iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43011

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Human Activity Recognition for Surveillance

Sakshi Rajput[1], Sachin Pande[2], Vibha Marda[3], Chetna Chandramore[4], Vaishnavi Deokate[5]
[1, 2, 3, 4, 5]*Department of Information Technology, Savitribai Phule Pune University*

Abstract: ***Human activity recognition is important in human-to-human contact and interpersonal relationships. One of the key objects of research in the scientific fields of computer vision and machine learning is the human ability to identify another person's activity. With the introduction of tiny sensor technologies that can be worn on the body, it is now possible to gather and retain data on various aspects of human mobility under free living settings. This technique has the potential to be employed in automated activity profiling systems that generate a continuous record of activity patterns over time. These activity profiling systems rely on classification algorithms to properly interpret body-worn sensor data and identify various activities. This article examines the many strategies used to classify normal activities and/or identify falls using body-worn sensor data. The study is organized according to the many analytical methodologies and highlights the wide range of approaches that have previously been used in this sector. Although tremendous progress has been achieved in this critical field, there is still much room for improvement, particularly in the application of sophisticated classification approaches to situations requiring a wide range of activities.***

Keywords: Deep Learning, Classification, Detecting Human activity, Image Processing, Feature Extraction, CNN, LR ect.

I. INTRODUCTION

Today physical protection is similarly crucial as cybersecurity, and it keeps you protected from theft, vandalism, burglary, etc. Surveillance and protection plan combines the best of both technology and specialized hardware resulting in developing a close connection. Because of the increasing sophistication of physical security via technology such as artificial intelligence (AI) and the Internet of things (IoT), IT and surveillance are becoming more inextricably linked, and as a result, security teams must collaborate to stabilize both physical and virtual assets. Video analytics technology effectively allows surveillance software to reason and detect irregular behavior and identify suspicious or unauthorized activity that might be ignored by a person. Video analytics operated by AI helps lessen the time spent on surveillance, allowing security operators to be more effective and successful in their work.

Some of the notable and impact creating projects powered by artificial intelligence in security are: -

1) Vehicle bomb detection on the underside
2) Home Security
3) Cameras for crime prevention
4) Large-scale event threat scanning
5) Border Control Lie Detector

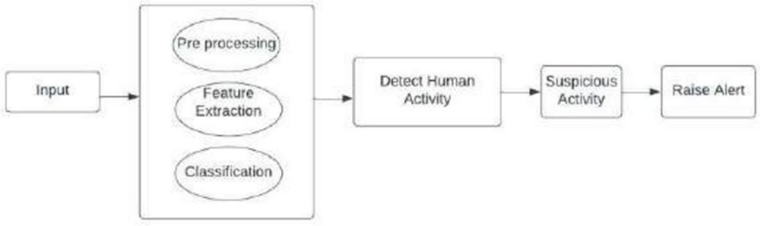


Fig.01. Data Flow Diagram

So, we designed a desktop application which can be used for detecting suspicious activity in a specific premise and then send an alert to the security personnel, which would further help the security personnel to specially check that individual and if found guilty, then take the action accordingly. Example, in some areas certain activities are prohibited like civilians are not allowed to walk on railway tracks, students are not allowed to use their phones at certain premises and many more We have used a media pipe holistic approach to draw landmarks and create datasets. If that comparison of live action and data set action crosses the probability of 0.7, then it would send an alert to the administrator and security head.

©IJRASET: All Rights are Reserved

II. LITERATURE SURVEY

1) Paper name:-Conditional regression forests for human pose estimation, in *Proceeding/CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2012)

Author:- M. Sun, P. Kohli, J. Shotton

Abstract:- Human activity recognition (HAR) has grown in prominence in recent years due to the embedded sensors in smartphones, with applications in healthcare, surveillance, human-device interactions, and pattern identification. An activity-driven hand-crafted neural network model for recognising human activities is presented in this study. Selecting meaningful features from the provided time and frequency domain characteristics is made easier with the help of an algorithm developed using neighborhood component analysis. Afterward, a four-layer deep neural network is utilized to classify the input data into several groups.

