© 2023 IJNRD | Volume 8, Issue 7 July 2023 | ISSN: 2456-4184 | IJNRD.ORG

IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

A Review of the Steps of Text Identification in Online Learning Communities

Vandana Devi
Computer Science & Engineering

Abstract: The brief texts seen in online learning communities are a crucial source of information for learning analysis. As a result, the research of learning analysis is significantly impacted by the brief text's quality. The expense of human identification and repair will be prohibitively high given the volume of text data in the learning community. The steps of text identification are identified and compared in this research, and various methods for text extraction from color photographs are examined. While this challenge is further broken down into text identification & localization, classification, segmentation, and text identification, two frequently used solutions for this topic are stepwise approaches & integrated approaches. This paper presents key methods applied to various stages along with their accompanying benefits, drawbacks, and applications.

IndexTerms: **Classification, Emotion, Segmentation, Text detection**

1. INTRODUCTION

In recent years, text recognition has become an important concern. This trend is the outcome of developments in the fields of computer vision & ML & a surge in apps depending on text identification. The international organization of numerous conferences, including the ICDAR, has sparked new advancements in the field of text processing from imagery. Attention is also being paid to text detection and identification from online pages and video captions. The area of text identification from images of natural situations has seen a lot of study. There are also numerous OCR methods available [1]. The issue of text detection and recognition has still not been fully resolved. Text segmentation and extraction from natural settings remain highly challenging tasks. The stages of text identification are examined in this research along with several methods that were utilized to complete each stage. It highlights the significance of each processing stage as well as the benefits, drawbacks, and applications of the many ways utilized by contributors to address these issues. This study also reviews a number of text identification applications. The organization of this article is as follows. Section II describes methodologies. Section III briefly discusses the significant roles of text identification. In Section IV, we describe the literature survey Applications of text recognition are explained in Section V. Lastly, conclusions are presented in Section VI.

2. METHODOLOGIES

The text detection stage and text recognition stage are the two stages that make up the recognition and detection process [2]. Text recognition involves turning the collected text into letters and words, while text identification assigns to extracting text from an input picture. Stepwise approaches and integrated methods are two classifications of approaches used for this motive. The steps of identification in stepwise methods are distinct, & they move from detection to classification to segmentation to identification. In order to recognize words from the available text, integrated algorithms share data among the detection and identification phases. Block diagrams for integrated and sequential techniques are shown in Fig. 1.

2.1 Stepwise methods

Stepwise approaches use classification, segmentation, recognition, and removal of background text from each level of text detection, localization, and classification. They could be used to distinguish text from pictures regardless of text size because they are independent of lexicon size.

