Deep Learning Assignment 2 - Report

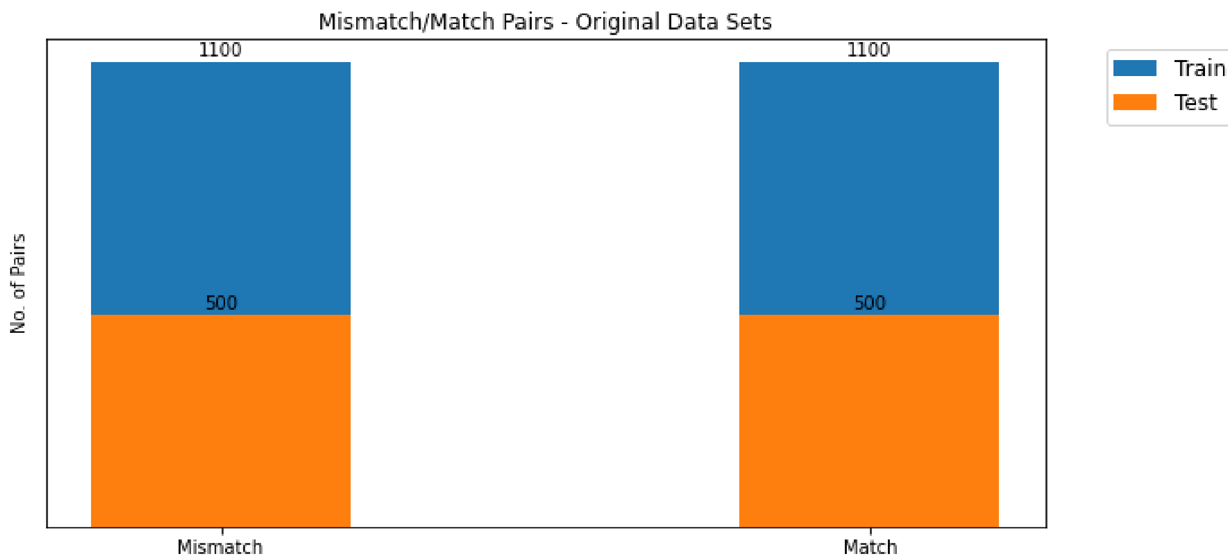
Students

- **Student name** - Leor Ariel Rose, **ID** - 208373365
- **Student name** - Aviv Rovshitz, **ID** - 307974162

Analysis of the dataset

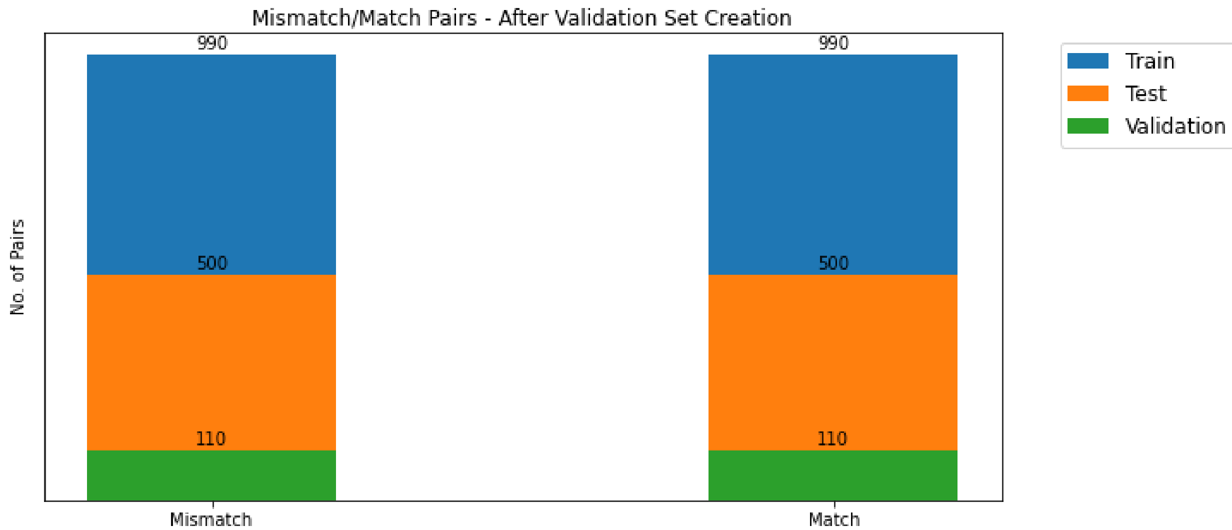
In this section we will present the analysis of the dataset by visual techniques (EDA) and by adding additional written information.

