CAPTCHA Images Decoder

Text-based CAPTCHA

Voice-over: Ze Hui Peng

Group Members:
- ❖ Ze Hui Peng
- ❖ Lijiangnan Tian
- ❖ Zijie Tan
- ❖ Yuxi Chen

CONTENTS

- ❖ **Introduction**
- ❖ **Data**
- ❖ **Our four algorithms: CNN, RNN, SVM, KNN**
- ❖ **Comparisons of results**
  - ➢ **Two segmentation-free algorithms**
  - ➢ **Two segmentation-based algorithms**
- ❖ **Conclusion**

Introduction

❖ CAPTCHA (*Completely Automated Public Turing test to tell Computers and Humans Apart*) has been a popular method to distinguish robots from human users

❖ Our goal is to use different algorithms to decode/recognise the text-based CAPTCHA images and compare their performance

Data

❖ **10,000 CAPTCHA images of size 200 × 50 pixels[1]**

➤ Each has darkened background & 6-character-long string

➤ Including 2 type of noise: shadow and fisheye effect

❖ **Generated by the Google Kaptcha Library[2][3]**

➤ In the name `{content}_{index}.jpg`

❖ **Supervised classification**

➤ Multiclass Labels (36 classes, including 10 digits & 26 letters)

A sample CAPTCHA image with content **882m62**

Methodology

❖ **Segmentation-free models**
  ➢ **Convolutional Neural Network (CNN)**
  ➢ **Recurrent Neural Network (RNN)**

❖ **Segmentation-based models**
  ➢ **Support Vector Machine (SVM)**
  ➢ **k-Nearest Neighbours (KNN)**

❖ **Measurement**
  ➢ **Performance is measured by <u>accuracy</u>**
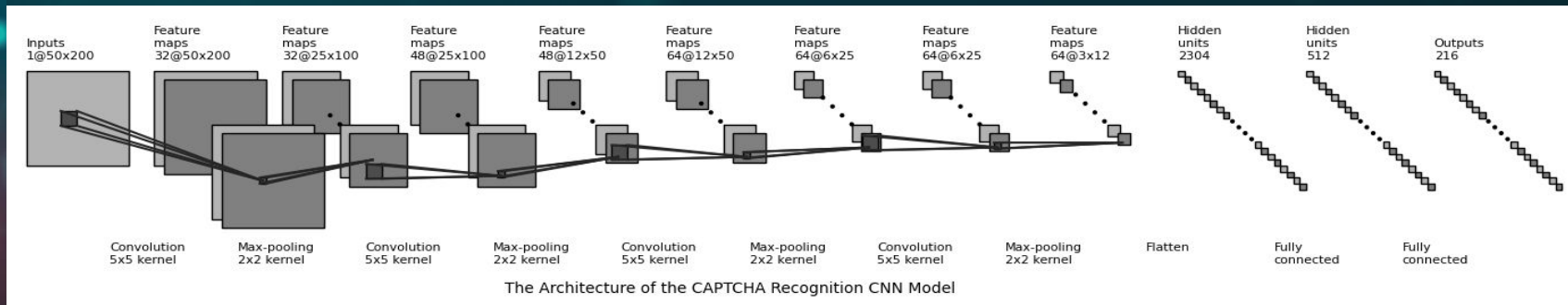 ➢ **Correctly predict all characters in a CAPTCHA?**

Segmentation-free method: CNN

❖ **Convolutional Neural Network**

  ➢ **Reduce large data volume images to small data volume**

  ➢ **Retains image characteristics**

❖ **Model**

  ➢ **ReLU activation function**

  ➢ **4 Convolution Layers, 4 Max-pooling Layers, 2 Fully-connected Layers**

  ➢ **Outputs a sequence of length 216 (6×36)**

| Inputs 1@50x200 | Feature maps 32@50x200 | Feature maps 32@25x100 | Feature maps 48@25x100 | Feature maps 48@12x50 | Feature maps 64@12x50 | Feature maps 64@6x25 | Feature maps 64@6x25 | Feature maps 64@3x12 | Hidden units 2304 | Hidden units 512 | Outputs 216 |

| Convolution 5x5 kernel | Max-pooling 2x2 kernel | Convolution 5x5 kernel | Max-pooling 2x2 kernel | Convolution 5x5 kernel | Max-pooling 2x2 kernel | Convolution 5x5 kernel | Max-pooling 2x2 kernel | Flatten | Fully connected | Fully connected |

