International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

LEVERAGING CNN AND TRANSFER LEARNING FOR VISION-BASED HUMAN ACTIVITY RECOGNITION

ADABALA TRINADH[1], S.K.ALISHA[2]

[1]MCA Student, B V Raju College, Kovvada, Andhra Pradesh, India.
[2]Associate Professor, B V Raju College, Kovvada, Andhra Pradesh, India.

Abstract:

One of the active research areas in computer system vision for several settings, including safety and security monitoring, healthcare, and human computer interface, is human activity recognition. Several methods for human action recognition using depth, RGB (red, green, and blue), and skeletal system datasets have been published in recent years. Most of the methods that have been developed for activity categorization utilising skeletal system datasets have limitations when it comes to representing characteristics, complexity, and performance. The challenge of providing a trustworthy approach to human action discrimination using a skeleton dataset remains, however. For optimal feature extraction for precise activity categorization, we use depth images as input one and a suggested moving joints descriptor (MJD) as input two. The MJD depicts the change in position of the body's joints over time. In order to train CNN networks with various inputs, we are getting ready to deploy neural networks for rating combination. We proposed running the code on publicly available datasets such as MSRAction3D.

Keywords: MJD, MSRAction3D, CNN, Human action.

INTRODUCTION

Not only is human activity recognition a challenging research problem, but it has also been a popular topic for quite some time due to its broad use in many different applications, such as smart security systems, human-robot communication, and home care systems [1]. The advancement of deep learning in recent years has led to widespread use of Convolutional Neural Networks (CNNs), which have achieved remarkable efficiency on monitoring, detection, classification, and detection tasks. CNNs are especially useful in computer vision and pattern recognition [2]. There are a few things to keep in mind while thinking about action acknowledgment. In order to keep up with technological advances and make it possible for people with disabilities to communicate with creators and understand their tasks through computer systems, standard methods of interaction are being developed as human-machine interaction grows in importance as a field of study in multimedia processing [3]. A lot of studies have attempted to use motion assessment to model and then identify people's behaviour. We focus on human behaviour assessment from video clips in this work. It's worth mentioning that a lot of information is hidden under gestures, fast movements, and walking speed. [1, 2] The expressive qualities provided by the two types of data have encouraged researchers in the field of human action recognition to focus on depth maps and body postures as representations of the action. Third The effectiveness of an activity recognition system relies on a detailed representation that provides unique characteristics of each task for categorization [5]. Because two movements could seem same from the front but distinct from the side, employing deepness map data for activity detection is still difficult for certain activities, leading to incorrect categorization [4]. When large obstructions are present, the depth maps captured by the depth cameras might be rather noisy, and the tracked joints' 3D settings can be completely off, leading to an increase in the actions' intraclass variances [5]. They propose "Skepxels," a spatial-temporal framework for skeletal series that makes full advantage of "regional" connections between joints by using CNN's 2D convolution bits. After using Skepxels to convert skeletal motion movies into salable images, they build a convolutional neural network (CNN) architecture for accurate human action identification using the produced images [9].

LITERATURE SURVEY

(Wang, Zhao-Xuan) [7] In order to validate the proposed technique, the following datasets were utilised: KTH and TJU, two well-known deepness datasets (MSR action 3-D and MSR daily task 3-D), and MV-TJU, a groundbreaking multiview multimodal dataset. Extensive experimental evidence in RGB and depth modalities demonstrates that this method outperforms popular, less expensive alternatives based on 2-D/3-D component models for a wide variety of human actions. I am C. Krishna Mohan: [8] Using a deep fully convolutional style, they suggested using action financial institution incorporates to identify human activities in videos. The similarities between the video and action bank movies are described by direct patterns called action bank characteristics. They are computed using an activity ban, which is a set of prescribed photographs. To Hiroshi Miki: [9] They provide an approach to recognition that examines the connection between human behaviour and items functions; the objective is to enhance our method by including human activities into dynamic item segmentation. To quote Ziaeefard and Maryam: [10] A state-of-the-art method for human action recognition based on normalized-polar histograms is proposed in this study. The process of amassing skeletal pictures was clarified.

