A Survey on Deep Learning Method to Detect Infection Caused by "Novel Corona Virus pneumonia" From High Resolution CT Images

A SURVEY REPORT

Submitted by

SUMIT RAJAK (GCECTB-R21-2038)

SOURAV SENAPATI (GCECTB-R21-2032)

SOUMYAPIA CHAKRABORTY (GCECTB- R21-2031)

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

Year: 2023-24

GOVT. COLLEGE OF ENGINEERING AND CERAMIC TECHNOLOGY

73, Abinash Chandra Banerjee Ln, Beleghata, Kolkata, West Bengal 700010

CERTIFICATE

Certified that the project's survey report "Deep learning method to detect infection caused by "Novel Corona Virus pneumonia" from High resolution CT images." is the bonafide work of "SUMIT RAJAK, SOURAV SENAPATI, SOUMYAPIA CHAKRABORTY" who carried out the project work at Government College of Engineering and Ceramic Technology under my supervision and guidance in the fulfilment of requirements of the 6th Semester, B. Tech (Information Technology).

Dr. Mausumi Maitra (Mazumdar)	**Mr. Pranay Adak**
Head of The Department	**Project-Guide and Supervisor**
Professor & Head of the Department	Assistant Professor
Information Technology	Information Technology

External Examiner

Acknowledgement

We would like to sincerely thank and express our gratitude towards all the concerned individuals and organizations for supporting us.

We extend my heartfelt gratitude to our project guide Assistant Professor Mr. Pranay Adak for him invaluable guidance and unwavering support throughout the completion of this project on "Deep learning method to detect infection caused by “Novel Corona Virus pneumonia” from High resolution CT images.." Him expertise and insightful suggestions profoundly shaped the direction of this endeavour. Mr. Pranay Adak’s mentorship, encouragement, and scholarly inputs were instrumental in understanding of machine learning models. Him dedication to academic excellence and commitment to fostering an environment of learning have been a constant source of inspiration.

Lastly, we would like to thank Government College of Engineering and Ceramic Technology, Department of INFORMATION TECHNOLOGY for providing us the opportunity to undertake this amazing project.

SUMIT RAJAK (GCECTB-R21-2038)

SOURAV SENAPATI (GCECTB-R21-2032)

SOUMYAPIA CHAKRABORTY (GCECTB-R21-2031)

TABLE OF CONTENTS

OBJECTIVE

The literature survey for the project "Deep Learning Method to Detect Infection Caused by Novel Coronavirus Pneumonia from High-Resolution CT Images" is to comprehensively review existing research in the field of medical image analysis, particularly in the context of COVID-19 diagnosis using CT imaging. By gaining familiarity with previous work, identifying research trends, evaluating and critiquing methodologies, selecting relevant articles, and informing model design and training, the survey aims to provide essential insights for the development of a robust deep learning system capable of accurately detecting COVID-19 pneumonia from CT scans. Through this process, researchers can leverage existing knowledge to address key challenges and contribute to the ongoing efforts in combating the pandemic.

ABSTRACT

The project "Deep Learning Method to Detect Infection Caused by Novel Coronavirus Pneumonia from High-Resolution CT Images" encapsulates a comprehensive approach aimed at leveraging deep learning techniques for the accurate detection of COVID-19 pneumonia from CT scans. With computed tomography (CT) emerging as a preferred diagnostic tool for COVID-19 pneumonia, the project proposes the construction of a system based on powerful deep learning algorithms for efficient and cost-effective diagnosis compared to conventional RT-PCR tests. The project entails a thorough literature survey to gather insights from existing research, critically evaluating methodologies, identifying trends, and selecting relevant articles, including studies utilizing prominent neural network architectures such as ResNet and UNet. By synthesizing insights from previous work and leveraging state-of-the-art deep learning architectures, the project seeks to develop a robust model capable of accurately identifying COVID-19 pneumonia patterns in high-resolution CT images. Through this interdisciplinary approach, the project aims to contribute to the ongoing efforts in combating the COVID-19 pandemic by providing a reliable and efficient tool for early and accurate diagnosis, thus facilitating prompt medical intervention and management of the disease.

