Ben-Gurion University of the Negev

Faculty of Engineering Sciences

Department of Software and Information systems Engineering

Deep Learning
Assignment 2

The purpose of the assignment

Enabling students to experiment with building a convolutional neural net and using it on a real-world dataset and problem. In addition to practical knowledge in the "how to" of building the network, an additional goal is the integration of useful logging tools for gaining better insight into the training process. Finally, the students are expected to read, understand and (loosely) implement a scientific paper.

Submission instructions:

a) The assignment due date: 25/4/2021
b) The assignment is to be carried out using Python V3.6 (or higher) and TensorFlow 2.0 (or higher).
c) Submission in **pairs** only. Only **one copy** of the assignment needs to be uploaded to Moodle.
d) Plagiarism of any kind (e.g., GitHub) is forbidden.
e) The entire project will be submitted as a single zip file, containing both the report (in PDF form only) and the code. It is the students' responsibility to make sure that the file is valid (i.e. can be opened correctly).

Introduction

In this assignment, you will use convolutional neural networks (CNNs) to carry out the task of facial recognition. As shown in class, CNNs are the current state-of-the-art approach for analyzing image-based datasets. More specifically, you will implement a one-shot classification solution. Wikipedia defines one-shot learning as follows:

"… an object categorization problem, found mostly in computer vision. Whereas most machine learning based object categorization algorithms require training on hundreds or thousands of samples/images and very large datasets, one-shot learning aims to learn information about object categories from one, or only a few, training samples/images."

Your work will be based on the paper [Siamese Neural Networks for One-shot Image Recognition](#) . Your goal, like that of the paper, is to successfully execute a one-shot learning task for previously unseen objects. Given two facial images of previously unseen persons, your architecture will have to successfully determine whether they are the same person. While we encourage you to use the architecture described in this paper as a starting point, you are more than welcome to explore other possibilities.

Instructions

1. Read the above-mentioned paper.

2. Use the following dataset - [Labeled Faces in the Wild](#)

    a. Download the dataset. Note: there are several versions of this dataset, use the version found [here](#) (it's called LFW-a, and is also used in the DeepFace paper).

    b. Use the following train and test sets to train your model: [Train](#) \ [Test](#). **[Remember - you will use your test set to perform one-shot learning. This division is set up so that no subject from test set is included in the train set]**. Please note it is often a recommended to use a validation set when training your model. Make your own decision whether to use one and what percentage of (training) samples to allocate.

    c. In your report, include an analysis of the dataset (size, number of examples – in total and per class – for the train and test sets, etc). Also provide the full experimental setup you used – batch sizes, the various parameters of your architecture, stopping criteria and any other relevant information. A good rule of thumb: if asked to recreate your work, a person should be able to do so based on the information you provide in your report.

3. Implement a <u>Siamese network</u> architecture while using the above-mentioned paper as a reference.

    a. Provide a complete description of your architecture: number of layers, dimensions, filters etc. Make sure to mention parameters such as learning rates, optimization and regularization methods, and the use (if exists) of batchnorm.

    b. Explain the reasoning behind the choices made in answer to the previous section. If your choices were the result of trial and error, please state the fact and describe the changes made throughout your experiments. Choosing certain parameter combination because they appeared in a previously published paper is a perfectly valid reason.

4. In addition to the details requested above, your report needs to include an analysis of your architecture's performance. Please include the following information:

   a. Convergence times, final loss and accuracy on the test set and holdout set

   b. Graphs describing the loss on the training set throughout the training process

   c. Performance when experimenting with the various parameters

   d. Please include examples of accurate and misclassifications and try to determine why your model was not successful.

   e. Any other information you consider relevant or found useful while training the model

   Please note the that report needs to reflect your decision-making process throughout the assignment. Please include all relevant information.

5. Please note that your work will not be evaluated solely on performance, but also on additional elements such as code correctness and documentation, a complete and clear documentation of your experimental process, analysis of your results and breadth and originality (where applicable).
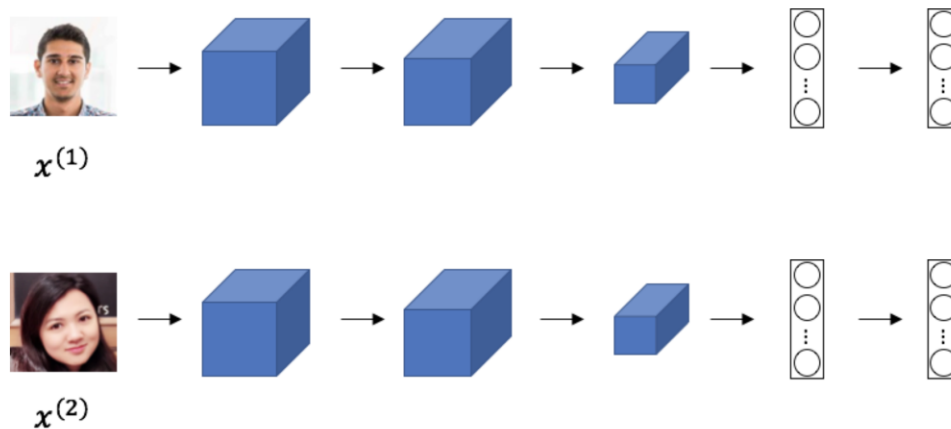

Figure 1 - Siamese network for facial recognition