Shubham Jain

Palo Alto, CA | (217)-904-9702 | shubhamjain1310@gmail.com | shubham1310

Education

University of Illinois Urbana-Champaign — GPA: **3.89**/4.0
Master of Science, Computer Science — May 2019

Indian Institute of Technology Kanpur — CPI: **9.9**/10.0
Bachelor of Technology, Computer Science and Engineering — June 2017
Institute Rank: **3**/820

Areas of Interest

Computer Vision, **Deep Learning, Machine Learning**, **Algorithms**

Work Experience

Facial Landmark Network: — San Jose, US

Software Intern in Deep Learning, Drive-IX team, ***NVIDIA*** — May - Aug 2018

- Trained a network (in **Theano**) which takes a face and gives **Facial Keypoints** on eyes, mouth, nose etc as output and also added **eye open/close detection** on top
- Converted this network to **Caffe** and defined a custom layer in Caffe for the Softargmax layer.
- Converted the Caffe network to **TensorRT** and defined custom layers for **Softargmax** and **Upsampling** layer in **fp16** which cut the run time by 50%
- Integrated the TensorRT model in the team's pipeline to be used with other features like face-detect, head-pose, eye open/close etc

Glimpsing Neural Network for Form Image Parsing: — Noida, India

Software Intern under Balaji K, ***Adobe*** — May - July 2017

- Aim of the project was to parse a high resolution form to detect various things such as widgets etc
- The underlying idea was taken from the paper **Show Attend and Tell** which is an attention model based on CNN - LSTM for captioning
- Created the dataset using **MNIST digits** on a black background and predicted the bounded box around the digits in the orderly manner

Sound/Acoustic Generation: — Montreal, Canada

Research Intern under Prof. ***Yoshua Bengio, MILA*** — May - July 2016

- Used **stacked GRU** as the basic model and **quantized** each audio sample as an embedding and added **weight normalization** to the GRU layers
- Applied **transformations** like pitch shifting, time stretch on audio which gave a better generalization
- Further added **conditional information** like different actors and categories of sound in the GRU model to generate a larger variety of sounds and also added **Recurrent Dropout** and **Recurrent Batch Normalization** to our model to make it more robust

Achievements

- Ranked 3rd in the batch of 2017 at IIT Kanpur out of 820 students — 2013-17
- Awarded **Academic Excellence Award, IIT Kanpur** thrice for the **first 3 years** — 2013-16
- Secured 60^{th} rank out of 250 teams in **ACM-ICPC Amritapuri Onsite Contest** — 2014-15
- Secured an **All India Rank** of **210** in **IIT-JEE** among 150,000 aspirants — 2013
- Secured an **All India Rank** of **92** in **JEE MAINS** among 1,400,000 aspirants — 2013
- Selected in nations top 1% in **National Standard Examination in Physics (NSEP, Olympiad)** out of 40,000 students conducted by **HBCSE** — 2012-13

Publications

SampleRNN: An Unconditional End-to-End Neural Audio Generation Model
Soroush Mehri, Kundan Kumar, Ishaan Gulrajani, Rithesh Kumar, Shubham Jain, Jose Sotelo, Aaron Courville, Yoshua Bengio
Published as a conference paper in International Conference on Learning Representations **(ICLR)**, 2017

Skills and Interests

- ***Languages:*** Python, C, C++, MATLAB, BASH
- ***Deep Learning platforms:*** PyTorch, Theano, Tensorflow, Keras
- ***Web Development:*** HTML, CSS, JS, PHP, MYSQL, C#
- ***Tools:*** LaTeX, GNU Octave, Git, GNU Plot, Vim, Verilog

Research Projects

Landmarks for Clothing Retrieval: UIUC

Master's thesis under Prof. Svetlana Lazebnik Aug'18 - Apr 2019

- We worked on In-shop Clothing Retrieval focusing on using features around landmarks (around neckline, shoulders, waistband etc) in the clothing
- We used Cascaded Pyramid Network for landmark detection with Tianchi dataset
- We used DeepFashion InShop dataset retrieval. We used a pre-trained VGG16 network (convolutions only) and used Zoomout to extract features around the landmarks which represented the clothing
- We then used a modified version of contrastive loss over these features to learn similarities between the clothing and fine-tuned the CNN model as well. We then used a K-Nearest Neighbour on these features to do retrieval.
- For the blouses subset, we got an increase of approx 15% in top 3 and top 5 retrieval results using the landmarks extraction methods combined with whole images compared to whole images