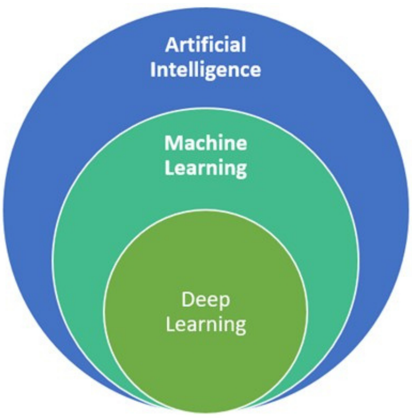
INTRODUCTION

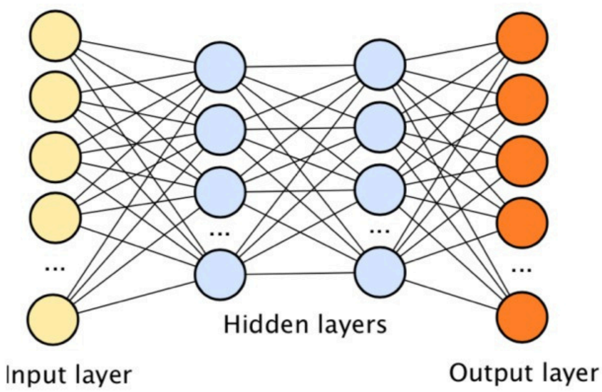
The project titled "Deep Learning Method to Detect Infection Caused by Novel Coronavirus Pneumonia from High-resolution CT Images" addresses a critical need in the current medical landscape, where computed tomography (CT) has emerged as the primary imaging modality for diagnosing COVID-19 pneumonia. With the aim of offering a cost-effective alternative to the conventional RT-PCR test, this project proposes the development of a deep learning-based system for the precise detection of COVID-19 pneumonia from high-resolution CT scans. By leveraging state-of-the-art deep learning algorithms and methodologies for medical image segmentation and identification, the project seeks to adopt a model-based approach to achieve accurate and efficient diagnosis. Central to this endeavor is the exploration of various network architectures through a comprehensive survey, ensuring the selection of the most suitable framework for the task at hand. Through the random selection and utilization of CT images from available datasets, this project endeavors to train a robust model capable of effectively detecting COVID-19 pneumonia, thereby contributing to improved diagnostic capabilities and timely intervention in the fight against the ongoing pandemic.

DEEP LEARNING

- **What is Deep Learning?**

- Deep Learning is a subfield of artifical intelligence and machine learning that is inspired by the structure of a human brain.Deep Learning is a cutting-edge subfield of artificial intelligence and machine learning that has revolutionized various industries by enabling computers to learn complex patterns and make decisions in a manner akin to human thinking. At its core, Deep Learning is inspired by the structure and functioning of the human brain, particularly its neural networks. These networks consist of interconnected nodes, or artificial neurons, organized in layers. Each layer processes information and passes it on to the next, with deeper layers extracting higher-level features from the input data. Deep Learning algorithms, typically implemented in deep neural networks, learn to perform tasks by adjusting the parameters of these interconnected layers through a process called backpropagation. Unlike traditional machine learning methods, Deep Learning algorithms can automatically discover intricate patterns in data, making them exceptionally proficient in tasks such as image and speech recognition, natural language processing, and even playing complex games like Go and chess at a superhuman level. The remarkable success of Deep Learning can be attributed to its ability to handle large volumes of data efficiently, its scalability to complex problems, and its capability to continuously improve with more data and computational resources. As technology advances and datasets grow, Deep Learning is poised to play an increasingly crucial role in shaping the future of artificial intelligence and its applications across diverse domains.





Deep Learning algorithms attemt to draw similar conclusion as human brain would by continuosly analysing data with a given logical structure called Nural Network

TYPES OF NURAL NETWORK

Neural networks are computational models inspired by the structure and function of the human brain, comprised of interconnected nodes or neurons organized in layers. Artificial Neural Networks (ANNs) constitute the foundation, with Multilayer Perceptrons (MLPs) being a common architecture featuring multiple layers of neurons, each layer connected to the next. Convolutional Neural Networks (CNNs) excel in analyzing grid-like data such as images by utilizing convolutional layers to extract features hierarchically. Recurrent Neural Networks (RNNs) are designed for sequential data processing, utilizing feedback loops to incorporate information from previous steps. Generative Adversarial Networks (GANs) employ a unique framework where two networks, a generator and a discriminator, compete to generate data indistinguishable from real data, fostering the creation of novel content. These diverse types of neural networks have revolutionized numerous fields including computer vision, natural language processing, and generative modeling, powering advancements in various applications such as image recognition, language translation, and artistic creation.

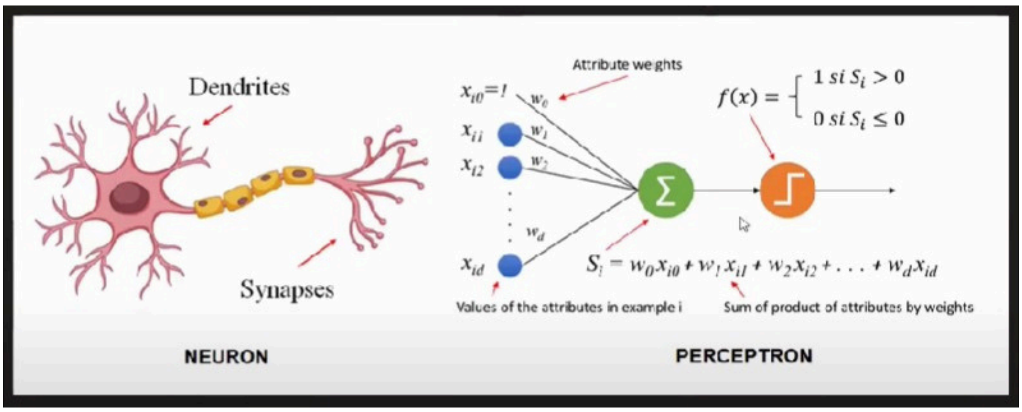
There are difference types of nural networks-

