Team Orange (Team 5)

Final Project Proposal

Ruoshui Li, Yi Feng, Jennie Sun, Shangwen Yan, Michael Tang

Project Title

Text Detection And Extraction Using OCR (Optical Character Recognition)

Question of Interest

Traditionally, text recognition in images involves intensive human labour and could be prone to human errors. Thus, this process is often inefficient and very costly to perform. Therefore, we are interested in exploring the implementation of OCR - a widely used computer vision-based application for text recognition. More specifically, it allows computers to automatically detect and extract alphanumeric data from image data without human intervention.

Project Motivation

The OCR focuses on how to get computers to locate and recognize textual data that appears in different parts of an image. OCR can not only help computers understand images better, it also has many other applications, such as assisting people that have vision impairment with a camera to identify the words on the street signs in front of them to prevent potential dangers. Researchers have also put in a tremendous amount of effort in the autonomous driving space, where OCR can help read speed limits to augment the car navigation systems and further enhance safety (e.g. a car reads the street signs to help drivers navigate to the destination).

Below lists a wide range of real life applications of OCR:

- Passport recognition
- Speed limit detection in autonomous driving
- Converting handwritten and typed texts to digital text
- Recipient address recognition in mail sorting
- Automatic Electricity Meter Reading
- Automatic license plate recognition
- Assistance for people who have impared vision
- Preprocessing for other NLP-related tasks

Datasets

- OCR handwritten characters dataset: http://ai.stanford.edu/~btaskar/ocr/
- Handwritten digits dataset:http://yann.lecun.com/exdb/mnist/

- The UFPR-AMR Electricity Meter Dataset: https://web.inf.ufpr.br/vri/databases/ufpr-amr/
- The Uber Text Dataset: https://s3-us-west-2.amazonaws.com/uber-common-public/ubertext/index.html
- Number Plate Dataset: https://platerecognizer.com/number-plate-datasets/
- OCR Open Dataset: https://github.com/xylcbd/ocr-open-dataset
- US Speed Limit Sign Dataset: https://www.kaggle.com/shaktip258/cropped-speed-limit-signs

Machine Learning Approach

We plan to implement a Convolutional Neural Network (CNN) algorithm and apply specific tasks including image recognition and classification for OCR, as CNN is the state-of-the-art technique for analyzing multidimensional signals such as images, and has the ability to capture the spatial dependencies in the input image through the application of relevant filters.

In order to perform photo OCR, we plan to implement the following machine learning pipeline below:

- **Model Training**: This step is needed to train a CNN-based classifier that could predict the class (0 - 9, A - Z) of a character with reasonable accuracy.
- **Image Preprocessing**: This step involves preprocessing image data so that the model could run smoothly.
- **Text Detection**: We will go through the images and find where there is text in the images.
- **Character Segmentation**: We will do character segmentation and segment it out into the individual characters.
- **Character Classification**: Finally, having segmented out into individual characters, we can then run a classifier. By doing so, eventually our algorithm will be able to recognize the words that appear in the images.

Performance Evaluation

Measuring OCR accuracy is done by taking the output of an OCR run for an image and comparing it to the original version of the same text. We will then either count how many characters were detected correctly (character-level accuracy), or count how many words were recognized correctly (word-level accuracy). Computational speed can also be used as a performance evaluation metric.

Team Roles and Responsibilities

- Yi Feng: 1-A, 2-A
- Ruoshui Li:1-B, 2-B
- Shangwen Yan: 1-C, 2-C
- Michael Tang: 1-D, 2-E

- Jennie Sun: 1-E, 2-D

List of References

- Understanding of a Convolutional Neural Network: https://www.researchgate.net/profile/Saad-Albawi/publication/319253577_Understanding_of_a_Convolutional_Neural_Network/links/5ad26025458515c60f51dbf9/Understanding-of-a-Convolutional-Neural-Network.pdf
- Optical Character Recognition: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.673.8061&rep=rep1&type=pdf
- Convolutional Neural Networks for Automatic Meter Reading: https://arxiv.org/pdf/1902.09600.pdf
- Deep Learning Detection of Inaccurate Smart Electricity Meters: https://arxiv.org/ftp/arxiv/papers/1907/1907.11377.pdf
- Automatic reading of domestic electric meter: an intelligent device based on image processing and ZigBee/Ethernet communication https://dl.acm.org/doi/abs/10.1007/s11554-013-0361-2

Weekly Meeting Time

- Every Saturday 7-10 PM, February 27 to April 17