Comparative study of deep learning methods for one-shot image classification (abstract)

Comparative study of Deep Learning methds for One-shot Image Classification (abstract)

Joep van den Bogaert[⋆@] **Haleh Mohseni**[⋆] **Mahmoud Khodier**[⋆] [@]J.V.D.BOGAERT@TUE.NL
Yuliyan Stoyanov[⋆] **Decebal Constantin Mocanu**[⋆] **Vlado Menkovski**
[⋆]Dep. of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands

Abstract

Training deep learning models for images classification requires large amount of labeled data to overcome the challenges of overfitting and underfitting. Usually, in many practical applications, these labeled data are not available. In an attempt to solve this problem, the one-shot learning paradigm tries to create machine learning models capable to learn well from one or (maximum) few labeled examples per class. To understand better the behavior of various deep learning models and approaches for one-shot learning, in this abstract, we perform a comparative study of the most used ones, on a challenging real-world dataset, i.e Fashion-MNIST.

1. Methods and Results

To perform the study, we proceed in a step-wise fashion approach. First, we split the training set of the Fashion-MNIST data (Xiao et al., 2017) in three subsets: a labeled training subset (S_1) consisting of all data (30000 images) from five classes (i.e. 1, 2, 3, 8, and 9); a labeled training subset (S_2) consisting in just one randomly chosen labeled example per class (5 images) from the remaining classes (i.e. 0, 4, 5, 6 and 7); and an unalbeled training set (S_3) consisting in all the remaining examples from the classes 0, 4, 5, 6 and 7 (29995 images). For testing we used the unaltered testing set of Fashion-MNIST.

To analyze the performance of the different approaches, secondly, we optimized a standard Convolutional Neural Networks model on S_1 (CNN_{S_1}) by performing a thorough hyperpameter optimization. Further on, we present the top five approaches that we followed: (1) we fine tune CNN_{S_1} on S_2 using transfer learning without data augmentation ($\text{CNN}^1_{S_1}$), (2) we fine tune CNN_{S_1} one S_2 using transfer learning and optimized data augmentation ($\text{CNN}^2_{S_1}$), (3) we trained a convolutional autoencoder on S_3 and we transfered its weights to a CNN which was fine tuned on S_2 (CNN^3_{AE}), (4) we trained Siamese Networks (Koch et al., 2015) (SN^4) on S_1 and then fine tuned them on an augmented version of S_2, (5) we have start performing a Divide and Conquer (DC^5) approach, where we divided the classification task in subtasks for which specialized autoencoders and CNNs were used. The results are reported in Table 1.

Table 1. Accuracy of the various models studied.

Model	Accuracy [%]
$\text{CNN}^1_{S_1}$	59.3
$\text{CNN}^2_{S_1}$	72.4
CNN^3_{AE}	64.0
SN^4	63.1
DC^5	66.5

2. Conclusion

The obtained results are to a degree unexpected. More exactly, in this specific scenario, they show that data augmentation and hyperpameter optimization can lead a simple CNN model to reach better accuracy then more complex neural networks models. As further work, we intend to develop further the Divide and Conquer approach, which seems to have a very good potential in the one-shot learning context.

Acknowledgments

This research has been done as part of the PDEng Data Mining module at JADS, Netherlands.

References

Koch, G., Zemel, R., & Salakhutdinov, R. (2015). Siamese neural networks for one-shot image recognition. .

Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms.

Submitted to DBDBD 2017, Einhdoven, Netherlands.