

Machine Learning with Text Recognition

Suraj A. Khandare

Department of Electronics

Veerмата Jijabai Technological Institute Matunga, Mumbai (M.S.) India 400019

Abstract - The prime objective of this Research is the development of effective reading skills in Machines. After reading the text and comprehending the meaning, it would self-program itself and according to the program it would implement the instructions. Here I am exploring a new era of computer vision and related Research. The current investigation presents an algorithm and software which detects, recognizes text and character with specific protocol in a live streaming video and programs itself according to the text. The protocol used for text consists of a Start word and an end word with embedded C instructions between them in special font.

The proposed algorithm first detects the frame having text from live video streaming. The character is at its best view which means no broken point in a single character and no merge between groups of characters. The image processor tallies the images and burns the main microcontroller accordingly. This technology can prove to be immensely important to various human activities in day to day life.

Keywords - Image and video Processing, machine learning, text detection, character Recognition. Cross compiler and integrated development environment (IDE).

I. INTRODUCTION

Technological advances in image processing have acquainted us with character Recognition and many such related technologies which have proved to be a milestone in computer vision. However, even years after the invention of these technologies we have not been able to achieve a technology by which machine can read, interpret and act according to the instructions and even update their database if required.

Here's an attempt to make these reality.

Machine replication of human functions, like reading, is a long awaited dream. However, over the last five decades, machine reading has transformed from a dream to reality. Text detection and character recognition, known as Optical Character Recognition (OCR) has become one of the most successful applications of technology in the field of pattern recognition and artificial intelligence. Numerous commercial systems for OCR exist for a variety of applications. [1]

Let us analyze how human body functions. In human body, the processes mentioned above are carried out in the following pattern as described in Fig.1.

1. The signals are first sensed by the eyes and then are sent to the brain.
2. The brain interprets these signals and stores them temporarily.
3. Brain then transmits these signals the specific nerves,

which in turn perform the action.

The operation of the machine is analogous to the processes carried out in human body.

1. Firstly, the machine acquires text signals through camera, which is sent to the image processor.
2. It then recognizes these signals and stores them in temporary memory of the processor in the form of images
3. Now these images are converted to equivalent characters by processor which is analogous to human brain.

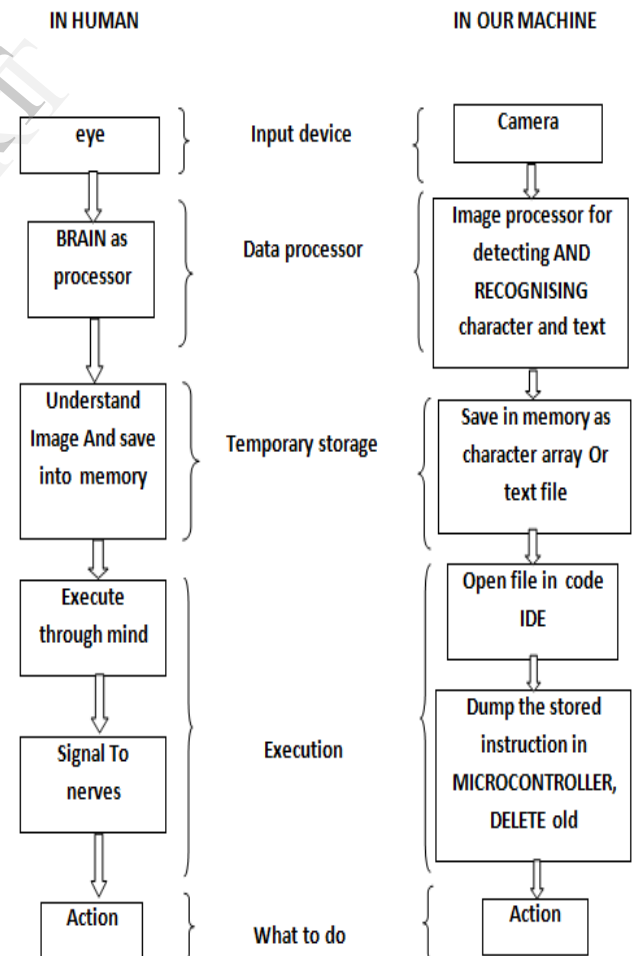


Fig.1. Algorithm for Reading in human and our machine

4. This generated text is stored as text file and verified for appropriate protocol.
5. The system gets programmed according to the instructions in text file.

We can well described this by above algorithm.