- **ANN**
- **Multilayer Perceptron (MLP)**
- **Convolutional Neural Network (CNN)**
- Recurrent Neural Network (RNN)
- Generative Adversarial Networks (GANs)

ANN

What is Perceptron?

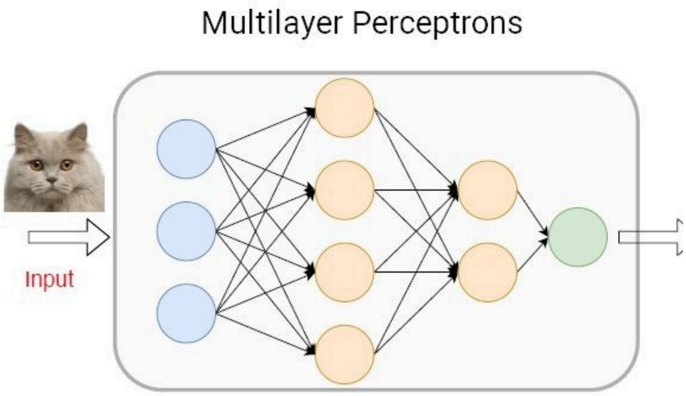
Perceptron is Machine Learning algorithm for supervised learning of various binary classification tasks.



TYPES OF NURAL NETWORK

Multilayer perceptron

A Multilayer Perceptron (MLP) is a type of artificial neural network consisting of multiple layers of interconnected nodes (neurons). It is a feedforward neural network, meaning that information moves in one direction—from the input layer through the hidden layers to the output layer—without any loops or cycles. Each neuron in an MLP is connected to every neuron in the subsequent layer, and each connection is associated with a weight that determines the strength of the connection. MLPs are commonly used for supervised learning tasks, such as classification and regression, and they employ activation functions to introduce non-linearity into the network, enabling it to learn complex patterns and relationships in the data. Training an MLP involves adjusting the weights of the connections using optimization algorithms like gradient descent, typically in conjunction with backpropagation to efficiently compute the gradient of the loss function with respect to the weights.



CNN

Convolutional Layers:
Convolutional layers are the building blocks of CNNs.They consist of a set of learnable filters (also called kernels) that slide across the input image, computing the dot product between the filter and the input at every position.The output of the convolution operation forms a feature map, highlighting the presence of different patterns or features in the input image.

