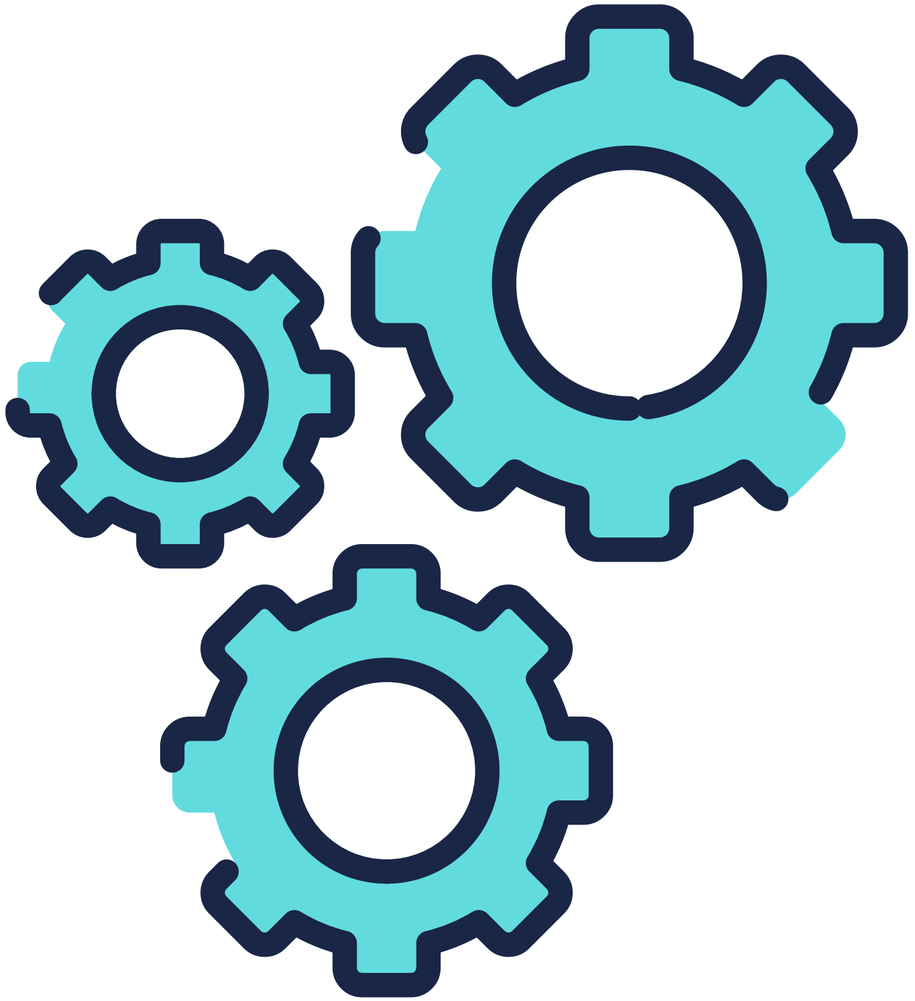
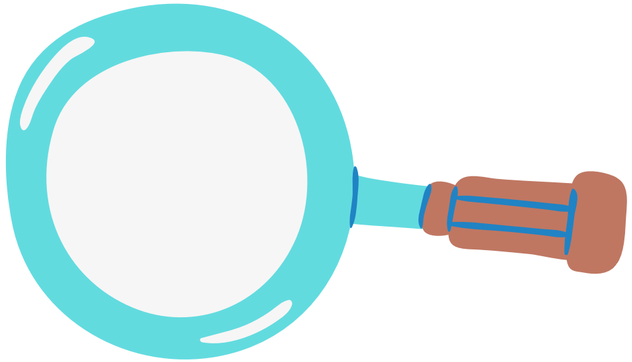
HAND GESTURE RECOGNITION

December 2022

Pierre Samaha

pierresamaha1998/deep_learning_hand_gesture_recognition: A deep learning model for hand gesture recognition (github.com)





PROBLEM STATEMENT

There is still a need to provide a traditional input method, such as keyboards and mice, to replace gestures and be able to communicate with computers and devices.

One of the main challenges in hand gesture recognition is the difficulty of accurately interpreting gestures. This can be due to a number of factors, such as the variability of gestures depending on the individual, cultural context, and other factors.

HOW CAN AI ADD VALUE?

- AI systems can be trained to recognize and interpret a wide range of gestures, making them well-suited to this task.

- Communication and interaction with computers and digital systems are crucial today, making the experience more like interacting with another person. By providing an alternative to traditional input methods, such as keyboards and mice, gesture recognition technology can make it possible for a wider range of people to use and benefit from technology.

- By providing an alternative to traditional input methods, such as keyboards and mice, gesture recognition technology can make it possible for a wider range of people to use and benefit from technology.

- AI-powered gesture recognition systems can enable to accurately recognize and interpret gestures in real-time

MARKET AND END USERS

Sign language translation

Controlling devices with hand gestures.

