C-M-009: Machine Learning Lab II '21-22

Lab. 1 (Given December 4, 2021; Due December 6, 2021)

Your answers must be entered in LMS by midnight of the day it is due. You must submit the code as well as a self-contained PDF which has the approach, an explanation of the implementation, the output as well as anything else asked by the question. Marks devoted to this Lab. are indicated in the "Syllabus" sheet that was provided the first day of class.

---

1. We want to design a system that can help a visually impaired person know what is around them. At present, they use a "white cane" that is used to tap and feel is there is an obstacle. We want to use machine learning to make life slightly better for them. However, given the limitation of time we will consider a subset, though a core part of the problem.

   The posture we will adopt is that the mobile phone camera can acquire images at some periodic intervals or when the individual wants to know what is in the field of view of where the camera is being pointed. Assume that there are only 10 objects that are there in the world (there is no class which is "everything else") and the field of view may contain one of those 10 objects.

   The image is fed into a network that you have to design that can recognize what the object is. Hopefully, a speech synthesizer can then announce it to the individual but we will not concern ourselves with the speech synthesis part.

   Use the CIFAR-10 dataset `https://www.cs.toronto.edu/~kriz/cifar.html`. This dataset has 10 objects and each image has an object. In a real system we will have to contend with segmentation problem but let us assume that step has been done for us. Design a CNN that can recognize the object (the category can be fed into a speech synthesizer that can announce the object in the filed of view but that is not part of the scope). Report the 5-fold cross validation accuracy and comment on some practical difficulties that you may face in implement this on a mobile phone