OCR Accuracy on Document Images utilizing Modern Enhanced APPROACH

Murtadha Abdulkareem Mahdi	Software Engineering Department , Engineering and Architecture Faculty , Altınbaş University mnee8542@gmail.com
Asst. Prof. Timur İnan	Software Engineering Department , Engineering and Architecture Faculty , Altınbaş University Timur.inan@altinbas.edu.tr

ABSTRACT

Electronic camera and versatile archive picture getting are progressing winning styles emerging in the area of texture location with Optical Character Recognition. Every so often, such correspondence joins different bends and conveys deficiently filtered texture or texture-photograph pictures and typical pictures, impelling a scheming OCR digitization. In here work, A remarkable nonparametric has been introduced with solo methodology to make up for bothersome record picture turns expecting to ideally additionally cultivate OCR accuracy. Our methodology depends upon an astoundingly practical heap of report picture upgrading strategies to recover bending of the whole record picture. Notwithstanding, we suggest a neighborhood marvel and detachment change strategy to manage lighting arrangements and the whimsical spread of picture enlightenment appropriately. Second, we utilize an upgraded greyscale change assessment to change our report picture to greyscale standard. At last, we hone the obliging data in the subsequent greyscale picture utilizing Un-sharp Masking strategy. At last, an ideal commonly binarizing strategy that implemented to setting up the final recorded image for OCR detection. The suggested strategy might usually further planting or refining optical character recognition precision as well texture discovery rate. For illustrating the accuracy of this approach, a serious experimentation upon a typical dataset has been presented.

Keywords:	optical character recognition, Improve OCR accuracy, text detection, document image enhancing, Document image distortion

1. Preface

With the progress of scientific research and the great scientific acceleration in the digital image processing field and the development of software and techniques used in computer science, the need has arisen to find modern techniques and approaches to detect phrases and words and distinguish them electronically. There are many applications for the technique of detecting scripts and written texts in many areas, including security, academic and various service areas.

Removing the texture from the report picture and handling it further for making it skilled to be perceived is the undertaking of advanced picture handling. Solid letter division, removable of different commotions, skeletonizing the letter and normalizing everything goes under the course of advanced picture handling. Preceding character

recognition, it is important to disengage the singular letters along texture picture. Computerized picture handling alludes handling of two-layered picture. Advanced Images are the photos that have been changed over towards a PC discernible twofold organization comprising of sensible 0's and 1's [4]. Improved image processing has an expansive amount of utilization in various areas, for example, the remote detecting along satellite units with other shuttle's, images transmissions as well bandwidth for a business applications, clinical handling, radar, sonar, and advanced mechanics and robotized investigation of modern parts [3]. The computerized picture handling activity might be comprehensively assembled towards five crucial classes, which are as per the following:

1. Image Portrayal and Demonstrating

Picture portrayal and schemes manage the ultimate common way of tracking down legitimate approaches to numerically address the picture. Fundamentally, there are two approaches for addressing an picture. The outside portrayal manages the limit of the picture also utilized when the essential spotlight is on the shape characteristics. The inward portrayal manages pixels inside the locale as well as utilized when the essential center is provincial properties, for example, shading with texture. In any case, once in a while utilizing both portrayal techniques is fundamental.

Picture improvement alludes to the approach involved with rolling out an improvement to the highlights of the picture like edges, limits, or differentiation to make a realistic presentation further valuable for show as well as examination [3]. The improvement interaction doesn't build the intrinsic data content of information yet it builds the powerful scope of the picked highlights so they might be recognized without any problem. Picture upgrade incorporates dark stage as well as differentiation control, sound decrease, edge crispening with honing, separating, interjection also, amplification, pseudocoloring, etc.

2.Noise Expulsion/Decrease

Commotion is an random variety of with

noticeable as image grains as well as image intensity. It might produce to emerge in the picture as impacts of essential physical science like lights photon nature or nuclear power of hotness innards the picture sensing units. It might deliver at the hour of catching or then again picture transmission [5]. In a picture, clamor alludes to the pixels showing unique force values rather than genuine pixel values. It is presented in the picture at the hour of picture procurement or transmission.

In a parallel picture, we allude commotion as a dark pixel that ought to be white or the other way around. The commotion ought to be taken out in beginning phases, on the off chance that not it is spread to the further strides of handling, then it might influence the activity of the framework. There are fundamentally two various kinds of clamor. Salt commotion, when the pixel should be dark however it is addressed by a white pixel. Pepper commotion, when the pixel should be white yet is addressed by the dark pixel. We might utilize different kinds of channels to eliminate commotion along the pictures, for example, mean channel, middle channel, Laplacian filter, mode channel and so on

Text location strategies might be utilized to separate or potentially record data to recover further data. In any case, naturally separating texture along archive pictures created by your advanced camera or versatile camera is an extremely challenging assignment. Empowering texture extraction along these records with high exactness will give the empowering innovation to various applications like better ordering, online accessible reports, indoor/outside scene understanding [10], as well as texture recognition in normal pictures [11]. OCR innovation gives preferred search as well as recovery capacities over was beforehand conceivable [12]. Dissimilar to exemplary examined reports, texture identification in record pictures delivered by computerized cameras with portable cameras is an extremely challenging errand, also a long way along a total arrangement. Complex foundations, lopsided lighting, with virtually bothersome texture textual styles, sizes, as

well as bearings present significant troubles in any event, for the ultimate current message recognition approaches.

In this study, we center around fostering a methodology that tackles the issue of separating text along report pictures. Our methodology enjoys a few benefits. To start with, we suggest to involve another scheme for enlightenment change based on Contrast Limited Adaptive Histogram Equalization (CLAHE) to work on the contradiction of the general items introduces in the handled report picture. Next, we use Luminance calculation to enhance grayscale change for textures extractions. At last, to improve textures subtleties with corners we utilizing Un-sharp concealing channel. As a last upgrade steps, Otsu Binarization calculations might be utilize as a strong strategy for Cleaning and brightening report foundation. The suggested technique is assessed utilizing a typical accessible datasets [13].

