

Abhishek Manoj Kumar, Student ID: 01675536

Date: 11/07/2016

Note: I have not mentioned any preference and would like to get your suggestions on which of these papers I could try and reproduce.

Project Survey:

Paper 1:

Reference:

Title: PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization.

About: Here an image is uploaded and it is compared to the database and a location of the image on a map is found.

Authors: Alex Kendall, Mathew Grimes, Roberto Cipolla.

Publisher: University of Cambridge.

DOI: [arXiv:1505.07427v4](https://arxiv.org/abs/1505.07427v4)

Journal: Cornell University Library.

Year of Publish: 2015

Code: The code is made available and makes use of matlab and python. It can be found at <https://github.com/alexgkendall/caffe-posenet>

Data: Two types of data is used here, one is for the outdoor scenes and the other for indoor. The outdoor scene data *Cambridge landmarks* was collected by them and publically available *7 scenes* dataset is used for indoor scenes. Both the datasets are made available here:

<http://mi.eng.cam.ac.uk/projects/relocalisation/>

VM's and Containers: No VM and container are provided, but a demo is provided which allows us to input image of our choice and get the results i.e. where it is located. I tried this and found that if the scene is present in the dataset, then this demo works.

Workflows: The documentation is limited but intermediate results are provided and the implementation details along with information about the dataset are provided.

Provenance: This research combines metric based and appearance based approaches which were previously used individually.

Reproducible rating: I would rate it at 3, as the code and dataset is made available but I couldn't find enough documentation on the workflow.

Paper 2:

Reference:

Title: We Are Humor Beings: Understanding and Predicting Visual Humor.

About: Here, an image scene is provided as input, and they try to train the machine to find if the scene is funny or not.

Paper URL:

http://www.cvfoundation.org/openaccess/content_cvpr_2016/papers/Chandrasekaran_We_Are_Humor_CVPR_2016_paper.pdf

Authors: Chandrasekaran, Arjun and Vijayakumar, Ashwin K. and Antol, Stanislaw and Bansal, Mohit and Batra, Dhruv and Lawrence Zitnick, C. and Parikh, Devi.

Journal: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

Year of Publish: 2016

Code: Python language is used and the code is made available at https://github.com/VT-vision-lab/abstract_scenes_v002

Data: The data is made available at https://github.com/VT-vision-lab/abstract_scenes_v002

VM's and Containers: No VM'S and containers are provided.

Workflows: A complete workflow is provided and information about all the dependencies like the python version, type of Jason files, etc.

Provenance: The main reason for this research was to improve the interaction between AI and humans, i.e. to enable computers to better understand humans.

Reproducible rating: 4 out of 5, since the code, data and extensive documentation is provided but a VM or containers is not used.

Paper 3:

Reference:

Title: Nonparametric Scene Parsing via Label Transfer.

About: The research is on image parsing and segmenting objects in an image.

Authors: Ce Liu, Jenny Yuen, Antonio Torralba

DOI: 10.1109/TPAMI.2011.131

Journal: IEEE Transactions on Pattern Analysis and Machine Intelligence

Year of Publish: 2011

Code: The code is implemented using Matlab and is made available at <http://people.csail.mit.edu/celiu/LabelTransfer/code.html>

Data: Datasets are made available along with the code in the same link.

VM's and Containers: No VM or Containers are made available.

Workflows: Details of how to implement the code and datasets is made available and also the system configuration is specified. How the experiment should be executed is provided in the experiment section of the paper and the intermediate results are also made available in the result section.

Provenance: The large database driven approaches for nonparametric methods was the motivation for this research.

Reproducible rating: 4 out of 5 since all the data, code and implementation details are provided but no virtual machines or containers exist.

Paper 4:

Reference:

Title: Image Question Answering using Convolutional Neural Network with Dynamic Parameter Prediction

About: The paper takes an input image and tries to answer the questions we may have about that image.

Authors: Noh, Hyeonwoo and Hongsuck Seo, Paul and Han, Bohyung

Journal: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

Paper URL:

http://www.cvfoundation.org/openaccess/content_cvpr_2016/papers/Noh_Image_Question_Answering_CVPR_2016_paper.pdf

Year of Publish: 2016

Code: Here standard xxhash code is implemented and can be found at <https://code.google.com/archive/p/xxhash/> . The code is available in different languages.

Data: Standard publically available datasets are used like DAQUAR ,COCOQA and VQA.

VM's and Containers: No virtual machines or containers are made available.

Workflows: Limited experimental details about datasets, code and implementation details are provided.

Provenance: This work was proposed to get closer to the holistic scene understanding. It extends the previous work which answered the basic questions like bird species, object, action etc. by answering questions like, How is the weather or Is this picture taken during the day etc.

Reproducible rating: 2 out of 5 is my evaluation of reproducibility, as there is very less documentation found which will make it hard to reproduce.

Paper 5:

Reference:

Title: A C++ Implementation of Otsu's Image Segmentation Method

About: The paper implements an algorithm named Otsu's image segmentation method to find the threshold that separates the image histogram into two classes.

Authors: Juan Pablo Balarini, Sergio Nesmachnow

DOI: 10.5201/ipol.2016.158

Journal: IPOL

Year of Publish: 2016

Code: The code is implemented in C++ language and is made available at http://www.ipol.im/pub/art/2016/158/?utm_source=doi

Data: There is no dataset; it's an implementation of the algorithm on any image.

VM's and Containers: No VM or containers are provided.

Workflows: The source code for the algorithm is provided along with the information on the implementation details. Also explains how the algorithm works.

Provenance: The research here is to emphasize the importance of mathematics in algorithm designs which are involved in image segmentation.

Reproducible rating: 4 out of 5, there is code and the implementation details are also provided.

Paper 6:

Reference:

Title: MoSIFT: Recognizing Human Actions in Surveillance Videos

About: The research here tries to recognize human actions obtained from surveillance cameras by coming up with an algorithm called MoSIFT.

Authors: Ming-yu Chen and Alex Hauptmann

Publisher: Carnegie Mellon University

Paper URL:

<http://www.cs.cmu.edu/~mychen/publication/ChenMoSIFTCMU09.pdf>

Year of Publish: 2009

Code: No code is provided but the whole software package is provided along with the .exe files and we just need to download the software and run it.

Data: User has to provide the data of any video in .avi format.

VM's and Containers: The software acts like a container which can be installed and run without a lot of efforts.

Workflows: The implementation details are provided along with the results obtained. How to use the software and other details are also provided.

Provenance: This research improves the accuracy obtained on the KTH dataset to 95.8 percent. Discusses about the other detector algorithms like 2D interest point detection algorithm and extending 2D harries corner detectors to 3D harries corner detectors.

Reproducible rating: 5 out of 5 since all we need to do is install the software and fallow the instructions and hopefully we will obtain similar results.