### A

**MINI PROJECT REPORT ON**

#### “Form Details Reader”

**Submitted by Jashkumar Dhameliya (19IT430)**

#### Harpesh Patel (19IT435) Akshay Kathiriya (19IT440)

**(Prof. Priyank Bhojak) Faculty Guide**



#### Information Technology Department

**Birla Vishvakarma Mahavidyalaya Engineering College (An Autonomous Institution)**

### AY: 2021-22, Semester II

|  |  |
| --- | --- |
|  | **Birla Vishvakarma Mahavidyalaya Engineering College**  **(An Autonomous Institution)**  **Information Technology Department**  **AY: 2021-22, Semester II** |

**CERTIFICATE**

This is to certify that project entitled with **“Form Details Reader”** has been successfully carried out by **Jashkumar Dhameliya (19IT430), Harpesh Patel (19IT435), Akshay Kathiriya (19IT440)** or the subject of **3IT31-Mini Project** under my guidance during the academic year 2021-22, Semester II. The Mini Project work carried out by the students of 6thsemester is satisfactory.

**Date:**

|  |  |  |
| --- | --- | --- |
| **Prof. Priyank Bhojak**  Faculty Guide IT Department BVM | **Prof. Nilesh Prajapati**  Course Coordinator IT Department  BVM | **Prof. Keyur Brahmbhatt**  Head of the Department IT Department  BVM |

First of all, we are indebted to the GOD ALMIGHTY for giving us an opportunity to excel in our efforts to complete this project on time.

We are extremely grateful to Dr. Indrajit Patel, Principal, Birla Vishvakarma Mahavidyalaya Engineering College and Dr. Keyur Brahmbhatt, Head of the Information Technology Department, for providing all the required resources for the successful completion of our project. Our heartfelt gratitude to our project guide Prof. Priyank Bhojak, Assistant Professor in Information Technology, for her valuable suggestions and guidance in the preparation of the progress report and implementation of project.

We will be failing in duty if we do not acknowledge with grateful thanks to the authors of the references and other literatures referred to in this project.

**Chapter 1: Introduction 1**

* 1. Aim of the Project
  2. Project Scope
  3. Project Objective
  4. Project Modules
  5. Project basic requirements

**Chapter 2: Analysis, Design Methodology and Implementation Strategy 3**

* 1. Comparison of Existing Applications with your Project
  2. Project Feasibility Study
  3. Project Timeline chart
  4. Detailed Modules Description
  5. Project SRS
     1. Use Case Diagrams
     2. Data Flow Diagrams
     3. Entity Relationship Diagrams
     4. Event Trace Diagram
     5. State diagram
     6. Class diagram
  6. Literature Review
  7. Database design

**Chapter 3: Implementation and Testing**



* 1. Software and Tools
  2. User Interface and Snapshot
  3. Testing using Use Cases

**Chapter 4: Conclusion and Future work 29**

* 1. Conclusion
  2. Future work

**Reference**

* Links of all research papers

**Fig 2.1 6 Fig 2.2 8**

**Fig 2.3 8 Fig 2.4 9 Fig 2.5 9 Fig 2.6 10 Fig 2.7 11 Fig 2.8 12 Fig 2.9 13 Fig 3.1 19 Fig 3.2 19 Fig 3.3 20 Fig 3.4 20 Fig 3.5 21 Fig 3.6 22 Fig 3.7 22 Fig 3.8 23 Fig 3.9 23**

|  |  |
| --- | --- |
| **Fig 3.10** | **24** |
| **Fig 3.11** | **25** |
| **Fig 3.12** | **26** |
| **Fig 3.13** | **26** |

**Fig 3.14 27**

**Fig 3.15 27**

# LIST OF TABLES

**Table 1: Advantages and Limitations of research paper 16 Table 2: Database Table 17**

# ABSTRACT

The application “Form Details Reader” is developed to convert physical form data into digital form data using OCR. To use this application user may need photos of form in their machine which are under than 5 MB if not than user can also capture image via their mobile device. By using this application, we can add physical form details into database. Verify extracted data and update that. This data is saved into database and user can see that data and fetch data in specific time period, can delete unwanted data. The main part of our application is that user can download extracted data in excel file so that he can further use it via excel file of using database.

**Chapter 1: Introduction**

#### Aim of the Project

Form details reader extracts details or information from image or scanned form and store that details into respective format

#### Project Scope

The scope of our product Optical Character Recognition on a grid infrastructure is to provide an efficient and enhanced software tool for the users to perform Document Image Analysis, document processing by reading and recognizing that are having large pool of documented, scanned images. Irrespective of the size of documents and the type of characters in documents, the product is recognizing them, and store it in particular format.

#### Project Objective

Optical Character Recognition, is a process of recognizing text inside images and converting it into an electronic form. These images could be printed text like documents, receipts, name cards, etc.

#### Project Modules

* + 1. Browse Image
    2. Capture Image
    3. Download Excel Sheet

#### Project Basic Requirement

Hardware:

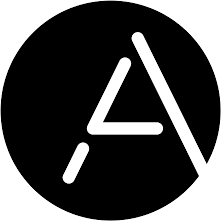
* Central Processing Unit (CPU) – Intel Core i3 8th Generation processor or higher. An AMD equivalent processor will also be optimal.
* RAM –8 GB minimum, 16 GB or higher is recommended.
* Graphics Card: 2GB or more Software:
* PyCharm IDE
* MS excel
* Operating System – Microsoft Windows

## Chapter 2: Analysis, Design Methodology and Implementation Strategy

#### Comparison of Existing Applications with your Project

**Anyline:**

Anyline makes data capture simple. By easily integrating our SDK into an app or website, our data capture solutions enable users to instantly scan and digitize analog data in the real world using any camera-enabled mobile device.



**Fig 1.1**

**Benefits:**

* + - Anyline’s mobile OCR solution for driver’s license scanning gathers all essential details in a couple of seconds.
    - From there it can provide you with instant scan results. This method of data collection has been shown to be up to 20 times faster than manual data entry.
    - This speed improvement will help you to optimize your processes, resulting in an improved user experience for your customers and employees.

