

Abhinav JAIN

Research Engineer

github.com/jabhinav

+91 951 918 9181 @ User ID : jabhinaviitk@gmail.com

Delhi, India

EDUCATION

APR 2017	INDIAN INSTITUTE OF TECHNOLOGY, KANPUR, INDIA
APR 2013	BACHELOR OF TECHNOLOGY (B.TECH) IN ELECTRICAL ENGINEERING Cumulative Performance Index, CPI : 9.0/10.0
APR 2013	JAWAHARLAL NEHRU SCHOOL, BHOPAL, INDIA
APR 2012	ALL INDIA SENIOR SCHOOL CERTIFICATE EXAMINATION, AISSCE, CLASS XII Percentage : 91.8%

PROFESSIONAL EXPERIENCE

PRESENT	STAFF RESEARCH SOFTWARE ENGINEER IBM RESEARCH LABS, INDIA
JUL 2019	DATA TRANSFORMATION, PROJECT LEAD <i>Project Manager</i> : HIMA PATEL <i>Description</i> : Developed a Programming-by-Example system to learn transformation programs for homogenising multiple synonymous labels of categorical features in structured datasets. The developed framework learns a transformation routine from few input-output example pairs provided by the user. Service provided as part of Data Readiness Toolkit hosted by IBM Research Labs. Symbolic System for Program Synthesis Automated AI Few Shot Learning PyTorch
JUL 2019	SOFTWARE DEVELOPER IBM RESEARCH LABS, INDIA
JUL 2017	SCANNED PDF-TO-HTML CONVERSION, WATSON COMPARE & COMPLY <i>Project Manager</i> : DR. SAMEEP MEHTA <i>Description</i> : Structured data extraction from scanned PDF documents and converting them into “consumable” representations (HTML/JSON) so that it can be served as input to Business Automation Tools. <ul style="list-style-type: none">> Reorientation of skewed text for reliable text extraction.> Multi-lingual support for converting documents when language of the text is changed.> Data extraction from hybrid PDF documents containing scanned and programmatic content.> Detect logos, bar-codes and signatures for downstream processing such as querying, retrieval. Text Image Processing Tesseract Gradle OpenCV Java
AUG 2016	INTERN IBM RESEARCH LABS, INDIA
MAY 2016	COHERENT VISUAL DESCRIPTION OF TEXTUAL INSTRUCTIONS <i>Project Manager</i> : DR. SAMEEP MEHTA <i>Description</i> : Developed a multi-stage framework to provide visual aid for a sequence of text-based instructions in the form of coherent images for better comprehension of the text. <ul style="list-style-type: none">> For each instruction, visualisable phrases consisting of head action verbs and noun phrases are mined using standard practices like POS tagging, Dependency parsing and Co-reference resolution.> For each visualisable phrase, a set of images is retrieved from sources such as WikiHow, Flickr.> Across instructions sharing common information in the form of latent/non-latent entities, coherency is maintained using a graph based matching method utilising Dijkstra’s algorithm. A user study was conducted to validate improvement in understanding of text instructions when phrases and images together dictate the action being conducted in the instruction. github.com/jabhinav/IBM-project textToImage Content Mining Graph Matching Convolutional Neural Networks Python

VIDEO REPRESENTATION LEARNING FOR FINE-GRAINED SCENE RECOGNITION AND RETRIEVAL

2018-19

github.com/jabhinav/Deep-Video-Understanding

Advertising in digital media often requires recognition of critical scenes in videos for smart placement of brand advertisements. These critical scenes raise viewer anxiety and are a part of some parent activity. We distinguish them from the rest of non-critical scenes using an order-preserving fine-grained similarity metric that learns the required representations. The learned metric is tested in two novel tasks : critical **scene recognition** and fine-grained **video retrieval**. To learn the metric, we proposed Pentuplet Loss [3] and later on, an improved and more robust Radial Loss [4]. These losses exploit ‘Quadlet Sampling’ to mine data where each training sample is a tuple of query, positive, intermediate and negative samples. Lastly, to ascertain the effectiveness of the loss in learning a deep metric for measuring similarities, we tested its performance against state-of-the-art baselines in the known tasks : fine-grained image retrieval and shot-boundary detection.

