Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/14/20230759

Analysis of convolutional neural networks and its application in object detection

Yiming Gao

Vincent Massey Secondary School, Windsor, ON, CA

Quintonice0@gmail.com

Abstract. With the development of times and technology, artificial intelligence-based neural network algorithms have been widely used in scientific research and life. Among them, Convolutional neural network (CNN) is the most classic and the most representative. CNN have been widely used in images classification in recent years. This article focuses on the basic information of the convolutional neural networks and its applications especially in object detection. Also, the advantages of the CNN and the possible future improvements will be told in this article. In the end of this article some research of experiments will be shown in order to prove the accuracy of convolutional neural networks in detecting objects and visualizing results. This article gathers basic knowledge of convolutional neural networks. As a result, it may help new starters to get to know the CNN and then make contributions to the development of deep learning especially the CNN according to the advantages and future improvements mentioned in the article.

Keywords: CNN, application, object detection.

1. Introduction

As the technology of computer and electrical engineering developing at a high speed nowadays, the images classification program has developed its own system which is based on the convolutional neural network. This article focuses on the theoretical basis of the object detection model which is convolutional neural networks, advantages of CNN, application of the convolutional neural networks and real cases of CNN's application. In addition, the conception of object detection model as well as the possible future improvement of the convolutional neural networks will be told in this article. The first part of this article will focus on the theoretical basis of object detection model. Then the second part will be telling the basic knowledge of the convolutional neural networks including its internal structure and working process. In the last part the conception of object detection model as well as how it uses the convolutional neural networks to detect objects will be told. In order to prove the accuracy of the research in this article, some real cases of CNN's application will be shown at the end of this article. This article is the summary of CNN's knowledge in application of object detection model. It may be helpful for starters to get to know about some facts of CNN and its application, then help to develop the CNN with some basic knowledge.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/14/20230759

2. Theoretical basis of object and key points detection model

Object detection model is a subset of neural networks based on common convolutional neural networks. It basically is a computer vision and image processing technology model that can identify objects in input digital images and videos. For example, an object detection program could find instances of screws on a factory floor, or saw blades on a table next to a workstation.

2.1. CNN basic theory

Convolutional Neural Network (CNN) is a subset of and machine learning. As a type of neural networks, it is made of at least one convolutional layer and followed by one or more completely connected layers just like a standard multilayer neural network [1]. The CNN's structure is designed for deep learning of algorithms in order to make the most of the flat structure of the input image and recognize objects in it. The convolutional neural network's structure is shown in figure 1 [2].

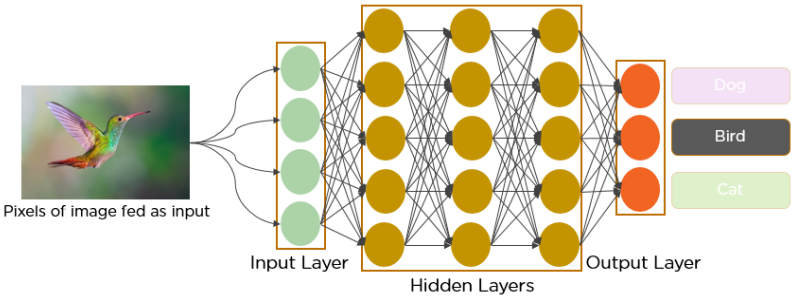

Figure 1. CNN's structure.

2.2. Advantages of CNN

The CNN is particularly useful in recognizing images. The reason is because CNN has the translation invariant feature as a result of the local connection, tied weights and some form of pooling. Convolutional neural networks are often used for image classification. By recognizing features of images, CNN can identify different objects and key points on images. This ability makes it possible to use CNN in medical purpose, for example, for MRI diagnostics. It can also be used in agriculture. The convolutional neural networks receive images input from satellites and then use the information to classify lands based on their level of cultivation. As a result, the output and data from the convolutional neural network can be used for predicting about the fertility level of the lands. Then people can develop a strategy for the optimal use of farmland based on the prediction [3]. Hand-written digits and handwriting recognition are early examples of the use of CNN in real life. Another reason of why CNN is suitable for recognizing images is that CNN is easier to be trained and it has many fewer parameters than normally, fully connected networks with the same number of hidden units.

Large learning capabilities. CNN can analyse data and process it completely only use its built-in intuition. As a result, you can expect it to develop even further over time. In addition, CNN doesn't require humans' supervision for the task to identify significant features, it can detect distinct features from images all by itself. That ability can be extremely helpful when users keep training it with the same kind of images input. Unlike the artificial intelligence, CNN will use its algorithm to interpret and visualize data instead of approaching it the same way every time. Therefore, CNN can get more room for its learning capabilities through that. In addition, CNNs are highly resembled as mammals, so they basically see the images as humans do [4]. That feature makes it more possible to be trained well.