The remaining of the paper is arranged as succeeding. Literature review has been presented in Section 2. The suggested strategy is illustrated in detail in Section 3. Section 4 presents the simulation results of the eveluated   performance examination have been interoduced. At last, in conclusions has been presented in Section 5,.

2.   Literature Review

A portion of the significant articles with papers that examined the texture recognition as well as discovery have been gathered with assessed in the accompanying sections.

Wang et al. [7] gave a beginning to STR terminal pipeline, where they achieve multi-scale texture area through SW grouping. Highlights are first removed by picked entries in a HOG descriptor handled at the window area. Then, Random Ferns is implemented to assess the likelihood of letter in the window area.

Container et al. [8] assessed the textures remaining conviction with scale data by means of SW. Along that point onward, a conditional random field (CRF) scheme is suggested to filter through the non-text parts.

Likewise, Mishra et al. [9] applied a typical SW approach against letter viewpoint proportion before recognize likely areas of letters in scene picture.

Wang et al. [10] implemented a convolutional brain network (CNN) scheme with SW intend to acquire up-and-comer lines of texture in given picture, also in such manner gauge texture areas.

Jaderberg et al. [11] likewise implemented CNN in SW design to figure texture saliency map, that remains a similar objective as the first picture through zero-cushioning. Behind such text bouncing boxes might be created based on these saliency maps.

Epshtein et al. [12] acquainted SWT administrator with register the width of the ultimate plausible stroke for picture pixel. Shrewd edge identifier is early utilized for edges finding in picture. Prior finding each edge pixels in the opposite inclination bearing, strokes are seen as fruitful with these pixels are gathered towards letter competitors.

Neumann et al. [13] provided a portrayal for letter ID issue, i.e., seeing all lining areas in picture with the ultimate objective that likelihood that the plan tends to texture has a neighborhood greatest. Based on the depiction, MSER classifier is prepared to notice locale including letters. At last, post-taking care of with affiliation rules are implemented to merge the competitor letters towards texture line. MSER procedure needs less priori data also is major areas of strength for further language also arranged texture. To resolve issues on hazy pictures or letters with low contrast, similar creators executed letter ID in every outer locale (ERs) rather than basically in MSERs [13,14]. They use steadily processed descriptors as highlights to prepare a consecutive classifier, that might minimize the high bogus certain rate continuously.

Yin et al. [15] suggested a quick MSERs pruning calculation, that might fundamentally minimize the amount of letter possibility to be dealt with.

Character applicants are bundled towards texture competitors by the single-associate gathering calculation, whose distance stacks also grouping breaking point might be

naturally scholarly. Such new MSER based methodology is further overwhelming with compelling for texture disclosure. CC systems, taking everything towards account, effectively convey with different non-text parts. Hence, precisely filtering through the misleading up-sides is basic to the achievement of this social event of procedures.

Huang et al. [16,17] implemented CNN to gain unquestionable stage highlights along the MSREs parts in picture. These parts show high discriminant capacity with strong strength against muddled foundation ones. Also, SW scheme with non-maximal covering (NMS) are consolidated in the CNN classifier to deal with the issue of different letters affiliation.

Gomez et al. [18] applied the MSER calculation to without skipping a beat acquire the underlying segmentations of image. After that they suggest a texture express specific inquiry system, that might pack the underlying areas by agglomerative gathering in an order where every center characterizes a potential word speculation. At last a positioned rundown of proposals zeroing in on the best speculations is obliged texture recognition.

Busta et al. [19] suggested a stroke marker, that initially notices stroke focal issues as well as afterward, by then, includes them to acquire stroke segmentation for scene texture. These illustrate the contrasted against the conventional MSER procedure, applying stroke express focal issues might perceive further letters with less area segmentations.

Zhang et al. [20] suggested a texture finder which exploits the equilibrium property of letter social affairs. Special corresponding to conventional approaches that primarily exploit the specifications of individual letters or strokes, such advanced finder might utilize setting info along scene image to execute texture lines extraction.

Cho et al. [21] introduced Canny texture marker utilizing multi-stage calculation. Trama focus technique is first used to separate letter competitors as numerous as possible, also the covered applicants are killed by NMS. Along that point forward, the competitors are named strong texture, feeble texture or non-text with twofold edge. Other than strong texture, competitors with low conviction, for instance powerless texture, are picked by hysteresis. At last, the persevering through message competitors are accumulated to make sentence.

Fabrizio et al. [22] introduced a hybrid texture finder, that embraces CC procedure to produce textures applicants as well furthermore implements surface examination to shape texture string or dispose of bogus up-sides. CCs in picture might be early gotten by using the switch planning morphological segmentations (TMMS) calculation. A shape descriptor based on quick wavelet crumbling is utilized to order every CC as letter or non-letter. Along that point onward, a movement of surface highlights are applied to prepare an assist vector with machining (SVM) for post-taking care of.

He et al. [23] made contrast-improvement maximally stable preposterous locales (CE-MSERs) marker, that expands the ordinary MSERs by upgrading power contrast between texture examples also foundation. Plus, they prepared a texture-consideration CNN that might extricate evident stage highlights including texture locale cover, letter name, with double texture/non-texture data. The two plans are integrated to approach a fruitful texture distinguishing proof scheme.

Segmentations issue. Those apply a FCN scheme depend upon comprehensively settled corner area (HED) to make worldwide guides, counting data of texture area, particular letters with their affiliation. Also, the suggested calculation might perceive multi-arranged as well as twisted textures in scene image. He et al. [24] presented the cascaded convolutional texture networks (CCTN), that applies dual systems to do coarse-to-fine segmentation for scene images. Noting such coarse networks yields a for each pixel heat-map showing the area with likelihood of texture occurrence, as well as the fine network yields two intensity maps for definite texture distinguishing proof.

