ISSN NO: 2770-4491
Date of Publication:08-07-2022

Text Recognition from Images

Iman Kadhim Ajlan [1]*, Ahmed fadhil Bin yusof [2*], Herlina Binti Abdul Rahim[3*]
Ahmed Jameel Ismael[4*]
1 UTM University, Malaysia and College for medicine University of Wasit, Iraq
2 International student office, UTM University, Malaysia
3 Research Management Center, U TM Skudai Johor Bahru Malaysia
4 School of Electrical Engineering Faculty of Engineering, UT M Skudai Johor Bahru, Malaysia.
* Corresponding author: ajlan@graduate.utm.my, ahmedfadhil@utm.my, herlina@utm.my, ahmedismael@graduate.utm.my

Abstract
Python has developed into one of the most popular and powerful languages in coding and can be used for various purposes such as language detection. Language detection's idea is based on the identification of the characters used in words and expressions in the text. The main aim is to identify words commonly used in a particular language such as of or to in English. Python has several modules to enhance language identification. Language detection is a common feature in any Web application or Social Network, which is usually combined with Machine Translation to enhance content accessibility and user experience. This develops a platform for other features including document analysis, articles/tweets/posts, and Machine Translation. When the document is fed, the language is not known. As a result, conducting segmentation is not possible since the language in the document is unknown, which makes using word base models impossible. Consequently, the best option is the use of "Observa." One research area that aims to establish a computer system that can identify texts from images is text recognition in images. In today's world, there is a growing demand to transfer information from paper documents to the storage disks in computers for easier reuse and access of the information through the process of searching. A simple method of achieving this is through scanning the paper documents and then storing them in image form. However, it is extremely difficult to reuse the information by reading the contents and manually searching for what is needed from the images. Various challenges associated are poor image quality and the font of the document's characteristics. As a result of the challenges, the computer cannot identify the characters and read the document. The paper discusses various ways to recognize texts from images. The aim of this work is to identify texts from images to ensure that the reader can apply a certain structure of distinct processing modules for easier comprehension.

1. Introduction

Language can be described as a system of conventional written, manual, or spoken symbols that are applied by humans, as a particular social group [1]. Users of a language express themselves in its culture. The main objectives of a language are emotional release, imaginative expression, play, identity expression, and communication. In the modern world, we find multiple images online displaying some written texts on them but we cannot comprehend what is being conveyed if we cannot understand the language used. However, it is impossible for an individual to comprehend each language on earth. Therefore, the identification of an image's text is made easier by the Observa. In today's world, the demand for software that can identify characters when a document is scanned by a computer has significantly increased since there are numerous mythological and historical newspapers and books that exist only in printed format. As time goes by, the prints get damaged due to improper handling and atmospheric changes [2, 3]. As a result, in today's world, there is an increasing demand to store the information found in search prints in computers and later apply the search process to easily reuse them. A simple method of achieving this is through scanning the paper documents and then storing them in image form. However, it is extremely difficult to reuse the information by reading the contents and manually searching for what is needed from the images [4]. Various challenges associated are poor image quality and the font of the document's characteristics. As a result of the challenges, the computer cannot identify the characters and read the document. The paper discusses various ways to recognize texts from images. The aim of this work is to identify texts from images to ensure that the reader can apply a certain structure of distinct processing modules for easier comprehension [5, 6].

2. Text Recognition System

This section will provide a brief description of the overall architecture of the system of text recognition as indicated in figure 1. A system of text recognition gets input in an image form that has text information. The system can be distributed into various modules including post-processing, text recognition, and pre-processing. The modules have been discussed in detail in the subsequent section [7, 8].

ISSN NO: 2770-4491
Date of Publication:08-07-2022

2.1. Pre-Processing Module

The paper document is changed into an image form by the optical scanner. A picture is a group of multiple picture elements referred to as pixels. During this stage, the data is stored as an image and the picture can go through further analysis to retrieve important information. As a result, various operations are conducted to improve the image's quality including normalization and noise removal [9, 10].

