

Hassan Akbari

Columbia University, Schapiro Research Building (CEPSR), Room 7LE3, 530 W 120th St., New York, NY 10027
ha2436@columbia.edu • <https://hassanakbari.com>

RESEARCH INTERESTS

- Computer Vision
- Deep Learning
- Multimodal Understanding (Text-Vision-Audio fusion)
- Visual Reasoning
- Digital Signal Processing (Multimedia, Biomedical Signals/Images)

EDUCATION

Columbia University in the City of New York, New York, NY, USA

Doctor of Philosophy (Ph.D.) in Electrical Engineering

Sep 2016 – Present

• **Adviser:** [Prof. Shih-Fu Chang](#)

• **Research:** I have recently started working on context-based visual reasoning on videos to extract probabilistic schema of events in videos. Prior to that, I had worked on methods for bridging vision with other modalities (e.g. text, audio) resulting in a fused common space representation which could be potentially used in a variety of applications such as Image-Phrase Grounding, Visual Question Answering (VQA), Video/Image Captioning, Video Perception, Video-Audio Translation, etc.

• **Cumulative GPA: 4.1/4.0**

Sharif University of Technology, Tehran, Iran

Master of Science (M.Sc.) in Biomedical Engineering (Bio-electrics)

Sep 2013 – Nov 2015

• Research Area: Blind Source Separation, Multilinear Feature Extraction

• **Cumulative GPA: 18.17/20**

Yazd University, Yazd, Iran

Bachelor of Science (B.Sc.) in Electrical Engineering (Telecommunication)

Sep 2009 – Jun 2013

• Research Area: Speech Recognition

• Graduated with College Honors.

• **Cumulative GPA: 18.01/20**

SKILLS

Programming Languages

- Proficient: Python ([Sample Link](#)), Matlab
- Familiar with: C/C++
- Touched: C#

Related Knowledge

- **Computer Vision:** Object Detection, Classification, Grounding, Visual Reasoning, Instance Matching, Weakly Supervised Detection/Grounding, Image-Text/Video-Audio Fusion.
- **Deep Learning:** Network Architecture Development including RNN, CNN, BiLSTM, CharCNN, Autoencoders, GANs, VAEs, Deep RL.
- **Deep Learning Libraries:** Tensorflow, Tensorboard, TensorHub, PyTorch, Caffe, Keras API
- **Algorithms and CS Knowledge:** Big-O Analysis, Data Structures, Sorting, Trees, Graphs, Hashing, Heaps, Dynamic Programming, Recursion, BFS, DFS.
- **Graph:** Graph Construction, Analysis, and Retrieval.

Technical Software/Library

- Matlab, T_EX, OpenCV, Numpy, Matplotlib, Jupyter, LMDB, RDFLib, SciPy, H5Py.

PROFESSIONAL EXPERIENCE

Research Intern at **Microsoft Research (MSR)**, Redmond, WA, USA

May 2019–Aug 2019

- Project: *Face-Speech Bridging by Cycle Video-Audio Reconstruction*
- We designed and developed an algorithm to reconstruct speech from silent video of a talking face and reconstruct the talking face solely from the speech sequence. We solved this problem by designing and training two mutual autoencoders for video and audio and used their bottlenecks as bridges for entering each modality from the other modality.
- The project was defined and performed in Applied Science Group for improving the services of Microsoft. A U.S. patent is filed for this project and will be issued by Nov. 2019.

RESEARCH EXPERIENCE

Digital Video and Multimedia Lab (DVMM), Columbia University, New York, NY, USA

Sep 2019–Present

- Project: *Temporal Reasoning on Video for Probabilistic Event Schema Extraction*

Apr 2018–Feb 2019

- Project: *Multi-level Multimodal Common Semantic Space for Image-Phrase Grounding*
- Accepted in **CVPR** 2019, [PDF](#)
- Github: <https://github.com/hassanhub/MultiGrounding>

Nov 2018–Present

- Project: *Transferable Neural Architecture Search Using Reinforcement Learning*
- To be submitted to **NeurIPS** 2020

Jan 2018–May 2019

- Project: *A Multi-media Multi-lingual Knowledge Extraction and Hypothesis Generation System*
- Accepted for **TAC** 2018, [PDF](#)

Neural Acoustic Processing Lab (NAPLab), Columbia University, New York, NY, USA

Apr 2017–May 2017

- Project: *Lip2AudSpec: Speech reconstruction from silent lip movements video*
- Best project **Idol Award** winner
- Accepted for **ICASSP** 2018, [PDF](#)
- Github: <https://github.com/hassanhub/LipReading>

