International Journal of Advanced Research in Engineering Technology & Science
ISSN: 2349-2819
Impact Factor: 7.10
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Email: editor@ijarets.org Volume-10, Issue-11 November – 2023 www.ijarets.org

STUDY ON DIGITAL IMAGE PROCESSING AND OCR USED FOR HANDWRITTEN IMAGES

Shepal Yogesh Ramdas

Research Scholar, NIILM University, Kaithal, Haryana

Dr. lokendra

Department of Computer Science And Engineering

NIILM University, Kaithal, Haryana

ABSTRACT

Among the many essential uses of digital information processing (DIP), some examples include the domains of medicine, the production of pictures and videos, the preservation and archive of documents, photography, remote sensing, and security monitoring. The aforementioned sources, in addition to additional sources, produce enormous quantities of digital image data on a regular basis. This quantity is well beyond the amount that could ever be humanly examined. The activity of utilising the capabilities of a digital computer in order to improve the quality of digital photographs is referred to as "digital image processing," and the term "digital image processing" . Utilising a computer to make changes to the features of a digital image is what is referred to as the process of transformation. It offers a method that enhances the visibility of an image by making it more detailed, sharper, and clearer than it was before. Images are more likely to be straightforward and well-organized when they are created by machines. In order to obtain the requisite input image, images, papers, and real-time devices such as tablets, tabloids, digitizers, and so on are scanned with the intention of obtaining the image. In its most elementary form, optical character recognition (OCR) methods include evaluating the text on a page and turning the characters or numerical image into code that may be used for the processing of data.

Keywords: *Digital , Image , Processing , OCR*

INTRODUCTION

The aforementioned sources, in addition to additional sources, produce enormous quantities of digital image data on a regular basis. This quantity is well beyond the amount that could ever be humanly examined. The activity of utilizing the capabilities of a digital computer in order to improve the quality of digital photographs is referred to as "digital image processing," and the term "digital image processing" . Utilizing a computer to make changes to the features of a digital image is what is referred to as the process of transformation.

Fundamentals of Digital Image Processing

Image- The actual item that exists is represented by a picture. It is possible to represent it by utilising two-dimensional functions X and Y, where the values of the function give the brightness of the image at any given place. This is a possibility. The ranges of X and Y will, of course, be determined by the intensity of the image; nonetheless, they are able to take any real value that falls between their minimum and maximum values. It is possible for the brightness values of an image to be any real number between 0 and 1, with 0 representing the darkest area and 1 representing the brightest area.

Digital image- The term "discrete" refers to an image that exists in digital form and functions independently of the values of x and y as well as f(x, y). They only accept integer values in the majority of situations, with the range of digital image values covering from 1 to 256 for each, and the brightness values similarly ranging from 0 (which represents black) to 255, which represents white. It is conceivable for it to be made up of a very large number of distinct dots, each of which is associated with a certain level of brightness.

Pixels- image elements are the individual dots that make up an image. A picture is composed of a number of these dots. A pixel's neighbourhood is comprised of the pixels that are directly adjacent to it and are referred to as its immediate neighbours. In the same way that the shape of a matrix may be used to describe it, the shape of an area can also be used to describe it:

Image processing- The process of modifying the behaviour of a picture is referred to as image processing. Image processing is a specialised technique. It serves two purposes: the first is to improve the graphical information of the image so that it can be interpreted by humans, and the second is to make it more suitable for the perception of machines that are able to operate on their own.

Optical Character Recognition (OCR)

Because it is necessary to convert scanned images into formats that are identifiable by computers, such as text documents, optical character recognition (OCR) technology has grown increasingly widespread. This is owing to the fact that the number of applications for this technology has increased by a significant amount. Optical character recognition (OCR), which is one of the most fascinating and challenging subfields in pattern recognition, has a wide variety of applications in the real world.

