Deep Learning Assignment I-B: Age Regression

Felix Wiewel

February 10, 2021

1 Introduction

This exercise is the second part of the regression homework in the course Deep Learning. While we covered a very basic regression task on synthetic data in the first part, we will now turn to the task of age regression. The task is to infer the age of a person given a portrait picture of him or her. Intuitively this task seems to be easy at first: you are shown a portrait of a person and need to estimate his or her age. But implementing such an estimator on a computer is rather challenging. If you want to use a model based signal processing algorithm, you would need some sort of model or set of rules, which maps the raw pixel values of an image to an estimated age. This would require a lot of expert knowledge and feature design. Fortunately we can use neural networks to learn such a mapping directly from annotated examples, i.e. portraits labeled with the age of the person in the picture. In this way we just need enough examples and don't have to design the probably very complex mapping from pixels to age. We will build on the mathematical formulation from the last exercise and take a slightly different approach for the implementation. So it is recommended to work through the first part of the exercise before starting with this part.

2 GPU support

In order to speed up calculations with Tensorflow, we need to change the runtime type of this notebook to GPU. For this click on "Runtime" in the top left menu and select "Change runtime type". Then choose "GPU" in the drop down list under "Hardware accelerator". This will enable Tensorflow to execute calculations on a GPU provided by Google Colab.

3 Dataset

The dataset we use in this exercise is called UTKFaces [1]. More information on this dataset can be found at https://susanqq.github.io/UTKFace/. It contains

Figure 1: Example images from the UTKFaces dataset. Image source

over 20000 images of people with ages from the interval $[0, 116]$. Although the individual images are labeled with the age, gender and ethnicity, we will only use the age annotation in this exercise. The images in this data set are available in two different sets. The "In-the-wild Faces" set contains images in their original size and orientation and the "Aligned&Cropped Faces" contains all images with the face aligned in the middle and the size cropped to 200×200 pixels. Since the aligned and cropped set is simpler to work with and requires less memory, we will use it in this exercise. Examples of images from this dataset can be seen in Figure 1. Note that the shown images are original ones that have not been aligned and cropped. In the provided code skeleton, however, aligned and cropped images are used.

4 Keras

In this exercise we will use a slightly different way to define our model when compared with the previous part. This time we will derive a new model class from the `Model` class provided by Keras. This comes with several advantages. First, we can use predefined layer classes from Keras in order to build our model. See the documentation for a list of all available layers and detailed information on how to use them. Second, we don't need to manage the variables which store the weights of our model manually. Keras takes care of this and provides a list of all trainable variables of the model. Third, we can use the `fit` and `fit_generatior` function of a Keras model in order to easily train our model. These functions offer a simple interface to implement a lot of different ways for training. But since sometimes one needs to implement a custom training loop, we will not use this function yet but rather implement our own training loop.

5 Tasks

In order to complete this assignment, please fill out the code skeleton provided with this task description and solve the following tasks.

1. Determine the sizes of the input, output and all feature map tensors.
2. Why can we use the ReLU activation for the output layer in this assignment? Justify your answer. What is the advantage of using a ReLU instead of a linear (no) activation?
3. The dataset contains only integer labels as the age of a person, therefore age estimation could be interpreted as a classification problem. State reasons for and against this idea.

Optional tasks.

1. Use the additional information, e.g. gender, provided in the dataset to build a better age estimator.

References

[1] S. Y. Zhang, Zhifei and H. Qi, "Age progression/regression by conditional adversarial autoencoder," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2017.