Exercise 4: Convolutional Neural Networks

Fabian Haas **Markus Reichl** **Florian Weingartshofer**

The goal of this exercise was to predict food classes from the **food 11** dataset by using different convolutional neural networks.

Data Preparation

We decided to use **PyTorch** for this exercise, so we needed to preprocess the dataset so it can be used by pytorch. For this, we implemented a DataLoader for pytorch that loads and transforms the data. The transformation works by resizing the images to 128x128 and normalizing them.

Models

- **Implemented**: these models have been implemented by us.
  - **Custom**: this is a simple convolutional neural network implemented by us and inspired by **ResNet**. It consists of a convolutional layer followed by a ReLu, a residual block and a max pooling layer followed by the same thing a second time and finished by a fully connected layer with a ReLu and another fully connected layer tham maps to the 11 outputs.
  - **LeNet**: this is an implementation of the **LeNet** mentioned in the lecture.
- **Pretrained**: we used these pretrained models: **ResNet50**, **ResNet152**, **GoogLeNet**,

Training and Evaluation

Due to the limited processing capabilities of our hardware, we limited the training epoches to 10, as more takes forever to train. We did not perform a cross validation to find good hyperparameters for this exercise, as the main goal was to try different implementations and it would just exponentailly increase training time. Instead we used **CrossEntropyLoss** as the loss function for all the models and the **Adam** optimizer with a learning rate of *0.001*. Figure 1, shows the *training loss* on the training data as wells as the *accuracy* on the validation dataset. We also evaluated the models in more details via the **classification_report** from **sklearn** which can be seen in Figure 2.

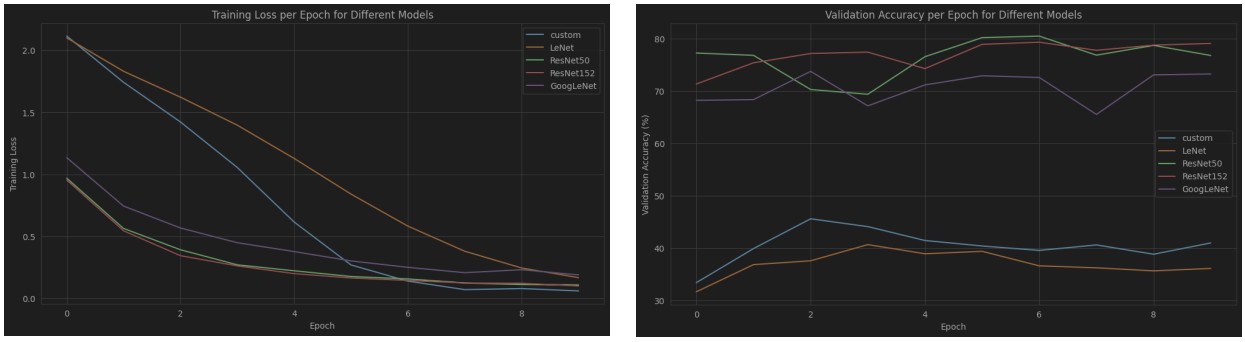

Figure 1: Training loss and Validation Accuracy

	precision	recall	f1-score	support
Bread	0.80	0.72	0.76	368
Dairy product	0.68	0.68	0.68	148
Dessert	0.77	0.76	0.76	500
Egg	0.86	0.80	0.83	335
Fried food	0.66	0.88	0.76	287
Meat	0.76	0.84	0.80	432
Noodles/Pasta	0.93	0.88	0.91	147
Rice	0.79	0.89	0.83	96
Seafood	0.93	0.74	0.83	303
Soup	0.94	0.92	0.93	500
Vegetable/Fruit	0.91	0.90	0.90	231
accuracy			0.82	3347
macro avg	0.82	0.82	0.82	3347
weighted avg	0.83	0.82	0.82	3347

Figure 2: ResNet50 classification report

As the results show, the **pre-trained** models vastly outperform the one we implemented ourselves. The results of the models could be improvement by optimizing the hyperparameters with e.g. **GridSearchCV** and training for more epochs. We could also improve our training strategy. like using StratifiedKFold. The **classification-reports** also show, that the models achieve different results for each of the categories and that the representation of the categories itself is not uniform, which is not ideal.