It is a pattern of movement that uses range and angle. The colour cyan is used to emphases one of the most important aspects of this task. The symbol is formed by encircling the core with all of the skeletal system model frames. A two-level multi-class support vector machine (SVM) was used to classify people's behaviours; the model was trained using generic functions first, and then with salient features. Mejdi DALLEL: [11] They revealed a large-scale RGB+S keleton action acknowledgment dataset called "Industrial Human Being Action Recognition Dataset (InHARD)". We have 4804 unique samples of industrial activity in our collection, spread out among 38 films representing 14 various categories. In order to evaluate our dataset using the proposed metrics, they finish building an end-to-end regression classification LSTM network. Si Yang: [12] By combining a pattern recognition semantic network with an autoencoder, they were able to construct a novel semantic network. human activities recognition using deep neural networks. Confirmation of the design they suggested came from The benefits of the model were discovered via experiments. By comparing results, they produced a readily accessible model. Many noteworthy achievements: The researchers uncovered a novel approach of merging many frames of data into a single image. This method has several applications. a method for automatically extracting features of human actions using a deep neural network Hello, Chen Xu, Lei Zong, and HongLin Yuan! [12] The current state of RF fingerprint identification is flawed because it uses a preset resolution formula, which has a small range of potential uses and requires a lot of previous information. A convolutional semantic network (CNN) RF fingerprint identification approach would be ideal for handling these issues. Three main aspects of the research are RF fingerprint extraction, convolutional neural network architecture, and wireless transmitter identification and verification. There is no fraudulent usage of fingerprint info [13]. Advanced convolutional neural networks (DCNNs) outperform traditional approaches that rely on hand-crafted functions. A new FLD approach called an upgraded DCNN with photo scaling is available, although many CNN designs suffer from taken-care-of-scale photos. The complexity matrix is used as an efficiency indicator in FLD for the first time. The amounts of the hypothetical results using the LivDet 2011 and LivDet 2013 datasets provide further evidence that our technique is more efficient at discovery than others. [14] A highly accurate computerized latent fingerprint recognition system is required to compare concealed fingerprints found at crime scenes to a large database of referral prints and provide a list of potential friends. Their Convolutional Neural Network (ConvNet)-based automatic unexposed fingerprint recognition method achieves 64.7% accuracy with the NIST SD27 dataset and 75.3% accuracy with the WVU dataset when tested against a reference data source of 100,000 rolled prints.

EXISTING SYSTEM

Occasionally, it may be more convenient for customers to browse merchandise rather than wait in queue to pay, as the latter takes up even more of their time. Now we're building this system that may serve the customer in every way and also the shop owner by drawing inspiration from this situation that was common in all the shops. Consequently, we devise a technique that allows the customer to comprehend their expenses as they add things to the basket. If a consumer buys anything from this grocery shop basket, they can quickly charge it, which is the greatest and most practical example.

PROPOSED SYSTEM

This method introduces innovative advancements compared to the current purchasing mechanism. Providing a web-based, centralised invoicing system is the main objective of this project. Not only does it have automatic billing, but it also has certain unique capabilities.The word "Supermarket Basket" is new to us.

MODULES EXPLANATION

1. **Data Collection:** The first step is to gather multivariate time series information from the phone's and the watch's sensing units. The sensors are tested with a continuous frequency of 30 Hz. Afterwards, the sliding home window method is utilized for segmentation, where the moment collection is split right into subsequent home windows of taken care of period without inter window voids (Banos et al., 2014). The sliding window technique does not call for pre-processing of the time series, and is therefore preferably matched to real-time applications.

2. Preprocessing: Filtering is done after that to remove loud worth and outliers from the accelerometer time series data, to ensure that it will certainly be appropriate for the feature extraction phase. There are two basic kinds of filters that are normally utilized in this action: average filter (Sharma et al., 2008) or average filter (Thiemjarus, 2010). Given that the kind of sound managed here resembles the salt and pepper sound found in images, that is, severe acceleration worth that happens in single snapshots scattered throughout the time collection. Consequently, an average filter of order 3 (window dimension) is put on remove this sort of noise.

3. Function Extraction: Below, each resulting section will be summed up by a fixed number of attributes, i.e., one attribute vector per segment. The used attributes are extracted from both time and frequency domains. Given that, numerous activities have a repetitive nature, i.e., They consist of a collection of motions that are done occasionally, like walking and running. This frequency of rep, likewise referred to as dominant regularity, is a detailed attribute and therefore, it has been taken into account.

4. Standardization: Because, the moment domain name features are gauged in (m/s 2), while the regular ones in (Hz), therefore, all attributes ought to have the exact same range for a reasonable comparison in between them, as some category formulas make use of distance metrics. In this action, Z-Score standardization is made use of, which will transform the credit to have absolutely no mean and unit variance, and is defined as

xnew = $(x - \mu)/ \sigma$ where μ and σ are the quality's mean and standard deviation respectively (Gyllensten, 2010).

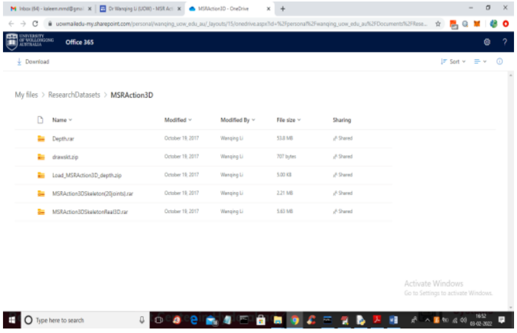
Human Activity Recognition from deepness maps and Poses utilizing Deep

Discovering In this paper, author is using the CNN (Convolution Neural Networks) algorithm to identify human action as this formula will remove crucial attributes by filtering the exact same information multiple times in order to make the most of possibilities of precise activity category, CNN networks are educated with different inputs features which will certainly not happen in existing RGB Deepness algorithm which will get train on two features such as pictures and skeletal system information.

As existing formulas are not reliable, so writers make use of the CNN formula which currently verifies its success in numerous fields such as image classification, weather and stock forecast etc

. To educate the CNN formula writer is making use of MSRAction3D skeleton dataset which has 20 various actions such as 'high arm wave', 'straight arm wave', 'hammer', 'hand catch', 'ahead punch', 'high throw', 'draw x', 'attract tick', 'attract circle', 'hand clap',' 2 hand wave', 'side-boxing', 'bend', 'ahead kick', 'side kick', 'jogging', 'tennis swing', 'tennis serve', 'golf swing', 'grab & throw'.

All this activities data is extracted from the MSRAction3D dataset and below are the display shots of that dataset

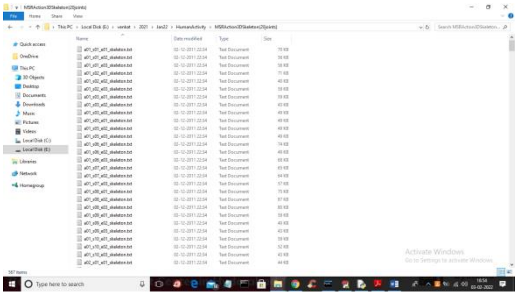
2. 

Inbox (84) - kaleem.mmd@gmail
Dr Wanqing Li (UOW) - MSR Act
MSRAction3D - OneDrive
uowmailedu-my.sharepoint.com/personal/wanqing_uow_edu_au/_layouts/15/onedrive.aspx?id=%2Fpersonal%2Fwanqing_uow_edu_au%2FDocuments%2FRese...
UNIVERSITY OF WOLLONGONG AUSTRALIA
Office 365
Download
Sort
Activate Windows
Go to Settings to activate Windows.
Type here to search
16:52
03-02-2022

My files > ResearchDatasets > MSRAction3D

Name	Modified	Modified By	File size	Sharing
Depth.rar	October 19, 2017	Wanqing Li	53.8 MB	Shared
drawskt.zip	October 19, 2017	Wanqing Li	707 bytes	Shared
Load_MSRAction3D_depth.zip	October 19, 2017	Wanqing Li	5.00 KB	Shared
MSRAction3DSkeleton(20joints).rar	October 19, 2017	Wanqing Li	2.21 MB	Shared
MSRAction3DSkeletonReal3D.rar	October 19, 2017	Wanqing Li	5.63 MB	Shared

Below is a screenshot of the dataset files. The dataset, named "MSRAction3DSkeleton(20joints)," was obtained from the aforementioned URL. Since it was recorded using DEPTH cameras, it will only record skeleton values.

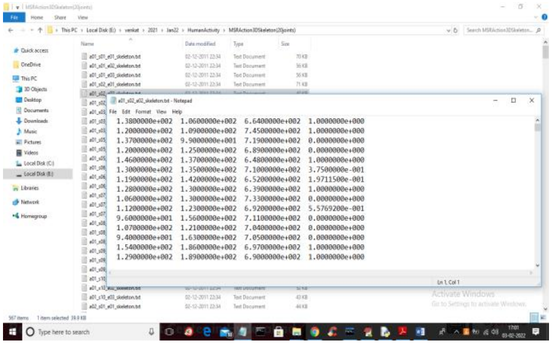


This PC > Local Disk (E:) > venkat > 2021 > Jan22 > HumanActivity > MSRAction3DSkeleton(20joints)