Home automation

Improve the accessibility of technology for people with physical disabilities or limitations

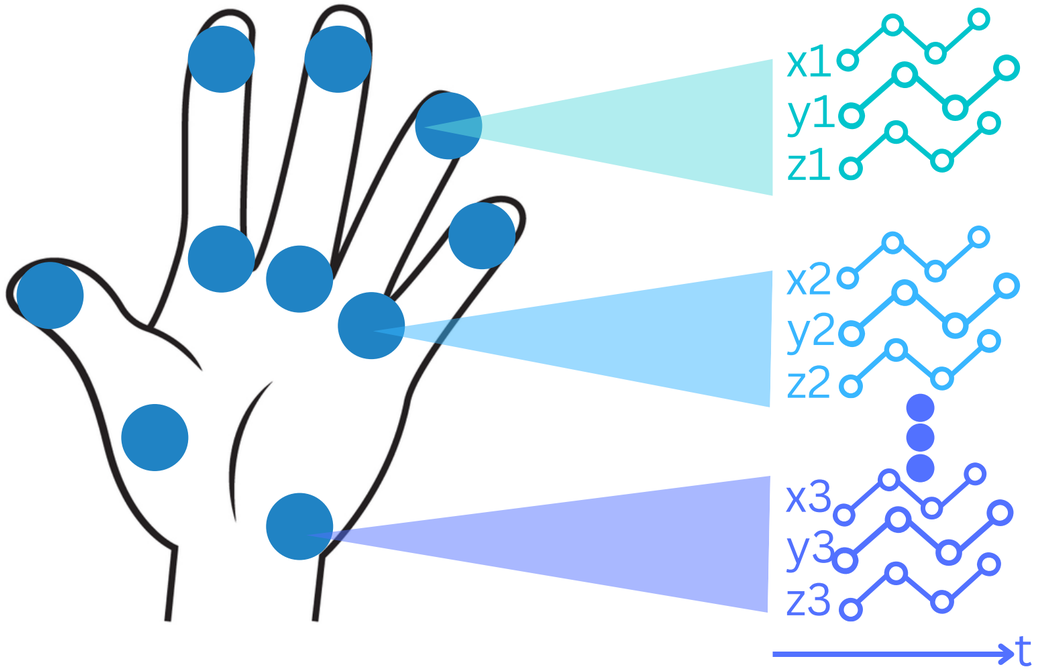
Interacting with one's environment in augmented reality or virtual reality

APPROACH

A deep learning models are used to classify hand gestures.
After researching existing gesture recognition techniques and algorithms such as temporal convolution 1D, we start to explore different machine learning algorithms and architectures, such as recurrent neural networks (RNNs), and experimenting with different methods for extracting motion features based on their temporal patterns.