The Architecture of the CAPTCHA Recognition CNN Model

Segmentation-free method: CNN

❖ **Flow of Data**

➤ **The model receives image data as input**

➤ **Outputs a sequence of length 216 (6×36)**

➤ **Sequence divided into 6 equal-sized segments (36, the same as the number of labels)**

➤ **Obtain the index of the label according to the index of the *maximum* value in segments**

The Architecture of the CAPTCHA Recognition CNN Model
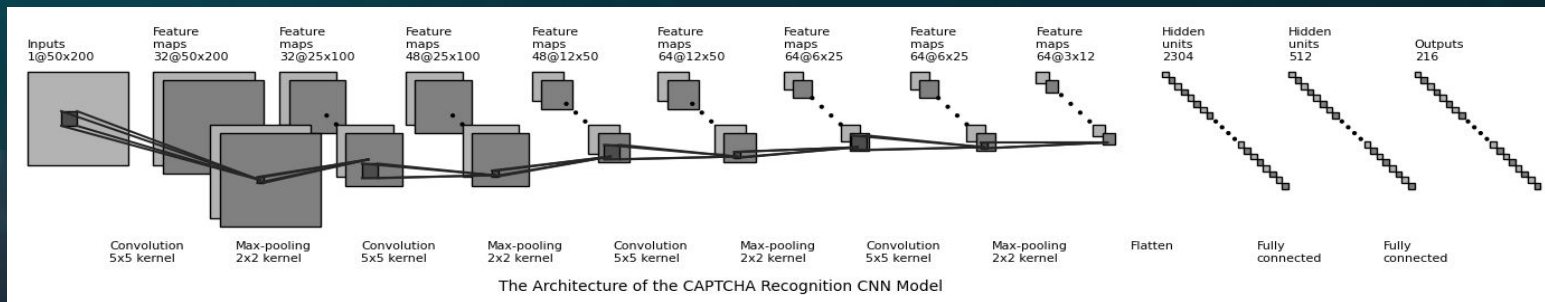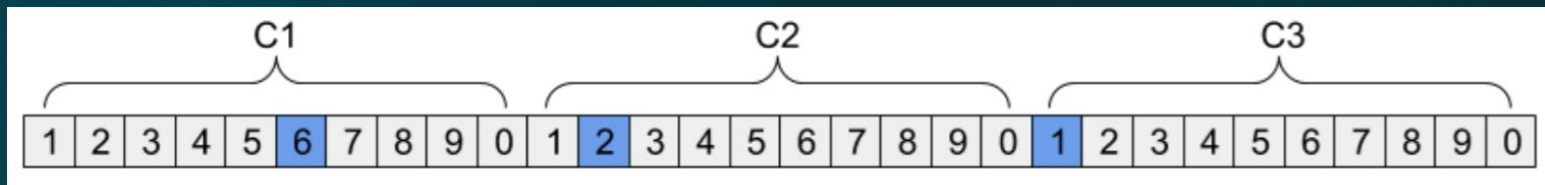
Segmentation-free method: RNN

