Statement of Purpose

I want to pursue a Ph.D. in Electrical and Computing Engineering and my career goal is to become a professor. My research interests include artificial intelligence, machine learning, music information retrieval, computer vision and natural language processing.

My motivation for pursuing a Ph.D. comes from my passion in doing research. Rather than school work, I always find it more enjoyable to do research on topics of my interest. I began my first research project early when I was in junior high school, and have completed four research projects (three in Mathematics and one in Physics) during high school, where two of them are co-advised by my high school teachers and a professor/researcher. In college, I worked with Prof. Hung Yu-Wei to build a game theoretic model for resource pricing in wireless mobile networks. After college, I have being working as a research assistant and doing machine learning research with Dr. Yi-Hsuan Yang for the past one and a half year. Through these experiences, I found myself most engaged when tackling unsolved problems that might improve the state of the art and have the potential to impact a great number of people. Hence, I am determined to pursue a career in academia, and I believe doing a Ph.D. is the best next step for my research career.

My interest in machine learning research first developed when I worked with Dr. Yang on automatic music generation. In my first project, I collaborated with my colleague, and we started by implementing a state-of-the-art deep learning model that can generate musical melodies. Being a challenge lover, I later proposed to turn the project to a more general yet less studied direction—generating polyphonic music of multiple instruments. This is challenging for it requires modeling both the multitrack interdependency and the temporal structure at the same time, and thus success in natural language generation and melody generation may not be readily applicable here.

As a first step toward this goal, we developed a novel multi-track sequential generative adversarial network (MuseGAN), which is the first model that can generate polyphonic multi-track music from scratch. We also extended the model to support automatic music accompaniment. Moreover, we designed several objective metrics and conducted a user study to evaluate the generated results. Although it is still preliminary in musical and aesthetic aspects, we show that the proposed model can start to learn some musical elements such as keys, scales, chords, bass lines and drum patterns. I published MuseGAN in a first-authored paper at AAAI 2018 [1]. In addition, we also published a new pianoroll dataset that we derived and used as the training data [2].

I later worked with Dr. Yang to further address the binarization issue we encountered in MuseGAN. I proposed to introduce an additional refiner network that learns to binarize the generated outputs. The experimental results show that the proposed approach outperforms common post-processing methods in several objective measures. I published this improved model in a first-authored paper at ISMIR 2018 [3]. Moreover, I also published Pypianoroll, an open source Python package I developed for working with music pianoroll data, in a demo paper at ISMIR 2018 [4].

In my most recent project, I turned to investigate a more general-purpose setting to model high-dimensional binary-valued data with generative adversarial networks (GANs). I proposed to adopt binary neurons at the output layer of the generator and train the network by backpropagation using gradient estimators for the binary neurons. The proposed BinaryGAN model can achieve competitive performance on generating binarized MNIST handwritten digits as compared to existing approaches. I wrote a first-authored paper for BinaryGAN and it is currently available on arXiv,

while we have not decided which venue to submit it to [5]. I am now working on adopting binary neurons as gate units to support conditional computation graphs for more advanced systems.

I have also been working on integrating the above models into a full music generation system that can 1) learn a latent space with a two-way mapping to and from the data space for music pianoroll data by a combination of generative adversarial networks and adversarial autoencoders, 2) generate multitrack music with temporal dependencies by training a recurrent neural network on top of the learned latent space and 3) dynamically activate/deactivate certain instruments in the middle of the generated music with the use of binary neurons.

In another ongoing project, I have been investigating applying the discriminative adversarial network (DAN) to multi-class and multi-label classification. As a first step, I implemented the model to classify MNIST handwritten digits. In this way, it provides us a systematic method to evaluate different GAN objectives. I am now working on applying it to multi-label classification. It deserves further investigations for it no longer requires a surrogate loss to be specified.

I have also been participating in an international collaborative project on melody harmonization benchmarking since September. We collaborate with Dr. Satoru Fukayama and Prof. Tetsuro Kitahara to compile a benchmark dataset and set up a standard procedure for model evaluation.

While I focused on music information retrieval in my recent research, I have a broad interest in general-purpose machine learning, computer vision and natural language processing. I have always been following the progress in these fields and reading the latest papers at top venues and arXiv. During my research, I have also learned and drawn inspirations from literature in these fields. Hence, I would be happy and willing to join in related projects for my graduate studies.

UC San Diego is especially attractive to me for the outstanding research projects carried out by its faculty members. There are several professors whose research are particularly appealing to me: Prof. Gert Lanckriet, Prof. Nuno Vasconcelos, Prof. Truong Nguyen and Prof. Ravi Ramamoorthi and Prof. Mohan Trivedi. In addition, your curriculum covers a wide range of research areas that I would like to explore. It would therefore be a privilege for me to join UC San Diego for my graduate studies, where I can learn from the exceptional faculty and contribute my skills and experience.

I am prepared and fully equipped for graduate studies in Electrical and Computer Engineering with a specialization in Machine Learning and Data Science. With my extensive research experience, I am confident that I am a good fit for your program, and that UC San Diego is certainly the best place for me to pursue a Ph.D. and continue my research career.

[1] Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang. MuseGAN: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, February 2018.

[2] Lakh pianoroll dataset. [link] https://salu133445.github.io/lakh-pianoroll-dataset/.

[3] Hao-Wen Dong and Yi-Hsuan Yang. Convolutional generative adversarial networks with binary neurons for polyphonic music generation. In *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, September 2018.

[4] Hao-Wen Dong, Wen-Yi Hsiao, and Yi-Hsuan Yang. Pypianoroll: Open source python package for handling multitrack pianoroll. In *Late-Breaking Demos of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, October 2018.

[5] Hao-Wen Dong and Yi-Hsuan Yang. Training generative adversarial networks with binary neurons by end-to-end backpropagation. *arXiv preprint arXiv:1810.04714*, October 2018.