ISSN 2229-5518

An Automatic System to Recognize Kannada Natural Sign Board Characters

[1]Pawan Kumar Ganjhu, [1]Pallavi Pratik, [1]Saurav Kumar, [1]B V Gowravi and [2]Dr. Purohit Shrinivasacharya

[1]Student of Information Science and Engineering, Siddaganga Institute of Technology, Tumakuru

[2]Associate Professor, Dept. of Information Science and Engineering, Siddaganga Institute of Technology, Tumakuru-752103, Karnataka, India

Abstract- Right from the dawn of human civilization, people have started migrating from places to places for various reasons. Our Country; India being a nation with diverse culture and languages, people on moving away from their native may find it very difficult to understand different languages. Let us say a non-native speaker of Kannada language comes to the state of Karnataka may not be able to read the Kannada script. So, this system which we propose will read Kannada texts present in natural scenes with the aim to provide assistance to the non-native speakers of Kannada language. The area of natural scene text recognition aims on the problem to recognize random texts in images. Examples of scene text may include street signs, name of shops, grocery item labels, and name plates etc. Smart phones and cameras are increasingly used for their intense feature and also for their ability to recognize and capture the text. The purpose of this project is to develop different ways for improving natural scene text recognition. Here, we try to focus on recognition of characters in such cases that are not handled effectively by traditional OCR (Optical Character Recognition) techniques. We design a database of images which contains annotated Kannada characters inclusive of alphabets, digits and other symbols. The problem is depicted as an object compartmentalization skeleton comprising of a database having characters as image. We fetch the potential of different attribute based on Support Vector Machine taxonomy and nearest neighborhood. We will do this by including new types of information into models and by considering how to compose simple components into highly active systems. Mainly we will focus on three areas of scene text recognition, each with a decreasing number of prior conjectures. Firstly, we will introduce two techniques for character recognition, where word and character bounding boxes will be used. Next we will look at word recognition, where only word bounding boxes will be used. We want to develop a new technique for fragmenting text for these images called bilateral regression segmentation. Lastly, we will remove the assumption that words have been located and describe an end-to-end system that detects and recognizes text in any natural scene image. We demonstrate the performance of the system designed using 200 sample natural scene images to train. It precludes the necessity of costlier gathering of data and annotating.

Keywords- Object Recognition, Character Recognition, Kannada Language, Text Extraction, Text Recognition, Text Documentation and Text Translation

1 INTRODUCTION

This paper aims to propose work towards automatic recognition, extraction and translation Kannada language natural sign board characters. In particular, it focuses on recognizing every single character in textual images. Figure 1, 2 and 3 apotheosizes the reason for this to be a flinty job. Even after ignoring the difficulty of text partitioning momentarily, the sources which brings problems are as follows: (i) font style and font size (ii) graphics in background and foreground (iii) camera alignment yielding to geometric distortion (iv) illumination (v) resolution of image (vi) removal of unwanted objects (vii) identifying threshold value for different images and (viii) edge detection. The above listed features give rise to the problem of object recognition. Henceforth, this technique cannot be used. Moreover, within the realm of possibility of such systems, OCR systems have been designed and developed for very less languages. Recognition of all languages of Indic origins are beyond their capability.

Figure 1: Sample Source Image in Our Dataset

Figure 2: Examples of Different Styles of Texts in Different Dimensions

Figure 3: Sample images of Kannada Characters belong to separate classes. As per Kannada Kagunitha, vowels on combining with consonants, it creates new character.

