Current Trends in Deep Learning

Yushu Yang
New York University, New York, 10012, United States

Abstract. This paper provides an overview of the current artificial intelligence (AI) and machine learning (ML) techniques, including Convolutional Neural Networks(CNN), Adversarial and Generative Techniques, Natural Language Processing (NLP), and Reinforcement Learning(RL). The paper discusses the background, applications, and future trends of these techniques, highlighting their potential for solving real-world problems. The purpose of the paper is to illustrate the trends that are emerging in these areas, as well as the challenges that must be addressed in order to fully realize their potential. By identifying several key areas, this paper concludes the future research and development.

Keywords: artificial intelligence, deep learning, machine learning, trends and development.

1. Introduction

In recent years, the field of machine learning has witnessed a rapid development with the advent of deep learning techniques and the availability of large datasets. Deep learning, first introduced by LeCun in 2015, has evolutionized the field by enabling the creation of more complex and accurate models. These models have found numerous applications in various fields, such as self-driving cars. This paper will discuss the fundamentals of deep learning, including convolutional neural networks, generative adversarial neural networks, natural language processing, and reinforcement learning. The current and future trends of these techniques will also be examined.

2. Convolutional Neural Networks

2.1 Background of Convolutional Neural Networks

Convolutional Neural Networks (CNN) have become an increasingly popular machine learning model for dealing with large input data with a huge amount of dimensions. With CNN, the neural networks can be generated more easily. “An image, for example, comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image. The second layer typically detects motifs by spotting particular arrangements of edges, regardless of small variations in the edge positions. The third layer may assemble motifs into larger combinations that correspond to parts of familiar objects, and subsequent layers would detect objects as combinations of these parts” (LeCun). In the paper of Deep Learning, LeCun introduced the specific roles of a multi-layered CNN model. Images, one form of data input, are dealt in three layers, each representing different types of the image information. CNNs work by spatially constraining and tying the weights for a given kernel in the hidden layers. It convolves the filter kernel with the input data by sliding the whole space to compute its dot products, and then forms activation maps to ultimately produce output data as the input of other neural networks. In this way, the representations are smaller and more manageable.

Since AlexNet, the first CNN proposed in 2012, CNN models have continuously enhanced their performance to reduce errors, as shown in Figure 1. By 2016, the error rate had decreased from 16.4% to 13.3% with the GoogLeNet-v4. The models have primarily focused on employing ample hidden layers, leading to the trend of adding more layers for lower errors. Infinitely many layers are good for accuracy. However, this can result in a significant amount of storage required to store the layers and maps. Hence, a more reliable future trend on the system side would be to find the most efficient way to store neural networks, enabling CNN run faster and train more effective

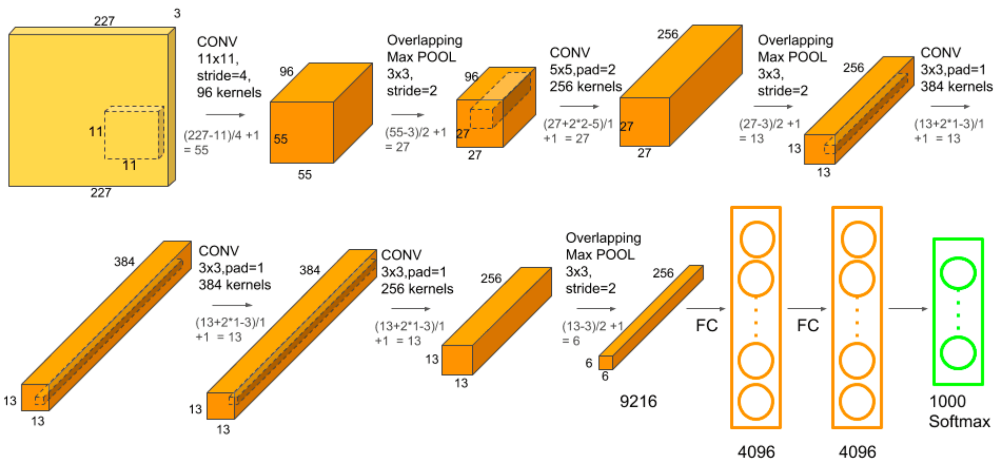

Figure 1. AlexNet which won the 2012 ImageNet LSVRC-2012 competition with 15.3% error rates.