2) Paper name:- Observing human-object interactions: using spatial and functional compatibility for recognition. IEEE Trans

Author :-A. Gupta, A. Kembhavi, L.S. Davis

Abstract:- The interpretation of photos and movies involving humans interacting with diverse objects is a terrifying endeavor. Understanding a scene or event entails examining human movements, recognising manipulable things, and seeing the influence of human motion on those items. While each of these perceptual requirements may be fulfilled alone, reputation charge improves when relationships between them are taken into account. Motivated by psychological studies on human perception, we offer a Bayesian technique for analyzing human-object interactions that incorporates several perceptual duties. Previous approaches to item and motion recognition rely on static form or appearance characteristic matching and, correspondingly, movement analysis. Our strategy goes beyond standard tactics by imposing spatial and practical limits on all perceptual elements in order to achieve cohesive semantic interpretation.

3) Paper name:- Learning spatio-temporal representations for action recognition: a genetic programming approach. IEEE Trans.

Author:- L. Liu, L. Shao, X. Li, K. Lu

Abstract:- The first and most important stage in human movement identification is to extract discriminative and strong features from video sequences. In this study, rather of using hand-crafted algorithms, we mechanically investigate spatio-temporal movement aspects for movement repute. This is accomplished by an evolutionary method known as genetic programming (GP), which develops the movement function descriptor on a population of basic three-dimensional operators (e.g., three-D-Gabor and wavelet).We intend to research data adaptive descriptors for unique datasets with more than one layer, which makes complete use of the knowledge to imitate the physical shape of the human visible cortex for movement recognition and concurrently reduce the GP looking area to successfully boost up the convergence of best solutions. In our evolutionary architecture, the common cross-validation classification error, that's calculated through an support-vector-device classifier at the education set, is followed because of the assessment criterion for the

GP fitness function. After the whole evolution manner finishes, the best-so-always answer decided on through GP appears because the (near-)best movement descriptor is obtained.

4) Paper name:- Real-Time Human Pose Detection and Recognition Using MediaPipe

Author:- Vedant Arvind Kumbhare & K. Arthi

Abstract:-Significance of human motion recognition has improved manifolds because of its wide-scale software in the subject of public security, gaming, etc., because of the creation of diverse new technologies. We suggest a framework that detects human motion under unique situations and viewing angles that permit q the identification of divergent styles based on one of a kind spatiotemporal trajectories. In this paper, we use new era consisting of MediaPipe Holistic which gives pose, face, and hand landmark detection models which parses the frames acquired via real-time tool feed the use of OpenCV via our MediaPipe Holistic version and offer a complete of 501 landmarks that's exported as coordinates to a CSV document upon which we educate a custom multi-elegance class version to recognize the connection among the elegance and coordinates to categorize and locate custom frame language pose. The system learning classification algorithms carried out on this paper are random forest, linear regression, ridge classifier, and gradient boosting classifier.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

III. SYSTEM ARCHITECTURE

The main objective of this project is to make a desktop application that will identify an anomalous activity by a human being. The flow will start with video preprocessing using Convolutional neural networks (CNN) for suspicious object detection and LR for suspicious activity detection. After this we will detect the movement in the premise, then categorize if there is any suspicious activity done by an individual or if he/she is carrying a suspicious i.e threat causing object. Application will track the activity made by that human, and if it is suspicious then an alert message will be generated for that activity. For example:- if a person is kicking or punching in a premise then the application will generate an alert message. We employed a mediapipe holistic approach that includes optimized face, hands, and posture components that enable for holistic tracking, allowing the model to recognise hand and body positions as well as facial landmarks. Creating a dataset utilizing these locations, taking into account various situations for various activities such as standing, kick, punch, squat, sitting, and suspicious things such as knife, scissors. Applying different testing algorithms such as Logistic regression, Random forest, Ridge regression, gradient boosting for better results. We have considered around 500 possibilities for each action and stored their landmarks in the dataset. If the probability of the activity detected is greater than 70% then it will email to the admin for that we have used smtp.