Various problems are to be solved so as to read Kannada texts from natural scene images in conjunction with finding text, identification of characters, partitioning of words. We try to highlight via this paper to recognize characters; documenting recognized characters into editable file format and lastly translating the phrase/ sentence recognized to another language (say English here). A standard database comprising of images/ characters of Kannada language is introduced. To measure the practicability of presenting the work as Kannada Character recognition, we establish a paradigm to measure the effectiveness of different attribute based on a standard data set as described above. The result delineates that indeed the confinement of character acknowledgment assignment is a dull work. The number of classes for Kannada characters is almost 619 counting consonants, vowels, digits and other images utilized with exceptionally small inter-class disparity as finished by Figures 2 and 3. This issue is eminently exceptionally ghastly for Kannada dialect where two typical characters in the letter set can contradict fair by substitution of a single speck, hyphen, bar, comma, accentuation like structure as appeared in Figure 4. The distinguished characters will be diverted to an editable record arrange. These writings can be afterward utilized for different purposes as required. The recognized characters can be archived in Latin script.

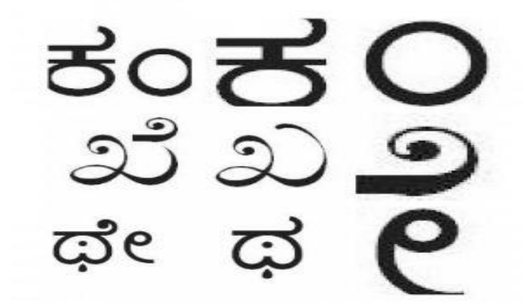

Figure 4: Characters followed by the attachments making it two different syllables

2 RELATED WORKS

The work of Kannada characters acknowledgment in characteristic scenes is related to issues considered in camera arranged record examination. Larger part of the work in scene content acknowledgment is particularly based [10], [9], [4] and [3] on finding and adjusting the content zones and taking after the OCR application strategies. [8] Such approaches are in this manner limited to environment where OCR works well. From this time forward such approaches are controlled to environment where OCR works appropriately. In expansion to amendment prepare, it does not straightforwardly relate to our work, as it points on discovery of printed characters. The edge detection is carried out by the technique described by J Canny, called canny edge detection technique [24] and also by an improved edge detection technique [23].

The technique for static recognition of hand written characters have been efficiently solved by intra-class variation due to non-identical styles of writing [15], [14]. Such scenes prototypically assume only a finite number of appearance classes, unable to resolve differences in foreground/background color and texture, especially the graphics present. This is achieved by identification and removal of extraneous graphics in a commercial OCR operation [25]. For occurrence, [16] we have utilized cognizance from NLP and display a Markov chain system

ISSN 2229-5518

for parsing pictures. [5] Presentation of composition machines for developing probabilistic progressive picture models. This makes a difference in obliging relevant connections. This approach permits re-usability of parts among different substances and non-Markovian disseminations. [16] Proposed a strategy that amalgamates picture highlights and dialect data a single demonstrate and coordinating disparity data between character pictures.

Acknowledgment of digits utilizing pipelines based on crude pictures classifications have been broadly utilized [12]. [21] By shape coordinating procedure, this is too done [1]. The classification is carried forward by HOG technique, known as Histogram Oriented Gradient. In this line by line detection of characters and words is done.

3 DATA SETS

We aim to recognize Kannada characters from natural scene images. To do so, we design a database containing images of natural scene having Kannada characters. These images have been gathered from the streets of Tumkur and Bangalore, India. The natural scene images comprises not only of street symbols but also of sign boards, hoardings, posters, pamphlets, banners, name plate, number plate etc. However collection and annotation of huge sample of images is a costly as well as costly job. So, we acquired a database of characters generated by computer fonts of size 72 as shown in Figure 5.

Figure 5: The Standard Dataset containing all possible Kannada language Characters and Digits

English language has characters separately in two cases namely upper case and lower case, but in case of Kannada Language it’s not the same. Kannada language alphabet does not have the system of upper case and lower case characters. It has 37 consonants and 16 vowels. By combining the vowels with consonants, it generates around 603 distinct classes. It has numerals from 0 to 9 which can further combined to generate infinite number of terms. Digits can be identified separately.

3.1 Data Set of Natural Image

We have captured 200 images from mobile phone and digital camera. Sample images are depicted in Figure 1.