In order to discriminate between printed or handwritten images, such as numbers and characters that contain digital copies of physical documents, the concept of optical character recognition (OCR) is implemented with technology; this allows for the recognition of these types of images. Among the types of papers that fall into this category are a scanned paper document, a photograph of a document, and a scene photograph (for instance, the text that is displayed on signs). In order to obtain the requisite input image, images, papers, and real-time devices such as tablets, tabloids, digitizers, and so on are scanned with the intention of obtaining the image. In its most elementary form, optical character recognition (OCR) methods include evaluating the text on a page and turning the characters or numerical image into code that may be used for the processing of data. This code can then be used to perform operations on the data. Other names for optical character recognition (OCR) include text recognition, character recognition, and numerical recognition of characters.

Preprocessing

The process of executing operations on images at the most fundamental level of abstraction is referred to as pre-processing, and it is one of the most frequent ways. Through the use of this technique, both the input and output of the data are images of intensity. It is also a step towards decreasing data that is not relevant and keeping image information that is valuable. All of these things are important. An intensity picture is often represented as a matrix of image function values (brightnesses), and these kinds of images are of the same kind as the original data that was obtained by the various media that were used to capture the data. In other words, intensity images are the same information as the original data. During the pre-processing stage, the primary objective is to enhance the image data by minimising undesirable distortions or boosting specific picture characteristics that are necessary for the upcoming stage of the process. The fundamental goal of pre-processing is to improve the image data, despite the fact that it is classed among multiple methods here because comparable approaches are used. This is because pre-processing is classified among numerous ways.

Image Enhancement

picture enhancement is the process of highlighting or enhancing certain visual qualities, such as boundaries or contrast, in order to make a picture more easily discernible and susceptible to analysis. This is done in order to make the image more interesting to study. This approach does not result in an improvement to the information content that is already present in the data. The article discusses a variety of approaches, including the modification of grey levels and contrast, the removal of noise, the crisping and sharpening of edges, the morphological filtering and deblurring filtering, the interpolation and magnification, the pseudo colouring, and other techniques.

Segmentation

When it comes to offline handwritten optical character recognition, segmentation is always a vital component that is included in the process. Character separation refers to a technique that requires dividing character graphics into discrete text lines. From these lines, characters are separated by employing connected component labelling. This technique is referred to as "character separation." For the aim of this investigation, a technique that is commonly referred to as the bounded box method was utilised in order to split lines, phrases, and characters that are contained within documents. The technique is built on the pixel histogram that was obtained, which acts as the foundation. The manner in which individual symbols are accurately segmented is the determining factor in determining the precision of the use of the character identification method. The breakdown of a picture of a sequence of characters into subimages of individual symbols is another application of this approach. This is accomplished by segmenting lines and phrases into their component parts.

The large character set of the Devanagari script, in addition to the enormous number of compound characters that are compatible with the Devanagari character, is one of the defining characteristics of this scripting system. Through the implementation of the segmentation process, it is possible to bring the error rate down to an acceptable level. The existence of characters that are skewed coupled with those that overlap makes the process of segmentation more complex. This in turn makes the process more difficult to complete. The following is a list of the steps that it is able to complete successfully.

At first, the text is divided into separate lines for easier reading. Next, the lines are broken down into words, and then, lastly, the words are broken down into characters. Characters are the final step in the process. The explanation can also be found in figure 12, if you so choose.

OBJECTIVES OF THE STUDY

1. To study on Fundamentals of Digital Image Processing
2. To study on Optical Character Recognition (OCR)

RESEARCH METHOD

The research method is based on One of the most important subtasks in a wide variety of natural language processing (NLP) applications is the perception of various word meanings. Despite the fact that the topic is theoretically straightforward and can be easily written as a conventional classification problem, achieving a high level of performance on this assignment is surprisingly difficult to do. Ambiguity in full phrases or clauses is referred to as structured syntactic ambiguity, but ambiguity in individual words is referred to as lexical semantic ambiguity. Assigning the right senses to words in a particular context is the primary goal of word sense determination (WSD), which is the process of resolving lexical semantic ambiguity. WSD is also known as word sense analysis.

