Wearing the Inside Out: Using Long Short-Term Memory Networks and Wearable Data to Identify Human Emotions

Carlos Aguirre

Maria Fernanda De La Torre
Kansas State University

Wearing the Inside Out: Using Long Short-Term Memory Networks and Wearable Data to Identify Human Emotion

Carlos Aguirre, Maria F. De La Torre and William H. Hsu

Department of Computer Science, College of Engineering

Abstract

Studying emotions may sound unusual in computer science, a field based on quantifiable data and rationality. Contrary to belief, studies have shown any decision is highly dependent on emotional input. To improve human-computer interaction, it is crucial to improve our understanding of human emotions and teach machines to identify them. With large amounts of information streaming available from our environment, identifying our current emotional state becomes challenging, even at the individual self-level. This project aims to identify indicative emotional temporal data from wearable devices. Using brain activity data from an EEG and smart watches that record data, such as heart-beat, physical motions and glucose-levels, we hope to find a correlation that will enable us to train a neural Long Short-Term Memory Network (LSTMN) that classifies the temporal physical-state data into the emotional state of the subject. LSTMNs allow the use of previous long- and short-term data points, expanding our understanding of what our body is telling us about our psyche.

Data

This data set is a perfect test bed for our experiment, since it has the same primary data-channel (heartbeat from smartwatches) and a secondary physiological data channel, in this case an oculographic sensor from JINS MEME smart glasses. Furthermore, it contains data about movement (gyroscope and accelerometer). The data set was collected by using a smartwatch and smart glasses and having subjects run through their routine life activities.

To prepare our data set for LSTM, we filtered the temporal data points that had action labels and features from both data-channels resulting in 5,873 data points. Moreover, the 24 initial classes were grouped into 7 general classes: fitness, entertainment, recreation, eat, sleep, transportation and work. The data points were split into training and testing sets.

Methodology

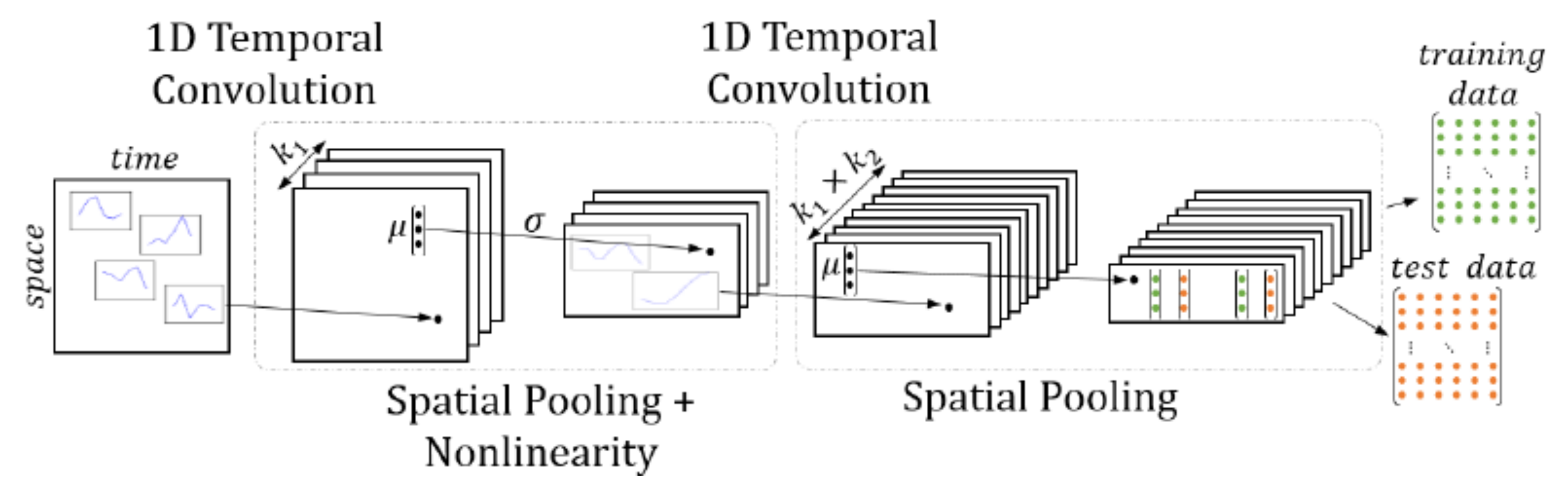


Figure 1. 1D Temporal Convolutional Neural Network which processes a batch of examples by kernel size. Combines linear and special pooling for temporal problems.