Sep 2016–May 2018

- Project: *Intelligible speech reconstruction from human auditory cortex using deep neural network models*
- Accepted for **Nature Scientific Reports** 2019, [PDF](#)
- Award: The **NYC Media Lab's Grand Prize** Winner
- News Coverage: 100+ news and science outlets ([Link](#))
- Notable Mention: *NIH Director's Blog* ([Link](#))
- Featured in *The Top-5 AI Stories* ([Link](#))
- Github: <https://github.com/hassanhub/DNNBrainSpeechRecon>

Laboratory Grenoble Image Parole Signal Automatique (GIPSA-lab), Grenoble, France

Aug 2015

- Project: *Fetal ECG extraction using π Tucker decomposition*
- Accepted for **IWSSIP** 2015, [Link](#)

Biomedical Signal and Image Processing Laboratory (BiSIPL), Sharif University of Technology

May 2014 – Feb 2015

- Supervised and unsupervised dimensionality reduction using Tucker decomposition
- Segmentation of biomedical images with intensity inhomogeneity/bias field correction

PUBLICATIONS

Recent

"Multi-level Multimodal Common Semantic Space for Image-Phrase Grounding"

Hassan Akbari, Svebor Karaman, Surabhi Bhargava, Brian Chen, Carl Vondrick, Shih-Fu Chang
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2019.

"Lip2AudSpec: Speech reconstruction from silent lip movements video"

Hassan Akbari, Himani Arora, Liangliang Cao, Nima Mesgarani

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2018

"Reconstructing intelligible speech from the human auditory cortex"

Hassan Akbari, Bahar Khalighinejad, Jose Herrero, Ashesh D. Mehta, Nima Mesgarani

Nature Scientific reports 9, no. 1 (2019): 874..

"GAIA - A Multi-media Multi-lingual Knowledge Extraction and Hypothesis Generation System"

Zhang, Tongtao, Ananya Subburathinam, ..., Heng Ji, Alireza Zareian, **Hassan Akbari**, et al.

In Proceedings of TAC KBP 2018, the 25th International Conference on Computational Linguistics: Technical Papers. 2018..

Old

"A robust FCM algorithm for image segmentation based on spatial information and total variation"
Hassan Akbari, Hamed Mohebbi Kalkhoran and Emad Fatemizadeh
9th Iranian Conference on Machine Vision and Image Processing, **IEEE**, **Best paper award**.

"FCM-based image segmentation with bias field and intensity inhomogeneity correction"
Hassan Akbari, Hamed Mohebbi Kalkhoran and Emad Fatemizadeh
The 15th **IEEE** International Symposium on Signal Processing and Information Technology (ISSPIT) 2015.

"Fetal ECG extraction using π Tucker decomposition"
Hassan Akbari, Mohammad B. Shamsollahi and Ronald Phlypo
International Conference on Systems, Signals and Image Processing (IWSSIP) 2015, **IEEE**

ACADEMIC HONORS & AWARDS

- Ranked 1st among all PhD students of Electrical Engineering
- Ranked 1st among all undergraduate students of Electrical Engineering
- Ranked 3rd among all graduate students of Biomedical Engineering
- In top 1% of all participants of Iranian nation-wide university entrance exam
- Achieved Exceptional Talent Award in Electrical Engineering

TEACHING ASSISTANTSHIP

- Advanced Computer Vision Spring 2019
- Speech and Audio Processing and Recognition Spring 2017
- Brain Computer Interfaces Lab Fall 2016
- Digital Signal Processing Spring 2015
- Statistical Learning Fall 2015
- Electronics I, II Fall, Spring 2012-2013

PROFESSIONAL SERVICES

- Reviewer**
- ACM Transactions on Multimedia Computing, Communications, and Applications
 - IEEE Transactions on Audio, Speech, and Language Processing
 - Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
 - International Conference on Learning Representations (ICLR)

PATENTS

Face-Speech Bridging by Cycle Video-Audio Reconstruction 2019
Patent No. USD407380S1

AWARDS

New York City Media Lab **Grand Prize** Winner (\$10k) 2019
Decoding the brain activities to perceivable speech

HOBBIES

Sports, Music, Art, Photography, Poetry

REFERENCES

Dr. Shih-Fu Chang, Richard Dicker Professor, Columbia University
sc250@columbia.edu

Dr. Nima Mesgarani, Associate Professor, Columbia University
nima@ee.columbia.edu

Dr. Liangliang Cao, Research Staff, Google
liangliang.cao@gmail.com