Learning algorithms are classified using both statistical and symbolic approaches. Symbolic methods, in contrast to statistical techniques, do not make explicit use of probability. Statistical methods are used to categorise learning algorithms. Illusion of Markov The following are some examples of symbolic methods: inductive logic programming, neural networks, genetic algorithms, clustering, support vector machines, and transformation-based error-driven learning are all examples of symbolic methods. Statistical learning techniques include models and log-linear models. Decision trees and decision lists are examples of symbolic methods.

Experimental procedure

The input nodes, which are words, are given appropriate weights in an attempt to construct semantic connections. This effort is undertaken in order to establish semantic links.

Database Creation

Two innovative fields, namely an educational system and a sports fraternity, were the components that comprised the corpus. It is possible to make a connection between the polysemous word and other words in any phrase by using the structure.

$$t_w\#\textit{is part of}\ \#t_{1,}\ t_w\#\textit{is a kind of}\ \#t_{2,} \ldots\ t_w\#\textit{contains}\ \#t_n$$
,

Performance Evaluation

For the purpose of determining the results of the experiment, we carried it out a predetermined number of times using learning sets of varying sizes. In proportion to the size of the learning set, the usual rates of learning and test mistakes are shown in Figure 1. Although the test error decreases as the size of the learning set increases, the learning error also increases with the size of the learning set increases.

A focus is placed on the fact that the experiment was successful in accomplishing its objectives with regard to the precision, recall, and correctness of the words. Table 1 displays the values of the metrics, which are arranged according to the frequency of their occurrence. A bar chart is shown in Figure 1 that illustrates how the indices vary with frequency for a specific pair of words belonging to a certain meaning. A comprehensive description of the supervised techniques and the WSI approach can be seen in Table 3.2. This table also provides ratings for recall, precision, and accuracy. The identical data that was shown is shown in Figure 3.9.

Table 1 The computation of precision, accuracy, and recall

Target/Polysemous word appearing in the input	Frequency	Accuracy % TP+FN/(TP+TN+FP+FN)	Precision % TP/(TP+FP)	Recall % TP/(TP+FN)
Book	19	75	74	69
Strain	6	64	67	65
Play	28	94	82	83
Board	21	77	76	67
Pen	2	78	76	75
Ring	5	80	71	69
Net	18	75	78	74
Note	36	97	81	78
Top	7	86	84	86
Deal	13	81	78	81

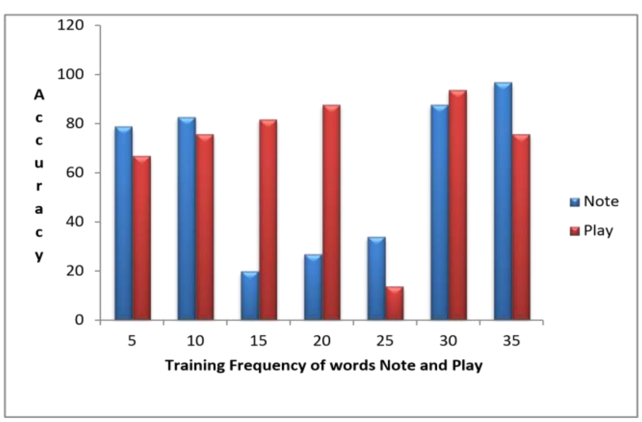

Fig 1 Plot of accuracy for words organised by frequency

DATA ANALYSIS

A comprehensive assessment of the suggested scheme was conducted in comparison to current supervised techniques, and the findings are shown in Table 2. The evaluation included accuracy, precision, and recall analysis. When contrasted with alternative methods, the data shown in Figure 2 demonstrate much higher recall rates and accuracy.

