Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 1915–1925

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

An Explainable AI driven Decision Support System for COVID-19 Diagnosis using Fused Classification and Segmentation

Niranjan K[a,*], Shankar Kumar S[a,*], Vedanth S[a,*], Dr. S. Chitrakala[a]

[a]*Computer Science and Engineering, College of Engineering Guindy, Anna University, Chennai, India*

Abstract

The coronavirus has caused havoc on billions of people worldwide. The Reverse Transcription Polymerase Chain Reaction(RT-PCR) test is widely accepted as a standard diagnostic tool for detecting infection, however, the severity of infection can't be measured accurately with RT-PCR results. Chest CT Scans of infected patients can manifest the presence of lesions with high sensitivity. During the pandemic, there is a dearth of competent doctors to examine chest CT images. Therefore, a Guided Gradcam based Explainable Classification and Segmentation system (GGECS) which is a real-time explainable classification and lesion identification decision support system is proposed in this work. The classification model used in the proposed GGECS system is inspired by Res2Net. Explainable AI techniques like GradCam and Guided GradCam are used to demystify Convolutional Neural Networks (CNNs). These explainable systems can assist in localizing the regions in the CT scan that contribute significantly to the system's prediction. The segmentation model can further reliably localize infected regions. The segmentation model is a fusion between the VGG-16 and the classification network. The proposed classification model in GGECS obtains an overall accuracy of 98.51 % and the segmentation model achieves an IoU score of 0.595.

Keywords: COVID-19; Explainable AI; Segmentation; Convolutional Neural networks; Class Activavtion Mapping; Gradient weighted Class Activation Mapping (Grad-CAM); Guided Grad-CAM.

1. Introduction

The coronavirus was was initially discovered in December 2019 in Wuhan, China. Since then, the virus has rapidly spread around the world, causing a global pandemic (COVID-19 Pandemic). Over 500 million Covid-19 instances have been documented globally as of April 2022, and more than 6 million of those cases had resulted in fatalities. The SARS-CoV-2 virus can spread through direct, indirect, and close contact with an infected person. Sneezing, coughing and speaking can result in the generation of aerosols which act as a medium for airborne transmission of the virus.

* Equal contribution to this work
E-mail address: niranjan0246@gmail.com

10.1016/j.procs.2023.01.168

Due to the high density of population in many countries and the ease of transmission of the virus, several countries have enforced nationwide lockdowns. This has subsequently had an impact on the nation's economy and people's quality of life.

The frequently utilised diagnostic procedure for COVID-19 that is the RT-PCR test. In order to gather samples for testing, the test involves inserting a 6-inch swab deeply into the nose. For some this method of testing can be highly uncomfortable. Moreover the RT-PCR test's sensitivity is insufficient to effectively prevent the pandemic. To guarantee that there is enough viral material to identify, the RT-PCR test should be performed 8 days following exposure or a suspected infection. However this 8-day period could prove to be costly since the virus is most infectious 2 days before the development symptoms.

In previous works, attempts have been made to identify COVID-19 from cough sounds [28]. Several machine learning approaches have also been explored which try to predict COVID-19 based on the symptoms shown by the patient [31]. However, the effectiveness of these methods is lower when compared to the work done using medical radiology imaging. Along with the RT-PCR test, CT imaging is a useful diagnostic technique for detecting COVID-19 infection. This is because the CT scans of coronavirus infected patients indicate some form of lung infection. Medical radiology imaging has been widely used for detecting various chest diseases such as Pneumonia, Cardiomegaly, Effusion and Atelectasis [18][1]. Despite the fact that chest X-rays are more accessible and less expensive, chest CT scans capture finer details, allowing for a more accurate diagnosis. Given the better resolution compared to X-Ray, CT scans also seem appropriate for a robust-scale screening.

CT scan scans are be capable of detecting COVID-19 infection in its early phases with very high sensitivity. The only drawback to chest CT-based diagnosis is the diagnostic time: radiologists need around 22 minutes to review each patient's data. In this case, artificial intelligence might be essential in automating the entire diagnostic procedure, saving radiologists' time and effort on visual inspection.

When it comes to the medical domain, the capacity to explain results is critical. The medical field frequently affects human life, hence any decision made must be supported by proper facts and rationale. Earlier works which deal with COVID-19 classification using medical images lack interpretability and therefore are less reliable. The proposed GGECS system makes use of GradCam and Guided GradCam to provide explainable COVID-19 classification results.

Our contributions, in summary are as follows:

- An end-to-end Covid-19 decision support system, called GGECS (Guided Gradcam based Explainable Classification and Segmentation) system is developed to perform interpretable classification and precise segmentation to help diagnose Covid-19 from CT Scans.
- A pixel space gradient visualization method called Guided GradCam which can effectively localize relevant image regions with fine grained importance is used.
- Further, Histogram equalization is applied to enhance the quality of CT scans by highlighting the lung boundaries.