Name	Date modified	Type	Size
a01_s01_e01_skeleton.txt	02-12-2011 22:34	Text Document	70 KB
a01_s01_e02_skeleton.txt	02-12-2011 22:34	Text Document	56 KB
a01_s01_e03_skeleton.txt	02-12-2011 22:34	Text Document	58 KB
a01_s02_e01_skeleton.txt	02-12-2011 22:34	Text Document	71 KB
a01_s02_e02_skeleton.txt	02-12-2011 22:34	Text Document	40 KB
a01_s02_e03_skeleton.txt	02-12-2011 22:34	Text Document	58 KB
a01_s03_e01_skeleton.txt	02-12-2011 22:34	Text Document	59 KB
a01_s03_e02_skeleton.txt	02-12-2011 22:34	Text Document	43 KB
a01_s03_e03_skeleton.txt	02-12-2011 22:34	Text Document	49 KB
a01_s05_e01_skeleton.txt	02-12-2011 22:34	Text Document	49 KB
a01_s05_e02_skeleton.txt	02-12-2011 22:34	Text Document	49 KB
a01_s05_e03_skeleton.txt	02-12-2011 22:34	Text Document	49 KB
a01_s06_e01_skeleton.txt	02-12-2011 22:34	Text Document	74 KB
a01_s06_e02_skeleton.txt	02-12-2011 22:34	Text Document	49 KB
a01_s06_e03_skeleton.txt	02-12-2011 22:34	Text Document	68 KB
a01_s07_e01_skeleton.txt	02-12-2011 22:34	Text Document	65 KB
a01_s07_e02_skeleton.txt	02-12-2011 22:34	Text Document	64 KB
a01_s07_e03_skeleton.txt	02-12-2011 22:34	Text Document	57 KB
a01_s08_e01_skeleton.txt	02-12-2011 22:34	Text Document	75 KB
a01_s08_e02_skeleton.txt	02-12-2011 22:34	Text Document	87 KB
a01_s08_e03_skeleton.txt	02-12-2011 22:34	Text Document	80 KB
a01_s09_e01_skeleton.txt	02-12-2011 22:34	Text Document	58 KB
a01_s09_e02_skeleton.txt	02-12-2011 22:34	Text Document	40 KB
a01_s09_e03_skeleton.txt	02-12-2011 22:34	Text Document	43 KB
a01_s10_e01_skeleton.txt	02-12-2011 22:34	Text Document	59 KB
a01_s10_e02_skeleton.txt	02-12-2011 22:34	Text Document	52 KB
a01_s10_e03_skeleton.txt	02-12-2011 22:34	Text Document	45 KB
a02_s01_e01_skeleton.txt	02-12-2011 22:34	Text Document	44 KB

567 items

All of the files in the dataset display basic information; for example, "a01" denotes activity 1 out of a possible twenty, "s01" is the subject ID, and "e01" is the circumstances ID. Following training, whenever we publish any kind of article, CNN will be able to use the aforementioned data to make predictions about future actions. As you can see in the screen capture below, every document will have skeleton values.


a01_s02_e02_skeleton.txt - Notepad
File Edit Format View Help
This PC > Local Disk (E:) > venkat > 2021 > Jan22 > HumanActivity > MSRAction3DSkeleton(20joints)

```
1.3800000e+002  1.0600000e+002  6.6400000e+002  1.0000000e+000
1.2000000e+002  1.0900000e+002  7.4500000e+002  1.0000000e+000
1.3700000e+002  9.9000000e+001  7.1900000e+002  0.0000000e+000
1.2000000e+002  1.2500000e+002  6.8900000e+002  0.0000000e+000
1.4600000e+002  1.3700000e+002  6.4800000e+002  1.0000000e+000
1.3000000e+002  1.3500000e+002  7.1000000e+002  3.7500000e-001
1.1900000e+002  1.4200000e+002  6.5200000e+002  1.9711500e-001
1.2800000e+002  1.3000000e+002  6.3900000e+002  1.0000000e+000
1.0600000e+002  1.3000000e+002  7.3300000e+002  0.0000000e+000
1.1200000e+002  1.2300000e+002  6.9200000e+002  5.5769200e-001
9.6000000e+001  1.5600000e+002  7.1100000e+002  0.0000000e+000
1.0700000e+002  1.2100000e+002  7.0400000e+002  0.0000000e+000
9.4000000e+001  1.6300000e+002  7.0500000e+002  0.0000000e+000
1.5400000e+002  1.8600000e+002  6.9700000e+002  1.0000000e+000
1.2900000e+002  1.8900000e+002  6.9000000e+002  1.0000000e+000
```

The following components were developed in order to carry out this assignment:

1) Following MSRAction3D For example, we may use this module to upload our activity dataset to an app.

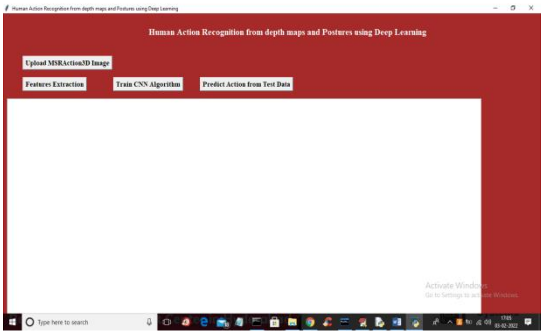
2) Includes Removal: This module will examine all files, delete functions (dataset values), and then visualise the results in a chart fashion. The action value will be considered the course title.

Third, we have the train convolutional neural network (CNN) algorithm, which takes in the extracted functions, trains it, and then uses test data from the experienced version to determine the correctness and complexity of the resulting matrix graph.