We have applied two methods of segmentations: canny edge detection technique and rectangular bounding boxes as shown in Figure 6. Both techniques are equally likely to be good. The canny edge detection technique detects objects other than characters and digits.

Figure 6: Canny Edge and Rectangular Bounding Boxes

Out of 200 images of natural scene in our database; here we have tabulated for 20 images which has 209 numbers of characters are present. The proposed system has identified 185 numbers of characters. The performance of the system is 88.52% and is tabulated separately in Table 1.

Image ID	No. of Characters present in an image	No. of Characters Identified	Performance (%)
1	6	3	50
2	2	1	50
3	4	4	100
4	8	8	100
5	7	7	100
6	49	49	100
7	49	49	100
8	2	1	50
9	3	0	0
10	10	10	100
11	14	14	100
12	9	9	100
13	5	5	100
14	6	6	100
15	4	4	100
16	4	1	100
17	10	10	100

ISSN 2229-5518

18	4	4	100
19	7	0	0
20	6	0	0
Total	209	185	88.52

Table 1: Chart Showing Success and Failure Ratio

3.2 Font Datasets

The standard datasets are created using the computer and the font size is set to 72. The sample of this dataset is shown in Figure 5. This dataset is referred as standard dataset for matching the characters from natural scenes.

4 REPRESENTATIONS AND FEATURE EXTRACTION

For question category acknowledgment, the well known procedure utilized to speak to picture substance is named as Bag-of-Visual-Words. The objects are spoken to as histogram of highlight tallies. It is accomplished by K-means calculation, (for other strategies of producing the lexicon see [7]) utilizing picture corpus. One can at that point outline each highlight extricated from an picture onto its closest visual word and speak to the picture by a histogram over the lexicon of visual words.

We total the set of visual words per class to frame the vocabulary.

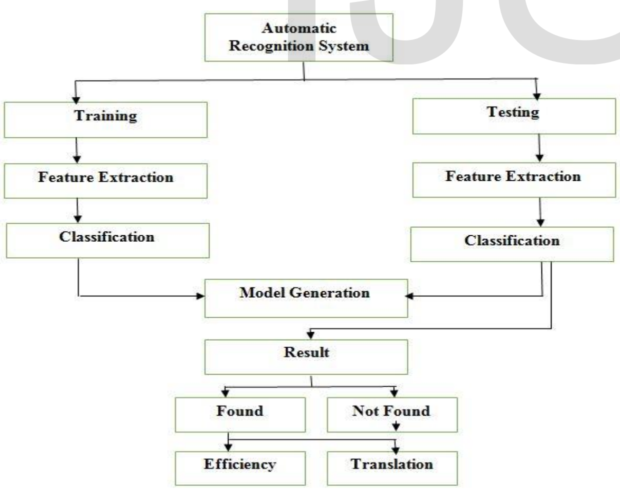

Figure 7: Flow of Training and Testing

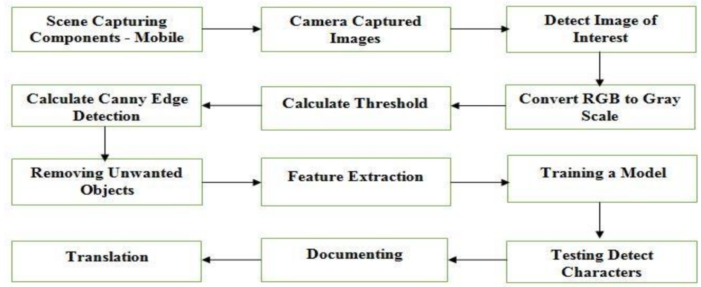

Figure 8: Sequential flow of execution of different features

4.1 Features

We have tried features like Shape Context and Geometric Blur, these are based on shape and edge detection. SVM classifier technique has also been used for its higher efficiency and faster processing unlike SIFT which is comparatively slow. The methods like canny edge detection, change of RGB image to Gray scale image, calculation of threshold for each image, highlights utilized for speaking to surface, such as channel reactions, patches and Spin Images have moreover been utilized [20]. We have also tried to remove the unwanted objects like graphics from foreground and background from the natural scene image. We have analyzed various commonly used parameters and feature detection technique used for each descriptor which have been described below. We have also tried to remove the unwanted objects like graphics from foreground and background from the natural scene image. We have analyzed different commonly utilized parameters and highlight location procedure utilized for each descriptor which has been depicted underneath.