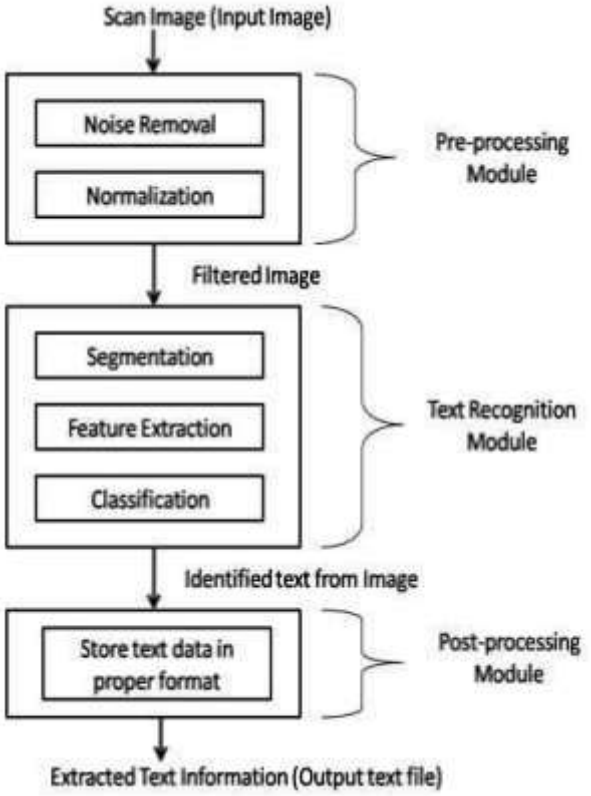

Figure 1: Architecture of text identification

a. ***Noise Removal***

One of the most essential processes is noise removal. As a result, the image's quality is significantly improved, which positively impacts the identification process for improved image recognition. Noise removal in images can be conducted in several ways including Gaussian filter, min-max filter, and mean filter.

b. ***Normalization***

Normalization is an essential aspect during the operation of identifying a text by a computer. The process ensures that the rotation, slant, and size of characters are made uniform.

2.2. Text Recognition Module

The module is essential in developing the output image's text identification of the pre-processing stage to develop output data that is in a form that can be understood by the computer [11, 12].

a. Segmentation

The most essential aspect in the module on text recognition is segmentation. The process is done to ensure that the image's characters are separated effectively.

b. Feature Extraction

Feature extraction can be described as retrieving data that is deemed most important from raw data. The most essential data is one that offers an accurate representation of characters. Distinct classes are developed to ensure that the different character features are stored. Multiple techniques are applied for feature extraction such as Chain Code (CC).

c. ***Classification***

Classification can be described as the process of recognition of every character and putting it in the accurate character class to ensure the image's texts are converted in a form that can be understood by the computer. The process utilizes the text image's extracted feature i.e. the process of feature's output is the input of classification. The stored pattern and the input feature are compared by classifiers to find out which class offers the best match. Multiple techniques are available for classification including Artificial Neural Network (ANN) [13, 11].

3. Related Work

1. Torchvision

The package of torchvision comprises common transformations of the image, model architectures, and popular datasets for computer vision. Torchvision is essential in identifying packages used in image loading [14].

ISSN NO: 2770-4491
Date of Publication:08-07-2022

2. NumPy

NumPy can be described as a numerical and open-source Python library. NumPy consists of an array that is multi-dimensional and structured on matrix data. NumPy can be used to do various mathematical operations including algebraic, statistical, and trigonometric routines [15].

3. Pillow

Pillow can be described as a Python Imaging Library (PIL) that enhances saving, manipulating, and opening images. The current Pillow version reads and recognizes numerous formats. The unintentional restriction is applied to write support to most common formats of presentations and interchange [16].

4. Torch

Torch is the torch's core package. The package is essential in providing Tesor or array that is N-dimensional and flexible that ensures support for basic routines for cloning, sharing storage, resizing, type-casting, transposing, slicing, and indexing. The object is applied by multiple packages, therefore, forming the library's core object [17].

5. SciPy

SciPy can be described as a library for an increased number of mathematical functions by employing NumPy. NumPy is used by SciPy as the structure for basic data, and also included the modules for tasks that are common and frequent in scientific programming. Such modules include signal processing, ordinary differential saving of equations, calculus, and linear algebra [18].

6. Scikit.image

Scikt.image, which was initially known as scikits.image, is an image processing library that is open for the programming language used by Python. It involves algorithms for feature detection, morphology, filtering, analysis, manipulation of color space, geometric transformations, segmentation, and more [19].