- Convolutional Neural Network (CNN) (figure 1) allows the model to find patterns from our spatial data and infer new features.
- The forget gate in a LSTM neuron (figure 2) enables the network to learn the features that define the action by deciding which raw and learned features to remember at each iteration.

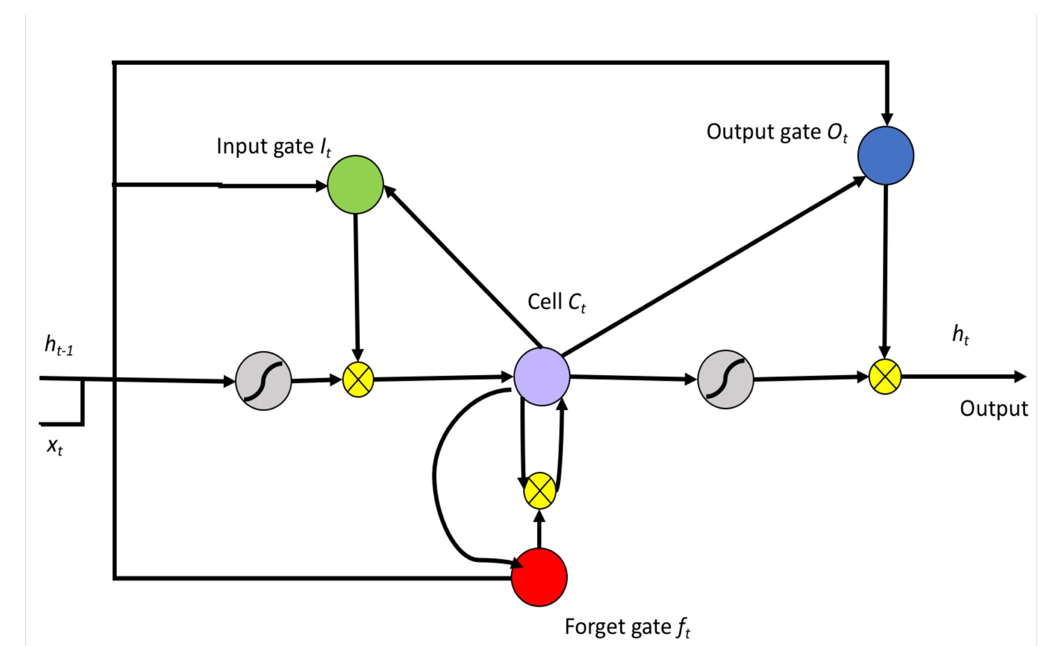


Figure 2. The structure of an LSTM neuron, containing an input gate it, a forget gate f_t, a cell c_t, an output gate O_t and an output response h_t.

Experiment Design

- Sequence classification is challenging due to the variation in the input over space and time. Combining both CNN layers and LSTM layers we can create a model that excels at learning spatial and temporal relationships.
- The CNN layer may be able to pick out invariant features for each action. These learned spatial features may then be learned as sequences by an LSTM layer.
- We apply a Dropout layer in between CNN and LSTM layers to avoid overfitting of the data.
- At the end, a dense, fully connected, layer is used to classify the output of the LSTM layer into the actions.

