Deep Learning - Project Proposal

FixMatch - a full-time labelers worst nightmare?

Ibrahim Abakar
Marcus Loberg
Magnus Tibbe

April 26, 2023

We are planning to implement the default project 1 about transfer learning on the Oxford-IIIT Pet Dataset. Since we are aiming for an A grade, we would like to explore the usage of semi-supervised learning, when fewer labelled training data is available, as an extension of the baseline. The paper [1] writes about leveraging unlabeled data to improve a model's performance by producing pseudo-labels to be used as a semi-supervised learning method. This approach achieved an accuracy of 94.93% using only 40 labels on the CIFAR-10 dataset, we want to reproduce this approach on the Oxford-IIIT Pet Dataset.

We are planning to mainly use Pytorch as a software package to implement the network structure, we want to use this Project as a chance to learn Pytorch, Transfer learning and a effective means of semi-supervised learning. As part of the extension, we will need to implement some advanced data augmentations and some more simple ones. The simple ones will be implemented from scratch from our knowledge of the past assignments and the advanced ones through the Python package Albumentations.

Milestones and grades

When using the Oxford-IIIT Pet database we want to make two tests, just classifying Dog/Cat and classifying the 37 different breeds.

E: A successfully implementation of a ConvNet with replaced final layer to work on the Oxford-IIIT Pet database. The accuracy achieved here will work as a baseline for the next implementations

The initial set of experiments we will run on our baselines (E-part) is CIFAR-10

D/C: Replication of the results of [1] on the CIFAR-10 dataset.

B/A: Transferring the network that we implemented based on the paper to the Oxford-IIIT Pet database and checking what amount of labelled data is needed to still get good classification results.

For initial testing we will use the CIFAR-10 dataset to compare our results to the results of [1] and make sure, that our implementation is correct.

A successful project would be achieved if the intended learning outcomes in table 1 are met and the results has the accuracy close to state of the art.

Magnus:	Learn about semi-supervised learning approaches and tuning large networks
Marcus:	Learn to manage pytorch pakages
Ibrahim:	Adjust implemented networks for your own projects

Table 1: Intended learning outcomes.

References

[1] K. Sohn, D. Berthelot, C.-L. Li, Z. Zhang, N. Carlini, E. D. Cubuk, A. Kurakin, H. Zhang, and C. Raffel, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," 2020.