Viterbi Internship - Progress Diary

Arka Sadhu

June 6, 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.

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.

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.