Individual lab report #6

Feb, 2, 2017

Xiaoyang Liu

Team F Rescue Rangers

Juncheng Zhang (Henry)

Karthik Ramachandran

Sumit Saxena

Individual Progress

- Modified the previous RGB human signature detection algorithm based on a series of problems.
- Explore on Neural Network for RGB.

Modified the RGB human signature detection algorithm

1) Problems we have in our previous algorithm

In the winter break, we reexamined our algorithm and find out several problems in it. The most severe one is the outfitting.

According to WIKEPIDEA*: In overfitting, a statistical model describes random error or noise instead of the underlying relationship. Overfitting occurs when a model is excessively complex, such as having too many parameters relative to the number of observations. A model that has been overfitted has poor predictive performance, as it overreacts to minor fluctuations in the training data.* As figure.1 shown below, the green line represents the overfitted model and the black one shows the regularized one. It turns out that even though the green line is best following the training data, it will generate random noise when doing prediction on test data.

In our case, we use 700 images (the number of observations) but the extracted HOG features(parameters) can be 7000 or more. This might cause the unexpected bias in our model.

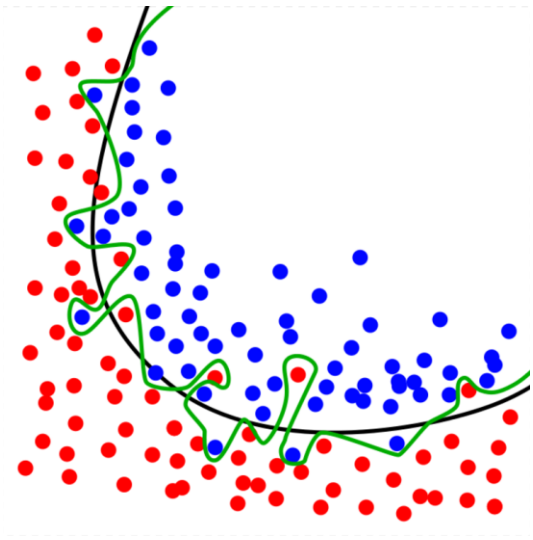

Figure.1 An overview of outfitting problem

The second problem we found out is that the preprocessing of input images is not sufficient. So the extracted features may not be very representative.

Also, as for the strategy of machine learning, we can try the F-Fold method to train our classifier in a more efficient way.

2) Solutions and Modification on the previous algorithm

- Solution for outfitting problem
  - Down sample the training images
  - Increase the HOG cells size
  - Increase number of training samples
  - *Regularization
- Solution for preprocessing
  - Binarize (already exited)
  - Erode, dilate…
- F-Fold Strategy
  - Randomly pick up ¼ images in the training set to be the validation set
  - Let the rest ¾ to be the new training set
  - Calculate the accuracy on the validation set based on the new training set
  - Iterate the whole process and find the most efficient training set with the maximum accuracy on the validation set.

3) Results

I conducted our experiment on a training set contains 700 training images, and a test set which has 219 images.

The result below shows the accuracy experimented on trained images with different HOG cell size and whether images are downsampled or not.

Table 1. The Comparison of different methods

#HOG cell size	%Down sample	Accuracy
4*4	**1**	**84.02%**
4*4	**4**	**83.86%**
8*8	**1**	**79.95%**
8*8	**4**	**86.36%**

Figure.2 The confusion matrix of test images before and after the modification

Figure.2 shows the confusion matrix of test images before and after the modification

4) Conclusion

The result above shows that the largest HOG cells combined with the downsampled images can give us the best result. The result kind of proves our deduction that the number of extracted features should be aligned with the number of training samples. Otherwise, the overfitting problem will occur.

During the process of modifying the original algorithm, I also use some image preprocess strategies to achieve a better result. The basic idea is that I try to get a binarized image which can show a clearer contour of human before sending it the learning classifier.

Exploration on Neural Network

With the extracted features, I plan to build a basic feedforward neural network next. I did some research on the activation functions, and finally choose to use ReLU.

The reason is that there are two major benefits of ReLUs. First of all, it has a sparsity and a reduced likelihood of vanishing gradient. The definition of a ReLU is h=max(0,a)h=max(0,a) where a=Wx+ba=Wx+b. So,

1) The reduced likelihood of the gradient to vanish. This arises when a>0a>0. In this regime, the gradient has a constant value. In contrast, the gradient of sigmoid becomes increasingly small as the absolute value of x increases. The constant gradient of ReLU results in faster learning.
2) Sparsity arises when a⩽0a⩽0. The more such units that exist in a layer the sparser the resulting representation. Sigmoid on the other hand are always likely to generate some non-zero value resulting in dense representations. Sparse representations seem to be more beneficial than dense representations.

Challenge

- I've tried the F-Fold method, but it turns out that it cannot increase the accuracy very efficiently. I was checking many times on my code, but it shows that I was implementing in a right way, so there must be some other reasons behind. I think that maybe it's because we have so few training samples that there would not be enough training samples to be used if we divide the training set into a new training set part and a validation set part. So, anyway we need to get more data for improving the performance of the algorithm
- It took a long time for us to finalize our test plan in Nardo Airport with our sponsor. The reason behind it is that we have to consider many factors like how to capture more various dataset, how to save time and flight passes while achieving the whole mission. Also, we are not familiar with how to actually fly a drone, and we are not aware of the topographic condition in Nardo Airport. Before the plan is finalized, we revised our plan several times after getting some feedbacks from our sponsor.

Team Work

In this week, we four together finalized the test plan in Nardo Airport. Juncheng (Henry) devoted most of his time exploring thermal thing. The result turns out that the HOG +SVM method can also be used to detect most of humans in one frame. Karthik worked on voice activity detection on sound samples. Sumit mainly focused on learning to work with tensorflow to develop neural net and literature study on fusion.

Future Work

Before the next progress review, Juncheng will do further researches and experiments on thermal images I will focus on neural network and PDB . Sumit will be in charge of modifying test plan parameters and exploring algorithms to detect location of signature from RGB image. The next checkpoint is FVE. So, in the coming week it's very urgent for us to integrate all the things together and test on large enough aerial dataset to verify the efficiency of our algorithm.