2.2 Applications of CNN

Convolutional Neural Networks (CNN) are widely used for images processing tasks such as colorization, face and emotion detection, and diseased cell recognization. The methods used for image classification consume training data. However, collecting sufficient training data is a challenge undoubtedly. To address this, researchers are exploring unsupervised methods to identify localized regions. CNN has also been applied to learn abstract human body poses, faces and emotions. For example, DeepPose is a CNN-based regression model that estimates body pose through body joint coordinates. By taking the overall body pose image as input and analyzing each body joint, DeepPose can accurately estimate body poses. In addition, CNN can also be used for body pose estimation in videos. Jain et al. incorporated RGB and motion features into a multi-resolution CNN architecture to improve accuracy. The input is a RGB image and its corresponding motion features. The output includes response maps of the joints that represent the energy for the presence of the joint at the pixel location (Gu, Jiuxiang, et al). Overall, CNN has diverse applications in image processing, including body pose estimation, and holds great potential for future research.

2.3 Trends of CNN

Despite significant advancements made in the field of CNN, there is still much room for improvement in terms of efficiency. The traditional approach of increasing the depth of the network for higher accuracy requires a large number of labeled training datasets, making the process time-consuming and laborious. Therefore, one of the major emerging trends in CNN is the development of unsupervised learning techniques that can operate with minimal labeled data. Additionally, to improve the efficiency of the training processes, researchers are exploring alternative algorithms that utilize CPU and GPU clusters beyond the popular stochastic gradient descent (SGD) algorithm (Gu, Jiuxiang, et al), which reduce the complexity of the process without sacraficing accuracy. Furthermore, hyper-parameters selection, including the learning rate, kernel sizes of convolutional filters, and the number of layers is also a key focus for further investigations that aim to improve current optimization techniques for CNN. These efforts are expected to enhance the efficiency of CNN models in real-world applications.

3. Adversarial and Generative Techniques

3.1 Background

Generative Adversarial Neural Networks (GANs) are a popular deep learning method. The main algorithm is to design two neural networks, through competing with each other to improve their predictions. The two agents are the generator and the discriminator, and their optimization is based on a zero-sum game, where one agent's gain is the other agent's loss. The generator is a convolutional neural network that produces synthetic data that closely resembles real data, so it matches its output distribution close to the reference distribution. The discriminator, on the other hand, is a de-convolutional neural network that distinguishes between real and synthetic data, outputting a value close to 1 when input appears to be from the reference distribution and 0 when input is from the generator distribution. GANs are known for their ability to generate high-quality outputs.

3.2 Applications

GANs have many useful applications, especially for visual content. They can be used to can fill in images from an outline, generate a realistic image from text descriptions, produce photorealistic depictions of product prototypes, and convert black-and-white imagery into color. In addition, GANs are feasible to be applied to video productions, such as modeling patterns of human behaviors and movement within a frame, generating subsequent video frames, and creating deepfake videos. With such diverse applications, GANs are becoming increasingly important in various fields, including entertainment, advertising, and design. The ability of GANs to generate realistic and high-quality data has opened up new possibilities for creating and enhancing digital content.

3.3 Current limitations and future trends of GAN

GANs have become a powerful tool in the field of deep learning, but it also has some limitations that need to be addressed for further development. One of the challenges is the low-diversity of generated outputs when the generator maps several different inputs to the same output. While there have been proposed solutions, more research is needed in this area to improve GAN's output quality. Additionally, when the generator produces highly accurate outputs, it becomes difficult for the discriminator to distinguish between real and fake data, which can result in a decline in overall performance, making the existence of the discriminator meaningless. As a result, GAN's output quality ultimately decreases. Further research can be conducted to address this fundamental problem. Another challenge for GANs is the computational intensity, a general limitation of many other neural networks as well. Since GAN requires a generator and a discriminator, the repetitive process makes the whole algorithm intensive in its computations and is time-consuming. Furthermore, GAN-based tools currently lack creativity when used for designing purposes (Hughes, Rowan T., et al.). GAN generator can only generate images based on its training data, so it cannot generate new images that it has not seen before. In the field of design, this should be considered as a limitation and need human interventions for more innovative outputs. Future developments should consider these limitations and continue to explore ways to improve GAN's performance and expand its applications.