2. Related Works

The most important step in building an effective diagnosis system is to quickly and accurately identify infected patients. Before the advent of COVID-19 pandemic, there existed a variety of systems for early identification and prevention of kidney disease [27]. Handling huge volumes of medical image data has been discussed in [20]. Several medical imaging systems have been developed to identify diseases such as pneumonia in chest radiology [15]. However, these systems as such cannot be used to identify Covid-19. [16] proposes a DarkCovidNet model for COVID-19 Detection using chest X-rays. [2] introduces a dataset composed of COVID-19 infected CT scans along with healthy and pneumonia infected scans. The DarkCovidNet architecture is inspired by the DarkNet architecture. The proposed model consists of 17 Convolution layers having different filtering on each layer. On binary classes, the system generates a classification accuracy of 98.08. [29] on the other hand performs three different binary classifications with four classes namely healthy, COVID-19, bacterial and viral pneumonia. In their work, five previously trained convolutional neural networks were re-trained on 7406 chest X-ray images. However, only 127 X-ray images in the dataset are diagnosed with COVID-19 thus making the proposed system less robust.

Other works have employed network capsules instead of CNNs to tackle the limitation of such network that requires a vast quantity of input and parameters. The authors of [3] chose this strategy. A capsule network to identify the presence of COVID-19 strains in CXR images was developed with an accuracy of 95.7%. The results were contrasted to those of Sethy et al. [25], who used SVMs to develop a model based on Resnet-50 and achieved 95.38% performance, 97.29% sensitivity, and 93.47% specificity. Although these results are promising, the proposed models were trained on X-Ray images. The low resolution in X-ray images makes it difficult to identify COVID-19 during the early stages of infection.

Image segmentation is the process of generating a pixel-wise mask of a particular object or several objects in an image. Image segmentation has been widely adopted and used in the medical domain. The U-Net [22] is one such image segmentation architecture which gained popularity for its effectiveness in performing segmentation on CXR and CT scans. A highly effective system for gland segmentation in cancer histopathological images has been proposed in [19]. Further development in the field resulted in 3D Unet [9] which performs segmentation on 3D volumes. The segmentation of infected regions in medical images for COVID-19 can help medical experts better understand the virus and the extent of infection. [9] proposes a DenseNet based Unet architecture to perform COVID-19 lesion segmentation on X-rays. Works such as Inf-Net [5] perform lesion segmentation on CT Scans.

Recent works combine the complex process of classification and segmentation into a single pipeline. [17] proposes an AI system to first perform lung segmentation using U-Net architecture and secondly lesion segmentation using DenseNet. The Dual branch combination network (DCN) proposed by [6] is able to perform subtle lesion detection. The DCN makes use of a segmentation branch based on U-net and a classification branch based on ResNet-50 with four residual blocks.

Explainable Artificial Intelligence (EAI) has lately been investigated because of its capacity to shed light on the behaviour and thought processes behind some difficult machine learning problems [7]. Decision trees and linear models can be used to describe models in a form that humans can understand and interpret, according to a number of studies [8]. [21] presented Model-agnostic Local Interpretable Explanations (LIME). This unique explanation method may explain any classifier's predictions in a comprehensible manner. Any deep learning model used for image classification may be understood, visualized, and interpreted using LIME [23]. [12] illustrated how explainable-AI can aid with AI/ML deployment in the medical field. Such interpretations of the trained AI models could help support the decision on COVID-19 classification.

One more explainable AI algorithm is the Gradient-weighted Class Activation Mapping (Grad-CAM). It is a generalization of CAM, which was introduced in [24]. Grad-CAM may provide post hoc local explanations with any kind of CNN architecture. Guided GradCam which is a fine grained image localization technique was also introduced by [24]. In Guided GradCam the visualizations generated by GradCam and GuidedBackprop are multiplied pointwise. [11] applies GradCam to a classifier detecting lymph nodes from medical images.

3. Methodology

The proposed GGECS is a diagnostic system for COVID-19 that combines the advantages of explainable classification and segmentation for better applicability. The system greatly simplifies and speeds up the diagnostic process for radiologists and other physicians by providing explainable classification findings along with related lesion segmentation. Fig. 1 depicts the block diagram of the proposed GGECS system.

3.1. Data Augmentation

Deep learning models require a large number of instances to avoid overfitting due to their complexity. However, data is scarce for a majority of the real-world challenges. In the case of medical data, things are even more complicated. Patient health data is protected by data laws in several countries, and most hospitals are reluctant to share their patient health data due to privacy concerns. Adding to this, most publicly available medical datasets often lack annotations. In this study, these issues are tackled by leveraging data augmentation. The process of data augmentation entails modifying images to increase the number of training examples while preserving semantic information. Three transformations are used on the training samples in this study namely: right and left flip, zooming. Such changes maintain the visuals and thus remain interpretable by a physician.