High accuracy in visualizing results. The CNN provides translation equivariance, which means a change in the input data will not alter the representation of the input but would shifts the input in the latent space linearly. As a result, it helps learning more robust representations. With the technology of neural networks developing in a high pace, the current advanced neural networks in image classification are not convolutional neural networks [5]. However, CNN has been dominating for a

DOI: 10.54254/2755-2721/14/20230759

very long time in recent years in most cases and tasks of image and video recognition and similar tasks. That is because the CNN usually shows higher accuracy than non-convolutional neural networks, especially when the input involved a lot of data.

To sum up, convolutional neural networks have a lot of advantages that make it useful for several kinds of purposes. Its large learning capability, similarity to human beings and high accuracy make it possible to be used in daily life and even medical purposes.

2.3. Future improvements

As the technology of convolutional neural networks developing at a high speed, CNN is now getting better than it was before. However, it still has some possible improvement that can be achieved in the future. Take internal structure as an example, the CNN requires a regular structure with fixed size. That requirement makes it less useful when dealing with data that has irregular geometry like graph data. Through the example it can be noticed that the convolutional neural networks are not suitable for all purposes. In order to solve the problem, increasing the dataset size is one possible way. CNN can be stronger and more accurate with a larger dataset size based on its massive learning capability. Additionally, improve the network design is necessary for developing convolutional neural networks. Make convolutional neural networks useful for more purposes is just one of its possible improvements. In the future CNN still has lots of parts that can be improved.

3. Application of CNN

As it is mentioned before, the convolutional neural network is particularly useful in recognizing images due to its translation invariant feature and structure. Furthermore, convolutional neural networks' large learning capability and high accuracy in visualizing results make it possible to deal with harder and more complex issue which is objects detection.

The convolutional neural networks applications in objects detection and recognition can be used widely in our daily life. Take decoding facial recognition as an example, the input is the image or video of faces shot by cameras. The convolutional neural network will detect some key features of faces such as eyes colour, nose tips, eyebrows, eye corners and so on. After that CNN can recognize the face in the image based on the dataset and then make a prediction from the recognition. A box structure that surrounds the face will put the face inside it. Then the program will say yes or no depends on the percentage it thinks that look like the right face feature which is the output. In addition, it is also mentioned that the convolutional neural networks can also be used in agriculture. However, convolutional neural networks can do much more than that. For example, for fighting climate. Experts can understand the climate changes through the change of the weather with the help of convolutional neural networks recognizing weather. CNN can even do advertising and be used to other internet fields.

3.1. Conception of object and key points detection

Object detection is a general term to describe a collection of tasks that needs the computer to detect objects and key points among images and videos. It is based on the image classification which is finished by convolutional neural networks. Many objects detection algorithms are based on deep learning such as convolutional neural networks. In traditional machine learning-based approaches, it usually starts to detect by identifying edges and contours by looking at an image then find its features and group the pixels that may belong to an object. An example picture of visualized result from the object and key points detection model is shown in figure 2 below [6].

DOI: 10.54254/2755-2721/14/20230759

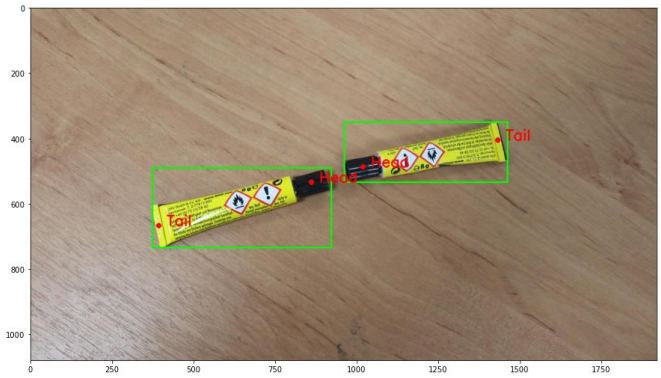

Figure 2. Example of object and key points detection.

Objects detection involves not only recognizing objects and key points, but also prediction of images. It combines the classification and localization together. As a result, it is much more complex than common application of convolutional neural networks. Object localization refers to identifying the location of objects and key points that may look like the target objects in an image, then make its prediction and visualize the result.