❖ **Recurrent Neural Network**

➤ **Potential to solve *variable-length* CAPTCHAs**

❖ **However**

➤ **Traditional RNN has poor performance on such an Image Recognition problem**

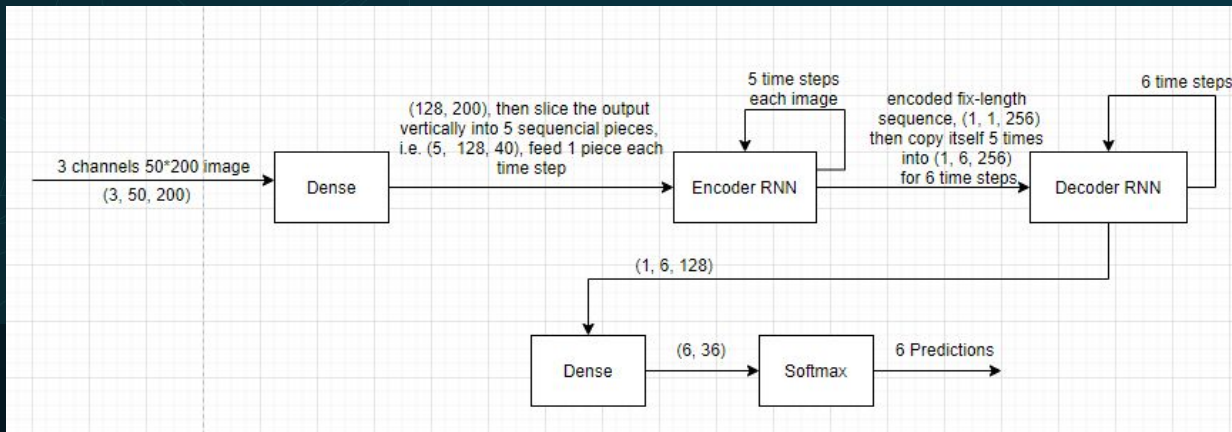■ **Accuracy 0.02% after 100 epochs**

➤ **Variation: Encoder-Decoder RNN**

Segmentation-free methods: RNN
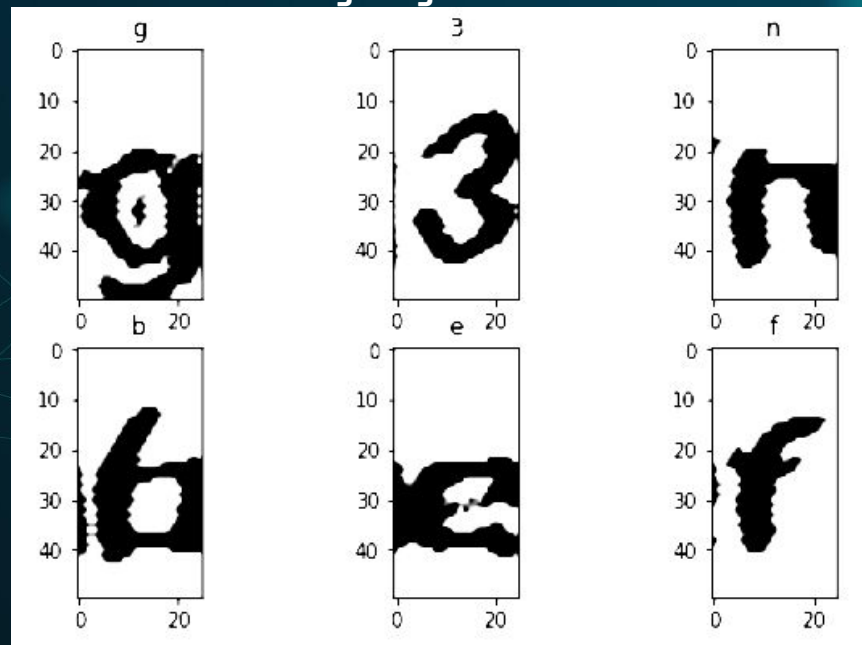
❖ **Encoder-Decoder Model (seq2seq)**
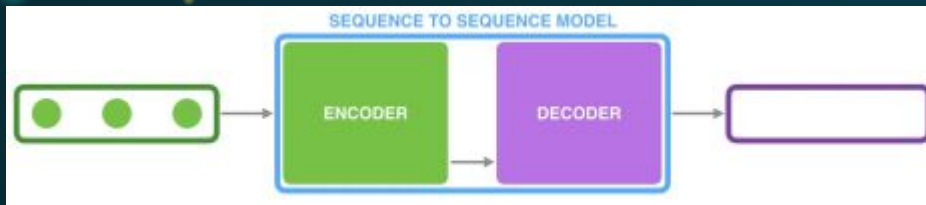
❖ **Flow of data**

Flow of Data for SVM and KNN

Image segmentation

Segmentation-based methods: SVM

❖ Support Vector Machine

❖ Tool used: `sklearn.svm.SVC`

❖ *Segmentation* is required: The CAPTCHA images need to be first segmented into single alphanumeric characters, thus a more complex image preprocessing is required than previous.

