**HUMAN ACTIVITY RECOGNITION USING GENERATIVE ADVERSERIAL NETWORK**

N. Indira, Amrutha Balumuri, Sana Rushitha Reddy, Usirike Swetha Reddy

+Department of Computer Science and Engineering,

+Panimalar Engineering college, Chennai,

Tamilnadu, India

[Rushithareddy105@gmail.com](mailto:Rushithareddy105@gmail.com)

[Swethareddysr15@gmail.com](mailto:Swethareddysr15@gmail.com)

[amruthasireebalumuri@gmail.com](mailto:amruthasireebalumuri@gmail.com)

*Abstract*—Human pose estimation is an important problem in the field of Computer Vision. Imagine being able to track a person’s every small movement and do a bio-mechanical analysis in real time. The technology will have huge implications. Pose Estimation is predicting the body part or joint positions of a person from an image or a video. Humans are flexible they are able to change their poses frequently. Now-a-days, this world is depending more on automation, so capturing all the activities are made in the surroundings using surveillances. It is difficult for an automatic computer to determine their poses for the analysis process. To analysis the human movement position Generative Adversarial Network (GAN) is used. GAN algorithm could help us to determine the exact pose of the human with highest efficiency rate, since this module uses generator and discriminator, in which the generator generates many reliable related images and the discriminator which tries to determine original and the fake. These two component make a efficient way for determining the exact pose the human and their gesture.

Keywords—Human interaction, Communication, TM, Security

# **1. Introduction**

Human-machine interaction is explained as a world is counting on challenges of computer vision for its automation. We face many problems during the detection of human pose. In advance, self driving car has a system to recognize the human pose which detects through cameras in the front and the rear. It is very difficult for it to recognize the human pose when the car is in motion. Those images get blurred and get poor quality images. In running state this system must be capable to recognize those blurred images and process and tune the images. In cricket, bowling action need to be perfect if not then that particular ball will not be considered to be counted.

The bowler must follow the rule which is the angle between the upper and lower arm during the bowling action as the arm passes above shoulder height, measure again when the ball released, the difference must be no more than 15 degrees. At the point of release the angle is almost always zero degrees. These actions must be processed soon to determine the ball for its validity. For these we need to use computer for faster processing the actions need to be captured using cameras. For a better computer vision for human pose recognition we use Generative Adversarial Network algorithm. The input source for our model is through video surveillances or through any cameras. The captured videos will undergoes following 3 processes.

1.Detection of person on the video feed.

2. Key point generation on the detected human

3. Processing with GAN algorithm module. Using the above three process we can recognize the human pose even with the blurred image

.

# **2.related works**

# Face aging with conditional generative adversarial networks It has been recently shown that generative adversarial networks (gans) can produce synthetic images of exceptional visual fidelity. In this work, the gan-based method for automatic face aging was introduced [1]. Contrary to previous work employing GANs for altering of facial attributes, a particular emphasize on preserving the original person’s identity in the aged version of his/her face was made. A novel approach for “identity-preserving” optimization of gan’s latent vectors is introduced. The objective of evaluation of the resulting aged and rejuvenated face images by the state-of-threat face recognition and age estimation solutions demonstrate the high potential of the proposed method.

“do as i do” motion transfer [2]: A source video is given for a person dancing, that performance is transferred to a novel (amateur) target after only a few minutes of the target subject performing standard moves. In this approach, this problem as video-to-video translation using pose as an intermediate representation. To transfer the motion, extract poses from the source subject and apply the learned pose-to-appearance mapping to generate the target subject. Two consecutive frames are predicted for temporally coherent video results and introduce a separate pipeline for realistic face synthesis. Although this method is quite simple, it produces surprisingly compelling results (see video). This motivated to provide a forensics tool for reliable synthetic content detection, which is able to distinguish videos synthesized by the system from real data. In addition, a first of- its-kind open-source dataset of videos that can be legally used for training and motion transfer was released.

3. Learning to discover cross-domain relations with generative adversarial networks [3]

Humans easily recognize relations between data from different domains without any supervision; learning to automatically discover them is in general very challenging and needs many ground-truth pairs that illustrate the relations. To avoid costly pairing, the task of discovering cross-domain relations given unpaired data is addressed. A method based on generative adversarial networks was developed that learn to discover relations between different domains.

