Pattern Recognition and Machine Learning

James L. Crowley

ENSIMAG 3 - MMIS Fall Semester
Lab Project 3: 6 January 2021

Face Detection with Convolutional Neural Networks

The objective of this project is to evaluate the effectiveness of convolutional neural networks for detecting faces in images. The written report for this exercise is due on 3 February 2021. Oral presentations by three project groups will be made on 27 January 2021.

Each programming team should

1) Implement a simple CNN similar to the example for recognizing MNIST digits shown on the course web site. This example has processes 28x28 pixel imagettes with a convolutional layer of 32 3x3 filters using relu, followed by 2x2 max pooling, a convolutional layer of 64 3x3 filters, using relu, followed by 2x2 max pooling, a flatten layer, dropout of 0.5 and a fully connected layer.
2) Use a balanced data set of face images to train your network over multiple epochs. Plot accuracy for your training set and an evaluation set and stop training when the accuracy scores diverge. Present results for Accuracy, Precision, Recall, F1, AUC for ROC curves, using a test set.
3) Demonstrate the effects of changing the number of convolutional layers, the number of filters per layer, the size of the filters at different layers, different types and sizes of pooling layers, number of fully connected layers, and dropout parameters by providing Precision, Recall, F1, AUC for ROC curves for each network.
4) Evaluate the effects of different optimizers on learning and accuracy.
5) Evaluate the effectiveness and computational cost for using your best networks to construct a sliding window detector.
6) Demonstrate the usefulness each network as a real time face detector using your computer web cam.
7) Demonstrate the use of your network for detecting face orientation, face identity and/or as a smile detector

Lab work will be reported with a written report in either French or English. Work will be evaluated based on the range of techniques tested, the effectiveness of the experimental evaluations, and the clarity and depth of the explanation of experimental results. You may use any framework or programming environment. The objective of this project is to evaluate effects of different variations in training data, training techniques and network architectures. Creativity and rigorous experimental method are encouraged and will be rewarded!

The following is an indicative barometer for grading. Actual grades will depend on a subjective appreciation for the amount of effort deployed and the depth of understanding displayed in the results.

Grade	Example of Criteria (Max grade is 20)
10	Implement a simple CNN similar to the example for recognizing MNIST digits shown on the course web site. Divide a balanced data set of face imagettes into a training set (80%), an evaluation set (10%) and a test set (10%). Use the training set to train your network over multiple epochs while displaying accuracy for your training set and your evaluation set. Stop training when the accuracy scores diverge. Present results for Accuracy, Precision, Recall, F1, AUC for ROC curves, using the test set.
+1 to +4	Using your initial network as a baseline technique, demonstrate the effects of changing the number of convolutional layers (+1) and the number of filters per layer (+1), different types and sizes of pooling layers (+1), and number of fully connected layers(+1), by providing Accuracy, Precision, Recall, F1, and AUC for ROC curves for each network.
+1	Demonstrate the effects for different values of dropout.
+1	Evaluate the effects of different optimizers and learning rates on learning and accuracy.
+1	Determine the effects of occlusion on your network by showing accuracy as function of percentage occlusion.
+1	Adapt your network for use as a smile detector.
+1	Adapt your network to detect if a person is wearing a face mask.
+1	Adapt your network to detect face orientation.
+1	Adapt your network determine the identity of a face from a small set of known faces
+1 to +5	Provide comparative performance evaluation of your best networks to any of the popular network architectures such as LeNet5, AlexNet, VGG, YOLO, etc by providing Accuracy, Precision, Recall, F1, AUC for ROC curves, using your training, evaluation and test set. (+1 point for each published architecture).
+1	Demonstrate the usefulness of one or more of your networks as a sliding window face detector for images. Discuss problems encountered in precisely estimating the size and position of the face in an image, and describe an evaluate solutions to these problems.
+1	Demonstrate and evaluate the usefulness of one or more of your networks for live images taken from your computer web camera.