International Journal of Innovative Research in Technology and Science
ISSN: 2321-1156 www.ijirts.org Volume 12 Issue 2, March 2024

A Review of various Deep Learning Models for Fire Detection

Puneet Nema, Sunil Phulre
LNCTU, Bhopal, India
Puneetnema11@gmail.com, sunilphulre@lnct.ac.in

Selection and peer review of this article are under the responsibility of the scientific committee of the International Conference on Current Trends in Engineering, Science, and Management (ICCSTEM-2024) at SAM Global University, Bhopal.

Abstract: In recent years, the deep learning (DL) computing paradigm has emerged as the gold standard within the machine learning (ML) community. Gradually, it has become the predominant ML computational approach, yielding remarkable results across various intricate cognitive tasks, often rivalling or surpassing human performance. One of the paramount advantages of DL is its capacity to glean insights from massive datasets. Machine learning tools, constituting algorithmic applications of artificial intelligence, endow systems with the capability to learn and enhance themselves with minimal human intervention. Concepts such as data mining and predictive modelling align closely with this paradigm, facilitating software to refine its predictive accuracy without explicit programming. This paper introduces convolutional neural networks (CNNs), the most prevalent DL network type, and delineates the evolution of CNN architectures alongside their salient features. The progression is elucidated by commencing with seminal works like the AlexNet network and culminating with advanced architectures such as the High-Resolution network (HR.Net).

Furthermore, the paper discusses prevalent challenges encountered in DL research and proposes potential solutions to bridge existing gaps in understanding. Subsequently, a compendium of major DL applications is presented. The computational tools underpinning DL, including FPGA, GPU, and CPU, are summarized, underscoring their pivotal role in advancing DL methodologies.

Keywords: Convolutional Neural Networks, Fire Detection, Machine Learning (ML), Deep Learning, Artificial Intelligence (AI)

1. INTRODUCTION

Deep Learning (DL) offers the remarkable ability to process vast amounts of data, making it a pivotal technology that has rapidly evolved in recent years. Its versatile applications span various domains, including cybersecurity, natural language processing, bioinformatics, robotics and control, and medical information processing. Notably, DL has surpassed traditional Machine Learning (ML) techniques in performance across numerous domains [1]. Despite the substantial volume of research reviewing the state-of-the-art in DL, existing works often focus on specific aspects, resulting in a fragmented understanding of the field. This paper adopts a comprehensive approach to address this gap to provide a foundational understanding of DL.

Additionally, it is essential to recognize that machine learning and deep learning models encompass different learning paradigms, including supervised learning, unsupervised learning, and reinforcement learning. Supervised learning relies on labelled datasets for categorization or prediction tasks, necessitating human intervention for accurate labelling. In contrast, unsupervised learning identifies patterns within unlabeled datasets, clustering data based on inherent characteristics. Deep learning is a subset of machine learning that operates within these paradigms [2,3].

2. VARIOUS DEEP LEARNING MODEL

Machine Learning Algorithm

Machine learning algorithms utilize structured, labelled data to make predictions. This means that specific features are defined from the input data and organized into tables. However, they can also process unstructured data with some preprocessing to organize it into a structured format [2]. In contrast, deep learning streamlines the data preprocessing typically required in machine learning. These algorithms can directly ingest and process unstructured data such as text and images, automating feature extraction and reducing the need for human intervention. For instance, suppose we have a collection of photos depicting various pets, and we aim to categorize them into "cat," "dog," "hamster," and so on. Deep learning algorithms can identify which features (e.g., ears) are crucial for distinguishing between different animals. In machine learning, human experts usually establish this feature hierarchy manually [3, 8]. Subsequently, through processes like gradient descent and backpropagation, the deep learning algorithm fine-tunes itself for accuracy, enabling it to make precise predictions about new animal photos.

Reinforcement Learning Model

Reinforcement learning is a process in which a model learns to become more accurate in acting in an environment based on feedback to maximize the reward.-1 shows different algorithm comparisons and conclusions cumulatively from different authors in past years [2]. The deep Learning Kit also supports convolutional neural networks. Its vision is to support other deep learning tools like Torch and Tensor Flow [1]. DL is derived from the conventional neural network but considerably outperforms its predecessors.

Moreover, DL simultaneously employs transformations and graph technologies to build multi-layer learning models. The most recently developed DL techniques have obtained outstanding performance across various applications, including audio and speech processing, visual data processing, and natural language processing (NLP) [1, 2]. Feature extraction is achieved automatically throughout the DL algorithms. This encourages researchers to extract discriminative features using the smallest possible human effort and field knowledge [1, 10]. These algorithms have a multi-layer data representation architecture, in which the first layers extract the low-level features while the last layers extract the high-level features. Note that artificial intelligence (AI) originally inspired this type of architecture, simulating the process in core sensorial regions within the human brain. Using different scenes, the human brain can automatically extract data representation. More specifically, the output of this process is the classified objects, while the received scene information represents the input. This process simulates the working methodology of the human brain. Thus, it emphasizes the main benefit of DL [7, 9].

Convolutional neural network

Due to its considerable success in the ML field, DL is currently one of the most prominent research trends. This paper presents an overview of DL and adopts various perspectives, such as the main concepts, architectures, challenges, applications, computational tools, and evolution matrix. Convolutional neural networks (CNN) are among the most popular and used DL networks [3, 4]. Because of CNN, DL is very popular nowadays. Compared to its predecessors, the main advantage of CNN is that it automatically detects significant features without human supervision, making it the most used. Therefore, we have dug deep into CNN by presenting its main components.

Furthermore, we have elaborated in detail the most common CNN architectures, starting with the AlexNet network and ending with the High-Resolution network (HR.Net) [6].

Several published DL review papers have been presented in the last few years. However, all of them have only been addressed one side focusing on one application or topic, such as the review of CNN architectures [2], DL for classification of plant diseases [2], DL for object detection [3], DL applications in medical image analysis [4, 11], etc. Although these reviews present good topics, they do not fully understand DL topics, such as concepts, detailed research gaps, computational tools, and DL applications. First, It is required to understand DL aspects, including concepts, challenges, and applications, and then go deep into the applications. To achieve that, extensive time and a large number of research papers are required to learn about DL, including research gaps and applications. Therefore, we propose a deep review of DL to provide a more suitable starting point for developing a full understanding of DL from one review paper. The motivation behind our review was to cover the most important aspects of DL, including open challenges, applications, and computational tools perspective.

Furthermore, our review can be the first step towards other DL topics [2]. This review aims to present the most important aspects of DL to make it easy for researchers and students to have a clear image of DL from a single review paper. This review will further advance DL research by helping people discover more about recent developments in the field. Researchers could decide on the most suitable work direction to provide more accurate alternatives to the field. Our contributions are outlined as follows: This is the first review that almost provides a deep survey of the most important aspects of deep learning. This review helps researchers and students to have a good understanding of one paper [2]. We explain CNN, the most popular deep learning algorithm, by describing the concepts, theory, and state-of-the-art architectures [3]. We review current challenges (limitations) of Deep Learning, including lack of training data, Imbalanced Data, Interpretability of data, Uncertainty scaling, Catastrophic forgetting, Model compression, Overfitting, Vanishing gradient problem, Exploding Gradient Problem, and Underspecification. We additionally discuss the proposed solutions to tackling these issues [3,4]. We provide an exhaustive list of medical imaging applications with deep learning by categorizing them based on the tasks, starting with classification and ending with registration [3]. We discuss the computational approaches (CPU, GPU, FPGA) by comparing the influence of each tool on deep learning algorithms [5].

Colour-Based Segmentation of U-Net

Deep Learning (DL)--based image processing is now used across various industries. In particular, methods like image classification (i.e., identifying what an image represents) and object detection (i.e., identifying an object and its location in the image) are being applied in use cases ranging from computer vision applications (e.g., autonomous driving) to identifying people in security videos) [3]. Semantic image segmentation can be achieved using a U-Net, a special Convolutional Neural Network (CNN) type. A U-Net adds an expansive path to generate classifications of the pixels belonging to feature(s) or object(s) found in the source image. In other words, it expands the output to a certain image size and forms the latter part of the U in the network. A U-Net allows us to go above and beyond normal image classification and object detection to classify the pixels of those objects in their exact shape [3]. To understand U-Nets, let's briefly review how their underlying foundation, the standard CNN, works [4]. We emphasize standard because there are several variations, including LeNet, AlexNet, and other CNNs, which all work off the same general principle of taking an image as input and convolving it using filters (aka kernels) to extract one or more feature maps, a process known as feature extraction. These feature maps are then down-sampled through pooling and passed to the next layer for further convolving and pooling. The feature map(s) produced in each subsequent convolving/pooling step extract higher-level features. The final down-sampled (pooled) feature map is then flattened and used as input to a fully connected neural network [4].