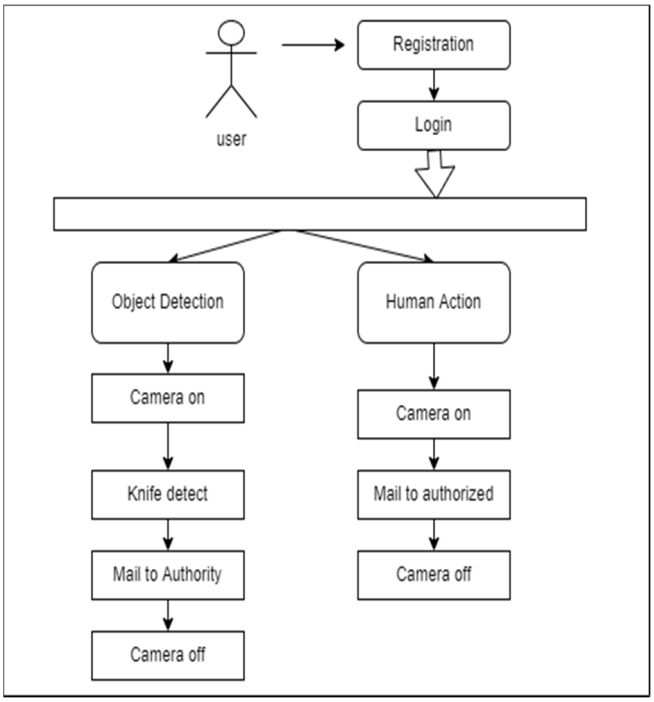


Fig.02. Architectural Diagram

A. *Pre-processing Data*

1) Obtain the dataset: The images in your camera feed may be of lower quality.Thus, it's essential to train the model to work in such conditions. A very elegant way of doing that is by performing a mediapipe holistic approach.

2) Formatting: Once the data set has been cleansed; it needs to be formatted.

3) Data preprocessing is a data mining technique which is used to transform the raw data in a useful and efficient format. Major tasks in data pre processing are-

a) Data cleaning - process to remove incorrect data, inaccurate data from the datasets, and it also replaces the missing values.

b) Techniques used: - Regression, Computing missing values

c) Data integration- process of combining multiple sources into a single dataset.

d) Techniques used: - Schema integration

e) Data reduction - process helps in the reduction of the volume of the data

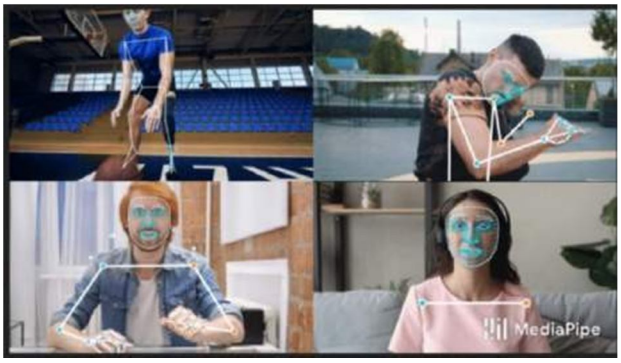
f) Techniques used: - Data compression

g) Data transformation - change made in the format or the structure of the data

©IJRASET: All Rights are Reserved

B. Mediapipe Holistic Pipeline

Simultaneous real-time recognition of human poses, facial features, and hand tracking on mobile devices enables a variety of influential applications, including: B. Fitness and sports analysis, gesture control and sign language recognition, augmented reality effects, etc. MediaPipe is an open source framework specifically designed for complex perceptual pipelines and designed to take advantage of accelerated inference (such as GPUs and CPUs) and is fast for these tasks. We already offer an accurate yet individualized solution.



MediaPipe Holistic consists of a new pipeline with optimized pose, face and hand components, each running in real time, minimizing memory transfers between inference backends, quality and speed. Further support for compatibility of the three components, depending on the trade-off. When all three components are included, MediaPipe Holistic provides a unified topology with over 540 breakthrough key points (33 poses, 21 manual work, 468 facial features).

