Review on COVID-19 Detection Method Based on Deep Learning

Yuhao Wang

University of California, Davis, USA, 550 Alumni Ln,Davis,CA 95616

wayne521025@gmail.com

Abstract. This article discusses the use of deep learning in combating Covid-19. It has been four years since Covid-19 ravaged the world. We hope to help researchers and medical staff through our investigation. Deep learning is a very effective method, which has a wide range of functions. Deep learning neural network is very suitable for fighting against covid. We will also introduce the information of this article in detail. Because neural network is mainly used to fight against CT in tumor and other diseases. First, I introduce the main image types. This is mainly CT type. I introduced the main features of this figure.Then we introduce the problem of image classification. Here we use convolutional neural network and its optimization scheme. At last, the problem of object recognition and image segmentation is introduced. It is hoped that it will aid in the rapid adoption of deep learning in COVID-19 research. I also try my best to contribute to this research. In recent years, in-depth learning has played a vital role in all fields. Especially in medicine, we hope that through this report, we can give you a deeper understanding of deep learning.

Keywords: medical staff, neural network, image classification, object recognition

1. Introduce

It will discuss the use of deep learning in image classification and target detection. Previous work has proved that people can improve accuracy through simple operations. For example, rotation and clipping. We will strengthen and improve accuracy by using deep learning. Let the neural network learn to improve the classifier, which will greatly improve the accuracy.

This paper will introduce the types of images processed, the classification of images, and the target recognition and segmentation of images.

Covid-19 is a viral infection caused by the Sars-cov-2 virus. Up to now, there are 607 million patients and 6.51 million deaths. This figure continues to grow. During my literature study, I discovered several examples of deep learning systems used against covid-19. Deep learning technology is widely used. For example, the classification of hyperspectral images defeated the go world champion.

2. Main image types in COVID-19 recognition method

Computed Tomography imaging is to judge the disease according to the different density of these tissues. Moreover, because of its CT imaging has a high density resolution and may frequently clearly depict soft tissue structures and organs with tiny density changes. also sensitively find lesions and display the characteristics of lesions, thus assisting in disease inspection and diagnosis. In addition, if someone already has a disease, CT scanning and PET scanning can help determine the severity of the disease. PET scanning can show the areas of increased activity in the body, while separate CT scanning can generate detailed images of tissues and organs in the body.[1] These methods allow doctors to see exactly where the problem is and how serious it is. This is a relatively painless process, which can measure the anatomy and metabolic function of the patient's body, because the image is captured in a single scan.

In fact, medical imaging inspection is to see the internal organs of the body by fluoroscopy outside the human body, so overlapping images are likely to occur. However, CT imaging is a sectional image,

so generally the tissue structure displayed does not overlap the image of the disease, As a result, the illness detection rate can be greatly increased.[2]

3. CT image classification method for COVID-19

Different from the conventional neural network, the The convolutional neural network's layers each include neurons organized in width, height, and depth are the three distinct dimensions. Convolution is a two-dimensional process, therefore understanding the breadth and height is easy.[3] However, a convolutional neural network's depth does not correspond to the depth of the entire network; rather, it refers to the active data volume's third dimension. A network's depth describes how many layers are there overall. To further grasp what width, height, and depth are, consider an example. The size of the input data volume is 32x32x3 if the CIFAR-10 picture is utilized as the convolution neural network's input (width, height and depth). As opposed to using the complete connection strategy, the neurons in the layer will only link with a tiny portion of the preceding layer. The convolution network used by CIFAR-10 to categorize the pictures' final output layer is 1x1x10. Because the last component of the convolution neural network construction involves compressing the full-size picture into a vector that contains the classification score and arranging the vectors in the depth direction.[4]

The 3D input data from the convolutional neural network is transformed into the 3D activation data of the neurons and output by each layer.[5] It has a 3 depth (representing three color channels: red, green, and blue). Following the convolution pooling layer is the activation value following convolution and pooling, also referred to as a neuron, in the blue area next to the red.[6]

The input layer and hidden layer are designed as "complete connections" in a typical neural network. It is possible, mathematically speaking, to extrapolate the characteristics of relatively tiny pictures from the entire image. However, from a calculational standpoint, learning the characteristics of a larger picture using this full connection network technique will take a lot of time. You must create a 10 to the power of 4 (100,000) design. There are 10 6th power parameters to learn if you wish to learn 100 characteristics. The 96x96 picture employs the forward transmission or backward transmission computation technique in comparison to the 28x28 block image. and the calculation process is also 10 times slower to the second power (=100).[7]