Image retrieval: UIUC

Course Project under Prof. Derek Hoeim Aug - Dec 2017

- We viewed this problem as a multi-class classification between objects of different categories taken from datasets like Caltech101 and Caltech 256
- We compared the tree retrieval method by Neister with a deep learning representation method
- For deep learning method the idea was to represent image as a 128vector and the vectors of the same class should be "close" and other class be "far off"
- We used dot product as well as contrastive loss for learning and then for retrieval we used a KNeighbourClassifier on the top of it
- The accuracy for CAL 101 was 83% for Dot product with deep learning and 81% for image retrieval, while for Cal 256 it was 50% and 20%

Image generation with refinement and NN: UIUC

Independent Study under Prof. Svetlana Lazebnik Aug - Dec 2017

- Aim was to tackle two problems that the Generative Adversarial Network (GAN) face which are producing centered images and lack of diversity
- The initial step was to make a network which takes in an image in which objects are cut-pasted and outputs a blended image so that the boundaries doesnt look offbeat and the image seem realistic.
- We used Coco Dataset and had two cut-paste datasets. A simple paste and other with poisson editing on top (to make the problem simple so that generator don't have to deal with edges)
- We tried several hyperparameters to train the model like learning rate, loss factors for generator, training discrimintor once in a while but nothing seemed to have helped till now

Human Centered Computing projects: IIT Kanpur

Course Project under Prof. Nisheeth Srivastava Jan - Apr 2017

We worked on the following (small) projects during the projects:

- **Using Statistics to choose the ideal cricket team:** Used various statistics of players performance to predict the outcome of the match and found out the most relevant attributes for the rankings
- **Studying the effect of a students Internet footprint on Academic performance:** Used browsing history data and categorized the sites using LDA to predict the person's CPI using the amount of time he/she spent on the different categories of sites
- **Call logs- The new personality test:** Tried to predict various personalities scores (eg. openness, neuroticism) of the person using his call logs which contained information about the time of call etc.

Key Courses undertaken

- Computational Photography
- Text Information Systems
- Recent Advances in Computer Vision
- Social Spaces on Internet
- Machine Learning Techniques
- Natural Language Processing
- Data Structures And Algorithms
- Macroeconomics
- Introduction to Economics
- Algorithms II
- Software Engineering I
- Computer Vision and Image Processing
- Human Centered Computing
- Applied Game theory
- Principles of Data Base Systems
- Principles of Programming languages
- Probability and Statistics
- Operating Systems
- Compiler Design
- Analytical Calculus

Other Research Projects

Prediction of dialogues in T.V. series: IIT Kanpur
Undergraduate Project under Prof. Gaurav Sharma and Prof. Vinay P. Namboodiri Aug - Dec 2016

- Aim of the project was to predict the next speaker using previous dialogues and frames.
- Baseline model was based on Bayes rule with the dependence on last 3 speakers
- Used LSTM model based on subtitles only and subtitles with video frames
- VGG-16 and skip-thought vectors were used for feature extraction from video frames and sentences
- We plan to predict the exact dialogue after adding information such as the topic of conversation etc.
- This can be used to find out how predictable a T.V. series is and that can be linked to its popularity

Relative Attributes for Zero-Shot Learning: IIT Kanpur
Course Project under Prof. Vinay P. Namboodiri Jan - Apr 2016

- Used Ranking Models like RankSVM and RankNet, which rank images by learning through features, and compared them
- Used HOG and CNN (VGG-16) based features and compared them on PubFig dataset
- For zero-shot learning made probabilistic (Gaussian) models of all the classes in attribute ranking space using images of seen class and relative (attribute) description of unseen classes

Sentence level grammatical error correction: IIT Kanpur
Course Project under Prof. Harish Karnick Aug - Dec 2016

- Aim of the project was to find out grammatical errors in the sentence and also predict the type of error such as insertion, deletion etc.
- Used Bi-directional LSTM to predict whether there is an error in the sentence at this word or not
- Input was concatenation of one hot representation for the word and the POS tag corresponding to it
- Tried a hierarchical encoder-decoder model for sentences which didn't work very well
- Used NUCLE 3.2 dataset for training and manually tested the results on NIPS 2015 papers

Action Recognition with Trajectory-Pooled Deep-Convolution: IIT Kanpur
Course Project under Prof. Gaurav Sharma Aug - Dec 2016

- Reproduced the results of the paper using the code provided on a subset the UCF101 dataset
- Used a new homography estimation method to speed up the improved trajectory method to find the flow and compared the speed and accuracy with the original improved dense trajectory method on the same dataset

Weakly Supervised Object Localization/Classification for Surveillance: IIT Kanpur
Course Project under Prof. Harish Karnick Jan - Apr 2016

- Tried to solve the problem of localizing and categorizing objects in the given surveillance video dataset which includes objects like cars, bikes, bicycles and pedestrians
- Used Structured Output Tracking with Kernels (Struck), which takes an initial bounding box and uses kernelized SVM to predict the subsequent bounding box locations. A budgeting mechanism is used to prevent the unbounded increase in the support vectors.
- Also tried SIFT matching and OpenCV methods like dilation and contouring which helps in finding minimum rectangular boundary for the objects which are then categorized.

