# Deep Learning for Visual Computing

## Assignment 1

### Group 32

### Amar Karabegovic 0164279

### Nils Kopali 01627943

## Report

#### Image Classification:

Image classification is a fundamental task in computer vision that involves categorizing an input image into one of several pre-defined classes. The goal of image classification is to teach a machine learning algorithm to recognize patterns in an image and to make accurate predictions about which class an image belongs to.

Image classification has a wide range of applications, from image search engines to medical diagnosis. In image search engines, image classification is used to categorize images and make them easily accessible to users. In the field of medical diagnosis, image classification is used to classify medical images such as X-rays, MRI scans, and CT scans into different classes, which helps in the detection of diseases and disorders.

Despite the success of CNNs in image classification, it is still a challenging problem. There are several factors that make image classification challenging, such as variations in lighting, viewpoint, and object occlusion. In addition, the dataset used for training the image classification algorithm must be representative of the real-world scenarios that the algorithm will be applied to. A dataset that is biased towards certain classes or contains images with poor quality can negatively affect the performance of the algorithm.

#### Loss Function:

A loss function is used to quantify how well a machine learning algorithm is performing. It measures the difference between the predicted output of the model and the actual output and provides a value that the algorithm can use to update its internal parameters and improve its performance. In image classification, the most used loss function is cross-entropy loss.

Cross-entropy measures the difference between the predicted probability distribution and the true probability distribution. The true probability distribution is a one-hot vector that represents the ground-truth label of the input image, while the predicted probability distribution is a vector of predicted probabilities for each class. The cross-entropy loss is a logarithmic loss that penalizes the algorithm for making incorrect predictions and rewards it for making correct predictions.

To support the cross-entropy loss, the ground-truth labels and predicted class-scores must fulfill certain criteria. The ground-truth labels must be one-hot vectors that represent the correct class of the input image. The predicted class-scores must be a vector of probabilities that sum to one, and each element of the vector must represent the probability that the input image belongs to a particular class. These criteria are ensured by using appropriate activation functions, such as the softmax function, and by encoding the ground-truth labels as one-hot vectors.

#### Training, Validation, and Test Sets:

To evaluate the performance of a machine learning algorithm, it is necessary to test it on data that it has not seen before. This is achieved by splitting the available data into three sets: training, validation, and test sets. The training set is used to train the algorithm, the validation set is used to evaluate the algorithm's performance during training and to select hyperparameters, and the test set is used to evaluate the final performance of the algorithm.

The training set is used to optimize the internal parameters of the algorithm and to minimize the loss function. The validation set is used to evaluate the algorithm's performance on data that it has not seen before, and to select hyperparameters such as the learning rate, the number of layers, and the number of neurons. The test set is used to evaluate the final performance of the algorithm and to provide an estimate of its generalization ability.

#### Conclusion:

In summary, image classification is a fundamental task in computer vision that involves categorizing an input image into one of several pre-defined classes. To train an image classification algorithm, a loss function such as cross-entropy is used to quantify its performance. The ground-truth labels and predicted class-scores must fulfil certain criteria to support the loss function, and appropriate activation functions and encoding methods are used to ensure this. Finally, to evaluate the performance of the algorithm, it is necessary to split the data into training, validation, and test sets. The training set is used to optimize the algorithm's internal parameters, the validation set is used to select hyperparameters, and the test set is used to evaluate the final performance of the algorithm.