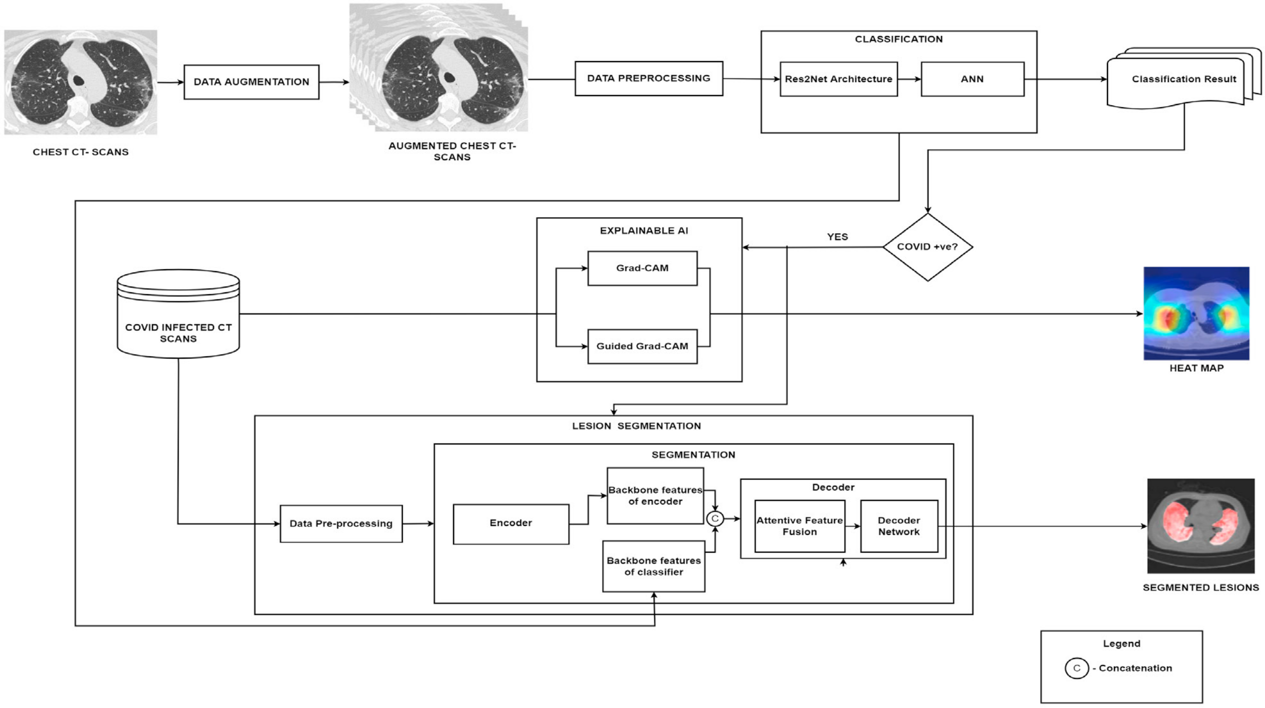

Fig. 1: GGECS architecture

3.2. Data Preprocessing

In computer vision applications, preprocessing is a prevalent practice. Preprocessing techniques can be effective for reducing undesired noise, highlighting parts of the image that can aid in recognition, and even assisting with the deep learning training phase. In this work, four preprocessing techniques are utilized before proceeding with the model training phase, 1) Image resizing 2) Noise removal 3) Image Enhancement 4) Normalization. The input images are resized to 256 x 256 to maintain compatibility with the network architectures. Normalization is an image standardization process by changing the different intensities of each image to a certain range. The noise present in the CT Scan is removed using the Median blur filtering which is a non-linear digital filtering technique. By altering the intensity frequency, the process of enhancement improves the sharpness of the image. This method can draw attention to areas of an image that are not clearly visible. The enhancement technique utilized is Contrast Limited Adaptive Histogram Equalization(CLAHE). Normalization is the process of bringing the varied intensities of each image within a defined range.

3.3. Classification of chest CT scans

The task of classification involves performing binary classification on the CT scans to differentiate between COVID-19 positive and negative patients. The proposed model used for classification is based on the Res2Net architecture. Residual networks offer an excellent balance of performance and parameter count, as well as faster training. Another added advantage of the residual network architecture is its capacity to feed images of different sizes than those with which it was trained. This is an important aspect of the training methodology used to train a high-performance network with a small number of epochs. Res2Net's novelty lies in the use of "4scale -(3x3)" residual, hierarchical architecture. This has enabled the Res2Net architecture to exhibit stronger multi-scale representation ability as compared to ResNet. ResNets in general have been widely used for various image classification tasks and proven to be highly effective [29] [15]. The fully connected layers at the end are removed and four layers of dimensions (20148,1024), (1024,512), (512,128), (128,2) with ReLu activation are added to the classification model.