3.2. Operating principle of the program

As an algorithm designed for images classification, CNN's have been extensively used to classify all sorts of images and videos. When the convolutional neural network gets an input image, it works as the following steps: First step is to divide the input image into at least several hundred parts in order to recognize them piece by piece. Then wrap some pieces form each region and gather them into a smaller image with feature that can be easily detected. Third, send the data to convolutional neural network and start to detect key points. During the third steps, the smaller image will be detected part by part according to their features. Last but the most important, make prediction and visualize results. The bounding boxes will be refined using bounding box regression so that the object is properly captured by the box [7]. Those steps are basically the operating principle of how convolutional neural networks detect objects in images. The detecting model can be more accurate after it gets trained.

3.3. Cases of CNN utilize in real life

The convolutional neural networks are particularly suitable for all forms of image classifications and objects detection. A lot of experts have taken all several kinds of experiments in order to test the possibility of using convolutional neural network in real image classification and object detection cases. For the past few years, the convolutional neural networks suffered a lot from recognizing Bangla handwriting due to its special structure and working principles [8]. However, a group of experts have solved it and made an experiment to prove that. They trained the detecting model with 118,698 images of Bangla dataset and many images from CMATERdb dataset for Bangla hand-written characters and digits. The proposed model has achieved the accuracy of 97.43% for classification with the average computational costs of 44.95 ms/f [9]. Besides, another experts group has discovered that the convolutional neural networks are particularly useful for detecting rice in the image shot by satellite. The result of the experiment verified the high accuracy of the measured data [10].

4. Conclusion

According to the research of the CNN and its application in this article, the convolutional neural networks are particularly used for images classification on account of its large learning capabilities and high accuracy in visualizing results. However, the convolutional neural networks still have some possible future improvements. The convolutional neural networks are not suitable for all purposes as a result of its structural issue, but that problem can be solved through many ways such as improving the

DOI: 10.54254/2755-2721/14/20230759

dataset of the program and so on. In the third part it is told that the object detection model is a program based on the convolutional neural networks. Additionally, it is easy to be seen that the CNN is getting more accurate when detecting objects according to cases in the last part.

References

[1] T. Li, Y. Yin, Z. Yi, Z. Guo, Z. Guo, and S. Chen, "Evaluation of a convolutional neural network to identify scaphoid fractures on radiographs," Journal of Hand Surgery (European Volume), p. 175319342211270, Oct. 2022.

[2] M. Mandal, "CNN for Deep Learning | Convolutional Neural Networks (CNN)," Analytics Vidhya, May 01, 2021. https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/

[3] E. Aldhahri et al., "Correction to: Arabic Sign Language Recognition Using Convolutional Neural Network and MobileNet," Arabian Journal for Science and Engineering, vol. 48, no. 2, pp. 2615–2615, Sep. 2022.

[4] M. Liu et al., "FocusedDropout for Convolutional Neural Network," Applied Sciences, vol. 12, no. 15, p. 7682, Jul. 2022.

[5] J. Sikora, R. Wagnerová, L. Landryová, J. Šíma, and S. Wrona, "Influence of Environmental Noise on Quality Control of HVAC Devices Based on Convolutional Neural Network," Applied Sciences, vol. 11, no. 16, p. 7484, Aug. 2021.

[6] M. Walia, "Object Detection, Image Classification, Keypoint Detection," Roboflow Blog, Sep. 28, 2022. https://blog.roboflow.com/object-detection-vs-image-classification-vs-keypoint-detection/

[7] Y. Li, M. Lei, Y. Cheng, R. Wang, and M. Xu, "Convolutional neural network with Huffman pooling for handling data with insufficient categories: A novel method for anomaly detection and fault diagnosis," Science Progress, vol. 105, no. 4, p. 003685042211354, Oct. 2022.

[8] M. H. Ghaffari et al., "Deep convolutional neural networks for the detection of diarrhea and respiratory disease in preweaning dairy calves using data from automated milk feeders," Journal of Dairy Science, vol. 105, no. 12, pp. 9882–9895, Dec. 2022.

[9] M. I. Khairul Islam, R. I. Meem, F. B. Abul Kasem, A. Rakshit, and Md. T. Habib, "Bangla Spell Checking and Correction Using Edit Distance," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, pp. 1-4, 2019.

[10] G.-S. Liu, P.-Y. Huang, M.-L. Wen, S.-S. Zhuang, J. Hua, and X.-P. He, "Application of endoscopic ultrasonography for detecting esophageal lesions based on convolutional neural network," World Journal of Gastroenterology, vol. 28, no. 22, pp. 2457–2467, Jun. 2022.