METHODOLOGY FOR RECOGNITION & EXTRACTING CONTENT FROM CAPTURED IMAGES USING GEOMETRIC AND STROKE WIDTH VARIATION PROPERTIES

R SHIVA SHANKAR1, VMNSSVKR GUPTA2, D RAVIBABU3, LV SRINIVAS4

1-4 S.R.K.R. Engineering College, Bhimavaram, West Godavari, Andhrapradesh, India

1 shiva@srkrec.ac.in; 2 guptavkrao@srkrec.ac.in; 3 ravibabu.devareddi@srkrec.ac.in; 4 lvs@srkrec.ac.in

ABSTRACT

Content detection and Extracting content in captured images is very important but having trouble to retrieve due to the variations inside the textual contents such as fonts, size, line orientation, complicated background in image and non-uniform illuminations. To overcome these problems, effective skills for textual content image recognition are used for extracting text from the images. Different OCR (Optical Character recognition) procedures are available but still this problem is not solved. The information present in the form of text in the captured images is highly useful and valuable so that this information should be portrayed accurately as it is in the pictures. Therefore it has became a most challenging task to extract the text from image due to text which can be of negative quality having different fonts and numerous styles and sometimes having a complex background etc. Various strategies have been launched but it is hard to solve. MSER (Maximally Stable External Regions) algorithm is proposed to detect the Text Region from the images. MSER is used to elevate the plain regions from text areas and non text areas by using geometric properties and stroke width variation properties.

Keywords: OCR; Text Detection; Text Retrieval; Scene Images; MSER; Geometric Properties; SWV Properties;

I. INTRODUCTION

Text detection has emerged as essential trouble inside the beyond few years. There are improvements which has supported text detection and retrieval within the subject of pc vision and machine learning further as growth inside the applications has resulted all through this trend. Nowadays we all recognize that how virtual cameras and smart cell telephones are easy to be had in a wider range, there is several research carried out on textual content detection in pixel captured with the aid of the digicam as text processing has viral attention on it. Text detection is an essential step for several document images processing tasks, like optical individual recognition (OCR), format analysis, and pre-processing algorithms [1].

Many techniques are targeting the text or constrained document cases, handiest little work was done on unconstrained camera-captured report images. For text detection in such inputs, address the challenges in natural (digicam-captured) photos and curved content lines. However, most conventional methods cantered on both cases, and additionally the text detection in unconstrained inputs wasn't completely addressed. Texts inside the photographs include valuable facts and deliver accuse about photos. So it is critical for a personality's furthermore as computer to comprehend the scenes. Many other applications include several procedures which are now developed in machine learning especially the features from unlabelled data and to show they allow constructing classifiers that are highly effective for detection in producing high accuracy by applying large-scale algorithms. Texts within the pictures deliver high-stage semantic records of the scene. These records in turn will increase large data in databases and in the web (data can be images). It should be a great challenge in developing strong strategies to perform operations on the resources by managing and retrieving them. It is a hard challenge to hit upon and phase textual content from photographs because of various styles and fonts of different sizes, different shades and sometimes there is a chance of having similar background color and text. After the
The completion of these stages, the text detection gadget will convert the image into readable textual content, for a complicated image it gives poor performance [2].

Picture colorization is the mechanism of rendering a gray scale image by taking a black and white image (a grayscale image) as reference and creating an object in RGB format (a color image). The method of coloring a picture has a lot more affect on fields such as celestial phenomena, security images, or images taken under an optical microscope, depending on what color spectrum is used. In most grayscale image, the amount of information accessible is insufficient to comprehend the picture. Thus, coloring the photographs provides one with additional detail necessary to comprehend the image's semiotics [3].

II. LITERATURE SURVEY

Text popularity from pictures remains active research in the subject of this pattern recognition. We have one kind of proposed technology to deal with the problems related to textual content reputation; each technique tries to deal with the problems in an exclusive way.

Nirmala Shivananda et.al, [1] suggested an effective method for discovering and extracting textual information from complicated historical colour record images. This technique employs the canny element detector algorithm, which is used to identify edges. If the edges have been collected, the procedure will be performed on them. As a result, it is concluded that the modules have gaps that must be kept to a limit. Connected additives with no holes were removed. Other non-textual content elements were removed by computing and interpreting the normal variance of each associated component. Noisy textual material regions were identified and processed at the end to improve the quality of the extracted foreground.

