Project Instructions

This project shows how advancement of data science deprecating security technologies. Specifically, we will using deep learning to break a CAPTCHA system called Really Simple CAPTCHA, which is a WordPress.org plugin. Try a demo here: https://contactform7.com/captcha/

CAPTCHA stands for "completely automated public Turing test to tell computers and humans apart", which is used to verify that a user is not a bot.

We require you to have some background knowledge in (1) Python programming and (2) Jupyter Notebook. If you do not already have the background, please check the following links for self-study:

- A quick tutorial on Python: http://cs231n.github.io/python-numpy-tutorial/
- A quick tutorial on Jupyter Notebook: http://cs231n.github.io/ipython-tutorial/
- How to set up your Python and Jupyter Notebook environment: http://cs231n.github.io/setup-instructions/

In this project, you will learn and use Python libraries including OpenCV, Scikit-learn, TensorFlow and Keras. OpenCV is used for image processing, Keras on top of TensorFlow is used for building deep learning models, and Scikit-learn is a machine learning library and we will use its preprocessing module.

Please open assignment.ipynb in the assignment folder, and following the instructions step-by-step to complete this notebook file. Most of the codes are provided but you are required to complete the code where you see

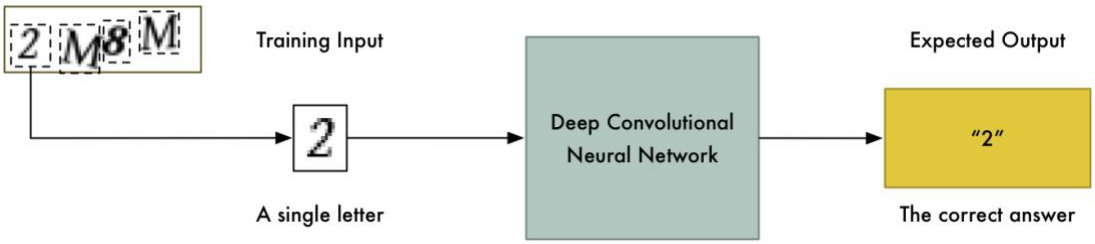
```
#########################################################################
# TODO: your code here ...                                              #
#########################################################################
```

You may also see some questions that you need to answer, marked by:

Type Your Answer Here:

The notebook file provides instructions that are self-explanatory, and can be divided into 3 steps:

1. Step 1: Extract single letters from the training CAPTCHA images provided;



2. Step 2: Train the neural network to recognize single letters, where the network is built using Keras;
3. Step 3: Use the model to break CAPTCHA images not seen before.

You are expected to answer all questions and complete all codes to fill in, and save the notebook file as an HTML file: