Identifying Hand Hygiene Using Neural Networks

James H. Edwards
advised by Dr. Valerie Galluzzi, Dr. Matthew Boutell, and Dr. Klaus Baer

A thesis submitted in partial fulfillment of the requirements for the
Bachelor of Science degrees in International Computer Science
at the Rose-Hulman Institute of Technology and Hochschule Ulm.

April 2, 2017

Contents

Abstract

Both machine and deep learning are growing fields of computer science that are rapidly increasing in relevance to our society. One compelling field of application is in healthcare and hospitals. Systems can be designed to help and improve the lives of patients with particular diseases or disabilities, and systems can even be trained to diagnose complicated symptoms or to otherwise aid doctors in their duties. The experiment used in this project was originally conducted by Dr. Valerie Galluzzi, who used custom 3-D wrist accelerometer sensors in order to measure healthcare worker compliance to hospital guidelines. My continuation of the experiment took the data and used neural networks to generate models that can predict when a novel sample is performing hand hygiene.

Chapter 1

Background

This part of the thesis provides an introduction to the concepts of deep learning and neural networks, followed by the Literature Review I conducted to gather information about other recent work in this field.

1.1 Hand Hygiene

–github
–link to thesis pdf, poster, and final slides
–keep code + such on repo

1.2 Deep Learning

Simulate the brain
Convolution
Testing vs training data

1.3 Literature Review

Chapter 2

Experiment

Now that the reader has an introduction to the neural networks and a view on the current research in deep learning in the healthcare field, I will describe the experiment I undertook. It begins with a discussion of the data and some issues I encountered, the models I developed to train, and other deep learning techniques I used to increase the accuracy of my models.

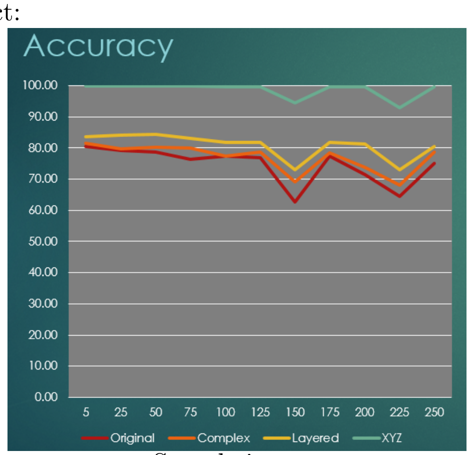
2.1 Initial Steps

My project is a continuation of the work of Dr. Valerie Galluzzi, who did her dissertation on using machine learning techniques to identify if various healthcare workers were being compliant with the guidelines of the World Health Organization. She and her team** worked to develop custom wrist-wearable acceleration sensors, which sent data to a separate device. The data could be offloaded for investigation. I was given most of the data she used, which consisted of X total samples from Y healthcare workers in Z hospital in A year. These acceleration values were taken at 100 Hz, and measured the X, Y, and Z values for each hand for varying lengths of time (the data for each hand was recorded on separate devices but the values were stored together).

Put in the data columns for the IPython notebook CSV part

I initially worked on reorganizing the data. It had been given to me in a JSON file, whereas it is much easier to work with CSV files that can be imported in Pandas, a Python module that can be combined with TensorFlow (Google's Deep Learning API). Once I had a Python script which converted the data to CSVs, I then made a second Python script to load the data from the CSV files and put the various acceleration data points into separate X, Y, and Z matrices. These matrices served as the inputs to the Tensorflow models.

The initial layout of the data is shown below, to the left of the data layout I used for the majority of the project:



Sample image.

DATA SLICING DISCUSSION

Because of the way the data was collected, I did not actually have a continuous stream of data; that is, each sample did not exactly or directly take place immediately before or after another sample. Therefore I had to work with the individual slices provided to me, rather than a "constant" stream of values I could divide up any way I wished.

DATA IMBALANCE DISCUSSION

One important aspect of the data to be discussed is that over 95% of the total length of all data samples were from Non-Hand-Hygiene samples. This imbalance led to various discussions about how it should be resolved, as my initial tests simply ignored the HH samples around 96% accuracy.

FINAL PIPELINE SUMMARY ??

