



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 3, March 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Virtual Physical Trainer Using Machine Learning

P. Tarun Raj¹, C.V. Sanjay², S. Santhan³, Dr D. Venkata Subramanian⁴

U.G Students, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai,
Tamil Nadu, India^{1,2,3}

Adjunct Professor, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai,
Tamil Nadu, India⁴

ABSTRACT: The world is adapting to the new normal of living remotely. Though it offers safety, it has its own disadvantages, especially physical and mental health concerns. Regularly exercising at home is pivotal for a balanced healthy life. Many people exercise, but due to improper poses, they do not get expected results. It may even lead to potential injuries. We intend to create an automated trainer which can assist the user by providing real-time meaningful feedback regarding the quality of exercise.

KEYWORDS: Machine Learning, Image Classification, TensorFlow, CNN, Deep Learning, Neural Networks, OpenPose.

I. INTRODUCTION

Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

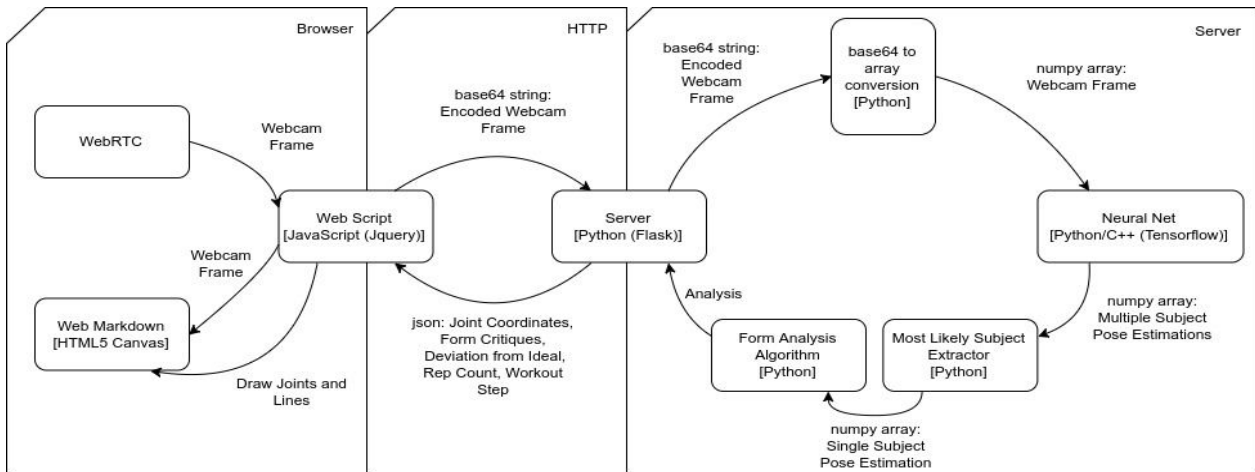
A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers; but not all machine learning is statistical learning.

The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. In its application across business problems, machine learning is also referred to as predictive analytics.

For simple tasks assigned to computers, it is possible to program algorithms telling the machine how to execute all steps required to solve the problem at hand; on the computer's part, no learning is needed.

For more advanced tasks, it can be challenging for a human to manually create the needed algorithms. In practice, it can turn out to be more effective to help the machine develop its own algorithm, rather than having human programmers specify every needed step.

The advent of artificial intelligence and its subsets (computer vision, machine learning, NLP, and more) is modernizing the health and fitness industry at an unprecedented rate. By making fitness machines, gadgets, wearables, and mobile applications smarter, this technology is helping people to stay fit and healthy. Right from helping businesses in this industry in improving their marketing and sales strategies to assisting people to reshape their day-to-day habits, AI is playing a big role in the fitness world.