In the field of pathology, morphology translation has great potential with the development of GANs, according to Tschuchnig, M. E.. Image-to-image translation involves mapping from one domain to another, and the difference between the domains is a result of variability in the generation process. However, detecting changes in morphological can be difficult due to the underlying tissue similarity (Tschuchnig, M. E.). Research has been conducted on translating data from different classes for data augmentation. A derivation of cycleGAN (Figure 2 and Figure 3) is used to differentiate the generated image and the real image. However, the recent architecture is not optimal for changed morphology, as it can lead to ambiguous mappings. Therefore, future developments in GANs should focus on developing more optimized architectures, such as translation between different imaging settings, to improve image qualities and classification accuracy. These advancements in GAN technology could greatly improve the efficiency and effectiveness of pathology diagnoses, leading to better patient outcomes.

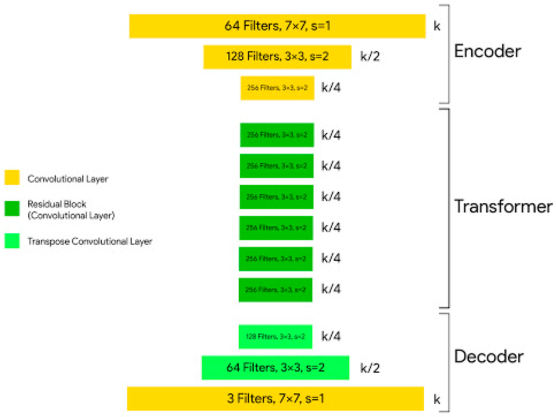

Figure 2. Each CycleGAN generator has three sections: encoder, transformer, and decoder.

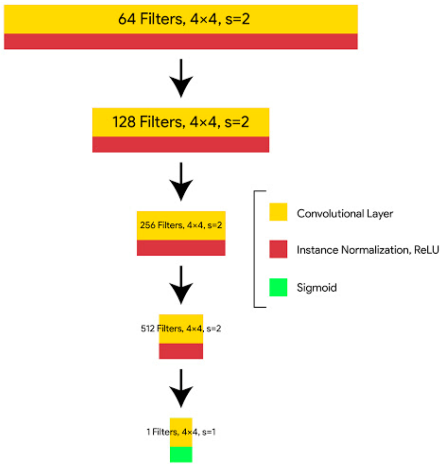

Figure 3. CycleGAN also has a discriminator that maps from a 256×256 image to a single scalar output or an NxN array of outputs X, which signifies “real” or “fake”.

Furthermore, GANs pose ethical concerns related to falsification. As the technology for generating images and contents becomes more advanced, it becomes increasingly difficult for people to differentiate between real and fake images. For instance, this issue becomes particularly problematic in cases of forged signatures or manipulated visual media. As new technologies continue to emerge, new ethical challenges will arise, and it is essential for future trends to be more focused on developing new technologies while addressing these ethical issues. A balance must be struck between advancing technology and ensuring that it is not used to deceive or manipulate individuals or society as a whole. Therefore, it is crucial for researchers and developers to be aware of these ethical implications and actively work towards developing GANs in a responsible and ethical manner.

4. Natural Language Processing

4.1 Background of Natural Language Processing

Natural language processing (NLP) is a critical subfield of machine learning with diverse applications across various domains. NLP deals with the interactions between computers and human language, with the primary objective of understanding language content and accurately distinguishing the hidden contextual nuances. Through NLP techniques, computers can analyze, interpret, and generate human language, enabling them to perform tasks such as language translation, sentiment analysis, chatbot development, and text summarization. The use of NLP has become increasingly important with the rise of new technologies, including voice assistants, chatbots, and smart devices, making it an essential component of the machine learning field.

4.2 Trends for NLP

The major trend of natural language processing (NLP) is the focus on the cognitive aspect. As a combination of cognition and NLP, language can be processed more abstractly and systematically that emulate intelligent behaviors. Cognition means acquiring knowledge and understanding through thought, experience, and senses. This helps with understanding the different meanings of the same words in different contexts. NLP algorithms are being developed to better understand the contextual nuances of language, which is essential for accurate interpretation of meaning.

