transparency	Explainability	Human-in-the-loop	Transfer Learning	Results
Cohen JP et al., 2020 [21]	Y	Y	N	N	MAE: 1.14 MAE: 0.78
Zou L. et al., 2022 [23]	Y	Y	N	Y	Precision: 0.52 Recall: 0.57
Nagaoka T. et al., 2022 [24]	Y	Y	N	N	Sensitivity: 84% Specificity: 76% Accuracy: 84%
Ieracitano C et al., 2022 [25]	Y	Y	Y	N	Sensitivity: 82.5% Specificity: 78.6% Accuracy: 81%
XAI-ICP	Y	Y	Y	Y	Accuracy: 92.14% AUC: 93.57%

**FIGURE 11.** XAI scores for XAI-ICP.

does not face any limitation in this designed model for classification.

F. DCNN PERFORMANCE MEASUREMENTS

The performance measurement for different CXR sections for infiltrate, cardiomegaly and effusion evaluation

is demonstrated in Figure 13. Even though the training required for different features has unequal samples quantities available, the evaluation time required is almost similar.

The infiltrate evaluation requires similar time with no big difference as the CXR image is divided into 4 sections. Similarly, for the effusion, the time required is similar as the CXR image is divided into 2 sections as left and right. In case of cardiomegaly, the time required is the lowest as the CXR is not divided into multiple sections and is processed independently. The comparison of XAI-ICP with different models having multiple features is presented in Table 8 with details. The dominance of XAI-ICP with different factors and results can be noticed clearly to achieve high classification.

G. CRITICAL ANALYSIS AND DISCUSSION

The open-box models provide transparency to the user, which helps to gain user trust and confidence in the XAI-ICP system. In addition, the features detected and evaluated are

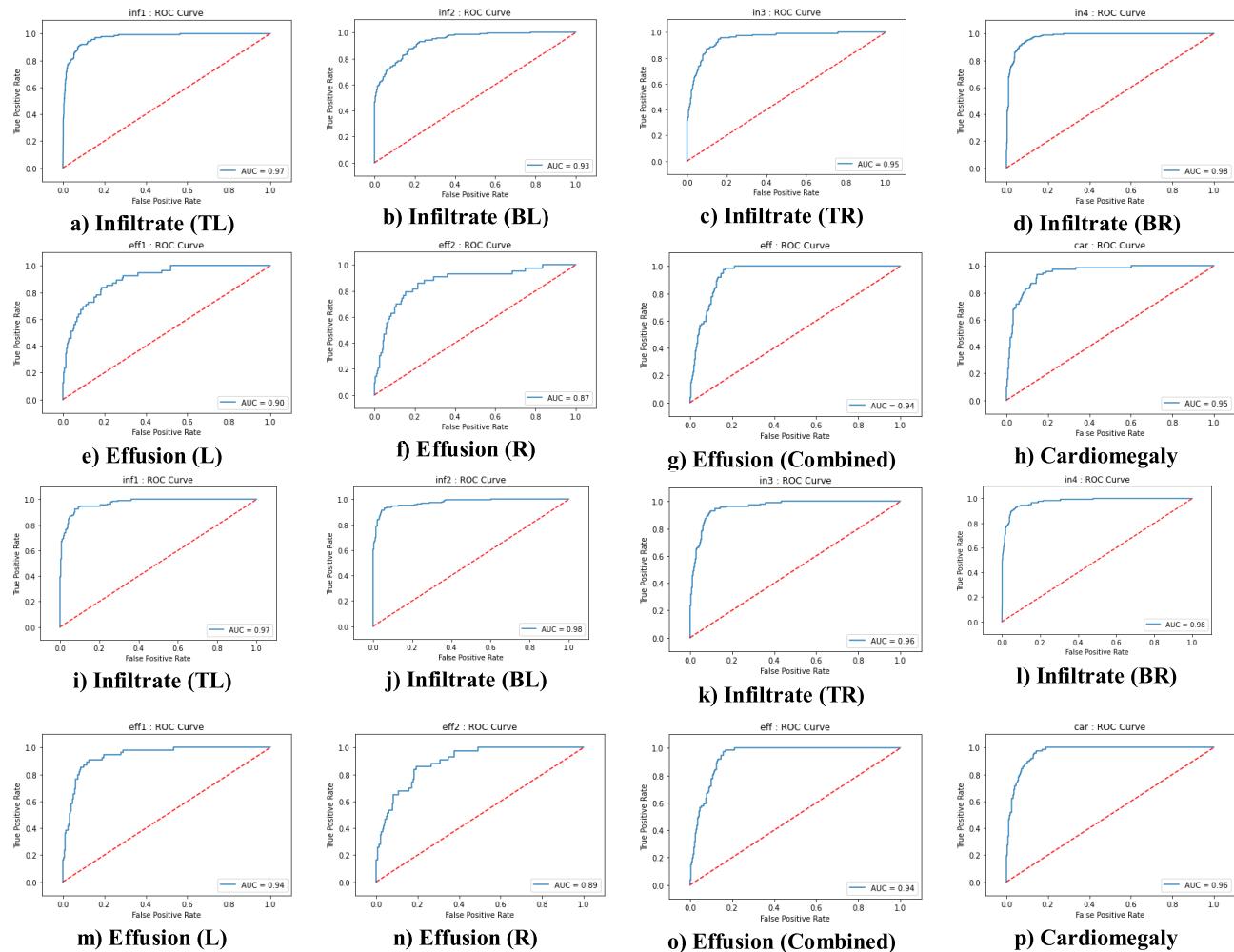


FIGURE 12. (a to h) ROC curve for TCVGH + Vindr classification without transfer learning and (i to p) ROC curve for TCVGH + Vindr classification with transfer learning.

