International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com

ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105

Volume 8, Issue 2
Febraury 2020.
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395

Received on: 04.02.2020 Published on: 28.02.2020 Pages : 52-56

DEEP LEARNING METHODS ANALYSIS AND IMPLEMENTATION OF HUMAN ACTION RECOGNITION

R. Maheswari [a*], K M Amira Thasneem [b], M.Maharajan[c]

[a,c] Department of Electronics and Communication Engineering,
Agni college of Technology. Chennai 600130 Tamil nadu India
[a] e-mail: maheswari.ece@act.edu.in
P.G. Student, Communication Systems, Agni College Of Technology, Chennai, India

Abstract

In recent times many deep learning methods like recurrent neural networks and convolutional neural networks have shown considerable results that automatically learns the necessary features from the sensor data that is given rawly. In this paper the difficulties of human action recognition are analysed and a suitable enhancement method is suggested by introducing a smart method for the identification of human activities from skeletal joint motions. Analysis of various methods used in human action recognition is also done.

Keywords: Deep learning methods , Neural networks, Human action recognition

1.0 Introduction

Human activity recognition involves prediction of the movement of a person based on the sensor data that is given and involves deep learning method from signal processing and the features that is taken from the raw data is fit in order to a machine learning model. The Streams of the sensor data are generally split into subs-sequences that are called as windows, and each of these window is associated with a next broader activity, approach called a sliding window . Convolutional neural networks and long short-term memory networks both together, are best suited to learn the features from the raw sensor data and predict the movement associated with it.Human activity recognition, or HAR for short, is a broad field of study concerned with identifying the specific movement or action of a person based on sensor data. Movements are activities that are performed in indoors like walking, standing and sitting. They may also be focused activities like types of activities that are performed in a kitchen. The data may be remotely recorded, such as video, radar, or other wireless methods. The data may be alternately recorded directly on this subject by carrying custom hardware or the smart phones that usually contain accelerometers and gyroscopes. The data for activity recognition was difficult and very expensive to collect and requires the custom hardware. For fitness and health monitoring phones and tracking devices are used that is cheap. As such, data from these devices is cheaper to collect, more common, and therefore is a more commonly studied version of the general activity recognition problem.

TECHNOLOGIES USED:

DEEP LEARNING PROCESS: Deep learning is a computer software that mimics the network of neurons in a brain. It is a subset of machine learning and is called deep learning because it makes use of deep neural networks.

Fig 1. Deep Learning Process

Deep learning algorithms are constructed with connected layers.

- The first layer is called the Input Layer
- The last layer is called the Output Layer
- All layers in between are called Hidden Layers. The word deep means the network join neurons in more than two layers.

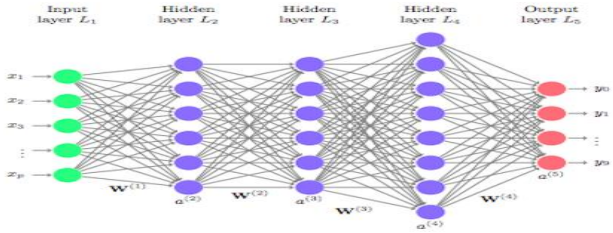

Fig 2. Deep Learning Layers

RECURRENT NEURAL NETWORKS (RNNS) RNN is a multi-layered neural network that can store information in the context nodes and allows it to learn the data sequences and output the number or any other sequence. It is suited for input sequence processing.

CONVOLUTIONAL NEURAL NETWORKS (CNN) CNN is a multi-layered neural network that has a very unique architecture designed to extract complex features of the data at each of the layer to determine the output. It is well suited for perceptual task.

2.0 Related Work

The algorithm outperforms existing state-of-the-arts significantly, which is measured in terms of recognition accuracy, receiver operating characteristic (ROC) curve and area under the curve (AUC)[1].The proposed method achieves real time performance with PCA dimension reduction because of sparse coding in lower dimensional space[5].Experimental results show that the HSTM can successfully classify human activities with higher accuracies on single-person actions (UCF) than other existing methods[3].Compared to the classical spectral geometry algorithms, such as PCA, LDA, GMSS, LSDA, and MFA, EMRP shows many attractive and competitive properties to measure the similarity between different acceleration signals[2].It achieves remarkable performance on high-level activity recognition, due to the joint contributions from all the context elements[4].

3.0 Proposed System

The objective of the paper is to recognize Human Activities using movements of skeletal joints. To design an intelligent human action and human gesture recognition system which can automatically recognize the human daily activities using human skeleton information, combining the techniques of image processing and deep learning. Here we are going to capture the different skeletal images of human actions and collect it as data sets ,train the data based on algorithms using deep learning.The proposed system discusses the development of an effective skeleton information based HAR recognition and Uses Logistic Regression algorithms to achieve the

purpose. Optimizing technique like SGD - stochastic gradient descent and regularization methods like ReLU and ELU to increase the accuracy.There are a number of difficulties when directly treating raw features as descriptors of activities, e.g., different scales, different ranges, weak discriminative power and descriptive capability.It is difficult to directly compare the envelope of acceleration signals in testing stage, because they are embedded in a nonmetric space.Depth videos are used and frames are combined to find out the motion which results in errors.

