Yoga Pose Classification Using Deep Learning

Prof. Ateeq Ahmed[1], Krishna Nandkishor Bhutada[2], Rohit Kattimani[3]
[1]*Professor, Department of Information Science and Engineering HKBK College of Engineering, Bangalore, India*
[2,3]*Department of Information Science Engineering HKBK College of Engineering, Bangalore, India*

Abstract: Human pose estimation is a deep-rooted problem in computer vision that has exposed many challenges in the past. Analyzing human activities is beneficial in many fields like video surveillance, biometrics, assisted living, at-home health monitoring etc. With our fast-paced lives these days, people usually prefer exercising at home but feel the need of an instructor to evaluate their exercise form. As these resources are not always available, human pose recognition can be used to build a self-instruction exercise system that allows people to learn and practice exercises correctly by themselves. This project lays the foundation for building such a system by discussing various machine learning and deep learning approaches to accurately classify yoga poses on prerecorded videos and also in real-time. The project also discusses various pose estimation and keypoint detection methods in detail and explains different deep learning models used for pose classification.

Keywords: - **Human pose estimation, yoga, openpose, machine learning, deep learning**

I. INTRODUCTION

Human pose estimation is a challenging problem in the discipline of computer vision. It deals with localization of human joints in an image or video to form a skeletal representation. To automatically detect a person's pose in an image is a difficult task as it depends on a number of aspects such as scale and resolution of the image, illumination variation, background clutter, clothing variations, surroundings, and interaction of humans with the surroundings. An application of pose estimation which has attracted many researchers in this field is exercise and fitness. One form of exercise with intricate postures is yoga which is an age-old exercise that started in India but is now famous worldwide because of its many spiritual, physical and mental benefits [2]. The problem with yoga however is that, just like any other exercise, it is of utmost importance to practice it correctly as any incorrect posture during a yoga session can be unproductive and possibly detrimental. This leads to the necessity of having an instructor to supervise the session and correct the individual's posture.

Since not all users have access or resources to an instructor, an artificial intelligence-based application might be used to identify yoga poses and provide personalized feedback to help individuals improve their form.

In recent years, human pose estimation has benefited greatly from deep learning and huge gains in performance have been achieved. Deep learning approaches provide a more straightforward way of mapping the structure instead of having to deal with the dependencies between structures manually. used deep learning to identify 5 exercise poses: pull up, swiss ball hamstring curl, push up, cycling and walking. However, using this method for yoga poses is a relatively newer application. YOGA POSE CLASSIFICATION USING DEEP LEARNING 2 This project focuses on exploring the different approaches for yoga pose classification and seeks to attain insight into the following: What is pose estimation? What is deep learning? How can deep learning be applied to yoga pose classification in real-time? This project uses references from conference proceedings, published papers, technical reports and journals. Fig. 1 gives a graphical overview of topics this paper covers. The first section of the project talks about the history and importance of yoga. The second section talks about pose estimation and explains different types of pose estimation methods in detail and goes one level deeper to explain discriminative methods – learning based (deep learning) and exemplar. Different pose extraction methods are then discussed along with deep learning based models - Convolutional Neural Networks (CNNs) and

Recurrent Neural Networks (RNNs).
Employers view users of the tool as flexible and well- prepared applicants because of its capacity to adjust to the constantly shifting demands of the labor market. Furthermore, by providing assistance with CV construction and practice interviews, the Virtual Assistant Bot elevates the training experience. In a virtual setting, users can hone their interviewing techniques and build their confidence, preparing