3.4. Explainable AI for COVID

The diagnosis of COVID-19 through CNN has always been a black box and in order to improve the transparency of the proposed decision support system, activation maps are generated to find the response region of the predicted result. Explainable AI algorithms like CAM, GradCam and Guided GradCam are applied. [30] introduced the Class Activation Mapping (CAM) technique for explaining CNN's. A layer called Global Average Pooling (GAP) has been added to replace the fully connected layers at the end of the architecture. The feature map from the last convolutional layer is reduced from size (H,W) to (1,1) after the GAP layer and hence loses its spatial representation ability. GAP simply averages each feature map's activation and outputs it as a vector. The final softmax loss layer receives a weighted sum of this vector. The response region is computed by a weighted sum of feature maps in the last convolutional layer and weights connected to the firing neuron in the last layer.

Grad-CAM is an enhanced variant of CAM that can generate visual explanations for any CNN, even if it has a series of fully connected layers. The gradients of the firing neuron are backpropagated till the extracted feature map's last layer. A significance score is computed with the help of previously calculated gradients, a response region showing the key component is generated. The significance score is calculated by calculating the gradient of output class c corresponding to the feature map in the last convolutional layer, and then these gradients are averaged across every feature map. The steps are detailed in Algorithm 1.

$$\alpha_c^k = \frac{1}{Z}\sum_i\sum_j \frac{\partial y_c}{\partial A_{ij}^k} \tag{1}$$

Here k represents the index of the feature map in the last convolutional layer, Z is the size of feature maps in the last convolutional layer (H * W). α_c^k represents the importance of 'k'th feature map for the required class 'c'. The computed significance score and feature maps are multiplied for each layer in the final feature map and summed together. Relu non-linearity is applied to the previously calculated sum to filter out only the pixels having positive correlation with the target class.

$$L_{Grad-Cam}^c = ReLu(\sum \alpha_k A^k) \tag{2}$$

$L_{Grad-Cam}^c$ is the coarse grained localization map. Guided GradCam generates localization maps by pointwise multiplication of GradCam and Guided Backpropagation results. In Guided Backpropagation, the negative gradients are suppressed or set to zero while backpropagating through the ReLu layers. The pixels that are detected by the neurons are only captured intuitively. Hence, Guided GradCam produces class discriminative as well as high resolution localization maps. The steps followed to generate Guided GradCam heatmap is detailed in Algorithm 2.

3.5. Segmentation by classification feature fusion

The segmentation's job is to precisely find lesion locations in suspected COVID-19 patients' chest CT scans. This network has a contraction path and an expansion path similar to a standard encoder-decoder network with the only difference being the encoder-decoder blocks connected by skip connections between symmetrical layers. This network's design is made up of a variety of blocks that serve a particular purpose. The flow of input is first through the Contraction Blocks (CB), then through a Feature Enhancement Module (FEM), Expansion Blocks (EB), and finally an Attention Feature Fusion module (AFF), which is responsible for fusion of the encoder and decoder outputs. The encoder model is VGG-16 based and includes five blocks. The encoder is fed with COVID-19 positive CT-scans and a max-pooling function is present at the end of each VGG block which downsizes the feature map by a specified factor. The decoder consists of feature fusion module to fuse feature maps from various stages and generate side outputs. The fusion module emphasises the relevance of the high level feature maps and filters out useful features through attention mechanism from the lower level feature map. The final prediction is based on the last output and is upsampled to the same resolution as the original CT image. The features extracted from the classification model improves the feature representation ability of the segmentation model. The fusion module merges feature maps derived from the encoder and classification model.

3.5.1. Feature Enhancement Module

In order to achieve higher performance, a Feature Enhancement module is included in the proposed encoder to increase its representational capability. After the last layer convolutional layer (conv5), the FEM module is introduced in the last block. It is made up of two Grouped Atrous Modules (GAMs), which extract larger receptive fields and more robust feature maps. The GAM module generates a feature map half the size of the VGG-16 backbone's coarsest feature map. It also improves the feature map given by the last encoder block in terms of representational quality.

The FEM is made up of two GAMs that are connected by a max-pooling function. Initially in the GAM, a 1 x 1 convolution layer expands the feature map's channel. After that, the feature map is divided into four equal groups. Atrous convolution, which is employed in place of trivial group convolution, increases to a greater extent the perceptual range of filters while preserving the same computing cost as normal convolution operation. The Squeeze-Excitation (SE) block in the network is employed to fully utilise valuable features, which entails recalibrating channel-wise convolutional feature responses using the attention mechanism. In the SE block, a sigmoid function is followed by two linear layers. After the SE block, a 1x1 convolution layer is added to cut down the number of output channels almost by two.