Thai V. Hoang et.al, [2] were defined the approach of text extraction from graphical images. Morphological Component Analysis (MCA) is used to describe the development of a sparse illustration framework. Two discriminative dictionaries have been primarily based on transform and curvelet transform. S A Angadi et.al, [4] are used the Discrete Cosine Transform (DCT) method to detect the content to picture analysis. The processing includes improving the sensitivity and the precision of the editing and parsing of neighbouring blocks of textual data and unprocessed zones. This technique is applied to low-resolution photographs taken indoors and outdoors. Additionally, this approach detects nonlinear research subject areas and can be derived from images in other languages with minimal adjustment and this entire flow was shown in Figure 1.

![Fig 1. Extracting Text Regions from Low Resolution](image_url)

R S Shankar et.al. [5] was used a technique LSTM model to sharpening blur images in different resolutions and generates a sharp image by converting it into the original format at the decoder. So that it reduces the time
complexity and gives an accurate output. The development of digital image processing, image interpretation and classification is facilitated by the use of CNN's in combination with a system. Additionally, we use tensor flow and an image data generator to prevent overfitting [6].

C.V Jawahar et.al, [7] also developed a framework for the devanagari script's accomplishment. They used a method for eliminated the need for text extraction, which is one of the most common sources of high word error rates. It was documented that the word error charges were reduced by more than 20% and the man or woman error charge was reduced by 9%, according to the OCR scenario. M.Sundaresan et.al, [8] suggested talks on the retrieval of English text from a blob in a comic picture. Text detection and extraction from comic images enable you to retain the text information and encoding throughout the phase transition, resulting in a rather high-quality text image. The RGB images are converted to binary images during pre-processing. Linked component photos use the CCL algorithm that subtracts colour information from their RGB channels to make finding the linked components easier. Extracting is used to identify non-text blobs, as well as to stay away from false positives. The recognition is mostly focused on blob size features which are shown in Figure 2.

Fig 2. Two blobs extraction method for extracting Digital English comic image

Karin Sobottka et.al, [9] have proposed the automatic technique for textual content area and identity the colored e-book. A clustering set of rules is used to reduce the number of colors turned into pre-processing step. For extracting textual content is divided into two methods. One approach is focused on top-down research, which divides picture regions sequentially. The alternative is to use bottom-up areas. Finally, both approaches are merged to distinguish “textual” &” non-textual” content. Text variables are routinely used to derive knowledge regarding the colour of the text.

Sachin et.al, [10] proposed a process for extracting text-like regions from printed documents that is both quick and accurate. To screen for natural language edges, the HAAR employs a Differentiated Text Ranking technique and employs the thresholding approach for non-text edges. To attach the remote candidate text after which line characteristic vector graph is generated primarily based on the threshold map.

Syed Saqib Bukhari et.al, [11] proposed the technique for colour textual content line that is extracted from the gray scale camera-captured record photos. These technique is based on total differential geometry, which makes the use of local path of gradients because the measure of text content. This gray scale textual content line is enhanced by using the usage of multi-oriented smoothing. Detection of central traces of curled textual content strains is determined the use of ridges. The orientation of gradients and derivatives are determined using the
Hessian matrix. The null crossover of the strongest directional derivatives of a smooth picture is observed using such knowledge.

R S Shankar et.al, [3] has implemented a technique to build an automated model for gray scale images using CNN. They extracted the features using feature extractor which combines the result of the encoder at fusion layer. Based on the performance measures they concluded that CCN will give better results to colorize the grayscale images. To remove the noise reduction for PGM images will be based on maximization of similar pixels. Image originality will be preserved to remove the noise effectively by using fuzzy filter [12].

Kohei Arai et.al, [13] has proposed the Reading digital comic on cellphone. It is advantageous to modify the current interactive comics web site. Rather than creating fresh material for cellular comedy. An automatic e-comic smart phone content material versioning technique for mechanically generating mobile comic content from a remote comic website repository is suggested. This methodology suggested is an optimal way of implementing real-time practices and the work flow was shown in Figure 3. The results of the experiments revealed that flat comic body extraction was 100 percent accurate, whereas non-flat comic body extraction was 91.48 percent accurate, and processing time was 90 percent faster than the previous procedure.

