**Logo

Description automatically generated**

**Teesside University**

**School of Computing, Engineering and Digital Technologies**

**MSc Artificial Intelligence**

**Research Methods**

**Literature review report**

Student Name : Tarun Tatakuntala

Student ID : B1231849

Module Name : Research Methods

Module Code : CIS4011-N-FJ1-2021

Module Leader : Dr. Zia Ush Shamszaman

Assignment Submission Date : 11th May 2022

Contents

1. Introduction…………………………………………………………………………………………………………3
   1. Research Question…………………………………………………………………………………………4
2. Database search…………………………………………………………………………………………………..5
   1. IEEE Database………………………………………………………………………………………………..5
   2. ACM Database……………………………………………………………………………………………….6
   3. Science direct Database…………………………………………………………………………………7
   4. Literature Evaluation……………………………………………………………………………………..8
3. Inclusion and Exclusion Process……………………………………………………………………………8
4. Conclusion……………………………………………………………………………………………………………10
5. References…………………………………………………………………………………………………………..11

**EFFICIENT MODEL FOR PRECISE TEXT EXTRACTION FROM VIDEO USING TEXT LABELLING & DEEP LEARNING MODELS**

1. **INTRODUCTION**

For multimedia comprehension and retrieval, video text extraction is essential. Individual frames have been used in most past of the existing models. The semantic signs concerning video material are provided by video text. The extraction of text from video is an important part of the analysis process. The majority of publications extract video text employing sensitive to the video backdrop stroke and intensity features. Because of the complicated background of video frames, extracting text characters is difficult. The structure of text characters can be found in their phase information, which is unaffected by backdrop. Deep learning-based text identification methods have been used in a variety of disciplines, including Intelligent Transportation Systems (ITS), Automated Driving Systems (ADS) and text recognition from Videos. Meanwhile, text detection and recognition in various contexts has received a lot of attention and research effort. The video text includes crucial information for analyzing, indexing, and recovering video footage. The primary technique for getting this information is the finding, verification and recognition of video text that is in numerous languages and types against difficult backdrops. An intense survey on recent works by different authors gives us scope for further enhancement in the proposed challenge.

Artificial text is sometimes known as motion text. It frequently contains important video content in a concise manner. In news videos, for instance, captions offer facts about the storey, but subtitles in speech videos are likely to summarise important themes. The goal of this study is to identify and characterise the text in video frames. OCR software applications for video employ a range of standard processing paradigms, including text detection, localization, extraction, and recognition. The first phase, called detection, is looking for text regions while the second phase, called localization, is more concerned with locating exact text lines. Extracting clear text from a difficult background is the focus of the extraction.

Deep learning is essential to the operation of driverless cars, enabling them to differentiate between a pedestrian and a lamppost, or even to detect a stop sign or detecting text from videos. Deep learning has seen a lot of recent publicity, and rightly so. It's showing positive results that were never achievable previously. Computers that use deep learning learn to classify things via photos, text, or sound. Deep learning models are used in the process of recognizing text from videos. Scene text can appear on any surface, in any position, and in any size or perspective, making extraction extremely challenging. Figure 1 depicts an example of a video text detection and extraction system.



Fig 1: Text Detection in a Video Frame

* 1. **Research Question**

What is the efficient model for precise text extraction from video using text labelling and deep learning models?

**Keywords:** Video Text Detection, Text Recognition, Corner Response Feature Map, Deep Learning, Precise Text Extraction, Text Labelling, Video Model.

1. **DATABASE SEARCH**
   1. **IEEE Database**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Search Number** | **Question** | **Keywords** | **Filters** | **Results** | **Notes** |
| 1 | What has been  published on  text extraction? | **“**text extraction” | None | 17312 | Relevant journals are displayed at the top of the  list however the list is  too long for review.  Keywords such as  Text extraction  appear in the abstracts. |
| 2 | What has been  published on  text extraction? | **“**text extraction” | Year:2010-2022  Includes journals and conferences | 1448 | Too many search results |
| 3 | What has been  published on  text extraction? | **“**text extraction” | Year:2017-2022  Includes journals and conferences | 8890 | Relevant journals  are displayed at the top of the list. However, the list is too long for review. |
| 4 | What has been  published on  text extraction? | **“**text extraction” | Year:2017-2022  Includes journals  only | 1138 | Some useful papers on text extraction from videos |
| 5 | What has been  published on  text extraction? | **“**text extraction” | Year:2017-2022  Includes journals  only  **Filter:**  IEEE Transactions on Image Processing | 44 | Mixed results are displayed |
| 6 | What has been  published on  text extraction? | **“**text extraction” | Year:2017-2022  Includes journals  only  **Filter:**  IEEE Transactions on Image Processing,  Text detection | 13 | Relevant papers are found. |
| 7 | What has been  published on  text detection? | “text detection” | No filter | 9622 | Some journals are displayed but there are too many search results. |
| 8 | What has been  published on  text detection? | “text detection” | Year:2017-2022 | 4892 | Relevant journals  are displayed at the top of the list. However, the list is too long for review. |
| 9 | What has been  published on  text detection? | “text detection” | Year:2017-2022  **Filter:**  Journals | 692 | Mixed results are displayed. |
| 10 | What has been  published on  text detection? | “text detection” | Year:2017-2022  **Filter:**  Journals  IEEE Transactions on Pattern Analysis &Machine Intelligence | 16 | Relevant journals found. |

* 1. **ACM Database**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Search Number** | **Question** | **Keywords** | **Filters** | **Results** | **Notes** |
| 11 | What has been  published on  text extraction? | **“**text extraction” | None | 38985 | Relevant journals  are displayed at the top of the  list however the list is  too long for review.  Keywords such as  Text extraction  appear in the abstracts. |
| 12 | What has been  published on  text extraction? | **“**text extraction” | Year:2010-2022  Includes journals and conferences | 26766 | Too many search results |
| 13 | What has been  published on  text extraction? | **“**text extraction” | Year:2017-2022  Includes journals and conferences | 14499 | Relevant journals  displayed at the top of the  list however the list is  too long for review. |
| 14 | What has been  published on  text extraction? | **“**text extraction” | Year:2020-2022  Includes journals  only | 6636 | Relevant journals  Displayed at the top of the  List. However, the list is  too long for review. |
| 15 | What has been  published on  text extraction? | **“**text extraction” | Year:2021-2022  Includes journals  only  **Filter:**  Journals | 3383 | Mixed results are appeared |

* 1. **Science direct Database**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Search Number** | **Question** | **Keywords** | **Filters** | **Results** | **Notes** |
| 16 | What has been  published on  text extraction? | **“**text extraction” | None | 240325 | Too many search results |
| 17 | What has been  published on  text extraction? | **“**text extraction” | Year:2010-2022  Includes journals and conferences | 126838 | Relevant  journals. Keywords such as  Text extraction  appear in the abstracts. |
| 18 | What has been  published on  text extraction? | **“**text extraction” | Year:2017-2022  Includes journals and conferences | 78997 | Relevant journals  appear at the top of the  list however the list is  too long for review. |
| 19 | What has been  published on  text extraction? | **“**text extraction” | Year:2020-2022  Includes journals  only | 30039 | Some useful papers on text extraction from videos |
| 20 | What has been  published on  text extraction? | **“**text extraction” | Year:2022  Includes journals  only | 11121 | Relevant papers found |

* 1. **Literature Evaluation**

Due to the continually advancing and creating nature of the computer and innovation businesses, it is vitally imperative that the larger part of data collected has been as of late distributed and covers innovation that's as of now utilized or in development for the longer term instead of obsolete software and equipment. For the larger part of looks a distributed date age extend was connected so that only the foremost later and up to date sources were recovered by the databases. In a few cases it may be suitable to see at more seasoned papers where innovation has appeared restricted progresses or the principles have not changed since the distribution of the source fabric.

Out of three databases, IEEE provides better search results as there are many filters present. Science direct database has more search results and less filter options which make it difficult to find the journals. IEEE has separate “Image processing” filter through which most of the journals are found easily.