Table 1 The computation of precision, accuracy, and recall

Target/Polysemous word appearing in the input	Frequency	Accuracy % TP+FN/(TP+ TN+FP+FN)	Precision % TP/(TP+FP)	Recall % TP/(TP+ FN)
Book	33	95	86	84
Strain	21	81	79	80
Play	29	95	87	83
Board	14	78	75	71
Pen	24	85	83	80
Ring	4	63	66	67
Net	9	66	64	68
Note	34	97	84	84
Top	8	67	70	71
Deal	17	82	74	76

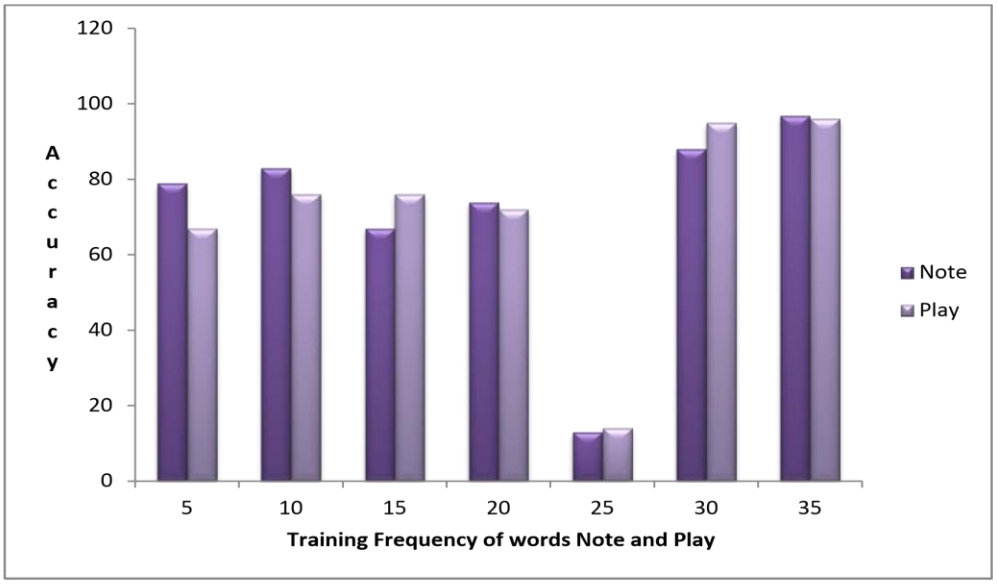

Fig. 2 Plot of accuracy for words organised by frequency

Table 2 Results comparison between supervised techniques and WSI's planned work

Total text = 135 sentences; 4063 relations;			
	Accuracy % TP+FN/(TP+TN+FP+FN)	Precision % TP/(TP+FP)	Recall % TP/(TP+FN)
Lesk Algo'	69.3	72.6	74
Lin Algo'	69.5	71.3	76
Resnik Algo	69	69	72
Heuristic – One sense per collocation	71	72.2	72.4
Heuristic – One sense per collocation	71.6	72	71.4
WSI	80.9	76.8	76.4

The results are confirmed by utilising the Lesk method, which states that "two different words are likely to have similar meanings if they occur in identical local contexts." This algorithm is designed to verify the results. This suggests that all words make use of the same knowledge sources, and rather than developing individual classifiers for each word, the present word is disambiguated based on how it has been used in the past in conjunction with other words. The use of the test kit words results in an asymmetrical accuracy path on the graph, in contrast to the WSI technique, which demonstrates a consistent increase in performance over the course of time.

Application of WSD In Event Trigger Detection

After collecting synonyms of keywords via the use of wordnet, we proceeded to apply the WSD method in order to unambiguously identify biological trigger words. In the course of this experiment, we make use of the optimisation technique that is based on NSGA-II for the purpose of selecting features from each subset of the overall dataset. Each of the subsequent steps is the whole of the method.

1. We compile two word dictionaries. Taken from the training dataset, one dictionary contains ambiguous terms and the other contains non-ambiguous trigger words.

