PhonoNet: Deep Learning for Raga Identification in Indian Classical Music

Chowdhuri, Sauhaarda (School: Westview High School)

Indian classical music is an improvisational form of music based on ragas, melodic frameworks which are passed down through a fading oral tradition. I aimed to provide computational prediction of ragas so singers can receive live feedback when learning and important features of the music can be preserved digitally in my PhonoNet system. First, my system computed the short-term Fourier Transform of the input audio data to form a chromagram representation of the notes being sung. This data was then augmented using a novel transpositional data augmentation algorithm and split into 150 second chunks used as training inputs for a deep convolutional neural network. The convolutional network's filters were analyzed using a novel saliency visualization algorithm and then modified with a recurrent layer to allow processing of full-length songs. The convolutional system achieved 78.9% accuracy for raga prediction on 150 second chunks. The visualization technique isolated characteristic raga features, identifying 72.8% more features in real audio than in random noise. The joint raga prediction system achieved a new state-of-the-art 98.9% accuracy for raga prediction on full-length songs. The PhonoNet system documents the structure of Indian music with deep networks and provides live feedback mechanisms for learning the art form. Future work can extend the proposed hierarchical system to other tasks with long temporal sequences and extend the data augmentation to different applications of music processing, in addition to extending the PhonoNet system to other forms of world music.

Awards Won:
Third Award of $1,000
Acoustical Society of America: Third Award of $600.00, plus students Mentor will be awarded $150.
GoDaddy: $1,500 Making the Best Use of Data Award