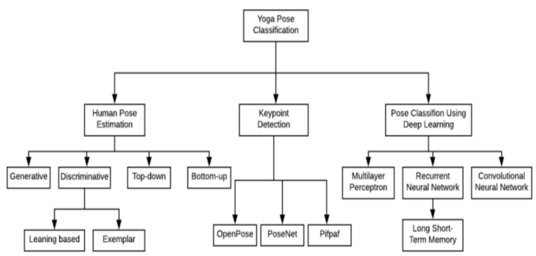

Fig-Architecture of model

II. LITERATURE SURVEY

AI strategies may maybe rely significantly upon heuristic human element extraction in everyday errands of location of social exercises. It is limited by human zone mindfulness. To talk this danger, creators have decided on a couple of techniques like profound learning methods. These strategies could consequently extricate explicit highlights during the reparing stage from crude sensor information, and afterward low-level fleeting qualities with significant level unique requests would be introduced. With regards to the freed application from profound learning approaches in fields like picture grouping, voice acknowledgment, preparing of regular language, and some others, it has been developing into a novel examination way in design location and to move it to an area of human action identification. Table 2 shows a couple of the current AI alongside the exactness of har.
Here, we have referenced just the techniques that give the best exactness for the most extreme number of subjects. For instance, on the off chance that the subject tally is less, at that point the precision might be better. Chen et al. proposed a system for distinguishing a yoga act utilizing a Kinect camera. Those assembled an amount of 300 recordings of 12 yoga stances from 5 yoga pros with each present performed on five events. At first, the closer view part is fragmented from the cut, and the star skeleton is used and acquired an exactness of 99.33%.
The creators in proposed a stance location system utilizing a quality Kinect camera with a goal of 640 X 480. They saw six postures performed by five volunteers. They separated 21000 casings from the Kinect camera by utilizing a foundation deduction technique with a 74% precision. Later the creators in proposed another posture discovery methodology utilizing Kinect camera. Also, Wang et al. proposed a position acknowledgment methodology using the Kinect camera. They isolated the human blueprint and utilized a learning vector quantization neural framework for five fundamental stances. The system accuracy was roughly 98 %. One the less, these outcomes have high precision rates, and they are seen as security prominent.

III. PROPOSED WORK

Human posture recognition has made huge advancements in the past years. It has evolved from 2D to 3D pose estimation and from single person to multi person pose estimation. uses pose estimation to build a machine learning application that helps detect shoplifters whereas uses a single RGB camera to capture 3D poses of multiple people in real-time. Human pose estimation algorithms can be widely organized in two ways. Algorithms prototyping estimation of human poses as a geometric calculation are classified as generative methods while algorithms modelling human pose estimation as an image processing problem are classified as discriminative methods. Another way of classifying these algorithms is based on their method of working. Algorithms starting from a higher-level generalization and moving down are called top-down methods, whereas algorithms that start with pixels and move upwards are called bottom-up methods.
Human posture recognition has undergone a remarkable evolution, progressing from basic two-dimensional analysis to the intricate estimation of 3D poses, even in scenarios involving numerous individuals. This technological leap not only enables diverse applications but also reflects a shift in how we extract information from images. Early systems were limited to a flattened perspective, whereas modern algorithms reconstruct the full physicality of human form within a 3D space. These advancements

empower diverse applications, including security solutions empowered to detect potential shoplifters – analyzing subtle behavioral cues beyond mere identification. Furthermore, systems capable of capturing multiple people's movements in 3D using conventional cameras revolutionize fields like motion analysis for sports, rehabilitation, and even large-scale animation for the entertainment industry. This system can be used in operating systems in a similar manner to Siri for effective information retrieval just by speaking various queries. Chatbots are computer programs that interact with users using natural languages. Just as people use language for human communication, chatbots use natural language to communicate with human users. In this paper, we begin by introducing chatbots and emphasize their need in day-to-day activities. Then we go on to discuss existing chatbot systems, namely ELIZA, ALICE and Siri. We evaluate what we can take from each of the systems and use in our proposed system. Finally, we discuss the proposed system. The system we intend to develop can be used for educating the user about sports. The database will be fed with sports related data which will be coded using AIML.
This system can be used in operating systems in a similar manner to Siri for effective information retrieval just by speaking various queries

IV. MODULES

1. Implementation Of Complete Framework
Support Vector Machine (SVM) in Python SVM, implemented using Python libraries like TensorFlow - Keras (with Theano backend), OpenPose, NumPy, and Scikit Learn, is a powerful supervised machine learning model commonly employed in classification tasks. In its essence, SVM serves as a binary classifier, but it can be extended to handle multiclass classification scenarios.

Overview of SVM: SVM operates by creating a hyperplane that maximizes the margin of separation between classes in the feature space. For multiclass classification, SVM constructs multiple binary classifiers, distinguishing each class from the rest. Two common strategies for multiclass SVM are one-vs-rest (OvR) and one-vs-one (OvO).

Kernel Function: The choice of kernel function significantly influences SVM's performance. The radial basis function (RBF) kernel, a popular choice, is employed in this implementation. RBF offers greater flexibility compared to linear and polynomial kernels, making it suitable for capturing complex relationships in the data.

Model Configuration: In this implementation, a default SVM model is trained on the provided training data. The RBF kernel is utilized, with a soft margin parameter (C) set to 1. The decision function is implemented using the one-vs-rest strategy. equal. The SVM model is executed on a system equipped with an NVIDIA Tesla 1080 GPU, boasting 4 GB of memory. This configuration facilitates efficient training and inference, particularly for computationally intensive tasks like machine learning on large dataset swinning open source natural language artificial intelligence college enquiry bot which utilizes AIML to form responses to queries.