Deep Metric Learning | Event Recognition | Content Based Retrieval or Ranking | Shot Boundary Detection | CNN-LSTM Siamese Nets | Loss Formulation

DEEP LEARNING FOR DOCUMENT IMAGE QUALITY ENHANCEMENT

2017-19

[github.ibm.com/abhinavj/Super-Resolving-Documents](https://github.com/abhinavj/Super-Resolving-Documents)

Poor quality scanned document images suffer from low token fidelity when an OCR engine such as Tesseract is used for token extraction. To remedy this, we leveraged deep learning based solutions for document quality enhancement [6] and delivered the same for public release as part of IBM’s Watson API. In [6], we formulated a novel ‘Text Quality Improvement Loss’ for the standard super-resolution convolutional neural network (SRCNN) to generate high-resolution text images. The proposed framework identifies text regions from images and minimizes additional MSE between such localised regions on top of the standard MSE, enforced by Single Image Super Resolution frameworks. This results in simultaneous optimisation of perceptual quality of the image and the OCR performance.

SRCNN for SuperResolution | Loss Formulation | OCR Boosting | UNLV Documents Dataset | Caffe

EDUCATION ENRICHMENT

2017-18

github.com/jabhinav/EducationEnrichment

Formal texts such as journal articles are composed of complex terminologies intended to be understood by targeted demographic. In absence of domain knowledge, they tend to be more ambiguous for general readers. To avail a complete semantic understanding of such texts for the readers, we proposed an enrichment system [2] that mitigates the problem of searching for required information through heaps of sources. The system augments given text with required concept definitions, applications and concept dependency graphs. Our framework extracts key-concepts (technical terms) based on user discretion via a sequence of filtering stages - Linguistic Filtering, BBC Pruning and StackExchange Pruning. It detects the presence of required information by classifying each associated sentence into definition/application of the key-concept using a CNN-LSTM network. The same framework also runs on a data source such as Wikipedia to return the concept’s missing definition or real-life application.

NLP | AI for Education | Sentence Classification | Information Extraction | Dependency Parsing

EVOLVING AI

2017-18

github.com/jabhinav/Model-Learning

In this project, we address the problem of re-training a deep neural network for a new class with limited training data (‘n’ to ‘n+1’ class learning) using a novel concept of Deep part embeddings (DPEs)[5]. DPEs are sub-networks of neuron activation extracted from a trained network identifying a visual and distinguishable element of a class. We identify visual elements that intuitively constitute a new class and extract the corresponding DPEs from the network pre-trained for the class sharing the identified visual element. Finally, we assemble them into a new network and re-train the model on limited samples of the new class and a subset of data from ‘n’ classes to achieve high accuracy on the new class without affecting the accuracy of n classes. We studied and produced results for DPE integration under two configurations : (i) sequential, when DPEs are sourced from different CNN architectures and (ii) shared; when DPEs are sourced from the same CNN architecture.

Few-shot learning | Knowledge Transfer | Model Learning | Deep Part Embeddings | CNN Visualisation | CNN Activation Filtering

PUBLICATIONS

- DEC 2017 | [1] S Mujumdar, N Gupta, A Jain, S Mehta, “Coherent Visual Description of Textual Instructions,” in *IEEE International Symposium on Multimedia (ISM)*.
- AUG 2018 | [2] A Jain, N Gupta, S Mujumdar, S Mehta, R Madhok, “Content Driven Enrichment of Formal Text using Concept Definitions and Applications,” in *Proceedings of the 29th on Hypertext and Social Media (HT)*.
- AUG 2018 | [3] N Gupta, A Jain, P Agarwal, S Mujumdar, S Mehta, “Pentuplet Loss for Simultaneous Shots and Critical Points Detection in a Video,” in *International Conference on Pattern Recognition (ICPR)*.

- APR 2019 | [4] A Jain, P Agarwal, S Mujumdar, N Gupta, S Mehta, C Chattopadhyay, "Radial Loss for Learning Fine-grained Video Similarity Metric," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- APR 2019 | [5] N Gupta, S Mujumdar, P Agarwal, A Jain, S Mehta, "Learning Convolutional Neural Networks with Deep Part Embeddings," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- SEP 2019 | [6] S Mujumdar, N Gupta, A Jain, D. Burdick, "Simultaneous Optimisation of Image Quality Improvement and Text Content Extraction from Scanned Documents," in *IEEE International Conference on Document Analysis and Recognition (ICDAR)*.
- [7] Paper on Programming-by-Example for Data Transformation using Reinforcement Learning (Manuscript submitted for publication)

PATENTS

- 2018 | **System and Method to Generate Dynamic Personalized Infographics**
 A system is proposed to automatically generate info-graphics from floating social media trends, personalised to user interests with the help of unstructured data available across different online platforms like Facebook, Twitter, News etc. The generated info-graphics are personalized by specifying the visual and non-visual elements specific to the user such as sketch or caricature view for the former and personalisation elements such as viewer age-group, emotions, sentiment, sarcasm etc. for the latter.

