Real-Time Note Maker

Krutarth Chovatiya[1], Vatsal Kacchia[2], Mani Butwall[3]

[1]Student VIII SEM, B.E., Computer Science & Engineering, Institute of Technology & Management Universe, Vadodara, Gujarat India

[2]Student VIII SEM, B.E., Computer Science & Engineering, Institute of Technology & Management Universe, Vadodara, Gujarat India

[3]Asst. Professor, B.E., Computer Science & Engineering, Institute of Technology & Management Universe, Vadodara, Gujarat India

---***---

Abstract - *While attending an Online Webinar/Lecture or browsing the internet sometimes you want to make notes or draw diagrams appearing on the screen but it will take too much time or you are too tired to do so. Also, in live online lectures or seminars, you cannot pause it and if you do so you will miss further content.*

Then you can use our "Real-time Note Maker" browser extension. Clicking on the extension will ask you to select a region after selecting it will save the screenshots in the device storage. The Make Notes feature on the browser extension will convert all the selected images to notes (into file extensions like pdf, doc, or ppt).

Key Words: *Note Making, Web API, Optical Character Recognition(OCR), Web Browser Extension*

1. INTRODUCTION

Our project is a web browser extension that will capture screenshot using various features and convert images of variable size and aspect ratios to an editable pdf with help of OCR and deep-learning techniques.

First, the extension will capture an image from the region of interest selected by the user then it will send to the server. The server will then convert the image to an editable pdf and send it back to the extension.

1.1 Overview

Note-making is a process of recording information from different sources and platforms. By taking notes, the writer records the essence of the information, freeing their mind from having to recall everything. From last year since the pandemic started most of the lectures, seminars and meetings are held by online means. So the number of people studying online has drastically increased for which we are presenting a new way to make notes. Our solution is a browser extension for online note-making.

Clicking on the extension will ask the user to select a region after the region is selected it will save the screenshots in the device storage. The Make Notes feature on the browser

extension will convert all the selected images to notes (into file extensions like pdf, doc, or ppt).

1.2 Advantages of Real-Time Note Maker

1. Faster note generation: As compared to the conventional way of note making which takes an average man kind to write approximately 5 to 20 wpm real-time note maker can note multiple pages (diagrams included) in seconds.

2. Easy to share: Generated notes are in digital copy in pdf/ppt/doc formats so they are easy to share.

3. Editable note: The generated notes could be edited in any way the user wants.

4. Image editing: Browser extension can perform image editing before notes generation

2. RELATED WORK

[1] Yuning Du, Chenxia Li, Ruoyu Guo, Xiaoting Yin, Weiwei Liu, Jun Zhou, Yifan Bai, Zilin Yu, Yehua Yang, Qingqing Dang, Haoshuang Wang

They propose a practical ultra-lightweight OCR system, i.e., PPOCR. The overall model size of the PP-OCR is only 3.5M for recognizing 6622 Chinese characters and 2.8M for recognizing 63 alphanumeric symbols, respectively. They introduce a bag of strategies to either enhance the model ability or reduce the model size.

[2] Linjie Xing, Zhi Tian, Weilin Huang, and Matthew R. Scott

This paper proposes convolutional character networks, referred as CharNet, which is a one-stage model that can process two tasks simultaneously in one pass. CharNet directly outputs bounding boxes of words and characters, with corresponding character labels. They utilize character as basic element, allowing us to overcome the main difficulty of existing approaches that attempted to optimize text detection jointly with a RNN-based recognition branch.

[3] Ray Smith

The Tesseract OCR engine, as was the HP Research Prototype in the UNLV Fourth Annual Test of OCR Accuracy, is described in a comprehensive overview. Emphasis is placed on aspects that are novel or at least unusual in an OCR engine, including in particular the line finding, features/classification methods, and the adaptive classifier.

3. SYSTEM ANALYSIS

3.1 Problem Definition

The classic method of making note of what appears on screen is very slow and time consuming. If the video sessions are live, then you can miss some points to note. It gets worse when you have to note diagrams. The machine generated notes of what appears on screen is the solution. The objective is to create a browser extension that will capture a selected portion from a video or a web page and make notes in pdf format.