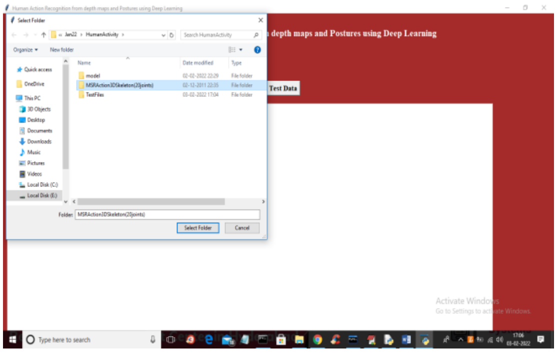
4. Anticipate Activity from Test Data: This component allows users to submit test data, which is subsequently processed by CNN. The network then checks the functionalities in the test documents and determines the activity based on those results.

RESULTS EXPLANATION

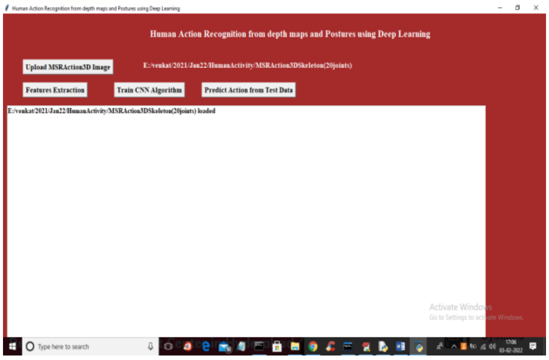
To run project double click on ‘run.bat’ file to get below screen


Human Action Recognition from depth maps and Postures using Deep Learning
Upload MSRAction3D Image
Features Extraction
Train CNN Algorithm
Predict Action from Test Data

To upload the dataset, go to the first screen and click on the "Upload MSRAction3D Image" button. Then, you'll see the second page.


Select Folder
depth maps and Postures using Deep Learning
model
MSRAction3DSkeleton(20joints)
TestFiles
Folder: MSRAction3DSkeleton(20joints)
Select Folder
Cancel

In the previous page, choose and submit the MSRACTION dataset. Then, to load the dataset, click on the "Select folder" button. This will bring up the following screen.

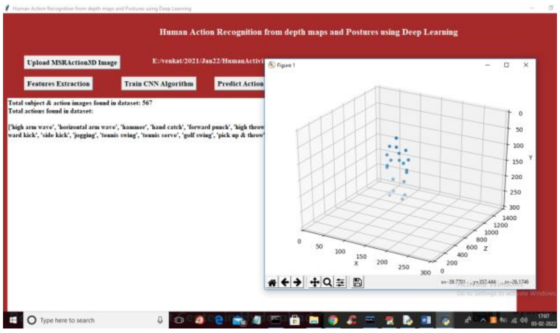


Human Action Recognition from depth maps and Postures using Deep Learning

E:/venkat/2021/Jan22/HumanActivity/MSRAction3DSkeleton(20joints)

E:/venkat/2021/Jan22/HumanActivity/MSRAction3DSkeleton(20joints) loaded

After reviewing all of the papers, building a functions array, and finally visualising one skeletal system image, I clicked the "Attributes Extraction" button in the aforementioned display dataset.

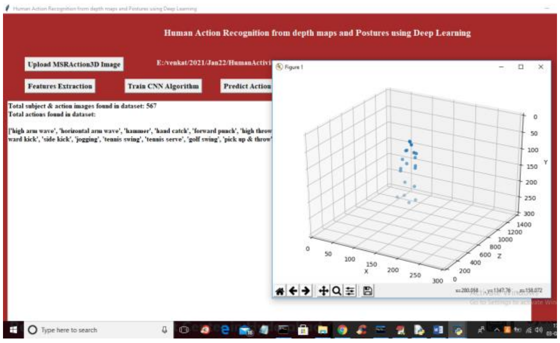


Human Action Recognition from depth maps and Postures using Deep Learning

E:/venkat/2021/Jan22/HumanActivi

Total subject & action images found in dataset: 567
Total actions found in dataset:

['high arm wave', 'horizontal arm wave', 'hammer', 'hand catch', 'forward punch', 'high throw
ward kick', 'side kick', 'jogging', 'tennis swing', 'tennis serve', 'golf swing', 'pick up & throw'

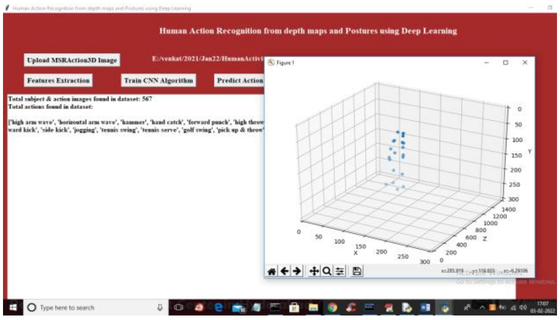


Human Action Recognition from depth maps and Postures using Deep Learning

E:/venkat/2021/Jan22/HumanActivi

Total subject & action images found in dataset: 567
Total actions found in dataset:

['high arm wave', 'horizontal arm wave', 'hammer', 'hand catch', 'forward punch', 'high throw
ward kick', 'side kick', 'jogging', 'tennis swing', 'tennis serve', 'golf swing', 'pick up & throw'



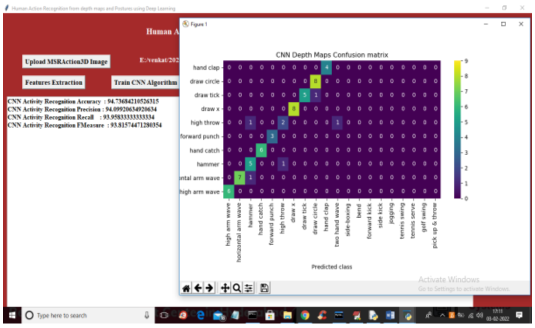
Human Action Recognition from depth maps and Postures using Deep Learning

E:/venkat/2021/Jan22/HumanActivi

Total subject & action images found in dataset: 567
Total actions found in dataset:

['high arm wave', 'horizontal arm wave', 'hammer', 'hand catch', 'forward punch', 'high throw
ward kick', 'side kick', 'jogging', 'tennis swing', 'tennis serve', 'golf swing', 'pick up & throw'

The dataset contains 567 documents, and I've shown 20 different motions, such "high arm wave" and "horizontal arm wave," on the screen above. The skeletal system is moving in the above graphic, which represents the activity of a human in the dataset. You may see the skeletal system activity on the graph after you post. After you've examined the graph, click the "Train CNN Algorithm" button to begin training the CNN and get the results shown below.



E:/venkat/202

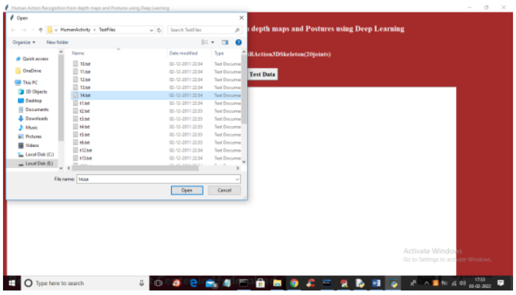
CNN Activity Recognition Accuracy : 94.73684210526315
CNN Activity Recognition Precision : 94.09920634920634
CNN Activity Recognition Recall   : 93.95833333333334
CNN Activity Recognition FMeasure : 93.81574471280354

CNN Depth Maps Confusion matrix

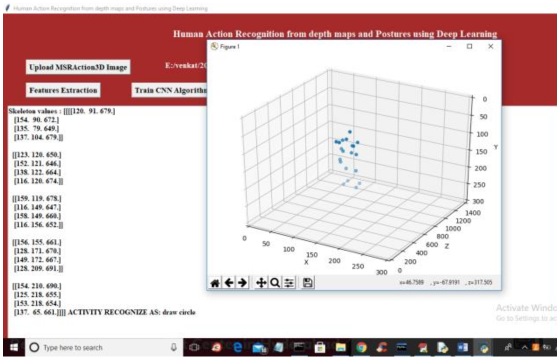
	high arm wave	horizontal arm wave	hammer	hand catch	forward punch	high throw	draw x	draw tick	draw circle	hand clap	two hand wave	side-boxing	bend	forward kick	side kick	jogging	tennis swing	tennis serve	golf swing	pick up & throw
hand clap	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0
draw circle	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0
draw tick	0	0	0	0	0	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0
draw x	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
high throw	0	0	1	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0
forward punch	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
hand catch	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
hammer	0	0	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ntal arm wave	0	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
high arm wave	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

In the above display, we can see that CNN achieved an action recognition accuracy of 94%. The x-axis represents expected activity classes, and the y-axis represents initial classes. All of the class prediction values shown in the diagonal boxes are correct predictions, and very few values are out of the diagonal, indicating that CNN is very efficient and also achieves a 94% precision. Now closed above the chart

To upload test data documents, click the "Predict Activity from Examination Information" button. CNN will then acknowledge activity based on that test file information.

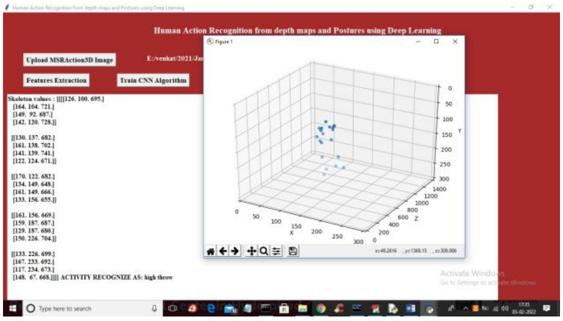


You may receive the following activity acknowledgment result by selecting the "14. txt" file in the previous screen, clicking the "Open" button, and then packing the examination data.