C. Models

1) *Landmark Model:* MediaPipe Holistic uses MediaPipe Pose, MediaPipe Face Mesh, MediaPipe Hands pose, face and hand landmark models, respectively, for a total of 543 landmarks (33 pose landmarks, 468 face landmarks, 21). Generates a hand landmark).
2) *Hand Recrop Model:* If the pose model is not accurate enough and the resulting hand ROI is still inaccurate, run an additional light hand retrimming model. It acts as a spatial transformer and accounts for only about 10% of the inference time. The cost of the hand model.

IV. SIMPLE MAIL TRANSFER PROTOCOL(SMTP) CONFIGURATION

SMTP: Sending email messages between servers. SMTP is typically used to send email messages from email From client to mail server. The SMTP design architecture is shown. As shown in the figure below:

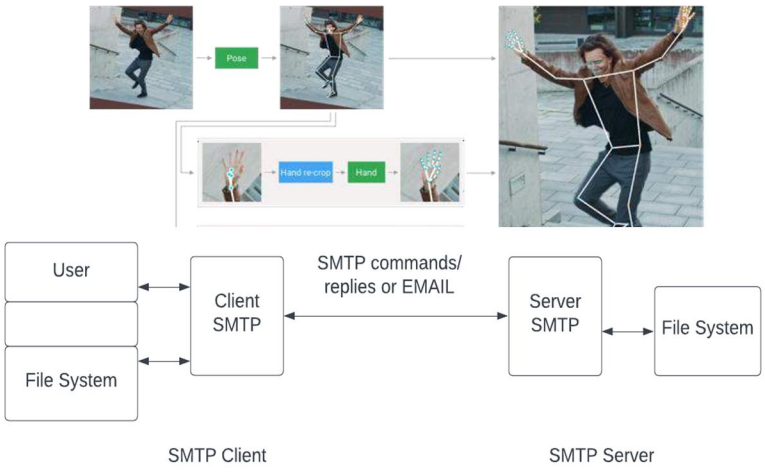


Fig. 03 SMTP Architecture

If there is a message to send to the SMTP client, the SMTP client will set up a Bidirectional transmission channel to the SMTP server. Or The task of the SMTP client is to send an email message.Send to one or more SMTP servers or report a failure. Here, I sent an email using my Gmail account.

Or

The settings of the Gmail SMTP server for sending mail are shown below.

§ Gmail SMTP server address: smtp.gmail.com
§ Gmail SMTP Username: Full Gmail Address
§ Gmail SMTP password: Gmail password
§ Gmail SMTP port: 465
§ Requires GmailSMTPSSL (Secure Sockets Layer): Yes
§ Gmail SMTP authentication required: Yes

V. ALGORITHM

A. Logistic Regression

It is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables. Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, True or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.

Logistic Regression is much similar to Linear Regression except that how they are used. Linear Regression is used for solving Regression problems, whereas Logistic regression is used for solving the classification problems. In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1)

B. Random Forest

It is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

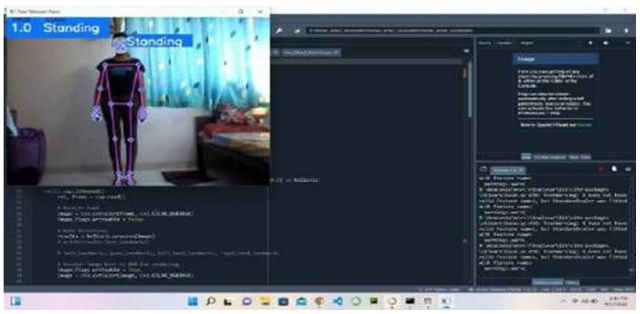
C. Ridge Regression

It is a model tuning method that is used to analyze any data that suffers from multicollinearity. This method performs L2 regularization. When the issue of multicollinearity occurs, least-squares are unbiased, and variances are large, this results in predicted values being far away from the actual values. Ridge regression is the method used for the analysis of multicollinearity in multiple regression data. It is most suitable when a data set contains a higher number of predictor variables than the number of observations. The second-best scenario is when multicollinearity is experienced in a set.