Convert RGB to Gray Scale [22] To convert RGB to grayscale, the average of all the three i.e. R, G and B is computed. To do so we add R with G with B and then divide it by 3 to obtain the grayscale. For example: Figure 9 and Figure 10 describes the conversion of RGB to Grayscale respectively.

Geometric Blur (GB) [2] Geometric blur is simply an average over geometric transformations of a signal. It is done by sampling method called feature extraction which is same as SC. It is isolated into distinctive locales and at that point the edge introductions are checked with distinctive obscure figure.

Calculating Threshold [28] Calculation of threshold of an image is done by separating a picture into closer view and foundation independently. This handle changes over the grayscale picture into binary picture. Example: Figure 11 shows how threshold image looks like.

Shape Contexts (SC) [1] SC is a feature descriptor. It is used for object recognition and description of shape that permits measuring shape similarity. We do it by Sobel edge detection technique using log-polar histogram. We use histogram of oriented gradients (HOG) technique as well.

Canny Edge Detection [24], [23] By the help of Canny Edge Detection algorithm, we detect edges of objects in an image. It is a multi-stage algorithm. It is useful as it extracts structural information.

Removal of Unwanted Objects [25] The subject of removing unwanted objects from natural scene images without generating any possible distortion has been handled.

Histogram of Oriented Gradients (HOG) [26], [27] It is a highlight descriptor utilized in computer vision and picture handling for the reason of protest location. The procedure tallies events of slope introduction in localized parcels of a picture. features = extractHOGFeatures(I) It returns extricated Hoard highlights from a truecolor or grayscale input picture, I . The highlights are returned in a 1-by-N vector, where N is the Hoard highlight length. The returned highlights encode neighborhood shape data from locales inside a picture.

Spin Image [11], [6] It is a two dimensional histogram encoding method for conveyance of picture brightness. The two dimensional of the histogram is d, separate from the center point, and i the escalated esteem. We have utilized d=11 and i=5 for concentrated esteem, coming about in 55-dimensional descriptors.

Maximum Response of Filters (MR8) [18] It is a surface descriptor based giving 8D vectors, on a set of 38 channels but as it were 8 responses.

Patch Descriptor (PCH) [19] It is the least complex thick include extraction strategy. For each position, the crude n × n pixel values are vectorized, producing an n2 descriptor. We utilized 5×5 patches.

5 RESULTS AND EXPERIMENTS

Here we describe the experiments under different schemes as mentioned above. The three classification schemes used are: (a) multiple kernels learning (MKL); (b) support vector machines (SVM); and (c) closest neighbor (NN) classification utilizing c2 measurement as a closeness degree (d) Histogram of Oriented Gradients (HOG).

The following outputs have been obtained using a sample natural scene image. The pictures of the output are given in sequential order.

