Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 9, Issue. 6, June 2020, pg.32 – 36

Number Plate Character Identification Using Image Processing and TesseractOCR

Mrs. Ponmani S[1]; Vijay Prasath D P[2]; Sundar G[3]; Vishal V[4]

[1]Prof, Dept. of CSE & Rajalakshmi Engineering College, Tamilnadu, India
[2]UG Student, CSE & Rajalakshmi Engineering College, Tamilnadu, India
[3]UG Student, CSE & Rajalakshmi Engineering College, Tamilnadu, India
[4]UG Student, CSE & Rajalakshmi Engineering College, Tamilnadu, India
[1] ponmani.s@rajalakshmi.edu.in ; [2] dpvijayprasath@gmail.com ;
[3] mailsundarg12@gmail.com ; [4] venvis1998@gmail.com

Abstract— The Number plate identifier is one of the most widespread and much needed system. The aim of such systems is to successfully identify the Characters in any number plate. An image of a vehicle is captured and processed using various algorithms. The number plate area is localized using a novel, "feature-based number plate localization" method which consists of many algorithms. But our study mainly focusing on the two fast algorithms i.e., Edge Finding Method and Window Filtering Method with the help of Opencv and Tesseract OCR engine .The proposed system achieves an approximate 78.12% accuracy.

Keywords— Automatic Number Plate Recognition, Gaussian Blur, Image Processing, Window Filtering, Edge Finding, Optical Character Recognition.

I. INTRODUCTION

In today's world Automobiles are becoming essential for everyone for transportation. As important these means of transportation are, a unique identification for every vehicle is provided as Number plates for easy identification of a vehicles details differentiated from million others.

Various Number Plate identification systems are being used all around the world for different context and purpose. Traffic cameras, Parking lots and Toll Gates are some of the most commonly used areas. The sole objective of this system is to successfully identify all the characters in any number plate. This is achieved using different methods and algorithms. Most of them vary from each other either in the very basic level or at final stages of processing. The system works by processing an input image through a series of operations and filters and finally segments each and every character and matches them with the trained model in order to provide the actual characters.

One of the biggest advantages of this system is reducing the requirement of man power or human resources in various fields. And moreover it can perform more efficiently than a naked eye.

One of the most popular algorithm used is "Edge finding" and "Window Filtering" method. In this system OpenCv is used for number plate localization and Tesseract OCR is used for character recognition.

II. RELATED WORK

Muhammad Tahir Qadri's paper shows how the OCR techniques is used for recognition of characters varying in all sizes and alignments. This can be used to advance the OCR recognition from various size and angles.

S.Kranthi, K.Pranathi's paper proposed that Automatic Number Plate Recognition (ANPR) is a method that catches the vehicle image and confirmed their license number that can be used in the identification of stolen vehicles.

Abd KadirMahamad's paper explained an automatic number plate inspection using image processing and optical character recognition that has been created of training interface using LABVIEW software.

Kuldeepak's paper has achieved an accuracy of 98% by taking all the parameters of a crowded street or road into account. They accomplished this with better precision streamlining. The issues like stains, blurred regions, smudges and sizes ought to be remembered.

AmrBadr's paper mainly focuses on Morphological operations, Histogram manipulation and Edge discovery Techniques for plate localization and characters segmentation. Artificial Neural Networks are used for Character classification and recognition.

III. DATASET AND REPROCESSING

The Method in this paper makes use of the Tesseract OCR (Optical character recognition) for identification of the characters. The Tesseract OCR consists of the trained model which has been trained with over 64.000 training data from [1].

OpenCv functions are used for localisation and edge detection. The tesseract OCR engine then matches the characters in the input image and recognizes the actual characters.

IV. METHODOLOGY AND EXPERIMENTS

A. Applying Gaussian Blur:

Gaussian Blur is applied to the input image in order to get rid of unwanted regions and noises which can intervene in the recognition process. The Gaussian blur is applied to the Grey scaled version of the input. The grey scale image in generated by the LUMA method in OpenCv.

