Vasu Sharma

Junior Year Undergraduate
Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur
Webpage: home.iitk.ac.in/~vasus

Email: sharma.vasu55@gmail.com
Phone: (+91) 9621364646
Address: Room D-101, Hall 1, IIT Kanpur

EDUCATIONAL QUALIFICATIONS

Year	Degree/ Certificate	Institute/ School, City	CPI/ %
2012-present	Bachelor Of Technology, Computer Science and Engineering, minor in Signal Processing(Electrical Engg.)	Indian Institute Of Technology, Kanpur	CPI=**9.8/10.0**
2012	AISSCE(Class XII CBSE Board)	St.Columba's School ,New Delhi	**97%**
2010	AISSCE(Class X CBSE Board)	St.Columba's School ,New Delhi	CGPA=**9.4/10**

PUBLICATIONS

- **"A Deep Neural Network Based Approach For Vocal Extraction From Songs"**
    - Accepted for IEEE's International Conference on Signal and Image Processing Applications 2015. (Single Authorship Paper)
    - This paper proposes several Deep Learning Frameworks for extracting Vocals from songs.
    - Deep Learning Frameworks like MultiLayer Perceptrons, Autoencoders, Restricted Boltzmann Machines and their extension to Deep Belief Nets were trained and their performance on this task was compared.

- **" AutoTag : Automatic Image Annotation and Content Based Image Retrieval"**
    - Submitted to European Conference on Computer Vision (ECCV) 2016, Amsterdam.
    - In this paper we propose an Automatic Image Annotation system based on a Deep Learning pipeline in conjunction with the FastXML algorithm to build a state of the art Image Annotation system on the IAPR-TC12, ESP-Game, MIRFlickr and several other benchmark datasets.
    - We also build an extremely fast content based image retrieval system by extending this approach.

- **"Automatic tagging and retrieval of E-Commerce products based on visual Features"**
    - Submitted to NAACL, Association for Computational Linguistics(ACL) conference, San Diego, 2016
    - This paper essentially uses a Deep Learning and Multi Label Classification based pipeline to identify important visual features in E-Commerce products and assigns them relevant tags. We then create a query based product retrieval system on top of it.

- **"Image Summarization using Topic Modelling"**
    - Accepted for IEEE's International Conference on Signal and Image Processing Applications 2015
    - This paper proposes a method to use topic modeling and clustering based techniques to summarize a large collection of images into a smaller, representative subset.
    - This paper describes how the techniques of topic modeling like Latent Dirichlet allocation, Latent semantic analysis along with methods like K-Means Clustering etc. can be used for image summarization.

- **"Analyzing Newspaper Crime Reports For Identification Of Safe Transit Paths"**
    - Accepted for NAACL, Association for Computational Linguistics(ACL) conference, Colorado, 2015
    - This paper proposes a method to find the safest path between two locations, based on the geographical model of crime intensities
    - This paper describes how the techniques of topic modeling like Latent Dirichlet allocation, Latent semantic analysis along with mathematical modeling and NLP techniques were successfully employed to identify the safest path based on model of crime intensities.

- **"Automatic Sign Language Recognition Systems based on Deep Neural Nets"**
  - Submitted to IEEE's Visual Communications and Image Processing, 2015. (Single Authorship paper)
  - This paper proposes several Deep Neural Network architectures and uses them to perform Automatic Sign Language Recognition. MultiLayer Perceptron, Autoencoders, Convolutional Neural Nets and Restricted Boltzmann machines were trained and state of the art performance was obtained by certain configurations
- "**A Gated Recurrent Neural Network for denoising speech signals**"
  - Submitted to IEEE's "Spoken Language Technology" Conference, Nevada, 2014
  - The paper focuses on Signal reconstruction and noise removal to denoise noisy signals by using deep and recurrent neural networks.
  - The neural network design is used to combine noise estimation and denoising into a single network design, effectively doing both simultaneously. The focus is to enhance intelligibility of speech signals by successfully removing the noise without distorting the original signal

