W23 CMPUT 412/503 Exercise 5

ML for Robotics

Primary TA's: Rizwan and Justin
Written report due: March 26th
Oral demonstration: March 29th/31st

Number detection using ML

1.1 MNIST Dataset and ML basic terminologies

This part is about read-and-learn. You are assigned two readings: one about backpropagation, and the other about MLP classification on the MNIST dataset. The former helps to understand an essential concept in deep learning, while the latter provides an example of python coding and some common practices in deep learning training. Each reading is accompanied by questions that you will need to answer in your written report.

Deliverable 1. Please read the article "Backpropagation Step by Step.". This article demonstrates one pass of backward propagation and the parameter update. You may notice that after the first pass, the prediction error is reduced. Please do another pass of backward propagation (calculation of weights) based on the updated parameters from the first pass and show the reduction in prediction error after the second pass of backpropagation. Include your work in your written report.

Deliverable 2. Please study the notebook "Multilayer Perceptron.ipynb" uploaded on eClass. One can take advantage of the GPU resource of Google Colab to run the code. If the Colab GPU is used for initial ML setup, it takes less than 20 seconds to run one epoch during training. Please modify the code as outlined below and compare the classification performance and runtime before and after each modification. Discuss/explain your observations in your written report.

1. What data augmentation is used in training? Please delete the data augmentation and rerun the code to compare.
2. What is the batch size in the code? Please change the batch size to 16 and 1024 and explain the variation in results.
3. What activation function is used in the hidden layer? Please replace it with the linear activation function and see how the training output differs. Show your results before and after changing the activation function in your written report.
4. What is the optimization algorithm in the code? Explain the role of optimization algorithm in training process
5. Add dropout in the training and explain how the dropout layer helps in training.

You're required to answer all the above questions in detail in your report.

1.2 Number Detection Node

Use the knowledge gained in section 1.1 to train a model on the MNIST dataset and create a ROS node that can detect the digits [0-9] from your Duckiebot's camera real-time.

You will need to write a node that:

1. Subscribes to the camera image topic
2. Use your camera calibration file to un-distort the image
3. Converts the image to black and white
4. Efficiently **detects the digits** on the April tag and **print** the corresponding digit to the terminal along with the Apriltag's position in the world frame (Apriltag positions are same as the one we used for exercise 3) *Note: we use the tag36h11 April tag family*
5. Keep track of the detected digits and properly **terminate your program** once all ten digits [0-9] have been found around the town
    a. Each Apriltag from Exercise 3 will have a unique number attached to it
    b. Your robot should be able to start from any position within the town and start exploring to find all the numbers

Tips:

- You're free to use any of the available platforms or libraries available online to complete the task. Please cite them in your reference section.
- It's suggested to select one of the following platforms for implementing ML code: Tensorflow, Pytorch or Keras.
- You can make use of GPU resources from googlecolab if you need for this task

Deliverable 3.

- A screen recording of rviz and your terminal showing the camera image detecting the digit (i.e., bounding box around the detected digit) with the correct digit and Apriltag location printed to the terminal window
    - You can obey the following flow: detect digit -> stop driving-> print to terminal location of tag and what digit is detected -> continue driving

Deliverables

Include in your written report:

- Deliverables 1, 2, and 3 as described above
- A brief paragraph describing your implemented strategy for number detection using ML
- A brief paragraph discussing your results; please answer these questions:
    - How well did your implemented strategy work? Was it reliable? In what situations did it perform poorly?

On eClass one member of your team will submit:

- A pdf printout of your published report
- A link to your team written report hosted on your website
- A link to your exercise 5 repository

For the oral demonstration duckiebot will be placed in some random position by TA. You need to run your program and make the duckiebot move around the town and stop the program after it successfully detects the first digit (you can stop it manually during the demo). Please show the rviz and terminal screen printing detected digit to the TA.

Useful Resources

- https://data-flair.training/blogs/python-deep-learning-project-handwritten-digit-recognition
- https://data-flair.training/blogs/image-segmentation-machine-learning/
- https://github.com/guptajay/Kaggle-Digit-Recognizer/blob/master/Digit_Recognizer_MNIST.ipynb
- https://colab.research.google.com/

You can use any material on the internet as long as you cite it properly. You are encouraged to collaborate with your labmates and if you develop a solution together please acknowledge who you worked with in your written report.