Viterbi Internship - Progress Diary

Arka Sadhu

June 22, 2017

Contents

1 Daily Progress

1.1 Week 1 : 15 May to 21 May

1.1.1 15 May : Monday

Accomplished :

- Read the paper Face Recog using deep multi pose representation halfway through.
- Watched and implemented tensorflow tutorials by Marvon Zhou till Lec15.

1.1.2 16 May : Tuesday

Target :

- Complete the Face Recog using deep multi pose representation.
- Watch NN course and complete till week 10 (curr status at week 6).
- Read Do We really need million faces.

Accomplished:

- Completed reading teh Face Recog using deep multi pose representation
- Started reading Do We really need million faces [upto page 4]

1.1.3 17 May : Wednesday

Target :

- Complete Do We really need million faces
- Complete the other two papers as well : A multi scale cascade fully convolutional network face detector, regressing parameters for 3DMM
- Finish upto week 10

Accomplished :

- Complted two papers : Do we really need million faces, and multi scale cascade fcn face detector, and started regressing parameters for 3DMM

1.1.4 18 May : Thursday

Target :

- Complete regressing 3DMM parameters
- Complete till week 10 from NN
- Also try to do the course CS231n Stanford : CNN for Visual Recognition.

Accomplished :

- Slight part of regressing 3DMM parameters is left, but will leaving it as is.
- Downloaded the new papers, and had a brief overview regarding that.

1.1.5 19 May : Friday

Target :

- Read all the 4 papers regarding MediFor.
- Complete till week 10 from NN
- Also see CS231n Stanford course.

Accomplished :

- Got a new Problem Statement : Given an image, need to develop a score map which can say wheather or not this image was taken in that location or not.
- Downloaded places365, need to start working with it.

1.1.6 20 May : Saturday

No work done.

1.1.7 21 May : Sunday

No work done.

1.2 Week 2: 22 May to 28 May

1.2.1 22 May : Monday

Target:

- Read im2gps paper, unsupervised visual representation learning by context prediction.
- Get places365 running nicely, and try to replicate the results.
- Complete till week 10 of NN course.
- Understand the problem statement once again.
- Get a hang of Caffe.

Accomplished:

- Read im2gps paper. Mildly interesting, mostly experiments. Didn't really explain well what it wanted to show.
- Read the unsupervised Visual representation learning by context prediction.
- Did the lec6e finally, and started lec7 of NN course.

1.2.2 23 May : Tuesday

Target:

- Get hang of how to use Caffe.
- Get Places-CNN working.

Accomplished:

- Got a good hang of caffe. Tried 3 hands on examples.
- The places365 website was down for some, reason, couldn't really do anything on that.
- Read first 3 lectures of CS231n. Tried few hands on examples.

1.2.3 24 May : Wednesday

Target:

- Get Places-CNN working on laptop. Train if required. See github for reference (site is also up).

Accomplished:

- Reality is harsher. It takes a lot of time to extract the relevant folders. Damn only if there was any way to make this process a bit more faster.

- Not feasible to train on the whole Places dataset. Will take enormous amount of time. More beneficial to use that time for testing purposes.

- Finally set up the PC here for caffe. Took a lot less time than it did on my computer.

- Tried the flickr_finetuning tutorial. The results as seen on Thursday were extremely bad. I have posted on caffe-users, but the community is not very responsive for some reason. Not exactly sure where the problem is. But still got a hang of fine tuning at least.

- Caffe documentation is seriously bad. Need some good tutorials for this. But still got a decent hang of caffe now. Need to start with the medifor dataset at the earliest.

- Seems like did a lot of not-so-really-useful-things today.

1.2.4 25 May : Thursday

Target :

- Try to get Places-CNN testing running some or the other way. Get a hold of the MediFor dataset. It will perhaps take time to preprocess.

- See CS231n for reference in the mean time, may get some useful ideas.