3.5.2. Classification Feature Fusion Module

The segmentation model's capacity to represent features is enhanced by the classification model's already learnt features. The objective of the fusion module is to merge the encoder and classification model's features. The segmentation model's encoder features are denoted by $M_1^E, M_2^E, M_3^E, M_4^E, M_5^E$. The Res2Net [40] backbone of the classification model has five stages, and is combined using the last feature maps A_k.

3.5.3. Attentive Feature Fusion

All input feature maps are treated the same by a traditional top-down decoder fusion technique. The features are merged using a special attention based fusion module. The smaller feature map is considered to be more crucial and important in the suggested fusion technique. Using 1x1 convolution layers, the input feature maps M_i^E, M_{i+1}^D are reduced to half their original size. To produce a feature map that is twice as large, bilinear interpolation is applied to M_{i+1}^D. A squeeze excitation block is used to combine both the above mentioned outputs to produce an improved feature map (also used in GAM). The feature map of the two fold up-sampled output from the previous stage is merged with the upgraded feature map. A Squeeze Excitation (SE) block is used to enhance the merged feature map. On the squeeze excitation block'soutput, a 3x3 convolution operation is applied. The feature map is constructed as an intermediate output of the current stage using a 1x1 convolution layer with a neuron.

Algorithm 1: Proposed GGECS

Input : Classification model: classification_model, CT scan: ct_scan, Segmentation model: segmentation_model
Output: Diagnosis Results

$L_{GradCam} \leftarrow 0$;
$L_{GuidedGradCam} \leftarrow 0$;
$L_{Segmentation} \leftarrow 0$;
$y_c \leftarrow classification\_model(testCT)$;
if $y_c = 0$ **then**
  $L_{GradCam} \leftarrow gradCam(ct\_scan)$;
  $L_{GuidedGradCam} \leftarrow guidedGradCam(ct\_scan)$;
  $L_{Segmentation} \leftarrow segmentation\_model(ct\_scan)$;
  return $L_{GradCam}, L_{GuidedGradCam}, L_{Segmentation}$;
else
  return COVID -ve;
end if

4. Experiments and Results

4.1. Dataset

A large CT scan dataset obtained by curating data from 7 public datasets [14] was used for training the model. These datasets have demonstrated their effectiveness in deep learning applications by being publicly used for COVID-19 diagnosis. The dataset contains 6,893 healthy CT scans and 7,593 CT scans from 466 patients. The segmentation dataset used is a subset of the large CT scan dataset consisting of 2729 CT Scan images with their corresponding ground truth masks.

4.2. Experimental Setup

In order for the model to yield better performance, the various hyper-parameters of the network are chosen through experimentation. Training and testing of the models were performed using a Nvidia Tesla P100 GPU along with 32-GB RAM. Regularisation techniques such as early stopping were is deployed to attain the best performance. The proposed deep learning framework has been implemented using Pytorch deep learning library.

For classification, the module consists of a Res2Net architecture, and the fully connected layers at the end are replaced with feed forward layers. Since the CT images are collected from different sources, the images are resized to 224x224 keeping in mind the computational capacity. The cross entropy loss was reduced using a SGD optimizer with learning rate of 0.0001.

For segmentation, the mini-batch size is chosen to be 2 for training. The images are again resized to 256 x 256 in order to fit the segmentation architecture. The backbone of the segmentation model is VGG-16. The optimizer used is Adam, with an initial learning rate of 0.001. The atrous rates in two consecutive GAN's are 1, 3, 6, 9 and 1, 2, 3, 4 respectively. The model was trained for 50 epochs before reaching saturation.

4.3. Performance Evaluation and Results

For classification, the widely used performance metrics are precision and recall. The precision of a class is the proportion of accurately predicted instances to all items that are predicated to fall into that particular class. The formula is given by Eq. 3.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Recall for a class is a ratio of correctly predicted instances among the total number of items actually belonging to that class. The formula is given by Eq. 4.

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

Both precision and recall are very important in estimating the system's robustness and performance. F1 Score provides an average of precision and recall given by Eq. 5. A score near 1 indicates a model with good discriminative ability.