Zhou et al. [25] additionally suggested a FCN based scheme for scene texture revelation. Various channels of pixel-stage texture score guide as well as math might be created in this

scheme, that is versatile to deliver either word stage or line stage figures. Also, a area mindful NMS with less period unpredictability is suggested for post-taking care of.

Zhang et al. [26] likewise execute texture area with coarseto-fine technique. They initially use a FCN (named Text-Block FCN) to anticipate the notable guide of texture blocks. Later to MSER procedure is implemented to remove multi-organized texture line competitors. At last, they learn other further modest FCN (named Character-Centroid FCN) to give the letter centroid info, based on that bogus texture line up-and-comers might be dispensed with.

Qin et al. [27] suggested a texture pointer based on the outpouring of two CNNs. Text locale of concern are early made by a FCN also afterward, by then, reexamined to a squared frame against settled area. The accompanying phase is the text recognition framework, i.e., preparing a YOLO-type networks to produce arranged rectangular skipping boxes for all texts. At last, a NMS step is executed to deal with covering skipping boxes.

He et al. [28] suggested a FCN design for multi-arranged scene texture area against dual errands. The arrangement mission completes down-examined segmentations among texture as well as non-text for entered picture, also the relapse mission concludes the vertex directions of quadrilateral texture limits along direct relapse.

Dai et al. [29] introduced a pointer based on joined texture segmentations network. Highlights of every picture is initially removed along a resnet-101 spine, as well as afterward, by then, amazed include maps are joined as well as dealt with to the locale suggested networks (RPN) for texture area of interest (ROI) age. The total engineering might execute texture distinguishing proof also segmentations at the same time as well as give gauges both in the pixel with text phase.

Deng et al. [30] suggested a scene texture finder (named PixelLink) based on example segmentations. The Single-Shot Indicator (SSD) [29] such engineering is applied to separate highlights also achieve text/non-text assumption with association estimate. The expected positive pixels are solidified towards texture occasions by expected positive associations. At last, texture skipping boxes are produced clearly along the segmentations outcomes without area relapse.

Li et al. [31] suggested the dynamic scale extension network (PSENet) for segmentations-based texture revelation. To deal with the eagerly nearby texture occasions, a dynamic scale development calculation is introduced. Energized by the possibility of expansiveness early-searching, the development begins along the pixels of different bits with iteratively merges the neighboring texture pixels until the biggest parts are researched.

Yang et al. [33] suggested an IncepText engineering based on example mindful segmentations, that might manage scene textures with enormous difference of scale, perspective proportion, as well as direction. ResNet-50 module is early utilized for highlight evocation, with Inception-Text module is attached after include blend.

Moreover, deformable PSROI pooling [32] is implemented to distinguish multi-arranged texture. Such social affair of procedures is reasonable for dealing with multi-oriented texture in genuine scene picture. At the point when texture occurrences in image are exceptional.

In this specific work, we center around further developing OCR exactness by pre-handling images of information reports. In contrast to the strategy suggested in [33], we suggest to utilize a somewhat further streamlined non-parametric methodology by working on the nature of the information archive, in this way utilizing a tried arrangement of preprocessing procedures. This enables us to further develop OCR precision autonomously of any prepared word reference, language explicit data, or language scheme.

3 Methodology
Removing the text from the report picture and handling it further for making it skilled to be perceived is the undertaking of advanced picture handling. Solid character division, removable of different commotions,

skeletonizing the character and normalizing everything goes under the course of advanced picture handling. Preceding character recognition, it is important to disengage the singular characters from text picture. Computerized picture handling alludes handling of two-layered picture. Advanced Images are the photos that have been changed over into a PC discernible twofold organization comprising of sensible 0's and 1's [4]. Advanced picture handling has an expansive range of use in different fields, for example, the remote detecting through satellites and other shuttle's, picture transmission and capacity for a business application, clinical handling, radar, sonar, and advanced mechanics and robotized investigation of modern parts [3]. The computerized picture handling activity might be comprehensively assembled into five crucial classes, which are as per the following:

3.1 Image Portrayal and Demonstrating

Picture portrayal and models manage the most common way of tracking down legitimate approaches to numerically address the picture. Fundamentally, there are two methods for addressing an picture. The outside portrayal manages the limit of the picture and utilized when the essential spotlight is on the shape characteristics. The inward portrayal manages pixels inside the locale and utilized when the essential center is provincial properties, for example, shading and texture. In any case, once in a while utilizing both portrayal techniques is fundamental. Picture improvement alludes to the method involved with rolling out an improvement to the highlights of the picture like edges, limits, or differentiation to make a realistic presentation more valuable for show and examination [3]. The improvement interaction doesn't build the intrinsic data content of information yet it builds the powerful scope of the picked highlights so they can be recognized without any problem. Picture upgrade incorporates dark level and differentiation control, sound decrease, edge crispening and honing, separating, interjection also, amplification, pseudocoloring, etc.