II. RELATED WORK

The fact that human appearance tends to stay unchanged through videos has been used in the past to aid pose estimation [1, 2, 3]. Ramanan et al. [2] train discriminative body part detectors by first detecting ‘easy’ poses (such as a ‘scissors’ walking pose) and then using the appearance learnt from these poses to track the remaining video with a pictorial structure model. Shen et al. [3] iteratively re-train a pose estimator from confident detections, and also include temporal constraints. In the same spirit as [2], we too initialize from high-precision poses and re-train a discriminative model (a ConvNet pose estimator) – but rather than training a personalized part detector from these poses alone, we first spread the initial annotation throughout the whole video [5] using image matching [4] and optical flow [6, 7] to generate far more annotated frames. Since long video sequences contain an abundance of data, we can simply delete poor pose annotations (rather than trying to correct them); and even if some of the remaining annotations are incorrect, our ConvNet is able to deal with the label noise. Prior work [9] has used evaluator algorithms to remove entire erroneous pose estimates, whereas here we evaluate individual body part annotations. Furthermore, in a similar manner to poselets [14, 15, 16, 17], our matching framework captures dependencies between both connected and non-connected body parts. More generally, the idea of starting from general classifiers, and then ‘personalizing’ them has been used in other areas of computer vision, such as pedestrian detection [8] or object tracking [10]. Kalal et al. [10] proposed a tracking-learning-detection paradigm for learning a set of object templates online, e.g. for vehicles. Typically, in this type of approach, object models are initialized from a single frame [12, 11], and then matched to the next frame before being re-learned. Supancic and Ramanan [13] improve this by proposing to revisit tracked frames and re-learn a model to correct previous errors. In a similar way, our approach utilizes the whole video to learn body part appearance, but differs to previous work in that we match to all frames in a video sequence in one step, and not just frame to frame.

III. PROPOSED SYSTEM

It provides a diverse array of basic exercises for the user. It also allows fitness centres to provide fitness trainer for their clients. It provides user with live visuals and stats during the workout. It provides a seamless user experience.

1. TENSOR FLOW: TensorFlow is an open-source end-to-end platform for creating Machine Learning applications. It is a symbolic math library that uses dataflow and differentiable programming to perform various tasks focused on training and inference of deep neural networks. It allows developers to create machine learning applications using various tools, libraries, and community resources. It uses Python to provide a convenient front-end API for building applications with the framework, while executing those applications in high-performance C++. TensorFlow allows developers to create *dataflow graphs*—structures that describe how data moves through a graph, or a series of processing nodes. Each node in the graph represents a mathematical operation, and each connection or edge between nodes is a multidimensional data array, or *tensor*. TensorFlow provides all of this for the programmer by way of the Python language.

Python is easy to learn and work with and provides convenient ways to express how high-level abstractions can be coupled together. Nodes and tensors in TensorFlow are Python objects, and TensorFlow applications are themselves Python applications. The actual math operations, however, are not performed in Python. The libraries of transformations that are available through TensorFlow are written as high-performance C++ binaries. Python just directs traffic between the pieces and provides high-level programming abstractions to hook them together.

2.OPEN POSE: OpenPose is a Real-time multiple-person detection library, and it's the first time that any library has shown the capability of jointly detecting human body, face, and foot key points. It was proposed by researchers at Carnegie Mellon University. They have released in the form of Python code, C++ implementation and Unity Plugin. These resources can be downloaded from OpenPose repository. In first step the image is passed through baseline CNN network to extract the feature maps of the input in the paper. In this paper these authors used first 10 layers of VGG-19 network. The feature map is then process in a multi-stage CNN pipeline to generate the Part Confidence Maps and Part Affinity Field. In the last step, the Confidence Maps and Part Affinity Fieldsthat are generated above are processed by a greedy bipartite matching algorithm to obtain the poses for each person in the image.