II. LITERATURE REVIEW

Machine learning is a scientific discipline that is concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases. A learner can take advantage of examples (data) to capture characteristics of interest of their unknown underlying probability distribution. Data can be seen as examples that illustrate relations between observed variables. A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. The character Recognition techniques can be implemented to achieve the objective of machine learning. [2] Character Recognition has been used to enter data automatically into a computer for dissemination and processing. The earliest of systems was dedicated to high volume variable data entry. The first major use of Character Recognition was in processing petroleum credit card sales drafts. This application provides recognition of the purchaser from the imprinted credit card account number and the introduction of a transaction. [3] The accurate recognition of Latin-script, typewritten text is now considered largely a solved problem on applications where clear imaging is available such as scanning of printed documents or their pictures. Typical accuracy rates on these exceed 99%; total accuracy can only be achieved by human review. [4]

III. OUR APPROACH

We have divided the whole research in two parts

1. Text acquisition.
2. Programming the microcontroller.

First phase is related to detection of the text by camera and converting the image into text format by using 'Character Recognition'. The Second phase deals with compiling the text and programming our main controller.

A. Text Acquisition

The text acquisition is a broad section which includes live streaming of continuous images (video), text detection, text cropping, recognizing and converting to text file. This process includes the following algorithm of 'Text Detection' And 'Image To Text Conversion'. This is described in section B and C respectively.

The image processor continuously takes the frame which is known as video streaming. The text detection block will detect the frame having readable text. And the output of this stage is structure of images. Image to text converter will convert this image into a text file. This file is stored as .c extension. This is nothing but the file acquired from live

video streaming. To match with the processing speed the streaming must have a delay between two frames. Fig.2. shows the algorithm for text acquisition

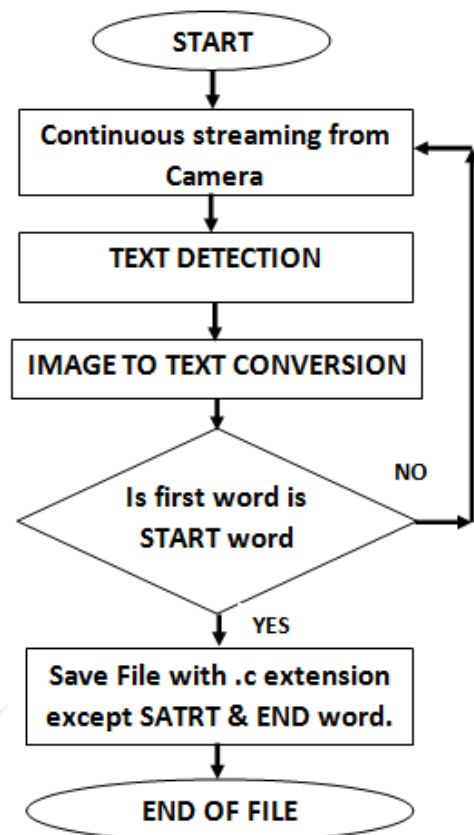


Fig.2. Algorithm For Text Acquisition`

B. Text Detection

This is the process of detecting the image frame from the continuous video stream having readable text. For understanding the logic shown in Fig.3 which we have applied, it needs to go through our protocol. We only want to detect the text which is in our protocol. Fig 3 gives the idea regarding this. It identifies the black colored rectangle surrounded by white color. Then it finds all the black shapes in that rectangle. It tallies weather the height of all the black shapes is same. Even though the width for characters '1' and 'l' is less than 1/2 the height of all remaining characters, they are accepted. If all these condition are satisfied it means the current frame is having a text in it. This text is cropped (Fig.4.) and it separates out all the black shapes in it which are nothing but the images characters.

Above cropped text is then processed to crop each character (Fig.5.) and saved as structure of images that is nothing but the structure of 2 dimensional matrices.

