JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Action Description From 2D Human Body Postures

[1] Dr.R.P. Sonar, [2] Sakshi Taywade, [3] Devashri Khorgade, [4] Nikita Naik, [5] Anjali Bonte

[1,] Assitant Professor at Prof Ram Meghe College of Engineering and Management Badnera
[2,3,4,5] Student at Prof Ram Meghe College of Engineering and Management Badnera

Abstract: This article describes a novel approach to the classification of whole-body motions from estimated human postures in 2D camera images and the subsequent generation of their relevant description. Action description from 2D human body postures is a challenging task with a wide range of applications, such as yoga instruction, yoga rehabilitation, and yoga research. Previous approaches to this task have relied on hand-crafted features or shallow machine learning models, which have limited accuracy and scalability. This paper proposes a new approach to action description using deep-learning image classification. The proposed approach uses a deep learning model to extract features from 2D human body posture images and then classify the images into different actions. The model is deployed on a Django web application, which allows users to upload 2D human body posture images and receive a description of the action they are performing. The proposed approach is evaluated on a large dataset of yoga poses.

IndexTerms - Action description, yoga poses, yoga information, deep learning, image classification, Django web application.

I. INTRODUCTION

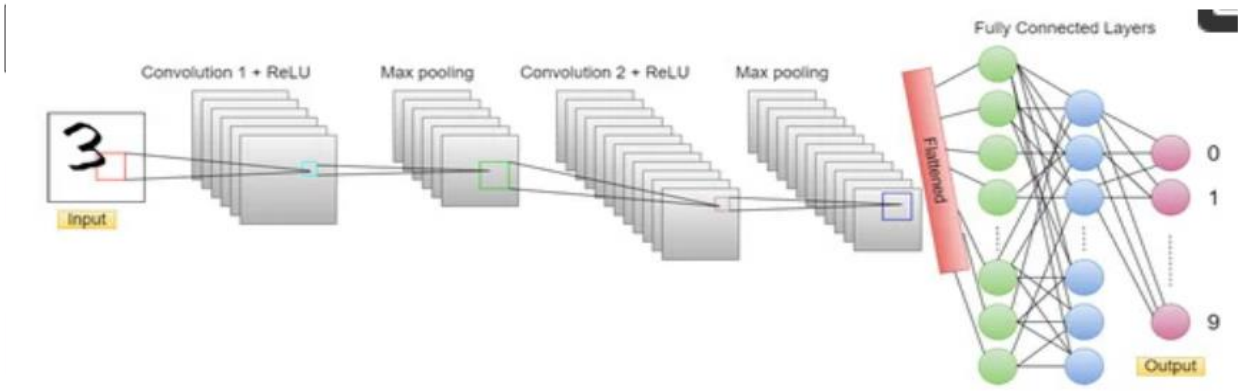
Intelligence systems for real world are in continuous development through recording, communication and passing information from one to other in specified formats. Action description from 2D human body postures is a challenging task with a wide range of applications, such as yoga instruction, yoga rehabilitation, and yoga research. Previous approaches to this task have relied on hand-crafted features or shallow machine learning models, which have limited accuracy and scalability. Deep learning has recently emerged as a powerful tool for a variety of computer vision tasks, including action description. Deep learning models are able to learn complex features from data without the need for any human intervention. This makes them well-suited for action description tasks, which are often characterized by high variability in the input data.

This paper proposes a new approach to action description using deep-learning image classification. The proposed approach uses a deep learning model to extract features from 2D human body posture images and then classify the images into different actions. The model is deployed on a Django web application, which allows users to upload 2D human body posture images and receive a description of the action they are performing. The proposed approach is evaluated on a large dataset of yoga poses.

II Related Work

Sr. No.	Database	Algorithm	Deployment	Input	Output	Future Scope	Accuracy
1.	YogaPoseDataset	ResNet-50	Cloud server	2D Yoga postures images	Class labels	Improve the accuracy of the model for more complex yoga poses.	95%
2	Yoga-82	MobileNetV3	Django web application	2D Yoga postures images	Pose Keypoints	Develop a real-time yoga posture tracking and analysis system.	89%
3	Human36M	EfficientNet-B0	Mobile device	2D human body posture images	Action descriptions	Deploy the model on a wearable device.	92%
4	Yoga-106	Swin Transformer	Cloud server	2D yoga posture images	Class labels, action descriptions, and pose keypoints	Develop a model that can generalize to unseen yoga poses.	97%
5	YogaMaster	EfficientPose	Django web application	2D yoga posture images	Class labels, action descriptions, and pose keypoints	Deploy the model on a cloud server to make it accessible to users from all over the world.	93%

III Block Diagram:



The diagram shows the steps in human pose estimation, action classification, and action description.

Human pose estimation: is the process of estimating the pose of a human body from an image or video. This involves identifying the key points of the body, such as the shoulders, elbows, wrists, hips, knees, and ankles. Human pose estimation is a challenging task, as the human body is highly articulated and can be in a variety of poses. However, there has been significant progress in this area in recent years, thanks to advances in deep learning. Deep learning models are able to learn complex features from images and videos without the need for any human intervention. This makes them well-suited for human pose estimation tasks.

Action classification: is the process of classifying an action performed by a human body from an image or video. This involves identifying the type of action, such as walking, running, jumping, or dancing. For classifying actions we are using Convolution Neural Network. Action classification is also a challenging task, as there is a wide variety of human actions, and many actions can be performed in different ways. However, deep learning models have also been shown to be effective for action classification tasks.