4.0 Implementation

To begin the process with the collection of data and constantly feeding data can improve the model's performance.From existing data a new set of training data can be created and image resizing and pixel scaling can be performed to interact with the model.Then the dataset can be separated by the proposed algorithm .Thus the action is recognised and checked with the fed dataset .The advancement of algorithms increase accuracy .ELU can be used to reduce the noises . We will be performing loss minimization such as cross entropy and finally validate the mode by giving real time inputs.

5.0 Module Description

- Dataset collection
- Dataset Augmentation
- Dataset pre-processing Module
- Feature Extraction with Hdf5 Dataset Generator
- Database Integration

DATASET COLLECTION:

A data set is a collection of data. Deep Learning method is used for solving various problems that are challenging. It is one of the method that is used for computer vision tasks.

DATASET AUGMENTATION:

The performance of deep learning neural networks is often improved with the amount of data available. Data augmentation is a technique that is used to artificially create new training data from the existing training data. This is done by applying particular domain-specific techniques to examples from the training data that creates a new and different training examples. Image data augmentation is perhaps the well-known type of augmentation and it involves creating the transformed versions of the images in the training dataset that initially belongs to the same class as the original image. It includes a range of operations in the field of manipulating image, such as flip, shift, zoom and many more operations.

DATA PREPROCESSING MODULE:

Dataset preprocessing module has three steps which involves dataset collection, data augmentation and also data preprocessing. Dataset augmentation encompasses a wide range of techniques used to generate new training samples from the original ones by applying random jitters and perturbations such that the classes labels are not changed.In the data preprocessing steps we will preprocess the data using different methods to get the exact output such as mean preprocessing, simple preprocessing, image to array conversion and label encoding.

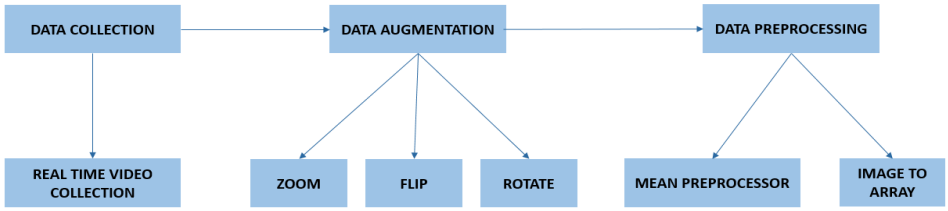

Fig 3. Data Pre-processing Module

FEATURE EXTRACTION WITH HDF5 GENERATOR:

HDF5 is binary data format that is created by the HDF5 group to store various gigantic numerical datasets on the disk (that are far too large to be stored in memory) which facilitates easy access and computation in the rows of the datasets. Data in the HDF5 is stored hierarchically, similarly to how a file system actually stores data. Data is first defined in groups, where a group resembles a container-like structure that can hold the datasets and other groups. Once a group is been defined, a dataset can be created within that group. A dataset can be thought of as a multi-dimensional array (i.e., a NumPy array) of a homogeneous data type (integer, float, unicode, etc.).Feature extraction includes several convolution layers followed by max-pooling and an activation function. For extracting the features we will be using VGG16 architecture.

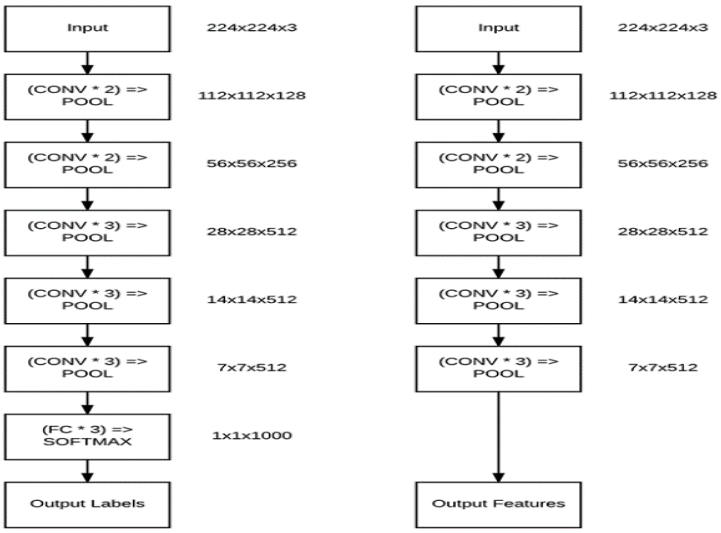

Fig 4. Feature Extraction With Hdf5 Generator

FINE TUNING WITH NETWORK SURGERY:

Fine-tuning is a powerful method used to obtain image classifiers from the pre-trained CNNs on custom datasets, which are even more powerful thanthe feature extraction in most of the cases. Fine-tuning is a type of transfer learning. We apply fine-tuning to deep learning models that have already been trained on a given dataset. Typically, these networks resembles the state-of-the-art architectures such as ResNet, VGG and Inception which have been trained on the ImageNet dataset. Instead of simply applying the feature extraction technique, we are going to perform network surgery and modify the original architecture so that we can re-train the parts of the network. AlexNet model is trained and using network surgery we will integrate it with a pretrained model output got from the VGG16 architecture.