3.2 Proposed System Feature

The proposed system can run on any Web browser which supports browser Extension. It will be able to take a screenshot of any web page. The screenshot functionality has two modes first is a full-page screenshot and second is the ability to take a screenshot of a selected portion of the web page. After taking a screenshot the user will have options to either edit the screenshot with the help browser extension or directly save the screenshot on the device. The proposed system is able to take an image from the local machine and then make notes in real-time. These notes will be generated in PPT/PDF/DOC format. The system performs its process with accuracy and precision to avoid problems. It is reliable and flexible. The system destroys the input data as soon as output data is created.

4. SYSTEM DESIGN AND IMPLEMENTATION
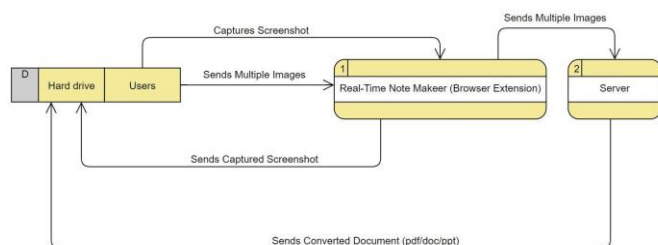
4.1 Proposed Real-Time Note Maker

Figure 1: Proposed diagram of Real-Time note maker

The planned model of the Real-Time note maker system is as shown in figure 1.

The proposed real time note maker has two components. First is the browser extension installed on the web browser of the user and second is the server side Flask API. The browser extension is responsible for taking the screenshot and then editing the screenshot in real time.

The browser extension is made with the help of Ajax, JavaScript, then it saves the images to the local machine. Extension also makes API call for notes generation. Second component is the web API which is the heart of this design. The server API is made using Tesseract OCR Engine, OpenCV and Flask. API performs OCR, and generates notes in form of pdf/ppt/doc. Then send it back to the sender in the form of a response.

4.2 Proposed Real-Time Note Maker's Functions

The Proposed Real-Time note maker System has the abilities to perform the following functions:

1. Convert images to ppt/pdf/doc.
2. Take a screenshot of the selected portion or full web page.
3. Edit the screenshot taken in real time.
4. Perform OCR.
5. It can run on any Web browser which supports browser Extension

4.3 Software

4.3.1 Front End Design

Talking about the software design the "web browser extension" which is used in this project and was our main aim that users could be able to use this while watching any informative videos or attending any online lecture for making notes. So in our front-end design, we have screen capture features by which users can grab screenshots from the webpage.

The screen capture includes two options one is a capture screen of a whole web page and the other is capturing a screenshot of the selected region where the user will be able to select the region with a nice user interface.

Also, we have included an image editor so that users can also use the editor for their own purposes like adding some important text, heading, or little note, and also use some tools like highlighting some text or some figures which users want to do. So basically it's a basic image editor which can be used after grabbing a screenshot from an extension. Our main aim was user will able to make notes of those images so user can then select the multiple images of any image format from local storage and then click on the "Make Notes" button so that the selected images will be sent backend for further task and responded pdf will be downloaded on user's local storage.

4.3.2 Back End Design

The backend of the system is a Flask API running on the server which is responsible for generating notes from an image and returning the notes in ppt/pdf/doc format. The API works in response to the call of browser extension. Receiving the images from sender extension API checks for supported image/file type then it performs OCR one the images. Finally, it generates the pdf/ppt file and it returns it to the browser extension.
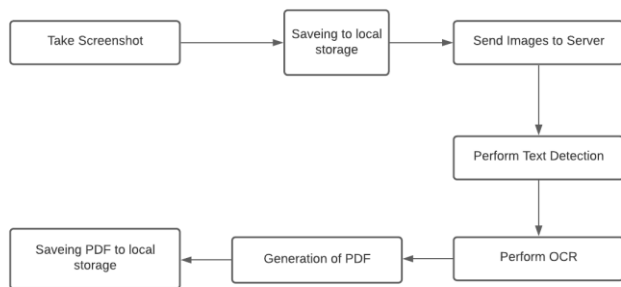
4.4 Implementation Setup

Figure. 2 sequence of activities

Figure 2 depicts the sequence of activities. The browser extension is on front end so the user can only interact with the browser extension. User is able to take a screenshot of full webpage or selected portion from the web page. Users can edit the screenshot after capturing it.