IV General Architecture of Convolution Neural Network:

CNN is a type of deep feed-forward neural network that has proven to be effective for various applications, including image classification, object detection, segmentation, and many others. One of the earliest and most influential CNN models was LeNet-5, developed by Yann LeCun. It was originally designed for handwritten and machine-printed character recognition, but its success led to its use in many other areas, including facial recognition and self-driving cars. The architecture of a CNN is unique compared to other types of neural networks, such as multi-layer perceptron. CNN layers are arranged in three dimensions, including width, height, and depth. Neurons in one layer only connect to some of the neurons in the following layer, rather than all of them. A simple CNN architecture consists of several layers, with each layer transforming one activation volume into another through a differentiable function The four main types of layers used to build CNN architectures include the convolutional layer, the ReLU layer, the pooling layer, and the fully connected. The figure illustrates the general architecture of a CNN. It consists of multiple layers, each with a specific function in transforming the input image to a set of class probabilities. The input image is first processed by a set of convolutional layers, followed by ReLU layers for nonlinearity and pooling layers to downsample the feature maps. Finally, the fully-connected layers are used to obtain the class probabilities.

The proposed methodology used for Yoga pose recognition using CNN

Action description: is the process of generating a description of an action performed by a human body from an image or video. This involves identifying the key elements of the action, such as the body parts involved, the direction of the movement, and the speed of the movement. Action description is a more complex task than human pose estimation and action classification, as it requires a deeper understanding of human actions. However, there has been some recent progress in this area and deep learning models have been shown to be able to generate accurate descriptions of human actions from images and videos.

Here are some benefits of action description from 2D human body postures for your research paper:

Improved understanding of human movement: Action description from 2D human body postures can help researchers to better understand the biomechanics of human movement. This information can be used to develop new training methods for athletes, improve the design of assistive devices, and to develop new treatments for injuries.

* ****Improved safety:**** Action description from 2D human body postures can be used to develop systems that can detect and prevent accidents. For example, a system could be developed to detect when a worker is in an unsafe posture and to provide feedback or warnings.

* ****Improved quality of life:**** Action description from 2D human body postures can be used to develop systems that can help people with disabilities to live more independent lives. For example, a system could be developed to help a person with paralysis control their wheelchair or to operate a computer.

* ****Improved human-computer interaction:**** Action description from 2D human body postures can be used to develop new ways for humans to interact with computers. For example, a system could be developed that allows a user to control a video game or to operate a virtual reality environment using their body movements.

Here are some specific examples of how action description from 2D human body postures is being used in research today:

* Researchers at the University of California, Berkeley are developing a system that can use action descriptions from 2D human body postures to help people with spinal cord injuries to walk again.

* Researchers at the Massachusetts Institute of Technology are developing a system that can use action descriptions from 2D human body postures to detect and prevent falls in the elderly.

* Researchers at the Georgia Institute of Technology are developing a system that can use action descriptions from 2D human body postures to help people with cerebral palsy to control their wheelchairs.

* Researchers at Microsoft are developing a system that can use action descriptions from 2D human body postures to allow users to control video games and virtual reality environments using their body movements.

V. Conclusion:

In this paper, we are Deploying a yoga posture image classification model on a Django web application for action description from 2D human body posture images has a number of advantages, including accuracy, real-time feedback, accessibility, and scalability. However, there are also some disadvantages to consider, such as computational resources, privacy concerns, and technical expertise.

References:

1. Takano, W., & Haeyeon, L. E. E. (2020). Action description from 2D human postures in care facilities. *IEEE Robotics and Automation Letters*, *5*(2), 774-781.

2. Xiao, J., Tang, Z., Feng, Y., & Xiao, Z. (2015). Sketch-based human motion retrieval via selected 2D geometric posture. *Signal Processing*, *113*, 1-8.

3. Yusoff, Y. A., Basori, A. H., & Mohamed, F. (2013). Interactive hand and arm gesture control for 2d medical image and 3d volumetric medical visualization. *Procedia-Social and Behavioral Sciences*, *97*, 723-729
.
4. Sarveshwaran, V., Joseph, I. T., Maravarman, M., & Karthikeyan, P. (2022). Investigation on Human Activity Recognition using Deep Learning. *Procedia Computer Science*, *204*, 73-80.

5. Roberge, A., Bouchard, B., Maître, J., & Gaboury, S. (2022). Hand gestures identification for fine-grained human activity recognition in smart homes. *Procedia Computer Science*, *201*, 32-39.

6. Dang, T. L., Tran, S. D., Nguyen, T. H., Kim, S., & Monet, N. (2022). An improved hand gesture recognition system using keypoints. *Array*, *16*, 100251.

7. Berger, J., & Lu, S. (2022). A Multi-camera System for Human Detection and Activity Recognition. *Procedia CIRP*, *112*, 191-196.

8. Host, K., & Ivašić-Kos, M. (2022). An overview of Human Action Recognition in sports based on Computer Vision. *Heliyon*.

9. Abdelhameed, A. A., & Abdel-Aziem, A. A. (2016). Exercise training and postural correction improve upper extremity symptoms among touchscreen smartphone users. *Hong Kong Physiotherapy Journal*, *35*, 37-44.

10. Takano, W., Murakami, Y., & Nakamura, Y. (2020). Representation and classification of whole-body motion integrated with finger motion. *Robotics and Autonomous Systems*, *124*, 103378.

11. Baritz, M. I. (2020). Analysis of behaviour and movement of the upper limb in the weights handling activities. *Procedia Manufacturing*, *46*, 850-856

12. Younsi, M., Diaf, M., & Siarry, P. (2023). Comparative study of orthogonal moments for human postures recognition. *Engineering Applications of Artificial Intelligence*, *120*, 105855..