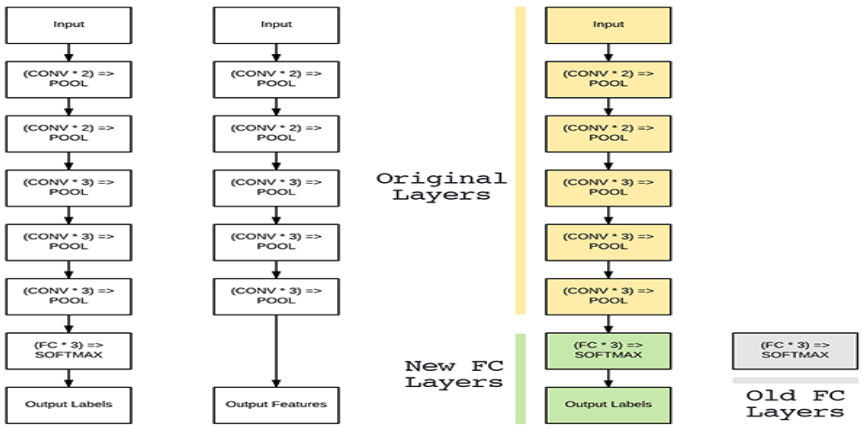

Fig 5. Fine Tuning With Network Surgery

DATABASE INTEGRATION:

MongoDB is an open-source cross-platform document-oriented database program that is used from the database management. The Database is a physical container for the collections. Each database gets its own set of files on the file system. A single MongoDB server has multiple databases in it .Collection is a group of MongoDB documents. It is the equivalent of an RDBMS table. A document is a set of key-value pairs. Documents have dynamic schema. Dynamic schema means that documents in the same collection does not need to have the same field sets or structure, and the common fields in a collection's documents that may hold different types of data in it.

6.0 Conclusion

Thus ideas to overcome difficulties in using image based human action recognition is proposed .In the coming future, we review the application of the human activity sensing technology in the recognition field and it can promote for all type of recognition with more accuracy compared to this project. In this field there are more chance to develop or convert this project in many ways. The accuracy of the prediction will be increased by using different efficient techniques and algorithms. A literature survey based on the Human Activity Recognition was successfully done. Based on this we arrived at a conclusion to develop a deep learning approach. The following are the modules that are proposed are Dataset collection, Dataset Augmentation, Dataset pre-processing Module, Feature extraction and Database Integration. The idea can be extended by creating a webpage and also the applications of this in various fields like medicine, surveillance etc.

References:

[1] Chhavi Dhiman, Member IEEE, Dinesh Kumar Vishwakarma, Senior Member, IEEE , "View-invariant Deep Architecture for Human Action Recognition using Two-stream Motion and Shape Temporal Dynamics"[2020, Vol No:0098-3063]

[2] Dapeng Tao, Lianwen Jin, Member, IEEE, Yuan, Senior Member, IEEE, And Yang Xue, "Ensemble Manifold Rank Preserving For Acceleration-Based Human Activity Recognition" [2016, Vol No: 2162-237X]

[3] Wanru Xu, Zhenjiang Miao, Member, IEEE, Xiao-Ping Zhang, Senior Member, IEEE, Yi Tian, "A Hierarchical Spatio-Temporal Model For Human Activity Recognition" [2017,Vol No: 1520-9210]

[4] Lei Wang, Student Member, IEEE, Xu Zhao, Member, IEEE, Yunfei Si, Liangliang Cao, Member, IEEE, and Yuncai Liu, Member, IEEE , "Context-Associative Hierarchical Memory Model For Human Activity Recognition And Prediction" [2016, Vol No: 1520-9210]*

[5] Jin Qi , Zhangjing Wang, Xiancheng Lin, And Chunming Li "Learning Complex Spatio-Temporal Configurations Of Body Joints For Online Activity Recognition" [2018, Vol No: 2168-2291]

[6] Seo Yul Kim, Hong Gul Han, Student Member IEEE, Jin Woo Kim, Sanghoon Lee, Senior Member IEEE And Tae Wook Kim, Senior Member IEEE , "A Hand Gesture Recognition Sensor Using Reflected Impulses" [2017, Vol No1558-1748]

[7] Kishore P.V.V,. Senior Member, IEEE, D.Anil Kumar, Student, Ieee, A.S.C.S.Sastry, Member, Ieee, And E.Kiran Kumar, Student, IEEE , "Motionlets Matching With Adaptive Kernels For 3D Indian Sign Language Recognition" [2018, Vol No: 1558-1748]

[8] Jianjie Lu And Kai-Yu Tong, Senior Member, IEEE "Robust Single Accelerometer-Based Activity Recognition Using Modified Recurrence Plot" [2019,Vol No: 1558-1748]