Chapter 7
A Comparative Study of Popular CNN Topologies Used for Imagenet Classification

Hmidi Alaeddine
https://orcid.org/0000-0002-2417-3972
Laboratory of Electronics and Microelectronics, Faculty of Sciences of Monastir, Monastir University, Monastir, Tunisia

Malek Jihene
Higher Institute of Applied Sciences and Technology of Sousse, Sousse University, Sousse, Tunisia

ABSTRACT

Deep Learning is a relatively modern area that is a very important key in various fields such as computer vision with a trend of rapid exponential growth so that data are increasing. Since the introduction of AlexNet, the evolution of image analysis, recognition, and classification have become increasingly rapid and capable of replacing conventional algorithms used in vision tasks. This study focuses on the evolution (depth, width, multiple paths) presented in deep CNN architectures that are trained on the ImageNET database. In addition, an analysis of different characteristics of existing topologies is detailed in order to extract the various strategies used to obtain better performance.

INTRODUCTION

Machine learning is a field that is interested in the development of technology that allows the interpretation and prediction of data in an automatic way. It was originally a result of pattern and object recognition research that seeks to make algorithms able to learn automatically from a huge amount of data. During the years of the field of machine learning, applications in this domain have gone beyond the scope of pattern and object recognition where they are widely used in various tasks such as filtering content on networks of social media or web searches. In the early days, the technique of carrying out this process relies on methods developed by experts who have knowledge in the field purpose of interpreting the

DOI: 10.4018/978-1-7998-3591-2.ch007

data. Today, the current methods are based on networks of artificial neurons that have the ability to perform the same role at the top level compared to the first techniques and without the need for human intervention. In the course of a long history, deep learning (Y. LeCun, 2015), (Schmidhuber, 2014) has undergone alternating periods of excitement and forgetfulness, under which it was known under various names and many researches. The first period of work on deep learning was for nearly 20 years from 1940 until 1960 and described simple neural networks with a supervised and unsupervised learning process (Rosenblatt, 1958). The second period began 19 years after the end of the first period (1979 to 1990) with fukushima's work Neocognitron (Fukushima, 1979) which can be considered as the first work that deserves to be described "deep" (Schmidhuber, 2014). In addition, this period has seen the application of back propagation to artificial neural networks (Werbos, 1981), it recalls us that the back propagation is an effective method of gradient descent for supervised learning has been developed in the years 1960 and 1970. The most modern period represents the golden period of deep learning started from 2006 as shown in the following works (G. E. Hinton, 2006), (Y. Bengio, 2007), where we can notice that in the previous last years, deep learning (Y. LeCun, 2015), (Schmidhuber, 2014) achieve huge success in different areas especially in computer vision. In addition, deep networks have won numerous competitions in image classification (Alex Krizhevsky, 2012), (Matthew D Zeiler, 2013), (Zisserman, 2015), (al, 2015), (He X. Zhang, 2015), (Komodakis, 2016), (Saining Xie, 2016), (Gao Huang, 2017), (Huang, 2016), (P. Sermanet, 2014), (Bengio, 1994) and object detection (P. Sermanet, 2014) such as the annual ILVRC competition Russakovsky. It is based on ANN consisting of various processing layers to learn representations of large sets of data using the back propagation algorithm to update the local parameters of each layer. Among various deep learning algorithms [1,2], Convolutional Neural Networks (CNN) has become the popular and frequently used neural model in deep learning it where has achieved a revolution in the tasks of artificial vision thanks to these capacities persising in which resist for many years. In addition, it offers good performance in terms of precision (Alex Krizhevsky, 2012), (Matthew D Zeiler, 2013), (Zisserman, 2015), (al, 2015), (He X. Zhang, 2015), (Komodakis, 2016), (Saining Xie, 2016), (Gao Huang, 2017), (Huang, 2016), (P. Sermanet, 2014), (Bengio, 1994).

This paper is organized as follows: the second section describes the background. The third section presents the comparative study of popular deep CNN topologies used for imageNET classification and the conclusions are drawn in the last section.

Background

Recent work on deep neural network approaches has improved the performance of these visual recognition systems. In section, we focus on the convolutional neural network, supervised learning and the ImageNet database.

Motivations

CNNs are described as a special type of artificial neural networks that is based on convolution, which is defined as a specialized type of linearity operation. They benefit from a capacity that can be controlled by varying these two elements: the depth and the width. In addition, they have a lower density of connections and parameters compared to the multilayer perceptron (MLP) which makes them easier to train.