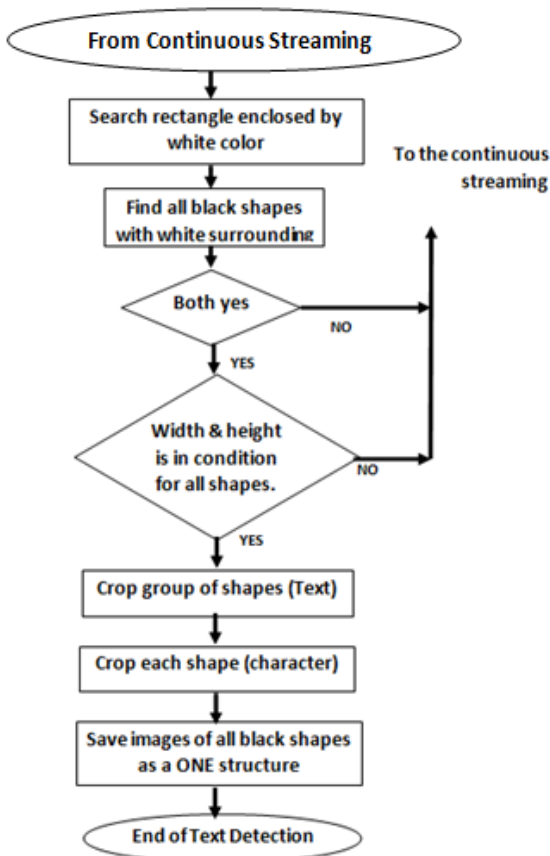


Fig.3. Flow Chart of Text Detection



Fig.4. Detected frame having Text in it.

```

# S T A R T
P O R T A = P O R T B
# E N D
  
```

Fig.5. Cropped Text

C. Image To Text Conversion

This is shown by Fig.6. Input to this stage is structure of images. Initially set the count i equal to 1 and take the i^{th} image from structure and apply 'Character Recognition' (explained in section D). The output of this is a character, now this character is saved in one array at i^{th} position,

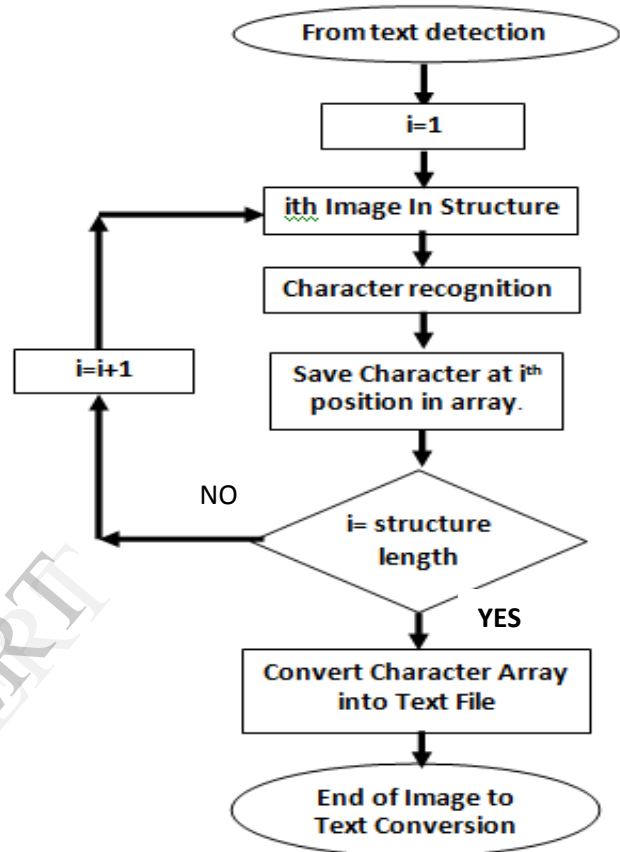


Fig.6. Algorithm For Image To Text Conversion

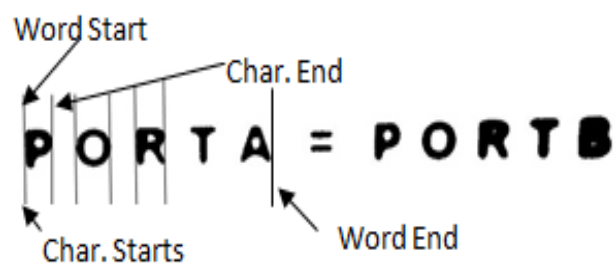


Fig.7. Separation of Characters [1]

This character array is converted to a text file and saved in temporary memory.

D. Character Recognition:

This is the most complicated Algorithm (Fig.8.). As we have a Specific format of text in our protocol we need not use fuzzy logic and neural network for Character Recognition. Here we are taking 23 samples (Fig.9.) in the complete image and comparing those samples to the samples in data base for

each character. The character to which maximum samples matches is the recognized character.