Then these pictures are sent to a server running the Flask API with the help of API call. The server contains a OCR engine which performs ocr on image and through this notes in pdf/ppt form are created. Server API sends generated notes as response to API call.
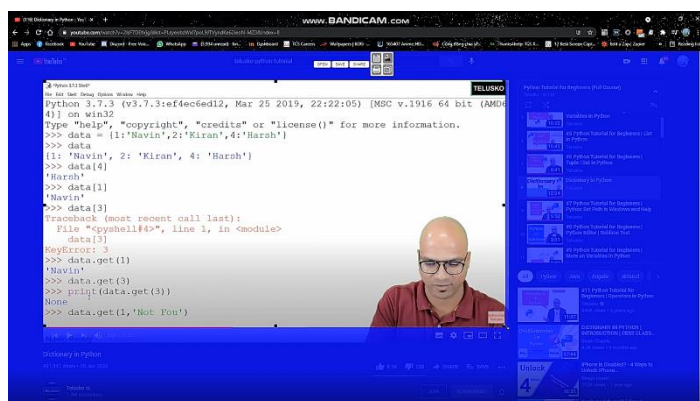
5. RESULTS

Figure 3: Screen Capture

Figure 3 depict the results when user takes the screenshot of live event i.e. online lectures, seminars, informative videos etc.
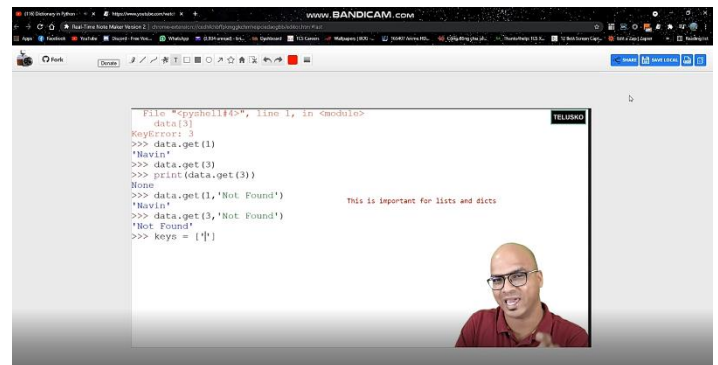

Figure 4: Real-Time Editor

Figure 4 shows that the tab which is opened when user wants to edit after taking screenshot.
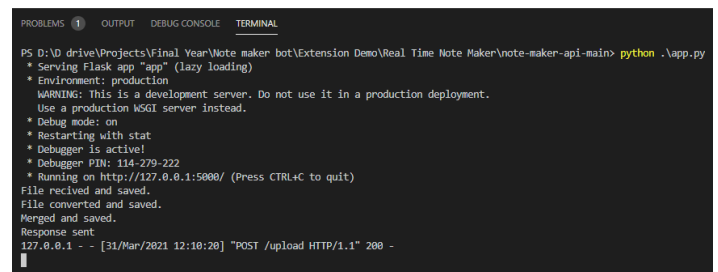