Rajib Lochan Das et.al, [14] Compensated world wide as well as adjacent characteristic extraction technique. Some technologies are involved in the global role, including six projection capabilities and four curvature capabilities. To obtain nearby functions, the image is divided into nine equivalent sections, with four gradient features measured for each row, yielding a total of 36 functions. For every institution new characteristic is calculated to characters inside the organization. They present the singular method to accurate discover text in coloration pixel in all likelihood with a complex background. First, they use complicated detection set of rules to extract text edge pixels. Second connected to evaluate the employed textual content candidate place and classify the non-text areas.

In terms of text recognition and extraction, Zeev Zelavsky et.al, [15] presented little challenges. In the majority of cases human inspection is needed for a full solution to be provided by Optical Character Recognition (OCR). They propose a novel text recognition algorithm that is entirely dependent on fuzzy logic principles and mathematical information about the evaluated font. In a series of blurry, mostly dependent laws, the latest technology incorporates letter statistics and correlation coefficients to allow twisted lyrics to be recognized properly. They concentrated on Rashi fonts correlated with bible commentaries, which are distinctively handwriting calligraphy was described in Figure 4. Yang, Jufeng et.al, [16] Developed a novel innovative approach focused on the wavelet filter. This method transforms through strategies quite quickly, providing the potential for real-time computing and cellular applications which was shown in figure 5. They examined the customized manner on images of complex scenes from a database.
Huang, K., et.al [17] introduce a novel approach for scene text detection focused on MSER that incorporates many novel techniques. To begin with, they suggest an algorithm that prunes MSERs very quickly while remaining precise, which allows us to locate and classify the majority of characters even though the picture is poor quality. Additionally, in machine learning, this suggests a novel simultaneous self-distance learning algorithm. Finally, they suggest a classifier that approximates the posterior likelihood of a nominee text and excludes candidates with very high likelihood of not being text. Eventually, they create a versatile text detection system, which overrides state-of-the-art approaches in many publicly accessible databases, by integrating these novel approaches.

Fig 4. Retrieving content from Rashi Semi Cursive handwriting via fuzzy logic

X. Chen, et.al. [18] provides a method for identifying and interpreting text embedded in natural photographs. This algorithm is for cityscape detection for the blind and visually disabled to help them find their way around the city. They begin by assembling a dataset of images captured in cities by blind and naturally sighted objects. They manually mark and remove text regions from this dataset. After that, the text regions are systematically evaluated to determine if the image characteristics are valid text measurements of low entropy (i.e. feature response is similar for all text images) inconsistent classifiers, where attribute responses and unqualified answers are combined. This ineffective classifiers are fed into a machine learning algorithm called AdaBoost, which is used to create efficient classification models. In the field, the multi-class technique they used the training of a strong decision tree of 4 feature maps containing a total of 79 features. To the regions chosen by the cascade classifier, an adaptive binarization and extension algorithm is implemented. An advertising OCR program is applied to interpret the text or to find rejected because of it is marked as an area of unreadability. On the test collection, the total algorithm achieves a performance rate of over 90% (as measured by full text identification and reading), and the unread text is usually tiny and far from the monitor.
Pan, Y. F. et.al, [19] offer a hybrid approach to scene text localization that incorporates area knowledge into a functional CC-based process. It should be noted that in addition to the fundamental property, the binary contextual relationship is introduced into a CRF model where both the supervised learning parameter and as well as the unary relationship are also play an important roles. The suggested technique shows to be applicable to scenes with any variety of scene structures, such as cluttered or organized, with loose or regular layouts, respectively. Applying regional information is highly effective for text segmentation and analysis, while it helps to group non-text elements into classes with CRF better. Using joint energy parameter minimization with CRF makes classifiers more robust.

Mirmehdi et.al, [20] worked on a portable text recognition platform that relies on MSERs for real-time text detection. By using the hierarchical structure derived from MESRs, the previous real-time algorithm is refined by the proposed strategy for producing more secure areas than the previous adaptive threshold approach. It overrides computationally other reported approaches and maintains a comparable efficiency in ICDAR data set text detection. In their future endeavours, they hope to pursue character recognition without relying on third-party applications and build on their already-deployed cascading filters to have even better performance.

III. METHODLOGY

OBJECTIVES

- **To achieve a personalized user experience**: This provides the user to retrieve the text from an image in a personalized manner. So, the user feels more contended.

