A CASE STUDY OF A SYNESTHETIC APPLYING NEURAL NETWORKS

Raminta Bartuliene[1], Ruta Narvilaite[1], Gustavas Davidavicius[1], Ausra Saudargiene[1,2], Sarunas Asmantas[1], Saulius Satkauskas[1]

[1] Department of Informatics, Vytautas Magnus University, Kaunas, Lithuania
[2] Neuroscience Institute, Lithuanian University of Health Sciences, Kaunas, Lithuania
raminta.bartuliene@vdu.lt

Synesthesia is a perceptual phenomenon in which one type of sensory stimulus automatically and involuntarily evokes another type of sensory sensation. There are many forms of synesthesia: sound-to-colour synesthesia, mirror-touch synesthesia, grapheme-colour, lexical-gustatory and many more. This study focuses on a case of chromesthesia. Chromesthesia is a type of synesthesia in which heard sounds evoke colour sensation. The study aims to distinguish evoked colours based on voice features.

The research subject was a female with visual deficiency. She claims that the grey silhouette of a person develops a colour after communicating with them. 39 participants (19 male and 20 female) attended in the study. Participants' voices were recorded with two-channel audio recording equipment and the colour which was seen by the research subject was registered. For audio analysis, our team extracted voice features using pyAudioAnalysis Python module. We extracted 68 auditory signal features: averages and standard deviations of Zero crossing rate, Energy, Entropy of Energy, Spectral Centroid, Spectral Spread, Spectral Entropy, Spectral Flux, Spectral Rolloff, 13 Mel Frequency Cepstral Coefficients, 12 Chroma Vectors, Chroma Deviation in 20 windows of 25ms length. For colour classification, we designed and trained a Multilayer Perceptron (MLP) neural network using Keras Python module. The MLP was constructed with 1 input layer (68 input neurons, activation - ReLU), 3 hidden layers (with 50-30-20 neurons in respective layers, activation - ReLU) and 1 output layer (1 output neuron, activation - sigmoid). To train the error backpropagation method with binary cross-entropy loss function was used. The MLP neural network was trained with data from 2 colour classes: pink and white females (6 pink and 5 white subjects). We dedicated 15% of the data for model testing. The rest of the data was split into training (80%) and validation (20%) sets.

The dominant registered colours of study participants were white (5) and pink (6) for females and blue (7) and black (3). There was also 12 multi-colour samples and the rest were 1-2 samples per colour.

White and pink females were classified with the accuracy of 93% by the neural network. F1 scores were 0.90 for the class of white-colour subjects and 0.95 for the class of pink-colour subjects.

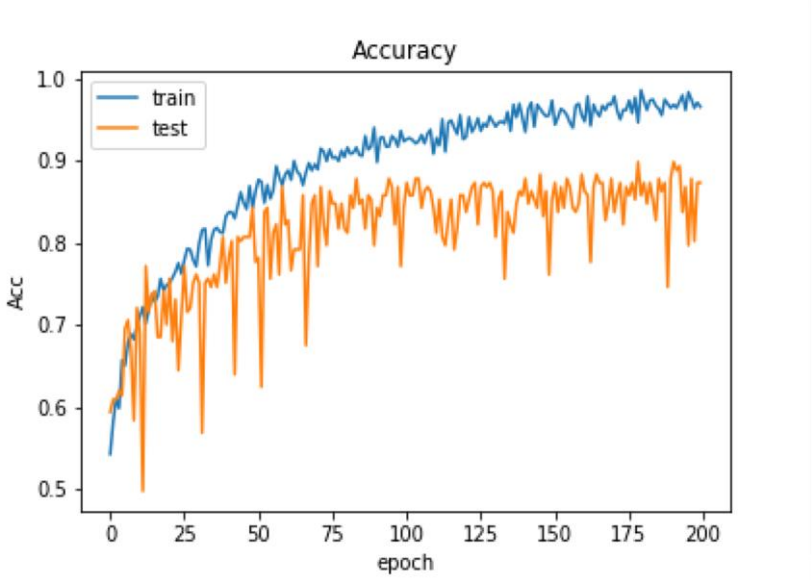

Fig. 1. The classification accuracy of the MLP model over 200 epochs.

After the study, we can claim that in this case, the evoked colour depends on the features extracted from recorded voice clips. Different classes were classified successfully using the MLP neural network. For better results, we need to do feature selection to remove irrelevant features. Our next goal is to design and train a neural network, that successfully classifies all participant colours and to determine which extracted features are key in evoking different colour sensation. For that purpose, collection of more voice data may be required.