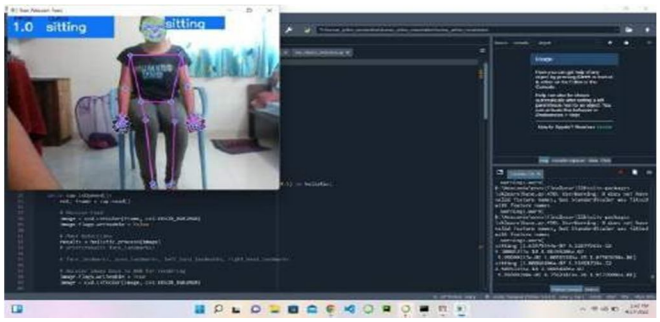
VI. RESULT

A. Detection of Human Activity

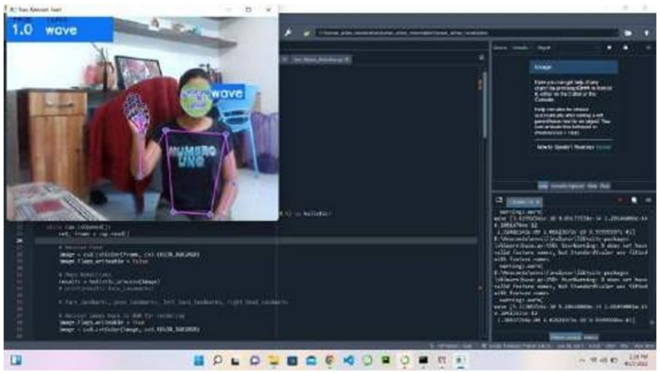
1. Standing



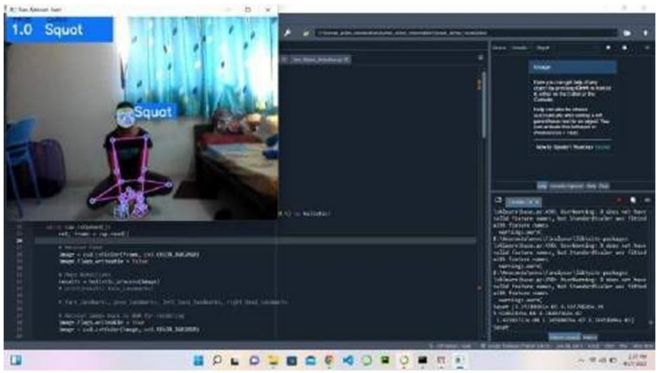
2. Sitting



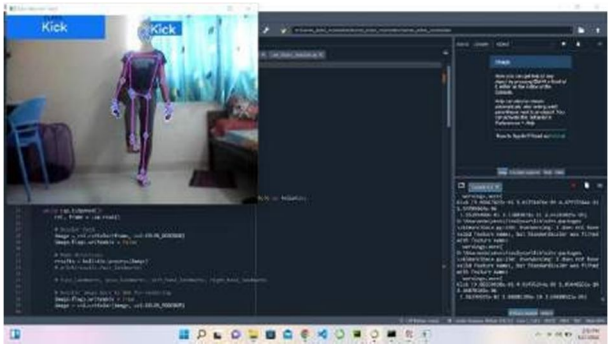
3. WAVE



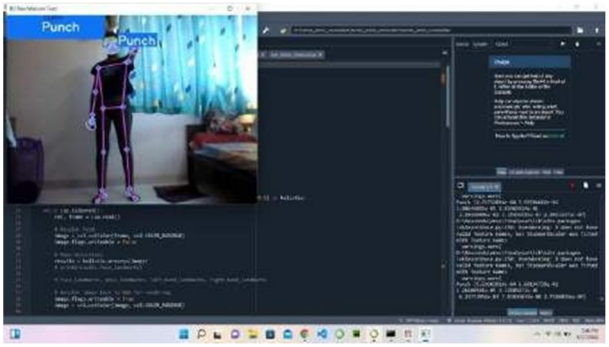
4. SQUAT



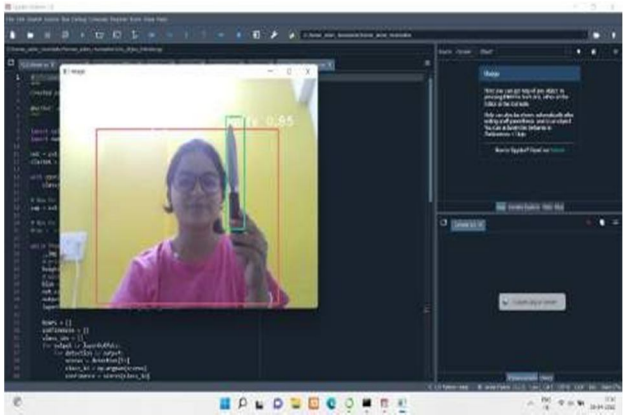
5. Kick



6. Punch



7. KNIFE DETECTION



8. EMAIL NOTIFICATION



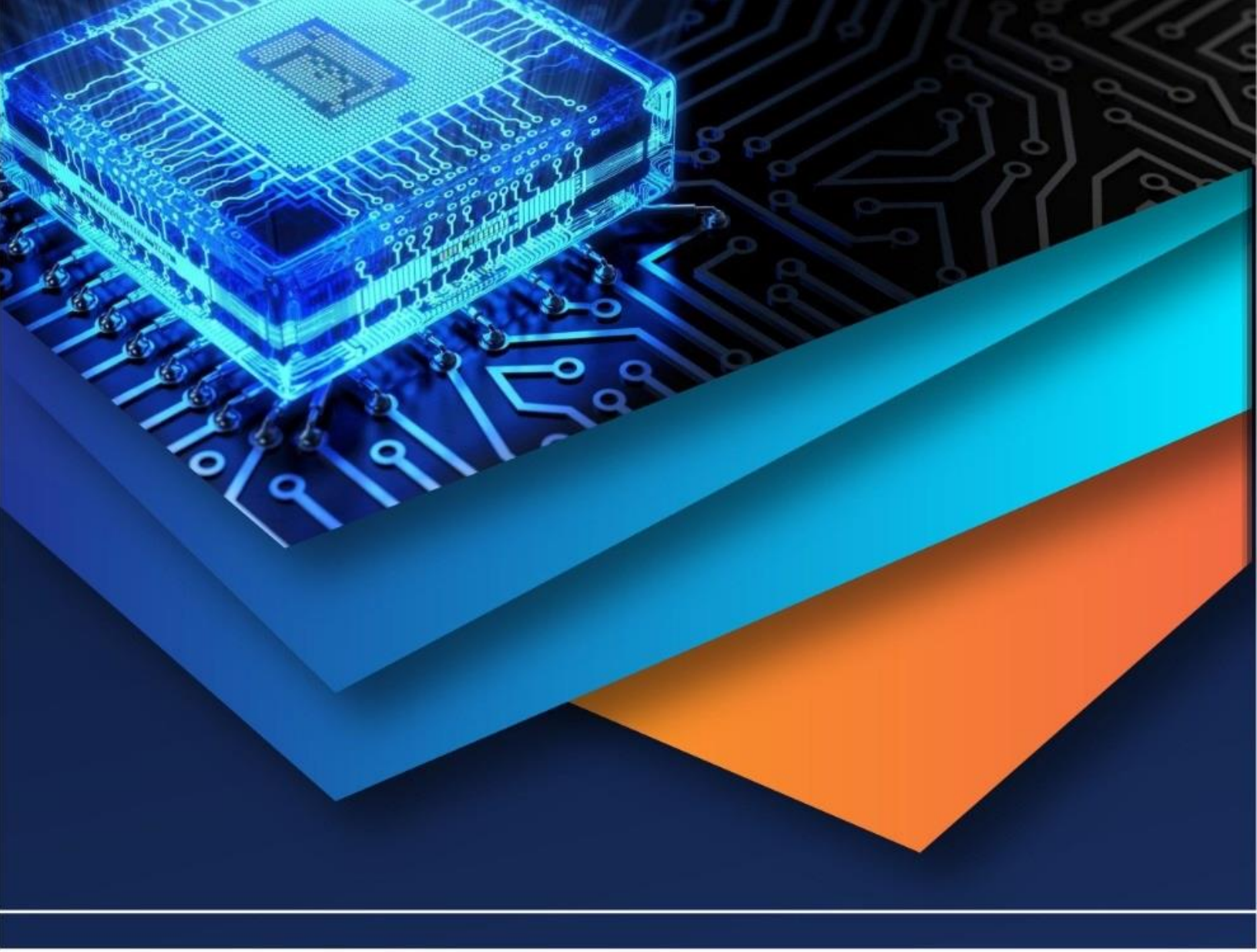
©IJRASET: All Rights are Reserved

VII. CONCLUSION

Human posture estimation is a hot study subject in computer vision that has recently emerged alongside the rise of deep learning. Early networks are generally shallow, employed in a fairly easy manner, and can only handle tiny pictures or patches because of limits in hardware device capacity and the quantity and quality of training data (Toshev and Szegedy, 2014; Tompson et al., 2015; Li and Chan, 2014). More current networks are stronger, deeper, and more efficient (Newell et al., 2016; Cao et al., 2016; He et al., 2017; Sun et al., 2019).We reviewed recent deep learning-based research on the 2D/3D human pose estimation problem from monocular images or video footage and classified approaches into four categories based on specific tasks: (1) 2D single person pose estimation, (2) 2D multi-person pose estimation, (3) 3D single person pose estimation, and (4) 3D