- **To be able to indulge more number of users**: with the embedded Text Extraction system, the user can easily get the exact text from any type of image. So, more number of users starts using the service

Our technique includes few steps which are written inside the figure. MSER (Maximally Stable External Regions) is set of rules detects the Text Region inside the photograph and exist in the result. MSER detects main text areas from natural scene image however along this moreover come across few non-text areas that are detected in next step. MSER makes use of two vital houses to get rid of non-text content a region by photograph.
first is Geometric Properties and followed by stroke width. By the non-textual content characters are merged by group of words. Finally the OCR detects the text contained in every bounding box.

**Step 1:** Detection of Text Regions: Here MSER algorithm is used to find text regions from all of the areas in the given photo. Their exits a lot of non text region and textual content can be eliminated. To detect textual content areas MSER first convert the coloration image into gray scale photograph.

**Step 2:** Removing Non-Text Regions involves onto the Geometric Properties: The MSER set of rules chooses most of the textual content and it moreover detects other stable regions within the photograph that aren't textual content. We can use a rule-based approach to locate and put off non-text areas present inside the photo. The non-textual content areas from image are filter by the usage of geometric homes. On beside we ready to maintain the system learning method by educate the main text both of geometric things and learning technique gives the best.

**Step 3:** On removing the Non-text regions on the basis of stroke width unusual metric used for it between the text and Non-text. The main variance is to be occurred by width of the curves and the contours making up of person are contained by non-text and the little bit by text regions.

**Step 4:** It outcomes by merging text regions are done by person textual content characters. To use these consequences for tasks, consisting of OCR, the separate textual content characters want to be merged by phrases or text lines. This permits recognition to the real phrases on a photo, which contain extra meaningful records than just the man or woman specifications.

**Step 5:** By detecting the textual content areas, use the OCR characteristic to understand the textual content within every bounding box.

**EXISTING SYSTEM:**
- There are different imaginative challenges which have been attained in the past few years by growing greater attention in the field of text based records in a photographs.
• Extracting the text from an image could be a challenging task as it can contain different kinds of variations in size and style and sometimes having massive text in picture and some blurry backgrounds which are having contrast lighting.

• The existing system is not that efficient for extracting the text from image and it contains the following steps:

Step-1: Detecting Features
Step-2: Finding and fixing using localization
Step-3: Tracking
Step-4: Text Extraction and Improvement
Step-5: Final Text recognition is done using function OCR

Disadvantages of Existing System: Due to complex background, and variations of font, size, color and orientation, textual content in natural scene images has to be robustly detected before being diagnosed and retrieved.

PROPOSED SYSTEM:
• As the First step we apply MSER algorithm in order to find the textual contents in the images and predict the results.

• This algorithm along with text parts in image also detects some of unwanted regions (thus may contains symbols or it can be a complete non text region) removal of such regions will be further done in next steps.

• The proposed algorithm mainly make use of two important those are geometric properties and stroke width variation properties these are used to eliminate most of the plain regions from the image.

• After removing the plain and unwanted regions like non text regions from the image we then combine all the identified characters to form phrases or in form of combined text words.

• The merged content is finally recognized by OCR which will identify the text within the bounding box. Later this will be displayed in the form of output to the user.

Advantages: This proposed process is designed to focus on low quality images which are used to find major text and most of the characters.

IV. IMPLEMENTATION

Image converter may be thanks to execute particular properties on enter photo, contains textual content and non-textual content with the intention to take out some advantageous statistics from the pictures or to beautify the pictures. Image converter may be a signal converter in which input takes the image and the output will be photographer features/capabilities related to that image. In our day to day life, image converter is fast growing technology.

Image converter involves the following three points:

• Importing the photograph

• Analyzing and arranging the photograph

• Output where the result is also altered picture or report that's based on image evaluation.
As an important representation from people language, visible texts are widely utilized in our day life. The text seems to be communicating on things round us. The textual content may be a use of statistics in our day life. Text in photos provides records approximately semantics of the image. Apart from this we analyze and use the MSER algorithm and its rules. The MSER algorithm discovers all letters from images even where the images are in either low or high resolutions. The photos contain essential names, locations, date and time, manufacturers of the products, street signs and symptoms, which may be helpful records to grasp picture. Finding quality manner to return across and extract differing kinds of textual content from complicated photographs is that the first goal of the text recognition from pictures.

Definitions of the phrases of textual content retrieval

- **Text Document:** In this the characters had some smooth history like a few scanning papers in a file text.