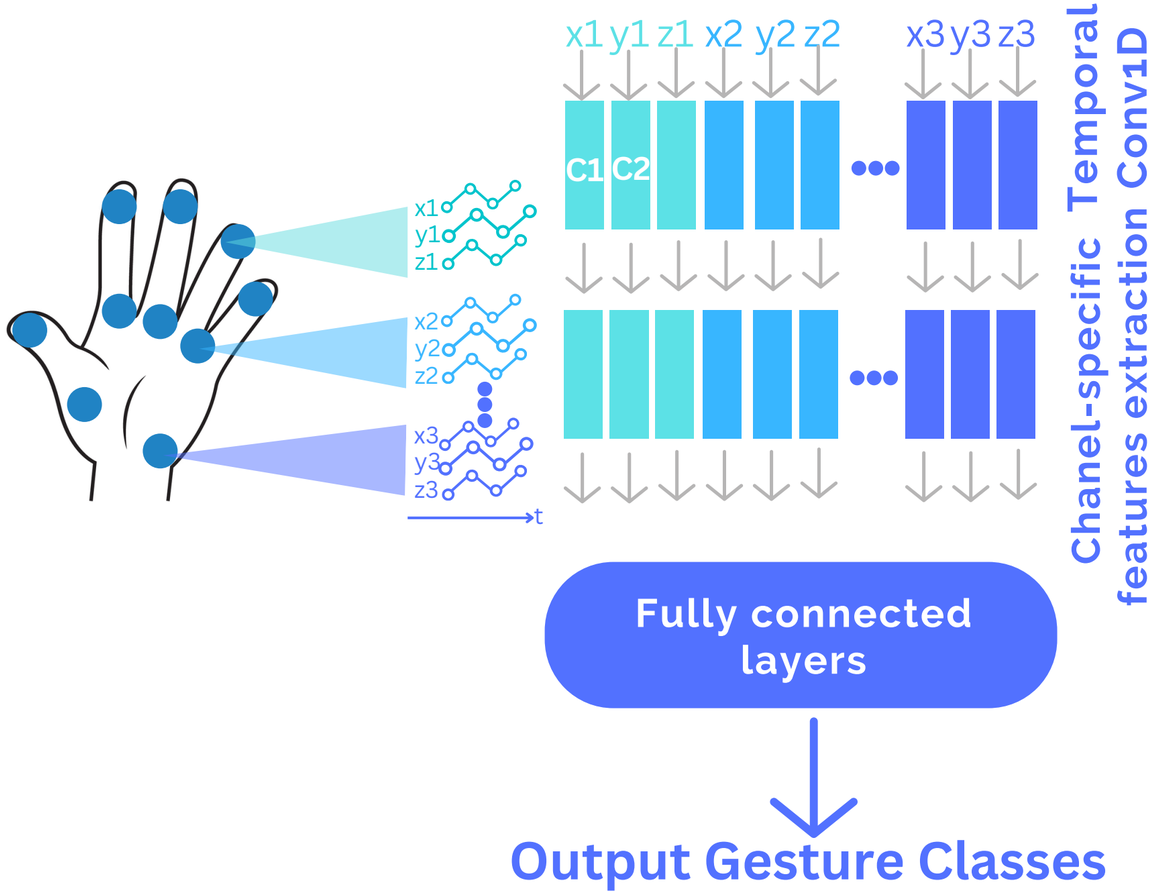
DATASET

Our dataset consists of the motion of the body's (skeletal) joint positions only. Some sensors directly provide streams of body skeletons e.g. Leap Motion, Kinect camera, or motion capture gloves. It is also possible to extract the pose information from videos using vision-based approaches like OpenPose, AlphaPose, or Google's Media Pipe at a good frame rate.



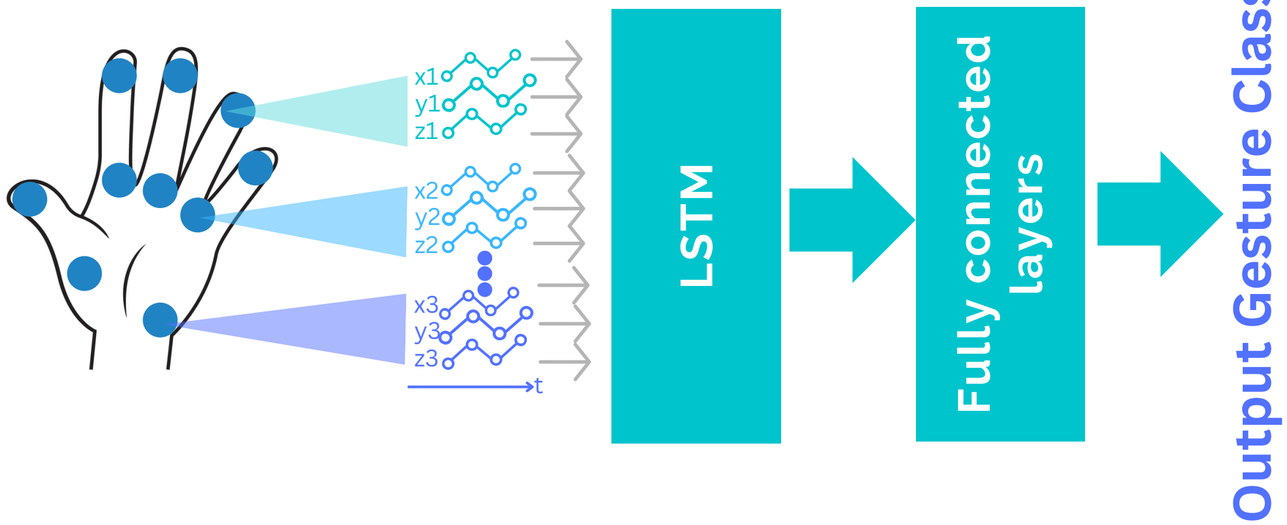
Such skeletal ("pose") representations are lightweight and very sparse compared to image and video representations.

Each hand joint typically has 3 dimensions, to represent its (x,y,z) position in space at a given timestep. A gesture is thus represented by a sequence over time of n_joints joints, or, equivalently by a sequence over time of n_channels.
The model uses such sequences as input. The model expects gestures to be tensors of the following shape: (batch_size, duration, n_channels).



OUR MODEL 1

The neural network extracts motion features, using a dedicated temporal feature extractor made of temporal convolutions for each individual 1D channel. These temporal features are finally used to determine the nature of the gesture performed. Once features have been extracted for each channel, they need to be “merged”. To that extent, they are all fed into a dense neural network (one hidden layer) which performs the final classification.



OUR MODEL 2

We also used LSTM model to solve the problem. The goal of LSTM model is to extract the temporal correlation between the n-joints by keeping a memory of past position of the joints. The LSTM layer is connected to a fully connected layer to get the classification output.

USE CASE

A pizza restaurant chain that uses hand gestures to allow customers to order and pay for their pizzas without having to touch menus or payment terminals. This could reduce the spread of germs and make the dining experience more convenient. This can be very helpful during this Covid-19 period.

This also could be used to allow users to create custom pizzas by selecting toppings and other ingredients. This could provide a fun and interactive way for customers to create their own pizzas. Hence, this increases the number of clients for the restaurant.

THANK
YOU