A simple way to restricting the link between the hidden units and the input units—each hidden unit can connect to a very limited number of the input units is one way to handle this type of problem using convolution layers. For instance, just a little portion of the input picture is connected by each concealed unit.[8] (There will also be some unique selection techniques for the "connection region" of the input signal connected to the single hidden layer for input forms other than picture input.) The subset of input units, for instance, that a hidden unit must link in a signal input method may only be the signal during a specific time period matching to an audio input.)[9]

The neurons in the convoluted layer also have depth since they are three-dimensional. A number of filters are included in the convolution layer's settings. Numerous filter output units share the same depth, which is trained in each filter.[10]

The size of the sample input unit is 32 32 3, as seen in the image below. The output unit has a depth of 5. The region’s connection to the source picture is same at the same position with varyingdepths of the output unit, but there are different parameters (filters).[11]

Each layer's function is really the convolution of the input layer and the weight if the application parameters are shared. This is where the convolutional neural network name came from.

Ignore convolution, please. Consider a size of 5 x 5 and a 3 x 3 for simplicity's sake. Let's just keep in mind that the convolution kernel in this case comprises a total of 9 parameters. The convolutional nucleus in this instance has 9 neurons, and its output is a 3. The 3-by-3 matrix is known as a characteristic graph. There is overlap! The first neuron is connected to the first 3 of the picture, while the second neuron is related to the second section. They are continually searching, much as your eyes are while you're looking around.[12]

4. Image target recognition and image segmentation for COVID-19

Target detection, sometimes referred to as target extraction, is the segmentation of a picture depending on the target's geometrical and statistical characteristics. Target identification and segmentation are combined into one procedure. Precision and real-time nature are two important aspects of the overall system. In complex situations where several targets must be handled in real-time.

Two series—the RCNN series and the YOLO series—are used for target detection. In addition to the well-known SSDs based on the enhancement of the first two series, the RCNN series and YOLO are exemplary algorithms based on area extraction and detection, respectively.

Candidate frame creation is a common component of target detection methods. Currently, the major methods for obtaining suitable frames are picture segmentation and region growth. The local region similarity (color, texture, etc.) of objects in the detected picture is the major cause of region expansion (merging). Technology advancements in object identification and picture segmentation support the efficient extraction of information from photos even more.[13]

The flow chart for the sliding window approach makes the fundamental notion very clear: first, the input picture may be moved using sliding windows of various sizes of windows from left to right and from top to bottom. Each time you slide, the classifier is applied to the currently active window. [14] If there is a good likelihood that the object will be categorised in the current window, it is considered to have been detected. The multiple windows' detected object marks will be obtained. There will be parts with a high repetition in these window sizes once each sliding window with a different window size is identified. After NMS filtering, the discovered items are finally acquired. The sliding window approach is straightforward and simple to comprehend, however searching photographs worldwide with various window sizes is inefficient, and the aspect ratio of objects should be taken into account while choosing the window size. For classifiers with high real-time requirements, the sliding window approach is thus not advised.

The exhaustive search of picture subregions is analogous to the sliding window approach, except often most of the image's subregions are object-free. In order to increase computing efficiency, researchers naturally considered limiting their search to the areas that are most likely to contain the image's items. The most well-known approach for extracting picture bounding boxes is selective search (SS), which Koen E. A. introduced in 2011.[10] The selection search algorithm's primary tenet is that the potential areas of items in the image should have some continuity or resemblance. In light of the foregoing concept, the selection search therefore employs the technique of sub region merging to extract bounding boxes. The input picture is first split to create a number of tiny subregions. Second, region merging and region iterative merging are carried out continually in accordance with the similarity between these subregions (the similarity criteria generally include color, texture, and size). Bounding boxes are created for these combined sub areas in each cycle. Candidate boxes is a term that is frequently used to describe these sub region bounded rectangles.[11]