3.2 Noise Expulsion/Decrease

Commotion is an arbitrary variety of picture Intensity and noticeable as grains in the picture. It might cause to emerge in the picture as impacts of essential physical science like photon nature of light or nuclear power of hotness inside the picture sensors. It might deliver at the hour of catching or then again picture transmission [5]. In a picture, clamor alludes to the pixels showing unique force values rather than genuine pixel values. It is presented in the picture at the hour of picture procurement or transmission. In a parallel picture, we allude commotion as a dark pixel that ought to be white or the other way around. The commotion ought to be taken out in beginning phases, on the off chance that not it is spread to the further strides of handling, then it might influence the activity of the framework. There are fundamentally two various kinds of clamor. Salt commotion, when the pixel should be dark however it is addressed by a white pixel. Pepper commotion, when the pixel should be white yet is addressed by the dark pixel. We can utilize different kinds of channels to eliminate commotion from the pictures, for example, mean channel, middle channel, Laplacian filter, mode channel and so on.
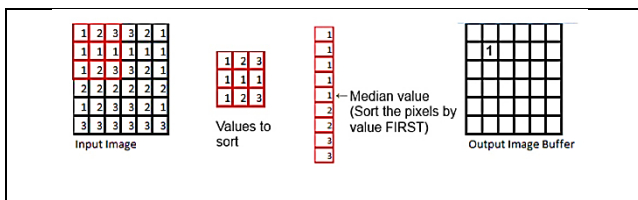
A. Median Filter

The middle channel is a basic and exceptionally powerful commotion evacuation channel. It is a non-direct channel furthermore, has a place with the class of edge-saving smoothing channels. While performing middle separating not entirely set in stone by the middle worth of all pixels in a chose neighborhood (veil, layout, and window) [6]. In the middle channel, an odd measured window like 3 X 3, 5 X 5, 7 X 7 and so forth is taken and the pixel values inside the window are organized in rising or sliding request. The middle worth is taken from the arrangement of values and supplant to the functioning pixel as shown in Figure 1.

B. Mean Filter

Mean filter, a least complex low pass straight filter, is straightforward and simple to use for

smoothing a picture. In this filtering, the amount of the relative multitude of adjoining pixels in the window is separated by the absolute number of pixels. The worth we get after the computation then replaces the worth of the pixel being handled. One of the fundamental issues of this filtering is the blurring of the edge. In the event that a sharp edge is required, this filter isn't the most ideal decision.


Fig. 1: Demonstration of the median filter [8].

3.3 Image Restoration

Whenever a picture is gained utilizing optical, electro-optical or electronic means, there is the chance of debasement because of detecting climate. The various sorts of debasement that could happen are as mathematical corruption, enlightenment and shading flaw, and obscure. Blurring is a type of transmission capacity decrease of an ideal picture inferable from the defective picture arrangement process. It very well may be brought about by relative movement between the camera and the first scene, or by an optical framework that is out of center [4]. The field of picture rebuilding (now and again alluded to as picture deblurring or picture deconvolution) is worried about the remaking or assessment of the uncorrupted picture from an obscured and uproarious one [4]. The general viability of the rebuilding filters relies upon information on debasement process and the filter plan. Picture reclamation isn't equivalent to picture improvement. Picture improvement measures are challenging to address numerically though the rebuilding issues might being evaluated unequivocally. Picture reclamation takes the recorded picture and delivers the best gauges of the first picture. Picture examination is the method involved with recovering significant data from pictures, remove measurable information. The element which is extricated

during picture investigation can be connected with tracking down shapes, recognizing edges, eliminating commotion, counting objects, estimating area and picture properties of an article. Thus, picture investigation incorporates processes, like division, skeletonization, incline rectification, and size standardization.

3.4 Skeletonization

Skeletonization is the aftereffect of the diminishing system, which strips the shade of the text until it arrives at the most average one-pixel width [8]. During this cycle, the pixels of an article are eliminated without influencing the overall state of the example. The result object after skeletonization ought to be one pixel thick, through the center of the item and the geography of the article ought to be saved. Figure 2 illustrates an example of the Skeletonization process.
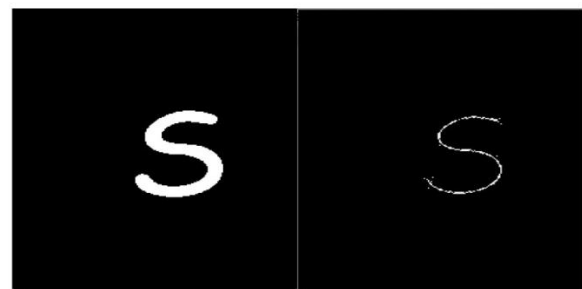

Fig. 2: Demonstration of the Skeletonization [10].

3.5 Plan & Construction

Configuration get rid of manages the method involved against laying the progression of data for a framework. The plan chiefly comprises of how the data sources are given to the framework and how the results are shown. It additionally comprises of the plan of the cycle for example how the handling modules happen, how the information travels through the module. The progression of information can be displayed with the assistance of the data flow diagram (DFD). A DFD outlines the progression of data utilizing various images like square shapes, circles, bolts, short text

marks and so on. It tends to be straightforward as well as staggered that digs dynamically more profound into how the information is taken care of. A Global Structure of the OCR scheme [12] is shown in Figure 3.
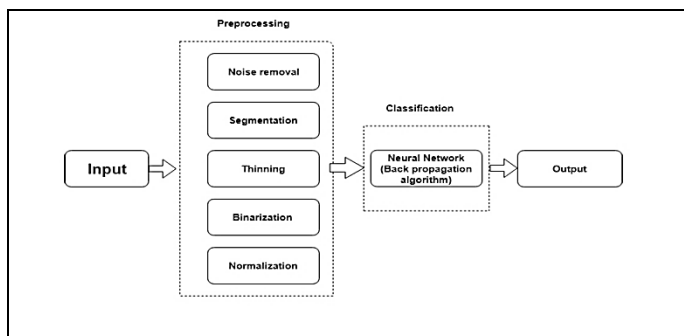

Fig. 3: Global Structure of the OCR scheme [12].