4. Human Pose Estimation

Human pose estimation aims to estimate the spatial configuration of body parts in given images. Most top methods without CNNs are based on the tree structured pictorial structures model. Yang and Ramanan model more complex joint relationships using a flexible mixture model. Sapp and Taskar further propose a multi-modal model that combines both holistic and local cues for mode selection and pose estimation. For more details, refer to the recent benchmark. Recent state-of-the art methods for pose estimation are based on CNNs. Toshev et al. present a cascade of deep neural networks (DNNs)-based pose regressors for pose estimation in a holistic fashion. Tompson et al. combine a CNN and an MRF and train both models jointly for human pose estimation and show state-of-the-art performance on the FLIC and LSP. However, they both need a reliable detector to generate a bounding box. Cao et al. propose a bottom-up approach to efficiently detect the pose of multiple people via introducing a non-parametric representation, named part affinity fields (PAFs), to learn the associations between body parts. It is fast and achieves state-of-the-art accuracies on multiple public benchmarks, but its ability declines when dealing with small persons. Moreover, they are all exacerbated by a limited computation cost or a restricted computing platform. The most related work to our pose estimation is to utilize a dependency graph to represent the relationships between reference points and sequentially estimate by multidimensional output regression forests.

**3. EXISTING SYSTEM**

Optical Character Recognition (OCR) is a type of document image analysis where a scanned digital image that contains either machine printed or handwritten script is input into an OCR software engine and translating it into an editable machine readable digital text format (like ASCII text). Exiting system provide less accuracy and less number of input taken.

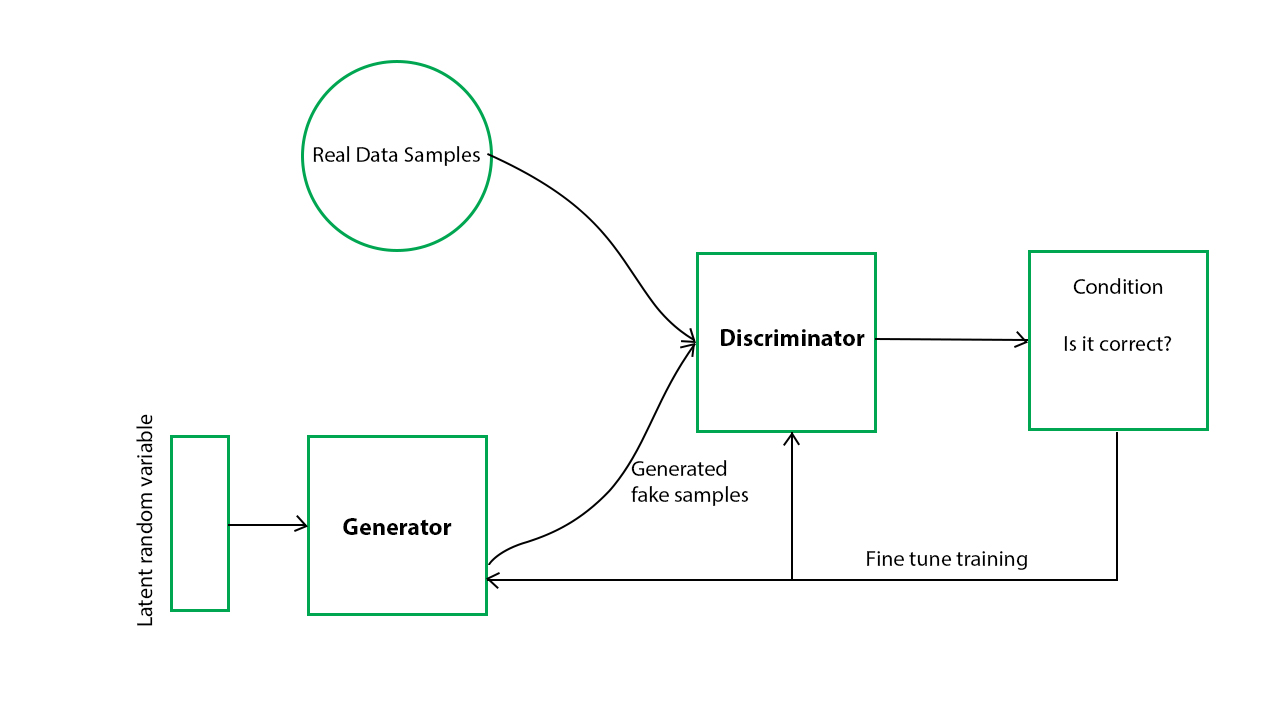
****

Figure 3.a

**4. PROPOSED SYSTEM**

Based on Deep learning, a lot of data set as an Input is provided to the software tool which will be recognized by the machine and similar pattern will be taken out from them. Octave as a building tool is used for this product but Octave is recommended in initial state as it’s free and easy to use. The Implementation of such a tool depends on two factors – Feature extraction and classification algorithm. Various classifiers can be used as available online and also read about basic feature extraction algorithm. The basic version of the product can be implemented in Octave with limited training data set and simple component analysis.