First let's see the number of examples in the original dataset (before splitting the train set into train-validation):



As we can see we have 2200 samples in the train set with 1100 as class "Match" and 1100 as class "Mismatch", 1000 samples in the test set with 500 as class "Match" and 500 as class "Mismatch".

Next, let's see the number of examples in the datasets after creation of a validation set:



As we can see we have 1980 samples in the train set with 990 as class “Match” and 990 as class “Mismatch”, 1000 samples in the test set with 500 as class “Match” and 500 as class “Mismatch” and 220 samples in the test set with 110 as class “Match” and 110 as class “Mismatch”. The validation set is 10% out of the training set.

Lastly, let's look at some of the match and mismatch pairs:

Approach 1 - Koch Et al. siamese neural network architecture

This approach is a loose implementation of the "Koch, Gregory, Richard Zemel, and Ruslan Salakhutdinov. "Siamese neural networks for one-shot image recognition." ICML deep learning workshop. Vol. 2. 2015." scientific paper.

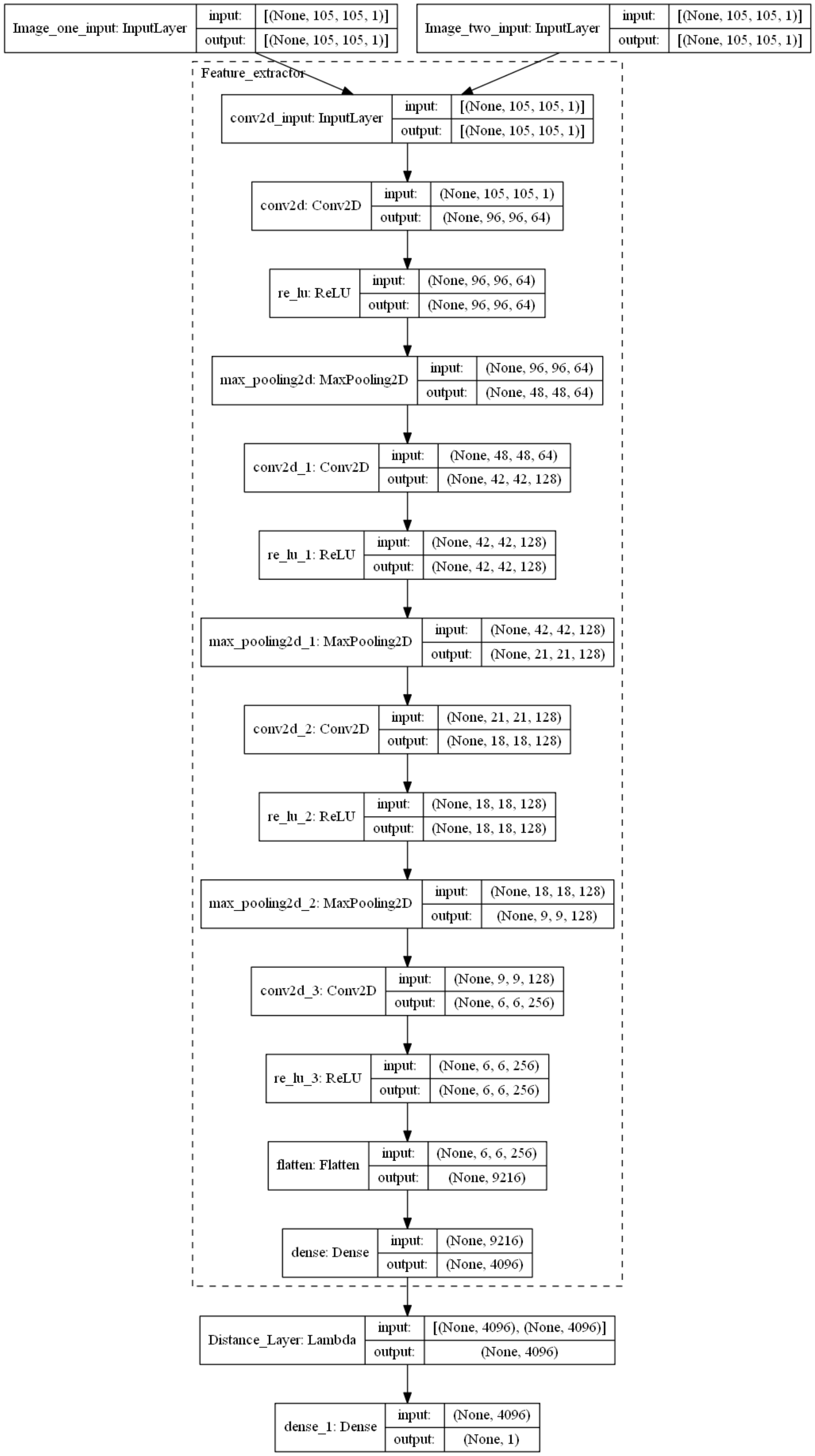
In our implementation we used the theses following elements the same:

- Model architecture.
- Weights and biases initialization.
- Encoded distance metric.

The reason this is loose implementation is because our implementation differs in the following:

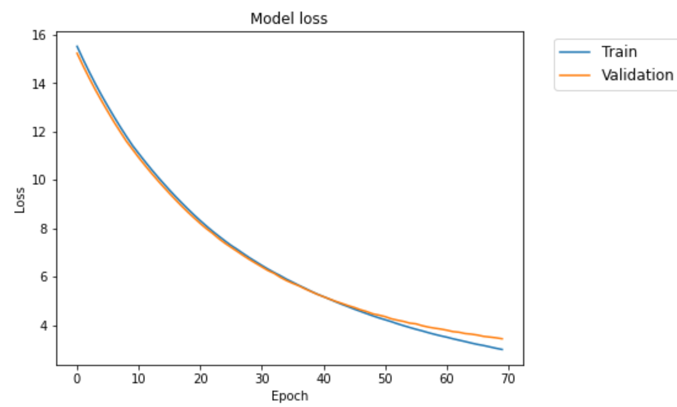
- We used a batch size of 32 and not 128 because of the dataset size.
- We used a learning rate of 0.0001 for a better convergence and in the paper they used between [0.1, 0.0001].
- Layer wise L2 regularization penalty was defined as 0.000002 for convolution layers and 0.00001 for dense layers and in the paper they used [0, 0.1]. The reason we used a lower regularization term is because the paper regulation terms were too high causing a weird loss.
- We did not perform a learning schedule and a constant learning rate was applied for simplicity.
- We used Adam optimizer for simplicity.
- Training Stopping criteria
    - Regular stopping - after 100 epochs.
    - Early stopping - When validation loss change (improvement) is less than 0.2 for 5 epochs.
- Training checkpoint - in the training process best model weights are saved by accuracy and are loaded after the end of the training step.

The following Model architecture was used:



Performance:

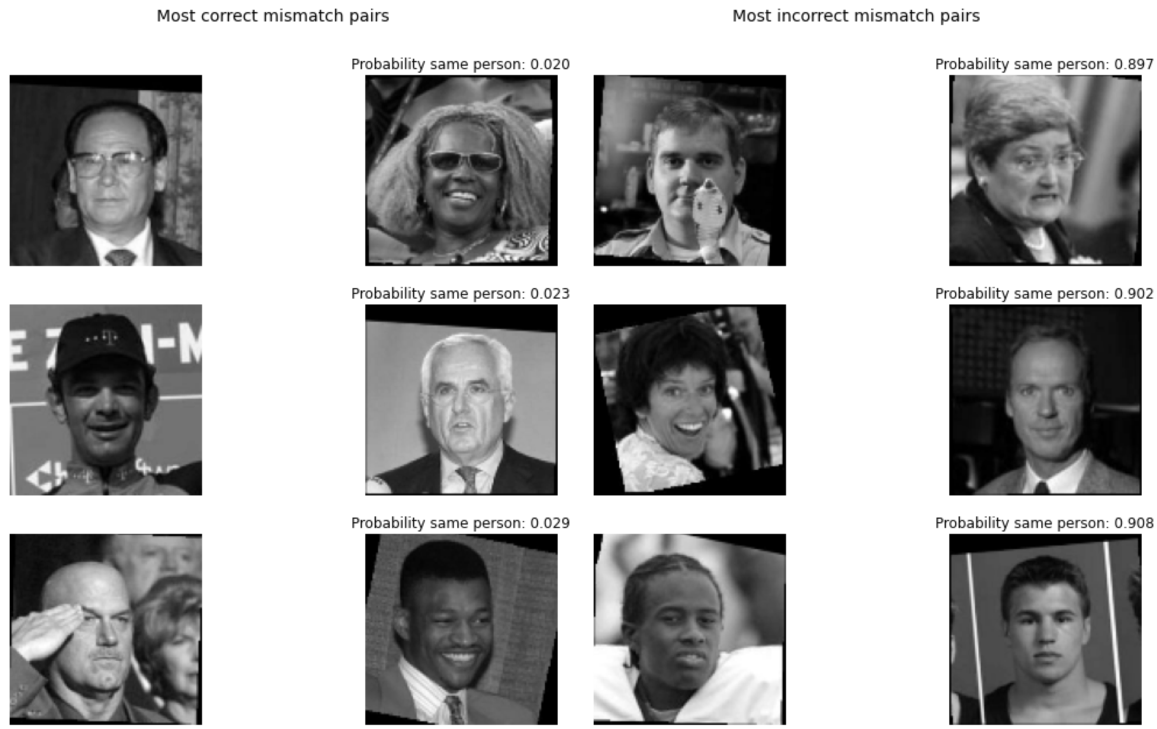
- Number of training epochs - 70
- Train loss - 3.6382
- Train accuracy - 0.9136
- Validation loss - 3.9274
- Validation accuracy - 0.6955
- Test loss - 3.9203
- Test accuracy - 0.6562
- Training process:



- Misclassification examples:

Most correct match pairs

Most incorrect match pairs



As we can see model performance isn't the best and there is a little overfitting. Also if we look at the misclassification examples we can see the model has problems identifying with different backgrounds and objects in the image.

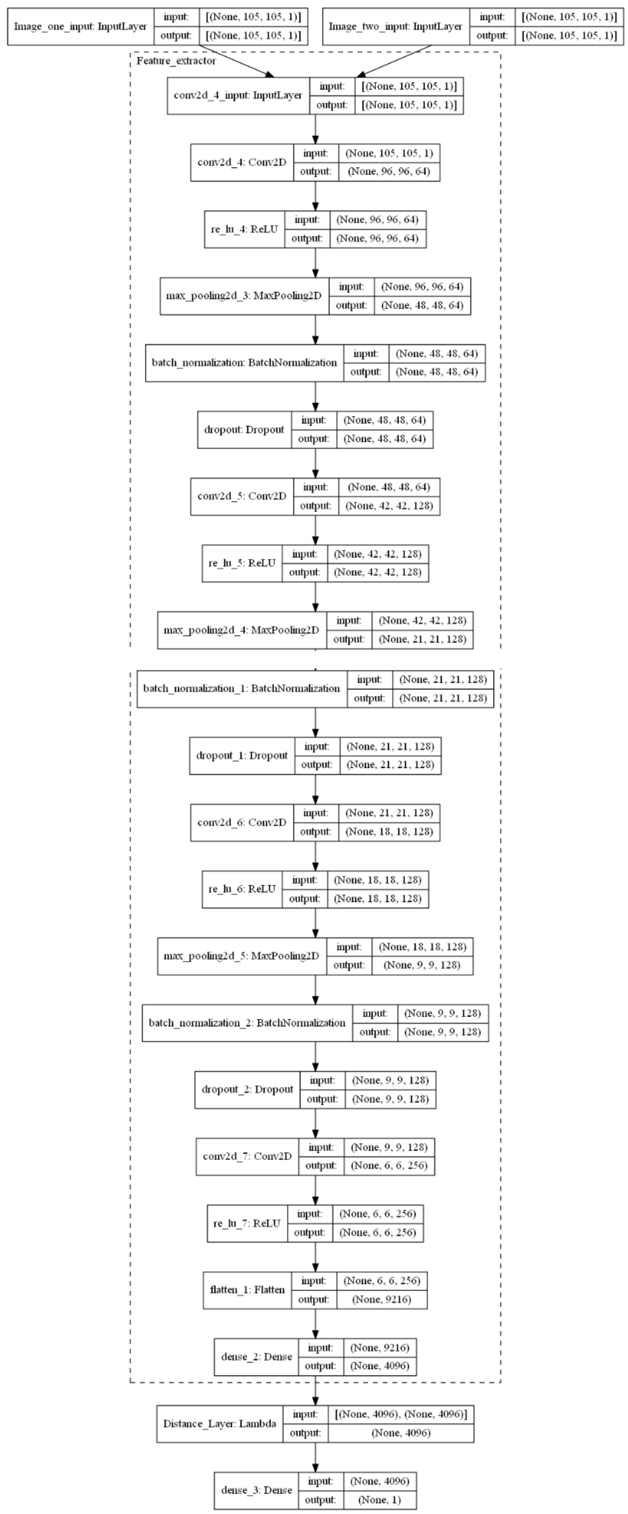
Approach 2 - Koch Et al. siamese neural network architecture improved