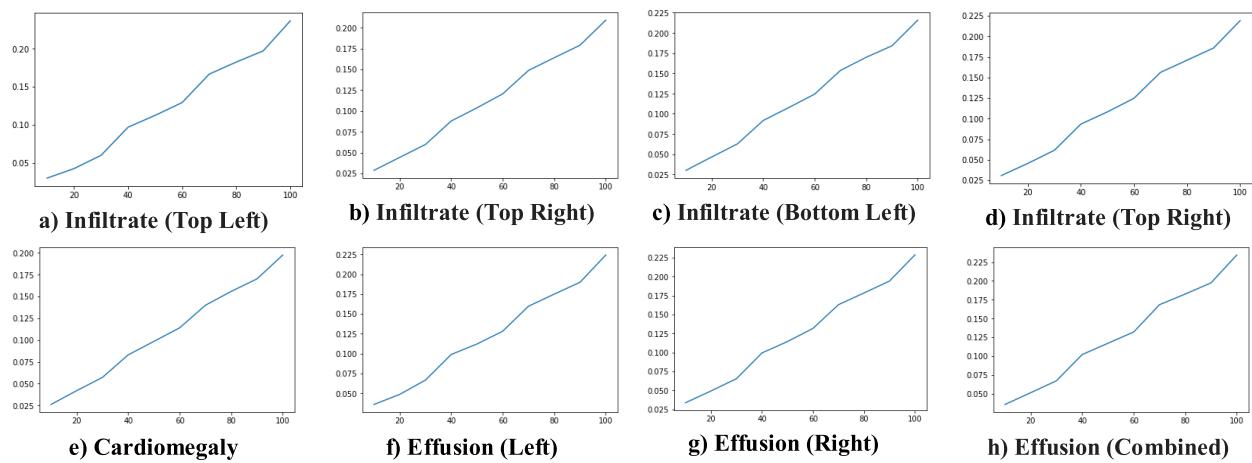


FIGURE 13. Performance measurement for different CXR features. (Measurements: X-axis = Time (ms) and Y-axis = Test Quantity).

provided in more detail as a part of explainable AI (XAI). The human-in-the-loop approach is useful to receive user

feedback and continuously improve the system. Nevertheless, the detailed evaluation provides a recently updated and

improved system model. The XAI-ICP impact can be analyzed by its transfer learning approach that adapts to the international dataset [58] and can be feasible to be used in different countries simultaneously. The novelty in design can be beneficial for its importance and distinctness in the market.

V. CONCLUSION

Pneumonia is one of the fatal infection known to affect the population worldwide. This work is suitable to the general ward patient as well as ICU ward patient's whose XAI based classification provides better understanding for helping medical examiner's to plan recovery treatment. The XAI-ICP

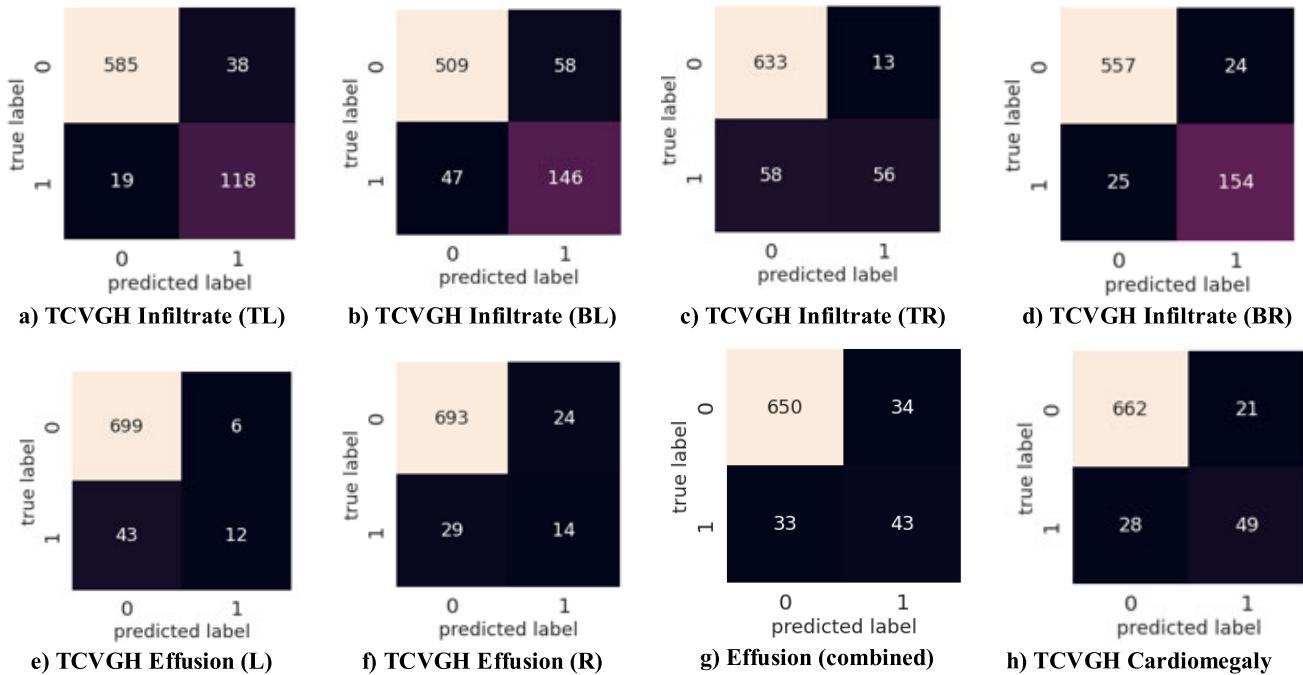


FIGURE 14. Confusion matrix for TCVGH + VinDr Classification without transfer learning and NIH pre-training.