Segmentation-based methods: KNN

- ❖ k-Nearest Neighbours
- ❖ **Tool used:** `sklearn.neighbours.KNeighborsClassifier`
- ❖ **Classifies each character based on the most common class of its k closest distance neighbours**
- ❖ **Distance is calculated using Euclidean Distance based on the image pixels**
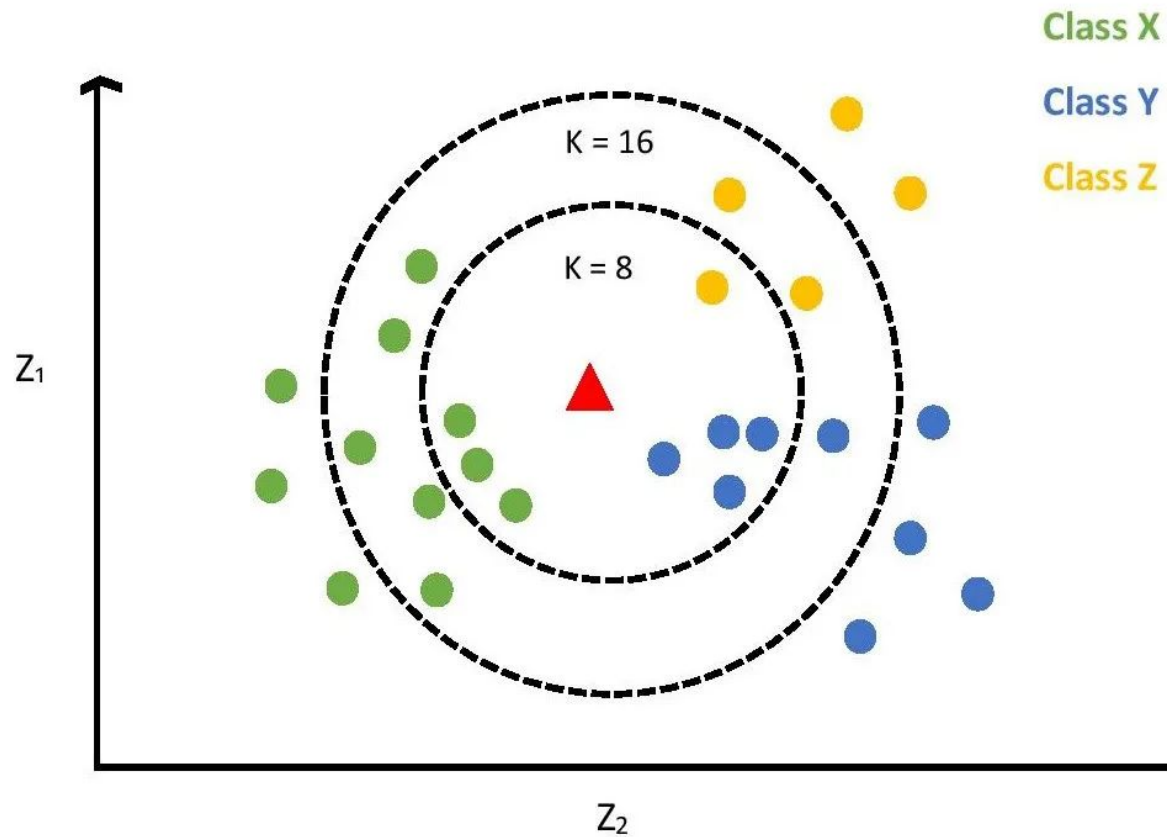
Results: segmentation-free