5. Discussion

Finding novel ways to annotate data or build tasks that show learning is a trend in medical image analysis. We looked at the intriguing advancement of text reporting, manual in-loop annotation, and lax supervision in picture text contrast learning. We are quite interested in what Zhang et al. [12] have discovered in their comparison of images and words. We mostly looked at applications in a particular field throughout the study. Solely processing text or only processing picture data is covered here. While also being a potential step towards "basic language acquisition," visual language representation learning. We advise readers interested in this line of inquiry to see Brisk et al. Performance on the glue benchmark has been increased thanks to visual language learning. An intriguing direction is Convirt [13], which offers medical image processing's linguistic "basis." CNN lacks the originality of a human brain, thus they must describe the instances they must handle. Developers of deep learning

typically attempt to address this issue with a method known as "data augmentation," in which they rotate or flip a small sample of pictures prior to training the neural network.

6. Conlusion

Artificial intelligence has been gradually integrated into our daily life. Deep learning is a great way to help us solve medical problems. In this report, we found that there were not many articles to introduce the detection and classification of reported pictures, which proved to be helpful to us. The deep convolution layer may provide us with a new idea to solve the problem. We can't ignore efficiency and accuracy, and different results may appear on some small differences. With the increase of its applications in the real world, there will be an exponential increase in demand for lightweight models that can be used on mobile and embedded devices. Although there is growing interest in this area, there are still many unanswered questions. In this essay, we show how the two-stage and one-stage detectors have developed step by step and surpassed the previous research. Two stage detectors are usually more accurate, but they are slower and difficult to meet apps that run instantly, like automated driving. However, one-stage detector research has accelerated recently has changed this situation, which can achieve the same performance as the former. As time goes by. I believe that the new crown will be eliminated in the near future. We will eventually return to normal life. I hope this investigation on the new crown can help everyone and scholars, and I also hope that we can solve the problems we are facing as soon as possible. Not only for this new crown, but also for the following research.

References

[1] Debesh R. Roy, Pratim K. Chattaraj. Conceptual density functional theory and aromaticity. 2021,,, 285-319. https://doi.org/10.1016/B978-0-12-822723-7.00009-1

[2] Debdutta Chakraborty, Pratim Kumar Chattaraj. Conceptual density functional theory based electronic structure principles. Chemical Science 2021, 123 https://doi.org/10.1039/D0SC07017C

[3] Aaron Keeley, László Petri, Péter Ábrányi-Balogh, György M. Keserű. Covalent fragment libraries in drug discovery. Drug Discovery Today 2020, 25 (6) , 983-996. https://doi.org/10.1016/j.drudis.2020.03.016

[4] Marco Franco-Pérez, José L. Gázquez, Paul W. Ayers, Alberto Vela. Temperature-Dependent Approach to Electronic Charge Transfer. The Journal of Physical Chemistry A 2020, 124 (26) , 5465-5473. https://doi.org/10.1021/acs.jpca.0c02496

[5] Fredrikson M, Jha S, Ristenpart T. Model inversion attacks that exploit confidence information and basic countermeasures; 2015. p. 1322–1333.

[6] Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shen D, Shi Y. Lung infection quantification of COVID-19 in CT images with deep learning; 2020. arXiv:2003.04655.

[7] Huang G, Liu Z, Weinberger KQ. Densely connected convolutional networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR); 2017. p. 2261–2269.

[8] Zoph B, Le QV. Neural architecture search with reinforcement learning; 2017. arXiv:1611.01578.

[9] Richa Arora, Upasana Issar. Computational Approach Toward Identification and Catalytic Degradation of Chemical Warfare Agents Using MOFs. 2022,,, 431-451. https://doi.org/10.1007/978-981-16-7959-9_17

[10] Eduardo Chamorro. Spin‐Polarized CDFT. 2022,,, 111-136. https://doi.org/10.1002/9783527829941.ch7

[11] Shanti Gopal Patra, Himangshu Mondal, Pratim Kumar Chattaraj. Variation in electrophilicity on electronic excitation. Journal of Physical Organic Chemistry 2022, 51 https://doi.org/10.1002/poc.4359

[12] Ranita Pal, Pratim Kumar Chattaraj. Electrophilicity index revisited. Journal of Computational Chemistry 2022, 1999 https://doi.org/10.1002/jcc.26886

[13] Zhang Y, Jiang H, Miura Y, Manning CD, Langlotz CP. Contrastive learning of medical visual representations from paired images and text; 2020. arXiv:2010.00747.