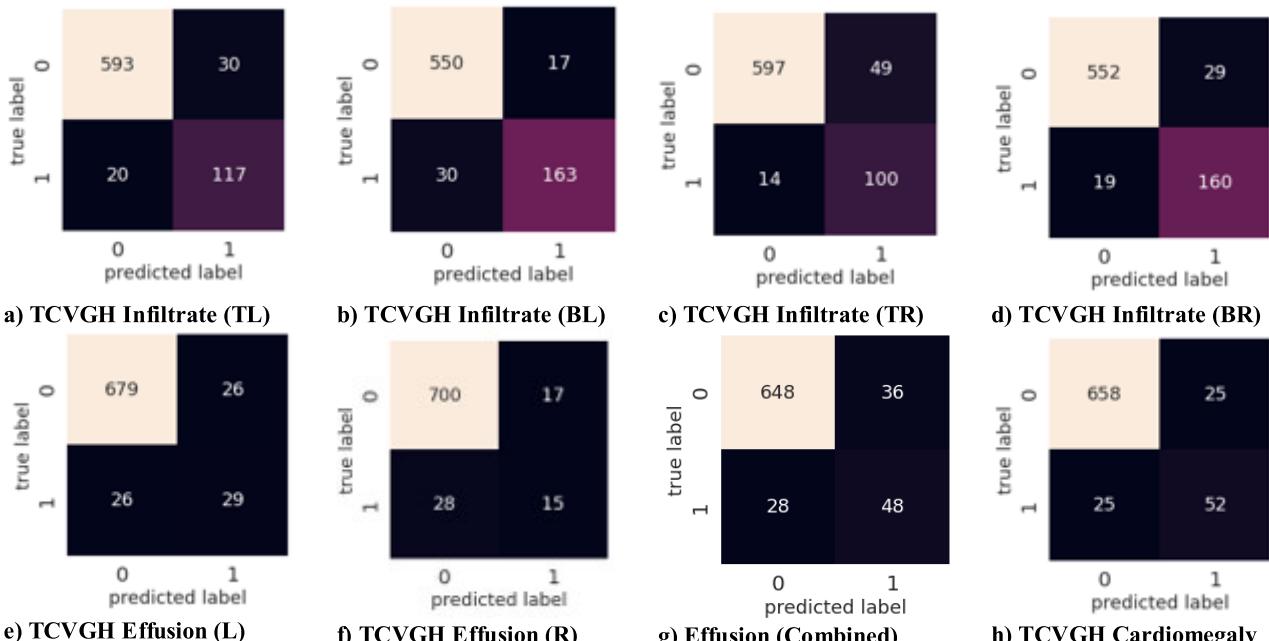
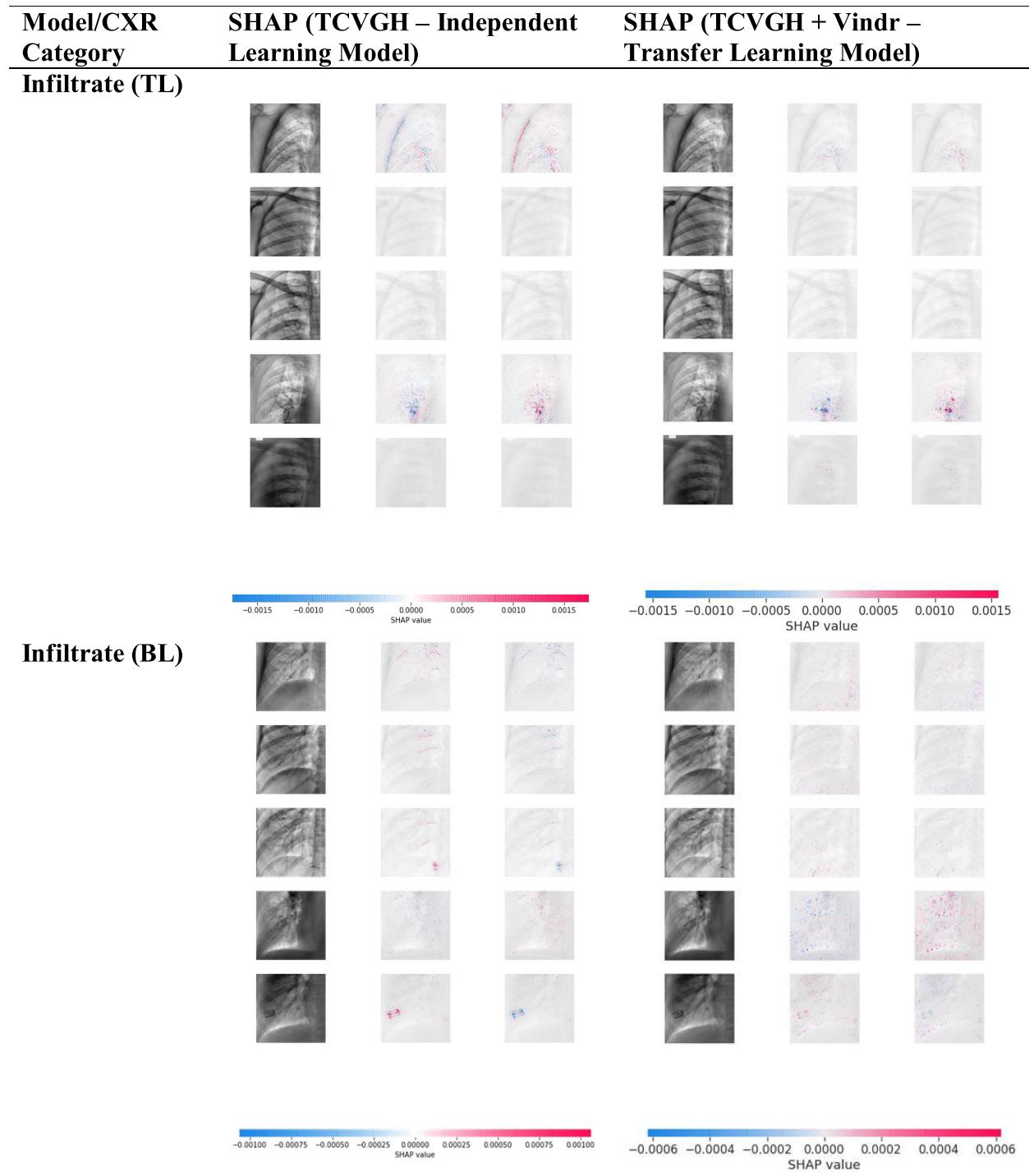
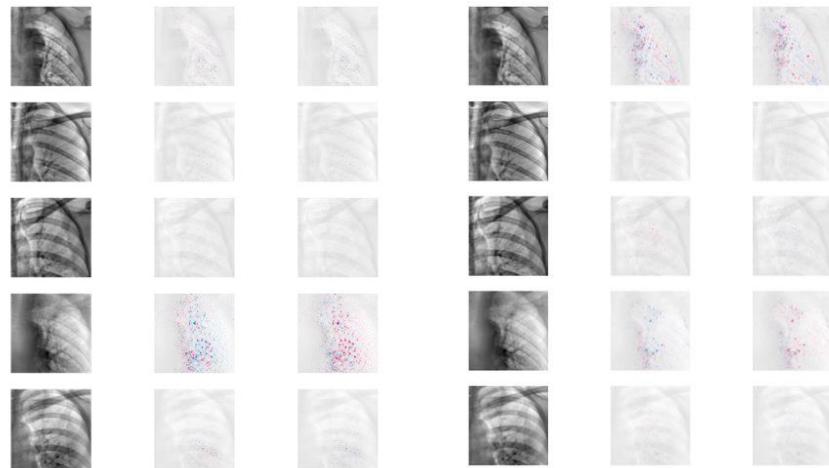


FIGURE 15. Confusion matrix for TCVGH + VinDr evaluation with transfer learning and NIH pre-training.