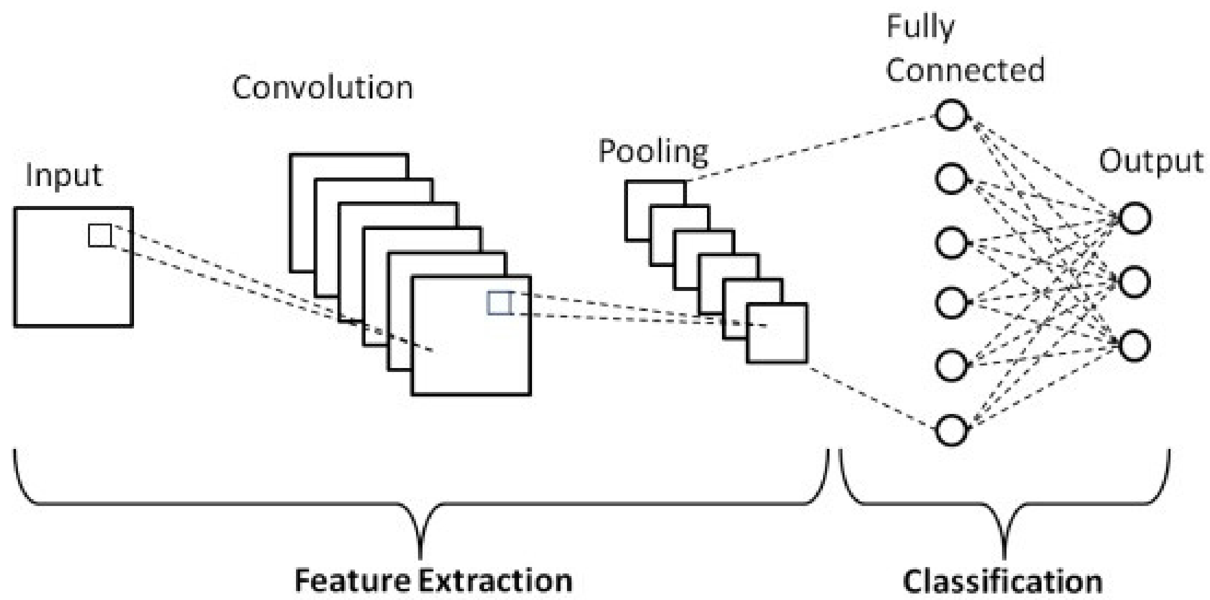
TYPES OF NURAL NETWORK

CNN

Pooling Layers:
Pooling layers are used to downsample the feature maps obtained from convolutional layers. Common pooling operations include max pooling and average pooling, which reduce the spatial dimensions of the feature maps while retaining the most important information. Pooling helps in achieving translation invariance and reducing computational complexity.

Activation Functions:
Activation functions introduce non-linearity into the network, allowing it to learn complex patterns.Common activation functions used in CNNs include ReLU (Rectified Linear Unit),sigmoid, and tanh.ReLU is preferred due to its simplicity and effectiveness in addressing the vanishing gradient problem.

Fully Connected Layers:
Fully connected layers are typically used towards the end of the CNN architecture to perform classification or regression tasks.They take the output of the convolutional and pooling layers and transform it into a vector representation suitable for the final task.These layers are often followed by a softmax activation function for classification tasks to produce class probabilities.



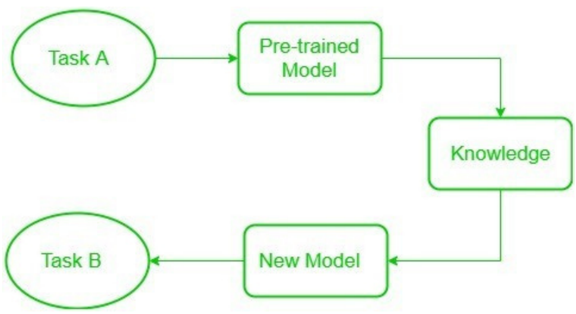
TRANSFER LEARNING

- **What is Transfer Learning?**

- Transfer learning is a technique in machine learning where a model trained on one task is used as the starting point for a model on a second task. This can be useful when the second task is similar to the first task, or when there is limited data available for the second task. By using the learned features from the first task as a starting point, the model can learn more quickly and effectively on the second task.

- This is what transfer learning is. Nowadays, it is very hard to see people training whole convolutional neural networks from scratch, and it is common to use a pre-trained model trained on a variety of images in a similar task, e.g models trained on ImageNet (1.2 million images with 1000 categories) and use features from them to solve a new task. When dealing with transfer learning, we come across a phenomenon called the freezing of layers. A layer, it can be a CNN layer, hidden layer, a block of layers, or any subset of a set of all layers, is said to be fixed when it is no longer available to train. Hence, the weights of freeze layers will not be updated during training. While layers that are not frozen follows regular training procedure. When we use transfer learning in solving a problem, we select a pre-trained
- model as our base model. Now, there are two possible approaches to using knowledge from the pre-trained model. The first way is to freeze a few layers of the pre-trained model and train other layers on our new dataset for the new task. The second way is to make a new model, but also take out some features from the layers in the pre-trained model and use them in a newly created model. In both cases, we take out some of the learned features and try to train the rest of the model. This makes sure that the only feature that may be the same in both of the tasks is taken out from the pre-trained model, and the rest of the model is changed to fit the new dataset by training.

- There are different type of pretrained model already exist For Example -
  1. RESNET 50 For Image Classification
  2. UNET For Image Segmentation.......



RESNET

ResNet50 is a deep convolutional neural network (CNN) architecture that was developed by Microsoft Research in 2015. It is a variant of the popular ResNet architecture, which stands for "Residual Network." The "50" in the name refers to the number of layers in the network, which is 50 layers deep.

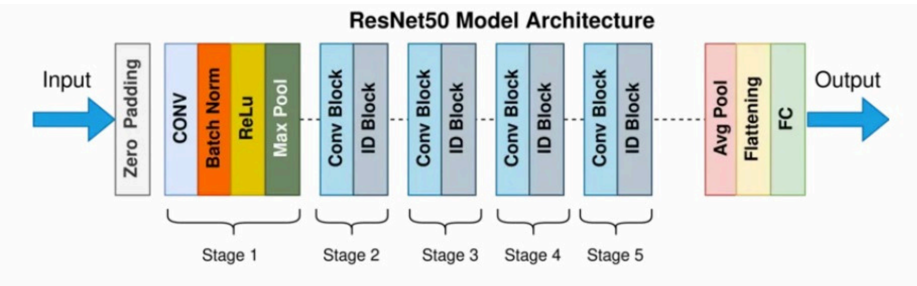
ResNet50 has been trained on large datasets and achieves state-of-the-art results on several benchmarks. It has been trained on the ImageNet dataset, which contains over 14 million images and 1000 classes. On this dataset, ResNet50 achieved an error rate of 22.85% which is on par with human performance, which is an error rate of 5.1%.