**Modules:**

* + - Register
    - Dashboard
    - Scan Documents

**How does our product differ from existing products?**

Our Project extract data from given form and store in particular excel file. While anyline detects data from file and verify from its own database.

#### Project Feasibility Study

A feasibility study is a high-level capsule version of the entire System analysis and Design Process. The study begins by classifying the problem definition. Feasibility is to determine if it’s worth doing. Once an acceptance problem definition has been generated, the analyst develops a logical model of the system. A search for alternatives is analyzed carefully. There are 3 parts in feasibility study.

1. Economic Feasibility:

As there are very system available for document scan in the market and as it is solving many problems faced by people. The organization can easily generate financial benefits from advertisement displayed on UI. And also, those with low or no formal education –lack the necessary skills and know-how. Economic feasibility generates the costs implication of developing an AI-OCR conceptualized model.

1. Technical Feasibility:

The inevitable issues raised that whether to go ahead with the proposed project or not are addressed during Feasibility Stage are:

* + Does the necessary technology exist to do what is suggested?

Yes, necessary technology exists to implement discussed above.

* + Will the proposed system provide adequate response to inquiries, regardless of the number of users?

yes, as the system is no matter how much load is there on the system with the number of images increasing to use this system it will serve every image within a good amount of time.

* + Can the system be upgraded if developed?

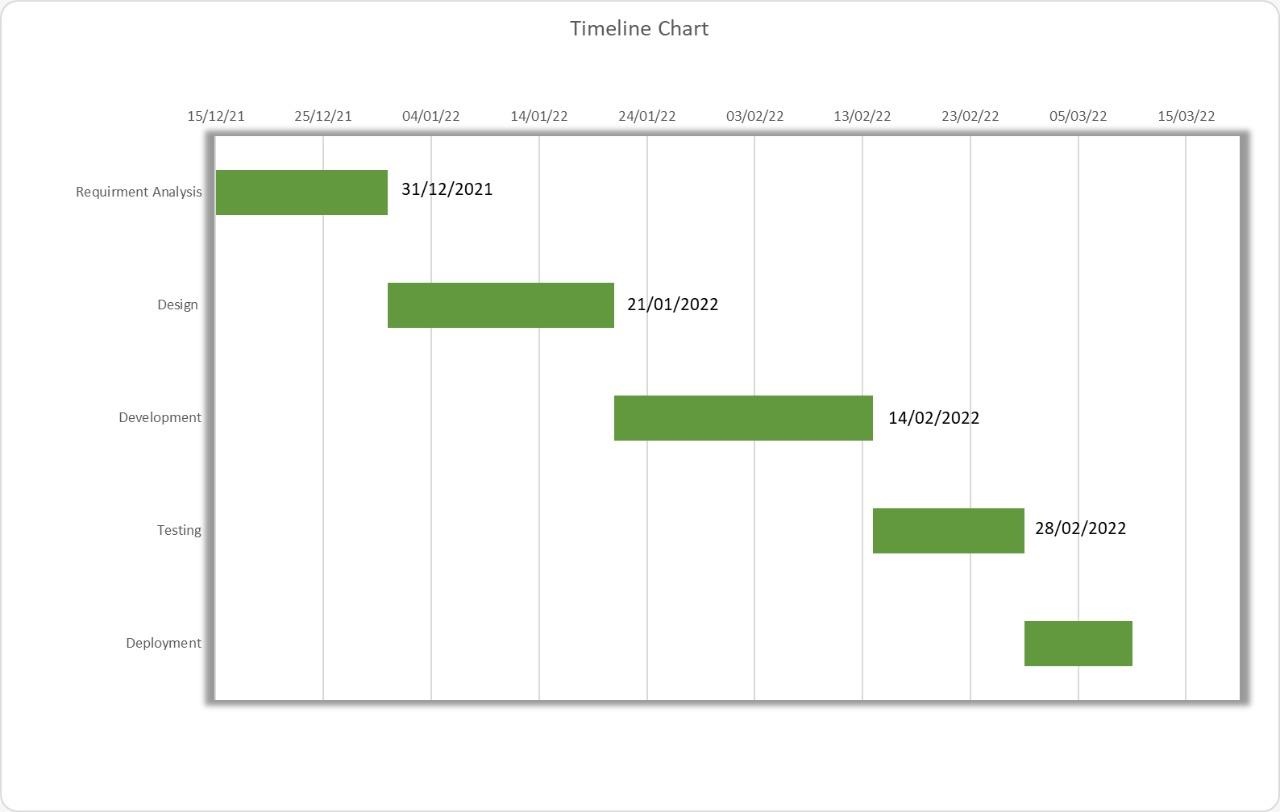
As the system is developed keeping scalability in mind it is developed with the well- structured and maintainable code it will be very easy to scale this system in the future

OCR (optical character recognition) is the use of technology to distinguish printed or handwritten text characters inside digital images of physical documents, such as a scanned paper document.

1. Schedule Feasibility:

By isolating the document template's entered text, an optical character recognition (OCR) engine can read and recognize much more accurately. pre-processing the image to remove noise also improves OCR performance.

#### Project Timeline chart



**Fig 2.1 Timeline Chart**

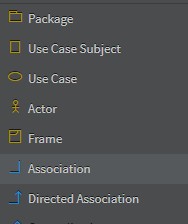
#### Detailed Modules Description

1. Browse Image
   * User will browse image and select it for data extraction.
   * After extraction a second screen will be shown and in that at left side of screen user will see tan image with extracted data, in the right side of screen user can see extracted data in form format in which he can further process.
2. Data Modification
   * User can modify extracted data.
3. Submit Data
   * If data is verified user can submit data and using back button user can brows another image.
4. Capture Image
   * User can capture image from smartphone using DroidCam application.
5. View Extracted data
   * User can view extracted data in table form.
   * In this window you can check data date vice or all data.
   * It has download button From Which you can download data in excel format.
   * Delete button to delete data.
   * Display button to display data between selected dates.
   * Back button to go to the main screen.