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{5}$$

Accuracy is the ratio of correctly predicted samples to the total number of samples given by Eq. 5.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

TN - True Negatives, TP - True Positives, FN - False Negatives, FP - False Positives

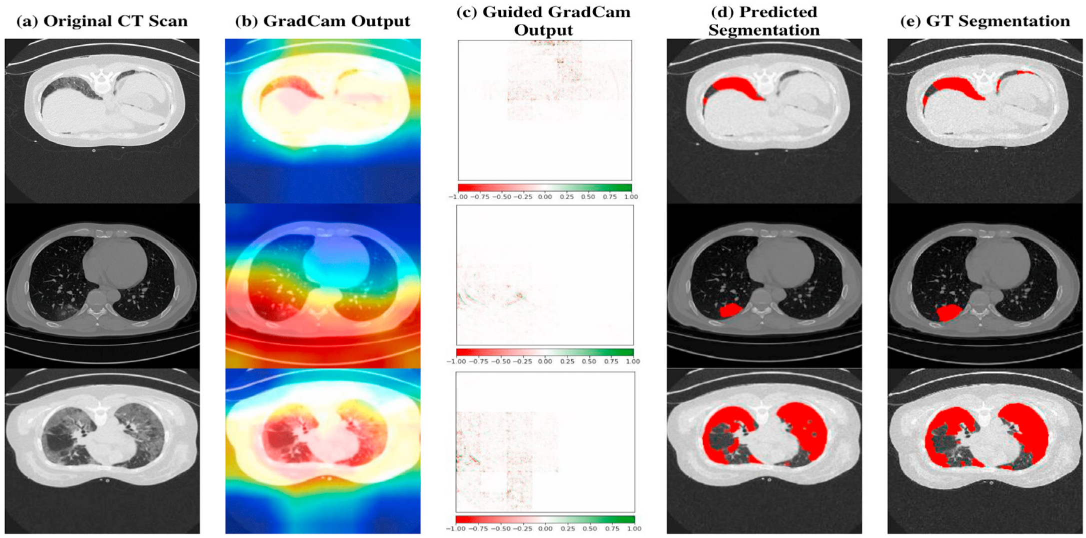

Fig. 2: Results from the proposed GGECS system

The Jaccard index, commonly known as the intersection over union (IoU) measure, is essentially a way to calculate the percentage of overlap between the target mask and the predicted output. IoU is formulated as shown in Eq. 7.

$$IoU = \frac{Area\ of\ Intersection\ of\ two\ boxes}{Area\ of\ Union\ of\ two\ boxes} \tag{7}$$

Fig. 2 shows the GradCAM, Guided GradCAM, and predicted segmentation output generated by the proposed GGECS system along with the ground truth. The classification model was trained first and after extensive experiments on the hyperparametars, a model with a test accuracy of 98.5 % was extracted. GradCAM is used to generate activation mapping on CT images. In the activation maps, the red potion represents the most important features, while the blue region highlights the least important parts. The GradCAM's output is found to be almost consistent with the very accurate lesion ground truth, but the activation map is very coarse-grained and it isn't very specific in localization. Guided GradCAM, obtained by pointwise multiplication of GradCAM and Guided Backprop results, provides fine-grained visualizations. It locates the response region of the prediction in a pixel-wise manner. The results generated by Guided GradCam are localised better compared to GradCam but not actually better than segmentation. The red regions in the predicted segmentation mask represent infected regions. The proposed segmentation model performs very well, as is evident from the predicted and ground truth masks in Fig. 2.

Precision, Recall, and F1 score for both COVID positive and COVID negative classes are produced using the suggested classification model, which is based on Res2Net. The occurrence of good precision and recall values for both classes demonstrates that there is no class imbalance in the dataset and that the model is not biased towards any one class in particular. If the model is skewed in favour of the minority class, the model is deemed to be inaccurate. Since F1-Score is the average of precision and recall, a value of 0.985 for the COVID Positive Class and a value of 0.984 for the COVID Negative Class indicates that the model is well balanced. The classification model achieved a training accuracy of 99.9% and a validation accuracy of 98.5%. Table 1. shows the classification performance.

The proposed Res2Net-based classification model performs better than all the other existing work done on the same dataset. The accuracy of the proposed classification system is compared with other works as shown in Fig. 3. The proposed model achieved a benchmark accuracy of 98.51 %, thus outperforming the performance achieved by the other models by a considerable margin. The proposed GGECS classification model performance is compared against other well knows pre trained CNN models such as ResNet50, Xception and VGG16 models which are trained on Imagenet [13]. Resnet50 has 50 deep layers [10]. The Xception model on the other hand has 71 deep layers

Table 1: Proposed Classification Model Performance

	Precision	Recall	F1 Score
Class 0(COVID +ve)	0.989	0.982	0.985
Class 1(COVID -ve)	0.980	0.988	0.984

