

HUMAN ACTIVITY RECOGNITION

PROJECT REPORT



SUBMITTED TO

Prof. Rajiv Kapoor

Mr. Nikhil

SUBMITTED BY

Akshat Tiwari 2K18/EP/008

Yash Saini 2K18/EC/195

Project Link [GitHub link](#)

AIM

We aim to develop a machine learning model architecture that uses CNN to recognize human activity.

MOTIVATION

Since many of smartwatch/fitness band only have a 3 axis accelerometer to detect human activity, which does not reveal accurate results.

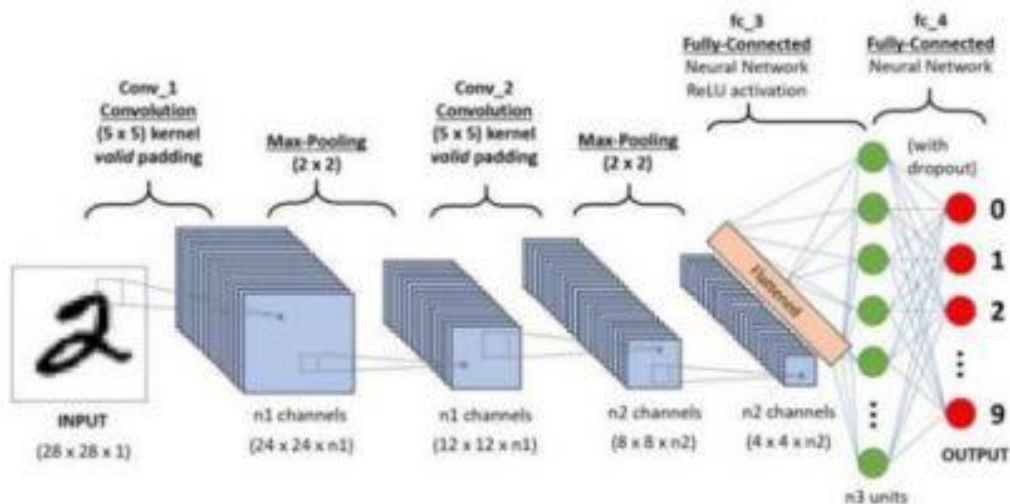
In this project we are using data from 4 different sensors. To predicting more accurately.

- 1) 3 axis accelerometer
- 2) Linear accelerometer
- 3) Gyroscope
- 4) Magnetometer

Convolution Neural Network

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand engineered, with enough training, ConvNets have the ability to learn these filters/ characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlaps to cover the entire visual area.



Technically, deep learning CNN models to train and test, each input image will pass it through a series of convolution layers with filters (Kernels), Pooling, fully connected layers (FC) and apply Softmax function to classify an object with

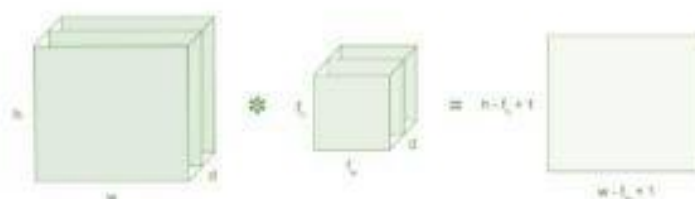
probabilistic values between 0 and 1. The below figure is a complete flow of CNN to process an input image and classifies the objects based on values.