RESNET ARCHITECTURE

The convolutional layers in ResNet50 consist of several convolutional layers followed by batch normalization and ReLU activation. These layers are responsible for extracting features from the input image, such as edges, textures, and shapes. The convolutional layers are followed by max pooling layers, which reduce the spatial dimensions of the feature maps while preserving the most important features.

The identity block and convolutional block are the key building blocks of ResNet50. The identity block is a simple block that passes the input through a series of convolutional layers and adds the input back to the output. This allows the network to learn residual functions that map the input to the desired output. The convolutional block is similar to the identity block, but with the addition of a 1x1 convolutional layer that is used to reduce the number of filters before the 3x3 convolutional layer.

The final part of ResNet50 is the fully connected layers. These layers are responsible for making the final classification. The output of the final fully connected layer is fed into a softmax activation function to produce the final class probabilities.



ResNet50 Architecture

UNET

What is U-Net?

U-Net is a convolutional neural network (CNN) architecture designed for semantic segmentation tasks. Its name is derived from its U-shaped architecture, which consists of a contracting path (encoder) followed by an expansive path (decoder). This unique structure allows U-Net to capture context at different scales while maintaining spatial information.U-Net is a widely used deep learning architecture that was first introduced in the “U-Net: Convolutional Networks for Biomedical Image Segmentation” paper.This network was designed to effectively leverage a smaller amount of data while maintaining speed and accuracy.

• UNET ARCHITECTURE

1.Contracting Path (Encoder)

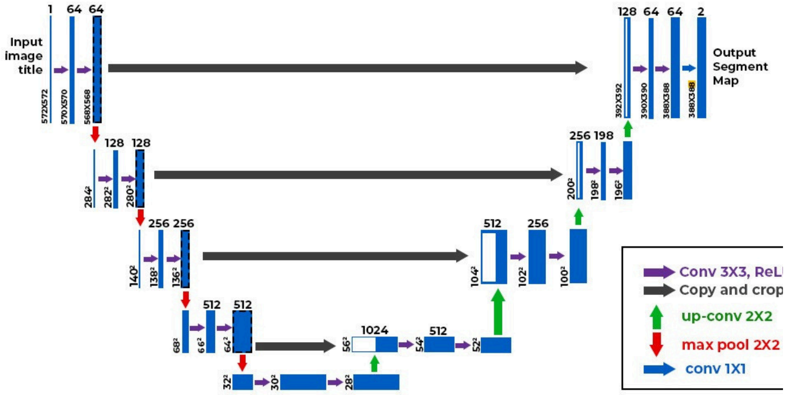
The encoder is responsible for capturing high-level features and reducing the spatial dimensions of the input image.It consists of repeated blocks of convolutional layers followed by max-pooling layers, effectively downsampling the input.

2. Bottleneck

At the center of the U-Net is a bottleneck layer that captures the most critical features while maintaining spatial information.

3. Expansive Path (Decoder)

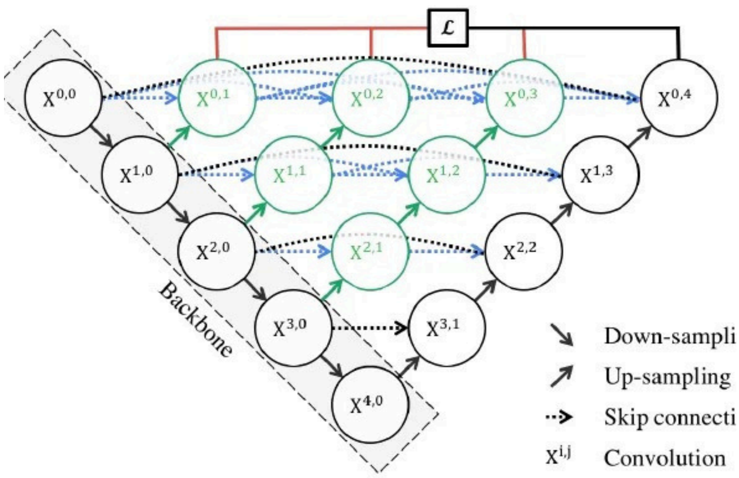
The decoder is responsible for upsampling the low-resolution feature maps to match the original input size.It consists of repeated blocks of transposed convolutions (upsampling) followed by concatenation with corresponding feature maps from the contracting path



