



## Data Extraction of Digital Text and Conversion into an Audio-Visual Output

Niharika Rajaram

Dept of Computer Science and Engineering

K S Institute of Technology  
Bengaluru, Karnataka, India

Rajeev Koushik Y G

Dept of Computer Science and Engineering

K S Institute of Technology  
Bengaluru, Karnataka, India

Monisha R

Dept of Computer Science and Engineering

K S Institute of Technology  
Bengaluru, Karnataka, India

K Venkata Rao

Dept of Computer Science and Engineering

K S Institute of Technology  
Bengaluru, Karnataka, India

Nikhila M.Y

Dept of Computer Science and Engineering

K S Institute of Technology  
Bengaluru, Karnataka

**Abstract**—Visual content provides an improved quality of learning. The proposed project is an automatic video generator, which creates a video from a textbook chapter. The video will consist of narration with subtitles of the whole chapter while showing images relevant to the current paragraph. In order to retrieve the images, web scraping will be used. Topic selection and keyphrase analysis will be done and the resulting keyphrase will act as a query for image search. Audio files are generated with the help of retrieved text. The images, audio and text are integrated to create a final video

**Keywords**—OCR, Data Extraction, Key phrase extraction, Text Processing.

### I. INTRODUCTION

The experience of being unable to follow lectures in classrooms, especially with regards to subjects of the theoretical nature, is quite relatable among students across the world. It is difficult to find study materials that are engaging, cover the syllabus and also aide in retaining the contents of the textbooks in memory. Therefore, there is a need for a product that aims at tackling exactly the above mentioned issues.

We propose an application that allows students to provide a PDF version of the prescribed source materials as an input and retrieve an output in the form of a video, complete with subtitles and audio. This video would contain images related to the text in PDF, which introduces an audio-visual

element into the learning process. This tailor-made output can greatly enhance the ability of students to grasp newer concepts quickly and help remember what they study for a longer duration of time.

### LITERATURE SURVEY

A. *A unified scheme of text localization and structured data extraction for joint OCR and data mining*

This paper discusses the integration of text detection and structured data extraction into a unified deep learning-based Image Text Extraction (ITE) scheme.

With the guidance of the above paper, we have been able to:

- i. Learn the concept of Image Text Extraction (ITE).
- ii. Understand the procedure of approaching text detection, text recognition and structured data extraction, along with the application of deep learning.
- iii. Recognize the various fields of applications of the ITE technology.

The paper gave us valuable insights into the applications of Optical Character Recognition (OCR).

B. *An efficient approach for Key phrase Extraction from English Document*

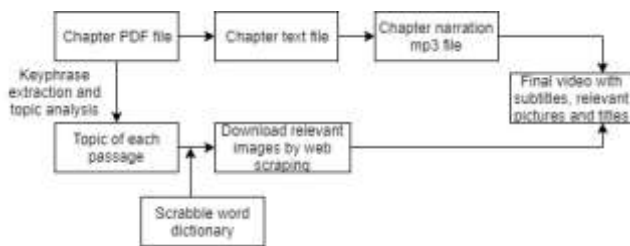
The above paper was helpful for us to:

- i. Define the concept of key phrases.
- ii. Study the methodology in key phrase nomination, weight measurement and key phrase ranking and selection.
- iii. Compare several different stemming algorithms.

We were determined to find a suitable system which yielded the highest efficiency in terms of searching for images relating to key terminologies.



## DESIGN AND DEVELOPMENT OF CROP YIELD PREDICTION APPLICATION



The project is presented in the form of a mobile application which is user friendly. The input to this application would be a textbook chapter PDF. This PDF is processed to create a video which contains images, subtitles and audio. The video is the output of the application.

### CONCLUSION

Students get bored when teachers read out passages from textbooks but are more enthusiastic to watch videos and pictures. The easiest way for students to grasp concepts is through visuals. The learning experience impacts significantly on understanding the concepts. This system auto generates narrated films from PDF/textbook chapter. Thus, students pay more attention in class and retain knowledge much better by associating pictures with information.

Further development can be made on topic analysis and the process of keyphrase extraction, to show more important and relevant images to the current paragraph.

### REFERENCES

[1] A unified scheme of text localization and structured data extraction for joint OCR and data mining, Yibin Ye, Shenggao Zhu, Jing Wang, Qi Du, Yezhang Yang, Dandan Tu, Lanjun Wang and Jiebo Luo Cloud BU, Huawei Technologies, University of Rochester (2018)

[2] An Efficient Approach for Keyphrase Extraction from English Document, International Journal of Intelligent Systems and Applications (2017)