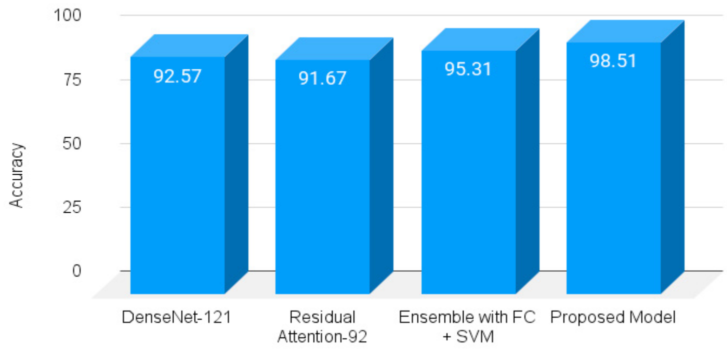

Fig. 3: Classification Model Accuracy Comparison

and proposed by [4]. VGG16 is a 16 layer deep convolutional neural network [26]. Table 2 shows the performance comparison between models. It is observed that the Ensemble with FC architecture has similar precision performance to the proposed GGECS classification model, however the later outperforms the former in both Recall and F1 Score. The proposed GGECS classifation model demonstrates better performance than traditional pretrained CNN models such as Xception, Resnet50 and VGG16.

Table 2: Classification Model Comparison

	Precision	Recall	F1 Score
VGG-16	0.891	0.887	0.889
ResNet50	0.953	0.961	0.957
Xception	0.925	0.931	0.928
DenseNet-121	0.8755	0.9442	0.9085
Residual Attention-92	0.9099	0.9047	0.9073
Ensemble with FC	0.9832	0.8984	0.9389
Ensemble (FC+SVM)	0.908	0.9793	0.9423
Proposed model	0.9852	0.9849	0.985

In this work, two different kinds of segmentation architectures were implemented: a segmentation network that was fused with the classification features obtained through the Res2Net-based classification model and a segmentation network that is purely an encoder-decoder network. The comparison between the two networks can be seen in Table 3. Since the segmentation model uses previously trained classification features, the fused model's IoU score is higher than the unfused model, as the encoder's output is fed to decoder module, these classification features improve the model's lesion localization ability. IoU score comparison between fused and unfused model is shown in Fig, 4.

5. Conclusion

A Guided GradCam-based Explainable Classification and Segmentation decision support system has been developed to aid doctors in their diagnosis of the coronavirus. Using GradCam, the system can detect the presence of

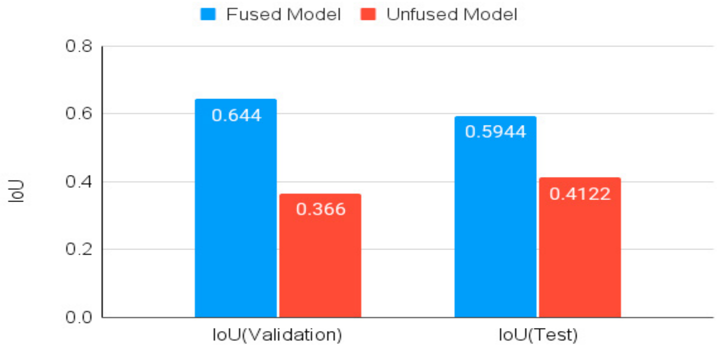

Fig. 4: Fused and Unfused Segmentation model comparison

Table 3: Fused and Unfused Model Comparison

	Fused Model	Unfused Model
IoU (Validation)	0.644	0.366
IoU (Test)	0.5944	0.4122

COVID-19 as well as provide an interpretation of the model's conclusions. The segmentation model's performance is enhanced further by merging the features obtained from the classification model.

The proposed GGECS system outperforms earlier efforts on the same dataset, resulting in a classification challenge benchmark performance. The proposed classification model has an accuracy rate of 98.51 %, while the segmentation model has an IoU score of 59.44. The suggested end-to-end system, which includes classification, segmentation, explainable AI, and a user-friendly UI, serves as a useful and supportive diagnosis method for diagnosing COVID-19 rapidly and effectively. Hence, the proposed GGECS system acts as an effective Covid-19 Decision Support System for the doctors and can thus aid in faster and efficient diagnosis. The findings of this study were confirmed by a leading physician in the field.

6. Acknowledgement

The authors of this paper thank Dr. T. V. Ramakrishnan, M.D., Professor Head, Accident Emergency Medicine, Sri Ramachandra Medical College and Research Institute for his invaluable contribution in examining and validating our work.

References

[1] Rahib H Abiyev and Mohammad Khaleel Sallam Ma'aitaH. Deep convolutional neural networks for chest diseases detection. *Journal of healthcare engineering*, 2018, 2018.

[2] Parnian Afshar, Shahin Heidarian, Nastaran Enshaei, Farnoosh Naderkhani, Moezedin Javad Rafiee, Anastasia Oikonomou, Faranak Babaki Fard, Kaveh Samimi, Konstantinos N Plataniotis, and Arash Mohammadi. Covid-ct-md, covid-19 computed tomography scan dataset applicable in machine learning and deep learning. *Scientific Data*, 8(1):1–8, 2021.

