UoEO Computing Department

Deep Learning Techniques - Literature Review

1. Introduction

Since the beginning of computers, we have seen tremendous developments in hardware and software, and we have completely changed the way we work. Every new decade brings ideas and prototypes that were previously thought impossible. Deep Learning is a technology capable of disrupting the way we collaborate with computers in the most exciting way.

This literature review is developed with a theoretical framework and will attempt to criticise the available literature of Deep Learning and the techniques that allow its utilisation. Additionally, it will attempt to pinpoint the way Deep Learning evolves; from the time researchers started developing it, up to the sectors we have applied it to today.

Finally, this review will try to identify the literature gaps found in the current literature to create the foundation for developing a solution proposal.

1.1 Literature's Review Audience

This literature review is intended to provide an outline of sources and current knowledge of the Deep Learning techniques used for processing data. The literature review will then be presented to the author's tutor and any further academic staff of the university.

The final MSc project will be targeted to IT professionals interested in expanding their knowledge in Deep Learning technology. Furthermore, IT students interested studying in

Deep Learning technology could consult the MSc project document for the theory gathered and its results.

1.2 Sources Collection Methodology

Before starting my actual academic research, I read articles considering the basic theory of Deep Learning, its techniques and applications, and its history. This method provided me with an insight into Deep Learning AI technology.

Following this initial research, I started searching with keywords (e.g., Neural networks, RNN, CNN, big data, etc.) on well known academic websites such as Google Scholar, IEEE computing articles, Google Books, the University of Essex Library for relevant up-to-date sources. In addition, I read individual research papers and journals uploaded to academia.edu and ResearchGate.

Further to my research, I reviewed the citations and references of these articles and journals to recognise the most common authors of the Deep Learning topic.

1.3 Literature Review Structure

This literature review is divided into three parts: the introduction, the available literature, the limitations, and the conclusion section.

In the introduction, I presented a preface of the literature's review topic, the aim of this review and a description of my research.

In the main body, I presented a part of the theoretical framework of the topic to provide the reader with a basic idea of the topic. I included the basic Deep Learning theory, along with the most common techniques used for initiating this technology, some of the applications of Deep Learning, its history, and its limitation.

In the conclusion, I presented the main findings and a summary of the available literature.

This literature review covers the topic of deep learning techniques rather than focus on a specific application of it.

2. Available Literature

Deep Learning is a subcategory of the artificial intelligence machine learning, which consists of architectures and algorithms acting just like a function of our brain called artificial neural networks (IBM, 2020). Deep Learning technology is divided into three or more layers, allowing it to "learn" and export information from large data (Dunn, 2021). For the neural network to operate, a user feeds the algorithm with an object (e.g., a picture) which the network then forwards to all its layers for processing. To "train" the machine, users need to feed it with large amounts of data to create classifications which then the system uses to recognise objects (Dunn, 2021). The more layers a neutral network has, the more accurate its predictions will be (IBM, 2020).

Deep Learning is separated from Machine Learning because machine learning operates mostly with structured data, while Deep Learning operates and analyses mostly unstructured data (Oprea et al., 2020). Structured data are clearly defined data organised in a searchable manner. These data include text, numbers, spreadsheets, databases (SQL), etc. (Kapoor, 2019). Unstructured data are data that are not structured in a predefined way and are hard to analyse. These data include audio files, videos, CCTV footage, etc. (Eberandu, 2016).

2.1 Deep Learning Architecture and Algorithms

Let's consider the Deep Learning technology as a basic Neural Network. The Deep Learning architectures are based on the different types of neural networks available today and allow the expansion and usage of Deep Learning technology from various sectors (Lisowski, 2020).

Deep Learning algorithms are commands and codes that allow the data to flow within the several “layers” of a Deep Learning architecture and provide the user with results. Simply put, Deep Learning algorithms utilise each architecture used and allow the networks to “run” (Shrestha & Mahmood, 2019).

2.2 Deep Learning architectures

Based on my research, multiple architectures were found to initiate a deep learning network. These include the most common technique which is using a neural network (NN), the recurrent neural network (RNN), the convolutional neural network (CNN), the long short-term memory (LSTM), and more. Additional architectures include Autoencoders, self-organising maps, etc. (Chen, 2021). Each model is developed with different functionalities to allow developers to use the more appropriate one for their solution (Vadapalli, 2020).