Human Action Recognition from depth maps and Postures using Deep Learning
Human Action Recognition from depth maps and Postures using Deep Learning
Upload MSRAction3D Image
E:/venkat/2
Features Extraction
Train CNN Algorithm
Figure 1
0
50
100
150
200
250
300
X
Y
Z
400
600
800
1000
1200
1400
x=46.7589 , y=-67.9191 , z=317.505
Activate Windows
Go to Settings to ac
Type here to search

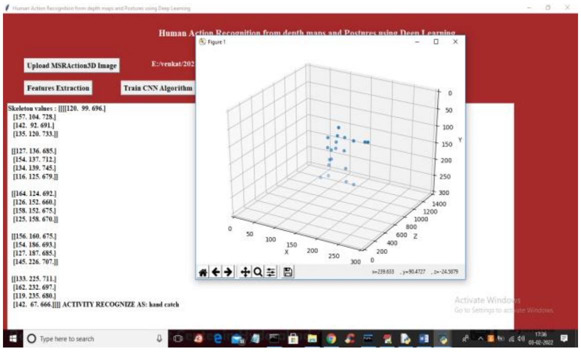
```
Skeleton values : [[[[120.  91. 679.]
  [154.  90. 672.]
  [135.  79. 649.]
  [137. 104. 679.]]

 [[123. 120. 650.]
  [152. 121. 646.]
  [138. 122. 664.]
  [116. 120. 674.]]

 [[159. 119. 678.]
  [116. 149. 647.]
  [158. 149. 660.]
  [116. 156. 652.]]

 [[156. 155. 661.]
  [128. 171. 670.]
  [149. 172. 667.]
  [128. 209. 691.]]

 [[154. 210. 690.]
  [125. 218. 655.]
  [153. 218. 654.]
  [137.  65. 661.]]]] ACTIVITY RECOGNIZE AS: draw circle
```

All of the values in square brackets in the previous presentation are skeleton values; the outcome is "Activity Acknowledged as 'draw circle'" on the final line, and the skeleton's activity is seen in the chart.


Human Action Recognition from depth maps and Postures using Deep Learning
Human Action Recognition from depth maps and Postures using Deep Learning
Upload MSRAction3D Image
E:/venkat/2021/Ja
Features Extraction
Train CNN Algorithm
Figure 1
0
50
100
150
200
250
300
X
Y
Z
400
600
800
1000
1200
1400
x=46.2616 , y=1369.15 , z=308.986
Activate Windows
Go to Settings to activate Windows.
Type here to search
17:35
03-02-2022

```
Skeleton values : [[[[126. 100. 695.]
  [164. 104. 721.]
  [149.  92. 687.]
  [142. 120. 728.]]

 [[130. 137. 682.]
  [161. 138. 702.]
  [141. 139. 741.]
  [122. 124. 671.]]

 [[170. 122. 682.]
  [134. 149. 648.]
  [161. 149. 666.]
  [133. 156. 655.]]

 [[161. 156. 669.]
  [159. 187. 687.]
  [129. 187. 680.]
  [150. 226. 704.]]

 [[133. 226. 699.]
  [167. 233. 692.]
  [117. 234. 673.]
  [148.  67. 668.]]]] ACTIVITY RECOGNIZE AS: high throw
```

In above screen, action is recognized as ‘high throw’


Human Action Recognition from depth maps and Postures using Deep Learning
Human Action Recognition from depth maps and Postures using Deep Learning
Upload MSRAction3D Image
E:/venkat/202
Features Extraction
Train CNN Algorithm
Figure 1
0
50
100
150
200
250
300
X
Y
Z
400
600
800
1000
1200
1400
x=239.633 , y=90.4727 , z=-24.5879
Activate Windows
Go to Settings to activate Windows.
Type here to search
17:36
03-02-2022

```
Skeleton values : [[[[120.  99. 696.]
  [157. 104. 728.]
  [142.  92. 691.]
  [135. 120. 733.]]

 [[127. 136. 685.]
  [154. 137. 712.]
  [134. 139. 745.]
  [116. 125. 679.]]

 [[164. 124. 692.]
  [126. 152. 660.]
  [158. 152. 675.]
  [125. 158. 670.]]

 [[156. 160. 675.]
  [154. 186. 693.]
  [127. 187. 685.]
  [145. 226. 707.]]

 [[133. 225. 711.]
  [162. 232. 697.]
  [119. 235. 680.]
  [142.  67. 666.]]]] ACTIVITY RECOGNIZE AS: hand catch
```

The aforementioned on-screen gesture is called a "hand catch," and it works just like any other file upload and testing tool.