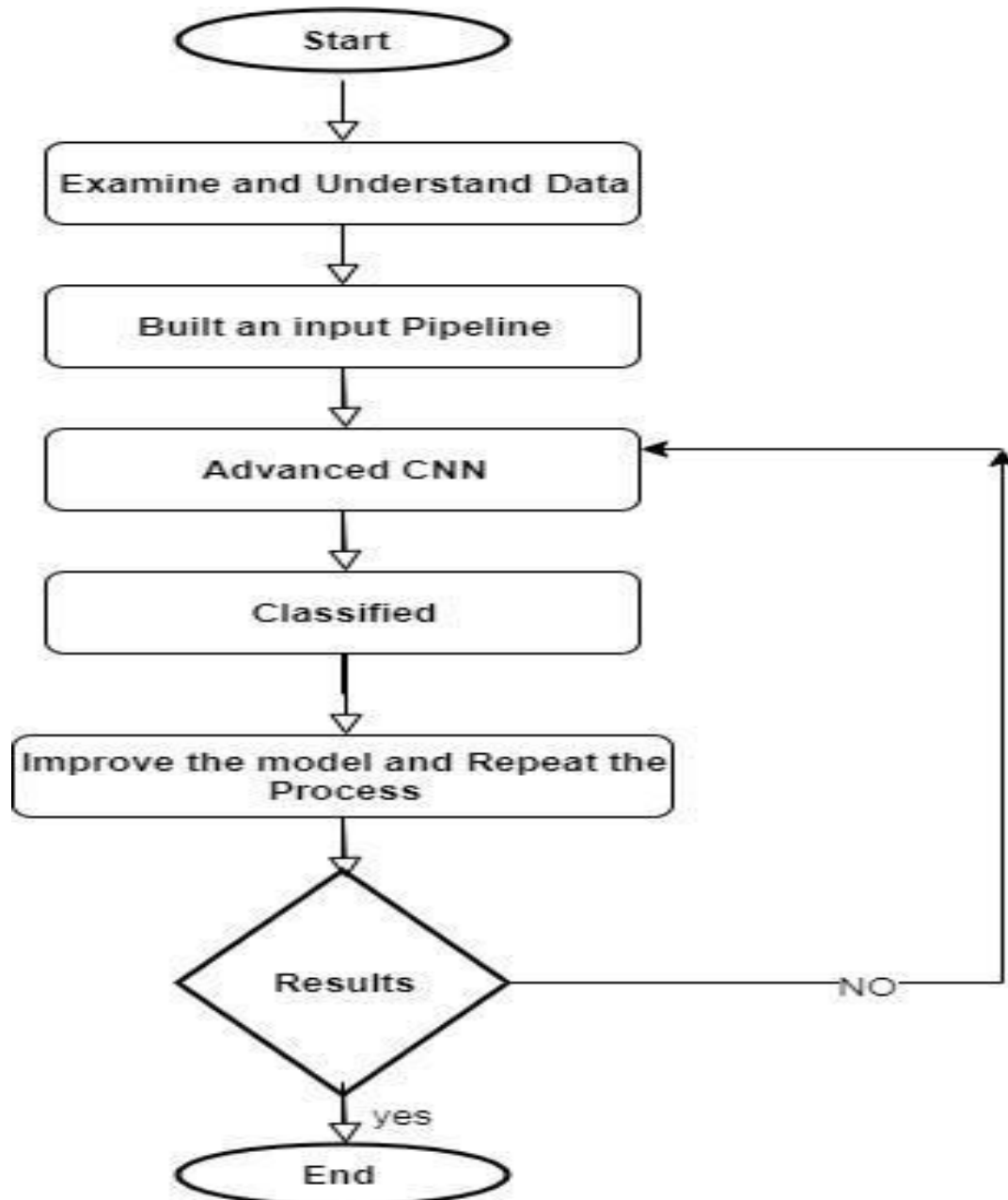


3.ml5.js:ml5.js aims to make machine learning approachable for a broad audience of artists, creative coders, and students. The library provides access to machine learning algorithms and models in the browser, building on top of TensorFlow.js with no other external dependencies. ml5.js is an open source, friendly high level interface to TensorFlow.js, a library for handling GPU-accelerated mathematical operations and memory management for machine learning algorithms. ml5.js provides immediate access in the browser to pre-trained models for detecting human poses, generating text, styling an image with another, composing music, pitch detection, and common English language word relationships, and much more. ml5.js aims to support broader public understanding of machine learning and foster deeper engagement with ethical computing, responsible data collection, and accessibility and diversity of people and perspectives in technology and the arts.

4.IMAGE FEATURE EXTRACTION: Feature extraction is done after the preprocessing phase in character recognition system. The primary task of pattern recognition is to take an input pattern and correctly assign it as one of the possible output classes. This process can be divided into two general stages: Feature selection and Classification. Feature selection is critical to the whole process since the classifier will not be able to recognize from poorly selected features. Criteria to choose features given by Lippman are: "Features should contain information required to distinguish between classes, be insensitive to irrelevant variability in the input, and also be limited in number, to permit, efficient