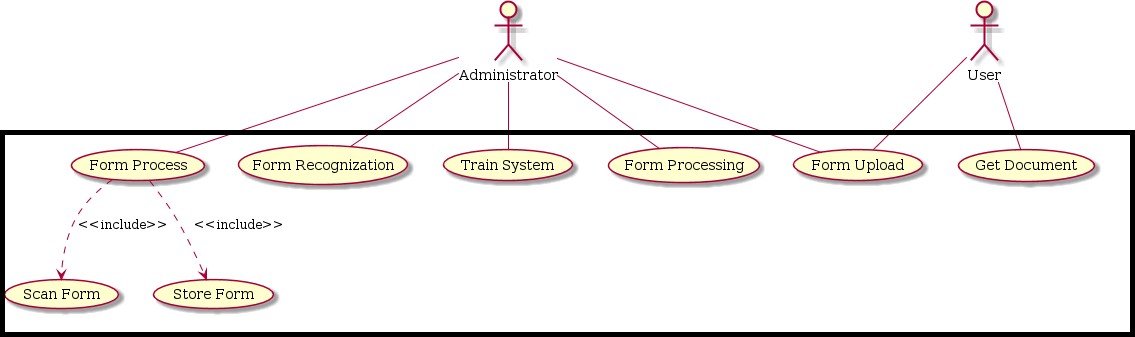
#### Project SRS

* + 1. **Use Case Diagrams**

A use case diagram is a graphical depiction of a user's possible interactions with a system. A use case diagram shows various use cases and different types of users the system has and will often be accompanied by other types of diagrams as well. The use cases are represented

by either circles or ellipses. The actors are often shown as stick figures.

**Fig 2.2 Symbols used in Use case diagram**

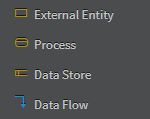


**Fig 2.3 Use case Diagram**

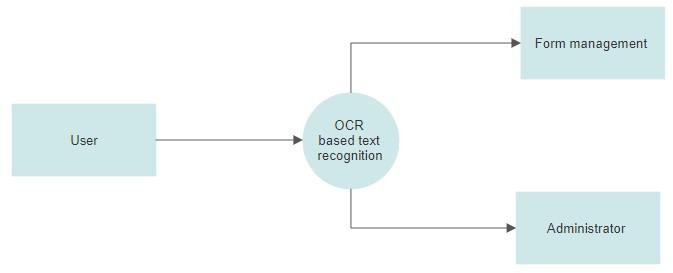
#### Data Flow Diagrams

A data-flow diagram is a way of representing a flow of data through a process or a system (usually an information system). The DFD also provides information about the outputs and inputs of each entity and the process itself. A data-flow diagram has no control flow — there are no decision rules and no loops. Specific operations based on the data can be represented by a flowchart.

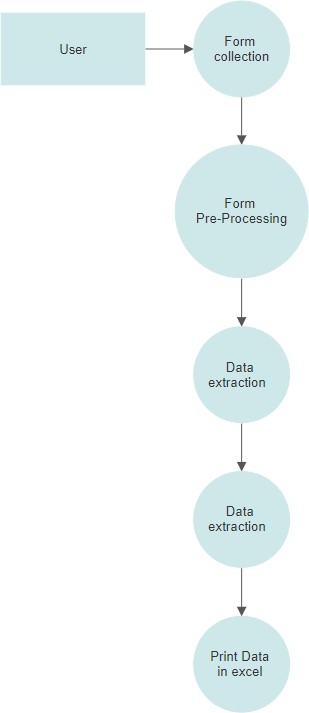
For each data flow, at least one of the endpoints (source and / or destination) must exist in a process. The refined representation of a process can be done in another data-flow diagram, which subdivides this process into sub-processes.



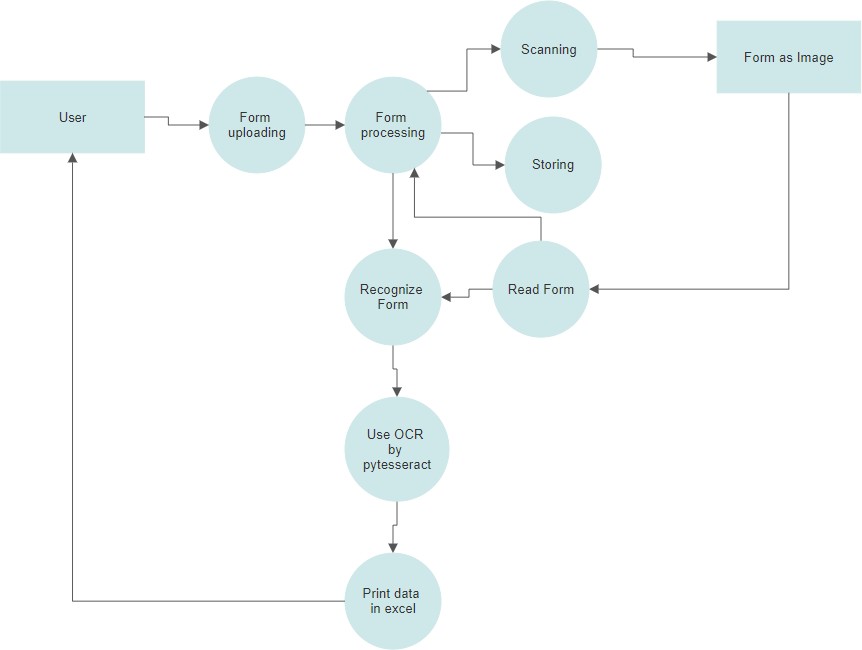
**Fig 2.4 Symbols used in Data Flow Diagram**



