Part 2: Report/Essay

Evaluate Deep learning in multimedia analysis

Executive summary

Artificial intelligence has become very popular in recent years. For example, the AlphaGo event not only popularized the latest progress of artificial intelligence to the public, but also made more people pay attention to the development of artificial intelligence. As a business analyst for a multimedia enterprise, I need to evaluate and select suitable and useful emerging analytics technologies for the organization to grow. As one of the highly recommended emerging technologies, deep learning can be used in many applications in the field of multimedia. This paper will evaluate the capabilities of deep learning in the field of multimedia and discuss data security and privacy issues in the application process.

Keywords: Artificial Intelligence · Deep Learning · Emerging Technology · Data Security · Privacy · Multimedia · Analytics Technology · Multimedia Analytics

Introduction

When artificial intelligence took off in 1956, people talked more about artificial intelligence. During the second wave, machine learning went mainstream. Now a third wave of A.I. has begun, this time with deep learning, self-learning and self-programming algorithms that can be used to solve more complex tasks. From a conceptual definition, deep learning should be a learning process that allows machines to apply what they have learned from one situation to learn new situations. Artificial intelligence is very much like people, but not like people. After being taught to recognize something, humans can recognize whether it is in a photo or not, and this process is very complicated through the brain.

In the era of mobile Internet, text, pictures, videos and other multimedia content outbreak, in order to maintain a good environment of multimedia platform, the content on the platform is very necessary to be classified, de-duplicated and quality evaluation. The basic manual audit can no longer meet the huge amount of content, so artificial intelligence algorithms based on computer vision will become a good helper for multimedia analysis. Ota, Dao & Mezaris et al (2017) mentioned in their research on mobile multimedia that deep learning is based on deep neural network, which can provide powerful recognition tools and efficiently process a large amount of data. Next, an in-depth analysis of the role of deep learning in multimedia analysis will be conducted to evaluate whether it will become the optimal choice of the company's emerging technologies.

Deep Learning in Multimedia Analytics

In 2012, AlexNet, a convolutional neural network designed by Alex Krizhevsky, emerged. It was sent to ImageNet, a prestigious object recognition competition. Before 2012, the highest accuracy of most models in this competition was about 75%, until the emergence of AlexNet (2012) improved the accuracy by about 10%. (Wei, 2019) Its appearance changed people's views on neural networks at that time and became a symbol of deep learning across The Times. Nowadays, multimedia data has become an indispensable part of big data. As an emerging technology, deep learning uses neural networks and other technologies to gradually improve the intelligence of multimedia analysis and bring convenience to related enterprises in management and operation. (Zhang,

Yao & Zhu et al., 2019)

Depth study of the application of image recognition has a lot of fun, not only can through the corresponding word to describe a piece of content is "a bird in the sky fly" image, and each term corresponds to the picture which position can be described clearly. Through this description, artificial intelligence can determine who issued what kind of task and perform a task more accurately. At present, the separation and recognition of massive images can be widely used. For example, Taobao, a Chinese online shopping software, has a function that allows consumers to know the price of a product by taking a picture of it with a mobile phone, which is convenient for consumers to purchase. In addition to recognition, retrieval can also be done by tagging photos with information. In addition, one of the more recent trends in AI, the generation of images from text, is also achieved through deep learning.

Under the influence of endless algorithms in recent years, the field of optical character recognition (OCR) has been the world of deep learning. The main role of OCR is to extract and recognize the content of images or videos, which is a very useful technology for our company. We can use OCR to supplement content features, such as adding titles to untitled content based on cover images and adding descriptions to products using information from images. In addition, we can also use OCR to identify product names and work names, identify advertising copies, and identify and review sensitive words.

Data Security and Privacy Issues

With the deepening of deep learning, the emergence of many applications has paid more and more attention to the problems caused by deep learning technology(Xu, Li & Ren et al., 2019), such as who is responsible for the behavior of machines in the event of an accident.

The craze for face-swapping apps began a few years ago with the launch of Deepfake, which replaces a face in one image with another. Not content with changing faces, some people began to use machines to generate unreal images of faces and photos of rooms, and even use contour information extracted from faces to make Mona Lisa and Einstein move. In the past, we always said that seeing is believing, but now some images generated by

artificial intelligence can be made to be true, leading to a decrease in people's trust in the authenticity of information (Agarwal & Varshney, 2019), which also leads to some controversial social issues such as social security. Therefore, the governance and ethical aspects of AI are also becoming more and more important.

Conclusions

In this paper, we have a deeper understanding of the role of deep learning in the field of multimedia and can find that deep learning can help the company in multimedia analysis and improve efficiency. However, as there are still some problems in deep learning that may lead to privacy disclosure or mislead readers and audiences, more attention should be paid to this aspect in the process of use to avoid potential problems as much as possible.

References

Ota, K., Dao, M. S., Mezaris, V. & Natale. F. De., 2017, Deep learning for mobile multimedia: A survey. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 13, no. 3s, pp. 34.

Zhang, W., Yao, T., Zhu, S. and Saddik, A.E., 2019. Deep learning–based multimedia analytics: a review. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 15(1s), pp.1-26.

Wei, J., 2019, 'AlexNet: The Architecture that Challenged CNNs', Towards Data Science, weblog post, 3 July, viewed 5 August 2022, <https://towardsdatascience.com/alexnet-the-architecture-that-challenged-cnns-e406d5297951 >.

Agarwal, S., & Varshney, L. R., 2019. Limits of deepfake detection: A robust estimation viewpoint, arXiv preprint arXiv:1905.03493, viewed 5 August 2022, < https://arxiv.org/abs/1905.03493>

Xu, G., Li, H., Ren, H., Yang, K., & Deng, R. H., 2019, Data security issues in deep learning: attacks, countermeasures, and opportunities. IEEE Communications Magazine, 57(11), pp. 116-122.