One key application of NLP is in intent-less AI assistance. The primary challenge of conversational AI assistance is accurately understand user intents. Languages are currently processed through a large training sample of user intents and responses are given without actually understanding the content. With large training data of human conversations available, responses can be generated by understanding what the user actually means rather than simply matching keywords, thus making user intent-driven training a secondary factor. This has great potential for improving the efficiency of AI service desk by reducing the repetitive works of human interventions. It goes through the content, distinguishes the category, has some basic analysis, and then decides the resolution mechanism. Therefore, the user’s problem can be solved virtually and efficiently. Moreover, more and more enterprises have shifted to hybrid mode after the pandemic. Therefore, enterprises need to keep employees engaged. They use their IT systems to solve employee queries. NLP can make the query process efficient and accurate, thus becoming another potential for further development.

Another area of focus in NLP is Natural Language Generation (NLG). Given a dataset, the algorithm first filters the main topic. Then, it uses machine learning models to understand and interpret the data, and gathers relevant content to summarize the topic. Finally, it can generate grammatically correct sentences by rewriting the sentences. Further research is needed to improve each step of the NLG process for more effective and accurate results.

5. Reinforcement Learning

5.1 Background of Reinforcement Learning

Different from supervised and unsupervised learning, reinforcement learning does not require labeled input or output. It involves finding a balance between exploring uncharted territory and exploiting current knowledge. The framework involves an agent interacting with the environment and strengthening or weakening its behavior based on the rewards it receives. This approach is similar to how animals learn and behave.

5.2 Applications

Reinforcement learning has numerous real-world applications, with one of the most notable examples being AlphaGo. Created by DeepMind, AlphaGo defeated human players by utilizing two neural networks. “One neural network, the ‘policy network’, selects the next move to play. The other neural network, the ‘value network’, predicts the winner of the game” (DeepMind). Two neural networks becomes powerful as they take actions according to the feedback it gets. This basic reinforcement learning framework has been applied to other fields such as robotics and self-driving.

5.3 Self-driving and its trends

Self-driving technology has been one of the most popular topics in the field of AI and machine learning for many years. However, due to safety concerns, it has not been fully implemented. The most challenging aspect of self-driving cars is handling difficult-to-predict situations such as sharing the road with erratic and distracted human drivers. To address this issue, Automated Vehicles companies are focusing on developing more extensive databases of trainable experiences through both simulation and on-road testing. Additionally, there is a shift in focus from robo-taxis to automated trucks and delivery vehicles. Although robo-taxis are not as demanding as before, but their autonomous techniques can be applied to delivery, which still has high demand. Another trend is improving performance in dense urban areas. While most current pilots are in suburban areas, urban areas present more difficulties due to smaller spaces and complicated roads. Future research on self-driving will focus on overcoming challenges in dense cities.

6. Conclusion

This paper presents an overview of several topics in the field of artificial intelligence and machine learning, including convolutional neural networks, generative and adversarial techniques, natural language processing, and reinforcement learning. The current trend in these areas is to simplify complex models while maintaining accuracy, and to develop applications that meet the needs of the users. In the future, more efficient training algorithms for both CNNs and GANs, and better optimization techniques for hyperparameters are some of the key areas that researchers can focus on. With the continued growth of data and advances in deep learning, the possibilities for applications of these techniques are endless. The key to success in this field will be to stay up-to-date with the latest trends and innovations and to remain committed to the ethical considerations that arise with the use of these powerful technologies.

References

[1] Gu, Jiuxiang, et al. "Recent advances in convolutional neural networks." Pattern recognition 77 (2018): 354-377.

[2] Alexnet - Neurohive.io. https://neurohive.io/en/popular-networks/alexnet-imagenet-classification-with-deep-convolutional-neural-networks/.

[3] Hughes, Rowan T., Liming Zhu, and Tomasz Bednarz. "Generative adversarial networks–enabled human–artificial intelligence collaborative applications for creative and design industries: A systematic review of current approaches and trends." Frontiers in artificial intelligence 4 (2021): 604234.

[4] Tschuchnig, Maximilian E., Gertie J. Oostingh, and Michael Gadermayr. "Generative adversarial networks in digital pathology: a survey on trends and future potential." Patterns 1.6 (2020): 100089.

[5] “Cycle Generative Adversarial Network (Cyclegan).” GeeksforGeeks, GeeksforGeeks, 23 June 2022, https://www.geeksforgeeks.org/cycle-generative-adversarial-network-cyclegan-2/.

[6] “AlphaGo.” DeepMind, https://www.deepmind.com/research/highlighted-research/alphago.