exercise 01 - Morteza Soltanpour 4149938, Driton Goxhufi 4233242

Samstag, 9. November 2019   22:13

1 Network design (1+2+2+2+2+1=10 Points)
  a)
  b) Theoretical questions about the training function in the exercise:
      1. Why is it necessary to divide the data into batches?
          i. memory issues, depending on the dataset, a graphic card cant handle a full dataset at a time.
          ii. faster iterations and feedback of learning(network) progression
      2. What is an epoch?
          i. an epoch is one whole iteration through the Dataset. e.g. if we have a dataset of 20 datapoints = N and use batches of 5 datapoints = B --> N/B = 4 Iterations = 1 Epoch.
      3. What do lines 43 to 48 do?
          i. we calculate the prediction of a batch by the current network(output of the network) and save all predictions into the variable "batch_out"
          ii. compute the difference between the prediction of the network and the ground truth which results in the loss through a cross entropy calculation
          iii. after that we backpropagate the loss through the network and do a optimize(update) step and adjust the weights in the network
  c) Which values do the three controls take?
      1. steer [-1, 0, 1] for right
      2. gas: [0.5, 0.0]
      3. brake: [0.8, 0.0]
  e) Can you achieve better results when changing the hyper-parameters?can you explain this?
      1. Yes you can achieve better results
      2. We changed these hyperparameters:
          i. epochs: increasing the number of epochs increases the time for the network to learn from data. reached limits after >500 epochs because of convergence.
          ii. Batchsize: the Idea was to Increase the batchsize to the max of memory capacity. It should improve the (update)steps of the optimizer.
          iii. Learning rate: as we increased the batchsize, we were able to increase the learning rate aswell and do bigger learning steps. Because theoretically, the network can calculate a better gradient with a bigger batch of samples.
  f) what is 'good' training data?
          i. good training data is, when your dataset covers all situations in the task.
          ii. balanced classes is good too.
      1. Is there any problem with only perfect imitations?
          i. yes, because the network doesn't know how to behave if he is out of the track.
              1) perfect imitations doesnt tell the network how to act with a failure.

  2. Network Improvements (2+2+2+2+2=10 Points)
      b) compare classificationnetwork with multiclassnetwork:
          i. multiclassnet > classnetwork
      e) Fine-tuning
          i. we tried batchnormalization
          ii. changing network architecture
              1) adding more layers
              2) increasing featuremaps(depth of convlayers)
          iii. CONCLUSION:
              1) it improves!!

training curve