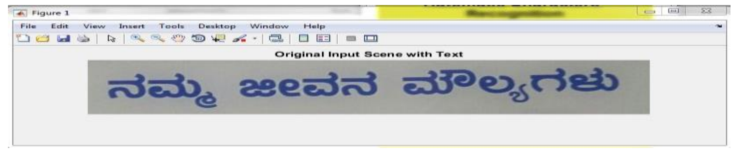

Figure 9: Original Image from Sample Database

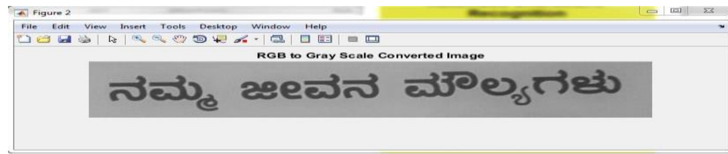

Figure 10: Converting RGB to Gray scale Image

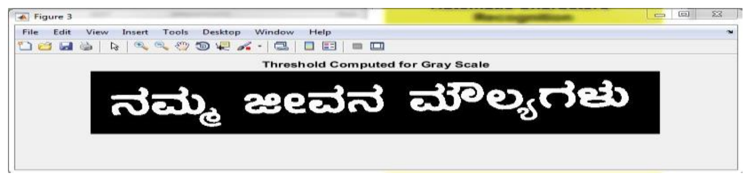

Figure 11: Calculating Threshold of Image

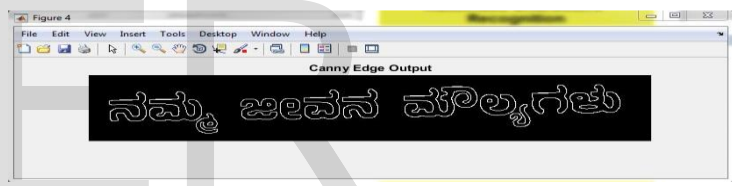

Figure 12: Canny Edge Detection

Figure 13: Removing Unwanted Objects less than 10 pixels

Figure 14: Feature Extraction

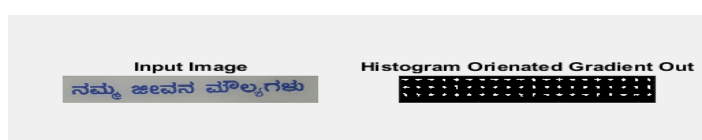

Figure 15: Feature Extraction – HOG

```
Computing descriptors for 251 training windows:   251

Training linear SVM classifier...
 Iteration   FunEvals     Step Length    Function Val        Opt Cond
         1          2     2.41653e-03     7.71065e-01     3.57309e+02
         2          3     1.00000e+00     8.29996e-02     2.10283e+01
         3          4     1.00000e+00     7.86730e-02     1.58662e+01
         4          5     1.00000e+00     7.01067e-02     1.26109e+01
         5          6     1.00000e+00     6.25725e-02     1.32498e+01
         6          7     1.00000e+00     5.59889e-02     6.99075e+00
         7          8     1.00000e+00     5.45956e-02     1.93974e+00
         8          9     1.00000e+00     5.43152e-02     1.88109e+00
         9         10     1.00000e+00     5.42197e-02     1.47152e+00
        10         11     1.00000e+00     5.41132e-02     2.59030e-01
        11         12     1.00000e+00     5.41052e-02     2.14562e-01
        12         13     1.00000e+00     5.40994e-02     1.30836e-01
        13         14     1.00000e+00     5.40953e-02     6.70054e-02
        14         15     1.00000e+00     5.40897e-02     1.24702e-01
        15         16     1.00000e+00     5.40791e-02     2.84086e-01
        16         17     1.00000e+00     5.40564e-02     5.05524e-01
        17         18     1.00000e+00     5.40010e-02     7.97585e-01
        18         19     1.00000e+00     5.38843e-02     1.22656e+00
        19         20     1.00000e+00     5.37165e-02     1.19194e+00
        20         22     1.67050e-01     5.36715e-02     1.32288e+00
        21         23     1.00000e+00     5.35028e-02     6.06172e-01
        22         24     1.00000e+00     5.34694e-02     1.39143e-01
        23         25     1.00000e+00     5.34663e-02     1.87681e-02
        24         26     1.00000e+00     5.34662e-02     9.18153e-03
        25         27     1.00000e+00     5.34662e-02     7.25667e-03
        26         28     1.00000e+00     5.34662e-02     9.40461e-04
Directional Derivative below TolX

Training accuracy: (249 / 251) 99.20%
```