UNET++

UNet++, an advancement over the original U-Net architecture, integrates DenseNet-inspired concepts to enhance performance in medical image segmentation tasks. Three key differentiators characterize UNet++: Firstly, it incorporates convolution layers within skip pathways, facilitating the amalgamation of semantic information between encoder and decoder feature maps, thereby mitigating the semantic gap. Secondly, it employs dense skip connections along skip pathways, optimizing gradient flow throughout the network, which aids in smoother and more efficient information propagation. Thirdly, UNet++ introduces deep supervision, allowing for multiple loss layers and enabling model pruning while maintaining or even surpassing performance comparable to using a single loss layer. By amalgamating these innovations, UNet++ not only enhances the segmentation accuracy but also ensures better gradient propagation, thus offering a refined solution for medical image segmentation tasks.

UNet++ uses the Dense block ideas from DenseNet to improve U-Net.
UNet++ differs from the original U-Net in three ways:
1) having convolution layers on skip pathways, which bridges the semantic gap between encoder and decoder feature maps.
2) having dense skip connections on skip pathways, which improves gradient flow.
3) having deep supervision, which enables model pruning and improves or in the worst case achieves comparable performance to using only one loss layer.



UNet++ starts with an encoder sub-network or backbone followed by a decoder sub-network.There are re-designed skip pathways (green and blue) that connect the two sub-networks and the use of deep supervision (red).

PREVIOUS RELEVANT WORK

Computed tomography (CT) is the preferred imaging method for diagnosing 2019 novel coronavirus (COVID19) pneumonia. We aimed to construct a system based on deep learning for detecting COVID-19 pneumonia on high resolution CT. For model development and validation, 46,096 anonymous images from 106 admitted patients, including 51 patients of laboratory confrmed COVID-19 pneumonia and 55 control patients of other diseases in Renmin Hospital of Wuhan University were retrospectively collected.

• DATA SET

Computed tomography (CT) is the preferred imaging method for diagnosing 2019 novel coronavirus (COVID19) pneumonia. We aimed to construct a system based on deep learning for detecting COVID-19 pneumonia on high resolution CT. For model development and validation, 46,096 anonymous images from 106 admitted patients, including 51 patients of laboratory confrmed COVID-19 pneumonia and 55 control patients of other diseases in Renmin Hospital of Wuhan University were retrospectively collected.

Training algorithm. Tis work is built on the top of UNet++, a novel and powerful architecture for medical image segmentation19, for the identifcation. Resnet-50 was used as backbone of UNet++ as previously described20. ResNet-5021 was pretrained using ImageNet dataset22, and all the pre-training parameters of ResNet-50 are loaded to UNet++.

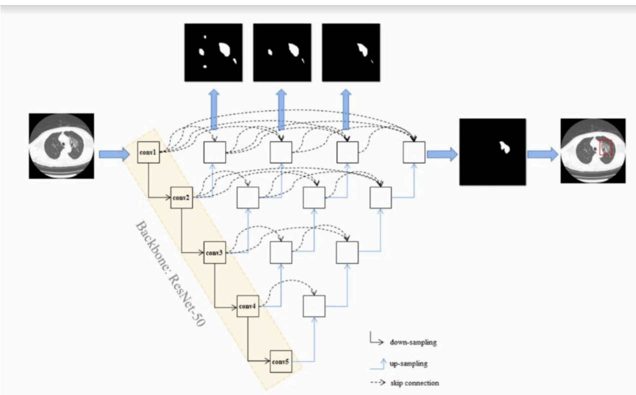


Figure 3. The network architecture of UNet++. UNet++ consists of encoder and decoder connecting through a series of nested dense convolutional blocks. The semantic gap between the feature maps of the encoder and decoder is bridged prior to fusion. The encoder extract features by down-sampling; the decoder map features to the original image by up-sampling, make classification by pixels, and thus achieve the purpose of segmentation. Resnet-50 was used as backbone of UNet++, and all the pre-training parameters of ResNet-50 are loaded to UNet++.

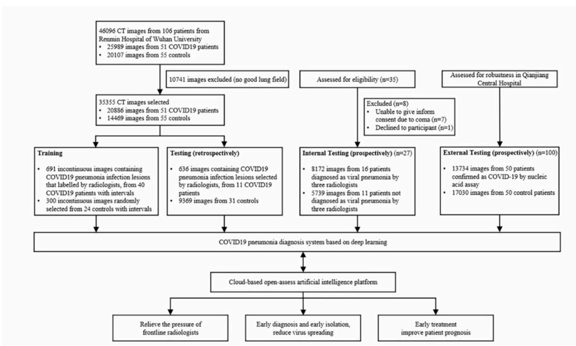


Figure - Workfow diagram for the development and evaluation of the model for detecting COVID19 pneumonia.