Figure .3, illustrates the overall plan for optical character recognition, OCR. It addresses the means associated with the development of the framework and the progression of information. The framework is separated into two unique principle stages. The early is preprocessing, it incorporates the calculations utilized for computerized picture handling. Once, the info is provided by the client to the framework. The preprocessing stage removes the characteristics of the character to be utilized in the following stage called, arrangement. For the arrangement, the brain network has been utilized. There are numerous calculations that can be utilized for the order. For this framework, back-proliferation calculation has been utilized. In this stage, the framework is prepared before it can perceive the characters. When every part of the characters have been prepared, the framework might perceive the characters with a result which might be created for the client. As currently expressed the cycle configuration is finished against the assistance of DFD's. Figure 4 presents the OCR context scheme [13].
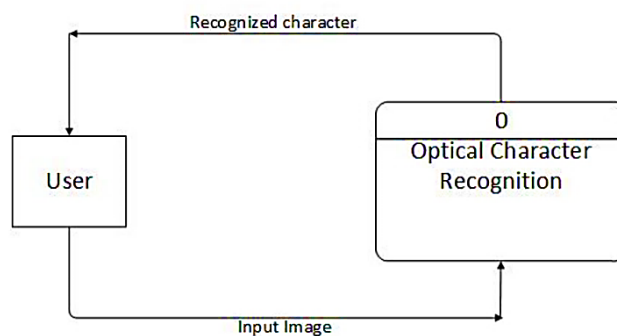
Fig. 4: OCR context scheme [13].

Figure 4, shows the context graph, which characterizes the outer substances that will associate with the framework. For the OCR, the outside element is the client. The client gives the contribution to the framework and sits tight for the result. This context graph can be additionally separated into more modest charts to show the cycles engaged with subtleties. Also, Figure 5 shows a Block diagram of Level 1 DFD for OCR. [14].
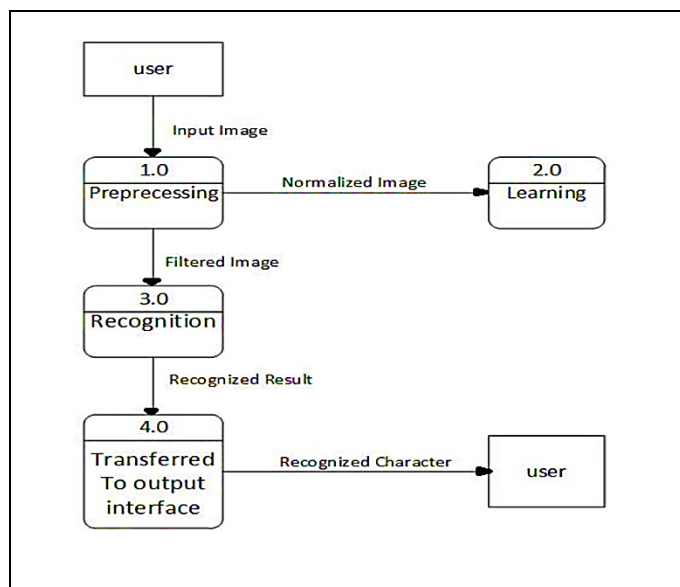
Fig. 5: Block diagram of Level 1 DFD for OCR. [14].

Figure 5 shows the Level 0 information stream graph, that is the breakdown of the context flowchart. OCR comprise of four primary cycles every numbered by its event with labeled by its capacity. The interaction amounts are provided as 1.0, 2.0 and so on that implies the cycle is the essential course of 0 level DFD or the second course of the 0 level

DFD. The info sequence provided by the client first enters the interaction named preprocessing. The preprocessing comprise of various advanced picture handling instruments which will change over the client picture to our necessary measures for the further handling. During this interaction, the picture goes through filtering of commotion, division, diminishing, binarization and standardization. The standardized picture is either moved to the learning system or the recognition interaction as per the necessity. The learning system creates the registered loads as result which is put away in loads data set. Essentially, the recognition cycle requires the loads from the loads data set as info. The result from the recognition cycle is the perceived character that is moved to the result connection point and given to the client.

4 Implementation Results

The OCR algorithm discussed in chapter three has been successfully implemented and examined utilizing MatLab2020 simulation program. The results have been evaluated and presented in form of the overall text recognition accuracy concerning the original with the processed character text images, as well as in form of the processed image brightness and clarity issues.

4.1 Accuracy Test results:

In this progression the text produced by the pre-owned page-perusing framework is coordinated with the right text to decide the base number of alter tasks (character inclusions, cancellations, and replacements) expected to address the produced text. This provides us with the quantity of mistakes. In the event that there are n characters in the right text, the character precision is characterized by (3)

$$\text{Accuracy} \% = \frac{n - n_{error}}{n} \times 100\%$$
$$(3)$$

Table. 3 presents the quantity of errors for our examination and the relating character Accuracy.

Table 3.1: : Character accuracy using Original and processed document image.

Docume nt ID	Original image		Processed image	
	Erro rs	Accura cy	Erro rs	Accura cy
1	10	99.34%	0	100%
2	60	99.01%	30	99.5%
3	312	77.17%	219	83.97%
4	89	97.64%	15	99.6%

As the table shows, we have tried our methodology on articles from, Magazine, Proceedings and diary. As we could find in table 3.1, the quantity of segments in an archive picture or report type doesn't immensely affect the end-product, however the nature of the information picture does.

As we might find in this yield text, the page per-user had the option to perceive all characters in the handled report picture aside along single character, though it couldn't perceive five characters in the first record picture.

4.2 Brightness & Clearance Results

In the other part of the obtained results, the processes character text image has been examined in form of its brightness and text content clearance as compared with the original one. The gained outcomes have been gathered and demonstrated in Figure 3, which illustrates the simulation of the project OCR text recognition results.

