

HAMMAD A. AYYUBI

hayyubi@cs.columbia.edu \diamond <https://hammad001.github.io>

EDUCATION

Columbia University

PhD in Computer Science, GPA: 3.97/4.0

Advisor: Prof. Shih-Fu Chang

New York, NY

Sep. 2020 - Present

University of California, San Diego

Master of Science in Computer Science, GPA: 4.0/4.0

Advisor: Prof. Gary Cottrell, Prof. Manmohan Chandraker

San Diego, CA

Sep. 2018 - June 2020

Indian Institute of Technology, Banaras Hindu University

Bachelor of Technology in Electrical Engineering, GPA: 8.7/10

Varanasi, India

May 2012 - May 2016

WORK EXPERIENCE

Google

Student Researcher

Mentor: Dr. Tianqi Liu

New York, NY

June 2023 - Nov. 2023

Microsoft Research

Research Intern

Mentor: Dr. Oriana Riva and Dr. Jianwei Yang

Redmond, WA

June 2022 - Aug. 2022

SRI International

Machine Learning Research Intern

Mentor: Dr. Yi Yao and Dr. Ajay Divakaran

Princeton, NJ

June 2019 - Sep. 2019

Soroco India Pvt. Ltd.

Software Engineer - Deep Learning & Computer Vision

Bangalore, India

Feb 2018 - Aug 2018

Citicorp Services India Pvt. Ltd.

Software Developer

Pune, India

July 2016 - Jan. 2018

PUBLICATIONS

1. **Hammad A. Ayyubi**, Rahul Lokesh, Alireza Zareian, Bo Wu, and Shih-Fu Chang. Learning from Children: Improving Image-Caption Pretraining via Curriculum. *ACL Findings*, 2023
2. **Hammad A. Ayyubi**, Tianqi Liu, Arsha Nagrani, Xudong Lin, Mingda Zhang, Anurag Arnab, Feng Han, Yukun Zhu, and Jialu Liu. Video Summarization: Towards Entity-Aware Captions. *In Submission CVPR*, 2024
3. **Hammad A. Ayyubi**, Christopher Thomas, Lovish Chum, Rahul Lokesh, Yulei Niu, Xudong Lin, Long Chen, Jaywon Koo, Sounak Ray, and Shih-Fu Chang. Beyond Grounding: Extracting Event Hierarchies from Multimodal Content. *AAAI*, 2024
4. Haoxuan You¹, Rui Sun, Zhecan Wang, Long Chen, Gengyu Wang, **Hammad A. Ayyubi**, Kai-Wei Chang, and Shih-Fu Chang. Idealgpt: Iteratively Decomposing Vision and Language Reasoning via Large Language Models. *EMNLP Findings*, 2023
5. Long Chen, Yulei Niu, Brian Chen, Xudong Lin, Guangxing Han, Christopher Thomas, **Hammad Ayyubi**, Heng Ji, and Shih-Fu Chang. Weakly-supervised Temporal Article Grounding. *EMNLP*, 2022
6. Haoxuan You, Zhecan Wang, Alireza Zareian, Liunian Harold Li, **Hammad A. Ayyubi**, Kai-Wei Chang, and Shih-Fu Chang. Learning Knowledge-aware Multimodal Representation for Visual Commonsense Reasoning. 2022

7. **Hammad A. Ayyubi***, Md. Mehrab Tanjim*, Julian McAuley, and Garrison W. Cottrell. Generating rationale in Visual Question Answering. *arXiv:2004.02032* 2019, 2020
8. **Hammad A. Ayyubi**. Leveraging Human Reasoning to Understand and Improve Visual Question Answering. *MS Thesis, UC San Diego*, 2020
9. **Hammad A. Ayyubi**, Yi Yao, and Ajay Divakaran. Progressive growing of Neural Ordinary Differential Equations. *ICLR Workshop on Integration of Neural Networks and Differential Equations*, 2020
10. Md. Mehrab Tanjim, **Hammad A. Ayyubi**, and Garrison W. Cottrell. Dynamicrec: A Dynamic Convolutional Network for Next Item Recommendation. *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*, 2020

ACADEMIC RESEARCH & PROJECTS

Action Reasoning via State Changes in Videos

June. 2022 - May 2023

Advisor: Dr. Oriana Riva and Dr. Jianwei Yang

- Proposed a new task to predict actions and state changing objects from start state and end state frames in videos. Also, proposed counterfactual probe tasks to test action reasoning.
- Proposed a transformer based slot attention model to detect state changing object and predict action.

