

# Nabarun Goswami

1125, 5-5-6 Toshima, Kita-ku, Tokyo, Japan, 114-0003

☎ +81-80-9548-3904 • ✉ [nabarungoswami@mi.t.u-tokyo.ac.jp](mailto:nabarungoswami@mi.t.u-tokyo.ac.jp)

## Education

---

### PhD Student at Machine Intelligence Lab

*The University of Tokyo*

**Tokyo, Japan**

10/2021 - present

### Master of Information Science and Technology

*The University of Tokyo*

**Tokyo, Japan**

9/2019 - 9/2021

GPA: 3.0/3.0

### Bachelor of Technology in Electronics and Communication Engg.

*Tezpur University*

**Assam, India**

2008 - 2012

CGPA: 8.18/10

## Skills

---

- Programming Languages: Python (current), Java (last used: 2018), C++ (last used: 2015)
- Deep Learning Frameworks: Pytorch (current), Tensorflow (last used: 2018)
- Languages: Assamese (mother tongue), English, Hindi, Japanese (Beginner)

## Publications

---

1. N. Goswami and T. Harada, "[SATTS: Speaker Attractor Text to Speech, Learning to Speak by Learning to Separate](#)," in *Proc. Interspeech 2022*, pp. 1203–1207, 2022
2. N. Goswami, "Unsupervised Style Modeling for Text to Speech Synthesis," Master's Thesis, University of Tokyo, Tokyo, Japan, 2021
3. N. Takahashi, S. Parthasaarathy, N. Goswami, and Y. Mitsufuji, "[Recursive Speech Separation for Unknown Number of Speakers](#)," in *Proc. Interspeech 2019*, pp. 1348–1352, 2019
4. N. Takahashi, P. Agrawal, N. Goswami, and Y. Mitsufuji, "[PhaseNet: Discretized Phase Modeling with Deep Neural Networks for Audio Source Separation](#)," in *Proc. Interspeech 2018*, pp. 2713–2717, 2018
5. N. Takahashi, N. Goswami, and Y. Mitsufuji, "[Mmdenselstm: An Efficient Combination of Convolutional and Recurrent Neural Networks for Audio Source Separation](#)," in *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, pp. 106–110, 2018
6. S. Sharma, R. Padhy, S. Kumar Choudhury, N. Goswami, and P. Sa, "[DenseNet with pre-activated deconvolution for estimating depth map from single image](#)," in *AMMDS 2017, Workshop on Activity Monitoring by Multiple Distributed Sensing*, September 2017
7. N. Goswami, T. Kalita, and R. Devi, "Video noise reduction based on statistical modelling of wavelet coefficients," Bachelor's Thesis, Tezpur University, Assam, India, May 2012

## Patents

---

1. N. Goswami, "[Method and system to generate one or more multi-dimensional videos](#)," Aug. 3 2021. Applied: Aug 18 2016. US Patent 11,082,754

2. N. Goswami and M. Sulibhavi, "[System and method for sharing multimedia content with synched playback controls](#)," Sept. 15 2020. Applied: Oct 10 2016. US Patent 10,778,742
3. P. Chintalapoodi, N. Goswami, H. Sadhwani, and M. Sulibhavi, "[System and method for processing video content based on emotional state detection](#)," Jan. 7 2020. Applied: Sep 9 2016. US Patent 10,529,379
4. N. Goswami, M. Sulibhavi, and P. Chintalapoodi, "[Device and method for generating a panoramic image](#)," May 21 2019. Applied: Nov 20 2015. US Patent 10,298,841

## Professional Experience

---

### **NABLAS Inc.**

#### ***Junior Research Engineer (Part Time)***

**Tokyo, Japan**

*October 2021 - Present*

- Research on Deep Fake generation technologies in speech and images

### **Tokyo Coding Club**

#### ***Programming and Robotics Tutor (Part Time)***

**Tokyo, Japan**

*February 2020 - October 2021*

- Teach middle and high school students basic robotics and programming

### **Sony**

#### ***Technical Specialist***

**Bangalore, India; Tokyo, Japan**

*November 2016 - February 2019*

- Research on Music Source Separation, Speech Enhancement and Speech Separation using Deep Learning in collaboration with Audio Technology Department, Sony Corporation, Tokyo.
- Participated and achieved the best scores in the Music (MUS) task of the sixth community-based Signal Separation Evaluation Campaign (SiSEC 2018).
- Developed custom multi-processing based data-loading mechanism for efficient training pipelines for large audio datasets, reducing training times by over 40%.

### ***Senior Software Engineer***

*August 2012 - October 2016*

- Developed the prototype for action camera highlight generation based on emotion of user detected using Electro-Dermal Activity sensors. Programming Language: Java, Python
- Developed prototype of Finger Tracking based Scroll Speed Control for Remote Control application for Sony Bravia TVs, which would improve user experience of mobile device-like scrolling on TV. Programming Language: Java
- Developed an automated CI platform based on Jenkins for centralized monitoring and reporting. Programming Language: Python
- Developed an in-house No Reference video Quality Measurement tool which cut the costs by almost €15000 annually (subscription cost of 3rd party tool). Programming Language: C++
- Developed a prototype for video comparison without access to original video, which reduced need for manual testing of streaming content. Programming Language: C++
- Developed image (TIFF, JPEG, BMP, PNG) metadata parsers for PlayStation 4 SDK which enabled implementation of in-game screenshots. Programming Language: C++

## Awards and Achievements

---

- Recipient of Japanese Government (MEXT) Scholarship for Graduate Studies at The University of Tokyo. *2019 - 2021*
- Winner of Music (MUS) task of the Signal Separation Evaluation Campaign (SiSEC), 2018.