Figure 16: Training

```
Searching 2096 detection windows...
  Image Scale 1.00, 384 windows - 384 matches total, 18.3% done
  Image Scale 0.95, 345 windows - 729 matches total, 34.8% done
  Image Scale 0.91, 273 windows - 1002 matches total, 47.8% done
  Image Scale 0.86, 240 windows - 1242 matches total, 59.3% done
  Image Scale 0.82, 209 windows - 1451 matches total, 69.2% done
  Image Scale 0.78, 153 windows - 1604 matches total, 76.5% done
  Image Scale 0.75, 128 windows - 1732 matches total, 82.6% done
  Image Scale 0.71, 105 windows - 1837 matches total, 87.6% done
  Image Scale 0.68, 84 windows - 1921 matches total, 91.7% done
  Image Scale 0.64, 65 windows - 1986 matches total, 94.8% done
  Image Scale 0.61, 48 windows - 2034 matches total, 97.0% done
  Image Scale 0.58, 33 windows - 2067 matches total, 98.6% done
  Image Scale 0.56, 20 windows - 2087 matches total, 99.6% done
  Image Scale 0.53, 9 windows - 2096 matches total, 100.0% done
  Image scale 0.51 is not large enough for descriptor, stopping search.
Image search took 5.52 seconds
```

Figure 17: Step wise process of Testing

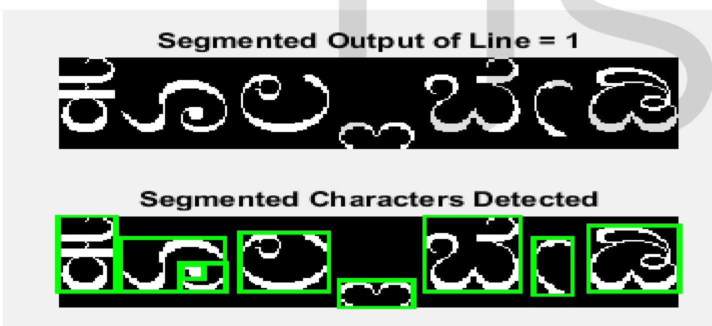

Figure 18: Testing

CONCLUSIONS

In this paper, we have handled the issue of recognizing characters in images of natural scenes. The database of test common scene images containing Kannada characters have been captured in Bangalore and Tumkur, India. Clarifying of normal pictures for training purposes can be costly and time expending. The techniques listed in neural network is not that effective when it comes to efficiency, so we have used SVM whose efficiency and computational time is more.

By distinguishing the downsides of the existing framework, we have handled the issue of recognizing Kannada characters in normal scene. The benefits of these experiments are- It increases the efficiency and effectiveness of work there by it saves time. Documents can be text searchable and editable. It can help non native speaker learn Kannada language and communicating with native people will be easy. It has got social benefits as well.

REFERENCES

[1] Belongie, S., Malik, J., and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[2] Berg, A. C., Berg, T. L., andMalik, J. (2005). Shape matching and object recognition using low distortion correspondence. In Proc IEEE Conf on Computer Vision and Pattern Recognition, San Diego CA, June 20-25.

[3] Brown, M. S., Sun, M., Yang, R., yun, L., and Seales,W. B. (2007). Restoring 2d content from distorted documents. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[4] Clark, P. andMirmehdi, M. (2002). Recognising text in real scenes. International Journal on Document Analysis and Recognition, 4:243–257.

[5] Jin, Y. and Geman, S. (2006). Context and hierarchy in a probabilistic image model. In Proc IEEE Conf on Computer Vision and Pattern Recognition, New York NY, June 17-22.

[6] Johnson, A. E. and Herbert, M. (1999). Using spin images for efficient object recognition in cluttered 3d scenes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(5):433–449.

[7] Jurie, F. and Triggs, B. (2005). Creating efficient codebooks for visual recognition. In Proc 10th Int Conf on Computer Vision, Beijing, China, Oct 15-21.

