



## Abhishek Jha

- 10 May 1993
- +91 8969562979
- abskjha.github.io
- github : //abskjha
- google scholar : Abhishek Jha
- abhishek.jha@esat.kuleuven.be

### Courses

- > Digital Image Processing
- > Statistical Methods in Artificial Intelligence
- > Computer Vision
- > Optimization Methods
- > Topics in Machine Learning

### Skills

**Languages:** Python, Shell Scripting, Php/MySQL, MATLAB, HTML/CSS, JavaScript.

**Libraries:** PyTorch, Keras, CUDA,  $\text{\LaTeX}$ , OpenCV, Matconvnet.

### Positions Held

#### Volunteer

17th R&D showcase 2018  
IIIT Hyderabad

#### Reviewer

- > 2nd IAPR International Conference on Computer Vision & Image Processing (CVIP-2017, September 10-12, 2017), IIT Roorkee.
- > 7th NCVPRIPG-2019, December 22-24, 2017, Hubli, India.

## Education

(Pre)Doctoral Student at ESAT  
KU Leuven, Belgium

2019 - Present

M.S. (Research) in Computer Science  
IIIT Hyderabad, India

2016 - 2019  
CGPA: 8.67/10

B.Tech in Electronics and Communication Engineering  
Indian Institute of Technology (IIT) Dhanbad, India

2010 - 2014  
CGPA: 7.57/10

## Publications

[1] **Abhishek Jha**, Vikram Voleti, Vinay Namboodiri and C.V. Jawahar, "Cross-Language Speech Dependent Lip-Synchronization", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* 2019.

[2] **Abhishek Jha**, Vinay Namboodiri and C.V. Jawahar, "Spotting Words in Silent Speech Videos : A Retrieval based approach", *Journal of Machine Vision Application (MVA)*, Springer, January 2019.

[3] Prajwal K. R., R. Mukhopadhyay, J. Philip, **Abhishek Jha**, Vinay Namboodiri and C.V. Jawahar, "Towards Automatic Face-to-Face Translation", *ACM Multimedia (MM)* 2019. (Oral)

[4] **Abhishek Jha**, Vinay Namboodiri and C.V. Jawahar, "Word Spotting in Silent Lip Videos", *IEEE Winter Conference on Applications of Computer Vision (WACV)* 2018.

[5] **Abhishek Jha**, Vikram Voleti, Vinay Namboodiri and C.V. Jawahar, "Lip-Synchronization for Dubbed Instructional Videos", *Fine-grained Instructional Video Understanding, CVPR Workshop* 2018.

[6] Sahil Chelaramani\* **Abhishek Jha\***, and Anoop Namboodiri, "Cross-modal style transfer", *25th IEEE International Conference on Image Processing (ICIP)* 2018. (\*equal contribution)

[7] Yashaswi Verma **Abhishek Jha**, and C.V. Jawahar, "Cross-specificity: modelling data semantics for cross-modal matching and retrieval", *International Journal of Multimedia Information Retrieval, Springer*, June 2018.

## Research Projects

[Current Work] **Multimodal Representation Learning**, Dr. Tinne Tuytelaars

- We are working on problems like: How to use different modalities to learn feature better suited for transfer learning? How to interpret these feature?

**Weakly supervised anomaly detection in action videos**, Dr. Partha Pratim Talukdar, Dr. Anirban Cakraborty, Dr. Anand Mishra

- Given an action video, we want to check its alignment with the predefined instructions.
- Exploring the use of knowledge graph as a weak supervision for video understanding task.

**Visual Dubbing: In the wild lip-motion synthesis**, Dr. Vinay Namboodiri, Prof. C.V. Jawahar

- Developed a lip synchronization model for dubbed instructional videos using Time-delayed LSTM and U-Net (encode-decoder) architecture.
- This facilitates lip-motion synthesis for English dialogue videos to be in lip-sync with dubbed foreign language, like Hindi, to reduce mental fatigue of user.
- Collected a dataset of Hindi lip-videos consisting of Actors speaking dialogue and Andrew Ng DeepLearning.ai tutorial videos to train our proposed architecture.
- Publications: FIVER, CVPR workshop 2018. [4] and ICASSP 2019 [1].

