

---

# Image Captcha Recognition

## Project Proposal

---

**Xiaoguang Wang**  
A53246911  
xiw503@ucsd.edu

**Qian Wang**  
A53235276  
qiw018@ucsd.edu

**Hao Zhang**  
A53234732  
haz351@ucsd.edu

### 1 Objective

Typically, when you are registering or logging into a website, you are required to input some verification code consisting of several digits of numbers or characters from a blur captcha image. It's annoying sometimes because even for a human, it's confusing to make out what is exactly on the captcha image. Here, we are going to building a system which can automatically recognize the content on the captcha images (in particular required format).



Figure 1: Some examples for captcha images

### Some primary ideas:

Here, in the automatic recognition system, we pretend to use a pretrained model (like GoogLeNet and VGG) to extract images feature from the captcha images since they have been well pretrained on large images dataset. Then we will employ soft-max layer to predict the exact number or character on each digit separately.

## 2 Dataset

We have searched the opening datasets from DSMLP, but we cannot find a suitable dataset (one at least should have sufficient captcha images and corresponding labels to train our network). But luckily enough, there is a Python library called *captcha* (<https://pypi.org/project/captcha/>), we can use the library to build our own dataset (including training set and test set). I have finished a demo to generate some captcha images.



Figure 2: Running demo for python captcha

We can see for each captcha image, we set their file name to be the content on the image, which serve as the supervised information(labels) during the training process. Also, since the dataset is generated and can be controlled strictly on the image format and dataset size, which is very convenient to verify our ideas and model.

61

62

### 63 **3 Challenges**

64

65 There still existing some obstacles which can be seen from this point before we actually get  
66 started to implement the system. We have also come up with some potential solutions.

#### 67 **3.1 Image noise**

68 As we can see from the demo captcha images, there are a lot of interference noise like spots  
69 and lines on theses images, not absolute numbers and characters. These noises may greatly  
70 influence the system performance if not handled properly.

71 **Some potential solution:** We can perform some noise cancellation operations from  
72 traditional CV like Gaussian Low Pass Filter before feeding them into CNN to compensate  
73 for this interference.

74

#### 75 **3.2 Multi-digits output**

76 Also, there are more than one verification code digits on each image. So, how to predict  
77 these digits separately is another challenge.

78 **Some potential solution:** At this point what we can images to solve this problem is to  
79 borrow the idea from multi-task learning, where we can design several output layer(soft-max)  
80 and use one layer to predict one digit separately.

81