4. Experimental Result

Figure 2: Install easyocr

Figure 3: Install the package required

ISSN NO: 2770-4491
Date of Publication:08-07-2022

Figure 4: All packages to enhance text identification

Figure 5: Process image in words in English

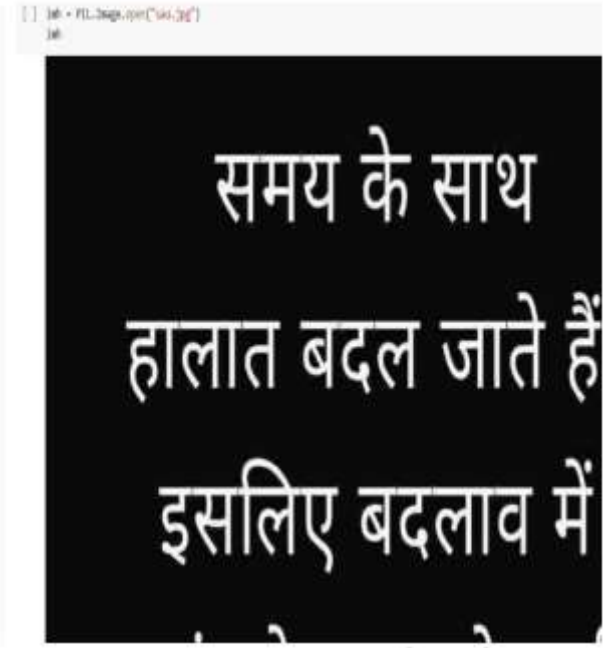

Figure 6: Process image in words in Hindu

ISSN NO: 2770-4491
Date of Publication:08-07-2022

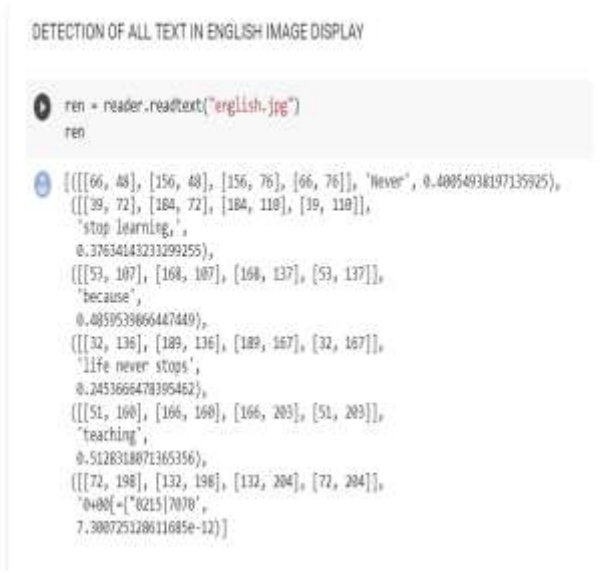

Figure 7: recognition of all English texts from the image

Figure 8: recognition of all Hindi texts from the image

ISSN NO: 2770-4491
Date of Publication:08-07-2022

Figure 9: The text's analysis and giving of borders in the red color

Figure 10: All text's analysis from the image in Hindi and giving a red border

ISSN NO: 2770-4491
Date of Publication:08-07-2022

Figure 11: The display of all texts recognized from the images

5. Application

The technology of recognition of texts can be applied through any type of industry. The main reason is that almost every industrial sector must employ the use of paper documents. Not all things can be recorded on computers. The technology of text identification is essential in revolutionizing the process of managing documents. Paper documents can be easily stored in computers through scanning and the technology on text recognition. Such action also makes it easier for the reader to find the details they want through searching. The technology is essential and effective in ensuring that scan documents can be changed from image files and change to documents that can fully be searched with the content of the texts in a form that a computer can recognize. With the technology's emergence, individuals will not be required to retype essential documents manually when putting them in electronic databases of the computer. Information stored in a computer cannot distort with time due to actions such as poor handling and change in weather. As a result, they will be always clearly readable while the reader can also search the document to find the information they require easily and effectively. The system of text recognition is essential in digging out information deemed important and automatically entering it electronically. The result is efficient and accurate processing of information in a short period. The following is an overview of various applications of the system of text identification.

1. Banking

The application of recognition of image text is different across various sectors. The technology has been widely used in banks. The processing of checks is a common activity in banks, which initially required the bank attendants to manually key in data from paper checks into the computer. However, the process today can be done without human supervision after the emergence of the technology. As a result, the speed of completing the processes is increased, which saves time for both bank employees and clients. When the check is inserted into a machine, the scanning of the paper is immediately conducted, which is the first step of the process to gain the image containing the texts. After the scanning, text identification is effectively and efficiently conducted and the funds are transferred to the client's account. The technology has also reduced cases of human error and increased accuracy. The process is fully accurate for checks that are printed and fairly accurate on those that are handwritten. In many cases, manual confirmation is required on handwritten checks. As a result, the processing of checks in many banks is enhanced and made more effective and efficient.