**Fig 2.5 DFD level-0**



**Fig 2.6 DFD level-1**



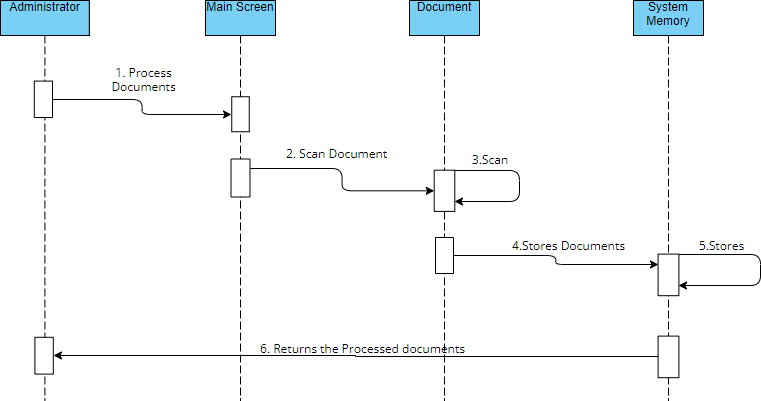
**Fig 2.7 DFD level-1**

#### 2.5.4 Event Trace Diagrams

A sequence diagram or system sequence diagram (SSD) shows object interactions arranged in time sequence in the field of software engineering. It depicts the objects involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of scenario. Sequence diagrams are typically associated with use case realizations in the logical view of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios.



**Fig 2.8 Symbols used in Event Trace Diagram**



**Fig 2.9 Event Trace Diagram**

#### Literature review

* + 1. **Text Extraction from Bills and Invoices**

**Publisher: IEEE**

**REVIEW: In this research paper author has said that**

This research tries to find out a methodology through which any data from the daily-use printed bills and invoices can be extracted. The data from these bills or invoices can be used extensively later on – such as machine learning or statistical analysis. This research focuses on extraction of final bill-amount, itinerary, date and similar data from bills and invoices as they encapsulate an ample amount of information about the users purchases, likes or dislikes etc. Optical Character Recognition (OCR) technology is a system that provides a full alphanumeric recognition of printed or handwritten characters from images. Initially, OpenCV has been used to detect the bill or invoice from the image and filter out the unnecessary noise from the image. Then intermediate image is passed for further processing using Tesseract OCR engine.

#### Text extraction from natural scene image: A survey

**Publisher: ELSEVIER**

**REVIEW: In this research paper author has said that**

With the increasing popularity of portable camera devices and embedded visual processing, text extraction from natural scene images has become a key problem that is deemed to change our everyday lives via novel applications such as augmented reality. Text extraction from natural scene images algorithms is generally composed of the following three stages: (i) detection and localization, (ii) text enhancement and segmentation and (iii) optical character recognition (OCR). The problem is challenging in nature due to variations in the font size and colour, text alignment, illumination change and reflections. This paper aims to classify and assess the latest algorithms.

#### OCR Using Computer Vision and Machine Learning

**Publisher: RESERCH GATE**

**REVIEW: In this research paper author has said that**

Being a well-researched area, optical character recognition or OCR has seen many advancements. Many state-of-the-art algorithms have been developed that can be used for the purpose of OCR but extracting text from images containing tables while preserving the structure of the table still remains a challenging task. Here, we present an efficient and highly scalable parallel architecture to segment input images containing tabular data with and without borders into cells and reconstruct the tabular data while preserving the tabular format. The performance improvement thus made can be used to ease the tedious task of digitizing tabular data in bulk. The same architecture can be used for regular OCR applications to improve performance if the data is in huge quantities.

#### Real time license plate detection using OpenCV and tesseract

**Publisher: Research gate**

**REVIEW: In this research paper author has said that**

This research paper which is related to license plate detection using OpenCV and tesseract is also base on machine learning and computer vision. In this research paper author has talked about the implementation of image to text conversion. The paper describes various steps required to extract text from any image file (jpeg/png) and create a separate text file consisting of information extracted from image file. The CV2 OpenCV library using Python language is used for image processing and Tesseract is used for text extraction from the processed image. The variable level of image processing ensures that different images get different levels of treatment in order to produce optimized text results. After the image processing step is employed the output text file are formatted by filtering out commas, semicolons, apostrophes, colons and other such characters using ASCII filtering.

#### A survey on various approaches of text extraction in images

**PUBLISHER: academia.edu**

**REVIEW: In this research paper author has said that**

There are many applications of a text extraction such as Keyword based image search, text- based image indexing and retrieval, document analysis, vehicle license detection and recognition, page segmentation etc. A number of methods have been proposed in the past

for extraction of text in images. These approaches considered the different attributes related to text in an image such as of size, font, style, orientation, alignment, contrast, color, intensity, connected-components, edges etc. These attributes are used to classify text regions from their background or other regions within the image. This paper provides a broad study of the various text extraction techniques and algorithms proposed earlier. Even though there are many numbers of algorithms, there is no single unified approach that fits for all the applications.

#### Advantages and Disadvantages of research paper

|  |  |
| --- | --- |
| Advantages | Limitations |
| Information of OCR can be readable with high degree of accuracy. Flatbed scanners are very accurate and may produce reasonably top quality images. | There is the need of lot of space required by the image produced. |
| The latest software can re-create tables also as original layout. | All the documents got to be checked over carefully then manually corrected. |
| This process is much faster as compared to the manual typing the information into the system | Handwriting must be learnt by the pc. |
| It is cheaper than paying someone amount to manually enter great deal of text data. | Quality of the ultimate image depends on quality of the first image. |

**Table 1: Advantages and Limitations of research paper**

**How does our application tend to solve some limitations imposed by revived research papers?**

* + - 1. The application which we are aiming to develop will be Portable so it can be made available on every platform.
      2. Users can generate excel sheet from form hence data will be stored in computers.
      3. Our project stores image in temp section for data extraction so picture quality can’t reduce.