2. Make m blank datasets that can be found using a trigger word; the size of the ambiguous and non-ambiguous dictionaries is equal to m.

3. The following steps are performed for every sample word in the dataset.

Table 3 Eliminating trigger words using a technique other than WSD

Phosphorylation	93.06	89.33	91.16
Positive_regulation	63.30	77.88	69.84
Gene_expression	79.85	88.61	84.00
Regulation	50.72	71.14	59.22
Binding	67.70	87.44	76.32
Localization	74.42	76.19	75.29
Entity	41.86	66.06	51.25
Negative_regulation	55.65	76.95	64.59
Transcription	49.57	90.63	64.09
Protein catabolism	100.00	91.30	95.45
Overall	63.99	80.33	71.24

a) The term is added to the relevant dataset if it is found in one dictionary.

b) Alternatively, if a wordnet synonym is present in one dictionary, it is added to the dataset where the match occurs.

c) Alternatively, submit to the default dataset.

4. For every dataset, we use NSGA-II for feature selection and build a classifier using Support Vector Machine.

We employ WSD on the BioNLP-2011 dataset for the purpose of trigger detection, and we put our WSD model through its paces on the development data. The results of the trials are shown in tables 4.4 respectively. Tables 4 respectively, provide the results of the trigger word extraction process when the WSD technique was not used and when it was utilised. The total findings are shown in Table-4, which reveals that the method that did not use WSD provided poorer results (a 5% increase in f-score value) than the method that used WSD for the detection of trigger words.

Table 4 Extracting trigger words using the WSD technique

Phosphorylation	90.27	92.85	91.54
Positive_regulation	70.28	83.56	76.35
Gene_expression	81.93	90.16	85.85
Regulation	60.28	76.36	67.37
Binding	76.65	88.73	82.25
Localization	76.74	82.50	79.51
Entity	53.48	66.66	59.35
Negative_regulation	62.50	82.03	70.94
Transcription	59.82	93.33	72.91
Protein catabolism	100.00	87.50	93.33
Overall	70.71	83.70	76.66