Conflict - Limited Resources and Incomplete Information IIT Kanpur
Course Project under Prof. Vimal Kumar Mar - Apr 2016

- Solved special cases for the game of conflict in which the payoff was related to the amount of weaponry used by the players.
- Solved the game in complete information setting with and without a constrainton the resources and also in the case of incomplete information.

Deep Lambertian Networks: IIT Kanpur
Undergradute Project under Prof. Vinay P. Namboodiri Aug - Dec 2015

- Aimed to solve challenges posed by illumination variation in visual perception, by using different reflectance models which gives the invariant for recognition
- The model uses Gaussian-Restricted Boltzmann Machine to determine the albedo, surface normal and the light source, which are then used to generate an image independent on illumination.

Minimum Spanning tree by Prims algorithms using various heaps: IIT Kanpur

Course Project under Prof. Surender Baswana May - Jul 2015

- Implemented Binary and Fibonacci heap to be used as a part of the algorithm in C language
- Compared their running time on large random graphs, varying the number of vertices and the average number of edges using G(n,p) random graph model
- Studied the research paper which uses Soft heap to further optimize the time complexity

Implementation Based Projects

[Software Development] iTrust

Course Project for Software Engineering I Sept - Dec 2018

- Worked with a team of 8 people on adding 5 use cases to the iTrust application. Specifically, we extended the system for obsterics care patients
- We added use cases like child birth, pregnancy report, office visit, email reminder etc
- The project included working with Java and HTML and we used JUnit for testing

[Search Engine] MovieLike

Course Project for Text information systems Jan - May 2018

- Scraped IMDB website to get movie reviews, plot and other movie details
- Added a search engine on top of the reviews to efficiently match user preferences to opinions from reviews
- Added things like multiple preferences and adjective similarity based query expansion

[Compiler] Making a full fledged compiler for Java (in python) to be run on a x86 machine

Course Project for Compiler Jan - Apr 2016

- Developed an Compiler for a subset of Java for x86 architecture
- Implemented Lexical Analyzer, Assembly Code Generator and constructed grammar rules for parsing.
- Constructed rules for converting Java code to Three Address Code
- Gave support of all the basic operation, conditional statements, Iterative statements, recursive and nested parameterized function, function overloading, scope handling, multidimensional arrays

[Operating System] NachOS

Course Project for Operating Systems Aug - Dec 2015

- Implemented System calls pertaining to Fork, Exec, Join, Sleep and Exit
- Implemented UNIX, First in First Out, Round Robin and Shortest Job First job scheduling algorithms
- Implemented Shared Memory, Semaphores, Condition variables and Demand Paging

[App Development] Project Sporada, Powering phones with sparse 2G:

3rd Inter-IIT Tech Meet Jan - Feb 2015

- Made a Java stand-alone application through Netbeans framework and used various python modules like sphinx
- Built a system to make relevant or popular content available on phones (offline)
- Managed the regular updates of the server, storage available and user's search experience

[App Development] Life Speaks

Microsoft's Code.fun.do, 24hrs competition Feb 2015

- Made a C# Application which converts text, images(like quotes), pdf(text-only pdf) to speech
- Created an audio stream and output speech using Microsoft's API
- Used API's which provides classes for Optical Character Recognition (OCR) that enables Windows Runtime apps to read and interpret text from images

[Interpreter] Making a Programming Language

Project under ACA, IIT Kanpur Dec - Apr 2014

- Developed an interpreter for the moves of a robot like turn, shoot, shield etc using Python 3.4
- Used the turtle module of Python for graphics and also added sound effects using PyGame module

[Web Development] Programming Club website

Project under Programming Club IIT Kanpur May - Jun 2014

- Made the front end of the website and added login, events tab, tutorial, project database (storing them on server side)
- Used Curl (PHP) for web scraping to make a database of top Codechef users from IITK
- Added technical tweets, Google Calendar, Phpbb forum; created a logo and slideshow using Javascript

Positions of Responsibility

- **TA for course Introduction to Computer Science CS101**
  To help the professor in making and grading the assignments and exams for 4 semesters during the masters degree.
- **TA for course Data Structures and Algorithms**
  To help the professor in making and grading the assignments and exams
- **Project Mentor**
  Mentored 5 first year students for a semester long project (under ACA, IIT Kanpur)