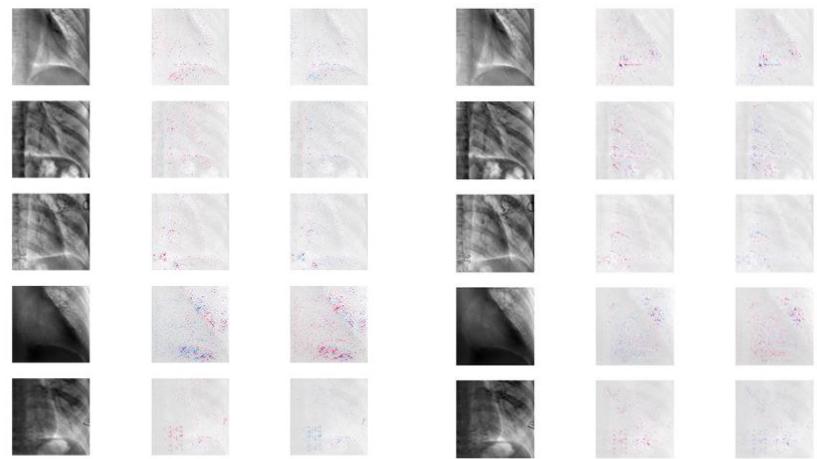
**FIGURE 16.** SHAP scores for the different CXR categories.

presents the data preprocessing model that can standardize the data, image segmentation and data augmentation to achieve better outcomes. The DNN model presented within this work is interpretable using SHAP and provides high transparency to the end user. The XAI classification algorithm provides deep insights within the CXR data processing.

Also the transfer learning in this work, can train and test on the open datasets to provide better understanding of the DCNN model behavior. In the final stage, an interpretable DCNN model is constructed to deal with the diagnosis classification for the pneumonia patient. The infected patient's severity can be analyzed, determined and classi-

Infiltrate (TR)

-0.0006 -0.0004 -0.0002 0.0000 0.0002 0.0004 0.0006
SHAP value

Infiltrate (BR)

-0.0015 -0.0010 -0.0005 0.0000 0.0005 0.0010 0.0015
SHAP value

-0.0008 -0.0006 -0.0004 -0.0002 0.0000 0.0002 0.0004 0.0006 0.0008
SHAP value

-0.0010 -0.0005 0.0000 0.0005 0.0010
SHAP value

FIGURE 16. (Continued.) SHAP scores for the different CXR categories.

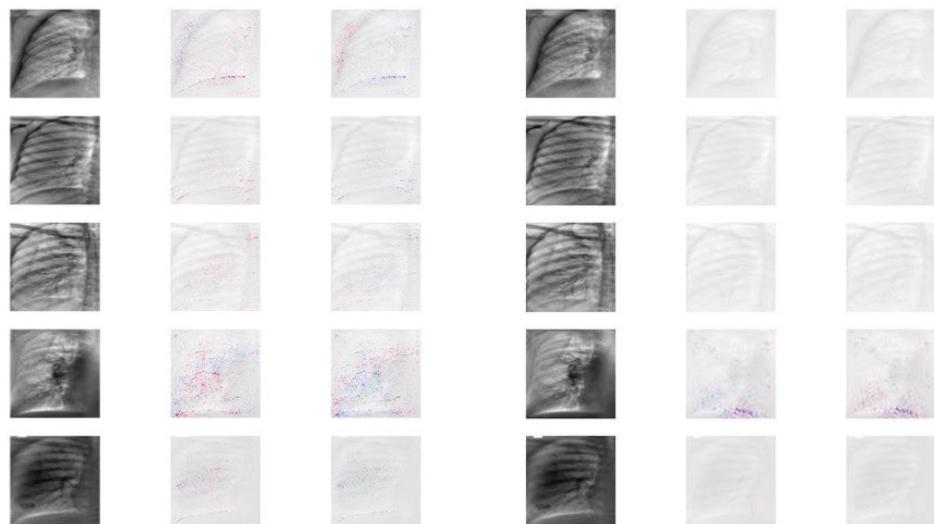
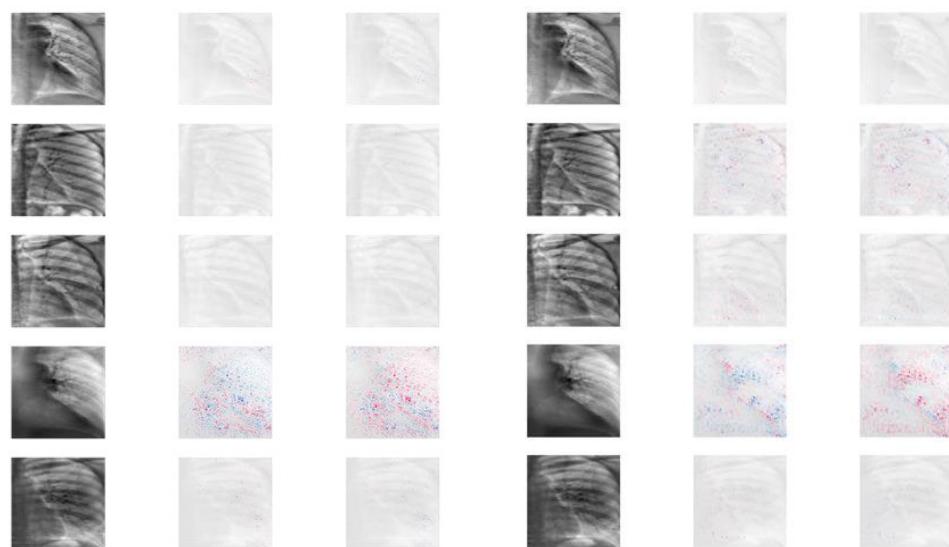
fied accurately. The method DCNN with TCVGH + VinDr dataset provides the accuracy of 92.14% and AUC of 93.57%, which surpasses the CNN models. The method of DCNN with transfer learning on TCVGH + VinDr dataset provides the accuracy of 93.29% and AUC of 95.43%. Therefore, a suitable re-configurable model design is approached to solve the pneumonia infection problem. The adaptation to multiple open datasets by the present DCNN model provides phenomenal improvements for the classification. In the future work, we have planned to add responsible AI by design. The motivation comes for handling the high risk areas and apply the privacy by design within each com-

ponent of the system architecture. Thus, the model will be free from any perturbation attack and fake training data.