2.2 Models

I began with 4 models, here listed with the names I gave them:

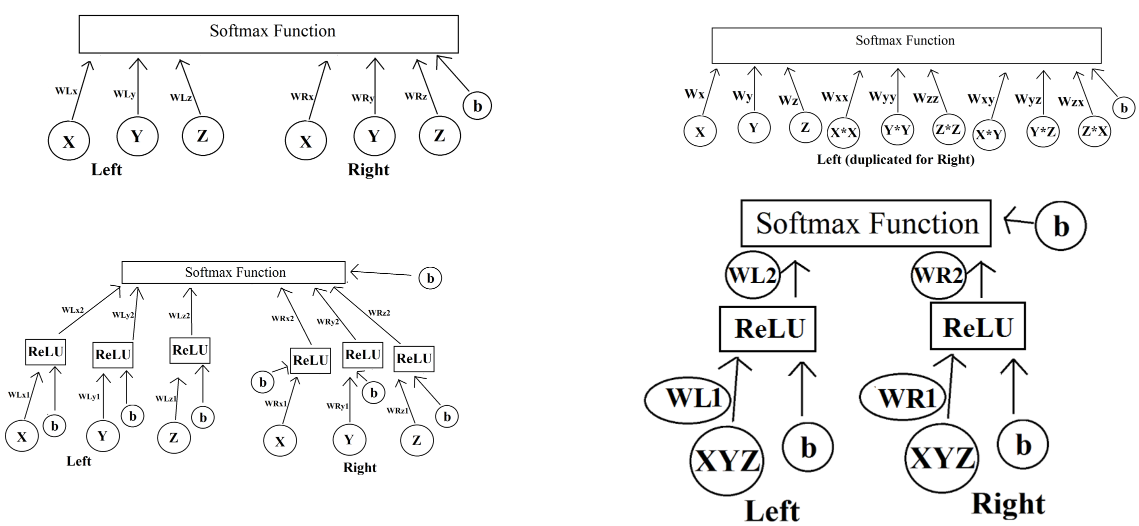
1. Original Model: A simple model with no hidden layer, which took in the X, Y, Z values.

2. Complex Model: A model with no hidden layer but used the $X, Y, Z, X^2, X^2, X^2, X * Y, Y * Z, Z * X$ values as inputs.

3. Layered Model: A model with 1 hidden layer but only the X, Y, Z values as inputs.

4. XYZ Model: A model with a hidden layer but the data was arranged with the concept of the *previous, now,* and *next* instances.

Following a standard convention for neural network models, I will illustrate the models below. In clockwise order beginning with top left, these are the visualizations for the Original model, the Complex model, the XYZ model, and finally the Layered model.



I then added a Convolutional Model, and I also added more layers to the Complex model and the Layered model in an attempt to improve accuracy.

The convolutional model ...

The XYZ model ...

$$\begin{bmatrix} x_1 & y_1 & z_1 & x_2 & \cdots & x_{n+1} & y_{n+1} & z_{n+1} & x_{n+2} & \cdots & x_{2*n+1} & y_{2*n+1} & \cdots & z_{3*n} \\ x_{n+1} & y_{n+1} & z_{n+1} & x_{n+2} & \cdots & x_{2*n+1} & y_{2*n+1} & z_{2*n+1} & x_{2*n+2} & \cdots & x_{3*n+1} & y_{3*n+1} & \cdots & z_{4*n} \\ \vdots & & (previous) & & \vdots & & (now) & & & \vdots & & (next) & \ddots & \end{bmatrix}$$

As mentioned before, I also simply added four hidden layers to the Layered Model and the Complex Model. I will not show the illustration here, but these hidden layers consisted of 512 nodes [1].

2.3 Other Techniques

With the models which had multiple layers, I also implemented L2 regularization. This technique attempts to prevent overfitting by reducing the total value of the weight matrices. Thus the model cannot become overtrained on the training data and score lower on the testing data [2].

Convolutional Model Attempt

XYZ Model Attempt

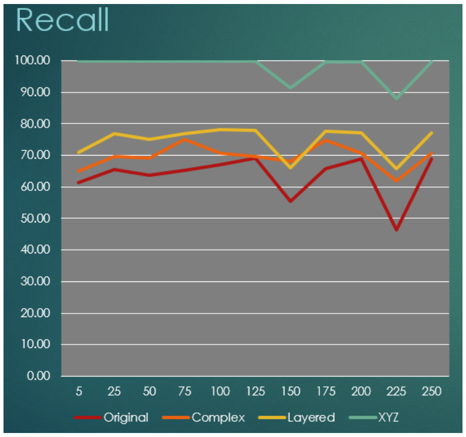
Chapter 3

Results

This part covers the results of the experiments I ran, described in the previous part of this thesis.