Accomplished :

- Damn, there seems to be some caution that needs to be taken care for External hard-disk. Note to self : in the future, if you are trying to untar, donot use the direct archive method. The problem isn't exactly keep on storing (writing) the data. In fact, if there is some trouble in the process of untaring, there is practically nothing one can do. I had to remove the Hard-disk from the PC. But then it threw the error that the disk is corrupted. My guess is that, while untaring, it still wanted to write some data, which it was not able to do so. I had to go windows and run chkdsk (probably short for check disk), most likely because it is an ntfs partition. At the very least, chkdsk solved the problem pretty quickly. Now trying to unzip using winRAR, and at the very least it shows ETA. For about 4.4GB test data set of the Places365 dataset, it is taking about 4 hours. Lets see how it goes.

- I am an idiot. I was trying to untar it on the hard disk. Untaring it on the SDD was like 1min or so. I seriously wasted the whole morning.

- I am once again reminded of the need to read the documentation. Read the complete caffe documentation, now I feel like I can do something.

1.2.5 26 May : Friday

Target :

- I think, I should go step by step. First I will try to read the existing caffe model, and try to run that on the places365 dataset. I should try to see the test results, and see if it is meeting the expected benchmarks.

- Parallely I will try to create some dataset related to the localization. Let's see how it goes from there on.

Accomplished :

- Got places365 to get working.
- Started reading the readme file of the Nimble 2016 project.

1.2.6   27 May : Saturday

No work done.

1.2.7   28 May : Sunday

No work done.

1.3   Week 3 : 29 May to June 4

1.3.1   29 May : Monday

Target :

- Get a hang of TensorFlow which would likely be used in the project subsequently.
- Watch NN video lectures side by side when running the code so as to fully utilize the time.
- Try to parse everything from the Nimble Challenge Dataset, and try to use Places365 model on it.
- Manually try to evaluate the results.

Notes to self :

- I need to work on the Nimble dataset. Professor suggested that my task would be on, weather the two images are the same.
- To achive that, one of the ideas would be to take a direct inner product between the feature vectors. I also need to think, weather to directly take the softmax probabilities or take the direct values from the previous layer, hoping that it would have captured some or the other thing relating to the semantics of the scene in question.
- Another thing I would need to work on is the image geolocalization.
- For both of these tasks, I would need to know which are the discriminative features, and which are not. For example, even though there are trees in both the images, but they may not be of the same tree. In fact, they may be completely different scenery. Need to clarify as to how to categorize them.
- For the image geolocalization problem, it is more crucial to know what the discriminative features are.
- I have decided to work primarily on caffe, and in case, I am not able to figure out what to do I shall shift to TensorFlow. As such I should be proficient in both TensorFlow as well as Caffe. In Caffe, the coding is mainly in the protobuff file, and mostly not much pains should occur.
- For prototyping, I think it would be advisable to work with AlexNet, but I should also try the ResNet once, to see how much memory it is occupying and all.

Accomplished :

- Parsed everything from Nimble dataset, and places365 model works on it. But the metric and results are very confouding.

1.3.2 30 May : Tuesday

Target :

- Try to interpret the results on Nimble Dataset.

- Try as many metrics as possible, and try to submit it to prof by the end of the day. Some metrics to try upon : SSD, SAD, NCC, census transform with hamming distance (the intuition behind this : assigning probabilities is like cardinal, but we might need ordinal, in the sense we as humans do not really assign probabilities, we rather give rankings : remember economics, but again this may not exactly correspond to the feature vectors, rather these may be for the final softmax probabilites)??.

- Try to understand (read papers) about convolutional neural nets, especially fully convolutional nets (FCN).

Notes to self:

- Output a csv file, containing the correlation found between the two images, : infact, not only correlation, give everything, if possible in dict form only (maybe useful later).

- For this, first try to use the multiprocessing thig, because even with different bash programs, I am not very sure, if they will be able to write to the same file. Try and confirm that the multiprocessing thing is working properly, only then proceed.

Accomplished:

- Tested for some simple results. The correlation (pearson's correlation) was quite high for similar images, and low for different images. But most of the images tried were control, so can't really judge the success.

- Still need to try it on difficult images, and also try to use the multiprocessing library. Check github issue

- Also started reading the documentation of tensorflow a bit. Should be helpful in the future.

1.3.3 31 May : Wednesday

Target :

- Try to make the multiprocessing thing work at the earliest.

- Output the csv file, as fast as possible.

- Get hands dirty with the nimble 2017 dataset.

Notes to self :

- Check git issue, to make multiprocessing work. Compare the speeds, and see how much memory is being used by the GPU.

- After that output csv, and possibly send it to prof (or host it online).

Accomplished :

- Multiprocessing issue resolved. Thanks to the git issue. Really helped a lot

- The output is definitely not up to the mark. In fact, I would say, very very bad results. I am truly disappointed at the appalling results.

- Wrote a mail to the prof asking for what can be done next. Lets see how it goes from here.

- Had a discussion, as to what could be done next. The obvious thing to do is to try it out using the last but one layer. Lets see how it goes from there.

1.3.4   1 June : Thursday

Target :

- Get hands dirty with the Nimble 2017 dataset, which is known to be more challenging.

- Evaluate if there is any better performance using the last but one layer features. (Intuitively should help, but again lets see)

- Try to also read the paper on Very Deep CNN (Zisserman) at the earliest.

Notes to self:

- Not sure why, but I was not at my peak performance today. Started because of the bloody cuda problem, and then with the hard disk. One after the other problem kept creeping up. At the least completed the basic TF tutorial.

- Nimble 2017 dataset is much (I would say exponentially tougher) than the Nimble 2016 dataset. It really tries to push the media forensics to its limits I would say. On the plus side, if I can get something out of this working, I will be quite the hero, and hopefully some paper publication.

- Now is where all the problem is. I need to segregate my work into serialized tasks, so that I can complete my quota today. Its been quite some time without much progress, lets see how it goes from here.

    - First and foremost, I should try to fasten the exisiting process on the Nimble 2016 dataset. It would be quite foolish to wait for 15 mins (this time probably much more than that) for every time I run the code. For this I will have to write a more structured code for the matter, mostly using the multiprocessing library and batch inputs efficiently.

    - Second on the agenda is to use one layer back and start getting things done. Immediately after that I should output the txt file giving the comparison.

    - Also I might need to change some of the incorrect jpg files to png files. I need to figure out a clean way to do them. This would be for the Nimble 2017 dataset.

    - Afterwards try the same code on the Nimble 2017. Before that I might first need to change the parsing of the references and index. Once parsing is done, the rest should be quite straightforward.

Accomplished :

- Wasn't able to do as much as I thought. The thing is I made terrible mistake in using multiprocessing. I cannot really call a global variable inside a process. I circumvented it by using a multiprocessing queue. But that did slow me down by quite an amount.

- I have only been able to complete the first part, i.e. getting the multiprocessing code running efficiently. Fixed up all the bugs (hopefully).

1.3.5   2 June : Friday

Target :

- Follow the tasks from previous note-to-self, and produce the output at the earliest.

- Also try to graph the output if possible possibly using matplotlib.

Notes to self:

- Include everything from June 1 notes to self.

- In addition to that, remember to split the images into two and then into four and check if there is at least some increase in accuracy.

Accomplished :

- Got some good results from using the fc8 layer. Also tried the splicing thing, and as expected the results were much better than those before.

- Also converted the troublesome jpg files to png. Was quite easy tbh.

- Tried the code on the Nimble Dataset 17 (with parsing and all done). Tried on the manipulated set, but it contained only 65 positive images. Need to also run it on the splice as well as the provenance set. Apparently the provenance set has quite a few examples.

- Next on the agenda would be to get the false positives as well, and get some kind of graphical relationship. Then I can try it on the bigger datasets like the one mentioned by the Professor, or the one that is Rama currently working on.

1.3.6   3 June : Saturday

Target :

- Finally working on a saturday. But for starters the target will be a bit low.

- Run the program on the provenance images. Though first may need to increase the disk space (may be).

- Also store the feature vectors into a .npy file, will be easier to get everything in the future.

- Then try to get the number of false positives.

- Also complete the NN course as much as possible.

Notes to self:

- To get the false negatives I would first need to store all the world feature vectors into a .npy file, so that I don't have to do the computations every time. For this, I should create a class which will store the file id along with the feature vector.

Accomplished :

- Only did the test on the provenance part. They are exactly the same as the Manipulated images with is_target as Y.

- Also saved fv1, fv2, fv3 using pickle. Pickle is an extremely useful tool for saving complete classes. And it took very less time. I am astounded that I didn't know about this library before.

1.3.7   4 June : Sunday

No work done

1.4   Week 4 : June 5 to June 11

1.4.1   June 5 : Monday

Target :

- Get results on the false positives and if possible plot a graph for the same. Also try to get some kind of intuition behind why it is or it isn't working.

- Also try to get a hold of the larger test set that the prof was talking about. It seems Rama was unable to give me the dataset. But I really really need some data to work with. I might have to talk to prof as to what else could be done.

Accomplished :

- For starters I got Elpy working with Emacs. Also installed Magit. Excited to use them.

- Got some false positives but very bad results. The main reason is the bad dataset. There are just too many manipulated data in the world dataset, so the matching is actually happening between two image datasets.

- Tried to do some manual interpretation, but I think I will stick to the idea of back tracking the actual image. For this I might have to spend a little more time.

1.4.2 June 6 : Tuesday

Target :

- Get the backtracking of the manipulations done.

- Another idea would be to try a classifier like the k-means to get the clusters. May need to use scikit-learn for the same.

- May also need to do some interpretation manually. Though I wonder why even with 1000 of them, it was not all correct.

Notes to self:

- Since the direct provenance results are quite decent, it might be quite worthwhile to try to implement a direct k-means classifier or something of that sort.

Accomplished :

- Damn the bug was a really prominent bug. Feeling satisfied to have fixed it. But damn, I did really loose my concentration half way through which is a real bad thing.

- Also used scikit-image and everything worked much better than expected to be honest.

- After looking at only the provenance node based I getting a whooping 64/65 correct answers and I am not even sure why that one answer is even incorrect. Really mysterious.

- Started the kmeans, but there is some error, and I think I have a vague idea for that, but still not very sure. Need to get this working at latest.

- Finally fixed the jupyter installation on the local laptop, which I think is quite an achievement.

1.4.3 June 7 : Wednesday

Target :

- Try to use the already written code on as many examples as possible. Especially check if kmeans is giving any decent results or not.

- Can also try to use the segmentation on Rama's matching dataset.

- Also need to evaluate on the eval dataset.

- Try a hobby project (like GAN) on Pytorch or Tensorflow.

Accomplished :

- Just realized that gcp is an extremely powerful tool. The file transfer speeds is quite enormous. (Around 100 MB/s)

- Kmeans gave decent results, and Prof correctly pointed out that further iterations wouldn't change the output much, because it converges to a local minima.

- Got the feature vectors for Rama's protest dataset.

- I think there is some problem with the eval dataset, and I mean it seriously. All images that I could get a hold of were only mask images. I seriously need to get a hold of the world set that they have.

- Haven't been able to try any project on GAN though I will come back to it sooner than later.

1.4.4 June 8 : Thursday

Target :

- Try to evaluate the protest dataset as soon as possible, and compile a report on it.
- Check the eval dataset properly and note if there really is anything useful. If not then get the world dataset either via the MediFor browser or by some other means.

Notes to self :

- Try to visualize the bounding boxes first if possible, before delving into the actual matching part. Perhaps can use scikit image for this, but this is the first step.
- Then I should try to understand the real problem and try to get the underlying bottleneck, and only then I should work on it.

Accomplished :

- Results are pretty bad for the protest dataset. In fact 0/10. I have no clue why.
- Trying with the histogram equalization, but I am quite skeptical that it will still be awful results, since the correlation between the two images was still quite low. And the problem is that I have no actual clue as to why it isn't working. The only distant explanation I can give is that after getting the rightful bbox, the neural net tries to capture the background information, which is what is causing the problem in some sense. This is a really wild and pathetic guess though. I wish I could gain some more insight.
- I am now trying to use kind of histogram equalization for comparison. I am around 80% sure that the algo is not illumination invariant. Another thing I will probably try is to do mean subtraction from the image as well, but I will have to calmly think as to how to go about that. But the preliminary results with histogram equalization were not too impressive.

1.4.5 June 9 : Friday

Target :

- Read the paper by ISI.
- Test with the hist equalization method, and any other tweaks to get the img correlation working.
- If time permits also download the NC2017 Dev Beta 3 version for evaluation on the larger dataset.

Notes to self :

- I have noticed that I am spending slightly more time on github, which I shouldn't do really. I need to be more focussed. I need to take the work more seriously. Time is extremely limited.
- First and foremost get the matching results using hist equalized images.
- After that try to shift on to the new eval set.
- Get time to read upon the paper as well.

Accomplished :

- Submitted 2 issues to skimage repo. And 1 to elpy. Hoping to contribute more. But lets see. To be honest it takes quite an amount of time to even fill in issue.

- After discussion with professor, I would need to ditch the task of evaluating on the protest dataset, and concentrate more on the NC2017 dataset. Try to get a hold of the dev3.

- Read the ISI report. The minimum spanning tree was quite interesting.

1.4.6 June 10 : Saturday

No work done.

1.4.7 June 11 : Sunday

No work done.

1.5 Week 5 : June 12 - June 18

1.5.1 June 12 : Monday

Target :

- First get a hold of the new evaluation dataset (the NC2017_Dev3_beta1)

- Then get the results on the new dataset. I seriously hope that the results are at least somewhat decent.

Notes to self :

- With the relatively decent results, there are a few ways to go to the next task.
    - First would be to get a provenance graph. This could be through the minimum spanning tree (need to read a bit more about it).
    - Second would be to tackle the donor detection (perhaps getting the bbox as well as matching the donor) would be the next step.
    - Third would be to get started on the grounding (Rama's project). The only problem I can remotely see is that I would need to use the older version of TensorFlow which I would want to avoid as far as possible. If required I would rather change the whole code to the newer version.

- For starters I should write the code for Minimum Spanning Tree and get some kind of provenance graph.

Accomplished :

- Tried the AlexNet (trained on ImageNet) on the dev1. Suprisingly gave 65/65. The only difference is that there were 61 of those ¿ 0.95.

- Also tried the AlexNet (trained on ImageNet) on the protest dataset. Got 7/10 correct.

- Got good enough results on the NC2017_Dev3 dataset using the alex365 (2155/2156) which is quite impressive.

- The next task would be to use an even further world set. Need to download the Eval set to see if there is anything worthwhile.

1.5.2 June 13 : Tuesday

Target :

- Try to use some kind of algo to get better graph than the one found by global thresholding.

- Get the direct bbox from Rama's dataset. Also try to come up with could be potentially be helpful for the direct matching of the two images. One thing that comes to mind is the direct AlexNet trained on ImageNet. But I would seriously like some better and more elegant method.

- Also try it on an even bigger dataset. But I have no clue where can I get the bigger dataset. So have a look at the eval dataset as well.

Notes to self :

- First start with the graph and try to come up with an algo while waiting for Rama's response.

- As soon as the response comes, ditch the graph and go for the matching.

- Also keep on thinking about possible donor matching methods.

- Also if possible keep up with the grounding things.

Accomplished :

- Couldn't really get anywhere using graphs. I need to search for some better graph algorithms.

- Experimented with Rama's dataset with the bounding boxes that he generated. The results are not impressive at all. It shows 5/10 but its more like 4/10 or even 3/10. The bbox he has given are not really good or even comparable to the ground truth.

- Also experimented with the bigger world set with around 17k images. Damn took a long long time : $14594 seconds = 4 hours$. Though I don't think the results will differ much. I am quite confident on this part.

- Also have the kept the code to get the alexnet trained on image net for running on the bigger world set. Need to compare the results, though I am pretty confident that nothing is really going to change.

1.5.3 June 14 : Wednesday

Target :

- Try the graph thingy, show evaluation results on the bigger (world datasets).

- Discuss with prof what else can be done using this.

Accomplished :

- For starters I experimented with different datasets. Especially with both bvlc alexnet and alex 365 on the Nimble 2017 dataset and got quite great results.

- I read up a bit on the R-CNN, but didn't go to completion.

- Also managed to run the test for trying to get all the baseline images, the average was around 49%.

- Had discussion with prof as to start with the splice dataset. Not really much interested but let's get this over with as well.

- Also given to read the disparity pdf, but didn't really read much of it.

1.5.4 June 15 : Thursday

Target :

- Read the report very carefully. Should be able to answer all the questions asked by the Prof.

- Await the code and data from Rama and then start working on them.

- Also keep on thinking about the donor detection and more importantly matching problem.

- If time permits read about the siamese network and other stuff.

Accomplished :

- This is yesterday's data. Complete detection of all baseline images : $1120/2157 = 0.51923$, with $48732/56223 = 0.8667$ accuracy of all the baseline images. I think the result is pretty decent.

- Got better results with the AlexNet trained on ImageNet : $1357/2157 = 0.6311$ with $49974/56223 = 0.8888$ accuracy of all the baseline images.

- Created a small dataset for manipulated images using gimp.

- Then wrote a pretty general code for generating manipulated images with one probe and one donor. Currently donor takes approximately $1/4^{th}$ of the total base image area.

- Also read a bit on the usc-disparity report. But didn't understand that very well. Need to review it once again.

- I am not really able to understand the splice dataset. The images seem to be uncorrelated.

1.5.5   June 16 : Friday

Target :

- Test on the code generated manipulated images.

- Read a bit on the bbox generation algorithm. I honestly think that black box generation is the main bottleneck.

- Then try to see how to improve matching as well.

Notes to self:

- Not much time left. Already week 5 is about to end. Need to seriously hurry up.

- If required I need to work overtime, though today is not possible.

- I need to work at home as well.

- Also imbibe all the basics into me at the earliest.

Accomplished :

- Kept the code to run and generate as many protest dataset as possible. Targetting 10k.

- Got some new tasks by the professor. He says I can start with the donor matching at the earliest, after generating the PR curve.

- Also need to get all the fv of all the places365 train data.

1.5.6   June 17 : Saturday

No work done.

1.5.7   June 18 : Sunday

No work done.

1.6 Week 6 : June 19 - June 26

1.6.1 June 19 : Monday

Target :

- Understand what a PR curve is and generate it on the Nimble Dataset at the earliest.
- Try to get the IOU in the newly generated protest dataset using the implemented code.
- Also let the places365 fv generation happen as much as possible and that too at the earliest.

Accomplished :

- Understood PR, ROC and to some extent CMC. Plotted basic ROC, but need to plot the advanced versions of the algo at the earliest.
- Got the IOU. It is around 0.67 even with all the predictions. But I still need to proof check that it is indeed the case, and that I am not making any blatant error.
- Fv generation is killing a lot of capacity, need to think of something better.

1.6.2 June 20 : Tuesday

Target :

- Complete the ROC, CMC plotting.
- Start with the donor matching part at the earliest.
- Read the paper on Matching Net (found on Arxiv).

Accomplished :

- ROC and CMC giving a bit baffling results.
- Corrected the IOU, now its 0.96, which is quite decent actually.
- Brainstormed about the donor matching part. Wrote quite a few points. Summarized below:
  - Might need a better dataset.
  - Can for now take the best IOU. Also need to take care that donor should be more or less be completely inside the bbx.
  - Still need to figure out a way to see if an image is manipulated, and then see where the manipulation has occured.
  - Can use image segmentation as well as brute force matching (still need to explore this idea a bit further).
  - Need to employ some or the other form of Image Normalization.
  - Specialized methods for text detection and person id. Can try to match it across donors.
- Read a paper on text detection. Interesting paper and the code is also available. Wasn't able to make the Cython part of the code run, but once done, it should be trivial from there on.

1.6.3 June 21 : Wednesday

Target :

- Try to get even better ROC plots.
- Try to run the text detection code at the earliest.

- If detected crop, normalize and check across protest datasets.

- May also try face detection, but not too sure on that part. (Bit behind on the list). Also Kan might already be working on it.

Notes to self :

- The text detection is pretty cool.

- Next task would be to crop the bbox regions, and then try to run a brute force search on the whole world dataset.

- Now I just need to get as many curves (roc, cmc) as possible, more importantly as good as possible.

Accomplished :

- Got the text detection code working. Giving better than expected results. But Rama has a fair point. It will be troublesome if the background images has a lot of text in it giving rise to a lot of false posiitives. Not a sure shot method but interesting results I would say.

- Have put the codes for CMC generation, but I am not sure if they will be useful if at all.

- Also text line bboxes generation code is ongoing.

- Read 2 ML blogs. Read about the Faster-RCNN and Mask-CNN. I think I will try to apply it on the image dataset that I have to see if I can actually get something out of it. I should try for implemented code (need not reinvent the wheel).

- After getting the text boxes I need to generate the fv for both alex, bvlc set.

1.6.4   June 22 : Thursday

Target :

- Since the text lines are not very great, not much interested in pursuing this any further very honestly.

- Around $(1/4)^{th}$ of them had no text detected. This is definitely not a viable approach.

- Should look at Mask-CNN. It may give some better insights (proposals).

- Discuss with professor as to what else can be done to improve it.

- Another idea would be to look at Deep Image Anthology (from arxiv) and can generate some very good examples of manipulations, much better than I would be able to generate on my own I think. If I can make this work, then Hurray.

Notes to self:

- First things first, check with the Text Proposals. But yes I should also give some thought to both CMC and ROC.

- For starters I can simply run the code for text proposals at the earliest, and while it is happening work on other things.

- I should look at both Mask-RCNN as well as the GAN paper.

- First I should complete the GAN tutorial (introduction). Then have a look at the paper briefly to get an idea of what is happening exactly. Then I can directly use it to generate some interesting results. May turn out to be a pretty neat dataset quite honestly.

- I should also give a look at the Mask-RCNN paper and see how good results can we actually achieve. The results might not be as good as predicted, need to be prepared for the worst.

2 Weekly Progress

2.1 Week 1

Accomplished:

- Read 4 papers on Face Recog, but then the project was itself scraped, and I was given a new project.
- Read 2 papers (2nd paper near to completion) on Friday.
- Understood the new project on MediFor, but that might be slightly changed in the near future.

2.2 Week 2

Target :

- Try to complete the NN course (all weeks[16]).
- Read papers on the image geolocalization topic.
- Try to replicate the Places-CNN paper results at the earliest.

Accomplished :

- Didn't really have the time for doing the NN course. Gonna postpone it to next week.
- Read the 4 papers on geolocalization topic.
- Replicated the Places-CNN papers, mainly the alexnet.
- Got a very good hang of Caffe. I think I am ready to fine tune my own dataset.
- Also did a few things on multiprocessing, but bash script seems easier.

2.3 Week 3

Target :

- The firt target would be to get a target. Define the problem clearly. More or less, I am gonna work on the Nimble Challenge dataset. I need to try to use and evaluate the results of Places-CNN on the Nimble Dataset. For this I would basically need to parse all the information as a first step.
- Try to complete NN course (all weeks[16]) this time without fail.
- Also keep on reading some or the papers on the topic Image Geolocalization as well as related to deep learning. Include visionbib and arxiv.

Accomplished :

- Couldn't complete (couldn't even touch) the NN course.
- Tried to read a few papers, but couldn't really get the time.
- The problem is a bit more clear, but I still need more test sets. The progress has been gradual but the results are not really. Seriously need to work harder.
- As such tried the image matching thing using the Places-CNN, and it is giving average results, but more importantly I need to get the number of False Positives also (I already have the false negatives). It looks like that it fails when the image is completely rotated and/or completely different entity is placed on it.

- Also started working on the Nimble Dataset 2017, which is quite an improvement over the previous 2016 dataset. The prof said that there is probably a much larger testset, which he would arrange for me at the earliest.