This approach is same as approach 1 with the following changes:

- Added dropout after each convolutional block in order to reduce overfitting.
- Added batch normalization after each convolutional block because it shows an improvement in the results.
- Removed L2 regularization in order to avoid too much regularization because of the dropout.
- Changed weights and biases initialized to the default (Xavier normal initializer) because it shows an improvement in the results and for simplicity.

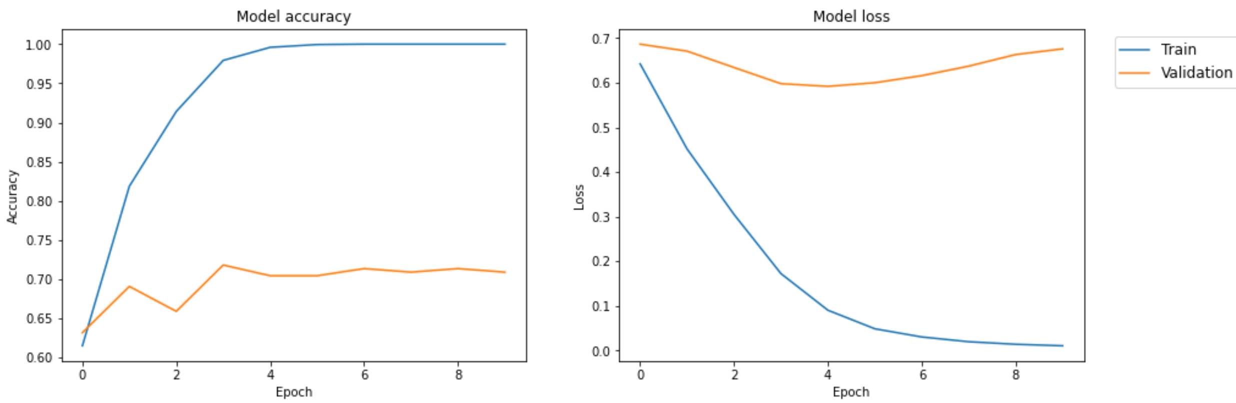
Most of these changes were performed by trial and error and showed an improvement on the original approach.

The following Model architecture was used:



Performance:

- Number of training epochs - 10
- Train loss - 0.3554
- Train accuracy - 0.9490
- Validation loss - 0.5979
- Validation accuracy - 0.7182
- Test loss - 0.5818
- Test accuracy - 0.7100
- Training process:



- Misclassification examples:

As we can see model performance improved and we needed less epochs in order to train but still this model isn't the best and there is a little overfitting. If we look at the misclassification examples we can see the model has problems identifying with different backgrounds and objects in the image.

Comparison

Approach	**Approach 1**	**Approach 2**
Number of training epochs	70	10
Train loss	3.6382	0.3554
Train accuracy	0.9136	0.9490
Validation loss	3.9274	0.5979
Validation accuracy	0.6955	0.7182
Test loss	3.9203	0.5818
Test accuracy	0.6562	0.7100

Conclusion

We can see that the additional changes added to Koch Et al. model improved the results but still we do not have a perfect model and maybe additional data or more quality data is the answer because as we can see the model has problems when there are additional objects in the image or the background\colors are too different.

Another additional step that can be done in the future is data augmentation in order to improve generalization.