1. **INCLUSION AND EXCLUSION PROCESS**

|  |  |  |  |
| --- | --- | --- | --- |
| **Paper title** | **Method** | **Include**  **Or**  **Exclude** | **Reason** |
| Ohya J, Shio A, Akamatsu S, “Recognizing characters in scene images,” IEEE Trans, pp 214-224, 1994. | Adaptive Thresholding, Relaxational operation | Exclude | Too old paper. Many advanced techniques are present. |
| Lienhart RW, Stuber F, “Automatic text recognition in digital videos,” Proc SPIE, 2666(3), pp. 180-188, 1996. | Segmentation, OCR | Include | 97% accurate on Gray scale images |
| Jae-Chang S, Dorai C, B- R, “Automatic text extraction from video for content based annotation and retrieval,” Pattern Recogn., | Generalized region labelling algorithm, OCR | Exclude | False positive is reduced from 21 to 2, handles text in horizontal and diagonally oriented |
| Hasan YMK, Karam LJ, “Morphological text extraction from images”, IEEE Trans Image Process 9(11), pp. 1978-1983, 2000. | Morphological technique , | Include | Robust to noise, insensitive to skew and text orientation  Good readability rate with 93%. |
| Palma D, Ascenso J, Pereira F, “Automatic text extraction in digital video based on motion analysis,” Image Analysis and Recognition, Springer, pp 588-596, 2004. | Geometric and spatial analysis, | Exclude | High performance but overlapping the text |
| Qixiang Ye, Qingming Huang, Wen Gao, Debin Zhao,” Fast and robust text detection in images and video frames,” Image and Vision Computing, Elsevier, Jan-2005. | SVM classifier | Include | Low false alarms |
| C. Mancas-Thillou, B.Gosselin, “Spatial and color spaces combination for natural scene text extraction,” IEEE Int. Conf. Image Process., pp. 985-988, 2006. | Log-Gabor filter+OCR | Exclude | Low resolution, poor contrast |
| R. minetto, N. Thome, M.Cord, J. Fabrizio, B. Marcotegui,” A multiresolution system for text detection in complex detection in complex visual scenes,” IEEE Conf, vol. 1, pp 3862-3864, 2010. | multiresolution algorithm, Snooper Text | Include | improved in multi-scale image |
| Lee.S, Cho.M, Jung.K, J.Kim, “Scene text extraction with edge constraint and text collinearity,” IEEE Int. COnf. COmput. Vis. Pattern Recognit., pp. 3983-3986, 2010. | K-means clustering +Markov random Fields | Include | Accuracy is high |
| Lukas Neumann, Jiri Matas, “Real-Time Scene Text Localization and Recognition,” IEEE conference on CVPR-2012, June-2012. | LSP, Bayesian classifier | Include | Robust |
| Shivakumara P, Sreedhar RP, Phan TQ, Shijian L, Tan CL, “Multi-oriented video scene text detection through Bayesian classification and boundary growing,” IEEE Trans (TCSVT), pp 1227-1235, 2012. | Deep NN | Include | High performance w.r.t non-horizontal data |
| Amit Panwar, Himanshu Suyal, “Fusion based text detection and localization for video image,” IJRASET. Vol. 4 Issue. IV ,ISSN: 2321-9653, April-2016. | Fusion based approach | Exclude | Not applicable for complex background |

1. **CONCLUSION**

This research highlights the extensive survey on text detection from video. We have surveyed numerous previous papers and compared the methods and techniques used. The problem of text detection can be divided into a) detection b) tracking c) segmentation d) extraction e) recognition. There are many algorithms to detect and recognize the text from video. However, there is no unique method to satisfy scene text, caption text and both caption and scene text.

Although, many algorithms have referred and implemented to resolve the issue of detection and recognition of text, there is no efficient system for users. Earlier OCR was used to recognize the text. However, it is not appropriate to imageries and video. To apply in real time applications, we need to work with more efficient methods like machine learning, hybrid techniques, fuzzy logic etc. Also, this paper described the datasets and performance measures used by different authors. This paper mainly conveys how to classify, analyze, and review these algorithms, and discuss the future work to be done.

1. **REFERENCES**

[1] Ohya J, Shio A, Akamatsu S, “Recognizing characters in scene images*,” IEEE Trans*, pp 214-224, 1994.

