DL Assignment 2 - Report

Shahar Bardugo 208496687
Leah Wachtfogel 206098931

Dataset Analysis

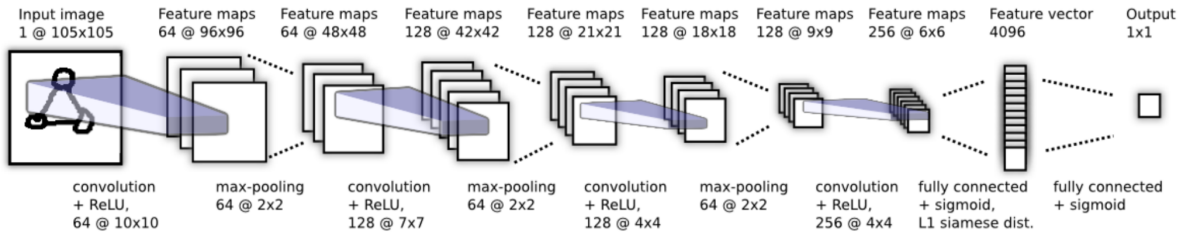
- Size : Total of more than 13,000 photos of 5749 distinct faces; 1680 with more than two photos of the same person and 4096 faces with a single photo.
- Train - Test samples:
  - 2 Classes - “similar” or “different”.
  - Train set 2,200 pairs, Test set over 1,000 pairs, both with class ratio of 1:1.
- Photo size: Each photo in the dataset is an image of size 250X250X3.

Experiment Setup

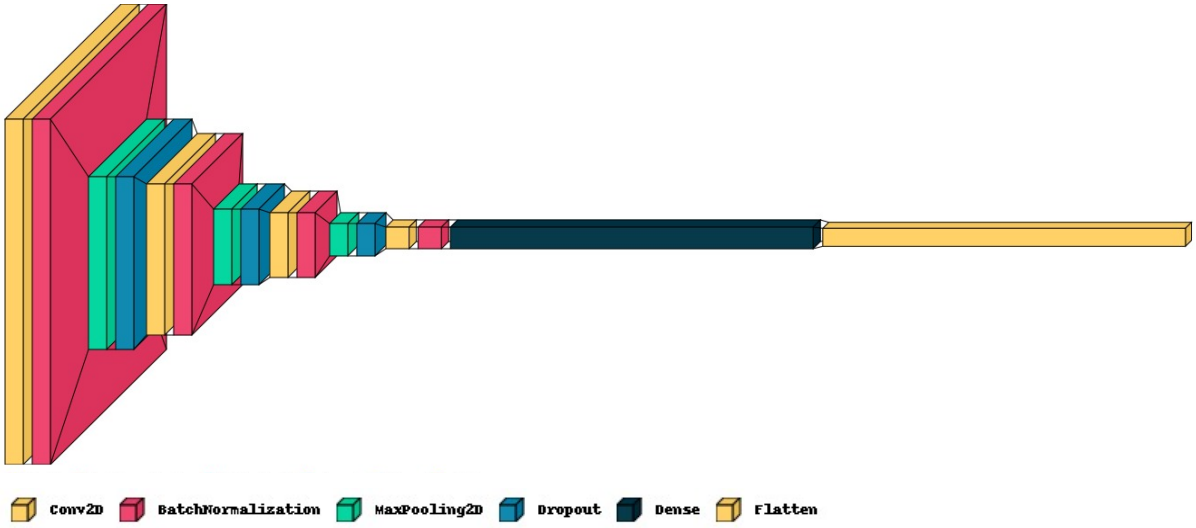
- Batch Sizes:
  - In Siamese Neural Networks for One-shot Image Recognition, the batch size with best performance was 128, since our dataset is much smaller we decided to start training with batch sizes of 16 and 32.
- Parameters:
  - Initialization was chosen to be the same as reported in the article:
    - weights ω - normal distribution, mean 0, std 10^-2
    - bias □ - normal distribution, mean 0.5, std 10^-2
    - On last layer bias is the same and weight initialization was set with mean 0, std 0.2
- Stopping Criteria:
  - No increase of more than 0.01 in validation loss after 5 epochs.
  - Maximum 100 epochs, this was chosen to limit the time performance to be reasonable.

Architecture

- Layers:
  The Layers used are similar to the architecture presented in Siamese Neural Networks for One-shot Image Recognition.