multi-person pose estimation. In addition, we described the most common human pose datasets and evaluation methodologies. Despite significant progress in monocular human pose estimation using deep learning, certain unresolved issues and gaps between research and practical applications remain, such as the effect of body part occlusion and crowded people. The most critical prerequisites for deep learning-based techniques are efficient networks and appropriate training data. Future networks should investigate both global and local settings for more discriminative human body traits, while also including human body structures into the network for previous limitations. Some useful network design strategies, such as multi-stage structure, intermediate supervision, multi- scale feature fusion, multi-task learning, and body structural limitations, have been confirmed in current networks. Network efficiency is also a significant consideration when implementing algorithms in real-world applications.

REFERENCES

[1] F. Adib, Z. Kabelac, D. Katabi, and RC. Miller, “3d tracking via body radio reflections,” 11th USENIX Symposiumon
[2] Networked Systems Design and Imple- mentation (NSDI’14), Vol. 14, pp.317-329, Apr. 2014.
[3] Sasakawa, N. Honma, T. Nakayama, and S. Iizuka, “Human activity estima- tion by height and RCS information detected by MIMO radar,” 2017 IEEE AP-S Symposium on Antennas and Propagation and USNCURSI Radio Science Meeting, in press.
[4] T. Asano, A. Miyata, and S. Honda, J. JSPE, vol. 77, no. 3, pp. 333-337, Sep. 2011.
[5] D. Sasakawa, N. Honma, T. Nakayama, and S. Iizuka, “Fast living-body local- ization algorithm using MIMO radar in a multi-path environment,” unpublished.
[6] R. Schmidt, “Multiple emitter location and signal parameter estimation,” IEEE Transactions on Antennas and Propagation, Vol. 34, Issue 3, pp. 276 - 280, Mar. 1986.
[7] Chen L, Nugent C D, Wang H. A Knowledge-Driven Approach to Activity Recognition in Smart Homes[J]. IEEE Transactions on Knowledge Data Engineering, 2012, 24(6):961-974.
[8] R. Schmidt, “Multiple emitter location and signal parameter estimation,” IEEE Transactions on Antennas andPropagation,
[9] Vol. 34, Issue 3, pp. 276 - 280, Mar. 1986.
[10] Y. Chen and C. Shen, ”Performance Analysis of SmartphoneSensor Behavior for Human Activity Recognition,”in
[11] IEEE Access, vol. 5, pp. 3095-3110, 2017.
[12] Chen C, Jafari R, Kehtarnavaz N. Improving Human Action Recognition Us- ing Fusion of Depth Camera and Inertial Sensors[J]. IEEE Transactions on Human-Machine Systems, 2015, 45(1):51-61
[13] Chen D, Yang J, Wactlar H D. Towards automatic analysis of social interaction patterns in a nursing home environmen from video[C]



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)