**(a) Original image   (b) Processed image**
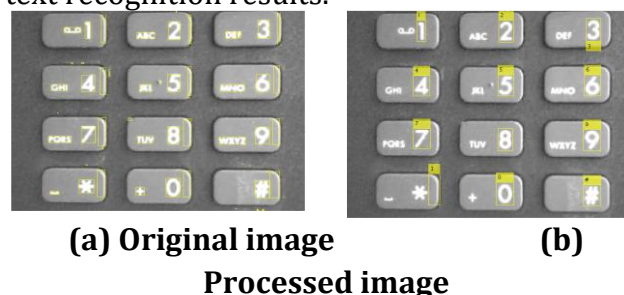
Figure 3: Simulation of the project OCR text recognition results.

We might finish up along the table that our technique was exceptionally effective upon further developing the OCR exactness in all cases handled in this study. The upgrade was along 2% to 10 %. We provide the subsequent

result text for the handled record utilized to illustrates the methodology presented in here project.

Also, the statistical histograms of both the original and the examined text images have been computed and presented in Figure 4, in which the Histogram of V channel has been applied, (a) for the original image converted to HSV color space, (b) for the enhanced image converted to HSV color space.


(a)   **(b)**
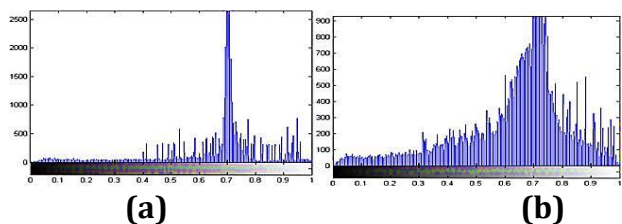
Figure. 4: Histogram of V channel, (a) for the original image converted to HSV color space, (b) for the enhanced image converted to HSV color space.

By referring to Figure 4.2, it is obvious that by applying the HSV color space technique, the concentrations of the histogram power are fixed around the center of the text brightness for the processed image whereas for the original one the power concentrations are poor. The characteristics of the processed image histogram have insuring the extraction of the text image details from the surrounding image background. By such technique, the brightness and clarity of the examined image have been satisfied and the text characters have been successfully extracted. Furthermore, the proposed structure along this project have introduce a technique for analyzing two types of information entrance associated with visual video data passing through cloud computing medium. Every kind of data have been analyzed as well tested through applying two separate and efficient algorithms, the Yolov3 and ANN respectively. The outcomes show an excellent characteristics concerning the efficiency , MSE, as well as the computational period. The results of the first sub program utilizing Yolov3 algorithm have been illustrated in Figure 5.
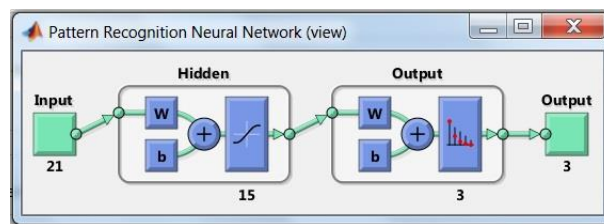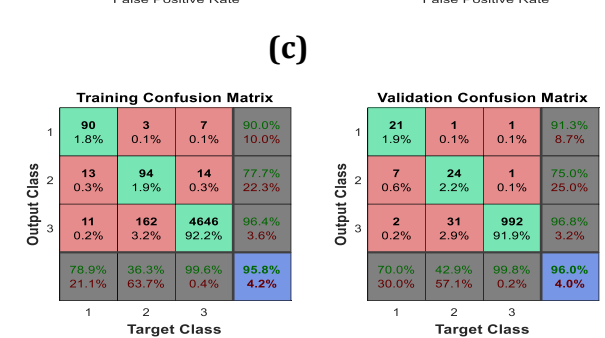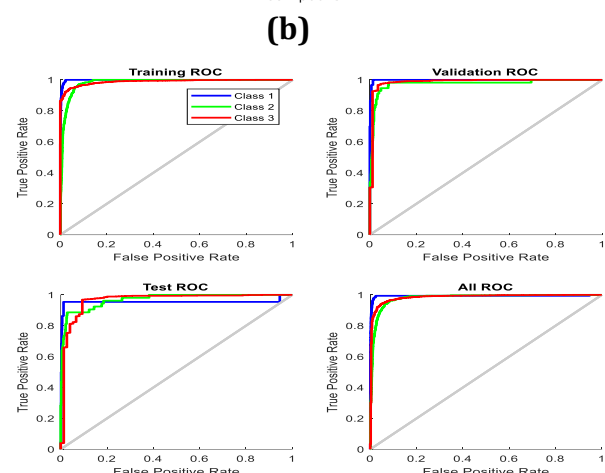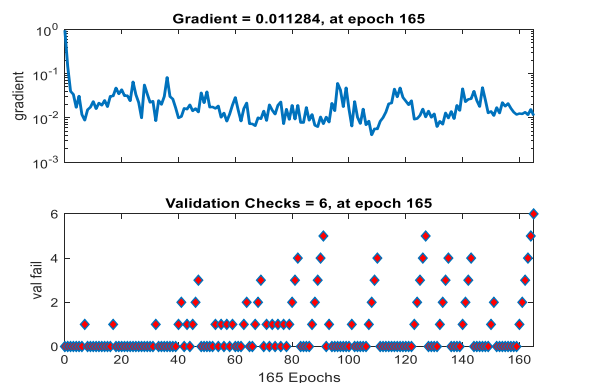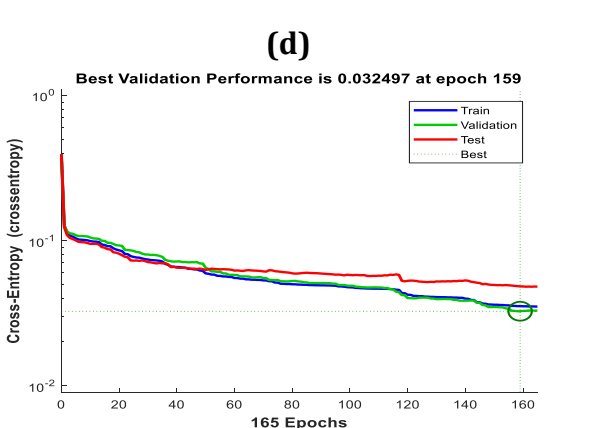

