

Face Detection using CNN

CS-725 PROJECT – MID TERM EVALUATION

APRIL 2, 2017

Overview

1. Introduction

- Project Description
- Paper referred

2. Collection of Dataset

3. Pre-Processing of data

4. Program and workflow

5. Results and Observations

6. Future work

Introduction

- Project Description

In this project we plan to solve the problem of face detection using a face detector. Once trained the face detector should be able to identify multiple faces within a picture.

- Paper Referred

Paper, which we are referring for our project is-

<https://arxiv.org/pdf/1502.02766.pdf>

Collection of Dataset

- We have used the same dataset as used by the paper <https://arxiv.org/pdf/1502.02766.pdf>
- The dataset is available at <https://lrs.icg.tugraz.at/research/aflw/>
- We had already asked for permission for this dataset and we have got the required access for this.

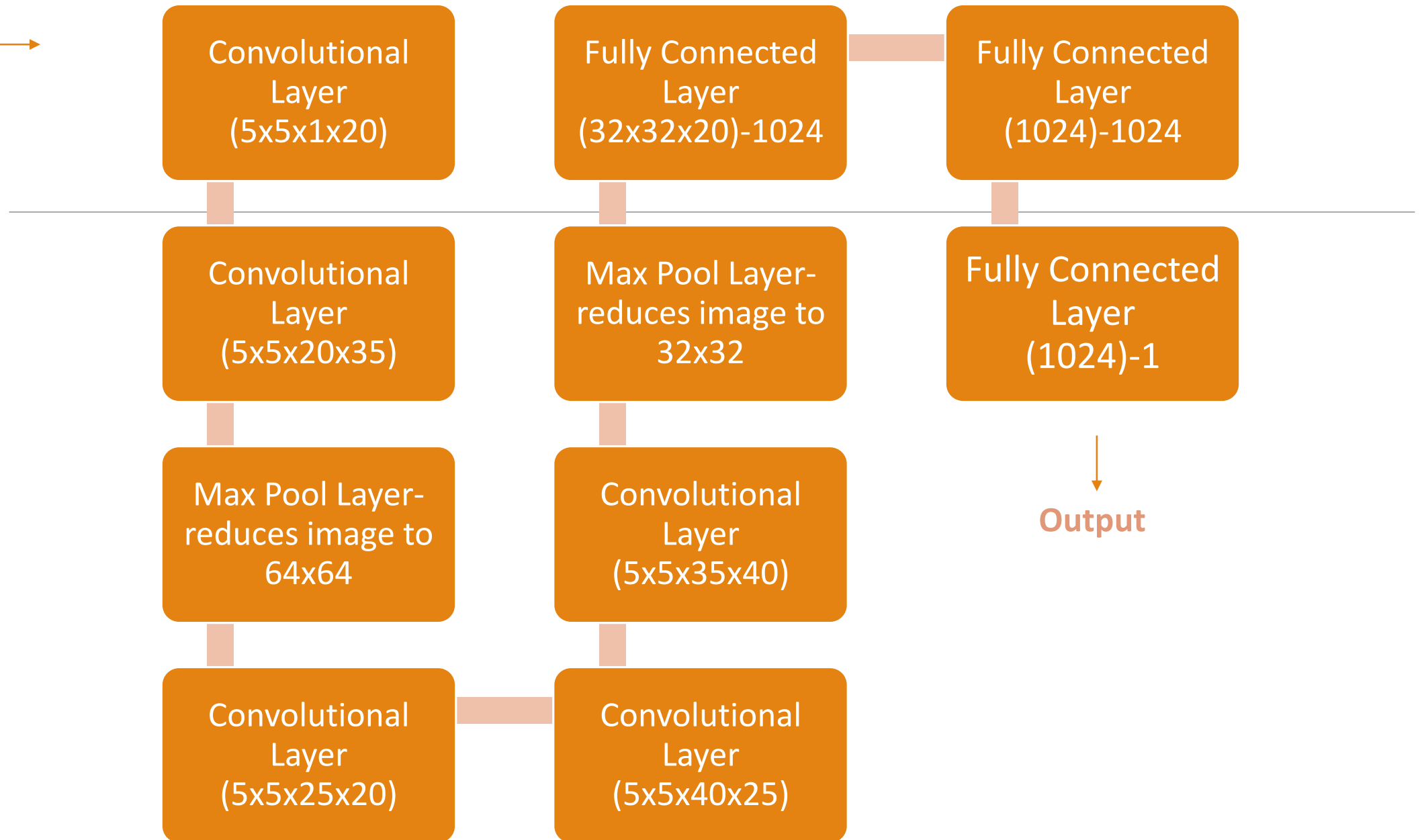
Pre-Processing of Data

- We have pre-processed the data that we received to get positive and negative examples for our training.
- We have constructed 24000(approx) black and white images from the data for faces, which are positive examples.
- We have also constructed 1,20,000(approx) negative images from the faces.
- For this, we have used sliding windows across each image.
- If the intersection over Union (IoU) of the sliding window with the positive image is lesser than 0.3. All the images have been taken to be 128x128.

Program

- We have used tensorflow for creating the model of Deep Convolutional Neural Network.
- Taking in so many images for training has been a challenge.
- We have used training batches of 120 each of which has 24 images and 96 positive images.
- We have used Cross Entropy Loss function in our program and are using the inbuilt 'Adam Optimizer' in Tensorflow.
- The layers that we have used in our shown in the flow diagram ahead.

Input →



-
- Our learning rate is $3e-3$ and we are using a regularization parameter of 0.01.
 - We have also employed the dropout feature of Tensorflow with the keepprob factor as 0.8.
 - Activation units at all the layers has been kept to be ReLU. At the last layer we have used sigmoid activation.
 - Stride for CNN used is 'same'.

Observations and Future Work

- With the above mentioned parameters, we have achieved an accuracy of 75% on training data.
- Further we will try to manipulate parameter values, like batch size, image dimension or stride in order to achieve higher accuracy.
- Network Architecture can be modified a little bit as per our requirement.