Visual Commonsense Reasoning

Jan. - June 2019

Advisor: Prof. David Kriegman

- Task: Answer a question, given an image, and also provide a rationale.
- Proposed novel end to end joint learning of answer and rationale prediction by using softmax, gumbel-softmax and reinforcement learning approaches to tackle non-differentiability.
- Implemented in PyTorch; used Docker and Kubernetes cluster for running multi-GPU tasks.

Generative Adversarial Network (GANs) Inspection

Apr. - June 2019

- Inspected latent manifold learned by DCGAN and PgGAN through various interpolation, extrapolation and vector arithmetics techniques.
- Proved that semantic relations are learned in the manifold.

Optical Character Recognition using Deep Learning

Feb. - June 2018

Advisor: Dr. Mohsen Malmir

- Researched various CNN models - U-Net, DeepLab v3+ - to segment text from images and PDFs.
- Used novel multi-task learning approach to instance segment, recognize and detect words.
- Achieved a recall of 97.8% on train set and 95% recall on test set.

Knowledge Distillation for Semantic Segmentation

July - Aug. 2018

Advisor: Dr. Mohsen Malmir

- Compressed U-Net into a much smaller model without decline in accuracy.
- Reduced model parameters by 84%, improved speed by 20% and reduced memory usage by 21%.

SKILLS

| | <i>Expert</i> | <i>Intermediate</i> | <i>Familiar</i> |
|----------------------------------|--|---------------------------|-----------------|
| Programming Languages | Python, Java, Matlab | C, C++, Bash, MySQL | Clojure |
| Libraries & Framework | PyTorch, Docker, Kubernetes, OpenCV, Numpy, Spring | Keras, Tensorflow, Hadoop | BOOST |

AWARDS & HONORS

| | | |
|------|-----------------------|--|
| 2019 | Runner's Up Award | Computer Vision Poster Competition, SRI Princeton |
| 2019 | People's Choice Award | Computer Vision Poster Competition, SRI Princeton |
| 2012 | JN Tata Scholar | JN Tata Endowment for the higher education of Indians |
| 2011 | Top 0.1%ile | Computer Science, Central Board of Secondary Education |

PROFESSIONAL SERVICES

| | | |
|------|-------------------|---|
| 2023 | Reviewer | ACL |
| 2022 | Reviewer | EMNLP |
| 2021 | Reviewer | EMNLP |
| 2021 | Program Committee | Visually Grounded Interaction and Language (ViGIL), NAACL |

RELEVANT COURSEWORK

| Graduate | Undergraduate | Online |
|---|--|------------------|
| Computational aspects of Robotics | Algorithms & Data Structures | DeepLearning.ai |
| Computer Vision | Artificial Intelligence & Expert Systems | Machine Learning |
| Deep Learning for sequence data | Calculus | |
| AI - Probabilistic Graphical Models | Linear Algebra | |
| Statistical Learning | | |
| Statistical Natural Language Processing | | |

TEACHING ASSISTANT

| | | | |
|------|-----------|--|---------------------------------------|
| 2022 | COMS-4995 | Deep Learning for Computer Vision | Computer Science Dept., Columbia Uni. |
| 2020 | CSE-291D | Advanced NLP | Computer Science Dept., UC San Diego |
| 2020 | CSE-250B | Introduction to AI: A Statistical Approach | Computer Science Dept., UC San Diego |
| 2019 | CSE-250A | AI - Probabilistic Graphical Models | Computer Science Dept., UC San Diego |
| 2016 | CS0 101 | Computer Programming | Computer Science Dept., IIT BHU |