- In the mean time, I have been doing some examples on TensorFlow, but I am still not upto the mark, unless I am actually able to code some project in it. I am wondering if I should try a hobby project on TensorFlow.

2.4 Week 4

Target:

- Carry out testing on all the available datasets. Getting hold of the larger dataset as well as trying it on the matching part of Rama would be the primary target.

- Also try to implement some ingenious methods like clustering on the smaller dataset. Use as many libraries as possible and avoid having to implement the algorithm on your own.

- Try to at least try to attemp the NN course.

- Have a look at the cool Pytorch implementations of the animeGAN for example, and try out some toy projects in free time.

Accomplished :

- The new dataset dev1, I got very nice results. But on Rama's dataset, the results were honestly too bad. I would actually say very very bad.

- Couldn't really spend time on other projects, but that's ok, as long as I can do something substantial.

- The work has more or less become stagnate. I need to break through this stagnancy.

2.5 Week 5

Target :

- Half way through. This is the point where I absolutely need to make a breakthrough. If not, then again this internship will be more or less a waste.

- Get the results on as many datsets as possible.

- Read papers, if required write code for the same.

- May be required to use some GAN concepts.

- Don't waste time here and there. Focus. Focus.

Accomplished :

- Some kind of breakthrough in the sense that I can start with somewhat different project, but at the same time I need to complete the assigned task at the earliest.

- Got results on quite a few datasets. Especially on the NC2017 dev3 as well as self generated manipulated images. But they gave disappointing results.

- Read papers, but didn't have time to write code for them. More truly I had the time but I didn't really utilize it that well.

- GAN is probably useless in this project.

2.6 Week 6

Target :

- The MAIN PLAN : Work to make a breakthrough. Work with taking breaks only when requried. Other than that don't stop. GO BEYOND! PLUS ULTRA!

- Need to get started with the donor matching as well. I don't have much clue as to how to go about it. But lets see.

- Gotta do extensive experimentation with the given datasets.

3 Month

3.1 Month : May

Accomplished :

- Didn't actually know what I should be working with. Initially the plan was with face id project. But then it was shifted to the MediFor project.

- In the MediFor project, I have read papers on geolocalization but the actual problem to be solved is extremely tough.

- Hence it was decided that I could use similar idea for image matching, and see the results.

- Instead of the medifor database, I am to work with Nimble Challenge.

- And here I am, the results are not up to the mark. Will try some tweaks and see if we can get somewhat better results.

- I am now more or less much more comfortable with Caffe. I am also trying to learn Tensorflow side by side. But for TF it would require a bit more time.

- Answered first question on stack exchange. But the number of questions regarding Caffe seem to be less despite its weak documentation for pycaffe.

- Had a discussion on github issue. The guy was extremely helpful and I have been able to speed up the process by a factor of 10. Though I still think that perhaps I could do bit faster by looking into threading and subprocess library.

- Haven't really been able to find out time for the NN course, and I am slightly disappointed in myself. Also couldn't really follow the CS231n course of Stanford because of lack of time management.

3.2 Month : June

Target :

- This the main/crucial month, and already one day has passed. Each and every day counts. July will flyby very quickly and I won't have time to do anything else other than simple documentation.

- Some broad (really broad) targets complementary to the project would be to complete the NN course (being delayed for quite a while) and complete the CS231n Stanford course. I might need to spend some of the holidays more efficiently. I might also need to manage time better. Also I should be quite familiarized with TensorFlow, and get more proficiency in Tensorflow.

- The main plan. I need to probably work harder to get any kind of results. First I would need to get the results on the Nimble Dataset 2017. It is a great revision to the 2016 dataset, and is more challenging. Much of the previous works may not necessarily yield very good results. But the good thing is, if I can get this working, I will actually be able to

contribute to the Vision community. Moreover this should help complement the MediFor dataset too.

- In addition, I should think about more loss functions or better correlators, or anything that will give me better results. I should also think a bit about Image Geolocalization, though even the simplest matching stuff is in itself still a big challenge. I should also read more papers related to the CNN architectures and deep neural networks. Specifically I think coding them would be a great bonus.