B. Sobel-Feldman Convolution Kernel:

The Horizontal and vertical Sobel-Feldman convolution kernels are applied to the image separately. These filters locate areas of high spatial frequency or larger pixel value differences and marks them as edges. Then, both these images are combined into a full edged image and binarized. An Aspect ratio Filter is then used to eliminate the unwanted edges and detects only the number plate.

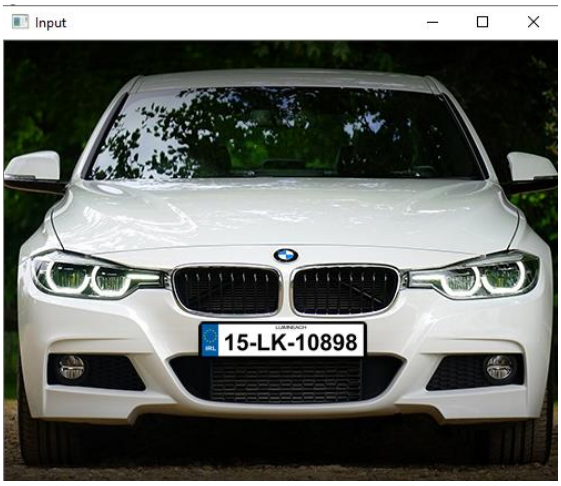

Fig. 1. Normal Image

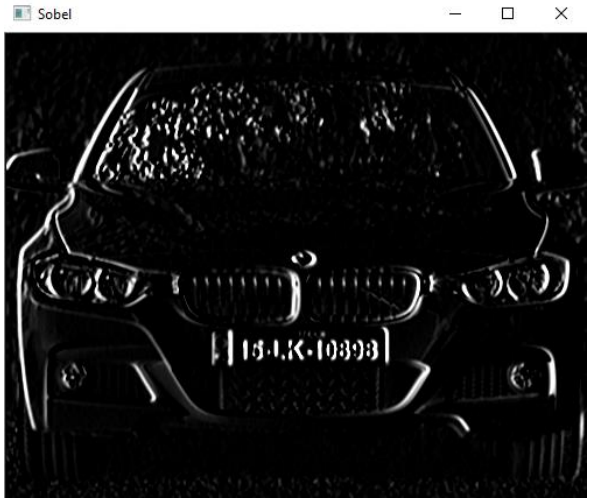

Fig. 2. Sobel-Feldman convolution

C. Optical Character Recognition:

After separating the plate region (Fig.3), the Tesseract OCR Engine is used to identify the individual characters from the region of interest. The Tesseract engine used has been trained to identify the characters with respect to the codes, suffixes and other components of a regular number plate. The Tesseract OCR Engine generates different font and styles of the characters, compiling a dictionary of possible orientations. Fig. 4 explains the proper flow of a Tesseract OCR Engine.