APPENDIX A

Figure 16 presents the SHAP based preprocessing analysis in detail. The 7 CXR sections are presented for 4 patient's each, evaluated by independent learning (TCGH) and transfer learning (TCVGH + Vindr).

It can be noticed that the transfer learning section highlights more features for pneumonia infection than independent learning model.

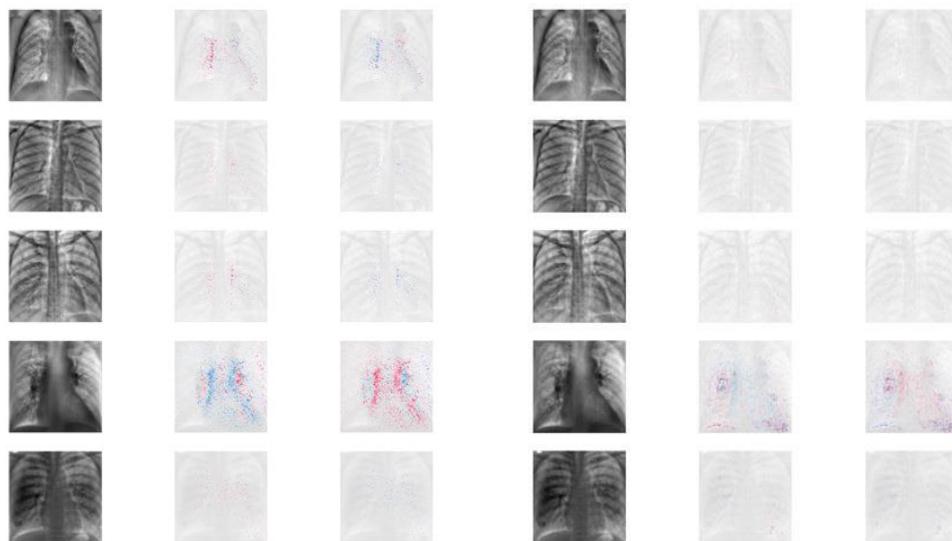
Effusion (L)**Effusion (R)****FIGURE 16. (Continued.) SHAP scores for the different CXR categories.****APPENDIX B****A. XAI-ICP NETWORK MODEL**

The Figure 17 presents the AI components within the XAI-ICP system model. The network model is first pre-trained with the NIH open dataset. Next, the input is taken from the medical examiner/user in the unknown size (#). The input CXR image is then segmented for identifying the chest thoracic(rib) cage within the CXR image, resize it

to fixed shape and divide it in the different portions for the preprocessing. The standard scalar then scales images and perform augmentation. Finally, the DCNN model consisting of DenseNet121 and interpretable layers are added for infection classification.

The XAI-ICP network parameters given in Table 9. The parameters include ADAM (adaptive moment estimation) with initial learning rate of 10^{-5} learning rate was reduced

Cardiomegaly



Effusion (Whole picture)

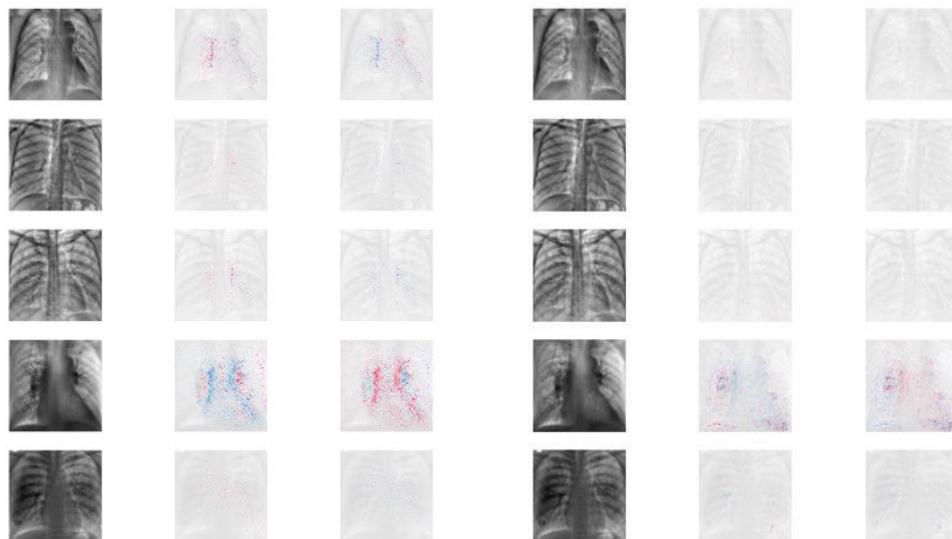
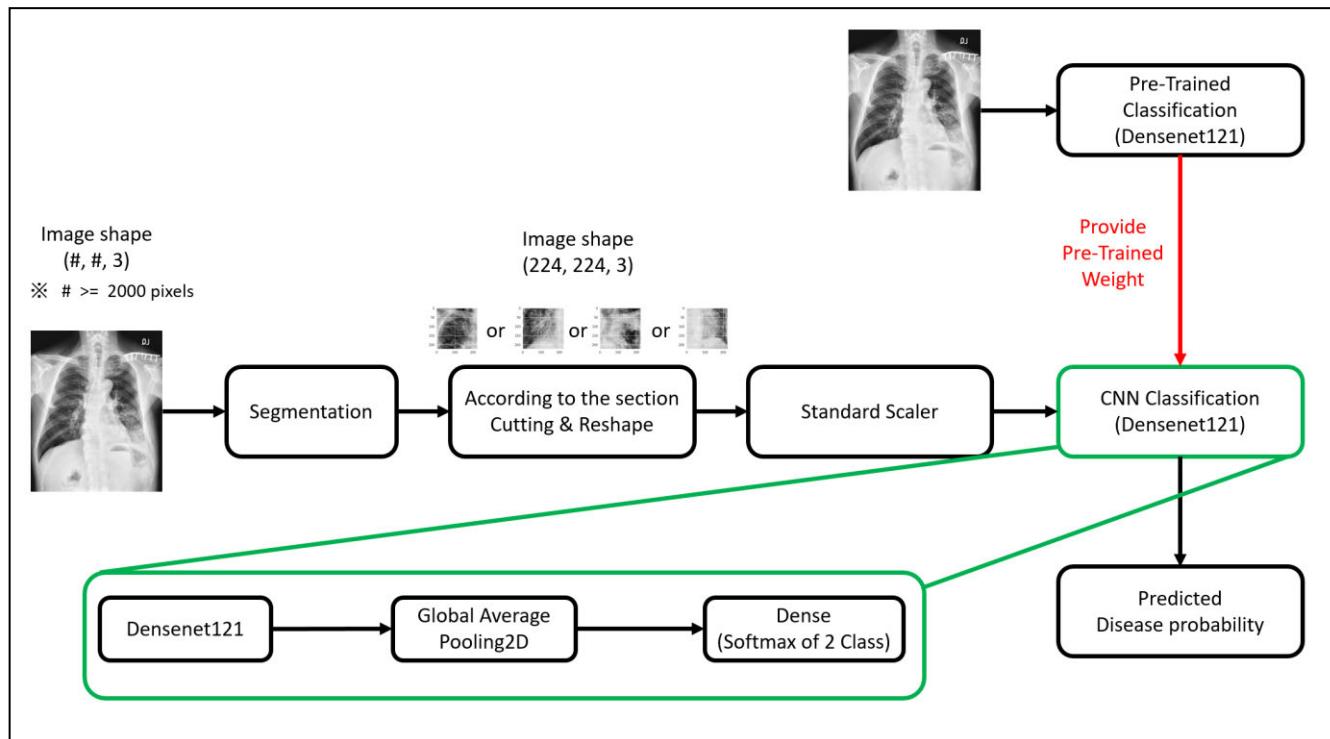