(a)
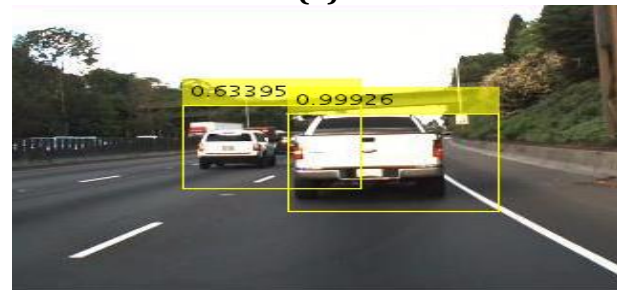
(b)
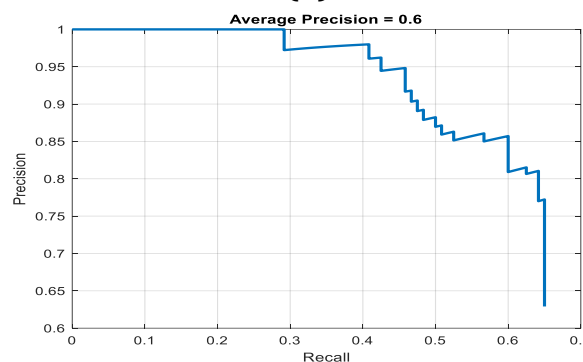
(c)

Fig. 5: Results of simulation, (a) Input video frame, (b) Object detected video frame, (c) Precision curve. Also, the results of the second sub program utilizing ANN algorithm have been illustrated in Figure 4.4.

(a)

(b)

(c)

(d)

(e)

Fig. 6: Results of sub program2, (a) ANN block diagram, (b) training state curve, (c) ROC curve, (d) Confusion matrix diagrams, € Performance curve.

5 Conclusions

In this article, we examine an extraordinary nonparametric and solo strategy to compensate for lamentable record picture goes planning to furthermore empower OCR exactness preferably. Our philosophy relies on an astoundingly functional heap of report picture overhauling systems to recuperate twisting of the entire record picture. In one or the other state, we propose a local wonder with separation change method to oversee lighting blends and the unpredictable spread of picture illumination fittingly. Second, we use a redesigned grey-scale change evaluation to change our report picture to grey-scale level. Third, we sharpen the obliging information in the ensuing grey-scale picture using Un-sharp Masking strategy. Finally, an ideal as a rule approach is used to set up the last record picture to OCR affirmation. The proposed approach could normally moreover develop text disclosure rate and optical individual affirmation precision. To show the supportiveness of our strategy, an escalated trial and error upon a standard dataset is presented. Finally, a couple of future models in video object affirmation are acknowledged to address the demands being referred to. The recommended model from this study have present a strategy for assessing double information entrance related with visual video information and cloud computing signals. Each kind of information have been broke down analyzed using two isolated also productive strategies, the Yolov3 and ANN separately. The outcomes show an incredible exhibition concerning the effectiveness , MSE, and the computational time.

References

1. Kanan, Christopher, and Garrison W. Cottrell. (2012) 'Color-To-Grayscale: Does The Method Matter In Image Recognition?'. PLoS ONE 7.1: e29740.

Web.

2. Grundland M, Dodgson N, (2007) Decolorize: Fast, contrast enhancing, color to grayscale conversion. Pattern Recognition 40: 2891–2896.

3. Cadik M, (2008) Perceptual evaluation of color-to-grayscale image conversions. Computer Graphics Forum 27: 1745–1754.

4. Pratt W, (2007) Digital image processing Wiley-Interscience.

5. N. M. Kwok, H. Y. Shi, G. Fang, and Q. P. Ha, (2012) "Intensity-based gain adaptive unsharp masking for image contrast enhancement," Image and Signal Processing (CISP), 5th International Congress on, On page(s): 529-533.

6. F. Y. M. Lure, P. W. Jones and R. S. Gaborski, (1996) "Multiresolution unsharp masking technique for mammogram image enhancement," Proc. SPIE Med. Imag., pp.830 -839.

7. OTSU N, (1979) "A Threshold Selection Method along Gray-level Histograms," IEEE Trans. Syst. Man Cybern, 9: 62-66.

8. Rafael C. Gonzalez, Richard E. Woods, (2007) "Digital Image Processing", 2nd ed., Beijing: Publishing House of Electronics Industry,

9. Zimmerman, JB, SM Pizer, EV Staab, JR Perry, W McCartney, BC Brenton, (1988) "An Evaluation of the Effectiveness of Adaptive Histogram Equalization for Contrast Enhancement", IEEE Trans. Med. Imaging, 7(4): 304-312.

10. Suprijanto, Gianto, E. Juliastuti, Azhari, and Lusi Epsilawati, (2012) "Image Contrast Enhancement for Film-Based Dental Panoramic Radiography," in International Conference on System Engineering and Technology, Bandung, Indonesia.

11. El Harraj and N. Raissouni, (2015): "Toward Indoor and Outdoor Surveillance Using an Improved Fast Background Subtraction Algorithm" International Journal of Computer, Control, Quantum and Information Engineering, vol 9, no. 4, pp 595-600

12. Gatos, I. Pratikakis, K. Kepene, and S. Perantonis, (2005) "Text detection in indoor/outdoor scene images", Proc of CBDAR, pp. 127-132

13. Y. Shao and C. Wang, (2011) "Text Detection in Natural Images Based on Character Classification. Advances in Multimedia Information Processing", PCM 2010 Lecture Notes in Computer Science,

14. Volume 6298/2011, 736-746

15. Z. Yalniz and R. Manmatha, (2011) "A fast alignment scheme for automatic ocr evaluation of books," in ICDAR.

16. N. Ezaki, M. Bulacu, L. Schomaker, (2004) "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons", Int.Conf. on Pattern Recognition (ICPR), vol. II, pp. 683- 686.

