Image Classification of Clothing Types

Dilmuratkyzy Zerde

Abstract—This report studies classification problem of predicting a label that describes the type of clothing within an image. Three types of clothing are choosen to classify: button downs, hoodies and pants. Two different neural networks are used for classification: AlexNet and GoogLeNet.After trying various experiments, over 96% accuracy on validation and testing sets, and 73.3% accuracy on different image testing set was achieved.

1 Introduction

In the big second hand clothing stores or charity thrift stores huge amount of clothes is collected each day, and they need to be sorted by types in order to put them for sale and display them in stores.The task of sorting clothes by its type is often very time consuming and physically tiring for workers. This task can be automated in many ways, but the core problem would be defining type of the clothing analyzing its image which can be implemented using image classification. It can be viewed as multiclass classification problem of predicting a label that describes the type of clothing within an image. Clothing type datasets will include images of clothing defined with a single label. Some of the challenges for this task include: some types of clothing can be small in size, types of clothing can share similar characteristics (e.g. bottom of the romper and regular shorts or pants), clothing can look very different depending on aspect ratio and angle.

2 Background / Formulation

To classify images of clothes two types of standard neural network models were used , AlexNet and GoogLeNet. AlexNet was also used for supplied data set.The reason these networks were chosen are that dataset contains colored images of clothings, and these networks are commonly used in image classification problems and has shown great results in accuracy and time. In both models base learning rate is set to 0.01, solver type is SGD(Stochastic Gradient Descent),and models are trained 25 with epochs. Numerous other learning rates and solving types were tested, but these showed the best result. AlexNet performed best with 25 epochs, achieving 96.6% accuracy, when adding more epochs decreased the accuracy by 1-2 percent.

3 Data Acquisition

Dataset is consisted of three classes button down, hooded sweatshirt and pants. Each class has pictures of 5 different pieces of clothing, 128 images for each, thus 640 images in total for each clothing type. Total number of images is 1920, 20 percent of images is used for validation, and 5 percent for test set.

Training set	1440 images
Validation set	384 images
Test set	96 images

Fig. 1. Sample images from the dataset

To collect the images, first, for each piece of clothing approximately a minute long video was captured using cell phone camera, while changing the angle and lighting . Then the video was cropped to a square, then images were extracted from the video. Finally all the images are converted to .PNG format, with 256x256 size. And an additional set of 15 test images (5 images for each class) was collected, this are images of clothes that are not in the training or validation set.

Fig. 2. Test images

4 Results

When training with AlexNet network model, on 25 th epoch validation accuracy was 96.6%. Classifying images from the testing set, which is 5% of the originally collected data, the model resulted 95.83% accuracy, 92 out of 96 images were classified correctly and it took 12 seconds.When classifying 15 new images of clothes that weren't in the training set this model showed 73.33% accuracy, 11 out of 15 images were classified correctly.It took 4 seconds to classify 15 images.

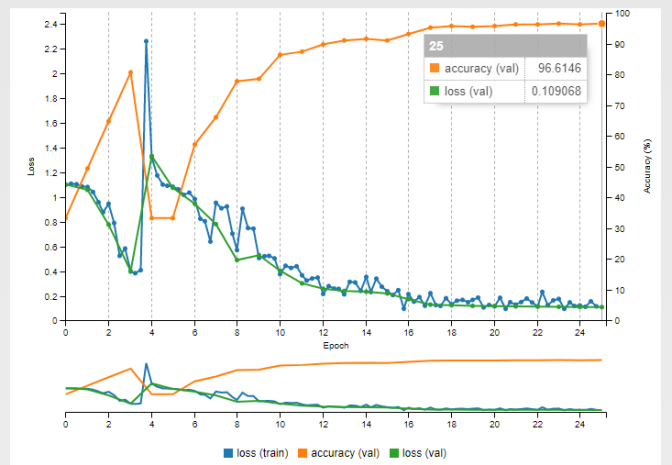

Fig. 3. AlexNet

	button down	hoodie	pants	accuracy
button down	5	0	0	100%
hoodie	2	2	1	40%
pants	0	1	4	80%

The model using GoogLeNet showed 100% accuracy starting from 12 th epoch, and resulted 100% accuracy on testing set, classifying all 96 images.Overal it took 58 seconds to classify the testing set.When testing on 15 other images the model showed 60% accuracy , 9 out of 15 images were classified correctly, and used 5 seconds to classify them.

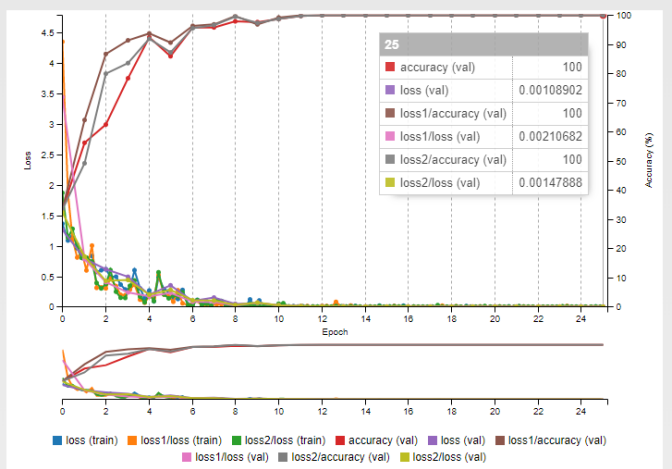

Fig. 4. GoogLeNet

	button down	hoodie	pants	accuracy
button down	2	1	2	40%
hoodie	1	2	2	40%
pants	0	0	5	100%

5 Discussion

As can be seen from the results even though GoogLeNet model showed better result when classifying test data set and validation set, AlexNet model resulted higher accuracy on test images. And AlexNet model showed more than 80% accuracy in 2 classes, where GoogLeNet resulted high accuracy only in one class. Additionally looking at inference time, although two model performed similarly when classifying 15 images, GoogLeNet model used 58 seconds to classify 96 images, when AlexNet model only took 12 seconds. Even though when considering the problem of sorting clothes there will be no need to classify large amount of images all at once, it might be key factor in other scenarios such as sorting images of clothes for e-commerce websites. Given the results, for this particular problem the AlexNet model might be more suitable.

6 Conclusion / Future work

Both network models showed similar results classifying validation and testing sets,but the model using AlexNet network resulted higher accuracy when classifing images of clothes that were not in the training set. Both models showed 40% accuracy classifying images of hoodies. It might be due to the fact that training set for hoodie class had images with mostly white/grey color tone. Collecting images of clothing pieces with different colors should improve the accuracy. To be commercially viable product both number of clothing types and the variety of clothes in one class with different color schemes and patterns need to be increased. Also when collecting data, along with angle and ratio,the way that clothing lays on the canvas(e.g. scrunched, folded, or pulled inside out) should be considered too. This types of classification can also be applied in areas like e-commerce, sorting out clothes by types or searching clothes by image.