[3] Parnian Afshar, Shahin Heidarian, Farnoosh Naderkhani, Anastasia Oikonomou, Konstantinos N Plataniotis, and Arash Mohammadi. Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images. *Pattern Recognition Letters*, 138:638–643, 2020.

[4] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017.

[5] Deng-Ping Fan, Tao Zhou, Ge-Peng Ji, Yi Zhou, Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao. Inf-net: Automatic covid-19 lung infection segmentation from ct images. *IEEE Transactions on Medical Imaging*, 39(8):2626–2637, 2020.

[6] Kai Gao, Jianpo Su, Zhongbiao Jiang, Ling-Li Zeng, Zhichao Feng, Hui Shen, Pengfei Rong, Xin Xu, Jian Qin, Yuexiang Yang, et al. Dual-branch combination network (dcn): Towards accurate diagnosis and lesion segmentation of covid-19 using ct images. *Medical image analysis*, 67:101836, 2021.

[7] Leilani H Gilpin, David Bau, Ben Z Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*, pages 80–89. IEEE, 2018.

[8] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5):1–42, 2018.

[9] Md Jahid Hasan, Md Shahin Alom, and Md Shikhar Ali. Deep learning based detection and segmentation of covid-19 & pneumonia on chest x-ray image. In *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, pages 210–214. IEEE, 2021.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[11] Tao He, Jixiang Guo, Nan Chen, Xiuyuan Xu, Zihuai Wang, Kaiyu Fu, Lunxu Liu, and Zhang Yi. Medimlp: using grad-cam to extract crucial variables for lung cancer postoperative complication prediction. *IEEE Journal of Biomedical and Health Informatics*, 24(6):1762–1771, 2019.

[12] Andreas Holzinger, Chris Biemann, Constantinos S Pattichis, and Douglas B Kell. What do we need to build explainable ai systems for the medical domain? *arXiv preprint arXiv:1712.09923*, 2017.

[13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

[14] Maede Maftouni, Andrew Chung Chee Law, Bo Shen, Zhenyu James Kong Grado, Yangze Zhou, and Niloofar Ayoobi Yazdi. A robust ensemble-deep learning model for covid-19 diagnosis based on an integrated ct scan images database. In *IIE Annual Conference. Proceedings*, pages 632–637. Institute of Industrial and Systems Engineers (IISE), 2021.

[15] Arpana Mahajan and Sanjay Chaudhary. Categorical image classification based on representational deep network (resnet). In *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, pages 327–330. IEEE, 2019.

[16] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U Rajendra Acharya. Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in biology and medicine*, 121:103792, 2020.

[17] Matteo Pennisi, Isaak Kavasidis, Concetto Spampinato, Vincenzo Schinina, Simone Palazzo, Federica Proietto Salanitri, Giovanni Bellitto, Francesco Rundo, Marco Aldinucci, Massimo Cristofaro, et al. An explainable ai system for automated covid-19 assessment and lesion categorization from ct-scans. *Artificial intelligence in medicine*, 118:102114, 2021.

[18] Chunli Qin, Demin Yao, Yonghong Shi, and Zhijian Song. Computer-aided detection in chest radiography based on artificial intelligence: a survey. *Biomedical engineering online*, 17(1):1–23, 2018.

[19] Priyanka Rastogi, Kavita Khanna, and Vijendra Singh. Gland segmentation in colorectal cancer histopathological images using u-net inspired convolutional network. *Neural Computing and Applications*, 34(7):5383–5395, 2022.

[20] Priyanka Rastogi, Vijendra Singh, and Monika Yadav. Deep learning and big datatechnologies in medical image analysis. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pages 60–63. IEEE, 2018.

[21] 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.

[22] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[23] Wojciech Samek, Thomas Wiegand, and Klaus-Robert Müller. Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint arXiv:1708.08296*, 2017.

[24] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626, 2017.

[25] Prabira Kumar Sethy and Santi Kumari Behera. Detection of coronavirus disease (covid-19) based on deep features. 2020.

[26] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[27] Vijendra Singh, Vijayan K Asari, and Rajkumar Rajasekaran. A deep neural network for early detection and prediction of chronic kidney disease. *Diagnostics*, 12(1):116, 2022.

[28] Alberto Tena, Francesc Clarià, and Francesc Solsona. Automated detection of covid-19 cough. *Biomedical Signal Processing and Control*, 71:103175, 2022.

[29] Fei Wang, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, and Xiaoou Tang. Residual attention network for image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3156–3164, 2017.

[30] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2921–2929, 2016.

[31] Yazeed Zoabi, Shira Deri-Rozov, and Noam Shomron. Machine learning-based prediction of covid-19 diagnosis based on symptoms. *npj digital medicine*, 4(1):1–5, 2021.