- **Text texture:** In this the text includes the pixels along with the media textual content.

- **Local history:** In this the nearby background carries particular region of the textual content. In the nearby historical past there is a mass of embedded text in the snap shots.

Application of the textual content recognition

Today, extra records are filed in different sorts of virtual shapes involving the snap shots. Using the text reputation the file carries vast characters to control that specific character. The text consisted in the photograph is robust source of data. The textual content additionally seems at the gadgets round us.

Scientific validation

In report text processing, the hassle is the way to extract the textual content from arbitrary heritage. The textual content in snap shots may additionally have diverse dimensions and fonts. We sincerely amplify the pictures to healthy the scale of popular methods of the record text. Text processing is shown in the below figure.

![Image of text processing](https://via.placeholder.com/150)

**A. THE TEXT DETECTION ALGORITHM**

Throughout this portion, we discuss about the text recognition and its set of rules that is MSER (Maximally Stable Extremal Region) algorithm. It could be a way for textual content detection and blob detection in images. The MSER rule eliminates the image's vast selection of covariant regions, i.e. text regions. First we should summarize the concept of stroke and then explain the stroke width transformation. MSER totally depends on the concept of places which live nearly the identical through a good scope of thresholds. Compared to the give threshold if the pixels are above to the give threshold then it consists as a black and if the pixels are below to the given threshold then it performs as a white. In MSER algorithm it has two properties to remove the text and non text regions that are Geometric Properties and the other is Stroke Width Variation Properties.
To begin using the MSER algorithm, we evaluate the text properties as follows:

a) Text in images often contains a large number of corners.

b) The text width is often larger than the text height.

c) The text dimension is often bounded.

d) This property's text has a unique function in that the texture is odd.

The following figure 7 depicts the MSER algorithm's flowchart. MSER algorithm identifies all language regions as well as certain non-textual content regions. Geometric properties are added to photographs to exclude non-textual areas. If Geometric Properties are unable to eliminate all non-textual content areas, use the Stroke Width Variation properties on photographs to eliminate any remaining non-textual content areas.

**Fig 7. Flowchart of the MSER algorithm**

**Geometric properties:**

Many of the photographs in a text are identified by the MSER module. On the other hand, non-text content was discovered. Geometric properties such as simplicity and randomness are most important in those places which have no text in them. Geometric Properties senses nearly all non-textual information areas in a photograph and removes it. We extend the Stroke width difference properties to any remaining non-text areas that aren't removed in Geometric Properties.

```
Algorithm 1: Finding stroke width
Input: binary image BW
Output: stroke width image SW
D:=Distance Transform(BW);
D:=round(D);
For p=each foreground pixel in D do
    PVal:=D(p);
    Lookup(p):=p's 8 neighbors whose value<PVal;
End for
MaxStroke:=max(D);
For Stroke=MaxStroke to 1 do
    StrokeIndex:=find(D==Stroke);
    NeighborIndex=Lookup(StrokeIndex);
    While NeighborIndex not empty do
        D(NeighborIndex):=Stroke;
        NeighborIndex:=Lookup(NeighborIndex);
    End while
End for
Return SW:=D;
```
Stroke Width Variation Properties:
These properties may be used to remove non-text areas as well Stroke width is a design metric used to indicate the shape and or complexity that a character possesses. In the non-text fields, there are few variables for stroke distance. To throw away the non text material, we need to set a threshold of width values. When the pixels were higher than the previous it is dark, when it is lower than the previous it is white.

Stroke width Filtering:
Stroke width is known as the distance from a text aspect to another from an orthogonal axis. This is mostly due to the fact that the breadth of the stroke algorithm remains almost identical in unmarried character; but, due to their irregularities, there is a common alternation in non-text areas.

Converting a group of letters into texted lines:
Here single characters i.e separate detected letters are merged in the form of words like phrases and line of text. These describe simple and understandable statistics when compared to single characters.

A. WORKING PROCESS
This shows the simplest way to find out detected regions on images that incorporate text is a typical project completed on unformed scenes. Pictures of discordant or irregular scripts are known as unformed scenes. For instance, you might step across detect textual information in recorded footage on a regular basis to warn a driver about an avenue sign. This is distinct from the focused scenes, which incorporate acknowledged scripts in which the location of textual content is known early. Partitioning the text from an unformed scene allows additional tasks together with Optical Character Recognition (OCR). The textual content detection set of rules in this case detects a big quantity of text vicinity applicants and step by step gets rid of those less possibly to comprise textual content.