[8] Kise, K. and Doermann, D. S., editors (2007). Proceedings of the Second International Workshop on Camerabased Document Analysis and Recognition CBDAR, Curitiba, Brazil. http://www.imlab.jp/cbdar2007/.

[9] Krempp, A., Geman, D., and Amit, Y. (2002). Sequential learning of reusable parts for object detection. Technical report, Computer Science Department, Johns Hopkins University.

[10] Kumar, S., Gupta, R., Khanna, N., Chaudhury, S., and Joshi, S. (2007). Text extraction and document image segmentation using matched wavelets and mrf model. IEEE Transactions on Image Processing, 16(8):2117–2128.

[11] Lazebnik, S., Schmid, C., and Ponce, J. (2005). A sparse texture representation using local affine regions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(8):1265–1278.

[12] Le Cun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324.

[13] Lowe, D. G. (1999). Object recognition from local scaleinvariante features. In Proc 7th Int Conf on Computer Vision, Corfu, Greece.

[14] Pal, U., Sharma, N., Wakabayashi, T., and Kimura, F. (2007). Off-line handwritten character recognition of devnagari script. In International Conference on Document Analysis and Recognition (ICDAR), pages 496– 500, Curitiba, PR, Brazil. IEEE.

[15] Plamondon, R. and Srihari, S. N. (2000). On-line and offline handwriting recognition: A comprehensive survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1):63–84.

[16] Tu, Z., Chen, X., Yuille, A. L., and Zhu, S. C. (2005). Image parsing: Unifying segmentation, detection, and recognition. International Journal of Computer Vision, Marr Prize Issue.

[17] Varma, M. and Ray, D. (2007). Learning the discriminative power-invariance trade-off. In Proceedings of the IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil.

[18] Varma,M. and Zisserman, A. (2002). Classifying images of materials: Achieving viewpoint and illumination independence. In Proceedings of the 7th European Conference on Computer Vision, Copenhagen, Denmark, volume 3, pages 255–271. Springer-Verlag.

[19] Varma,M. and Zisserman, A. (2003). Texture classification: Are filter banks necessary? In Proc IEEE Conf on Computer Vision and Pattern Recognition, Madison WI, June 18-20, volume 2, pages 691–698.

[20] Weinman, J. J. and Learned Miller, E. (2006). Improving recognition of novel input with similarity. In Proc IEEE Conf on Computer Vision and Pattern Recognition, New York NY, June 17-22.

[21] Zhang, H., Berg, A. C., Maire, M., and Malik, J. (2006). SVM-KNN: Discriminative nearest neighbor classification for visual category recognition. In Proc IEEE Conf on Computer Vision and Pattern Recognition, New York NY, June 17-22.

[22] C. Saravanan (2010) Second International Conference on Computer Engineering and Applications titled Color Image to Grayscale Image Conversion.

[23] Bing Wang and ShaoSheng Fan (2009) Second International Workshop on Computer Science and Engineering on An improved CANNY edge detection algorithm.

[24] John Canny (1986) IEEE Transactions on pattern analysis and machine intelligence, Vol. Pami-8, No. 6 on A Computational Approach to Edge Detection.

[25] Ray R. Hashemi, Arlie Epperson, Steve Jones, Leij In, John Talbrut on Identification and Removal of extraneous graphics in a commercial OCR operation.

[26] Jinse Shin, Dongsung Kim, Christoph Ruland (2014). Content based image authentication using HOG feature descriptor.

[27] Ankit Kumar Sah, Showmik Bhowmik, Samir Malakar (2017). Text and non-text recognition using modified HOG descriptor.

[28] Sheng Lu, Member, IEEE, Xinnian Chen, Jørgen K. Kanters, Irene C. Solomon,and Ki H. Chon, Senior Member IEEE (2008). Automatic Selection of the Threshold Value r for Approximate Entropy.