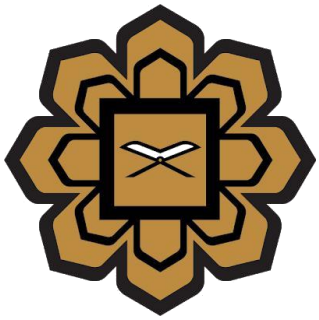


الجامعة الإسلامية العالمية ماليزيا
INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
يونيورسيتي إسلام انتارابغسا ملايسيا

Proposal Paper On:

"Image To CODE''

Supervised By:

AMELIA RITAHANI BT. ISMAIL

Organized by

Sanjir Inam Salsabil

Table of Contents

Introduction

This project is about converting images with text-code to text, the dataset will be collected from several sources such as Google images, Pixabay etc. Image processing includes some basic operations namely image restoration/rectification, image enhancement, image classification etc. Image classification forms an important part of image processing. The objective of image classification is the automatic distribution of the image to topical classes. There are two types of classification are supervised classification and unsupervised classification. The process of image classification includes two more steps, training of the system followed by testing. The training process intends to take the characteristic properties of the images (from a class) and form a unique description of a particular class.

The process will be done for all classes relying on the type of classification problem; binary classification or multi-class classification. The testing step means to categorize the test images under several classes for which system has been trained. This assigning of a class will be done based on the separation between classes based on the training features.

Project Objectives

1. To study a machine learning model for image classification.
2. To collect the data obtaining text-code.
3. To apply a deep learning algorithm (**convolutional neural network**) for image processing.
4. To analyze the results based on the model developed.

Exploring a different kind of Machine learning problems and algorithms. Deep understanding of the neural network ML algorithm. Identify the difference between the fully connected and the convolutional neural network. Using the CNN with the problems involving image data.

Expected Output/Results

1. A data set of text-code image.
2. Converted text-code images to text.
3. Image classification with convolutional neural networks model

A machine learning model which is able to produce a completely accurate source file containing the code which is presented in the images. Either this code was digital one or handwritten code.

LITERATURE REVIEW

Texts in an image directly carry high-level semantic information about a scene, which can be used to assist a wide variety of applications, such as image understanding, image search and indexing, navigation, and human-computer interaction. A number of approaches for text detection in images have been proposed into the past. Automatic detection and translation of text in images done using different techniques proposed. These methods aim to detect the characters based on the general properties of character pixels.

- It is a challenging task to detect and segment text from captured images due to two main issues:
  - Different variety of text patterns like sizes, fonts, orientations, colors, and
  - presence of background outliers similar to text characters, such as windows, bricks, and character-like texture

Digital image processing

processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be applied to the input data.

Image processing is a process to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal released in which input is an image, a video frame or photograph and output may be image or characteristics associated with that image.

Text Detection on images

Many recent methods have been proposed to design a better feature representation and models on text detection on images. A research paper from Stanford University by Adam Coates said that there are two key components of the most system which are (i) text detection from images and (ii) character recognition. They applied a method that developed recently in machine-learning which is the large-scale algorithm and it about learning the features automatically for unlabeled data and show that they allow to construct highly effective classifier for both detection and recognition to be used in high accuracy end-to-end system and they come to conclusion that larger banks of features they achieved increasing in accuracy with top performance compared to other systems.

Minhua Li, Meng Bai also proposed that to detected text on images based on image complexity analysis. And this approach adopts an image complexity analysis step to classify image complexity into three categories: low complexity, middle complexity, and high complexity. Then images with different complexity adopt different methods to extract image edges. The proposed text detection method takes a course to fine

detection strategy which combines the edge-based method, connected component-based method, and the texture-based method into a framework.

Jain and Zhong use a neural network (NN) to discriminate between text, graphics, and halftones in document images. Zhong analyzes local spatial variations in a gray-scale image and locates regions with a high variance as texts. They also combine a texture-based method with a CC-based method. Li extract the wavelet features of small windows in images, then classify them using NNs. Texture-based methods are known to perform well even with noisy, degraded, textured, or complex texts and backgrounds, however, they are usually time-consuming as texture classification is inherently computationally dense.

Synthesis Analysis -

A text information in images serves as an important clue in different applications. It provides instructions for assistive reading and content-based image retrieval and so many applications.

TEXTS in images include useful information for the automatic annotation, indexing, and structuring of images. Text detection is the process of detecting and locating those regions that contain texts from a given image and is the first step in obtaining textual information. However, text variations related to size, style, orientation, and alignment, as well as low contrast and complex backgrounds make the problem of automatic text detection extremely challenging.

Experimental Setup

Step 1_ Collecting the Dataset
In order to train our model, we need a huge amount of data so that our model can learn from them by identifying out definite relations and common features related to the objects.