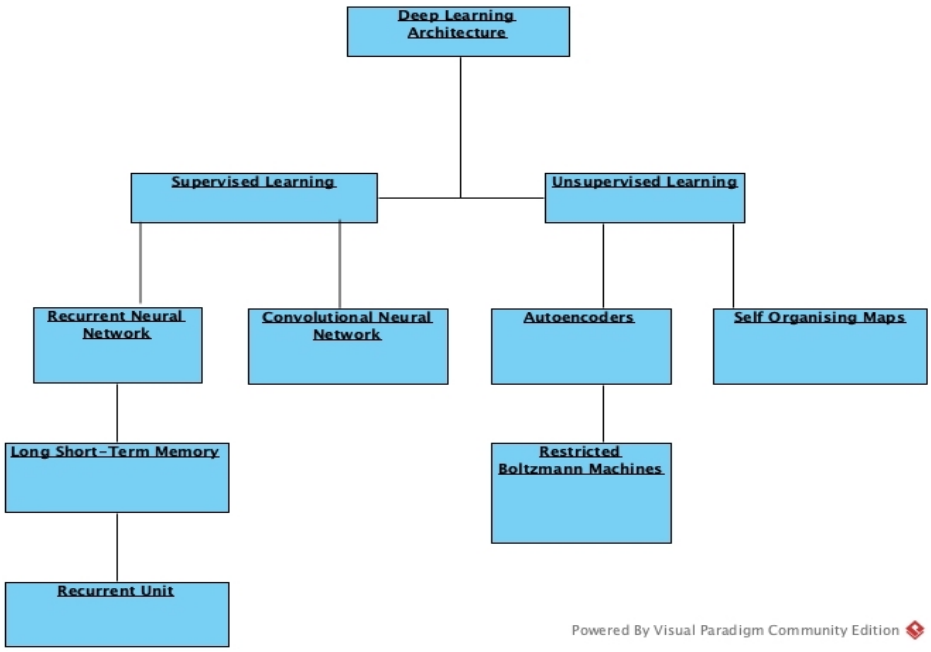
The first Convolutional Neural Networks (CNN) was developed in 1989 by Yann LeCun and it was able to recognise hand-written digits. Convolutional Neural Networks consist of multiple layers and are designed mostly for image processing and object detection (Mathew and Sivakumari, 2020).

The Long Short Term Memory (LSTM) network is part of the Recurrent Neural Network (RNN) and is used to memorise previously inputted dependencies. They can predict time

series and they are mainly used for speech recognition, music composition and by pharmaceutical companies for medicine development (Mathew and Sivakumari, 2020).

The Recurrent Neural Network (RNN) consists of direct cycles and transforms the outputs from the Long Short Term Memory (LSTM) network into inputs to the current phase. RNNs are mostly used for machine translation, natural-language processing, and time-series analysis (Mathew and Sivakumari, 2020).

Diag. 1 *Deep Learning Architecture*



Convolutional Neural Networks (CNN)

The Convolutional Neural Networks (CNN) are used for their ability to process pictures, audio inputs and speech. They are divided into three layers: Convolution layer, Pooling layer and Fully-connected layer.

The Convolution layer is the main layer where most of the computation occurs. It requires the user inputted data, a filter, and a feature map (IBM, 2020). All three parameters allow the system to create a "neuron" that will compute the data.

The Pooling layer will receive the data from the convolution layer and will work as a filter to reduce the input (i.e., previously output) to the number of the parameter of the original input to provide a unified result (IBM, 2020).

Finally, the fully-connected layer works as a classification system that connects each calculated output to the user input allowing the results to be displayed within the same format (IBM, 2020).

Recurrent Neural Network (RNN)

The Recurrent Neural Network (RNN) are a type of Deep Learning architecture that allow a neural network to select inputs from previous outputs (Amidi & Amidi, n.d). They are used for ordinal or temporal problems like speech recognition, natural language processing, image captioning, etc. and are highly used by voice assistants such as Siri, Google Now, etc. (IBM, 2O20). RN networks are divided into the input layer, the hidden layers, and the output layer (IBM, 2020). RN networks operate using a Feedforward Neural Network where the data flow in one input without being able to roll back. This limits the network to operate only with logical inputs (IBM, 2020).