17. J. Park, G. Lee, E. Kim, J. Lim, S. Kim, H. Yang, M. Lee, S. Hwang, (2010) "Automatic detection and recognition of Korean text in outdoor signboard images", Pattern Recognition Letters.

18. G. Piccioli, E. De Micheli, P.Parodi, and M. Campani, (1996): Robust method for road sign detection and recognition. Image and Vision Computing, 14:209-254

19. S.K. Kim, D.W. Kim and H.J. Kim, (1996). A recognitionof vehicle license plate using a generic algorithm based segmentation. Inproceedings of ICIP, pages 661-664

20. T. Gotoh, T. Toriu, S. Sasaki, and M. Yoshida, (1988). A flexible vision-based algorithm for book sorting system. IEEE Trans, PAMI, 10:393-399,

21. J. Zhou, D. Lopresti, and T. Tasdisen, (1998). Finding text in color images. In Proceedings of SPIE,

22. Document Recognition V, Pages 130-140

23. R. Lienhart and F.Stuber, (1996). Automatic text recognition in digital videos. In Proceedings of ACM Multimedia, Pages 11-20,

24. Moalla, F. Lebourgeois, H. Emptoz, A.M. Alimi, (2006): Image Analysis for Paleography Inspection. Document Analysis Systems VII, 25-37

25. M. Ben Jlaiel, S. Kanoun, A.M Alimi, R. Mullot, (2007): Three decision levels strategy for Arabic and Latin texts differentiation in printed and handwritten natures. 9th International Conference on Document Analysis and Recognition,

1103-1107

26. T. Phan, P. Shivakumara and Chew Lim Tan, (2009), "A Laplacian Method for Video Text Detection," The 10th ICDAR, pp.66-70.

27. Yi, and Y. Tian, (2011) "Text String Detection from Natural Scenes by Structure-based Partition and Grouping," IEEE Trans. on Image Processing

28. X. Chen and A. L. Yuille, (2004) "Detecting and reading text in natural scenes," Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 2, pp. II-366 – II-373

29. Y. Shao and C. Wang, (2011) "Text Detection in Natural Images Based on Character Classification. Advances in Multimedia Information Processing", PCM 2010 Lecture Notes in Computer Science,

30. Volume 6298/2011, 736-746

31. J. Edwards, Y.W. Teh, D. Forsyth, R. Bock, M. Maire, G. Vesom, (2004): Making Latin Manuscripts Searchable using gHMM's. Advances in Neural Information Processing Systems, 17, 385-392.

32. H. Cecotti, A. Belaîd, (2005): Hybrid OCR combination approach complemented by a specialized ICR applied on ancient documents. 8th International Conference on Document Analysis and Recognition, 7, 1045-1049

33. A.M. Namboodiri, P.J. Narayanan, C.V. Jawahar, (2007): On Using Classical Poetry Structure for Indian Language Post Processing. 9th International Conference on Document Analysis and Recognition, IEEE Computer Society- Volume 02, 1238-1242

34. Ringlstetter, K. Schulz, S. Mihov, K. Louka, (2005): The same is not the same – post correction of alphabet confusion errors in mixed-alphabet OCR recognition. 8th International Conference on Document Analysis and Recognition, 1, 406-410

35. L. Zhuang, X. Zhu, (2005): An OCR Post-processing Approach Based on Multi-knowledge. 9th International Conference on Knowledge-Based Intelligent Information and Engineering Systems, 346-352

36. S. P. Chowdhury, S. Mandal, A. K. Das, and B. Chanda, (2007) Segmentation of Text and Graphics from Document Images, Proc. of ICDAR, 619-623.

37. We-Te Chu and Fan liu, (2013) Mathematical formula detection from heterogeneous document Images, Proc. of CTAAI, 140-146.

38. Kale, P., Phade, G.M. ; Gandhe, S.T. ; Dhulekar, P.A, (2015) "Enhancement of old images and documents by digital image processing techniques" International Conference on Communication,

39. Information & Computing Technology (ICCICT) pages: 1-5

40. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, (2002) Color plane interpolation using alternating projections, IEEE Trans. Image Process., vol. 11, no. 9, pp. 997-1013.

41. X. Lin (2007), On modeling interchannel dependency for color image denoising, International Journal of Imaging Systems and Technology, vol. 17, no. 3, pp. 163-173.

42. KD.Gennetten, (1993).RGB to CMYK conversion using 3D barycentric interpolation[J]. Proc. SPIE 1909, pp. 116-126.

43. R Lukac, K N Plataniotios, (2007). Color image processing : methods and applications [M]. New York : Taylor & Francis Group, LLC.

44. Hsueh-Yen Yang, Yu-Cheng Fan ; Hen-Wai Tsao, (2008) Algorithm and architecture design of adaptive brightness correction for video application. IEEE Transactions on Consumer Electronics. 54: 1858 – 1866.

45. https://commons.wikimedia.org/wiki/File:German_text_example_-_Em_sentence_spaced_-_1907.jpg

46. http://www.mediateam.oulu.fi/downloads/MTDB/

47. K. Kukich, (1992) "Technique for automatically correcting words in text," ACM Computing Surveys, vol. 24, no. 4, pp. 377–439.

48. X. Tong and D. A. Evans, (1996) "A statistical approach to automatic OCR error correction in context," in Fourth

Workshop on Very Large Corpora (WVLC-96), pp. 88–100.

49. Z. Yalniz and R. Manmatha, (2011) "A fast alignment scheme for automatic ocr evaluation of books," in ICDAR.