#### Database Design

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Name** | **Datatype** | **Field Length** | **Constraint** |
| **1** | ID | Int(11) | 11 | Primary Key |
| **2** | Name | Varchar(100) | 100 | Not null |
| **3** | ID\_No | Varchar(10) | 10 | Not null |
| **4** | Mobile | Double |  | Not null |
| **5** | Branch | Varchar(225) | 225 | Not null |
| **6** | Admission Year | Varchar(11) | 11 | Not null |
| **7** | Email | Varchar(100) | 100 | Not null |
| **8** | AddressLine1 | Varchar(200) | 200 | Not null |
| **9** | AddressLine2 | Varchar(200) | 200 | Not null |
| **10** | Declaration | Int(1) | 1 | Not null |
| **11** | Entdate | Date |  | Not null |

**Table 2: Database Table**

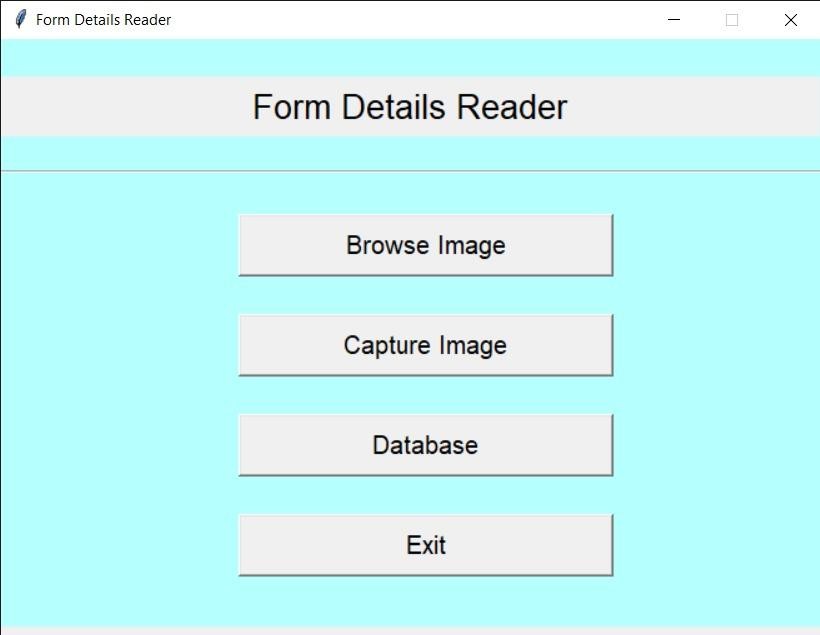
## Chapter 3: Implementation and Testing

#### Software and Tools

* + 1. **Hardware Requirements**
       1. Central Processing Unit (CPU) – Intel Core i3 8th Generation processor or higher. An AMD equivalent processor will also be optimal.
       2. RAM –8 GB minimum, 16 GB or higher is recommended.
       3. Graphics Card: 2GB or more
    2. **Software Requirement**
       1. PyCharm IDE
       2. MS excel
       3. Operating System – Microsoft Windows 10.

#### User Interface and Snapshot

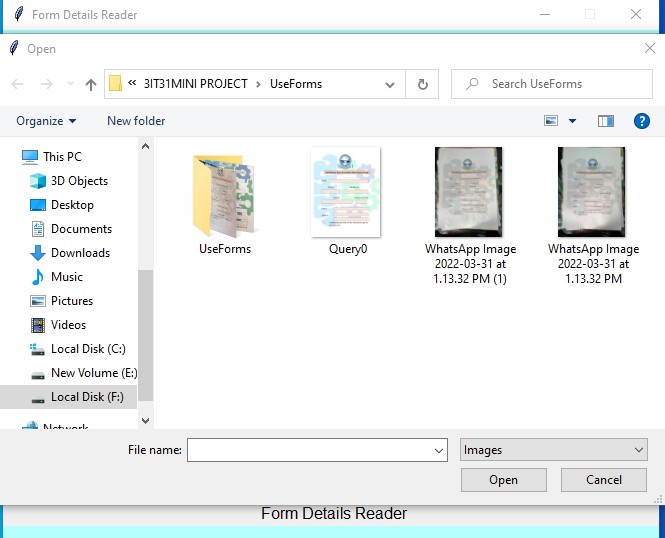
* + 1. **Main window**



**Fig 3.1 Main Window**

* This window will appear when we start program.

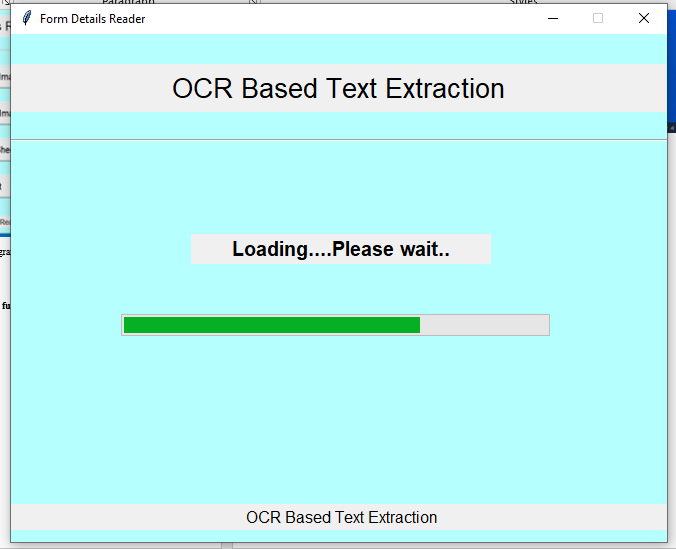
#### Browse Image



**Fig 3.2 Browse Image**

* + User Have to Select Image.

