Chapter 14
Summary and Outlook

With the tremendous success of deep learning in recent years, the field of data science has undergone unprecedented changes that can be considered a "revolution". Despite the great successes of deep learning in various areas, there is a tremendous lack of rigorous mathematical foundations which enable us to understand why deep learning methods perform well. In fact, the recent development of deep learning is largely empirical, and the theory that explains the success remains seriously behind. For this reason, until recently, deep learning was viewed as pseudoscience by rigorous scientists, including mathematicians.

In fact, the success of deep learning appears very mysterious. Although sophisticated network architectures have been proposed by many researchers in recent years, the basic building blocks of deep neural networks are the convolution, pooling and nonlinearity, which from a mathematical point of view are regarded as very primitive tools from the "Stone Age". However, one of the most mysterious aspects of deep learning is that the cascaded connection of these "Stone Age" tools results in superior performance that far exceeds the sophisticated mathematical tools. Nowadays, in order to develop high-performance data processing algorithms, we do not have to hire highly educated doctoral students or postdocs, but only give TensorFlow and many training data to undergraduate students. Does it mean a dark age of mathematics? Then, what is the role of the mathematicians in this data-driven world?

A popular explanation for the success of deep neural networks is that the neural network was developed by mimicking the human brain and is therefore destined for success. In fact, as discussed in Chap. 5, one of the most famous numerical experiments is the emergence of the hierarchical features from a deep neural network when it is trained to classify human faces. Interestingly, this phenomenon is similarly observed in human brains, where hierarchical features of the objects emerge during visual information processing. Based on these numerical observations, some of the artificial neural network "hardliners" even claim that instead of mathematics we need to investigate the biology of the brain to design more sophisticated artificial neural networks and to understand the working principle of

artificial neural networks. However, when neuroscientists (especially computational neuroscientists) were asked why the brain extracts such hierarchical features, it was surprising to find out that they usually rely on numerical simulations with artificial neural networks to explain how hierarchical properties arise in the brain. From a mathematical point of view, this is a typical example of "circular proof", an apparent logical fallacy.

Then, how can we fill in the gap between empirical success and the lack of the theory? In fact, one of the lessons we learn from the history of science is that the gap between the empirical observation and the lack of theory is not the limiting factor, but suggests the birth of a new science. For example, during the "golden age of physics" in the early twentieth century, some of the most exciting empirical discoveries in physics were quantum phenomena. Experimental physicists discovered many exotic quantum phenomena that could not be explained by either Newtonian or relativistic physics. In fact, there was a serious lag in the theoretical physics that could explain newly discovered quantum phenomena. Mathematical models were further developed, questioned, and refuted by the empirical observations. Even the greatest Albert Einstein said that he could not believe quantum physics since "God does not play dice with the universe." During these intense intellectual efforts to explain the seemingly unexplainable empirical observations, the new theory of quantum mechanics was rigorously formed, which led to numerous Nobel laureates; and new mathematics such as functional analysis, harmonic analysis, etc., has become mainstream in the modern mathematics. In fact, these efforts by scientists completely changed the landscape of physics and mathematics.

Similarly, now there is a great need to develop mathematical theories to explain the enormous empirical success of deep neural networks. In fact, computer scientists and engineers who work on the implementation are like the experimental physicists who give endless inspiration, and the mathematicians and signal processors are like theoretical physicists who try to find the unified mathematical theory to explain the empirical discoveries. Therefore, contrary to the false belief that we are in the dark age of mathematics, we are now actually living in the "golden age", ready to discover the beautiful mathematical theory of deep learning that can completely change the field of mathematics. Therefore, this book has aimed to explore the mathematical theory of deep learning to crack open the black box of deep learning and open a new age of mathematics.

The field of deep learning is interdisciplinary in nature, and includes mathematics, data science, physics, biology, medicine, etc. Therefore, collaborative research efforts between mathematics and other fields are crucial. This is because empirical results not only give the inspiration for the mathematical theory, but provide a means to verify whether a mathematical theory is correct. Therefore, although this book primarily focuses on discovering the fundamental mathematical principles of deep learning, it is hoped that it will play an instrumental role promoting the basic sciences in physics, biology, chemistry, geophysics, etc. using deep learning, and enable readers to be inspired by new empirical problems to obtain better mathematical models.