1. Knowledgebase creation for Chabot
Interactive Question-Answering systems turn the focus to the communication between the user and the program, not just the question answering. These systems allow either the user to drive the dialogue or the system to play a greater role by suggesting related materials or even refinements to a user's query. An interactive layer in a QA system also allows for more complex queries to be issued by the user as mistakes made by the system can be rectified through user input; or the system can engage the user to check its understanding of a user's query is correct. Text to Speech: The bot also speaks out the answer.
2. Integration and Connectivity
As that defined in bot to bringing together the all data to collect from the knowledgebase and place in SQLite3 database. That connected with the application and gets the proper query of college. Display the conversation between user and the bot sure ask the query and bot will be answring.as we can see the conversation user say hi and answer of hi display by the bot, .It's totally depends on user which type of query user will asked.
3. Knowledgebase creation for Chabot
Interactive Question-Answering systems turn the focus to the communication between the user and the

program, not just the question answering. These systems allow either the user to drive the dialogue or the system to play a greater role by suggesting related materials or even refinements to a user's query. An interactive layer in a QA system also allows for more complex queries to be issued by the user as mistakes made by the system can be rectified through user input; or the system can engage the user to check its understanding of a user's query is correct. Text to Speech: The bot also speaks out the answer.

V. USES AND APPLICATION

A yoga pose classifier using deep learning can enhance yoga practice by providing real-time feedback on form and alignment, guiding beginners and assisting in personalized workout recommendations. It aids in injury prevention and rehabilitation, tracks progress over time, and supports virtual yoga classes. Additionally, it enables research into the effects of yoga on health conditions, promotes accessibility for individuals with disabilities, and contributes to health and wellness applications.

VI. SCOPE

The scope of a yoga pose classifier using deep learning is broad and promising. It extends across various domains including fitness technology, healthcare, education, and research. With advancements in deep learning algorithms and the proliferation of wearable devices equipped with sensors, the potential applications are vast.In fitness technology, such a classifier can revolutionize how people practice. yoga, offering real-time feedback and personalized workout routines. This not only enhances the user experience but also promotes safer and more effective exercise habits.In healthcare, the classifier can be utilized for injury prevention, rehabilitation, and monitoring of physical therapy sessions. It can also contribute to research on the therapeutic effects of yoga for various health conditions. In education, the classifier can serve as a virtual instructor, making yoga more accessible to a wider audience. It can provide guidance and support to learners of all levels, from beginners to advanced practitioners. Moreover, the scope extends to areas such as sports performance analysis, where understanding yoga poses and their biomechanics can enhance athletes' flexibility, strength, and overall performance.

VII. CONCLUSION

The research outlined in this study capitalizes on deep learning methodologies to identify and rectify improper yoga postures, thereby providing users with guidance on enhancing their form by pinpointing areas of misalignment. Within the proposed system, individuals have the autonomy to select their desired yoga poses for practice, subsequently uploading recorded videos of their sessions for analysis. Integral to this approach is the extraction of angular measurements from monitoring activities, which are then utilized as features after being appropriately scaled. However, it's worth noting that in instances where key points undergo rotation, the angles may remain unchanged, resulting in less reliable outcomes. To mitigate this issue, the system prioritizes angles relative to the ground rather than between joints. Consequently, even slight rotations of key points can lead to variations in angle measurements, ensuring greater accuracy in posture assessment. Central to the effectiveness of this system is the training of a multilayer perceptron (MLP) model, which attains an impressive accuracy rate of 99.58% when evaluated against testing datasets. Comparative analysis with existing methodologies reveals the robustness of the proposed approach. While SVM achieves a test accuracy of 93.19%, CNN and CNN + LSTM models achieve rates of 98.58% and 99.38%, respectively. Although the MLP model exhibits lower computational power than CNN and CNN + LSTM configurations, its modified feature set facilitates a remarkable accuracy rate of 99.58%. These findings underscore the efficacy of the proposed methodology, which not only outperforms existing techniques but also maintains a low computational complexity.