PROJECTS AND INTERNSHIPS

- **Summer Internship '14 at School Of Computer Science, Carnegie Mellon University**
  (Mentors: Dr. Rita Singh , Dr. Bhiksha Raj)   *(May –Jul '14)*
  - **"Application of Deep Learning for Audio and Speech Signal Processing"**
  - The project focused on construction of a **Deep Recurrent neural network** to achieve signal reconstruction by denoising noise corrupted signals by **dynamic spectral subtraction.**
  - Techniques used: **Recurrent and Time Delay Neural Networks, Spectral Subtraction, Multi Layer Perceptrons and other Deep Learning techniques**
- **Winter Internship '15 at Xerox Research Labs, Europe**
  (Mentors: Dr. Larlus Diane and Dr. Albert Gordo)   *(Sep –Dec '15)*
  - Worked on '**Large Scale Image Recognition using Deep Nets**'
  - The projects primarily focused on constructing Deep Learning frameworks for Image Recognition. Coded AlexNet, GoogleNet and several other architectures from scratch and also worked on designing some new Deep Learning frameworks for the image recognition task on the ImageNet dataset. Also made extensive use of GPU's and the popular Caffe library for training Deep Convolutional Neural Nets.
- **Summer Internship '15 at Xerox Research Labs**
  (Mentors: Dr. Vivek Tyagi)   *(May –Sep'15)*
  - Worked on 3 projects during this internship: **"Application of Deep Learning for Automatic Speech Recognition", " A comprehensive analysis of Activation Functions in Deep Nets"** and **"A new hashing technique to enhance Deep Net performance ".** Also got the Best Project award for the same.
  - The projects primarily focused on constructing Deep Learning frameworks for Speech Recognition. The Internship provided me extensive research and coding experience of how to efficiently train Deep Nets. **Several research papers** and **patents** are also in process of being written on the same.
- **'Real Time Video Surveillance using Deep Convolutional Neural Networks'**
  (Course Project Machine Learning Techniques under Prof. Harish Karnick   *(Jan –Apr'16)*
  - Implemented a real time video surveillance system performing image segmentation and entity recognition using a **Deep Region based Convolutional Neural Networks**.
  - Applied the concept **of Transfer Learning** by training the model on Pascal VOC and using it on the large amount of unlabeled video surveillance data. The model yielded **state of the art performance** and was fast enough to be able to perform the processing in **real time**.
- **'Identifying safest path in Real time based on crime records'**   *(Mentors: Dr. Bhiksha Raj , Dr. Rita Singh)*
  Carnegie Mellon University's Winter School   *(Dec '13)*
  - Awarded **Overall Best project.**
  - Selected for **Gandhian Young Technological Innovation Award**   *(Mar '14)*
  - This project was an attempt to **counter the growing crime rates** throughout the world by ensuring commuter safety by providing him with the safest transit path between 2 locations computed in real time.
  - Techniques used: **Latent Dirichlet Allocation, Latent Semantic Analysis, Named Entity Recognition, k nearest neighbors , Bag of Words model, Probabilistic modeling etc.**.

- '**Image Summarization using Topic Modelling**' *(Mentors: Dr. Bhiksha Raj , Dr. Rita Singh)*
  Carnegie Mellon University's Winter School *(Dec '13)*
    - o Awarded **Overall Best project.**
    - o The aim of this project was to design an **Automatic Image Summarization software** which would choose a subset of a large collection of photos choosing only the **best images representative of the entire set** hence automatically creating an image summary.
    - o Techniques Used: **Scale Invariant Feature Transforms, Latent Dirichlet Allocation, K means clustering, Bag of Words Model etc.**
- **'Online File Sharing System with Collaborative Editing'**
  (Course Project Computing Laboratory-II under Prof. Arnab Bhattacharya) *(Sep -Nov'14)*
    - o Implemented file/folder sharing among multiple users, allowing upload/download from remote server.
    - o Modified ShareLatex to integrate with the system, allowing multiple users to collaboratively edit and render shared Tex files collaboratively. Used python client to synchronize the local and remote file systems.
- **'Extension of NachOS'**
  (Course Project Operating Systems under Prof. Mainak Choudhari) *(Sep -Nov'14)*
    - o Designed various functionalities in NachOS - instructional software in C++ to run as secondary OS on linux.
    - o Implemented various system calls (fork, join) , various scheduling algorithms (FIFO, RR, Unix Scheduler), various techniques for synchronization (semaphores, condition variables), demand paging, shared memory and various page replacement algorithms(FIFO, LRU, LRU Clock).
- **'Education based Webapp for Primary education'** *(Mar '14)*
    - o Secured **2nd Position** in El-Eduvate (**Techkriti, Technical Festival, IIT Kanpur**)
    - o The webapp was designed with several interactive applications and a highly attractive GUI to provide primary school students an exciting platform to learn new things away from the monotonous classroom environment.
    - o Used HTML, CSS, Javascript, PHP to design webapp.
- ' **Karaoke Generator Webapp with vocal comparison'** *(Yahoo! 's R & D team)*
    - o **Honorable mention.** *(Oct '13)*
    - o Given a song, the webapp muted the vocals, fetched the lyrics and allowed the user to sing along it.
    - o Techniques Used: **Audio Processing, Vocals extraction and removal, Heuristic based vocal comparison.**

TECHNICAL SKILL SET

- **Programming Languages**: C, C++, Python + Theano, Caffe, Torch, Pylearn2, Matlab, R, SQL, HTML, Javascipt, PHP, Bash, Perl, OpenCV

RELEVANT COURSES DONE

Code	Course Name	Grade	Code	Corse Name	Grade
CS210	Data Structures And Algorithms	A	CS335	Compiler Design	Ongoing
MTH102	Linear Algebra and Ordinary Diff. Eqn's	A*	CS345	Advanced Algorithms	A
MSO201	Probability and Statistics	A*	EE200	Digital Signal Processing	A
CS771	Machine Learning Techniques	Ongoing	MTH552	Data Mining and AI	Ongoing
CS202	Intro. To Logic	A	CS330	Operating Systems	A*
Coursera	Digital Image and Video Processing	Distinction	Coursera	Data Science Specialization (7courses)	Completed