FIGURE 16. (Continued.) SHAP scores for the different CXR categories.

by a factor of 10 if the validation loss did not improve for 2 consecutive epochs. The early stop is 15, if the validation

loss did not improve for 15 consecutive epochs then model will stop training.

**FIGURE 17.** XAI-ICP network model.**TABLE 9.** XAI-ICP network parameters.

Network Parameters/ Model	Pre-trained Model	XAI-ICP Model
Loss	Binary cross-entropy	Binary cross-entropy
Optimizer	ADAM	ADAM
Initial Learning Rate	10^{-5}	10^{-5}
Batch Size	32	40
Epoch	22	50
Early Stop	N.A.	15

ACRONYMS

- International Classification of Diseases (ICD),
- Intensive Care Unit (ICU)
- Machine Learning (ML), Deep Learning (DL),
- Gradient-weighted Class Activation Mapping (Grad-CAM)
- A customized deep CNN termed as (CovNNet)
- International Skin Imaging Collaboration (ISIC)
- Area Under the Receiver Operating Characteristic (AUROC).
- Convolutional Neural Network (CNN)
- Visual Geometry Group (VGG)
- BiLingual Evaluation Understudy (BLEU)
- Deep Convolutional Neural Network (DCNN)
- Dense Convolutional Network (DenseNet121)
- Residual Neural Network (ResNet152)
- Area Under the Curve (AUC)
- Receiver Operating Characteristic (ROC)

ACKNOWLEDGMENT

The authors would like to thank AI center for providing necessary resources for this work. Also, several discussions with medical examiner's of TCVGH hospital is appreciable for setting the experiments criteria and improvements within this work.

REFERENCES

- [1] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England J. Med.*, vol. 380, no. 14, pp. 1347–1358, 2019.
- [2] J. Yanase and E. Triantaphyllou, "A systematic survey of computer-aided diagnosis in medicine: Past and present developments," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112821.
- [3] B. Sahiner, A. Pezeshk, L. M. Hadjiiski, X. Wang, K. Drukker, K. H. Cha, R. M. Summers, and M. L. Giger, "Deep learning in medical imaging and radiation therapy," *Med. Phys.*, vol. 46, no. 1, pp. e1–e36, Jan. 2019.
- [4] A. Singh, S. Sengupta, and V. Lakshminarayanan, "Explainable deep learning models in medical image analysis," 2020, *arXiv:2005.13799*. Accessed: May 11, 2021.
- [5] R. Shad, J. P. Cunningham, E. A. Ashley, C. P. Langlotz, and W. Hiesinger, "Medical imaging and machine learning," 2021, *arXiv:2103.01938*. Accessed: May 11, 2021.
- [6] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Mach. Intell.*, vol. 1, no. 5, pp. 206–215, 2019.
- [7] C. Rudin and J. Radin, "Why are we using black box models in AI when we don't need to? A lesson from an explainable AI competition," *Harvard Data Sci. Rev.*, vol. 1, no. 2, p. 1162, 2019.
- [8] O. F. Ahmad, A. S. Soares, E. Mazomenos, P. Brando, R. Vega, E. Seward, D. Stoyanov, M. Chand, and L. B. Lovat, "Artificial intelligence and computer-aided diagnosis in colonoscopy: Current evidence and future directions," *Lancet Gastroenterol. Hepatol.*, vol. 4, no. 1, pp. 71–80, Jan. 2019.
- [9] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, R. Wang, H. Zhao, Y. Chong, J. Shen, Y. Zha, and Y. Yang, "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 6, pp. 2775–2780, Nov. 2021.