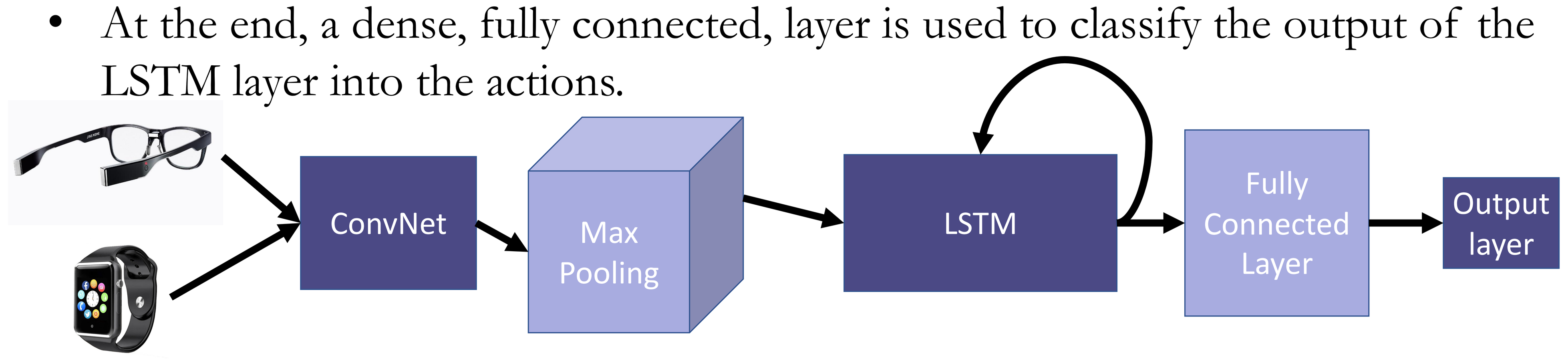
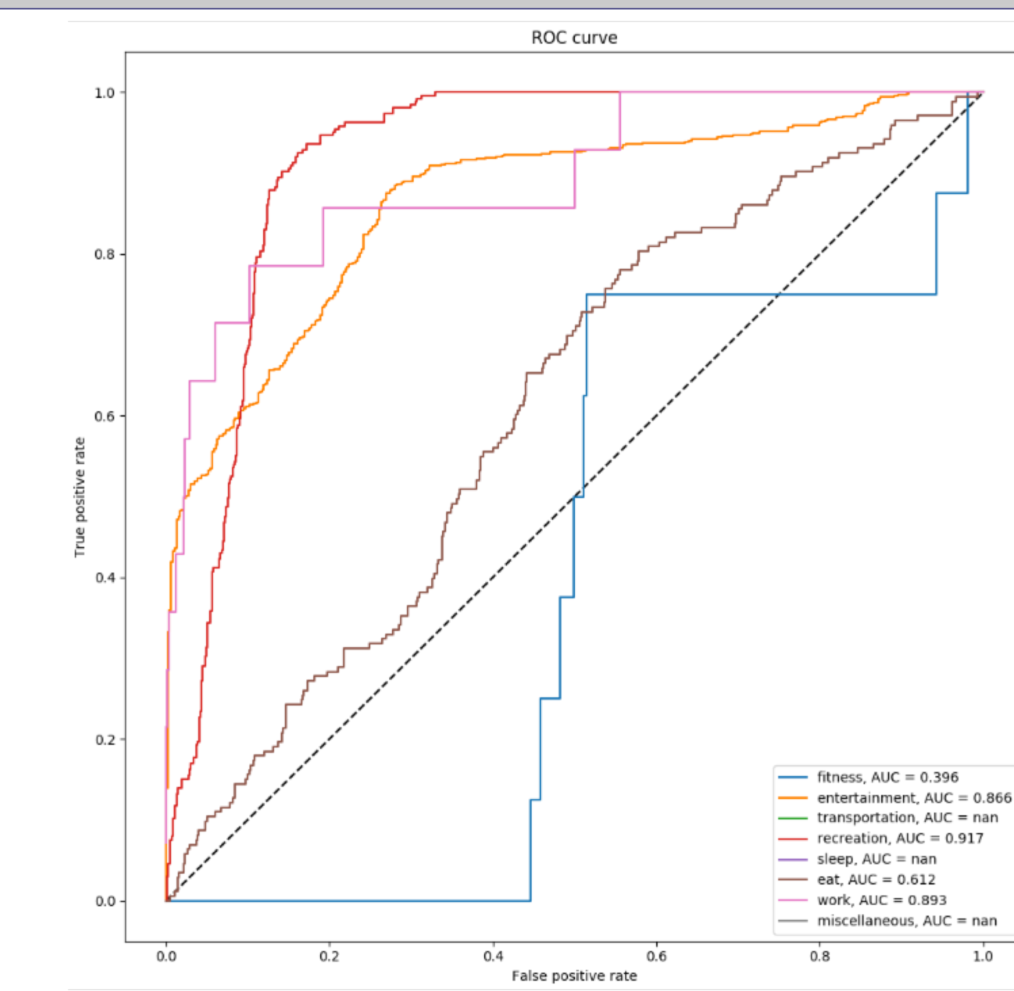


Figure 3. Our model incorporating the 1D Temporal Neural Network, Max pooling, and LSTM layers

Results



We used Area Under the Receiver Operating Characteristic Curve AUC ROC as a measure of performance in our classification problem. Since we have a multiclass problem, we treat it as a binary classification problem that is considered one vs all for each action. Our results indicate the need for further data collection and labeling as gaps in the sequential data result in class imbalance.

Future Work

Our future work involves collecting both smartwatch and EEG data, and using our model to find a correlation between the two data-channels and classify the arousal level of a particular emotional state. Potential inducers of emotion and arousal will include music, videos or photographs. We are also exploring other hierarchical deep learning models such as stacked recurrent neural networks and LSTMNs.

References

Sébastien Faye, Nicolas Louveton, Sasan Jafarnejad, Roman Kryvchenko, Thomas Engel. An Open Dataset for Human Activity Analysis using Smart Devices. 2017.