computation of discriminant functions and to limit the amount of training data required” Feature extraction is an important step in the construction of any pattern classification and aims at the extraction of the relevant information that characterizes each class. In this process relevant features are extracted from objects/ alphabets to form feature vectors. These feature vectors are then used by classifiers to recognize the input unit with target output unit. It becomes easier for the classifier to classify between different classes by looking at these features as it allows fairly easy to distinguish. Feature extraction is the process to retrieve the most important data from the raw data. Feature extraction is finding the set of parameter that define the shape of a character precisely and uniquely. In feature extraction phase, each character is represented by a feature vector, which becomes its identity. The major goal of feature extraction is to extract a set of features, which maximizes the recognition rate with the least amount of elements and to generate similar feature set for variety of instance of the same symbol. The widely used feature extraction methods are Template matching, Deformable templates, Unitary Image transforms, Graph description, Projection Histograms, Contour profiles, Zoning, Geometric moment invariants, Zernike Moments, Spline curve approximation, Fourier descriptors, Gradient feature and Gabor features.





5.CNN: CNN is a deep [neural network](#) originally designed for image analysis. Recently, it was discovered that the CNN also has an excellent capacity in sequent data analysis such as [natural language processing](#). CNN always contains two basic operations, namely convolution and pooling. The [convolution operation](#) using multiple filters is able to extract features (feature map) from the data set, through which their corresponding spatial information can be preserved. The pooling operation, also called subsampling, is used to reduce the [dimensionality](#) of feature maps from the convolution operation. Max pooling and [average pooling](#) are the most common pooling operations used in the CNN.