Beyond its technical prowess, the proposed approach holds significant promise for practical application, particularly in the realm of self-guided yoga practice. By leveraging the system's ability to detect incorrect postures, users can preemptively address alignment issues, thereby mitigating the risk of chronic problems stemming from improper form. Moreover, the system's low computational complexity enhances its accessibility, rendering it suitable for integration into individuals' busy lifestyles. Ultimately, the experimental results gleaned from this research offer compelling evidence of the system's efficacy in facilitating safe and effective yoga practice, underscoring its potential to empower individuals to prioritize their physical well-being amidst the demands of modern life.

REFRENCE

[1] L. Sigal. "Human pose estimation", Ency. of Comput. Vision, Springer 2011.

[2] S. Yadav, A. Singh, A. Gupta, and J. Raheja, "Real-time yoga recognition using deep learning", Neural Comput. and Appl., May 2019. [Online]. Available: https://doi.org/10.1007/s00521-019-04232-7

[3] U. Rafi, B. Leibe, J.Gall, and I. Kostrikov, "An efficient convolutional network for human pose estimation", British Mach. Vision Conf., 2016.

[4] S. Haque, A. Rabby, M. Laboni, N. Neehal, and S. Hossain, "ExNET: deep neural network for exercise pose detection", Recent Trends in Image Process. and Pattern Recog., 2019.

[5] M. Islam, H. Mahmud, F. Ashraf, I. Hossain and M. Hasan, "Yoga posture recognition by detecting human joint points in real time using microsoft kinect", IEEE Region 10 Humanit. Tech. Conf., pp. 668-67, 2017.

[6] S. Patil, A. Pawar, and A. Peshave, "Yoga tutor: visualization and analysis using SURF algorithm",Proc. IEEE Control Syst. Graduate Research Colloq.,pp. 43-46, 2011.

[7] W. Gong, X. Zhang, J. Gonzàlez, A. Sobral, T. Bouwmans, C. Tu, and H. Zahzah, "Human pose estimation from monocular images: a comprehensive survey", Sensors, Basel, Switzerland, vol. 16, 2016.

[8] G. Ning, P. Liu, X. Fan and C. Zhan, "A top-down approach to articulated human pose estimation and tracking", ECCV Workshops, 2018.

[9] A. Gupta, T. Chen, F. Chen, and D. Kimber, "Systems and methods for human body pose estimation", U.S. patent, 7,925,081 B2, 2011. YOGA POSE CLASSIFICATION USING DEEP LEARNING 38

[10] H. Sidenbladh, M. Black, and D. Fleet, "Stochastic tracking of 3D human figures using 2D image motion", Proc 6th European Conf. Computer Vision, 2000.

[11] A. Agarwal and B. Triggs, "3D human pose from silhouettes by relevance vector regression", Intl Conf. on Computer Vision & Pattern Recogn.pp.882–888, 2004.

[12] M. Li, Z. Zhou, J. Li and X. Liu, "Bottom-up pose estimation of multiple person with bounding box constraint", 24th Intl. Conf. Pattern Recogn.,2018.

[13] Z. Cao, T. Simon, S. Wei, and Y. Sheikh, "OpenPose: realtime multi-person 2D pose estimation using part affinity fields", Proc. 30th IEEE Conf. Computer Vision and Pattern Recogn,2017.

[14] A. Kendall, M. Grimes, R. Cipolla, "PoseNet: a convolutional network for real-time 6- DOF camera relocalization", IEEE Intl. Conf. Computer Vision, 2015.

[15] S. Kreiss, L. Bertoni, and A. Alahi, "PifPaf: composite fields for human pose estimation", IEEE Conf. Computer Vision and Pattern Recogn, 2019.

[16] P. Dar, "AI guardman – a machine learning application that uses pose estimation to detect shoplifters". [Online]. Available: https://www.analyticsvidhya.com/blog/2018/06/ai-guardman-machine-learningapplication-estimates-poses-detect-shoplifters/

[17] D. Mehta, O. Sotnychenko, F. Mueller and W. Xu, "XNect: real-time multi-person 3D human pose estimation with a single RGB camera", ECCV, 2019.

[18] A. Lai, B. Reddy and B. Vlijmen, "Yog.ai: deep learning for yoga". [Online]. Available: http://cs230.stanford.edu/projects_winter_2019/reports/15813480.pdf

[19] M. Dantone, J. Gall, C. Leistner, "Human pose estimation using body parts dependent joint regressors", Proc. IEEE Conf. Computer Vision Pattern Recogn., 2013. YOGA POSE CLASSIFICATION USING DEEP LEARNING 39

[20] A. Mohanty, A. Ahmed, T. Goswami, "Robust pose recognition using deep learning", Adv. in Intelligent Syst. and Comput, Singapore, pp 93-105, 2017.