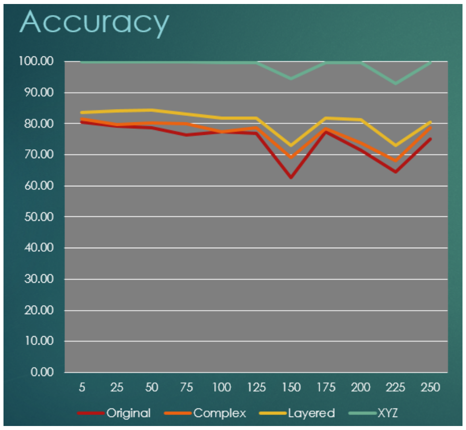
3.1 Recall

For all the tests I ran, I recorded statistics gathered from the confusion matrix. One of the main values one can gather from a confusion matrix is recall, which is $\frac{truepositives}{truepositives+falsenegatives}$ [3].



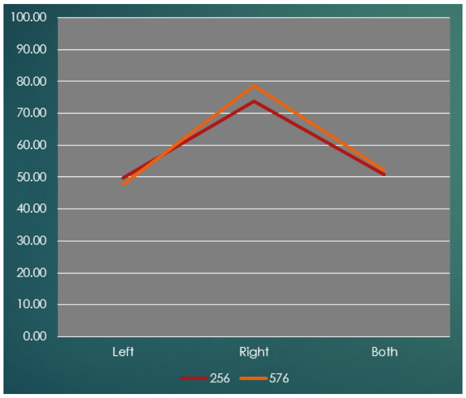
Sample image.

3.2 Accuracy



Sample image.

3.3 Other Views



Chapter 4

Discussion

Now that the reader is familiar with my experiment and the results of it, I will enter into some discussion about the impact of what I have done as well as future work that could be done in this vein.

4.1 Impact

The main motivation for this work was to find a better way to ensure that healthcare workers are adequately washing their hands to prevent the spread of diseases in hospitals. While my best result of 85% is by no means a perfect 100%, I feel that this system would be an important first step to increasing the amount of times a healthcare worker would wash his or her hands.

One important factor is using the wrist-sensor system compared to any other would be the cost. These sensors would not be too expensive to produce, and would definitely be cheaper (and perhaps more accurate) than installing special sensors near every sink or hand sanitizer dispenser and then constructing a system to measure the amount of time a doctor or nurse is within a certain distance of a hand-washing location.

4.2 Future Work

If one were interested in further developing the physical system, it could be interesting to investigate using different sensors, such as velocity, relative location, and/or angular acceleration, and determining if a particular combination is more accurate at measuring hand hygiene.

Of course, one could also try to develop a more complicated neural network model or implement future deep learning techniques in order to improve the recall and accuracy of the system.

Another important area to look into would be utilizing a video input of hand movements, perhaps with a depth camera or just an RGB camera. The video input could also be combined with acceleration data for increased effectiveness [5].

Something else to look into would identifying proper hand hygiene technique, compared to simply detecting "hand hygiene or not" for a particular sample. Measuring the effective of the hand hygiene may need many more sensors, as discussed above. It could also be difficult to express proper technique in a way that would generalize for many subjects [4].

$$\begin{bmatrix} x_1 & x_2 & ... & x_n \\ x_{n+1} & x_{n+2} & ... & x_{2*n} \\ \vdots & & \ddots & \end{bmatrix}$$

Bibliography

[1] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[2] N. Hammerla, S. Halloran, and T. Ploetz. Deep, convolutional, and recurrent models for human activity interaction using wearables. In Subbarao Kambhampati, editor, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 2016.

[3] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(3):436–444, may 2015.

[4] Natalia Neverova, Christian Wolf, Graham W. Taylor, and Florian Nebout. Moddrop: adaptive multi-modal gesture recognition. *CoRR*, abs/1501.00102, 2015.

[5] Michael A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.