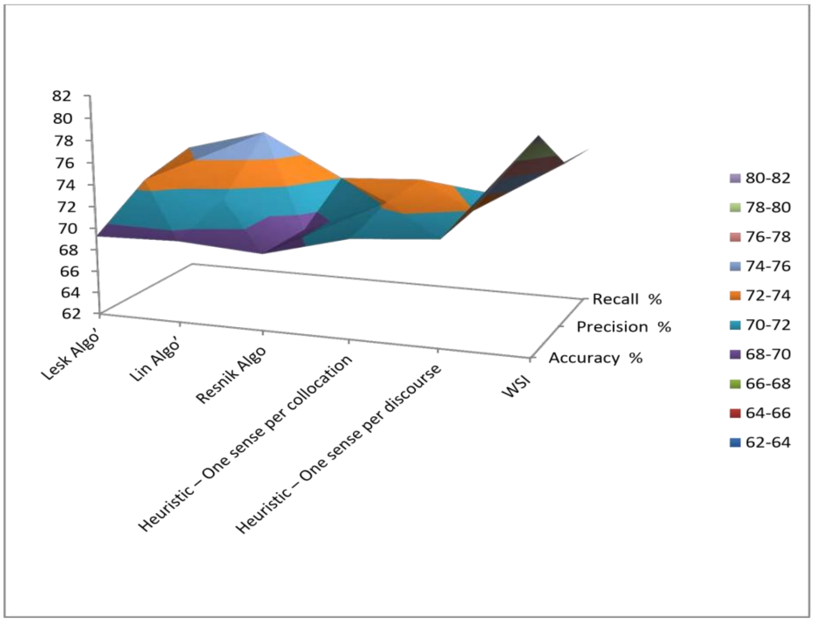

Fig. 3 Suggested Tasks vs. Supervised Approaches

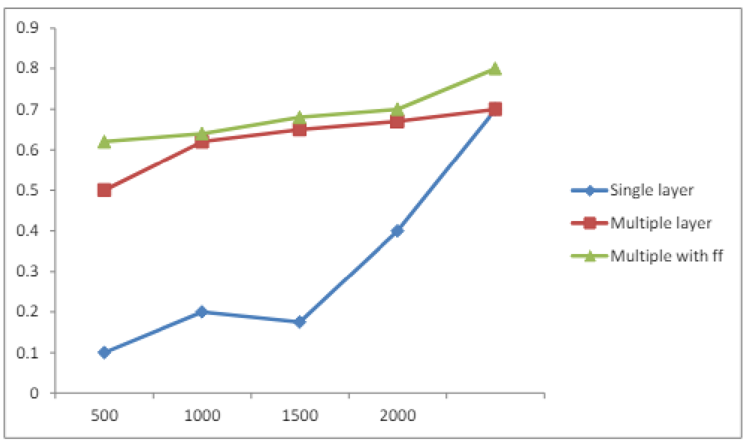

Fig. 4 Acquiring Knowledge

CONCLUSION

The extraction of biological events and the interpretation of word meanings based on biomedical data. In this part, a summary of the results of the study is presented, some conclusions are drawn from the thesis, and a number of potential directions for additional research are suggested. In order to get the most out of our experiment, we adjusted the parameters of the algorithms that we used and compared the outcomes with and without optimisation to determine how well the system worked both with and without optimisation. Additionally, we used a Genetic Algorithm in order to fine-tune the class-weight system. The results of our experiments indicate that the F-score value for event extraction on the development dataset is 57.54%. Documentation exists for each and every one of these court cases.

REFERENCES

[1] Agirre, Eneko & German Rigau. 1996. "Word sense disambiguation using conceptual density", in Proceedings of the 16th International Conference on Computational Linguistics (COLING), Copenhagen, Denmark, 1996.

[2] Agirre E. and Soroa A. Using the multilingual central repository for graph based word sense disambiguation. In Proceedings of LREC '08. Marrakesh, Morocco, 2008.

[3] Agirre E. and Soroa A. Personalizing page rank for word sense disambiguation, In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL-09). Athens, Greece, 2009.

[4] David. Yarowsky, 1992. "Word sense disambiguation using statistical models of Roget's categories trained on large corpora", in Proceedings of the 14th International Conference on Computational Linguistics (COLING), Nantes, France, 454- 460, 1992.

[5] David. Yarowsky, 1994. "Decision lists for lexical ambiguity resolution: Application to accent restoration in Spanish and French", in Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics (ACL), Las Cruces, U.S.A., 88-95, 1994.

[6] David. Yarowsky, 1995. "Unsupervised word sense disambiguation rivaling supervised methods", in Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics (ACL), Cambridge, MA, 189- 196,1995.

[7] Fass D., Guo C., McDonal J., Plate T. and Slator B. Wilks Y., 1993. Providing Machine Tractablle Dictionary Tools, in Semantics and the Lexicon (Pustejovsky J. ed.), 341-401. Yarowsky, D. 1992.

[8] Foltz, P., Kintsch, W., And Landauer, T. 1998. The Measurement Of Textual Coherence With Latent Semantic Analysis. Disc. Proc. 25, 2–3, 285–307.

[9] Frawley, W. 1992. Linguistic Semantics. Lawrence Erlbaum Associates, Hillsdale, NJ.

[10] Hatzivassiloglou, V., Klavans, J., And Eskin, E. 1999. Detecting Text Similarity Over Short Passages: Exploring Linguistic Feature Combinations Via Machine Learning. In Proceedings Of The Conference On Empirical Methods In Natural Language Processing And Very Large Corpora. 203–212.

[11] Hecht-Nielsen. R (1988) Counter propagation networks, Neural Networks 1, 131 - 139.