

Bachelor- oder Masterarbeit in Bioinformatik

AG Bioinformatik - Prof. Dr. Dominik Heider
Ansprechpartner: Theodor Sperlea
theodor.sperlea@staff.uni-marburg.de

Thema: Analyse des Trainingsverhaltens von tiefen neuronalen Netzen mit Hilfe von Phi-Delta-Diagrammen (Machine Learning)

Das Feld der Bilderkennung und -klassifikation hat durch die Entwicklung von tiefen neuronalen Netzen im Allgemeinen und *convolutional neural networks* (CNNs) im speziellen große Fortschritte machen können. Analog zur Bildklassifikation wurden CNNs in den letzten Jahren in vielen Gebieten in der Bioinformatik, hauptsächlich zur Klassifikation von DNA-Sequenzen, eingesetzt [1]. Wegen der Komplexität der Modelle sind deren Klassifikationsentscheidungen nicht oder nur in Einzelfällen nachzuvollziehen, was aber in vielen Anwendungen wünschenswert wäre [2].

In unserer Arbeitsgruppe wurden in vergangenen Projekten sogenannte Phi-Delta-Diagramme zur Bewertung der von CNNs gelernten *features* entwickelt und in das Keras-Framework für Deep Learning eingearbeitet. Diese Form der Darstellung erlaubt es, verschiedene Eigenschaften mehrerer Klassifikatoren oder *features* auf einen Blick darzustellen und somit nachvollziehbar zu machen [3].

In dieser Abschlussarbeit soll untersucht werden, wie sich das Verhalten einzelner Neuronen und *convolution layers* im Verlauf des Trainings eines CNNs verändert. Dazu sollen CNNs u.a. auf dem MNIST-Datensatz trainiert werden und nach jedem Trainingsschritt mit Phi-Delta-Diagrammen ausgewertet werden. Die so gesammelten Informationen sollen dann nach Interessanten und für die verschiedenen Klassen relevanten Verhaltensänderungen untersucht werden. Außerdem sollen verschiedene Architekturen (d.h. z.B. verschiedene *hidden layer*-Größen) mit Phi-Delta-Diagrammen evaluiert werden, um womöglich unnötige Neuronengruppen zu identifizieren. Durch diese Untersuchungen sollen die internen Abläufe in tiefen neuronalen Netzen besser verständlich gemacht werden.

Literatur

[1] Min, X., Zeng, W., Chen, S., Chen, N., Chen, T., & Jiang, R. (2017). Predicting enhancers with deep convolutional neural networks. *BMC Bioinformatics*, 18(S13). <https://doi.org/10.1186/s12859-017-1878-3>

[2] Zeiler M.D., Fergus R. (2014). Visualizing and Understanding Convolutional Networks. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) *Computer Vision – ECCV 2014*. ECCV 2014. Lecture Notes in Computer Science, vol 8689.

[3] Armano, G., Giuliani, A., Neumann, U., Rothe, N. and Heider, D. (2018). Phi-Delta-Diagrams: Software Implementation of a Visual Tool for Assessing Classifier and Feature Performance. *Mach. Learn. Knowl. Extr.*, 1(1), 121-137

Bitte melden Sie sich bei Interesse per Email bei:
Theodor Sperlea - theodor.sperlea@staff.uni-marburg.de

13.01.20