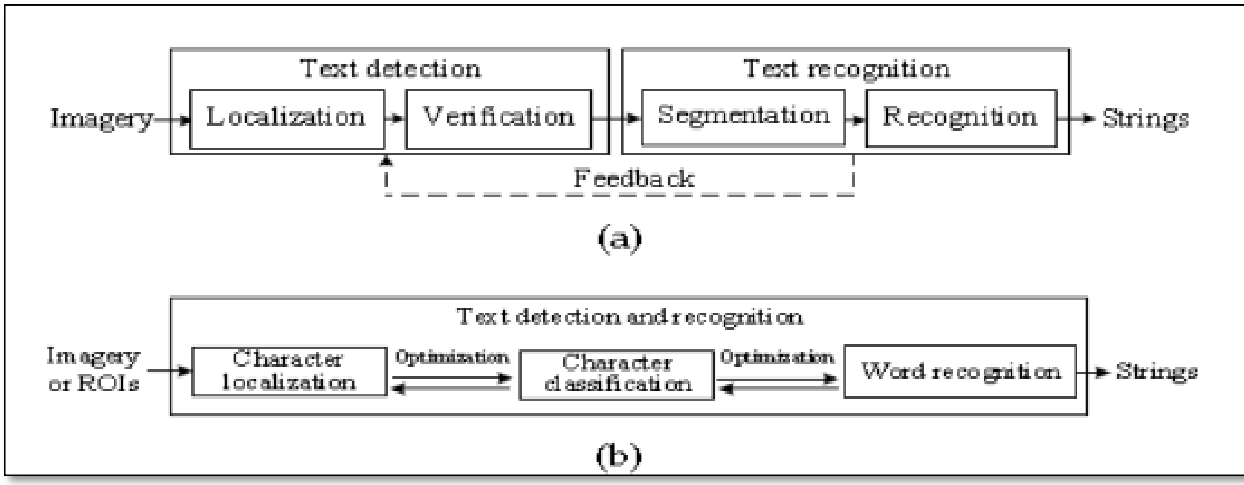

Fig.1: Block diagram of a) Stepwise approaches b) Integrated approaches [1]

2.2 Integrated methods

Integrated approaches concentrate on isolating certain words from visuals. The segmentation stage is frequently skipped in favor of the word recognition or matching stages in integrated approaches. These techniques are employed in applications that recognize a fixed set of words from a restricted lexicon.

3. Important stages of text identification

The 4 significant steps of text identification are covered in this section. Text recognition, segmentation, classification, and text localization are discussed along with their functions and significance. This section also explains the various strategies used for these stages.

3.1 Text detection and localization

While text localization pinpoints the location of the text & creates groups of text areas by removing the majority of the background, text identification assigns with identifying the existence of text in the input picture. Using linked component analysis or region-based approaches, the text is detected and localized.

3.2 Classification

False positive results from the text identification & localization stages may include non-text areas as well as text areas. Using classification techniques, the classification stage confirms text regions and removes non-text sections. This phase is also known as the verification phase. There are two types of classification algorithms: supervised and unsupervised. Before classifying text, controlled algorithms are aware of its various characteristics, including color, size, texture, etc. Unsupervised algorithms don't already know the characteristics of text.