❖ **CNN model has considerably better performance than RNN on this problem**

❖ **Possible reason:**

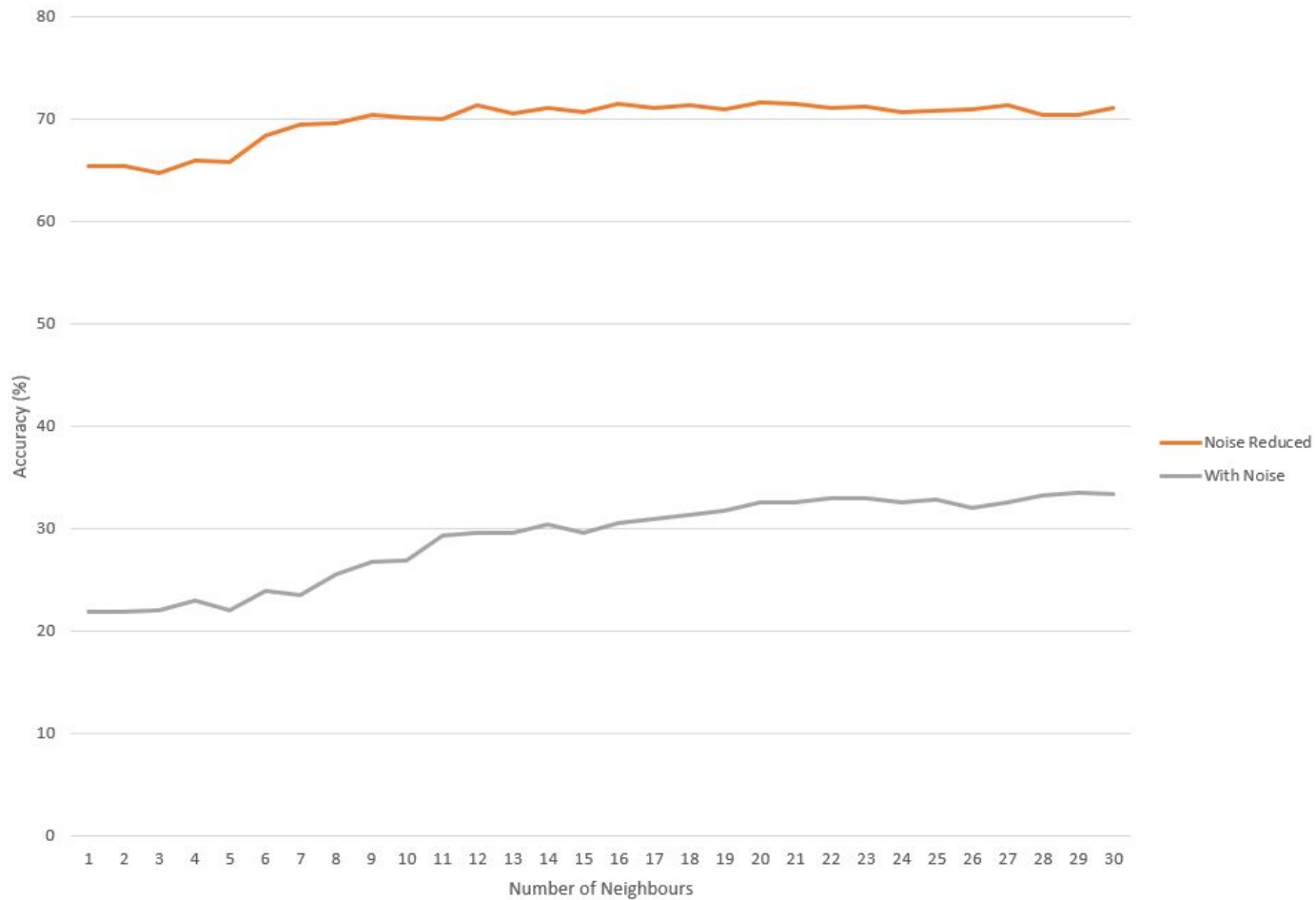➢ **Spatial properties of CAPTCHA images?**
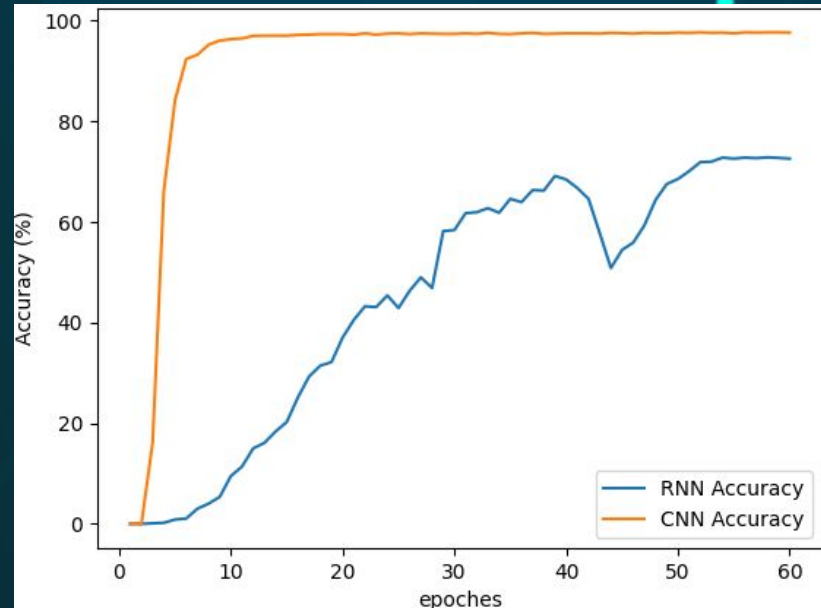
Result: segmentation-based

❖ **For SVM, accuracy can reach around 80%. The *kernel type* plays a particularly important role in the accuracy**