- [10] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022.
- [11] F. K. Došilovic, M. Brčić, and N. Hlupic, "Explainable artificial intelligence: A survey," in *Proc. 41st Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO)*, 2018, pp. 210–215.
- [12] N. Petrick et al., "Evaluation of computer-aided detection and diagnosis systems," *Med. Phys.*, vol. 40, no. 8, 2013, Art. no. 087001.
- [13] L. Wu, R. Huang, I. V. Tetko, Z. Xia, J. Xu, and W. Tong, "Trade-off predictivity and explainability for machine-learning powered predictive toxicology: An in-depth investigation with Tox21 data sets," *Chem. Res. Toxicol.*, vol. 34, no. 2, pp. 541–549, Feb. 2021.
- [14] L. Nanni, S. Ghidoni, and S. Brahnam, "Handcrafted vs. non-handcrafted features for computer vision classification," *Pattern Recognit.*, vol. 71, pp. 158–172, Nov. 2017.
- [15] J. Zhang, Y. Xia, Y. Xie, M. Fulham, and D. D. Feng, "Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 5, pp. 1521–1530, Sep. 2018.
- [16] J. D. Fuhrman, N. Gorre, Q. Hu, H. Li, I. El Naqa, and M. L. Giger, "A review of explainable and interpretable AI with applications in COVID-19 imaging," *Med. Phys.*, vol. 49, no. 1, pp. 1–14, Jan. 2022.
- [17] G. Yang, Q. Ye, and J. Xia, "Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond," *Inf. Fusion*, vol. 77, pp. 29–52, Jan. 2022.
- [18] J. Tang, J. Liu, M. Zhang, and Q. Mei, "Visualizing large-scale and high-dimensional data," in *Proc. 25th Int. Conf. World Wide Web*, Apr. 2016, pp. 287–297.
- [19] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," 2013, *arXiv:1312.6034*.
- [20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "GRAD-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 618–626.
- [21] J. P. Cohen, L. Dao, K. Roth, P. Morrison, Y. Bengio, A. F. Abbasi, B. Shen, H. K. Mahsa, M. Ghassemi, H. Li, and T. Duong, "Predicting COVID-19 pneumonia severity on chest X-ray with deep learning," *Cureus*, vol. 12, no. 7, Jul. 2020, Art. no. e9448.
- [22] Y. Oh, S. Park, and J. C. Ye, "Deep learning COVID-19 features on CXR using limited training data sets," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2688–2700, Aug. 2020.
- [23] L. Zou, H. L. Goh, C. J. Y. Liew, J. L. Quah, G. T. Gu, J. J. Chew, M. P. Kumar, C. G. L. Ang, and A. Ta, "Ensemble image explainable AI (XAI) algorithm for severe community-acquired pneumonia and COVID-19 respiratory infections," *IEEE Trans. Artif. Intell.*, early access, Feb. 25, 2022, doi: [10.1109/TAI.2022.3153754](https://doi.org/10.1109/TAI.2022.3153754).
- [24] T. Nagaoka, T. Kozuka, T. Yamada, H. Habe, M. Nemoto, M. Tada, K. Abe, H. Handa, H. Yoshida, K. Ishii, and Y. Kimura, "A deep learning system to diagnose COVID-19 pneumonia using masked lung CT images to avoid AI-generated COVID-19 diagnoses that include data outside the lungs," *Adv. Biomed. Eng.*, vol. 11, pp. 76–86, 2022.
- [25] 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.
- [26] B. Hu, B. Vasu, and A. Hoogs, "X-MIR: Explainable medical image retrieval," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2022, pp. 440–450.
- [27] Z. Tang, K. V. Chuang, C. DeCarli, L.-W. Jin, L. Beckett, M. J. Keiser, and B. N. Dugger, "Interpretable classification of Alzheimer's disease pathologies with a convolutional neural network pipeline," *Nature Commun.*, vol. 10, no. 1, p. 2173, Dec. 2019.
- [28] G. Zhao, B. Zhou, K. Wang, R. Jiang, and M. Xu, "Respond-CAM: Analyzing deep models for 3D imaging data by visualizations," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Cham, Switzerland: Springer, 2018, pp. 485–492.
- [29] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.
- [30] S. Lapuschkin, S. Wäldchen, A. Binder, G. Montavon, W. Samek, and K. R. Müller, "Unmasking clever hans predictors and assessing what machines really learn," *Nature Commun.*, vol. 10, no. 1, p. 1096, Dec. 2019.
- [31] W. Samek, G. Montavon, A. Binder, S. Lapuschkin, and K.-R. Müller, "Interpreting the predictions of complex ML models by layer-wise relevance propagation," Nov. 2016, *arXiv:1611.08191*.
- [32] A. W. Thomas, H. R. Heekeren, K.-R. Müller, and W. Samek, "Analyzing neuroimaging data through recurrent deep learning models," *Frontiers Neurosci.*, vol. 13, p. 1321, Dec. 2019.
- [33] L. Arras, F. Horn, G. Montavon, K.-R. Müller, and W. Samek, "What is relevant in a text document?: An interpretable machine learning approach," *CoRR*, vol. abs/1612.07843, pp. 1–17, Dec. 2016.
- [34] L. Hiley, A. Preece, Y. Hicks, S. Chakraborty, P. Gurram, and R. Tomsett, "Explaining motion relevance for activity recognition in video deep learning models," Mar. 2020, *arXiv:2003.14285*.
- [35] O. Eberle, J. Buttner, F. Krautli, K.-R. Müller, M. Vallorani, and G. Montavon, "Building and interpreting deep similarity models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 3, pp. 1149–1161, Mar. 2022.
- [36] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?: Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, Aug. 2016, pp. 1135–1144, doi: [10.1145/2939672.2939778](https://doi.org/10.1145/2939672.2939778).
- [37] H. Lakkaraju, E. Kamar, R. Caruana, and J. Leskovec, "Faithful and customizable explanations of black box models," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc.* New York, NY, USA: Association Computing Machinery, Jan. 2019, pp. 131–138.
- [38] P. Zhu and M. Ogino, "Guideline-based additive explanation for computer-aided diagnosis of lung nodules," in *Interpretability of Machine Intelligence in Medical Image Computing and Multimodal Learning for Clinical Decision Support*, K. Suzuki, Eds. Cham, Switzerland: Springer, 2019, pp. 39–47.
- [39] A. M. Nguyen, J. Yosinski, and J. Clune, "Multifaceted feature visualization: Uncovering the different types of features learned by each neuron in deep neural networks," *CoRR*, vol. abs/1602.03616, pp. 3–4, May 2016.
- [40] Y. Yan, J. Zhu, M. Duda, E. Solarz, C. Sripada, and D. Koutra, "GroupINN: Grouping-based interpretable neural network for classification of limited, noisy brain data," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA: Association for Computing Machinery, Jul. 2019, pp. 772–782.
- [41] B. Kim, M. Wattberg, J. Gilmer, C. Cai, J. Wexler, F. Viegas, and R. Sayres, "Interpretability beyond feature attribution: Quantitative testing with concept activation vectors (TCAV)," in *Proc. ICML*, vol. 80, J. G. Dy and A. Krause, Eds., 2018, pp. 2673–2682.
- [42] M. Graziani, V. Andreczky, and H. Müller, "Regression concept vectors for bidirectional explanations in histopathology," in *Understanding and Interpreting Machine Learning in Medical Image Computing Applications*, D. Stoyanov, Eds. Cham, Switzerland: Springer, 2018, pp. 124–132.
- [43] A. Ghorbani, J. Wexler, J. Y. Zou, and B. Kim, "Towards automatic concept-based explanations," in *Advances in Neural Information Processing Systems*, vol. 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. D'Alché-Buc, E. Fox, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2019, pp. 9277–9286.
- [44] Y. Qin, K. Kamnitsas, S. Ancha, J. Nanavati, G. Cottrell, A. Criminisi, and A. Nori, "Autofocus layer for semantic segmentation," *CoRR*, vol. abs/1805.08403, pp. 1–8, Jun. 2018.
- [45] R. Fong and A. Vedaldi, "Interpretable explanations of black boxes by meaningful perturbation," *CoRR*, vol. abs/1704.03296, pp. 1–8, Jan. 2017.
- [46] G. Alain and Y. Bengio, "Understanding intermediate layers using linear classifier probes," Nov. 2016, *arXiv:1610.01644*.
- [47] E. Varol, A. Sotiras, K. Zeng, and C. Davatzikos, "Generative discriminative models for multivariate inference and statistical mapping in medical imaging," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Cham, Switzerland: Springer, 2018, pp. 540–548.
- [48] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad, "Intelligible models for HealthCare: Predicting pneumonia risk and hospital 30-day readmission," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*. New York, NY, USA: Association Computing Machinery, Aug. 2015, pp. 1721–1730.
- [49] S. P. Praveen, P. N. Srinivasu, J. Shafi, M. Woźniak, and M. F. Ijaz, "ResNet-32 and FastAI for diagnoses of ductal carcinoma from 2D tissue slides," *Sci. Rep.*, vol. 12, no. 1, Dec. 2022, Art. no. 20804, doi: [10.1038/s41598-022-25089-2](https://doi.org/10.1038/s41598-022-25089-2).
- [50] M. Woźniak, J. Siłka, and M. Wieczorek, "Deep neural network correlation learning mechanism for CT brain tumor detection," *Neural Comput. Appl.*, pp. 1–6, Mar. 2021. [Online]. Available: <https://link.springer.com/article/10.1007/s00521-021-05841-x#citeas>

- [51] D. Schneeberger, K. Stöger, and A. Holzinger, "The European legal framework for medical AI," in *Proc. Int. Cross-Domain Conf. Mach. Learn. Knowl. Extraction*. Cham, Switzerland: Springer, Aug. 2020, pp. 209–226.
- [52] R. Srinivasan and A. Chander, "Explanation perspectives from the cognitive sciences—A survey," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, Jul. 2020, pp. 4812–4818.
- [53] H. Q. Nguyen et al., "VinDr-CXR: An open dataset of chest X-rays with radiologist's annotations," *Sci. Data*, vol. 9, no. 1, p. 429, Jul. 2022, doi: [10.1038/s41597-022-01498-w](https://doi.org/10.1038/s41597-022-01498-w).
- [54] N. Hasan, Y. Bao, A. Shawon, and Y. Huang, "DenseNet convolutional neural networks application for predicting COVID-19 using CT image," *Social Netw. Comput. Sci.*, vol. 2, no. 5, p. 389, Sep. 2021, doi: [10.1007/s42979-021-00782-7](https://doi.org/10.1007/s42979-021-00782-7).
- [55] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191).
- [56] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," Dec. 2017, *arXiv:1705.02315*.
- [57] R.-K. Sheu and M. S. Pardeshi, "A survey on medical explainable AI (XAI): Recent progress, explainability approach, human interaction and scoring system," *Sensors*, vol. 22, no. 20, p. 8068, Oct. 2022, doi: [10.3390/s2208068](https://doi.org/10.3390/s2208068).
- [58] D. Demner-Fushman, M. D. Kohli, M. B. Rosenman, S. E. Shooshan, L. Rodriguez, S. Antani, G. R. Thoma, and C. J. McDonald, "Preparing a collection of radiology examinations for distribution and retrieval," *J. Amer. Med. Inform. Assoc.*, vol. 23, no. 2, pp. 304–310, Mar. 2016.



KAI-CHIH PAI received the Ph.D. degree from the Graduate Institute of Educational Information and Measurement, National Taichung University of Education, in May 2019. He was a Researcher with the National Taichung University of Education, from 2015 to 2018. He has been an Assistant Professor with the College of Engineering, Tunghai University, Taiwan, since 2019. His current research interests include data analytics, cloud computing, intelligent learning environments, and educational technology.



LUN-CHI CHEN received the M.S. degree in computer science from Tunghai University, Taiwan, in 2005, and the Ph.D. degree from the National Chung Hsing University, in 2015. Since 2006, he has been with the National Center for High-Performance Computing, Taiwan, where he was an Associate Researcher, from 2007 to 2018. He has been an Associate Professor with the College of Engineering, Tunghai University, since 2018. His current research interests include data analytics, deep learning, text mining, and cloud computing.



RUEY-KAI SHEU received the M.S. and Ph.D. degrees from the National Chiao Tung University, in 1998 and 2001, respectively. He joined W&Jsoft Inc., as a Research and Development Leader, from 1999 to 2007. He has been with the Department of Computer Science, Tunghai University, Taichung, Taiwan, since 2014, where he is currently an Associate Professor. His current research interests include the landing of machine learning or artificial intelligence applications, cloud systems, and data leak protection technologies.



CHIEH-LIANG WU received the M.D. and Ph.D. degrees. He has been with the Department of Automatic Control Engineering, Feng Chia University, Taichung, Taiwan, and the Department of Industrial Engineering and Enterprise Information, Tunghai University, Taichung. He is currently with the Department of Critical Care Medicine, Taichung Veterans General Hospital, Taichung.



MAYURESH SUNIL PARDESHI received the B.E. degree in information technology from Mumbai University, in 2010, the M.Tech. degree in computer science and engineering (CSE) from the Walchand College of Engineering, Sangli, India, in 2013, and the Ph.D. degree in electrical engineering and computer science (EECS-IGP) from the National Chiao Tung University, Hsinchu, Taiwan. He is currently a Researcher with the AI Center. His current research interests include artificial intelligence and security in distributed systems. He is serving as a reviewer and a sub-reviewer for many international conferences/journals.



WEI-CHENG CHEN received the degree in computer science from Tunghai University, Taiwan, where he is currently pursuing the master's degree with the Department of Computer Science. His research interests include computer vision and deep learning.

• • •