2. Legal

ISSN NO: 2770-4491
Date of Publication:08-07-2022

The legal sector has also experienced increased demand to store paper documents digitally. The industry involves a lot of paper works, which can be tiresome to manually type in computers for digital storage. As a result, text identification technology is highly effective and recommended in this sector. The technology is essential in increasing the speed of storing paper works digitally and enhancing accuracy. To remove the requirement to sift through paper files in multiple boxes and files and save space, it is essential to scan documents and use the text identification process to store the information on computer databases. Moreover, the process of text recognition further makes it easier for legal professionals to find the documents they require since they are searchable from the computer's database. As a result, it becomes easier to work with and locate the documents after being stored on a computer database. As a result, legal professionals have an easy and quick way to access a large library of the document using a computer and can easily find what they require by typing a few words on the computer's keyboard.

3. Healthcare

Healthcare is also a sector that heavily relies on paperwork. As a result, the sector is heavily reliant on the application of the process of text identification to ensure that paper documents are effectively and efficiently stored in a digital form. In recent years, the health sector has adopted technology in the provision of healthcare including electronic medical records. The records cannot distort with time due to various aspects such as poor handling techniques and changes in weather. Moreover, they can be easily accessed by simply searching through the computer database by typing the information included in them. Over the years, medical practitioners have always had to deal with huge amounts of forms from patients. Such forms include those on general health and insurance forms. Keeping up with such paperwork information can be tiresome and consuming. In this sector, accuracy is essential in ensuring the safety and wellbeing of patients. A small error can cause devastating effects on the health of the patients including increased cost of treatment, prolonged stay in hospitals, or even loss of lives. As a result, it is essential to ensure that relevant data is stored in digital form. The storage using the technology on text identification ensures accuracy and also makes it easier and quicker to find the information required by a health professional at any time. The application of image identification is effective and efficient in extracting information from forms and storing it in computer databases. As a result, the data of each patient is recorded promptly and accurately. Consequently, medical experts can ensure the provision of the best possible healthcare services to each client. Overall, the new technology has ensured the provision of improved healthcare in many hospitals.

4. Image text identification in other sectors

The technology on recognition of image text has also been widely used in other multiple sectors such as government agencies, finance, and education. In the education sector, the new technology has been widely used to store information from various academic articles digitally, including historical articles. Such actions are essential in eliminating the distortion of the information found in the articles due to several reasons such as weather change and poor handling methods. Various government agencies have also used image text identification to ensure that information from their paperwork is effectively and efficiently stored digitally. As a result, they can easily and quickly access the information accurately and can also use searches to eliminate manual searching which is tiresome and time-consuming. The technology has ensured the online availability of countless texts, saving money for learners in the education sector and ensuring easier and quicker sharing of knowledge. Many businesses apply invoice imaging to quickly, efficiently, and accurately process invoices to prevent the piling up of payments backlog and keep an eye on financial tracks. In various independent organizations and government agencies, image text identification has been used to simplify data analysis and collection among other uses. With the increasing advancement in technology, image text identification continues to be subjected to more applications such as improved use of recognition of handwritten paperwork

6. Conclusion

The paper aims to develop a new framework for image text identification using Machine Learning and Artificial Intelligence. Python has developed into one of the most popular and powerful languages in coding and can be used for various purposes such as language detection. Language detection's idea is based on the identification of the characters used in words and expressions in the text. The main aim is to identify words commonly used in a particular language such as of or to in English. Python has several modules to enhance language identification. Language detection is a common feature in any Web application or Social Network, which is usually combined with Machine Translation to enhance content accessibility and user experience. This develops a platform for other features including document analysis, articles/tweets/posts, and Machine Translation. When the document is fed, the language is not known. As a result, conducting segmentation is not possible since the language in the document is unknown, which makes using word base models impossible. Consequently, the best option is the use of "Observa." One research area that aims to establish a computer system that can identify texts from images is text recognition in images. In today's world, there is a growing demand to transfer information from paper documents to the storage disks in computers for easier reuse and access of the information through the process of searching. A simple method of achieving this is through scanning the paper documents and then storing them in image form. However, it is extremely difficult to reuse the information by

ISSN NO: 2770-4491
Date of Publication:08-07-2022

reading the contents and manually searching for what is needed from the images. Various challenges associated are poor image quality and the font of the document's characteristics. As a result of the challenges, the computer cannot identify the characters and read the document. Lastly, the subject requires more research from various aspects including the language used in documents and other input mediums that prove difficult to recognized compared to images. Therefore, future works should consider the document's language identifier.