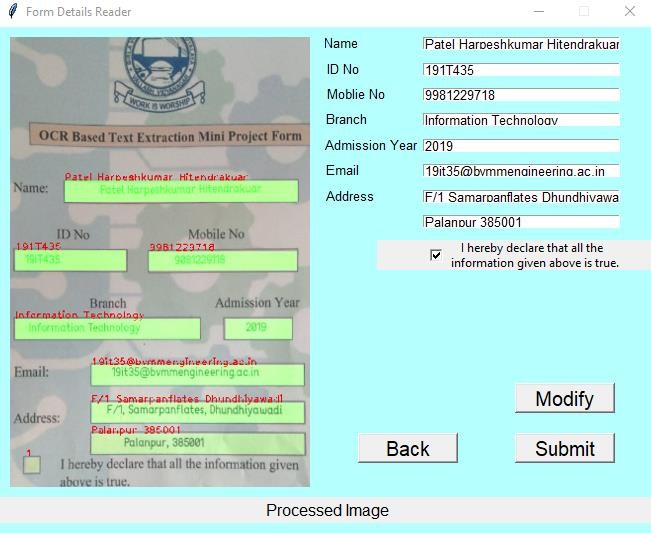
#### Loading Screen



**Fig 3.3 Loading screen**

* Data reading process will show on progress bar.

#### Data will show to user for further operation



**Fig 3.4 Extracted data**

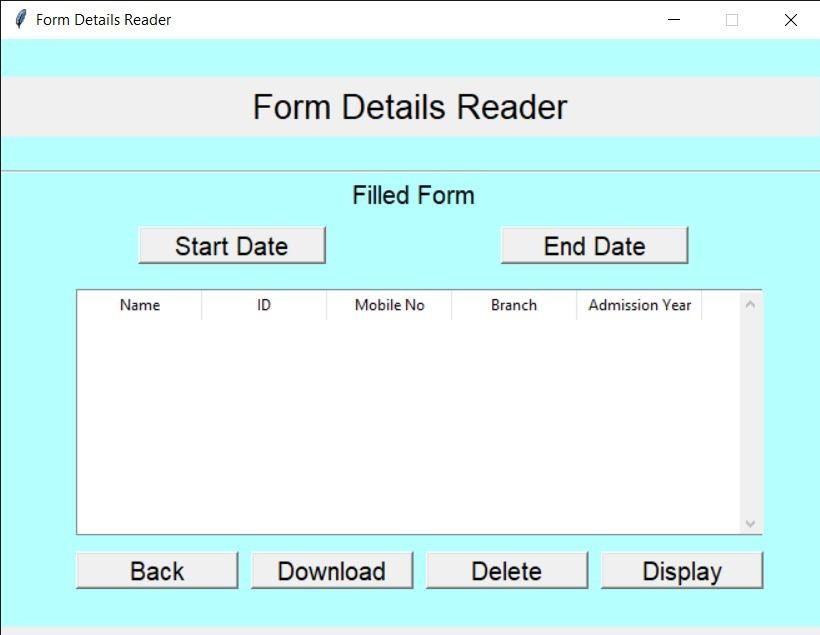
#### Capture Image



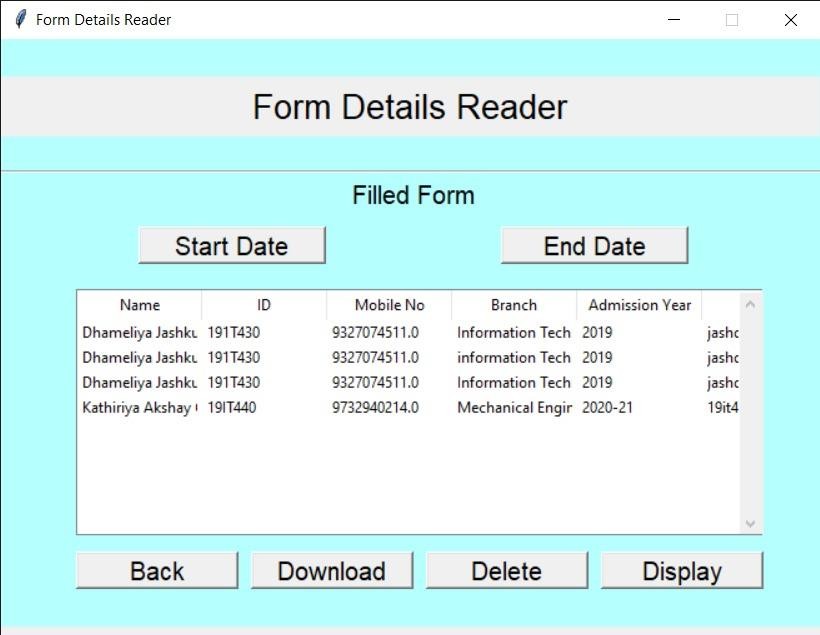
**Fig 3.5 Capture Image**

* image from mobile phone.