Step 2 —Splitting the Dataset
we will split our data into two parts training and test sets.

Step 3 — Building the CNN
This is the most important step for our network. It consists of three parts:

- Convolution
- Polling
- Flattening
- The main purpose of Convolution is to produce features from the input image. Convolution preserves the locative relationship between pixels by learning image features using small squares of input data.
- Pooling minimizes the dimensionality of each feature map but keeps the most important information.
- The flattening step is necessary so that you can make use of fully connected layers after some convolutional layers. Fully connected layers do not have a local limitation like convolutional layers. This means you can merge all the found local features of the previous convolutional layers.

Step 4 — Full Connection
Full connection is a step where our convolutional network is connecting to a neural network and then compiling our network.
We will make 2 layers of the neural network with a sigmoid function as an activation function for the last layer as we need to find the probability of the object.

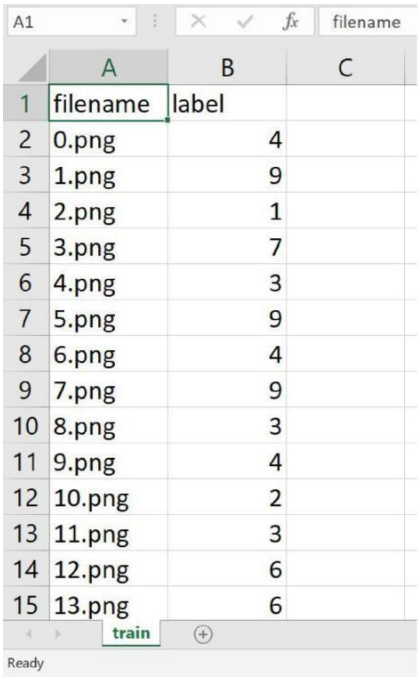
Step 5 — Data Augmentation
While training our data, we need a lot of data to train upon. To get more data, we just need to make small alterations to our existing dataset. Small changes such as flips or rotations. Our neural network would think these are discrete images anyway.
Data augmentation is a way we can reduce overfitting on models, where we increase the amount of training data using the information only in our training data. The domain of data augmentation is not new, and in fact, a lot of data augmentation techniques have been applied to specific problems.

Step 6 — Training our Network
After completing all the steps of construction, its time to train the model.

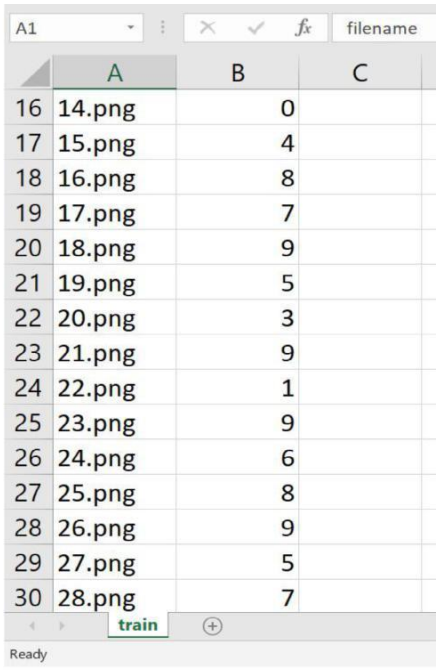
Step 7 — Testing
We will start to test our model with random images.

So, after creating a simple Image Recognition Classifier. The same concept can be applied to a diverse range of objects with a lot of training data and an appropriate network.

Sample of Dataset (.csv)



	A	B	C
1	filename	label	
2	0.png	4	
3	1.png	9	
4	2.png	1	
5	3.png	7	
6	4.png	3	
7	5.png	9	
8	6.png	4	
9	7.png	9	
10	8.png	3	
11	9.png	4	
12	10.png	2	
13	11.png	3	
14	12.png	6	
15	13.png	6	



	A	B	C
16	14.png	0	
17	15.png	4	
18	16.png	8	
19	17.png	7	
20	18.png	9	
21	19.png	5	
22	20.png	3	
23	21.png	9	
24	22.png	1	
25	23.png	9	
26	24.png	6	
27	25.png	8	
28	26.png	9	
29	27.png	5	
30	28.png	7	

Summary

The Project aim is to convert images containing text which is basically a code to an actual source file which can be compiled directly. This image can have any kind of code written in any programming language and it can be digital characters or even handwritten code

References

1) http://matlab-code.org/image-segmentation-algorithm-based-on-self-orga nized-kohonens-neural-maps-and-tree-pyramidal-segmented.
2) http://www.engpaper.net/image-processing-using-matlab.htm.
3) https://www.researchgate.net/topic/Image-Processing.