Using Deep Learning for Alcohol Consumption Recognition

Joseph Bernstein¹, Brandon Mendez², Peng Sun³, Yang Liu³, Yi Shang³
UMBC¹, University at Buffalo², University of Missouri³

Introduction

- With machine learning, computers have proven effective at recognizing objects in images (99.7% accuracy)².
- Translating audio to words and sentences.
- Recognizing musical chords.
- Research Questions:
  - Can alcohol be detected based off body signals such as Heart Rate, Breath Rate, Skin Temperature.
  - Can machine learning be used to solve the task?
- Our Approach:
  - See Figure 1.
- Impact:
  - Demonstrates how machine learning can be used to detect alcohol consumption.

Results

- Training Set:
  - 80 images, parameters found in Figure 5.
  - 40 positive (drink).
  - 40 negative (no-drink).
- Testing Set:
  - 20 images.
  - 10 positive (drink).
  - 10 negative (no-drink).
- Performance Metrics:
  - Training Loss, Figure 6a.
  - Descriptor of how effective the model is at learning features of the spectrograms. Lower loss is better.
  - Our training loss tended to 0, indicative of learning many features.
- Training Accuracy, Figure 6b.
  - Descriptor of the prediction rate of the model based on the training dataset.
  - Our training accuracy tends to 100%, implying that the model is over-fit to the data.
- Testing Accuracy, Figure 7.
  - Descriptor of the prediction rate for data the model has never seen. Test accuracy after the model is finished training.
  - Overall accuracy of 75% (15/20 images).
  - Features vectors (Figure 8).
  - Features that are picked up from spectrogram set.
  - Useful in determining where the model believes the significant information is.

Method and Technology

- Sensor Data Source:
  - Two types of data, comes from ADA², survey and sensor data.
- Sensor Data, measures:
  - Heart Rate.
  - Breath Rate.
  - Skin Temperature.
- Activity.
- Survey Data, measures:
  - Points where alcohol is consumed.
  - Number of drinks.
- Spectrogram:
  - Data is converted to a spectrogram (Figure 2).
  - Spectrogram generation is configurable, allowing for all parameters to be adjusted as necessary.
  - Algorithm to generate spectrograms (Figure 3).
- Convolutional Neural Network (CNN):
  - Pretrained and finetuned pre-existing CNN, AlexNet (Figure 4).
  - Located on AWS EC2 Server w/ GPU access.
  - Pretrained on ILSVRC 2012 data.

Figure 2. The spectrogram, and associated waveform, of a heart rate.

Figure 3. The algorithm used to generate the spectrograms used.

The ‘Drinking’ Spectrogram Algorithm

GET data FROM excelSheet
PUT data INTO table
DROP EACH row IN table WHERE there is no instance of drinking FOR EACH row IN table of drinking instances
GET all points IN original dataset
REMOVE all in dataset EXCEPT those within 60 minutes of the drink point.
PLOT points.

The ‘No Drinking’ Spectrogram

GET data FROM excelSheet
PUT data INTO table
DROP FROM table WHERE row was used for Drinking Spectrogram
NUMBER of ‘No Drink’ Spectrograms = NUMBER of ‘Drink’ Spectrograms
GET random data/time, select random 60 minute time window
IF (random data/time NOT missing data)
PLOT points.

Figure 4. AlexNet is a pretrained CNN. “Pretrained” means it has already been fed hundreds of thousands of data points.

As such, it takes less time to train, as well as fewer example images.

AlexNet (Krizhevsky et al. 2012)

The class with the highest likelihood is the one the CNN selects.

Using Deep Learning for Alcohol Consumption Recognition.

When AlexNet is processing an image, this is what is happening at each layer.

Conclusions and Looking Forward

- Successfully created deep learning pipeline to take in sensor information and predict whether or not alcohol was consumed.
- Created configurable system to allow various parameters to be changed, so that different input can be tested to find optimal results.
- Achieved prediction accuracy of 75% which outperforms random guessing.
- Looking forward, create more spectrogram options such as:
  - Change scaling from linear to log based.
  - Zoom in on certain frequency bands.
  - Search for which frequency bands have the best results.
- Combine predictions to incorporate different signals at one time.

Acknowledgements

We would like to give thanks to the NSF for funding all of the costs to stay in Mizzou for our research.

References