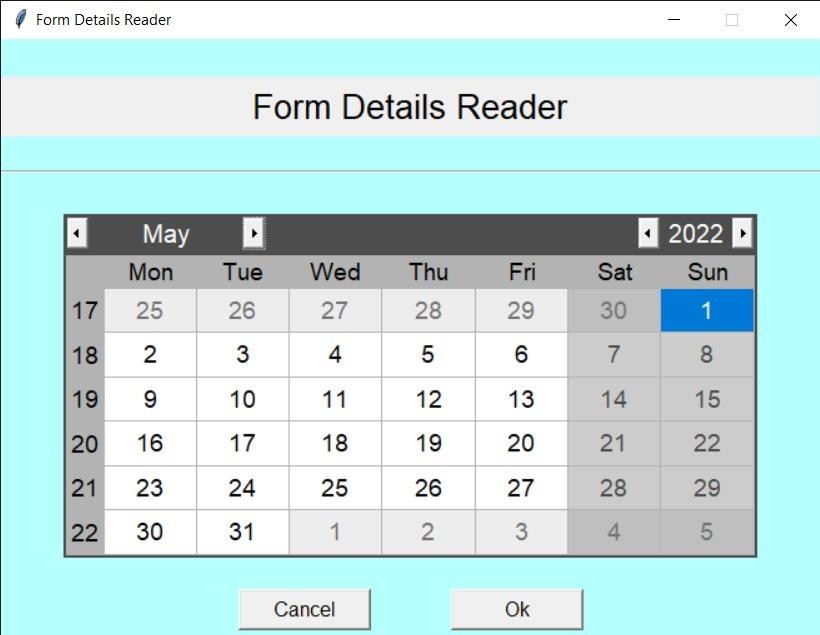
#### Data Base

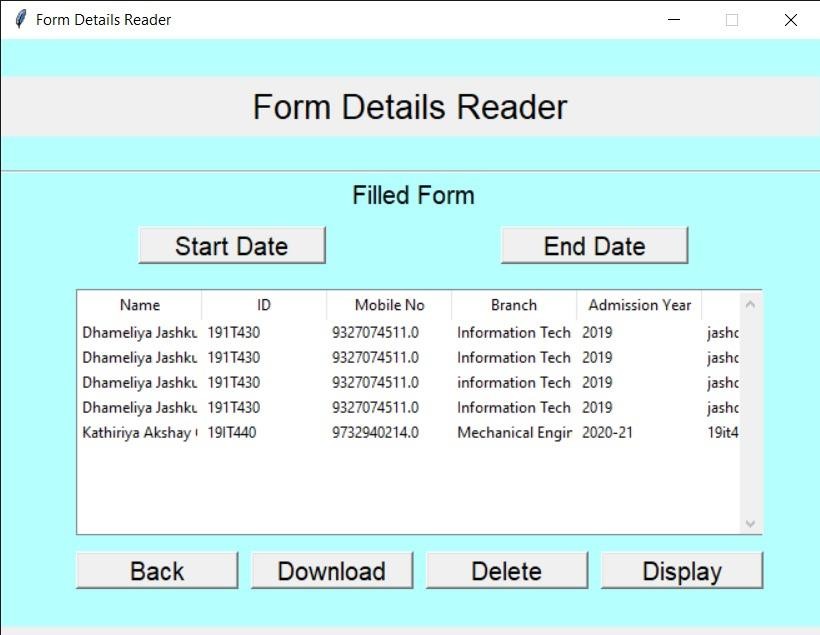


**Fig 3.6**

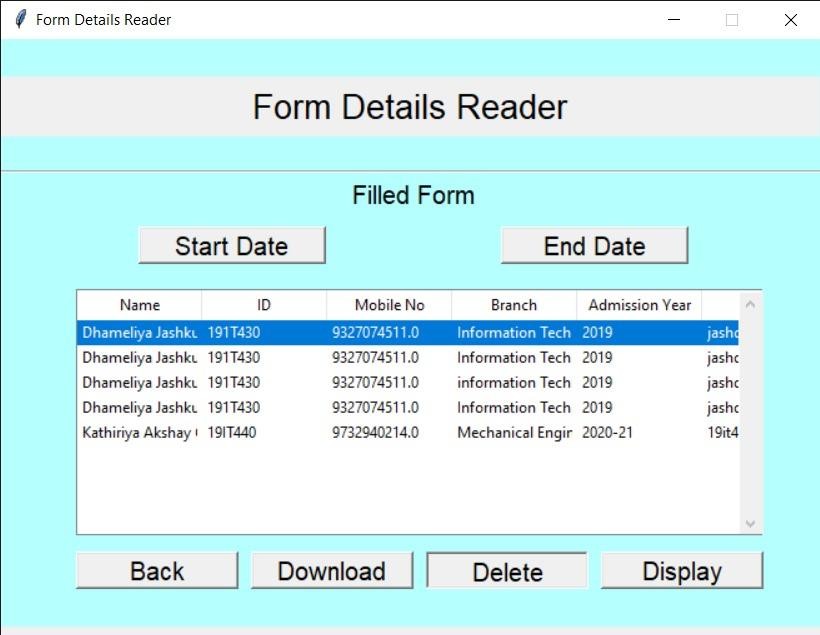


**Fig 3.7**



**Fig 3.8**

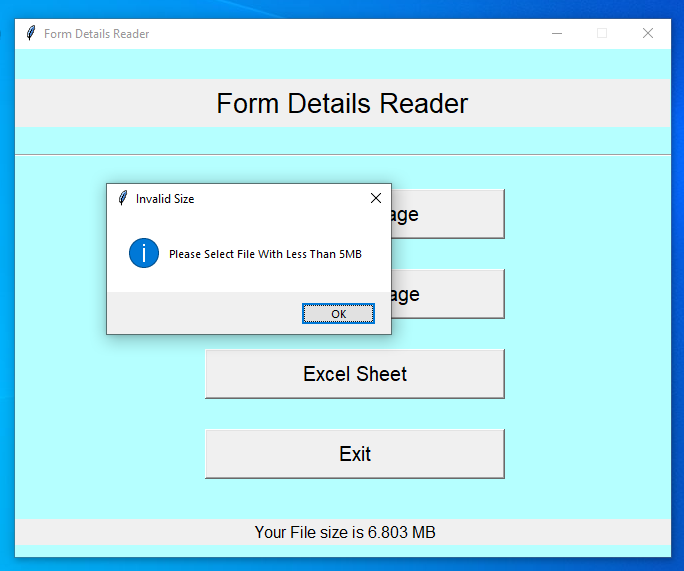
**Fig 3.9**



**Fig 3.10**

#### Testing using Use Cases

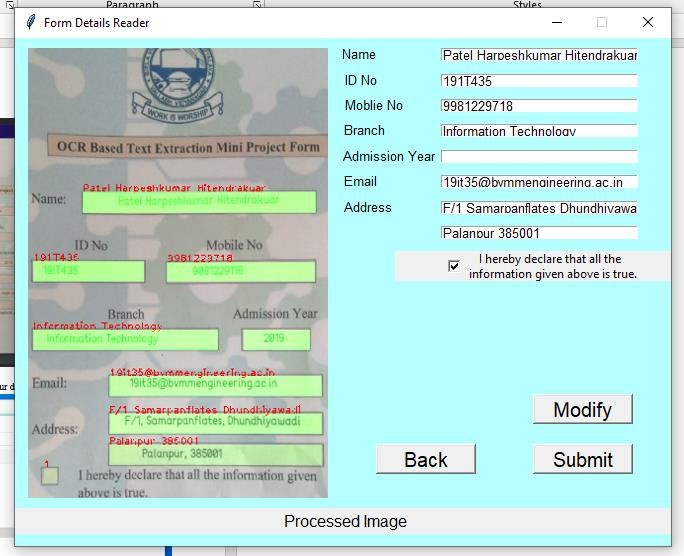
* + 1. **If image size is more than 5MB.**



