Machine Learning Engineer Nanodegree
Capstone Proposal
Dalin Wang
December 20th, 2016

Detect and Recognize 2D shapes with Python and OpenCV

Domain Background

In this section, provide brief details on the background information of the domain from which the project is proposed. Historical information relevant to the project should be included. It should be clear how or why a problem in the domain can or should be solved. Related academic research should be appropriately cited in this section, including why that research is relevant. Additionally, a discussion of your personal motivation for investigating a particular problem in the domain is encouraged but not required.

There is already numerous research related to this topic. Tools and techniques are readily available in Python and OpenCV. OpenCV is an open source computer vision library used for image processing and feature detection of images and videos. Historically, applications have been created to facilitate drawing, security screening in airport. The intended goal is to create an intelligent agent that is able to recognize shapes and draw close contours in an image. My motivation is to later use this agent create a system that help kids learn shapes.

Problem Statement
(approx. 1 paragraph)

In this section, clearly describe the problem that is to be solved. The problem described should be well defined and should have at least one relevant potential solution. Additionally, describe the problem thoroughly such that it is clear that the problem is quantifiable (the problem can be expressed in mathematical or logical terms) , measurable (the problem can be measured by some metric and clearly observed), and replicable (the problem can be reproduced and occurs more than once).

The problem I am going to solve is to create a two-dimensional shapes detection and recognition system that can first, recognize shape and separate it from the background; then, closely draw boundaries around the shape; third, identify the shapes, namely, whether the shape is a circle or a rectangle, etc. I am going to use OpenCV-Python for this project, and I am going to try SVM and Neural-Networks on this problem and see how well each algorithm performs. The success of the project can be measured by the accuracy of the predictions in the testing set. The training and testing sets come from BabyAIShapesDataset(http://www.iro.umontreal.ca/~lisa/twiki/bin/view.cgi/Public/BabyAIShapesDatasets). The dataset is made up of several training and testing sets generated by corresponding python files. The datasets I am of particular interest of are shapeset2_1cspo_2_3.10000.train.py, shapeset2_1cspo_2_3.5000.valid.py, and shapeset2_1cspo_2_3.5000.test.py. The problem can be easily reproduced using either of three datasets by passing the labeled data to the classifier system.

Datasets and Inputs
(approx. 2-3 paragraphs)

In this section, the dataset(s) and/or input(s) being considered for the project should be thoroughly described, such as how they relate to the problem and why they should be used. Information such as how the dataset or input is (was) obtained, and the characteristics of the dataset or input, should be included with relevant references and citations as necessary It should be clear how the dataset(s) or input(s) will be used in the project and whether their use is appropriate given the context of the problem.

The BabyAIShapesDataset datasets are composed of python files that can generate 32 * 32 images with one shape and random uniform background (one color for the shape and one color for the background). The shapes vary in position, size, color and position. The particular datasets of shapes I will be using are shapeset2_1cspo_2_3.10000.train.py, shapeset2_1cspo_2_3.5000.valid.py and shapeset2_1cspo_2_3.5000.test.py. The training set consists of 10000 samples, the validation set consists of 5000 samples, and the testing set has 5000 samples. The shapes are composed of general shapes such as rectangles, ellipses and arbitrary triangles. I intend to preprocess the images to black and white with the shape being black and background as white and then feed the processed image to my classifier.

Solution Statement
(approx. 1 paragraph)

In this section, clearly describe a solution to the problem. The solution should be applicable to the project domain and appropriate for the dataset(s) or input(s) given. Additionally, describe the solution thoroughly such that it is clear that the solution is quantifiable (the solution can be expressed in mathematical or logical terms) , measurable (the solution can be measured by some metric and clearly observed), and replicable (the solution can be reproduced and occurs more than once).

I am going to employ openCV with Python and machine learning algorithms such as SVM and Neural Networks to tackle this problem. I will use the training and validation sets for training, and the testing set for testing. The detail of the datasets is outlined in the question above. The SVM method map the input space to into a high-dimensional feature space and try to find the hyperplane that best separates different types of shapes. For the neural network method, I plan to try out different parameters like the number of hidden layers, and some different transfer functions. The solution is quantifiable as we can calculate the accuracy of the prediction.

Benchmark Model
(approximately 1-2 paragraphs)

In this section, provide the details for a benchmark model or result that relates to the domain, problem statement, and intended solution. Ideally, the benchmark model or result contextualizes existing methods or known information in the domain and problem given, which could then be objectively compared to the solution. Describe how the benchmark model or result is measurable (can be measured by some metric and clearly observed) with thorough detail.

Based on the characteristics of datasets which contain three types of classification. The benchmark model states that the accuracy of the prediction should be no lower than 33.3%, that is, betting on one of the three classes for all samples. Based on this research paper in 2013, the researchers are able to use simple metrics like ratio of area of the object to its bounding box to obtain a near 100% accurate predicting model for simple shapes (https://pdfs.semanticscholar.org/7dce/cc82b4acc57c692d842817e3726b77480654.pdf). I am more inclined to avoid using pre-defined metrics and have the intelligent agent figure out the metrics itself from reading large amount of images. However, I will do some image processing prior to feeding the images to the intelligent agent.

Evaluation Metrics
(approx. 1-2 paragraphs)

In this section, propose at least one evaluation metric that can be used to quantify the performance of both the benchmark model and the solution model. The evaluation metric(s) you propose should be appropriate given the context of the data, the problem statement, and the intended solution. Describe how the evaluation metric(s) are derived and provide an example of their mathematical representations (if applicable). Complex evaluation metrics should be clearly defined and quantifiable (can be expressed in mathematical or logical terms).
The evaluation metric is the accuracy of the predictions. All correct predictions of 2D shapes divided by the attempts of predictions. The metric is selected based on the nature of the data itself. The data, as examined, is balanced distribution over three classification classes: ellipse, rectangle and triangle.

Project Design
(approx. 1 page)

In this final section, summarize a theoretical workflow for approaching a solution given the problem. Provide thorough discussion for what strategies you may consider employing, what analysis of the data might be required before being used, or which algorithms will be considered for your implementation. The workflow and discussion that you provide should align with the qualities of the previous sections. Additionally, you are encouraged to include small visualizations, pseudocode, or diagrams to aid in describing the project design, but it is not required. The discussion should clearly outline your intended workflow of the capstone project.
The project requires a lot knowledge in image processing. So first, I will learn how to detect shape edges and draw contours around the shapes. And then transform the image to black and white with black shapes and white background. I intend to use this simplified image as the input feature to my classifier. Then I should apply the two algorithms, support vector machine and neural networks to the input space to classify the shapes.

Before submitting your proposal, ask yourself. . .

- Does the proposal you have written follow a well-organized structure similar to that of the project template?

- Is each section (particularly **Solution Statement** and **Project Design**) written in a clear, concise and specific fashion? Are there any ambiguous terms or phrases that need clarification?
- Would the intended audience of your project be able to understand your proposal?
- Have you properly proofread your proposal to assure there are minimal grammatical and spelling mistakes?
- Are all the resources used for this project correctly cited and referenced?