---

### Student Server Administrator

Ada: High Performance Computing(HPC) Clusters (2017-Present)

IIIT Hyderabad

---

### Server Administrator

Cosmos: High Performance Computing (HPC) Clusters (2016-Present)

CVIT,IIIT Hyderabad

---

### Organizing Team

Summer School on Computer Vision 2016

IIIT Hyderabad

---

## Scholastic Achievements

Presented a talk on "Neural Turing Machine and Attention model" at KCIS, IIIT Hyderabad, March 2017.

Recipient of Merit cum mean scholarship (Govt. of India) for the year 2011-14 from IIT Dhanbad.

Pre-placement Job offer by Tata Steel Ltd, based on performance during Internship 2013.

Secretary (2012-13), IEEE-Student branch, IIT Dhanbad.

Qualified IIT-JEE 2010 by IIT at 98.91 percentile, with All India Rank of 5090 (out of around 470,000 students)

---

## Hobbies

Web Designing

Writing Poetry

Biking

Traveling

---

## References

**Prof. Tinne Tuytelaars**

Visics, KU Leuven

**Prof. C.V. Jawahar**

IIIT-Hyderabad

**Dr. Vinay Namboodiri**

IIT Kanpur

**Lip-reading: Audio-visual speech recognition**, Dr. Vinay Namboodiri, Prof. C.V. Jawahar

- Analyzed different CNN, RNNs, and attention based architectures for spatio-temporal modeling of lip-videos.
- Collected a dataset containing 100K annotated video-clips of spoken English word from Obama's public Speech.
- Developed a word-spotting method for recognition-free retrieval for lip-videos giving more than, 30% improvement over SOTA on largest dataset in the domain: LRW dataset.
- Publications: WACV 2018 [3] and MVA journal [2].

**Cross-modal style transfer**, Dr. Anoop Namboodiri

- Developed a method to photo-realistically stylized images with text as style signal.
- Given a style-text we first searched relevant images in word2vec space in YFCC100M dataset. They are then ranked according to their similarity with the content image.
- Formulated a style vector in VGG-16 feature space using weighted combination of retrieved images. Research based on the work has been published in ICIP 2018. [5]

**Unsupervised Sketch simplification for SBIR**, Dr. Vineeth Gandhi

- Developed an unsupervised method of sketch simplification: Formulated a loss function based on Fourier descriptors for human induced artifact reduction in sketch.
- Trained an encoder-decoder network with above loss function on TU-BERLIN dataset and used the bottle-neck representation to show sketch based image retrieval (SBIR) on Caltech-256 dataset.

**Cross-modal multimedia retrieval**, Dr. Yashaswi Verma, Prof. C.V. Jawahar

- We proposed a metric learning method "Cross-Specificity" to capture similarity between text and images, which agrees with the human notion of cross-modal similarity.
- We showed better retrieval performance over SOTA for both text2image and image2text retrieval. Research based on the work published in IJMIR, Springer [6].

---

## Work Experience

**Research Assistant**, VISICS, PSI, KU Leuven

Nov 2019-Present

- PhD student working on Multimodal representation learning.

**Project Assistant**, IISc Bangalore

Feb 2019-Oct 2019

- Visiting student collaborating with Dr. Partha Pratim Talukdar and Dr. Anirban Charabarty on a project on anomaly detection in videos using Knowledge Graphs (KG).

**Research Assistant**, IIIT Hyderabad

July 2016-Jan 2019

- Master's Student working on Audio-visual speech recognition and Synthesis with Prof. C. V. Jawahar and Dr. Vinay Namboodiri.

**Research Fellow**, CVIT, IIIT Hyderabad

July 2015-June 2016

- Developed an cross-modal multimedia retrieval system, by leveraging the semantics of the joint embedding between visual and textual modality.
- Research paper based on the work has been published in International Journal of Multimedia Information Retrieval, Springer [6].

**Graduate Trainee & Manager, Project Planning**, Tata Steel Ltd.

July 2014-Aug 2015

- Improvement project on plant automation system for coal beneficiation at Bhelatand Coal Preparation Plant. Subsequently Promoted as 'Manager Project planning' in July 2015.

---

## Teaching Experience

**Teaching Assistant**, Topics in Machine Learning(CSE975)

June 2018 - Dec 2018

**Teaching Assistant**, Info. Sys. and Signal Proc. (B-KUL-H09M0A)

Feb 2020 - June 2020

**Mentor**, Foundations of AI and ML, Industry-oriented course

Jan 2018-May 2018