

St Vincent Pallotti College of Engineering and Technology

Department of Information Technology

Academic Year: 2019-20

“CREATION OF SPELLCHEKER FOR IMAGE PROCESSING ON TESSARACT-OCR and GOOGLE DOC”

1) Abstract:

Optical character recognition or optical character reader (OCR) is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, or from subtitle text superimposed on an image

Image processing is an absolutely crucial area of work for those groups and industries that are working in areas such as medical diagnostics, astronomy, geophysics, environmental science, data analysis in laboratories, industrial inspection, etc. Although not originally developed for users who are visually impaired, Optical Character Recognition (OCR) technology has become an aid for inputting documents quickly by and for users with vision impairments. A complete OCR system consists of a scanner, the recognition component, and OCR software that interacts with the other components to store the computerized document in the computer. The process of inputting the material into the computer begins with the scanner taking a picture of the printed material. Then, during the recognition process, the picture is analyzed for layout, fonts, text and graphics. Finally, the picture of the document is converted into an electronic format that can be edited with an application software.

2) Hardware & Software requirements

Linux 16.01 or higher versions.v

Anusaaraka should be installed in Operating Systems.

Tesseract-OCR to be installed.

3) Literature Survey

A better understanding of the Anusaaraka tool is proposed by its developers and made it freely available on https://ltrc.iiit.ac.in/Dictionaries/Dict_Frame.html. All the Dictionaries, Translated Dictionaries for Hindi, Marathi, Kannada, Bengali and other traditional languages are provided. For understanding the working of these dictionaries it is important to debug its automation codes. For every root dictionary there is a methods by which it is incorporated in the tool and made accessible to the Translator and Alignment tools and different phases of the model. The different layers of the tool are supported by different set of dictionaries. And at every layer, the root dictionaries are equally important for correct and errorless translation of the given sentence.

4) References:

- [1] https://ltrc.iiit.ac.in/Publications/anu_brief.html
- [2] <http://sampark.iiit.ac.in/anusaaraka/>
- [3] ANUSAARAKA: MACHINE TRANSLATION IN STAGES : By Dr. Vineet Chaitanya, P. Kulkarni Rajeev Sangal Satyam, from School of Applied Information Systems , IITH.

Roll No	: 20 ,29,53	Name of Student :	Parineeta Warade Shruti Bangre Rahul Raipurkar
Signature	:	Date	:
Name of the guide	: Prof. Vikas Bhowate		
Signature with date	:		
Status	: Approved / Approved with modification / Not Approved		
Remarks	:		
Sign of HOD :	Date		:
Remarks	:		

St Vincent Pallotti College of Engineering and Technology

Department of Information Technology

Academic Year: 2019-20

Guidelines for writing abstract

- All paragraphs will be **single spaced** with a **double space** between each paragraph. Each paragraph will begin with a **five-space indentation**
- Seminar **titles** should be **bold** with **14 pt** typed in all **CAPITAL** letters and should be aligning at the **center** of the page.
- Main part of abstract should be **Times Roman, 12 pts , justified and 1 (Single) line spacing**
- Creating a reference list or bibliography **Few Standard format** given below:
 - **Journal Articles**

Example

1] Graham, Elspeth and Boyle, Paul. Editorial introduction:(re)theorising population geography: mapping the unfamiliar. *International Journal of Population Geography*. Vol. 7, no. 6, 2001: 389-394

2] Lundmark, Linda. Economic Restructuring into Tourism in the Swedish Mountain Range. *Scandinavian Journal of Hospitality and Tourism*. Vol. 5, no. 1, 2005: 23–45. doi: 10.1080/15022250510014273.

- **Parts of a Book**

Example:

Ellet, Elizabeth F. L. By rail and stage to Galena. In *Prairie state: impression of Illinois, 1673-1967, by travelers and other observers*, Paul M. Angle (ed.), 271-79. Chicago: University of Chicago Press, 1968.

- **E-Books**

Example

Bowen, Natasha K. & Guo, Shenyang. *Structural equation modeling*. New York: Oxford University Press, 2012. E-book.

- **Internet Documents**

Example

Organisation for Economic Co-operation and Development (OECD). Health: OECD says governments must fight fat. 2010. http://www.oecd.org/document/35/0,3343,en_21571361_44315115_46064099

St Vincent Pallotti College of Engineering and Technology
Department of Information Technology

Academic Year: 2019-20