PREVIOUS RELEVANT WORK

Comparison between the efficiency of radiologist with or without the assistance of AI. In the frst time the expert radiologist read CT scan images of the 27 prospective patients, the average reading time forhim to determine whether each patient has viral pneumonia was 116.12 s per case (IQR 85.69–118.17).

After 10 days of wash out period, the same expert radiologist re-read the CT images of the 27 prospective patients with the assistance of the AI model. Te results for determining whether each patient has viral pneumonia were not changed, while the average reading time of the expert was greatly decreased by 65%. Tis indicates that the efciency of radiologist could be greatly improved with the assistance of AI.

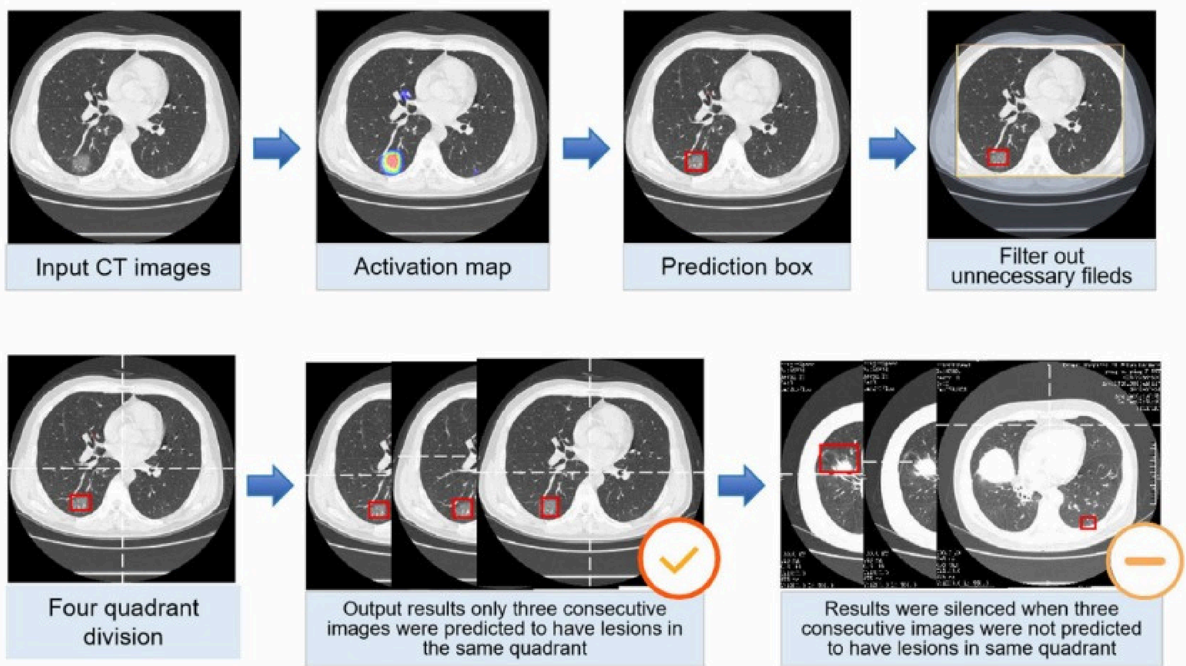


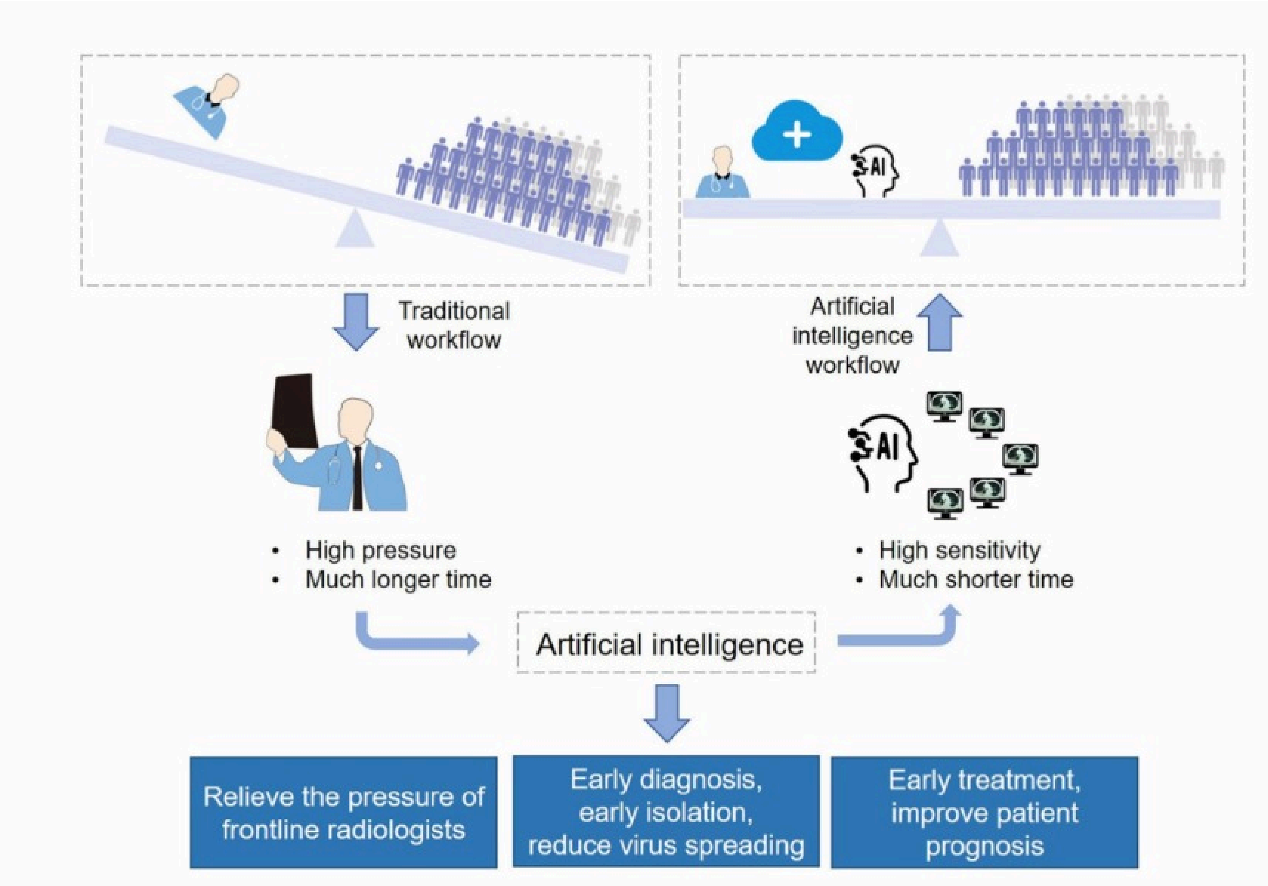
Figure 4. Processing and prediction schematic of the model. Raw images were firstly input into the model, and after processing of the model, prediction boxes framing suspicious lesions were output. Valid areas were further extracted and unnecessary fields were filter out to avoid possible false positives. To predict by case, a logic linking the prediction results of consecutive images was added. Computed tomography (CT) images with the above prediction results were divided into four quadrants, and results would be output only when three consecutive images were predicted to have lesions in the same quadrant.

CONCLUSION

Computed tomography (CT) is the most effcient modality for screening and clinically diagnosing COVID-19 pneumonia. However, compared to the needs of the patients, the number of radiologists is quite small.

After enrolling artifcial intelligence in identifying COVID-19 pneumonia in CT images, the efciency of diagnosis is greatly improved. The artifcial intelligence holds great potential to relieve