❖ **For KNN, maximum accuracy is around 72% with k = 20 neighbours, but the algorithm is highly susceptible to *noise***
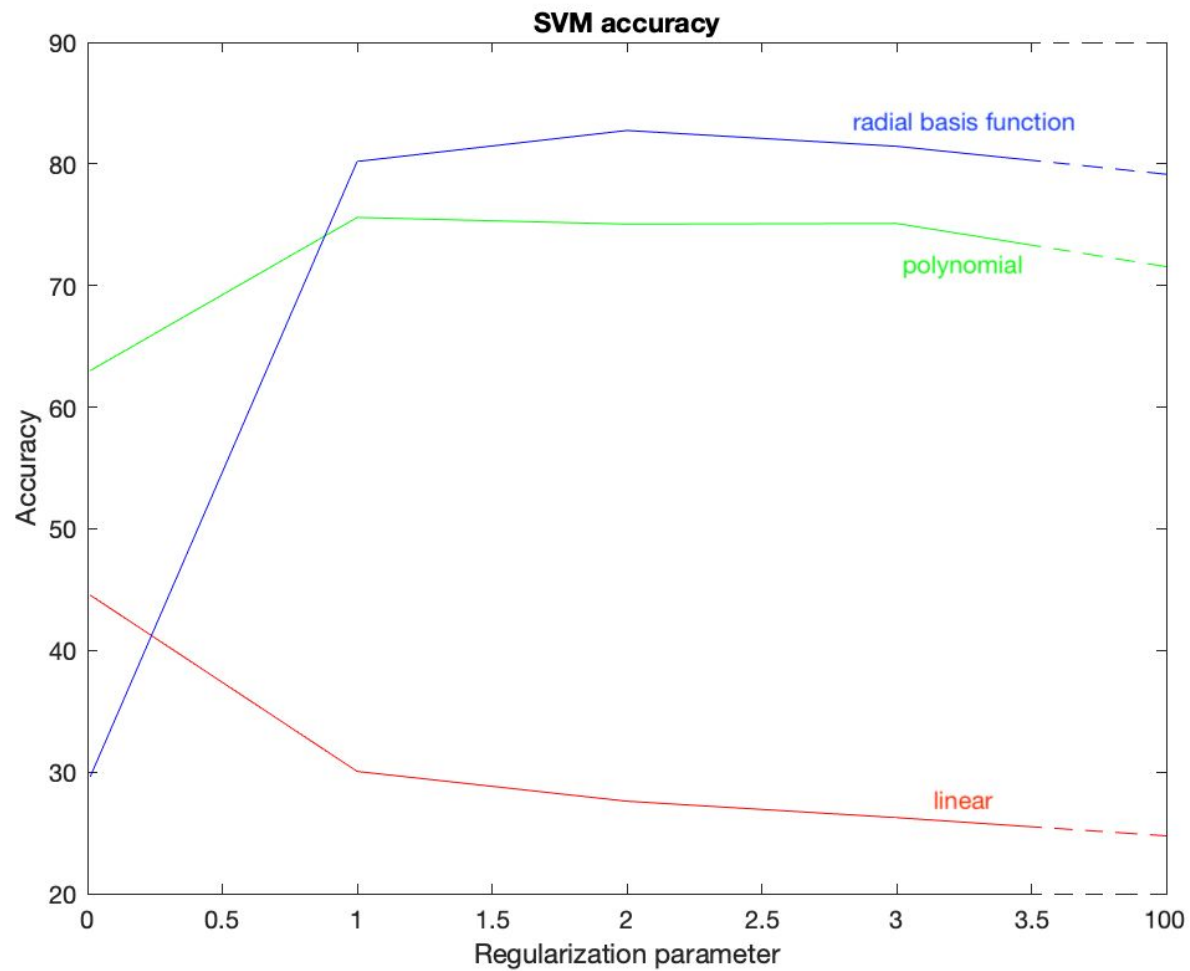
KNN performance chart

Discussion, Conclusions, Future work

CNN	RNN	SVM	KNN
Champion Best performance Robust	Hope of variable-length CAPTCHA	Good performance even with rudimentary segmentation algo	Good performance with little noise

Thank You for watching!

☺

Group Members:
- ❑ Yuxi Chen
- ❑ Zijie Tan
- ❑ Lijiangnan Tian
- ❑ Ze Hui Peng

Voice-over: Ze Hui Peng

RESOURCES

1. https://www.kaggle.com/ethan404/captcha6digits
2. https://code.google.com/archive/p/kaptcha/
3. https://github.com/Ethan707/CAPTCHA-Generator