VGG-16 Model

VGG16 is a convolution neural net (CNN) architecture used to win the ILSVR(Imagenet) competition 2014. It is considered one of the most excellent vision model architectures to date. The unique thing about VGG16 is that instead of having a large number of hyper-parameters, they focused on having convolution layers of 3x3 filter with a stride one and always used the same padding and maxpool layer of 2x2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. It has 2 FC(fully connected layers) and a softmax for output [4]. The 16 in VGG16 refers to 16 layers that have weights. This network is pretty large and has about 138 million (approx) parameters [3,4].

Model Checkpoint helps us save the model by monitoring a specific model parameter. In this case, validation accuracy is monitored by passing val -acc to Model Checkpoint. The model will only be saved to disk if the validation accuracy of the model in the current epoch is greater than what it was in the last epoch. Early Stopping helps us to stop the model's training early if there is no increase in the parameter set to monitor Early Stopping. In this case, validation accuracy is monitored by passing val_acc to Early Stopping. Here, set patience to 20, meaning the model will stop training if it doesn't see any rise in validation accuracy in 20 epochs [4,5]. Using model.Fit generator: using Image Data Generator to pass data to the model and passing train and test data to fit generator. In fit generator steps per epoch will set the batch size to pass training data to the model, and validation steps will do the same for test data. You can tweak it based on your system specifications [4,5].

ReSeg Models

We propose a structured prediction architecture that leverages the local generic features extracted by Convolutional Neural Networks (CNNs) and the capacity of Recurrent Neural Networks (RNNs) to capture distant dependencies. The proposed architecture, ReSeg, builds upon the recently introduced ReNet model for image classification. We adapt and expand it to tackle the more complex semantic segmentation task. Each ReNet layer comprises four RNNs that traverse the image horizontally and vertically in both directions, encoding patches or activations and providing relevant global information. Furthermore, ReNet layers are stacked atop pre-trained convolutional layers, thereby harnessing generic local features. Following the ReNet layers, upsampling layers are employed to restore the original image resolution in the final predictions. The proposed ReSeg architecture is characterized by its efficiency, flexibility, and applicability to various semantic segmentation tasks. We evaluate ReSeg on several widely-used semantic segmentation datasets, including Weizmann Horse, Oxford Flower, and CamVid, achieving state-of-the-art performance. Our results demonstrate that ReSeg is a suitable architecture for semantic segmentation tasks and holds promise for addressing other structured prediction problems [5].

3. CONCLUSION

Much of the AI we encounter today operates on deep learning: a machine is presented with a dataset and a desired output, generating its algorithm to achieve this. The process repeats, perpetuating itself. While both machine and deep learning analyze data and learn from it, deep learning specifically endeavours to mimic the functions of the human brain when concluding. Deep learning employs artificial neural networks to execute complex computations on vast datasets. It represents a branch of machine learning inspired by the structure and function of the human brain. Deep learning algorithms train machines by learning patterns within data. The primary objective of deep learning is continual improvement with each new piece of data. This involves the ability to adapt its underlying structure to interpret data accurately. Once the network is thoroughly developed using test data, greater personalization through customer analytics is enabled. Deep learning techniques shine in scenarios lacking domain expertise for feature introspection as they require less concern about feature engineering. Deep learning excels in solving complex problems such as image classification, natural language processing, and speech recognition.

REFERENCES

[1]. Goodfellow, Ian, et al. "Deep learning." Nature 521.7553 (2015): 436-444.

[2]. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." Nature 521.7553 (2015): 436-444.

[3]. Deng, Li, and Dong Yu. "Deep learning: methods and applications." Foundations and Trends® in Signal Processing 7.3–4 (2014): 197-387.

[4]. Schmidhuber, Juergen. "Deep learning in neural networks: An overview." Neural Networks 61 (2015): 85-117.

[5]. Hinton, Geoffrey, et al. "Deep neural networks for acoustic modelling in speech recognition: The shared views of four research groups." IEEE Signal Processing Magazine 29.6 (2012): 82-97.

[6]. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification

with deep convolutional neural networks", Advances in neural information processing systems, 25 (2012): 1097-1105.

[7]. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[8]. He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[9]. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

[10]. Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[11]. Jin, Shuting, et al. "A self-adaptive wildfire detection algorithm by fusing physical and deep learning schemes." International Journal of Applied Earth Observation and Geoinformation 127 (2024): 103671.