[2] Lienhart RW, Stuber F, “Automatic text recognition in digital videos,” *Proc SPIE*, 2666(3), pp. 180-188, 1996.

[3] Jae-Chang S, Dorai C, B- R, “Automatic text extraction from video for content based annotation and retrieval,” Pattern Recogn.

[4] Hasan YMK, Karam LJ, “Morphological text extraction from images”, *IEEE Trans Image Process 9(11)*, pp. 1978-1983, 2000.

[5] Palma D, Ascenso J, Pereira F, “Automatic text extraction in digital video based on motion analysis,” Image Analysis and Recognition, *Springer, pp* 588-596, 2004.

[6] Qixiang Ye, Qingming Huang, Wen Gao, Debin Zhao,” Fast and robust text detection in images and video frames,” Image and Vision Computing, *Elsevier*, Jan-2005.

[7] R. minetto, N. Thome, M.Cord, J. Fabrizio, B. Marcotegui,” A multiresolution system for text detection in complex detection in complex visual scenes,” *IEEE Conf, vol. 1*, pp 3862-3864, 2010

[8] Lee.S, Cho.M, Jung.K, J.Kim, “Scene text extraction with edge constraint and text collinearity,” *IEEE Int. Conf. Comput. Vis. Pattern Recognit.,* pp. 3983-3986, 2010.

[9] Lukas Neumann, Jiri Matas, “Real-Time Scene Text Localization and Recognition,” *IEEE conference on CVPR-2012*, June-2012.

[10] Shivakumara P, Sreedhar RP, Phan TQ, Shijian L, Tan CL, “Multi-oriented video scene text detection through Bayesian classification and boundary growing,” *IEEE Trans (TCSVT)*, pp 1227-1235, 2012.

[11] Amit Panwar, Himanshu Suyal, “Fusion based text detection and localization for video image,” *IJRASET. Vol. 4 Issue. IV* ,ISSN: 2321-9653, April-2016.

[12] D. Ciresan, U.Meier, J.Schmidhuber, “ Multi-column deep neural networks for image classification,” *IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 3642-3649, 2012.

[13] A. Krizhevsky, I. Sutskever, G.E.Hinton, “Imagenet classification with deep convolution neural networks,” *Advances in neural information processing systems*, pp. 1097-1105, 2012.

[14] K. Simonyan, A.Zisserman, “Very deep convolutional neural networks for large-scale image recognition,” 2014.

[15] C.Szegedy, W. Liu, Y.Jia, P.Sermanet, S.Reed, D.Anguelov, D.Erhan, V.Vanhoucke, A. Rabinovich, “*Going Deeper with convolutions,”* 2014.

[16] M. Zeiler D, R.Fergus, “Visualizing and understanding convolutional networks,” *ECCV 2014*, pp. 818-833, 2014.

[17] A. Yilmaz, O.Javed, M.Shah, “Object tracking: A survey,” *ACM, Comput. Vis.* pp. 776-783, 2014.

[18] C. Yao, X.Bai, W.Liu, “A unified framework for multioriented text detection and recognition,” *IEEE Trans. Image Process., vol. 23, no. 11*, pp. 4737-4749, 2014.

[19] C. Yao, X.Bai, W.Liu, Y.Ma, Z.Tu, “Detecting texts of arbitrary orientations in natural images,” *CVPR*, pp. 1083-1090, 2012.

[20] X. C.Yin, W.Y.Pei, J.Zhang, Hong Wei Hao, “Multi-orientation scene text detection with adaptive clustering,” *IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9,* pp. 1930-1937, 2015.

[21] L.Kang, Y.Li, D.Doerman, “Orientation robust text line detection in natural images*,” CVPR*, pp. 4034-4041, 2014.

[22] B. Epshtein, E.Ofek, Y.Wexler, “Detecting text in natural scenes with stroke width transform,” *CVPR*, pp. 2963-2970, 2010.

[23] Xu-C Yin, X Yin, Kujang, Hong Wei Hao, “Robust text detection in natural scene images,” *IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 5*, pp. 970-983, 2014.