IV. RESULT

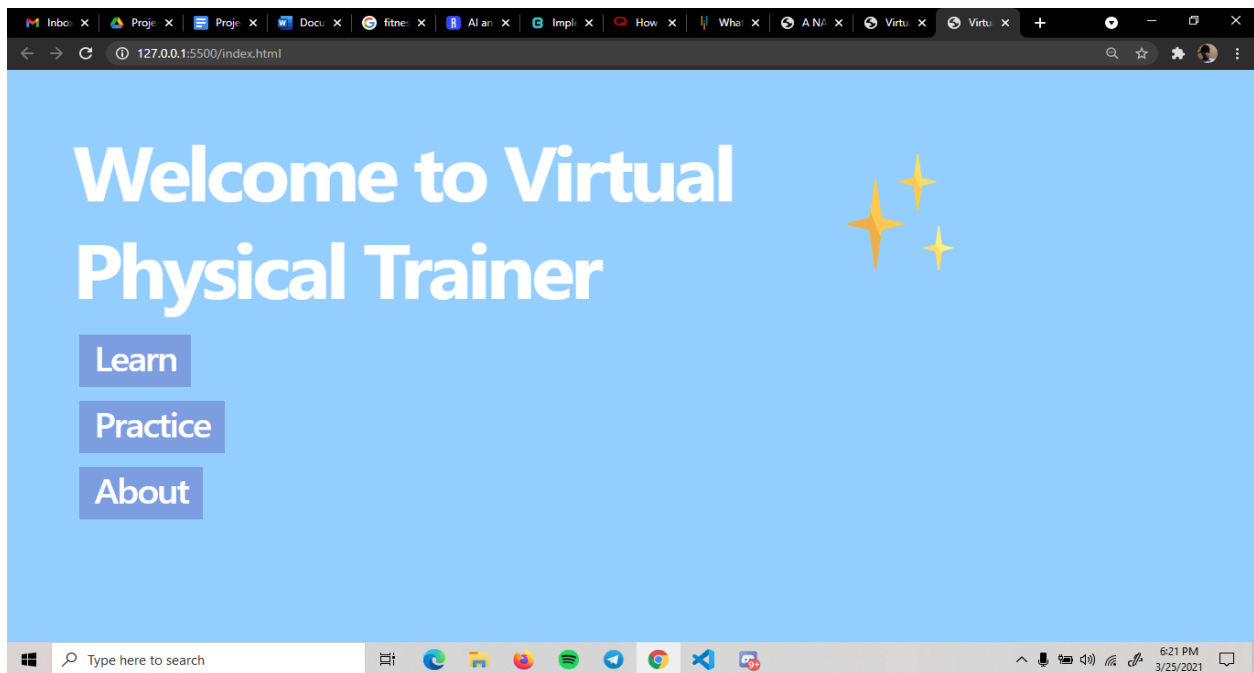


Fig 1. Home Page

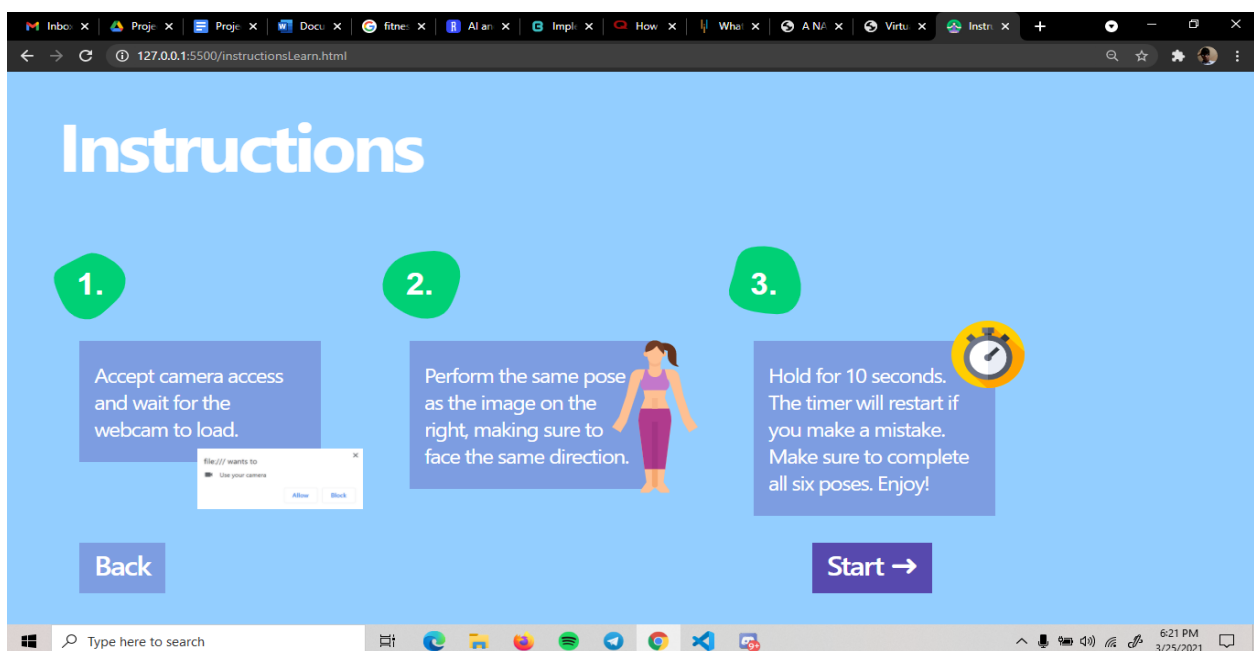


Fig 2. Instruction Page

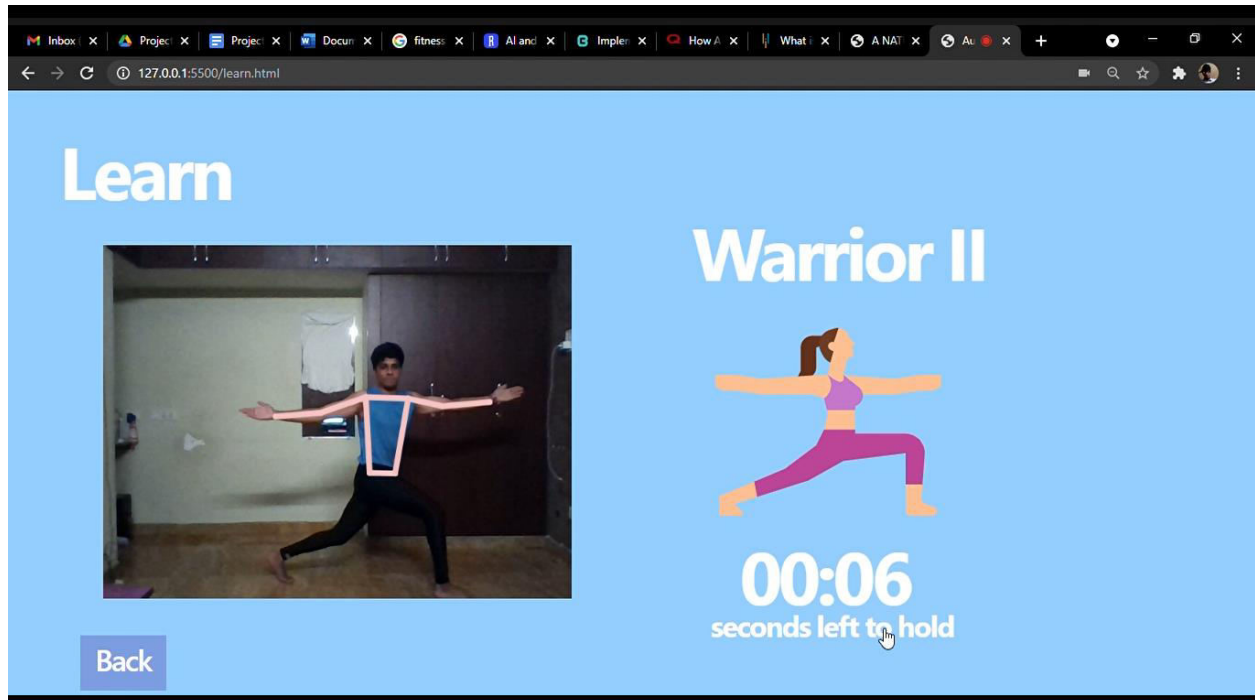


Fig 3. Virtual Training Page

V. CONCLUSION

With the automated physical trainer, one can exercise in the correct position and increase the quality of every movement yielding better results in much less time. It's a less expensive measure, as we no more need a gym coach. This model can prove to be very useful for users exercising from homes, due to the improved convenience and professional guidance.

REFERENCES

- [1] S. Amin, P. Muller, A. Bulling, and M. Andriluka. Test-time " adaptation for 3d human pose estimation. In Proc. GCPR, 2014.
- [2] D. Ramanan, D. Forsyth, and A. Zisserman. Strike a pose: Tracking people by finding stylized poses. In Proc. CVPR, 2005.
- [3] H. Shen, S.-I. Yu, Y. Yang, D. Meng, and A. Hauptmann. Unsupervised video adaptation for parsing human motion. In Proc. ECCV. 2014.
- [4] J. Sullivan and S. Carlsson. Recognizing and tracking human 3071 action. In Proc. ECCV, 2002.
- [5] A. M. Buchanan and A. W. Fitzgibbon. Interactive feature tracking using K-D trees and dynamic programming. In Proc. CVPR, volume 1, pages 626–633, 2006.
- [6] J. Charles, T. Pfister, D. Magee, D. Hogg, and A. Zisserman. Upper body pose estimation with temporal sequential forests. In Proc. BMVC, 2014.