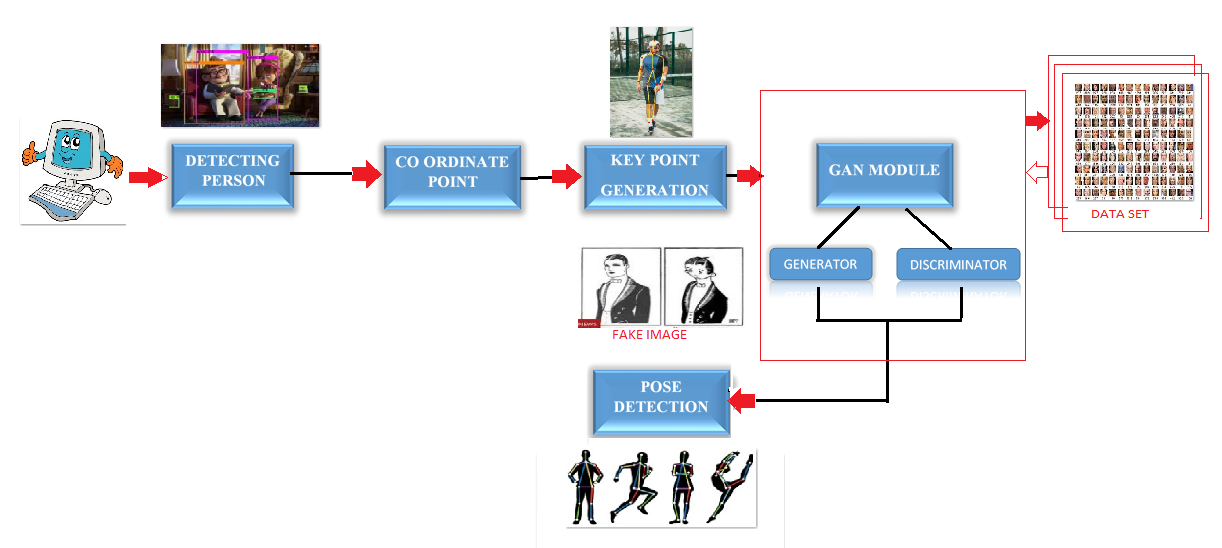


Figure 4.a

4.1 MODULES:

Now-a-days, Recognising the Human pose is the important problem in the field of computer science, As humans are more flexible to change their poses and its being more difficult to estimate the poses. It is difficult for an automatic computer to determine their poses for the analysis process. To analyze this human movement the following modules are used.

* Detecting the person
* Coordinate separation or key point generation
* GENERATIVE ADVERSARIAL NETWORK (GAN) algorithm
* Pose detection

4.1.1 DETECTING THE PERSON:

To detect a person on the frame, using cluster grouping algorithm on a set of detection areas which define a set of features based on spatial, color and temporal information for each detection. Then using these features, we cluster the detections. We finally define a measure to calculate the actual number of people within each cluster to infer the final estimation of the number of people in the scene. Can also classify the gender using colors.

4.1.2 COORDINATE POINT or KEY POINT:

To determine the pose of the human and to classify the pose structure, we need to determine the coordinate points such as elbow, shoulder, neck, knee, hip, toe, etc., which might help us to determine exact pose of the human.