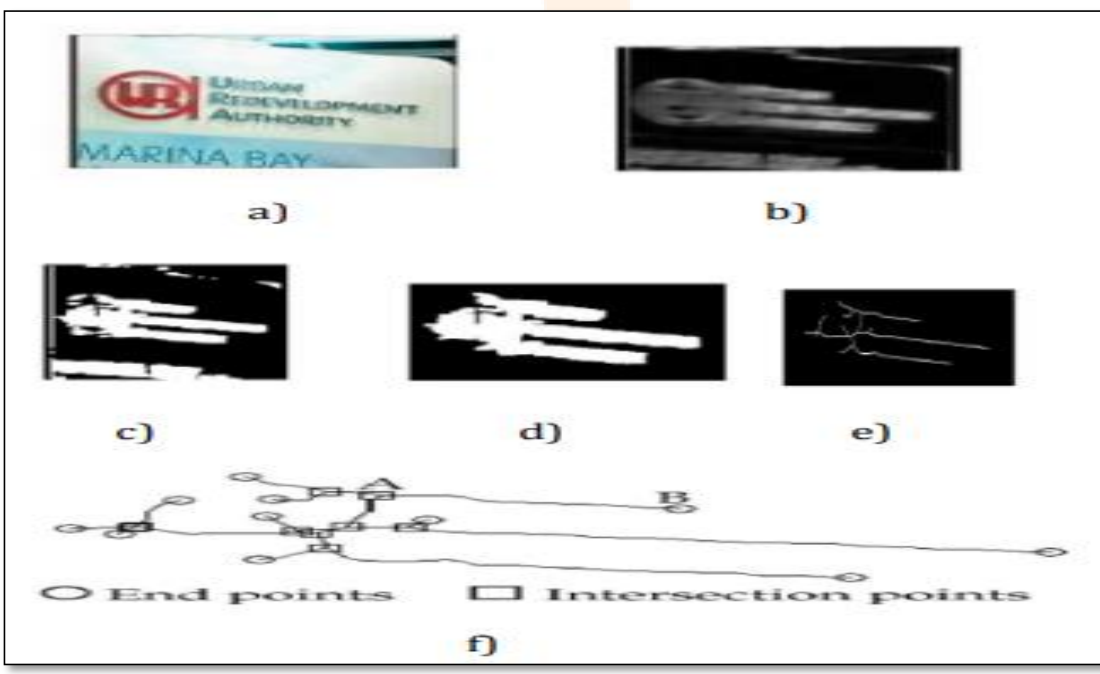

Fig. 2.a's: Extraction of linked elements. maximum difference map, the original image, the text cluster, the connected component (CC), and the skeleton of the CC f) CC's intersection and endpoints [3]

3.3 Segmentation

Text, as well as background, are separated using the segmentation technique, and bound text is extracted from images [4]. Stepwise approaches go through segmentation to acquire accurately extracted characters that are supplied to the recognition stage, unlike integrated methods that focus on word identification which frequently merge the difficult segmentation phase with the recognition phase.

3.4 Text Recognition

Text pictures are transformed into strings of characters or words during the text recognition phase. The conversion of text images into words is crucial since a person uses words as a basic entity for visual recognition [5]. Character recognition and word recognition are two various approaches to recognition.

Character recognition techniques separate a text image into several single-character cutouts. For these strategies, the space between neighboring characters is crucial.

Table 1: Comparison of approaches used for Text identification

Approaches	Features of the approach
Stepwise Approaches	Different components for identification; phases for localization, classification, segmentation, & detection; and capacity for large-scale word identification [6].
Integrated Methods	The recognition and detection modules are not independent. They are able to forego segmentation or use word recognition in its stead.

3. LITERATURE SURVEY

Yuan et al., (2020) Make a suggestion for a reliable scene text detector. The detector begins by picking up on carefully crafted, semantically significant key elements. Then, for each text instance, the main points are taught to be linked together to form a hexagon. The detector then forecasts curves along the edge of the text field after starting from a critical point. More precise text region prediction is achieved by using a straightforward heuristic post-processing method, than the predicted curve. In order to fix mistakes made during the search process, the projected important locations are employed as anchors. Given that it only requires one stage of keypoint detection and straightforward post-processing, the suggested method is effective. On numerous benchmark datasets, it performs at the cutting edge or at a level comparable to it [7].

Fayu Pan et al., (2021) By determining whether the text contains errors, tracking down the locations of those faults, and then producing repairs, you may demonstrate an intuitive multitask learning technique. A pair of classifiers are supplied to serially identify sentence-level and token-level faults since errors only take a few parts in typical datasets. In contrast to conventional methods, adopt a non-autoregressive decoder and only create words that are necessary to fix the faults that are found, hence enhancing the efficiency of the correction step. A comparison of our approach's inference speed and GEC performance with the conventional approach reveals that our approach can be up to ten times faster [8].

Tan et al., (2020) based on the structural modification of BERT, suggest a CharacterPhonetic BERT paradigm. This achieves end-to-end spelling error correction by using the BiLSTM network to locate erroneous words & and then adding the pinyin previous knowledge of the error position to the BERT network. Compared to the Bert-Finetune model without taking into account the pinyin information of the error setting, the Character-Phonetic BERT system in this paper enhances the impact by about 5%, according to the experimental results. It also increases by 2.1% when compared to the Bert-Finetune model which does take into account the pinyin data of the error position. Chinese writing is produced that is more accurate and automatically fixes mistakes [9].

Thilagavathy et al., (2021) outline a two-step process for identifying and categorizing the input information. CNNs were employed as the categorization architecture. The recognition of input text initiates the procedure. The language that was utilized to create the input number is determined in the second stage. Additionally, the handwritten words in the MNIST dataset are analyzed using Python. The simulation results demonstrate an extraordinarily high efficiency and error-free recognition rate. 1.4% training loss, 99% testing accuracy, and 99.6% training accuracy were attained by the suggested work [10].