Dropout layer was added after each MaxPooling and BatchNorm layer was added after each convolution layer as follows:



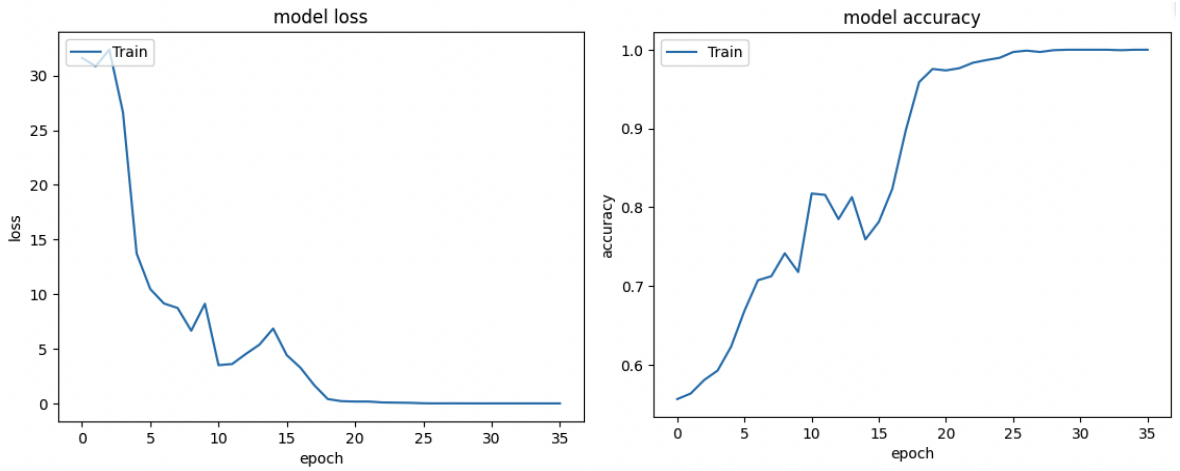
- Dimensions:
  The main difference between the dimensions in the model and the dimensions in the architecture shown above is the input image.
  A single image in the input was reshaped from an RGB image 250X250X3 to a reduced size, greyscale image of 105X105X1 for better time performance and to avoid crashing on running time.
  The other layer dimensions are as shown above from Here.
- Filters: Filter sizes in each layer are as shown in the above graph, varying between 10X10 and 2X2.
- Learning Rate:
  - Similar to the learning rate described in the paper, we used an exponentially decaying learning starting at value 10^-4, with a decaying rate of 0.01.
- Optimization:
  - We used a commonly used built-in optimizer, Adam, that is mentioned in the article and has shown good performance on simple image analysis models.

- Regularization
  - BatchNorm layer was added after each convolutional layer for regularization.
    Attempts to run the model without BatchNorm caused no learning at all (accuracy of ±0.51, similar to random).
  - Dropout layer with a dropout rate of 0.1 was added after each max-pooling layer for regularization.
    Attempts to run with no dropout or with a dropout rate higher than 0.1had worse performance than all other hyper-parameter trials.

Evaluation

#model	dropout	batch size	initial lr	lr decrease	loss train	acc train	loss test	acc test
1	0.1	256	e^-4	0.9	0.2875	0.9629	3.8179	0.615
2	0.1	64	e^-4	0.9	0.1637	0.9662	0.4732	0.677
3	0.1	32	e^-4	0.9	0.5133	0.9184	4.685	0.641
4	0.1	32	e^-4	0.99	0.0324	0.9909	4.5934	0.675
5	0.1	64	e^-4	0.99	0.1044	0.9838	4.4534	0.677
6	0.1	64	e^-2	0.99	4.42E-04	0.9994	9.7482	0.692
7	0.1	32	e^-2	0.99	6.98E-05	1	6.34	0.7
8	0	32	e^-2	0.99	1.60E-03	1	1.7815	0.683

Training loss and accuracy for model 8:



Evaluation examples

Accurate classification :

- True positive - the two pictures are of the same woman and our model labeled them as the same even though in one picture the woman smiles and not in the other. we can see that in both pictures the pose is very close to the middle - we assume that affects positively at the performance.

- True negative - the two pictures are of two different men and our model labeled them as different. We can see that they look very different from each other. one man with hair the other is bald. One with glasses and the other is not and their color is different so we assume the model learned to separate those features. We can see that in one picture the pose is very close to the middle and the other is not - it might help the model say the two pictures are not of the same person but not for a good reason.

Misclassification

- False positive - the two pictures are of two different men and our model labeled them as the same. We can see that they have some shared features like both of them with glasses, both of them wearing suits and both of them are almost bald and both have similar shape of the face.

- False negative - the two pictures are of the same man and our model labeled them as different even though the features of the man look the same in both pictures. We can see that in both pictures the pose is very different and that may affect negatively at the performance.