- [7] S. Zuffi, J. Romero, C. Schmid, and M. Black. Estimating human pose with flowing puppets. In Proc. ICCV, 2013.
- [8] K. K. Htike and D. C. Hogg. Efficient non-iterative domain adaptation of pedestrian detectors to video scenes. In Proc. ICPR, 2014.
- [9] N. Jammalamadaka, A. Zisserman, M. Eichner, V. Ferrari, and C. Jawahar. Has my algorithm succeeded? an evaluator for human pose estimators. In Proc. ECCV. 2012.
- [10] Z. Kalal, K. Mikolajczyk, and J. Matas. Tracking-learning-detection. IEEE PAMI, 2012.
- [11] Z. Kalal, J. Matas, and K. Mikolajczyk. Pn learning: Bootstrapping binary classifiers by structural constraints. In Proc. CVPR, 2010.
- [12] S. Duffner and C. Garcia. Pixeltrack: a fast adaptive algorithm for tracking non-rigid objects. In Proc. ICCV, 2013.
- [13] J. S. Supancic and D. Ramanan. Self-paced learning for long-term tracking. In Proc. CVPR, 2013.



- [14] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In Proc. ICCV, 2009.
- [15] G. Gkioxari, P. Arbelaez, L. Bourdev, and J. Malik. Articulated pose estimation using discriminative armlet classifiers. In Proc. CVPR, 2013.
- [16] L. Pishchulin, M. Andriluka, P. Gehler, and B. Schiele. Poselet conditioned pictorial structures. In Proc. CVPR, 2013.
- [17] B. Sapp and B. Taskar. Multimodal decomposable models for human pose estimation. In Proc. CVPR, 2013.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

Impact Factor:
7.512

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details