Jin Jiang et al., (2021) Offer a new paradigm for error detection that uses a binary classifier built on the BERT. Negative samples of sentences that contain errors are used to fine-tune this classifier. To further fine-tune, the authors present a pair-shuffle training strategy. 1st experiment Results from https://github.com/jiangjin1999/Sentence-level-detection-on-CSC show that our paradigm and training strategy outperforms the SOTA model on the SIGAHN2015 database [11].

Chen Chen et al., (2023) In order to address a number of flaws and inaccuracies that occur during the detection process, suggest an innovative approach. This technique also streamlines a few laborious detection-related tasks, enhancing the model's overall performance. We have run numerous tests on a variety of benchmarks to confirm the performance and efficacy of our suggested model, which has surpassed state-of-the-art techniques in means of accuracy & speed of findings [12].

Mithilesh et al., (2020) the automatic DCA method for TDS issues, called AADGen. With this technique, DCA takes hardly any time at all and the procedure of data annotation is error-free. Through the provision of an ongoing, on-demand supply of training data, the DCA method will assist us in developing online learning system training. have performed numerous tests on the data annotation procedure and have attained an annotation accuracy of more than 99%. offer a data annotation method that is quick, scalable, flexible, and error-free [13].

Rafi Dwi Rizqullah et al., (2022) described the approach that can be taken to solve certain issues. The method employs a transformer model and a language detection module. Language detection was accomplished using a BERT model tagger, while normalization was accomplished using two ByT5 models. The study demonstrates that the ERR score for the proposed technique is 1.01 percent lower than the baseline [14].

Meifang Zhang et al., (2023) The semantic analysis method built around NLP is proposed to optimize the English translation models in order to boost the efficacy of the error correction as well as the accuracy of English machine translation. This is done by developing the NLP corpus while developing a semantic feature analysis framework for English translation text error correction. To increase the precision of English translation, semantic analysis technology built on NLP will be used. To identify the error category, the stated error class's pre-training sentence pairings are created, and by auto-regressive decoding, the masked persons in sentences of different tongues are anticipated. In addition, the large-scale sentences are transformed into corresponding correct sentences in the case of unknown scripts. The ability to read with assistance from text detection and identification can greatly improve the quality of life for blind or visually impaired people. Additionally, it is utilized for automatic cheque signature reading. Another use for text recognition is automatic document scanning [15].

Denis Eka Cahyani et al., (2022) compared word embedding algorithms like Word2Vec, BERT, and GloVe while utilizing CNN to build emotion recognition in text. The poem distinguishes between five different emotional states, including pleased, angry, sad, afraid, and astonished. The commuter line, transjakarta, and commuter line+transjakarta data types are utilized. In comparison to Word2Vec+CNN and GloVe+CNN, BERT+CNN produces the most accurate results. BERT+CNN produced accuracy values of 86.47%, 87.23%, and 86.18% for the commuter line, transjakarta, & commuter line+transjakarta data, respectively. The second-best accuracy number is held by Word2Vec+CNN, and GloVe+CNN comes in last. BERT+CNN has the highest score when compared to other approaches according to the findings of the comparison of Precision, Recall outcomes. This illustrates that text emotion recognition using the combination of CNN and BERT embedding approaches works well [16].

Deng et al., (2023) To identify all linked emotions conveyed in a given text, a MEDA is developed. MC-ESFE and ECorL make up the majority of MEDA. MEDA uses the MC-ESFE module, which is made up of numerous channel-wise ESFE systems, to capture the underlying emotion-specific properties. Through a hierarchical structure, each channel in MC-ESFE is dedicated to the FE of a specific emotion from the level of the sentence to the level of the context. An emotion sequence predictor in ECorL is used to implement emotion correlation learning using underlying characteristics. We also provide a brand-new loss function called multi-label focused loss. By combining the estimation of both positive & negative feelings with this loss feature, the system can concentrate more on misclassified positive-negative emotion pairs & enhance overall production. RenCECps and NLPCC2018 datasets are used as the emotional corpus for the evaluation of the suggested MEDA architecture. The experimental outcomes show that the proposed method can do this work more effectively than state-of-the-art approaches [17].