Convolution Layer

Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes two inputs such as

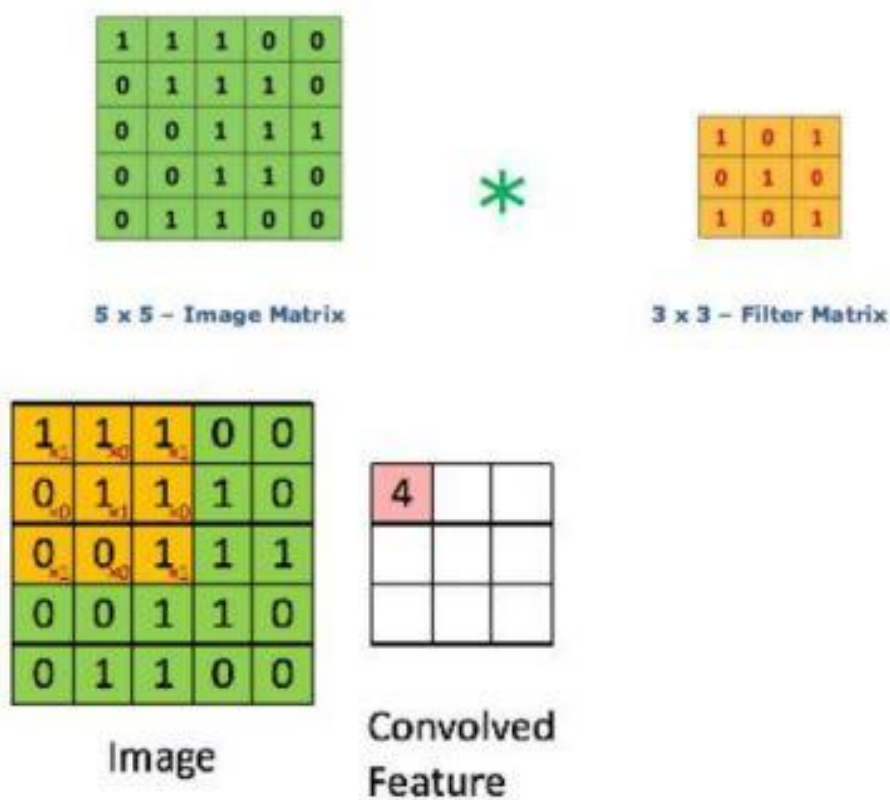
image matrix and a filter or kernel.

- An image matrix (volume) of dimension $(h \times w \times d)$
- A filter $(f_h \times f_w \times d)$
- Outputs a volume dimension $(h - f_h + 1) \times (w - f_w + 1) \times 1$



Consider a 5 x 5 whose image pixel values are 0, 1 and filter matrix 3 x 3 as shown in below

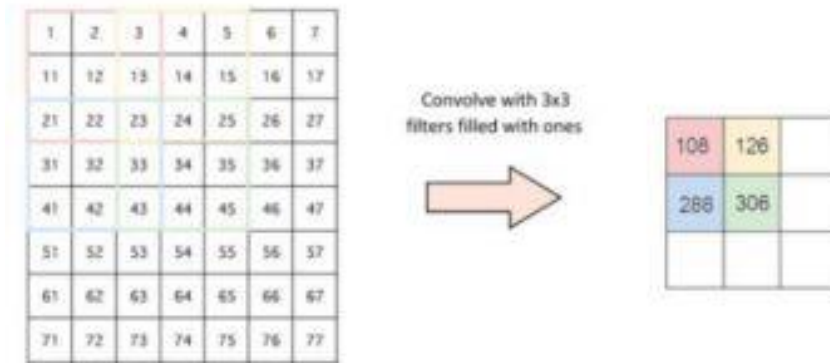
Then the convolution of 5 x 5 image matrix multiplies with 3 x 3 filter matrix which is called “Feature Map” as output shown in below



Convolution of an image with different filters can perform operations such as edge detection, blur and sharpen by applying filters. The below example shows various convolution images after applying different types of filters (Kernels).

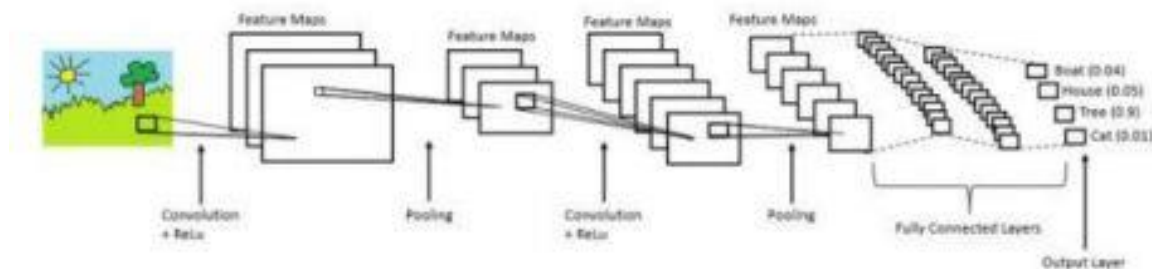
Strides

Stride is the number of pixels shifted over the input matrix. When the stride is 1 then we move the filters to 1 pixel at a time. When the stride is 2 then we move the filters to 2 pixels at a time and so on. The below figure shows convolution would work with a stride of 2.