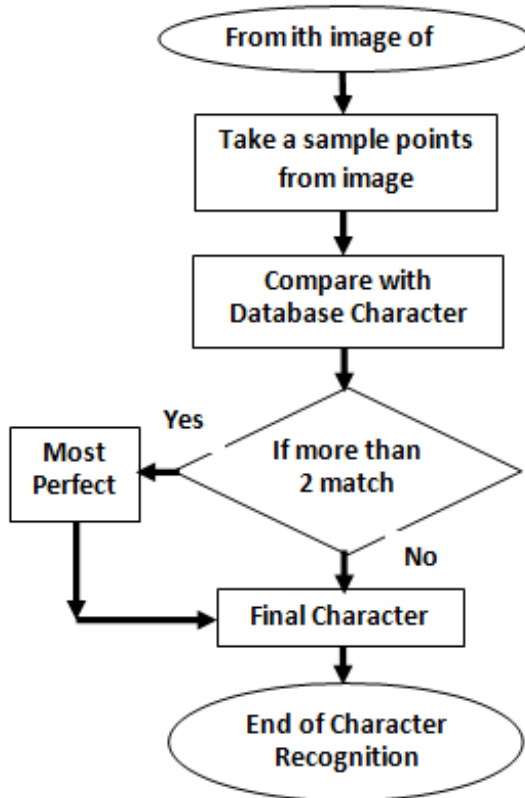


Fig.8.Character Recognition

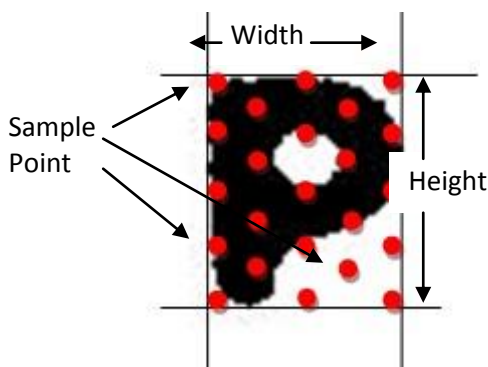


Fig.9. character Recognition by Sampling of Images.

At the end of this process we get a character array which is converted to a text file and returned to 'Image To Text Conversion' algorithm. The described algorithms showed the efficiency and difference when compared with other algorithms.

E. Programming The Main Controller

This is the most important and the final step to achieve our goal. The text acquired from continuous live streaming is then dumped into the microcontroller. This is achieved by opening the text file through Integrated Development Environment (IDE). First, it scans all the words. If all words are present in library, i.e. the machine knows the meaning of

those words, then there is no problem, but if some mismatch occurs then it becomes complicated to find the error and rectify it. We have achieved this by simple algorithm explained in Fig.10. If the number of mismatches exceeds 1 then IDE opens the last executable file which was saved at specific memory location when controller was programmed last time. If 1 mismatch is found then the software will find three tentative matches to the word and these three words replace the mismatched word one by one. After checking all the three words if there still exists an error then the software will compile the last saved program. If there is no mismatch then there is negligible possibility of errors. This will compile the same program three times in the controller then finally last saved program will get dumped into the controller.

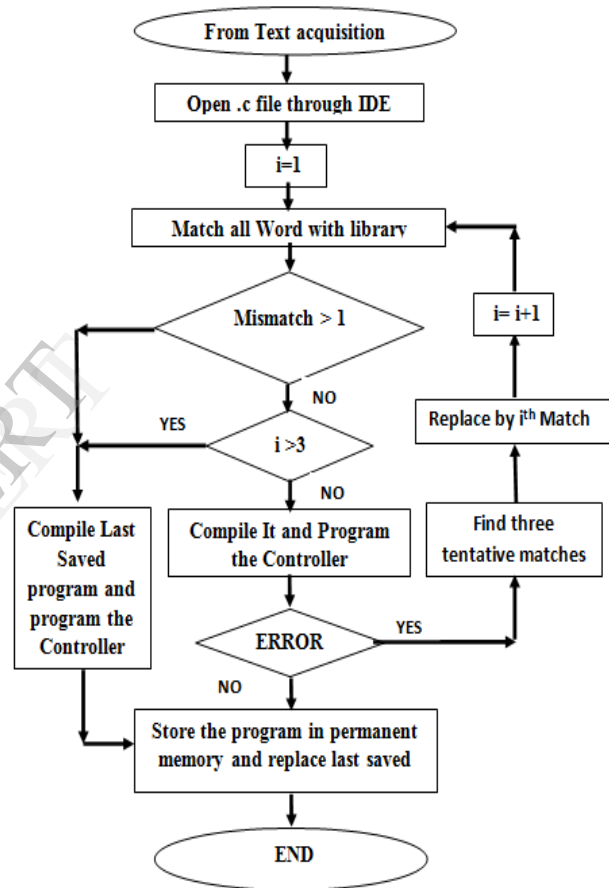


Fig.10. Algorithm For Programming The Main Controller

IV. FUTURE PROSPECTS

This research is a step towards bringing Machine world closer to human world; this research has the capabilities to give the large number of applications in future. Tomorrow's world is a world of machines and robots, man will use robots and machines to extract work from them. In order to work like man, machines must have human skills and reading is one of the most important skills. This research gives the reading ability to the machine and the capability of understanding and implementing meaning of text in it. Let's take an example of robot in year 2030, suppose that you have a robot in your

