CNN LABELING PROJECT - NOTES

1. Introduction

The aim of the project is to create a labeling clothes with Convolutional Neural Networks. There is a main model to classify Glasses/Sunglasses, Trousers/Jeans oraz Shoes on images. After that, regarding the prediction of a main classier, one of these three models can be chosen:

- Glasses and Sunglasses Model
- Model with labels: Trousers Male, Jeans Male, Trousers Female, Jeans Female
- Gender model, and then Male model with labels: Boots, Trainers/Sneakers, Sandals/Flip Flops/Slippers, Formal shoes, Others and Female: Boots, Ballerina shoes, Trainers/Sneakers, High heels, Sandals/Flip flops/Slippers/Others

2. Main model

Labels:

- 0: 'Glasses/Sunglasses'
- 1: 'Trousers/Jeans'
- 2: 'Shoes'

Chosen architecture and hyperparameters:
Conv2D(32, 3, “relu”) -> MaxPooling2D(2,2) --> Conv2D (32, 3, “relu”)--> Flatten() -- > Dense(3))

Accuracy : 0.999

3. Glasses and Sunglasses

Labels:

- 0: “Glasses”
- 1: “Sunglasses”

Chosen architecture and hyperparameters:
Conv2D(32, 7) -> MaxPooling2D(2,2)--> Conv2D(32, 3) --> MaxPooling2D(2,2) --- > Flatten -- > Dense(256) --- > Dense(2)

Accuracy : 0.968

4. Trousers and jeans

To train the best model, two approaches were studied:

4.1. One classifier

Labels:

- 0 - Male trousers
- 1 - Female trousers
- 2 - Male jeans
- 3 - Female jeans

4.1.1. First model: Conv2D - > Flatten - > Dense

Chosen architecture and hyperparameters:
Conv2D(32, 3, ‘relu’) ---- > Flatten() - - > Dense(4)
Accuracy : 0.6960

4.1.2.Second model: Conv2D -> MaxPooling2D - > Conv2D -> MaxPooling2D - > Flatten - > Dense

Chosen architecture and hyperparameters:
Conv2D(32, 5) -> MaxPooling2D(2,2) - > Conv2D(32, 5)-> MaxPooling2D(2,2) - > Flatten - > Dense(4)
Accuracy : 0.756

4.1.3. Third model: Conv2D -> MaxPooling2D - > Conv2D -> MaxPooling2D - > Flatten - > Dense -> Dense

Chosen architecture and hyperparameters:
Conv2D(64, 3) -> MaxPooling2D(2,2) - > Conv2D(64,3) -> MaxPooling2D(2,2) - > Flatten - > Dense(256) -> Dense
Accuracy : 0.766

4.2. Gender classifier and two classifiers for separate genders

4.2.1. Gender classifier

Chosen architecture and hyperparameters:
Conv2D(96, 7) -> MaxPooling2D(2,2) - > Conv2D(96,5) -> MaxPooling2D(2,2) - > Flatten - > Dense(256) -> Dense(2)
Accuracy : 0.9

4.2.2. Male classifier

Chosen architecture and hyperparameters:
Conv2D(32, 3) -> MaxPooling2D(2,2) - > Conv2D(32,5) -> MaxPooling2D(2,2) - > Flatten - > Dense(512) -> Dense(2)
Accuracy : 0.892

4.2.3. Female classifier

Chosen architecture and hyperparameters:
Conv2D(64, 3) -> MaxPooling2D(2,2) - > Conv2D(64,5) -> MaxPooling2D(2,2) - > Flatten - > Dense(256) -> Dense(2)
Accuracy : 0.824

4.3. Comparing two methods

In “Trousers and jeans - comparing methods” two approaches were compared: The results:
Correct prediction for first method: 0.876
Correct prediction for second method: 0.876

Because of the same result, the first method with 4 labels was chosen - third architecture.

4.4. Adding L2 regularization to the third model

L2 regularization is a technique to avoid overfitting. It adds a penalty term into loss function.

After training third model with various lambda values for L2 regularization, the best accuracy was for lambda = 0.01, but the accuracy decreased: 0.762 (without L2: 0.766)

4.5. Adding Dropout to the third model

Dropout technique helps avoid overfitting by dropping a part of nodes in a deep learning process.

After training a third model with 0.1, 0.2, 0.3, 0.4, 0.5 nodes dropped, there was no correction of the accuracy (0.762).

4.6. Adding Data augmentation

Data augmentation is about manipulating training images by rotating and resizing them to make larger training set and better the accuracy and avoid overfitting. There was a decrease in accuracy (0.554).

5. Shoes

5.1. Building gender classifier

Chosen architecture and hyperparameters:
Conv2D(32, 5) -> MaxPooling2D(2,2) - > Conv2D(32, 5)-> MaxPooling2D(2,2) - > Flatten - > Dense(256) - >Dense(2)

Accuracy : 0.84168

5.2. Male classifier

Chosen architecture and hyperparameters:
Conv2D(96, 3) -> MaxPooling2D(2,2) - > Conv2D(96, 3)-> MaxPooling2D(2,2) - > Flatten - > Dense(256) - >Dense(5)
Accuracy : 0.9112

5.3. Female classifier

Chosen architecture and hyperparameters:
Conv2D(64, 3) -> MaxPooling2D(2,2) - > Conv2D(64, 3)-> MaxPooling2D(2,2) - > Flatten - > Dense(512) - >Dense(6)
Accuracy : 0.9112