References

[1] L. Weidinger, J. Mellor, M. Rauh, C. Griffin, J. Uesato, P.-S. Huang, M. Cheng, M. Glaese, B. Balle, and A. Kasirzadeh, "Ethical and social risks of harm from language models," *arXiv preprint arXiv:.04359,* 2021.

[2] J. Yang, K. Wang, J. Li, J. Jiao, and J. Xu, "A fast adaptive binarization method for complex scene images," in *2012 19th IEEE International Conference on Image Processing*, 2012, pp. 1889-1892: IEEE.

[3] H. Salim, J. S. Qateef, A. M. Alaidi,and R. M. Al_airaji, "Face Patterns Analysis and recognition System based on Quantum Neural Network QNN," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 9, 2022.

[4] H. TH., and A. M. Alaidi, "Automated Cheating Detection based on Video Surveillance in the Examination Classes," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 10, 2022.

[5] S. Dutta, N. Sankaran, K. P. Sankar, and C. Jawahar, "Robust recognition of degraded documents using character n-grams," in *2012 10th IAPR International Workshop on Document Analysis Systems*, 2012, pp. 130-134: IEEE.

[6] B. Majeed, Haider TH. Salim "Computational Thinking (CT) Among University Students," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 10, 2022.

[7] M. Rhead, R. Gurney, S. Ramalingam, and N. Cohen, "Accuracy of automatic number plate recognition (ANPR) and real world UK number plate problems," in *2012 IEEE international carnahan conference on security technology (ICCST)*, 2012, pp. 286-291: IEEE.

[8] S. H. Abbood, H. N. Abdull Hamed, M. S. Mohd Rahim, and A. H. M. Alaidi,, "DR-LL Gan: Diabetic Retinopathy Lesions Synthesis using Generative Adversarial Network," *International Journal of Online Biomedical Engineering*, vol. 18, no. 3, 2022.

[9] K.-S. Son, J.-W. Kim, and J.-H. Lim, "Convergence CCTV camera embedded with Deep Learning SW technology," *Journal of the Korea Convergence Society*, vol. 10, no. 1, pp. 103-113, 2019.

[10] B. Shi, M. Yang, X. Wang, P. Lyu, C. Yao, and X. Bai, "Aster: An attentional scene text recognizer with flexible rectification," *IEEE transactions on pattern analysis machine intelligence*, vol. 41, no. 9, pp. 2035-2048, 2018.

[11] S. Malakar, S. Halder, R. Sarkar, N. Das, S. Basu, and M. Nasipuri, "Text line extraction from handwritten document pages using spiral run length smearing algorithm," in *2012 international conference on communications, devices and intelligent systems (CODIS)*, 2012, pp. 616-619: IEEE.

[12] S. Stoliński and W. Bieniecki, "Application of OCR systems to processing and digitization of paper documents," *Information Systems in Management VIII*, vol. 102, 2011.

[13] K. Ntirogiannis, B. Gatos, and I. J. I. T. o. I. P. Pratikakis, "Performance evaluation methodology for historical document image binarization," vol. 22, no. 2, pp. 595-609, 2012.

[14] F. Albardi, H. D. Kabir, M. M. I. Bhuiyan, P. M. Kebria, A. Khosravi, and S. Nahavandi, "A comprehensive study on torchvision pre-trained models for fine-grained inter-species classification," in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2021, pp. 2767-2774: IEEE.

[15] S. J. A. P. Gholizadeh, "Top Popular Python Libraries in Research," 2022.

[16] A. J. Clark, "Pillow (pil fork) documentation," 2015.

[17] S. Marcel and Y. Rodriguez, "Torchvision the machine-vision package of torch," in *Proceedings of the 18th ACM international conference on Multimedia*, 2010, pp. 1485-1488.

[18] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, "Probabilistic programming in Python using PyMC3," *PeerJ Computer Science*, vol. 2, p. e55, 2016.

[19] H. Singh, "Basics of Python and Scikit image," in *Practical Machine Learning and Image Processing*: Springer, 2019, pp. 29-61.