Assignment-3: Mid & High-Level Vision

Instructions:

[1] Any plagiarism will lead to award of **F grade** STRICTLY
[2] Use python only for the implementation of all the assignments
[3] Use NumPy to represent the vector and array
[4] Do not use the inbuilt functionality of any library including NumPy until suggest so.
[5] PyTorch must be used to implement the deep learning-based methods.
[6] One mark will be deducted for each late day.
[7] Submit via Moodle only. Email submissions won't be considered.

No.	Question	Marks
1	Data source: sample video from UCF101 Write a program to compute the optical flow for all the pairs of consecutive frames in the video and display the frames and optical flow in subplots. #use of library function for optical flow are not allowed.	10
2	Data source: MNIST	35
	I. Implement a 1-hidden layer neural network in NumPy, use sigmoid function as activation in all layers and train it using MSE and SGD with manually calculated gradients.	10
	II. Implement a 1-hidden layer autoencoder using tanh activation in hidden layer and linear in output layer in PyTorch.	5
	III. Implement a 3 layer MLP using PyTorch use sigmoid function as activation in all layers and train it using MSE and SGD with autograd.	5
	IV. Implement a 3 layer MLP using PyTorch use sigmoid function as activation in all layers and initialize the first 2-layers with pretrained weights as per the deep belief network using autoencoders then fine tune it using MSE and SGD with autograd. Use appropriate learning rate or a momentum on weights. Compare the results of the III and IV for training vs validation losses and accuracies. Plot the train and validation curves in a single figure.	10
	V. Implement LeNet and compare with the best of above.	5
3	Data source: CIFAR10 Write a program in PyTorch to	30
	I. Implement a 3 CONV + 1 FC layers CNN for classification and train with CIFAR10 dataset. Fine tune it for the filter size, activation function, pooling, batch normalization, data augmentation, and other model and optimization hyperparameters.	5

	II. Implement a program to train AlexNet, VGG16, GoogLeNet, ResNet152, EfficientNet-b1 from scratch on CIFAR10. Fine tune each for best hyperparameter. Compare the results losses and accuracies in two single figures each for train and val.	25
4	Data source: Sample images from ImageNet Dataset	25
	Write a program without any library to	
	I. Implement clustering-based segmentation algorithms using k-means, GMMs, Mean-Shift algorithms with various information used to form a graph. Compare the results on a sample image in a single figure for a single algorithm but different graphs. Then compare different algorithm on best performing graph.	15
	II. Implement Graph-based segmentation algorithms using Min-Cut and Normalized Min-Cut algorithms with various information captured in the feature vector. Compare the results on a sample image in a single figure for a single algorithm but different feature vector. Then compare different algorithm on best performing feature vector.	10