Deep Learning Projects (CSE 641)

1. Deep Learning for Mobile Platforms

OBJECTIVE
Current works in the area of object recognition and semantic segmentation rely on expensive computational resources like distributed CPU clusters and high end GPUs. Today's mobile platforms have become powerful enough to run traditional Machine Learning algorithms. Your task is to port an existing object detection framework (eg. SSD) to a mobile platform (preferably android). Evaluation will be done on the basis of the time taken for detecting objects, number of objects correctly detected, class confidence score and RAM usage. You may use different model compression techniques for benchmarking.

DATASET
You can use any publicly available dataset for training. The testing images would be a small set of images containing many objects taken from the Internet.

REFERENCES
1. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg “SSD: Single Shot MultiBox Detector”, ECCV 2016.
2. https://developer.android.com/studio/index.html
3. https://www.slideshare.net/anirudhkoul/squeezing-deep-learning-intomobile-phones

2. Road Sign Detection, Parsing and Understanding

OBJECTIVE
Your task is to detect road signs in an image and recognize the text present in it. The sign may contain both Hindi and English texts as well as road traffic symbols.

DATASET
Dataset Link: http://192.168.22.72/HindiSignboardDataset.zip
Any publicly available dataset can be used for training the model (sign board detection and recognition). The linked dataset has been created from Indian roads. This dataset does not contain any annotation. You can use this dataset to do your own evaluation. The test set is similar to the provided dataset.

REFERENCES
1. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg “SSD: Single Shot MultiBox Detector”, ECCV 2016.
2. http://caffe.berkeleyvision.org/gathered/examples/mnist.html
3. http://www.ee.surrey.ac.uk/CVSSP/demos/chars74k/

3. Smart Signal

OBJECTIVE
Your task is to train an agent to operate a smart traffic signal in order to maximize overall traffic flow. You will have to use/create a simple simulator as shown in this video. An existing simulator is Simulation of Urban Mobility (SUMO). Design the MDP that you need to solve using *only* the visual input and reward to train the agent using Deep RL.

EVALUATION

Compare your model's performance with a manually tweaked fixed time traffic signal. Your test cases should include different traffic densities and speeds and the evaluation metric should contain average travel time, average cumulative delay and average queue length as in [2].

REFERENCES

1. Mnih et al. Playing Atari with Deep Reinforcement Learning
2. Genders and Razavi, Using a Deep Reinforcement Learning Agent for Traffic Signal Control

4. Low Resolution Face Recognition

OBJECTIVE

To train a deep learning model that works with low-resolution face recognition. For this purpose, you can use the Labeled Faces in the Wild dataset and synthetically generate low-resolution images for training and testing your model.

DATASET

Labeled Faces in the Wild (LFW) with synthetically downsampled images. As is standard procedure, train and test splits should have different identities.

EVALUATION

Report accuracy and cumulative match characteristic (CMC) scores for different factors of subsampling. For the original resolution, the accuracy should meet the state-of-the-art (acc_{orig}) and for a lower-resolution image with a subsampling factor of *k*, the accuracy should reduce gracefully as $acc_{orig} - k*x$. Your goal would be to achieve the smallest x. Report accuracies for k= 1, 2 and 3.

REFERENCES

1. Erik Learned-Miller, Gary B. Huang, Aruni RoyChowdhury, Haoxiang Li, and Gang Hua. **Labeled Faces in the Wild: A Survey.**

5. Wildlife Detection and Recognition in Camera Trap Images

OBJECTIVE

The use of camera trap images for monitoring animals in the wild is on the rise. This results in a huge number of images, which have to be classified on the basis of species. Use the provided dataset to train an animal species detector.
Given an image, perform the following tasks:
1. Draw bounding boxes around the animals present in the camera trap images.
2. For each bounding box, classify the species of the animal present in it.
3. Report accuracies for the test set provided.

DATASET

To be provided.

REFERENCES:

1. Stefan Schneider∗ , Graham W. Taylor† , Stefan C. Kremer∗, "Deep Learning Object Detection Methods for Ecological Camera Trap Data", ArXiv, March 2018.
2. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg "SSD: Single Shot MultiBox Detector", ECCV 2016.

3. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”

6. Domain Adaptation for Image Segmentation

OBJECTIVE

Training a CNN based Semantic Segmentation model requires ground-truth annotations at the pixel level. Such annotations are readily available in synthetic datasets generated from game engines like Unity, Unreal, etc. However, upon training DNN models on synthetic data and testing it on real data, there is a significant domain shift. In this project, your goal is to use state-of-the-art unsupervised domain adaptation techniques for performing well on real-world images.

DATASET

Train on Synthetic dataset like Synthia and test it on Camvid
Download dashcam videos for Indian roads (e.g., from youtube) and show qualitative results on a test set.

REFERENCES

1. Sankaranarayanan et al. Learning from Synthetic Data: Addressing Domain Shift for Semantic Segmentation
2. Hoffman et al. FCNs in the wild: Pixel-level adversarial and constraint-based adaptation
3. Bousmalis et al. Unsupervised Pixel–Level Domain Adaptation with Generative Adversarial Networks

General Guidelines for All Projects

1. Decide your project by Friday, 13th April, 2018.
2. We will have in class presentation on 19th April, 2018. You will be required to present your model and training strategy.
3. The final presentations will be on May 4th, 2018, for which you will be submitting a short report on what you have done (with code, evaluation strategy and accuracy).