Kim et al., (2020) present a virtual emotion detection design that aims to deliver timely emotion-based IoT services in advanced smart cities and has two-way enabled delay bound. Additionally, the authors formally describe a challenge whose goal is to create two-way enabled virtual emotion barriers with the least amount of virtual emotion detection delay possible. Then, using comprehensive simulations in a variety of situations and circumstances, suggest a unique Two-Way-Enabled-Border-Slab approach & assess its production. We also talk about potential research problems, difficulties, and upcoming projects that are pertinent to the suggested system.

Chen et al., (2019) To make use of a significant amount of unlabeled data, suggest LLEC. First, the feature layer and the decision layer are used to examine the unlabeled data. By combining the similarity theory with the entropy approach, this study offers a hybrid label-less learning system capable of autonomously categories information without human interaction. Then, to clean up the automatically labeled data, build an improved hybrid label-less learning method. authors use improved hybrid label-less learning for multimodal unlabeled emotion information in a way to maximize the accuracy of the emotion detection design & the use of unlabeled data. Finally, in order to assess the LLEC algorithm, we construct a real-world test bed. The test findings demonstrate that the LLEC method could greatly increase the accuracy of emotion recognition.

Li et al., (2021) To help computers better reliably identify the emotions of images, we present a unique spatial & channel-wise attention-based emotion prediction design, SCEP. On the grounds that the spatial attention system can improve the contrast among salient areas and potentially irrelevant regions & that the channel-wise weight mechanism may emphasize informative functions while suppressing less useful functions, SCEP integrates both spatial attention & channel-wise weight mechanisms into a classical CNN layer framework to estimate image feelings. Numerous studies have demonstrated that SCEP could increase the accuracy of emotion prediction (as measured by the concordance correlation coefficient) by 3%–5% in the arousal domain and by 3-6% in the valence domain compared to base models (i.e., VGG and ResNet) without spatial attention or channel-wise methods. As a result, draw the conclusion that SCEP can improve emotion prediction accuracy.

5. Applications of Text Identification

With improvements in approaches for image processing, a number of applications for word identification or recognition from photos and videos have appeared in recent centuries. Applications for text identification are growing as a result of improvements in various embedded systems and growing research in computer vision and machine learning. Industries utilize text detection and recognition to read package labels, numbers, etc. It is utilized to recover specified text contents from web pages as well as video captions. It is utilized at toll booths for automatic number plate identification & for reading street signs when unmanned vehicles are involved. The ability to read with assistance from text detection and identification can greatly improve the quality of life for blind or visually impaired people. Additionally, it is utilized for automatic cheque signature reading. Another use for text recognition is automatic document scanning.

6. CONCLUSION

The processes of text identification, as well as several strategies utilized to do so, are given in this study. Character identification and localization, categorization, and segmentation are additional divisions of this process. This study presents these steps and compares the methods utilized to proceed through the stages indicated above. Analyses of the benefits, drawbacks, and applications of various strategies have also been carried out here.

REFERENCES

[1] Qixiang Ye, David Doermann, “Text Detection and Recognition in Imagery: A Survey” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014.

[2] K. Elagouni, C. Garcia and P. Sbillot,” A Comprehensive Neural-Based Approach for Text Recognition in Videos using Natural Language Processing,” in Proc. ACM Conf. Multimedia Retrieval, 2011.

[3] P. Shivakumara, T.Q. Phan and C.L. Tan,” A Laplacian Approach to Multi-Oriented Text Detection in Video,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 33, no. 2, pp. 412-419, 2011.