Comparison of Convolutional Neural Networks (CNN) and Recurrent Neural Network (RNN)

Both architectures are commonly used. CNN are mostly used for spatial data such as audio and pictures while RNN are mostly used for sequential data such as text.

Additionally, in CNN the size of the input to output is fixed, while in RNN it can vary (Petersson, 2021).

CNN is considered to be more powerful mostly because it allows for wider utilisation of data (Vadapalli, 2021).

Nowadays, we live in an era of a data revolution. Every day, more and more data are generated and because of the low cost of storage devices, governments and companies can store huge amounts of information in data warehouses (Goodfellow et al., 2016). Deep learning along with Big Data technologies can allow computer and software engineers to easier process these unstructured data and create information useful for the organisations (Wang and Alexander, 2016).

Furthermore, by "training" a deep learning network, a software engineer is creating a way to process data in a completely unsupervised way. This can result in new useful information, or in successfully predicting outcomes (Murnane, 2016).

2.3 Applications of Deep Learning

Deep Learning is an all-around technology providing application to various sectors. Some examples include Virtual Assistants, Streaming Services, Fake News Detection, image colouring, image captioning, etc. (Wani et al., 2020).

Virtual Assistants are cloud-based applications able to understand human voices and complete tasks for the users. These applications include Amazon Alexa, Siri, Google Assistant, etc. (IBM, 2020).

Streaming services, such as Netflix learn from the users' choices and recommend content personalised for each user (Wani et al., 2020).

With Deep Learning networks being so versatile, it is no wonder how almost all sectors are eager to develop faster and smarter networks.

2.4 Deep Learning History

Even though deep learning is still a developing technology, its history goes back more than 70 years.

The Deep Learning concept can be dated back to 1943 when Warren McCulloch and Walter Pitts developed a computer model based on the neurons of the human brain. The model was based on a combination of mathematics and algorithms, and it was only mimicking the way the brain works (Foote, 2017).

In 1970, the first artificial (AI) network was released, but because it did not function as expected, it resulted in a kick-back of Deep Learning and Artificial Intelligence research (Foote, 2017).

In 1989, Yann LeCun and Geoffrey Hinton released the first operational artificial neural network able to read handwritten numbers (Dunn, 2021).

During the 1990s, following Yann LeCun's and Geoffrey Hinton's success, AT&T developed a convolutional neural network able to read cheques (LeCun et al., 1998).

Nowadays, with the evolution in computing power, computational speeds are vastly improved, allowing for further applications of Deep Learning. In 2009, the ImageNet website was launched, where over fourteen million pictures were uploaded to provide input for

"training" the Deep Learning networks. Furthermore, in 2012, Google Brain developed a system that took ten million screenshots from YouTube to feed his system (Foote, 2017). There is no doubt that with the constant improvement in computing power, we will witness further breakthroughs in Deep Learning.

It is notable to mention that all sources share the same theoretical characteristics despite the year of publication. For instance, LeCun's paper from 1998 mentions the same Deep Learning techniques as all the recent papers. This shows us that even though the technology is evolving in terms of computational power, the processes of Deep Learning remain almost the same.

3. Deep Learning Limitations

With Deep Learning still being a developing technology, software engineers are still coming across multiple limitations (Chollet, 2017). Specifically, deep learning algorithms are operating in an input-to-output mapping, lacking the ability to process reasoning problems. Reasoning problems include programming tasks, long-term planning, emotion recognition, etc. (Chollet, 2017). Therefore, deep learning machines are still limited in input-to-output calculations and do not have a human understanding of their input (Papernot et al., 2016).

4. Conclusion

It is not unreasonable to say that Deep Learning along with AI networks has the potential to become one of the most ground-breaking technologies of our times. Today, Deep Learning is adopted by several sectors, providing numerous benefits to consumers and citizens, in ways that we do not even realise.

However, most of the currently available literature is focused on the old non-reasoning Deep Learning techniques. Even though many things have changed since the late 1980s, mostly in the Deep Learning applications, the range of its functions will always be limited, until researchers find a functioning way to introduce programs able to process reasoning problems.

Literature findings and gaps

After completing my literature review, I reached certain assumptions about Deep Learning technology and how professionals perceive it:

a. Deep Learning is still developing along with Big Data, but it is already a part of our daily life.
b. Its uses can be unlimited, however, there are still certain limitations.
c. The increase in the computation power increases Deep Learning's usability.