4.1.3 GENERATIVE ADVERSARIAL NETWORK (GAN) ALGORITHM:

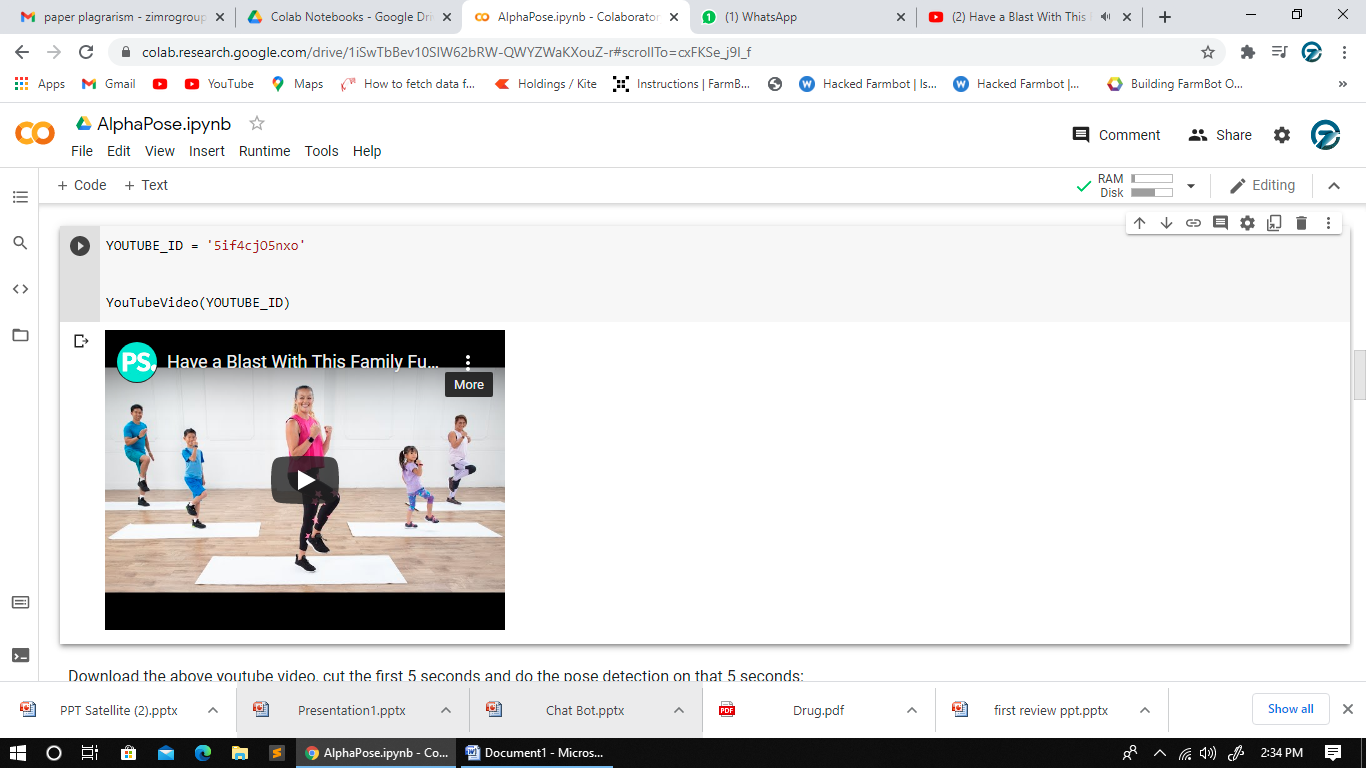
GAN has two parts generator and discriminator the generator generates fake samples of data (be it an image, audio, etc.) and tries to fool the Discriminator. The Discriminator, on the other hand, tries to distinguish between the real and fake samples. The Generator and the Discriminator are both Neural Networks and they both run in competition with each other in the training phase. The steps are repeated several times and in this, the Generator and Discriminator get better and better in their respective jobs after each repetition.