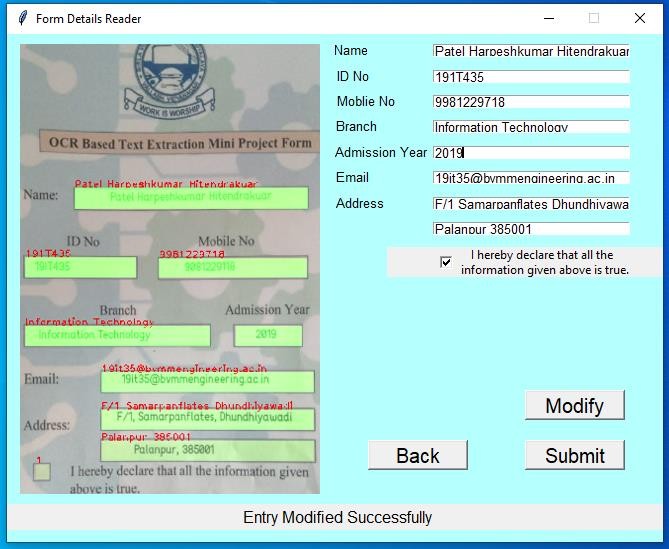
**Fig 3.11**

* + 1. **User Modification**

When Data is not correct then user can modify the data using given Modify function.

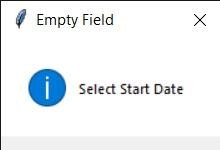


**Fig 3.12**

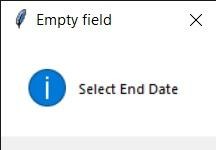


**Fig 3.13**

#### If database date is not selected



**Fig 3.14**



**Fig 3.15**

**Chapter 4: Conclusion and Future work**

#### Conclusion

* + - In this project, machine learning algorithm has been implemented to extract particular data from the pre-defined form. In this implementation OpenCV, py tesseract library are used as a main library. System has performed step by step procedure to read the data and found the character between the parameters. In this project with the use of nural network the given image is equalized with the reference image and the text is detected with some amount of dots projection on given image. After that the extracted data is displayed to the user for further data modification by the user. By user confirmation the data will be added to mySQL database. User can fetch data from database by selecting starting and ending date.

#### Future work

* + - In this application we can further increase its accuracy.
    - Implement new module so that we can add different forms and select and use them.
    - For image capturing only already defined user can capture using its added device.
    - UI can be more attractive.
    - The text extraction speed can be increase by some code optimization.

## References

* [**https://www.researchgate.net/publication/357990718\_Optical\_Character\_Re**](https://www.researchgate.net/publication/357990718_Optical_Character_Recognition_OCR_based_Vehicle%27s_License_Plate_Recognition_System_Using_Python_and_OpenCV)[**cognition\_OCR\_based\_Vehicle's\_License\_Plate\_Recognition\_System\_Using\_Pyth**](https://www.researchgate.net/publication/357990718_Optical_Character_Recognition_OCR_based_Vehicle%27s_License_Plate_Recognition_System_Using_Python_and_OpenCV)[**on\_and\_OpenCV**](https://www.researchgate.net/publication/357990718_Optical_Character_Recognition_OCR_based_Vehicle%27s_License_Plate_Recognition_System_Using_Python_and_OpenCV)
* [**https://www.researchgate.net/publication/334167173\_Text\_Extraction\_from\_Bills**](https://www.researchgate.net/publication/334167173_Text_Extraction_from_Bills_and_Invoices)

[**\_and\_Invoices**](https://www.researchgate.net/publication/334167173_Text_Extraction_from_Bills_and_Invoices)

* [**https://stackabuse.com/automating-processing-pdf-invoices-in-python-with-borb/**](https://stackabuse.com/automating-processing-pdf-invoices-in-python-with-borb/)
* [**https://www.google.com/search?q=text+extraction+from+bills+and+invoices+usin**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)[**g+python&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)[**031&ei=g-**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)[**h4YrjKAcqUseMPu8ep0Aw&oq=text+extraction+from+bills+and+invoices+usin**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)[**g+py&gs\_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAx**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)[**AnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKAB**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)[**SgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTm**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)[**YAQCgAQHIAQTAAQE&sclient=gws-wiz**](https://www.google.com/search?q=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpython&sxsrf=ALiCzsaGCjyslHF2GiX3AwNHV3B8xskJWQ%3A1652091011031&ei=g-h4YrjKAcqUseMPu8ep0Aw&oq=text%2Bextraction%2Bfrom%2Bbills%2Band%2Binvoices%2Busing%2Bpy&gs_lcp=Cgdnd3Mtd2l6EAMYADIFCCEQoAEyBQghEKABOgcIIxCwAxAnOgcIABBHELADOgUIABCABDoGCAAQFhAeOgQIIRAVOgcIIRAKEKABSgQIQRgASgQIRhgAUJIDWNkdYLAqaAFwAXgAgAGiAogBvBCSAQMyLTmYAQCgAQHIAQTAAQE&sclient=gws-wiz)
* [**https://nanonets.com/blog/handwritten-character-recognition/**](https://nanonets.com/blog/handwritten-character-recognition/)
* [**https://projectgurukul.org/python-text-detection-extraction-opencv-ocr/**](https://projectgurukul.org/python-text-detection-extraction-opencv-ocr/)