ComposeGAN: A Conditional GAN Approach to Symbolic Music Style Transfer

Lu, Conan (School: Redmond High School)

This project introduces ComposeGAN as a generative adversarial network that learns and transfers the style elements of tonal music genres. Early attempts at symbolic music style transfer have faced the challenge of preserving specific content features, such as chords, notes, and musical surface, while simultaneously changing style, including aspects of music like rhythm and ornamentation. ComposeGAN presents a novel method to differentiate between style and content by embedding chroma feature extraction within the training process. A chroma is a visual representation of music that only records note density, explicitly separating content and style. ComposeGAN uses conditional GAN architecture and loss functions, paralleling Pix2Pix image-to-image translation algorithms. Pop, jazz, and classical datasets were used for training purposes. To evaluate the success of the transfer, I used two metrics: Tonnetz distance, to measure harmonic similarity, and a separate genre classifier, to measure transfer realism. The success of the transfer is evidenced by the high independent genre classifier accuracy rate and near-zero Tonnetz distance, demonstrating a convincing style change and a conservation of content, respectively. Using chroma feature extraction offers significant advantages, as the format improves transfer consistency while requiring less data to train. ComposeGAN can be utilized as a tool for musicians to study compositional techniques. In addition, experiments using lead sheets in conjunction with ComposeGAN present the possibility of automatic music generation.

Awards Won:
Association for the Advancement of Artificial Intelligence: Honorable Mention