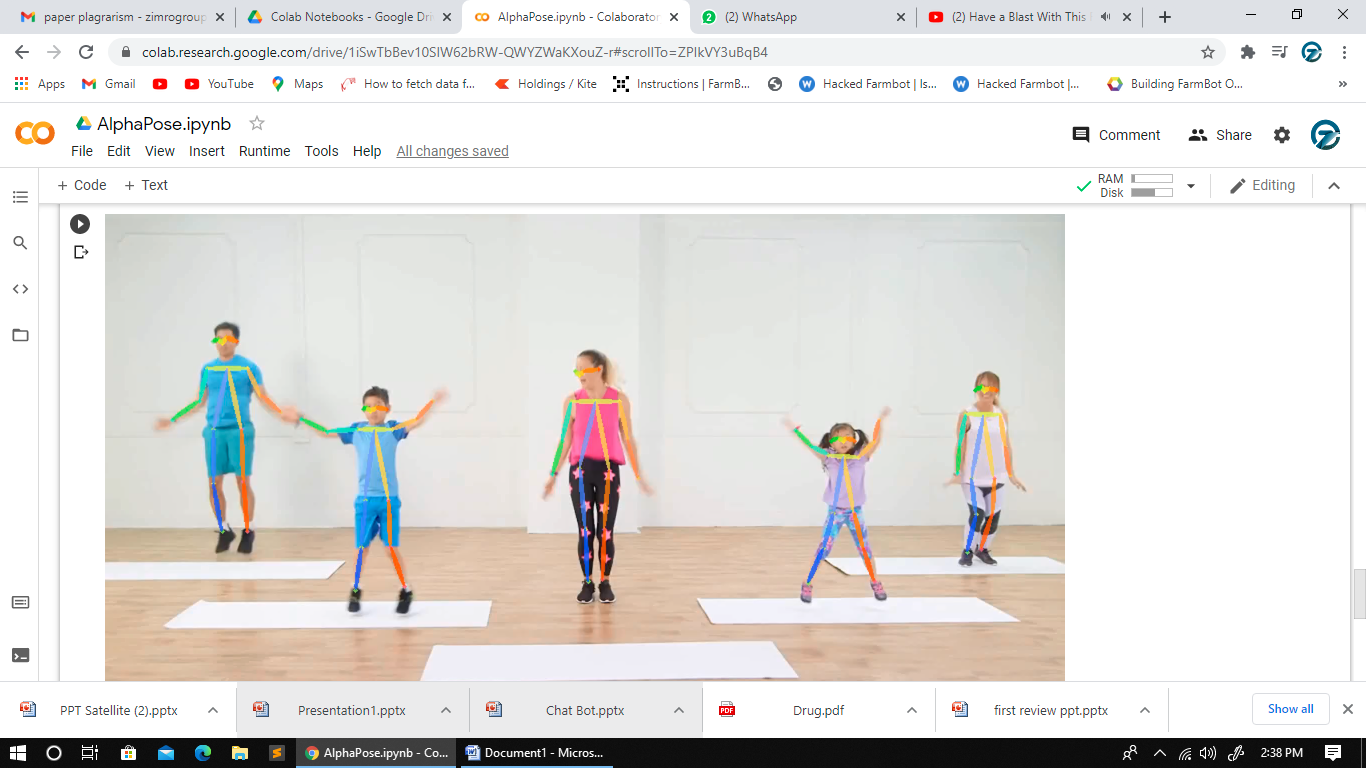
4.1.4 POSE DETECTION:

After determining all the possible position of the human image using generator and discriminator, now it would make us very simple to determine exact position of the human and to understand the human gesture

4.1.4 RESULTS:

By verifying the key points with the data sets the result can be obtained as sleeping, running, walking, etc.





##### 5 CONCLUSION

##### From the above observations the proposed system addressed and endured the difficulties in capturing the images with GAN based machine learning model and optimizes the blurred images. This proposed system provides novelty about performance measures related to the quality of human pose detections. This model concise about small variations that capable of increase the amount of real data applications involving live human datasets.

##### REFERENCES

1. S. Zhang, R. Benenson, M. Omran, J. Hosang, and B. Schiele, “How far are we from solving pedestrian detection?” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 1259–1267.
2. [2] P. Dollar, Z. Tu, P. Perona, and S. Belongie, “Integral channel features,” ´ in Proc. Brit. Mach. Vis. Conf., 2009, pp. 91.1–91.11.
3. [3] W. Nam, P. Dollar, and J. H. Han, “Local decorrelation for improved ´ pedestrian detection,” in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 424–432.
4. [4] S. Zhang, R. Benenson, and B. Schiele, “Filtered channel features for pedestrian detection.” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1751–1760. Y. Yang, Z. Wang, and F. Wu, “Exploring prior knowledge for pedestrian detection.” in Proc. Brit. Mach. Vis. Conf., 2015, pp. 176.1–176.12.
5. J. Cao, Y. Pang, and X. Li, “Pedestrian detection inspired by appearance constancy and shape symmetry,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 1316–1324.
6. S. Zhang, C. Bauckhage, and A. B. Cremers, “Informed haar-like features improve pedestrian detection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 947–954.
7. P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627–1645, 2010.