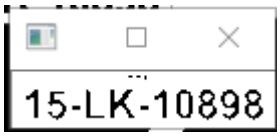

Fig. 3. Plate region

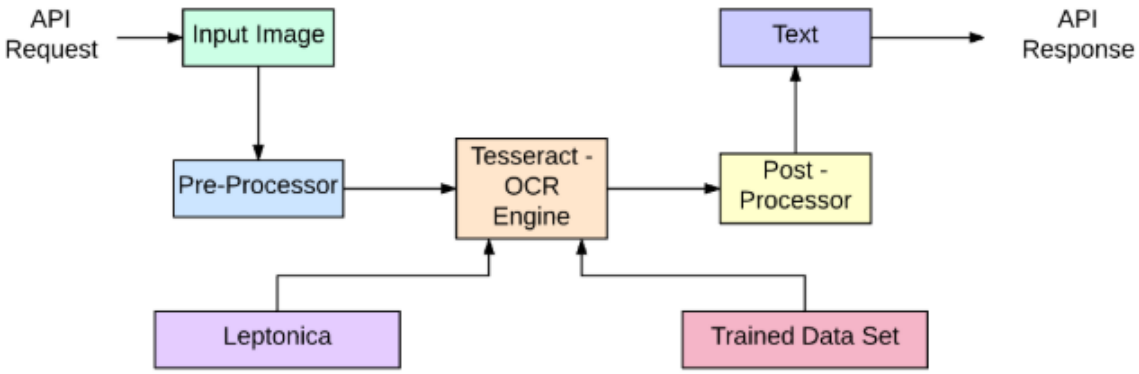

Fig. 4. OCR Process flow

V. FINAL RESULT

The Proposed System proves to be 78.16% accurate in images with a few abnormalities or distortion. Images without any disturbance or Distortion are identified with an approximate efficiency of 82.6%. The few errors that occur are ultimately due to unclear image or in some cases due to the number plates being decorated by unwanted stickers. This infers that the number plate has to be standard without any abnormalities in its design or positioning.

The Figure 5 shows the flow diagram of the entire system and its process.

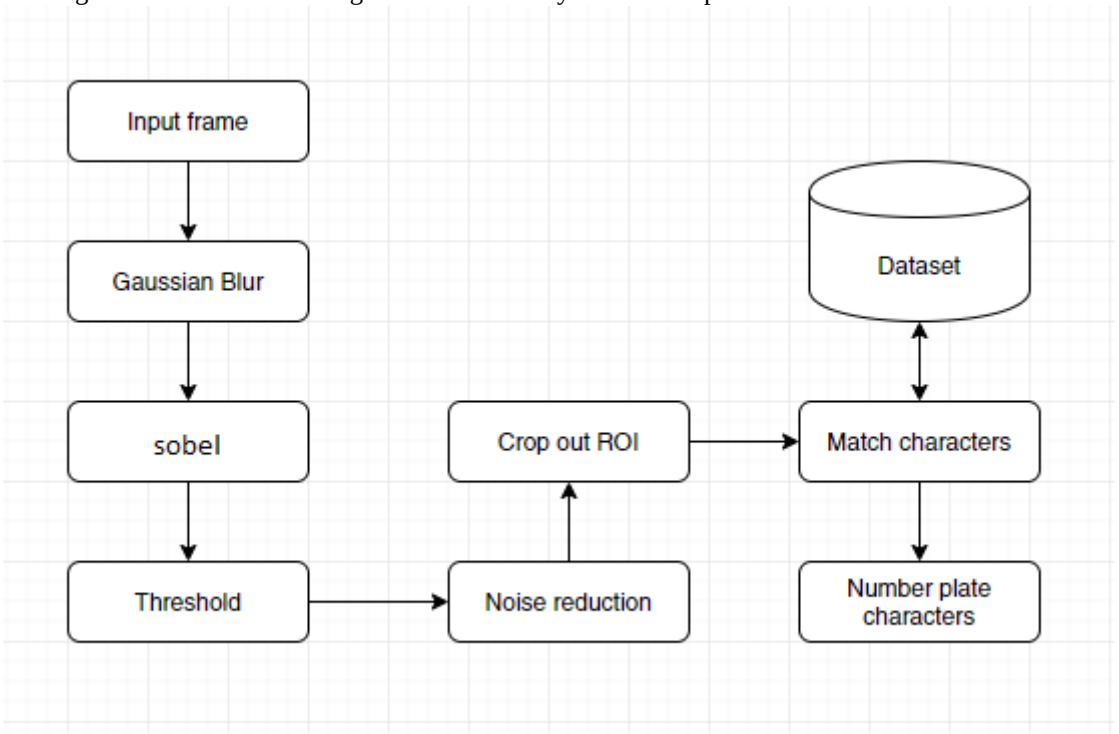

Fig. 5. Architecture

Fig. 6. Detected plate

VI. CONCLUSIONS

The Paper presents a Number plate character recognition system which operates at an average accuracy of 78.12% under various circumstances.

Name	Specification	Model
CPU	Intel	I7 5th Gen
GPU	Intel HD graphics	5500
RAM	DDR3	8GB
OS	Windows	10(64Bit)

ACKNOWLEDGEMENT

Our Sincere Acknowledgement to MNIST and Kaggle for providing with dataset which is used for training the model with such efficiency and specificity.

REFERENCES

[1] https://dataturks.com/projects/devika.mishra/Indian_Number_plates