home as cook and you are out of station for important work. Your kids are at home and regular microwave oven is not in working condition. You are having another oven but robot is not programmed for that. Our technology gives the power to robot or machine to read user manual of other microwave oven and program him as instructions in manual to operate new Microwave oven. Again when robot will walk on roads like man it will have to be follow instruction written on roadside like human such as 'DON'T HORN', 'GO SLOW' or follow some speed limit. Our research gives the ability to program him according to the instructions. The Research is in primary stage that's why we are using some prototype, after advancement robots will be able to understand human languages.

V. CONCLUSION

Through this paper we have tried to make machine more capable and thus introduce machinization in the day-to-day activities to lesser human dependence to program them. Employing the above technique we can make machines more intelligent and reliable so that machines recognize (read) instructions and bring them into action.

ACKNOWLEDGEMENT

I am extremely grateful to Prof. R. D. Daruwala, Veermata Jijabai Technological Institute Matunga, Mumbai, for his guidance which have ameliorated my pains and provided a head start to my research work. I am also thankful to my friend Mr. Harshad Mahajan for his priceless contributions.

REFERENCES

- [1] Journal of ELECTRICAL ENGINEERING, VOL. 57, NO. 5, 2006, p 258-267.
- [2] Tom M. Mitchell (1997) Machine Learning p.2
- [3] Alpha Drive Pittsburgh, Optical Character Recognition (OCR). AIM., 63 Pa 15238-2802, USA.
- [4] Suen, C.Y., et al (1987-05-29), Future Challenges in Handwriting and

Computer Applications, 3rd International Symposium on Handwriting and Computer Applications, Montreal, May 29, 1987, retrieved 2008-10-03.

- [5] Ye, Q.—Gao, W.—Wang, W.—Zeng, W. : A Robust Text Detection Algorithm in Images and Video Frames, the 4th International Conference on Information, Communications & Signal Processing — 4th IEEE Pacific-Rim Conference On Multimedia (ICICS-PCM2003), Singapore, 2003.
- [6] Okun, O. Morphological filter for text extraction from textured background. In Vision Geometry X, Longin J. Latecki, David M. Mount, Angela Y. Wu, Robert A. Melder, Editors, Proc. of SPIE.
- [7] Ye, Q.—Gao, W.—Wang, W.—Zeng, W. : A Robust Text Detection Algorithm in Images and Video Frames, the 4th Conference on Information, Communications & Signal Processing — 4th IEEE Pacific-Rim Conference On Multimedia (ICICS-PCM2003), Singapore, 2003.
- [8] Yuan, Q.—Tan, C. : Text Extraction from Gray Scale Document Images Using Edge Information, Proc. Sixth Int. Conf. on Document Analysis and Recognition, 302-306.
- [9] Hara, S. OCR for Japanese Classical Documents Segmentation of Cursive Characters: National Institute of Japanese Literature, PNC2004 Annual Conference on Conjunction with PRDLA, 2004.
- [10] Mohanad Alata — Mohammad Al-Shabi Text Detection And Character recognition Using Fuzzy Image Processing Journal Of Electrical Nginieering 57, No. 4, 2006.
- [11] Husni A. Al-Muhtaseb, Sabri A. Mahmoud, and Rami S. Qahwahi- A Novel Minimal Script for Arabic Text Recognition Databases and Benchmarks International Journal Of Circuits, Systems And Signal Processing Issue 3, Volume 3, 2009 p145-153.
- [12] Naresh Kumar Garg, Lakhwinder Kaur, M. K. Jindal -Segmentation of Handwritten Hindi Text2010 International Journal of Computer Applications (0975 – 8887) Volume 1 – No. 4
- [13] Seiichi Uchida, Hiromitsu Miyazaki, Hiroaki Sakoe- Mosaicing-by-recognition for video-based text recognition. Pattern Recognition 41 (2008) 1230 – 1240.
- [14] Video text recognition using sequential Monte Carlo and error voting methods Datong Chen, Jean-Marc Odobez. Pattern Recognition Letters 26 (2005) 1386-1403.



Suraj A. Khandare received B.Tech degree in Electronics and Telecommunication from Amravati University in 2011. My current research interest includes biomedical embedded systems.