CONCLUSION

It has been proposed to use deep convolutional neural networks to recognise human actions based on depth maps and posture data. By combining the outputs of the three convolutional neural network (CNN) networks, we were able to optimise attribute extraction utilising two activity representations and three channels from convolutional semantic networks. The method has been tested on three publicly available, industry-standard datasets. When compared to state-of-the-art methods that rely on deepness or posture data, the three datasets' category accuracy is light years ahead. The claim made in this work is that various representations of actions provide different clues. Unlike the other depictions, one of them has action functions.

ACKNOWLEDGMENT

We thank CMR Technical Campus for supporting this paper titled "**HUMAN ACTION RECOGNITION FROM DEPTH MAPS AND POSTURES USING DEEP LEARNING**", which provided good facilities and support to accomplish our work. I sincerely thank our Chairman, Director, Deans, Head of the Department, Department Of Computer Science and Engineering, Guide and Teaching and Non- Teaching faculty members for giving valuable suggestions and guidance in every aspect of our work.

REFERANCES

[1] C. Schuldt, I. Laptev, and B. Caputo, "Recognizing human actions: A local SVM approach," in Proc. IEEE 17th Int. Conf. Pattern Recognit., vol. 3. Cambridge, U.K., Aug. 2004, pp. 32–36.

[2] J. Sun, X. Wu, S. Yan, L.-F. Cheong, T.-S. Chua, and J. Li, "Hierarchical spatio-temporal context modeling for action recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Miami, FL, USA, Jun. 2009, pp. 2004–2011.

[3] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, "Learning realistic human actions from movies," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Anchorage, AK, USA, Jun. 2008, pp. 1–8.

[4] W. Li, Z. Zhang, and Z. Liu, "Action recognition based on a bag of 3D points," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, San Francisco, CA, USA, Jun. 2010, pp. 9–14.

[5] J. Wang, Z. Liu, Y. Wu, and J. Yuan, "Mining actionlet ensemble for action recognition with depth cameras," in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pp. 1290–1297, IEEE, 2012.

[6] Q. Ke, M. Bennamoun, S. An, F. Sohel, and F. Boussaid, "Skeleton optical spectra based action recognition using convolutional neural networks," arXiv preprint arXiv:1703.03492, 2016.

[7] Liu, An-An; Su, Yu-Ting; Jia, Ping-Ping; Gao, Zan; Hao, Tong; Yang, Zhao-Xuan (2015). Multipe/Single-View Human Action Recognition via Part-Induced Multitask Structural Learning. IEEE Transactions on Cybernetics, 45(6), 1194–1208. doi:10.1109/tcyb.2014.2347057

[8] Ijjina, Earnest Paul; Mohan, C. Krishna (2014). [IEEE 2014 13th International Conference on Machine Learning and Applications (ICMLA) - Detroit, MI, USA (2014.12.3-2014.12.6)] 2014 13th International Conference on Machine Learning and Applications - Human Action Recognition Based on Recognition of Linear Patterns in Action Bank Features Using Convolutional Neural Networks. , (), 178–182. doi:10.1109/icmla.2014.33

[9] Miki, Hiroshi; Kojima, Atsuhiro; Kise, Koichi (2008). [IEEE 2008 Second International Conference on Future Generation Communication and Networking (FGCN) - Hainan, China (2008.12.13-2008.12.15)] 2008 Second International Conference on Future Generation Communication and Networking - Environment Recognition Based on Human Actions Using Probability Networks. , (), 441–446. doi:10.1109/fgcn.2008.62.

[10] Y. Kim, J. Chen, M.-C. Chang, X. Wang, E. M. Provost, and S. Lyu, “Modeling transition patterns between events for temporal human action segmentation and classification,” in Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on, vol. 1. IEEE, 2015, pp. 1–8.

[11] C. Chen, R. Jafari, and N. Kehtarnavaz, “Action recognition from depth sequences using depth motion mapsbased local binary patterns,” in Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on. IEEE, 2015, pp. 1092–1099.

[12] J. Koushik, “Understanding convolutional neural networks,” arXiv preprint arXiv:1605.09081, 2016.

[13] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, and G. Wang, “Recent advances in convolutional neural networks,” arXiv preprint arXiv:1512.07108, 2015.

[14] E. Park, X. Han, T. L. Berg, and A. C. Berg, “Combining multiple sources of knowledge in deep cnns for action recognition,” in Applications of Computer Vision (WACV), 2016 IEEE

Winter Conference on. IEEE, 2016, pp. 1–8.

[15] J. Zhang, W. Li, P. O. Ogunbona, P. Wang, and C. Tang, "Rgb-d-based action recognition datasets: A survey," Pattern Recognition, vol. 60, pp. 86–105, 2016.