As Deep Learning is still developing, certain gaps and limitations exist:

a. Training of the neuron networks takes a long time to complete, resulting in time constraints and accuracy problems (Shrestha & Mahmood, 2019).

b. Deep Learning requires a vast amount of data to train itself, therefore, further data warehouses should be implemented for supporting these networks (Goodfellow et al., 2016).
c. Big Data is still a developing technology, resulting in limited support for Deep Learning (Goodfellow et al., 2016).

Additional research should transpire for creating new and modern Deep Learning techniques compatible with today's emerging technologies, to expedite the full utilisation of Deep Learning technology.

5. References

Amidi, A. and Amidi, S. (n.d) Convolutional Neural Networks Available from: https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks [Accessed 26 September 2021].

Chi-Hua, C. (2021) Deep Learning for Internet of Things Available from: https://ieeeaccess.ieee.org/open-special-sections/deep-learning-for-internet-of-things/ [Accessed 16 September 2021].

Chollet, F. (2017) *The limitations of deep learning*. Deep Learning With Python.

Dunn, T. (2021) *Deep learning*. Salem Press Encyclopedia of Science.

Eberendu, A.C. (2016) Unstructured Data: an overview of the data of Big Data. *International Journal of Computer Trends and Technology*, 38(1), pp.46-50.

Foote, K. D. (2017) A Brief History of Deep Learning Available from: https://www.dataversity.net/brief-history-deep-learning/ [Accessed 17 September 2021].

Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep learning*. MIT press.

IBM Cloud Education (2020) Convolutional Neural Networks Available from: https://www.ibm.com/cloud/learn/convolutional-neural-networks [Accessed 26 September 2021].

IBM Cloud Education (2020) Deep Learning Available from: https://www.ibm.com/cloud/learn/deep-learning [Accessed 17 September 2021].

IBM Cloud Education (2020) Recurrent Neural Networks Available from: https://www.ibm.com/cloud/learn/recurrent-neural-networks [Accessed 26 September 2021].

Kapoor, A., (2019). Hands-On Artificial Intelligence for IoT: *Expert machine learning and deep learning techniques for developing smarter IoT systems*. Packt Publishing Ltd.

LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11), pp.2278-2324.

Lisowski, E. (2020) Deep Learning Architecture Available from: https://addepto.com/deep-learning-architecture/ [Accessed 26 September 2021].

Mathew, A., Amudha, P. and Sivakumari, S., (2020) Deep Learning Techniques: An Overview. *International Conference on Advanced Machine Learning Technologies and Applications* pp. 599-608.

Murnane, K. (2016) What Is Deep Learning and How Is It Useful? Available from: https://www.forbes.com/sites/kevinmurnane/2016/04/01/what-is-deep-learning-and-how-is-it-useful/?sh=3633574ad547 [Accessed 17 September 2021].

Oprea, S., Martinez-Gonzalez, P., Garcia-Garcia, A., Castro-Vargas, J.A., Orts-Escolano, S., Garcia-Rodriguez, J. and Argyros, A. (2020) A review on deep learning techniques for video prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Papernot, N., McDaniel, P., Jha, S., Fredrikson, M., Celik, Z.B. and Swami, A. (2016) The limitations of deep learning in adversarial settings. *IEEE European symposium on security and privacy (EuroS&P)* pp. 372-387.

Petersson, D. (2021) CNN vs. RNN: How are they different? Available from: https://searchenterpriseai.techtarget.com/feature/CNN-vs-RNN-How-they-differ-and-where-they-overlap [Accessed 26 September 2021].

Shrestha, A. and Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE Access* 7, pp.53040-53065.

Vadapalli, P. (2020) Top 10 Deep Learning Techniques You Should Know About Available from: https://www.upgrad.com/blog/top-deep-learning-techniques-you-should-know-about/ [Accessed 16 September 2021].

Vadapalli, P. (2021) CNN vs RNN: Difference Between CNN and RNN Available from: https://www.upgrad.com/blog/cnn-vs-rnn/ [Accessed 26 September 2021].

Wang, L. and Alexander, C.A. (2016). Machine learning in big data. *International Journal of Mathematical, Engineering and Management Sciences*, 1(2), pp.52-61.

Wani, M.A., Luo, F., Li, X.A., Dou, D. and Bonchi, F., (2019) *19th IEEE International Conference on Machine Learning and Applications*. ICMLA.