ACADEMIC PROJECTS

- Nov 2016 | **TOPIC MODELLING**
 Aug 2016 | *Supervisor* : Dr. Gaurav Pandey, Dr. Piyush Rai, I.I.T KANPUR
Course : UNDER-GRADUATE PROJECT
- > Discover underlying latent themes(topics, meta-topics) of Newsgroup Corpora using topic modelling.
 - > Formulated a survey report comparing hitherto proposed state-of-the-art algorithms - Poisson Factor analysis (PFA), PFA with Sigmoid Belief Networks, Deep PFA (DPFA) and Discriminative-DPFA.
 - > Compared aforementioned algorithms based on their formulation, underlying assumptions, advantages-disadvantages and their document label classification performance.
- Topic Modelling | Document classification | Probabilistic Machine Learning | Gibbs Sampling
- Jul 2015 | **VISION BASED SURVEILLANCE AND TRACKING SYSTEM FOR UAVS**
 May 2015 | *Supervisor* : Dr. N.K. Verma, I.I.T KANPUR
Course : SUMMER PROJECT
- > Real-time implementation of the following vision based object tracking algorithms to gain hands-on experience - Optical Flow Approach, Background Subtraction, Interest Point Tracking and Real-Time Compressive Tracking.
 - > Extensive testing and comparison of the algorithms' robustness against the following factors - pose variation, illumination, occlusion, and motion blur.
 - > Integrated the best performing, Real-Time Compressive Tracking with UAV for real-time testing.
- Unsupervised Object Tracking | Survey | Computer Vision
- Dec 2016 | **BAYESIAN OPTIMIZATION FOR HYPER-PARAMETER TUNING**
 Aug 2016 | *Supervisor* : Dr. Purushottam Kar, I.I.T KANPUR
Course : OPTIMIZATION TECHNIQUES
- > Studied and compared Bayesian Optimization frameworks - Gaussian Process based and Deep Network for Global Optimization, DNGO.
 - > Demonstrated the improvements offered by the Bayesian optimization selection strategies against random selection and hard coding done by experts for automatic hyper-parameter tuning in the task of MNIST digit classification via sparse auto-encoders stacked with softmax classifier.
- Bayesian Optimisation | Gaussian Process | DNGO | Spearmint | Matlab-Python

FUTURE COURSEWORK

Logic And AI
Deep Reinforcement Learning
Learning based Methods in Computer Vision
Designing AI to cultivate Human Well-Being
Representation Learning in Computer Vision

COMPLETED COURSEWORK - IIT KANPUR

Data Structures and Algorithms Signals, Systems & Networks
Probabilistic Machine Learning Probability and Statistics
Introduction to Stochastic Processes Image Processing
Partial Differential Equation Visual Recognition
Online Learning and Optimization Linear Algebra

EXTRACURRICULAR ACTIVITIES - IIT KANPUR

2015 Event Coordinator, Antaragni Leadership Initiative(ALI)
2013-14 Runner-Up, Football Freshers Inferno
2014-15 Secretary, Fine Arts Club
2013-15 Member, Formula SAE

SKILLS

Java ● ● ● ○
Python ● ● ● ●
PyTorch ● ● ● ●
Keras ● ● ● ●
OpenCV ● ● ● ○

AWARDS AND ACHIEVEMENTS

- 2019 | **RDA, IBM RESEARCH LABS**
IBM's Research Division Award (RDA) for "Document Conversion and Table Understanding"
- 2019 | **CERTIFICATE OF APPRECIATION, IBM RESEARCH LABS**
Certificate was issued in honour of my contributions to IBM Watson Compare & Comply project which led to significant business and research impact
- 2016 | **BEST POSTER AWARD, IBM RESEARCH LABS**
Achieved recognition for outstanding presentation carried out during internship
- 2014-15 | **ACADEMIC EXCELLENCE**
Awarded for exceptional performance in academics at I.I.T Kanpur
- 2013 | **JOINT ENTRANCE EXAMINATION**
Secured **All India Rank 655** (amongst 150,000 students) in JEE-Advanced 2013
- 2012-13 | **BEST ALL ROUNDER AWARD**
Awarded Student of the year in the senior year of high school for all-round excellence