Figure 5: Server

Figure 5 shows the server when the browser extension makes API call and server generates the notes, then after it sends generated note in response.
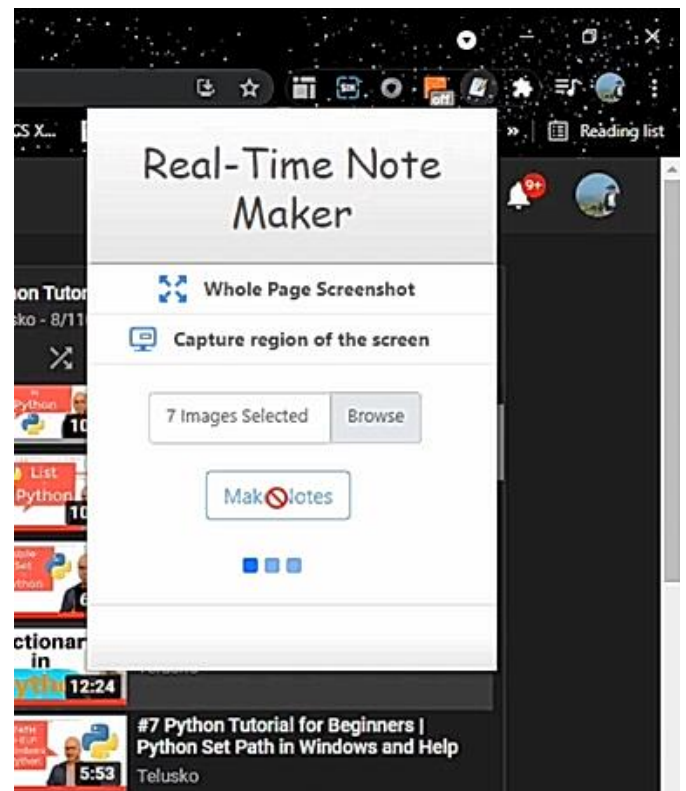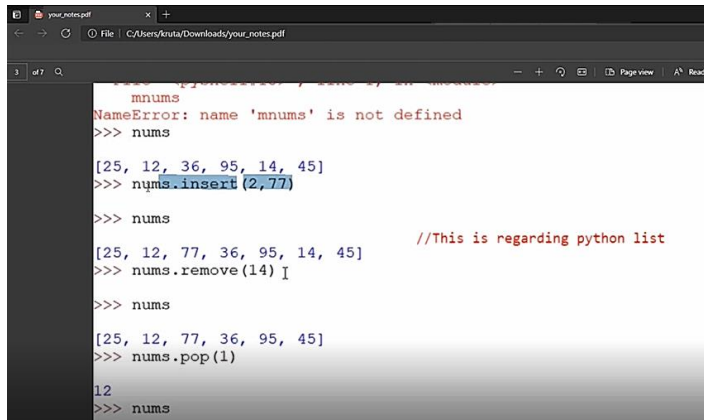

Figure 6: Downloading PDF

Figure 6 Shows the web browser extension while it is waiting for server response. After the response is received the notes are downloaded in user's local device.

Figure 7: Generated Notes

Figure 7 shows us the generated notes which is in pdf format. These notes allow user to copy and edit the text content of file.

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusion

This project has some tremendous power to revolutionize the Virtual Note Making process. The feasibility of this project mostly depends on the accuracy of the Text detection and OCR implementations. We would like to conclude by saying that in this paper we are bringing into use some of the technologies that are currently trending and are in high demand and will likely rise up in the future – from using concepts of AI and ML for OCR and Text Detection all the way to building User Interface for Web Browser Extension

6.2 Limitations

The Output may not work on Handwritten texts. It's not very accurate at the current state. Also the Text should not be too big or too small otherwise accuracy will be a bit low. This may not work on complex mathematical equations.

6.3 Future Work

Project will be able to work in hand written text. Real-time multiple screenshot capture by single button of a selected region every time. Resetting of extension during conversion will be prevented.

ACKNOWLEDGEMENT

REFERENCES

[1] Smith, R. (2007). An Overview of the Tesseract OCR Engine. Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) Vol 2. doi:10.1109/icdar.2007.4376991

[2] Mori, S., Nishida, H., & Yamada, H. (1999). Optical character recognition. New York: Wiley.

[3] Xing, L., Tian, Z., Huang, W., & Scott, M. (2019). Convolutional Character Networks. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). doi:10.1109/iccv.2019.00922

Du, Y., Li, C., Guo, R., Yin, X., Liu, W., Zhou, J., . . . Wang, H. (2020, October 15). PP-OCR: A Practical Ultra Lightweight OCR System. Retrieved from https://arxiv.org/abs/2009.09941v3

BIOGRAPHIES

KRUTARTH CHOVATIYA
Pursuing B.E. in Computer Science & Engineering at Institute of Technology & Management Universe, Vadodara, Gujarat, India

VATSAL KACCHIA
Pursuing B.E. in Computer Science & Engineering at Institute of Technology & Management Universe, Vadodara, Gujarat, India

MANI BUTWALL
Asst. Professor in Computer Science & Engineering at Institute of Technology & Management Universe, Vadodara, Gujarat, India.