Step 1: Detects Text Regions Using MSER.
The MSER algorithm function detector works nicely for locating textual areas. It works well for text due to the fact the regular coloration and large comparison of text lead to stable intensity profiles. Apply the come across MSER properties to locate and plot all of the regions in the picture Due to the existence of multiple text fields, be sure to look for non-text elements in the visual data.

Step2: Removing the Non-Textual Regions by using Geometric Properties.
The MSER algorithm and its set of rules choose out a maximum of textual content, it additionally detects the large number of strong areas inside the photo that are not text. To remove non-text regions we use the rule-based properties. Typically, a mixture of the two strategies produces better results. Several geometric homes can be appropriate for distinguishing between text content and a non-textual content region includes:

- Aspect ratio: It is the width-to-height ratio.
- Eccentricity: It depends on the circular nature of the given areas.
- Euler wide variety: It is a feature of the binary photograph
- Extent: The place and length of the rectangle.
- Solidity: It is the percentage of the pixels within the raised shape area that are likewise in a given place. It is calculated by, Area / raised area

Apply region properties to degree a few of these properties and then do away with areas primarily based on their property values.

Step 3: Use Stroke Width Variation to exclude non-textual regions from a picture.
Stroke width props are another popular parameter used to differentiate between textual and non-textual material. A measurement of the extent of stroke width that measures the amount of curves and strokes is called stroke width. Textual areas have the potential to have few stroke interpretations, while non-textual information areas tend to have a lot of them. Examine the stroke width of each of the detected MSER areas to help understand how to utilize the stroke width to eliminate non-textual subject regions.

**Step 4: Merging of Textual Regions for Text Detection Results.**

In this step, all the detection consequences are composed of person textual content. One technique for arranging character textual areas into text traces is to first find neighbouring textual regions and then shape a bounding container around those areas. To locate neighbouring areas, increase the bounding bins computed in advance with area props. This makes the surrounding boxes of neighbouring textual content regions.

**Connected Component Analysis**

- Once area surroundings have been detected, it's miles often beneficial to eliminate areas that are not separated.
- Any group of pixels that is not distinguished by a border is referred to as a name unit.
- A connection between the pixels of a maximal position is called a joint component.
- Combining all of the parts of an image to form a distinct subset of the collection of components for certain image processing systems, segmentation is highly useful.

**V. RESULT ANALYSIS:**

This developed system automatically describes the text of images which may be a challenging task. It could provide more accurate and compact text present in images.

The performances of result analysis for some of the text images are:

The process that is initially happened when the image is given as input to the system is explained by considering an image as an example,

The next step is to transform it to a grayscale picture, which means that the input color image is turned into a black and white image by feature extraction through MSER.
Due to the removal of such non-text regions dependent on the regions are that contain both text and non-text the geometric properties. One of the regions in random is automatically considered and stroke width variation is applied to it. Later that region is converted into a skeleton image to easily classify the characters. Here from the above image one of the characters named D is recognized.

We applied subplot between the region image and the stroke width image.

We obtain the following picture after extracting all non-text regions from the image using stroke width variation.
Expanding bounding boxes such that each character will be under each bounding box.

The finally detected text by expanding the individual bounding boxes in order identify the complete word.

OUTPUT SCREEN:

```
'HAND DICAPPED
PARKING
SPECIAL PLATE
REQUIRED'
```
VI. CONCLUSION

In this task, we have diagnosed a textual content from photograph using an algorithm MSER and both the Geometric as well as Stroke width Variation. MSER contains an appropriate textual reputation performance. We have used MSER to detect the textual content areas after which OCR is used to become aware of the characters.

The future upgrades are:

Extension to Video. The strategies that are utilized in this challenge are designed for recognition of textual content in images. It have to be cited that videos have dependency amongst frames, for example, if a body include a phrase, the next frame is more likely to comprise the identical phrase in almost equal location. Modelling such dependencies and the use of inherent spatio-temporal cues in a power minimization framework is every other possible manner to broaden the proposals of this thesis.

Multi-Script Scene Textual content understanding. In many countries like India, multiple scripts are utilized in diverse regions of the country. Robustly recognizing many of these scripts remain an open ended hassle for printed text domain, while scene textual content popularity is even greater challenging.

REFERENCES