Padding

Sometimes filters do not perfectly fit the input image. We have two options: Pad the picture with zeros (zero-padding) so that it fits Drop the part of the image where the filter did not fit. This is called valid padding which keeps only the valid part of the image.



DATASET USED

[Dataset Link](#)

The dataset consist of x, y, z readings from 4 different sensors in total of 12 readings and 8 different output instances

'walking', 'standing', 'jogging', 'sitting', 'biking', 'upstairs', 'downstairs', 'upstairs'

RESEARCH METHODOLOGY

Step 1:- Downloading and extracting dataset.

Step 2:- Data Pre-processing

Step 3:- Model Validation

Step 4:- Model Development

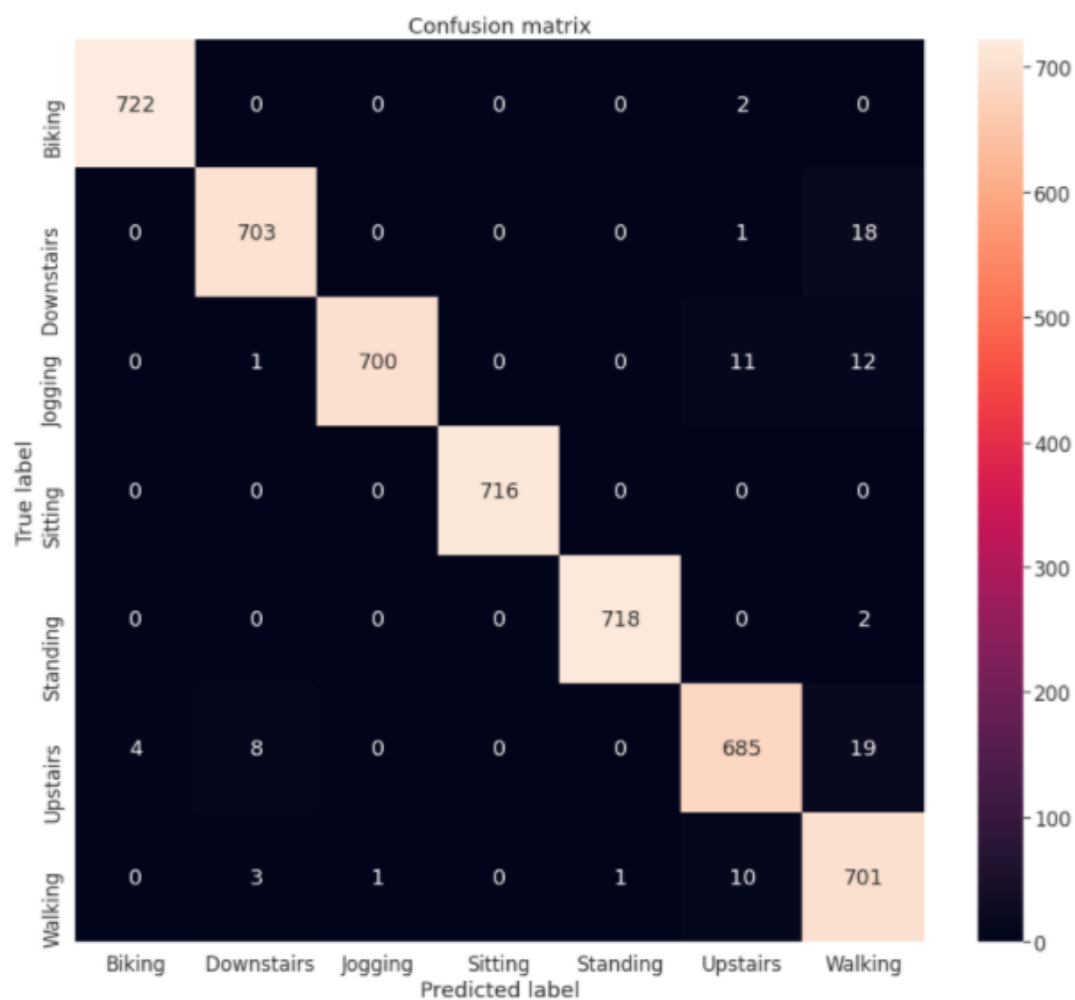
Step 5:- Making Prediction

Step 6:- Performance Measures

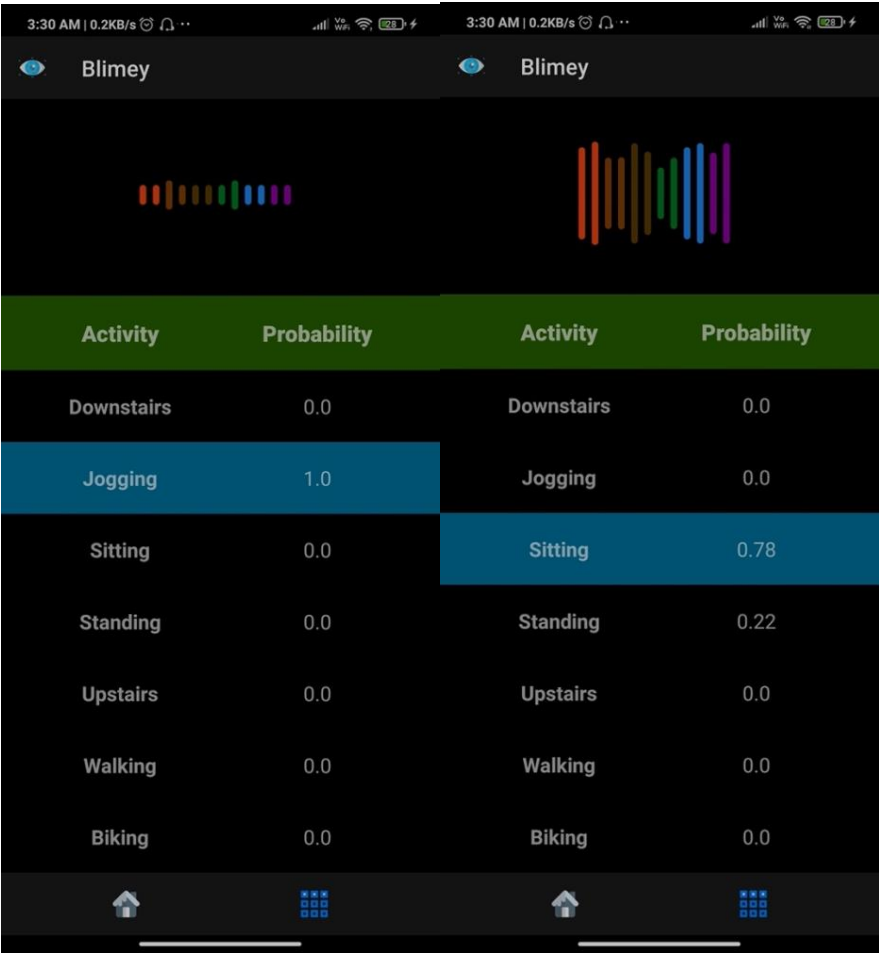
Result

Achieved accuracy of 98.15%

[GitHub link](#)



App
[App Link](#)



REFERENCES

- 1) Peizhen Zhang, Yongyi Tang, Jian-Fang Hu, and Wei-Shi Zheng (2020), *Fast Collective Activity Recognition Under Weak Supervision*, In *IEEE TRANSACTIONS ON IMAGE PROCESSING*, VOL. 29, 2020.
- 2) <https://developer.android.com>
- 3) www.tensorflow.org