- **the pressure of frontline radiologists,**
- **accelerates the diagnosis,**
- **isolation and treatment of COVID19 patients** and therefore contribute to the control of the epidemic.



REFERENCES

1. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography
Jun Chen, Lianlian Wu, Jun Zhang, Liang Zhang, Dexin Gong, Yilin Zhao, Qiuxiang Chen, Shulan Huang, Ming Yang, Xiao Yang, Shan Hu, Yonggui Wang, Xiao Hu, Biqing Zheng, Kuo Zhang, Huiling Wu, Zehua Dong, Youming Xu, Yijie Zhu, Xi Chen, Mengjiao Zhang, Lilei Yu, Fan Cheng & Honggang Yu

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-76282-0
Correspondence and requests for materials should be addressed to L.Y., F.C. or H.Y.
Reprints and permissions information is available at www.nature.com/reprints.

2. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. IEEE conference on computer vision and pattern recognition. 770–778, (2016).

3. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. & Liang, J. Unet++: a nested u-net architecture for medical image segmentation. Preprint at. https://arxiv.org/abs/1807.10165 (2018).

4. Chen, L., Papandreou, G., Schrof, F. & Adam, H. Rethinking atrous convolution for semantic image segmentation. Preprint at.
https://arxiv.org/abs/1706.05587 (2017).

5. J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3431–3440, 2015.

6. O. Ronneberger, P. Fischer, and T. 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.

7. Automatic prediction of COVID− 19 from chest images using modified ResNet50 Marwa Elpeltagycorresponding author1 and Hany Sallam2

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8095476/#:~:text=In%20this%20study%2C%20the%20efficacy,extract%20the%20best%20discriminative%20features