

MASTERING YOGA POSTURE USING DEEP NEURAL NETWORKS

Dataset:

We have multiple datasets for key points detection like the Common Objects in Context (COCO) [1] dataset and Max-Planck-Institut für Informatik (MPII) [2] human pose estimation dataset. We have read both the datasets into python environment and visualized some images with annotations to get the overview of those datasets. In this process we found that COCO dataset is a collection of images from various classes and images containing humans is one of the classes of that dataset. MPII dataset is especially for human pose estimation.

Data Providers:

COCO -Microsoft with other sponsors (like Facebook, MightyAI)

MPII- Max Planck Institute for Informatics, Germany

Library: we have used the following libraries to read the dataset.

- API tensorlayer [3] a framework developed by google for MPII dataset.
- pycocotools [4]- COCO API assists in loading, parsing, and visualizing annotations in COCO library for reading COCO dataset.

Dataset assessment:

COCO Dataset:

Images are broadly classified into 12 super-categories and further classified into 80 sub-categories and 'person' is one of the categories of the 12 main categories which makes the percentage of useful data less when compared to total size of the data.

Bias: - We need a dataset with images containing humans and their annotations containing the respective postures to effectively train our model. But COCO dataset has less than 10% of its total data categorized as person data which means that more than 90% of the data is not affectively useful for training our model. This shows a clear imbalance of useful data for our project in COCO dataset. This dataset does not contain any special category for yoga positions which is very critical for our project.

Data Annotation: - Images in the data set have been annotated with the following features which makes the data usable for your purposes.

	segmentation	num_keypoints	area	iscrowd	keypoints	image_id	bbox	category_id	id
0	[[267.03, 243.78, 314.59, 154.05, 357.84, 136....	8	28292.08625	0	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...	537548	[267.03, 104.32, 229.19, 320]	1	183020
1	[[640, 408.25, 639.54, 4.33, 546.25, 5.29, 426...	2	39122.63310	0	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...	117891	[206.77, 1.44, 433.23, 408.73]	1	183030

Figure 1

MPII Dataset:

MPII Human Pose dataset has 25000 images which contains more than 40000 people with annotated body joints. This dataset is a collection of images containing every day human activities, to be specific 410 human activities with activity labels. These images were extracted from various YouTube videos.

Bias: - We did not find any noticeable bias in the dataset as it contains a wide variety of human postures, but our only concern is that this dataset does not concentrate on yoga postures which limits the number of yoga postures in the dataset.

Data Annotation: - The following annotated features in the images makes this dataset validly usable for your purposes.

	filename	train	head_rect	is_visible	joint_pos
0	015601864.jpg	1	[627.0, 100.0, 706.0, 198.0]	{'6': 0, '7': 1, '8': 0, '9': 0, '0': 1, '1': ...	{6: [610.0, 187.0], 7: [647.0, 176.0], 8: [637...
1	015601864.jpg	1	[841.0, 145.0, 902.0, 228.0]	{'6': 0, '7': 0, '8': 0, '9': 0, '0': 1, '1': ...	{6: [979.0, 221.0], 7: [906.0, 190.0], 8: [912...

Figure 2

Yoga Position:

We would collect image dataset from the blog sample projects posted by Dhruv Jawalkar. These images are downloaded from google using a keyword 'asana' and cropped to retain the human subject in the images.[6]

This dataset consists of images of people doing 107 different asanas in 5994 total images.

Bias: - We did not find any bias in this dataset as it contains a wide variety of yoga positions and multiple images for each position with a different human subject.

Pre-Processing

The data is in two files. All the images are in jpg format in one file and the annotations are in json file. We can use the above-mentioned library to associate annotations with the respective images.

Image Resolution -The images in both the datasets have different resolution. We must convert them into a standard resolution and evaluate them once the processing is done.

Ethical principles

Consent: We will ask the user for their consent to give us permission to store their login profile details, for recording and storing the video while they are practicing yoga using our application for analytical and training purposes.

Clarity: Data will be used to track, analyze and report the progress of the user by analyzing their data.

Consistency: User data will be stored in well secured data centers which are equipped with state of art security firewalls and maintained as per government regulations.

Control: Users have total control over their data. Then can update or delete their data just by logging in to the application. They can even delete their account using the application which results in deletion of all their data with us without a possibility for recovery.

Consequences: User data is not used for any application which may have harmful consequences for the users. User data will not be shared with any other individual or organization for any purpose without proper user consent.

License to use data:

COCO: This data is licensed by CREATIVE COMMONS PUBLIC LICENSE and the data can be used without any cost or liability for no commercial purpose. In case of any modification or updating the dataset then it cannot be shared publicly distributed or made available digitally in any form.[5]

MPII: The data is openly available and can be used for educational and research purposes. There are no specific terms and conditions cited in the website regarding the use of this dataset, we will contact the collectors of the dataset to obtain the appropriate permission for any commercial use in future.[2]

Yoga: Being an open source dataset it's free to use for educational and research purposes.[6]

Ability to use multiple data sources

We have one constraint to use multiple data sources, the number of key points annotated in each image differ from one data source to other. For example, in MPII and Sports Leads dataset the number of key points annotated are 14 and in COCO dataset the number is 17. As MPII dataset is specially collected for human pose estimation we are using MPII dataset for training our model.

Once our model is trained and ready to deploy then we will get the pose estimations for yoga dataset using our model. Then we will use these estimates as the benchmarks to compare the human pose detected in the real time.

Solutions to be implemented based on the size of the data

The size of the data is 12 GB which is large when compared to the computational power of our laptops which have 16 GB RAM. So, we are going to split the data into batches of images and then feed them to our model for training.

Future data needs and challenges

In real time the data would come in different formats and resolution as it will be captured on different devices. We must process this data into a standard format and resolution before feeding it to the model and this process should be fast to reduce any lag in the pose estimation.

Currently we are working with images and small videos not the real time data. So, the data preprocessing steps should be taken care in the final application. We are training our model with images and videos and it should be compatible to work continuously with real-time data.

Data encryption for security reasons when the model is deployed in the server.

Code

The code to read the datasets is in the following files:

- The file MPII.html contains the code for reading MPII human pose estimation dataset.
- The files COCO.html and COCO_Eploration.html contain the code for reading COCO dataset.

Reference

- [1] <http://cocodataset.org/#download>
- [2] <http://human-pose.mpi-inf.mpg.de/#download>
- [3] <https://tensorlayer.readthedocs.io/en/latest/>
- [4] <https://pypi.org/project/pycocotools/>
- [5] <https://creativecommons.org/licenses/by-nc-sa/2.0/legalcode>
- [6] <https://dopelemon.me/yoga-poses.html>