[4] F. Ren and K. Matsumoto, “Semi-Automatic Creation of Youth Slang Corpus and Its Application to Affective Computing,” IEEE Trans. Affect. Comput., vol.7, no.2, pp. 176-189, 2016.

[5] D. Sznycer, and A.W. Lukaszewski, “The emotion–valuation constellation: Multiple emotions are governed by a common grammar of social valuation,” Evol. Hum. Behav., vol. 40, no. 4, pp. 395-404, Jul. 2019.

[6] P. Xu, A. Madotto, C.S. Wu, J.H. Park, and P. Fung, “Emo2vec: Learning generalized emotion representation by multi-task training,” arXiv preprint arXiv:1809.04505, 2018.

[7] Yuan, K., He, D., Yang, X., Tang, Z., Kifer, D., & Giles, C. L., “Follow the Curve: Arbitrarily Oriented Scene Text Detection Using Key Points Spotting And Curve “, IEEE International Conference on Multimedia and Expo (ICME).

[8] Fayu Pan;Bin Cao, “Efficient Grammatical Error Correction with Hierarchical Error Detections and Correction”, 2021 IEEE International Conference on Web Services (ICWS).

[9] Tan, M., Chen, D., Li, Z., & Wang, P. (2020). Spelling Error Correction with BERT based on Character-Phonetic. 2020 IEEE 6th International Conference on Computer and Communications (ICCC).

[10] A. Thilagavathy; Karuturi Hemanth Suresh; Katari Tejesh Chowdary; Meka Tejash; Vijayarao, “Text Detection based on Deep Learning”, International Conference on Innovative Data Communication,2023.

[11] Jin Jiang; Yanquan Zhou, “A Sentence Classification Method for Chinese Spelling Error Detection Based on BERT”, International Conference on Asian Language Processing (IALP),2021.

[12] Chen Chen; Fang Zheng; Shaohui Zhang; Li Wei, “A New Strategy for Reducing Errors in Scene Text Detection”, 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL),2023.

[13] Mithilesh Kumar Singh, “AADGen: Automatic Annotated Data Generation for Training Text Detection and Recognition Models”, AADGen: Automatic Annotated Data Generation for Training Text Detection and Recognition Models 2020 IEEE International Conference for Innovation in Technology (INOCON),2020.

[14] Rafi Dwi Rizqullah; Indra Budi, “Text Normalization on Code-Mixed Twitter Text using Language Detection”, Seventh International Conference on Informatics and Computing (ICIC),2022.

[15] Meifang Zhang; Yanhong Zhang, “The Study on NLP-based Semantic Analysis Technology to Improve the Accuracy of English Translation “, IEEE International Conference on Control, Electronics and Computer Technology (ICCECT)**,**2023.

[16] Denis Eka Cahyani; Aji Prasetya Wibawa; Didik Dwi Prasetya; Langlang Gumilar, “Emotion Detection in Text Using Convolutional Neural Network”, 2022 International Conference on Electrical and Information Technology (IEIT),2022.

[17] Deng, J., & Ren, F., “Multi-label Emotion Detection via Emotion-Specified Feature Extraction and Emotion Correlation Learning”,IEEE Transactions on Affective Computing, Volume: 14, Issue: 1,2020.

[18] Kim, H., Ben-othman Jalel, Mokdad, L., & Bellavista, P., “A Virtual Emotion Detection Architecture with Two-way Enabled Delay Bound toward Evolutional Emotion-based IoT Services”, IEEE Transactions on Mobile Computing, Volume: 25, Issue: 6,2020.

[19] Chen, M., & Hao, Y., “Label-less Learning for Emotion Cognition”, IEEE Transactions on Neural Networks and Learning Systems, 1–11, Volume: 31, Issue: 7,2019.

[20] Li, B., Ren, H., Jiang, X., Miao, F., Feng, F., & Jin, L., “SCEP—A New Image Dimensional Emotion Recognition Model Based on Spatial and Channel-Wise Attention Mechanisms”, IEEE Access, 9, 25278–25290,2021.