الكلية الجامعية للعلوم التطبيقية
University College of Applied Sciences
Leader of Innovation رائـــــدة الإبـــــداع

University College of Applied Sciences

(Data Science and Artificial Intelligence)

Digital Image processing - final project

Questionnaire Analysis

Table of contents	

➢ Introduction :

In an increasingly digital world, simplifying data collection is essential. This project introduces an innovative **Questionnaire Analysis System**, designed to revolutionize the way surveys are processed. Users can easily **design and print questionnaires**, then manually fill them out before scanning and uploading them into the system.

The software automatically analyzes the scanned forms, detecting **marked answers** with precision. It generates the question IDs and corresponding responses without the need for manual intervention, making the entire process faster and more efficient.

By incorporating **advanced image pre-processing**, the system ensures that each scanned document is properly aligned and of optimal quality, allowing for accurate answer detection. This **automation** not only accelerates the data collection process but also minimizes errors, empowering users to focus on the insights rather than the process itself.

➢ Methodology :

describes the step-by-step approach used to develop the **Questionnaire Analysis System**, which automates the process of detecting and extracting answers from scanned questionnaires :

1. Designing the Questionnaire and bubble sheet ::

- A questionnaire template was created with multiple-choice questions.


Student Name : Grade:

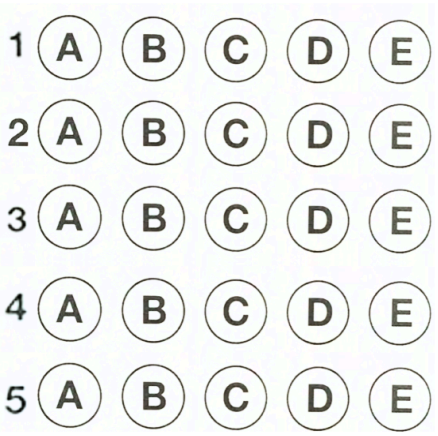
1. What is the capital of Palestine ?
   A. Gaza City
   B. Ramallah
   C. Jerusalem
   D. Nablus
   E. Bethlehem

2. Which sea borders Palestine to the west ?
   A. Red Sea
   B. Arabian Sea
   C. Mediterranean Sea
   D. Dead Sea
   E. Caspian Sea

3. What is the official currency of Palestine ?
   A. Jordanian Dinar
   B. Israeli Shekel
   C. Palestinian Pound
   D. Egyptian Pound
   E. US Dollar

4. What is the name of the holiest mosque in Palestine ?
   A. Al-Masjid an-Nabawi
   B. Al-Aqsa Mosque
   C. Sultan Ahmed Mosque
   D. Al-Umari Mosque
   E. Grand Mosque of Gaza

5. Which Palestinian territory is located on the coast ?
   A. West Bank
   B. Gaza Strip
   C. Nablus
   D. Hebron
   E. Jenin



- The user filled out the printed forms manually, marking their answers with a pen.

2. Scanning and Preprocessing the Image

- Once the form was filled, it was scanned and saved as an image on the computer.
- Extracting and Detecting Questions : Using **OCR (Optical Character Recognition)**, the text from the image was extracted (**using** `pytesseract.image_to_string(image)`)

```
4. What is the capital of Palestine ?
Gaza City

Ramallah

Jerusalem

Nablus

Bethlehem

moomp>

2. Which sea borders Palestine to the west ?

A. Red Sea
B. Arabian Sea
C. Mediterranean Sea
D. Dead Sea
E. Caspian Sea
3. What is the official currency of Palestine ?

A. Jordanian Dinar
B. Israeli Shekel

C. Palestinian Pound
D. Egyptian Pound
E. US Dollar

‘4. What is the name of the holiest mosque in Palestine ?

A. AbMasid an-Nabawi
B. ALAgsa Mosque
```

then using regular expressions to extract the question and store it in a table for display the final results

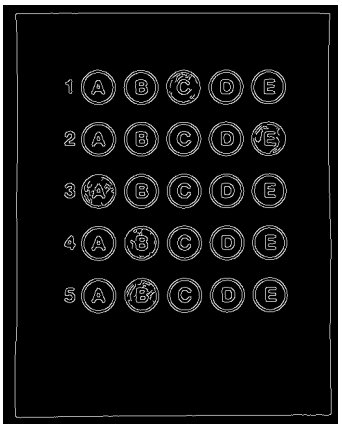
```
[5]  # Regular expression pattern to capture the question
     pattern = r'\b\d+[.,]?\s*([^\?]+?\?)'

     # Find and extract all matching questions
     questions = re.findall(pattern, text)

     for idx, question in enumerate(questions, 1):
         print(f"{idx}. {question}")
```

```
1. What is the capital of Palestine ?
2. Which sea borders Palestine to the west ?
3. What is the official currency of Palestine ?
4. What is the name of the holiest mosque in Palestine ?
5. Which Palestinian territory is located on the coast ?
```

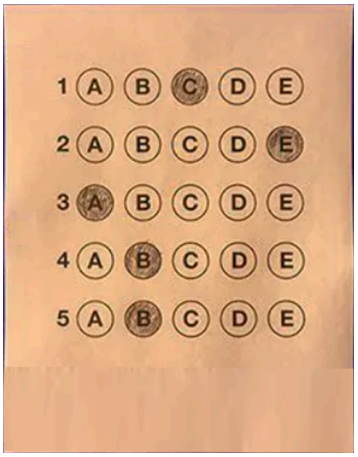
- The image underwent preprocessing to ensure optimal quality:
  - **Grayscale Conversion:** To simplify the image and make processing easier.

- **Blurring:** Applied to reduce noise and smooth the image.

- **Edge Detection:** Used to highlight the marked answers clearly.

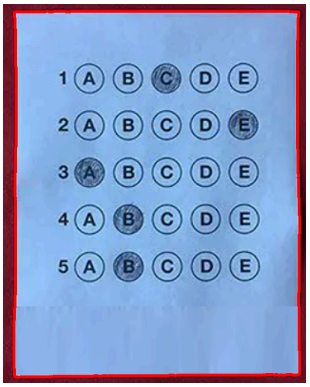


- **Orientation Correction**: The image was adjusted to fix any tilt or skew.

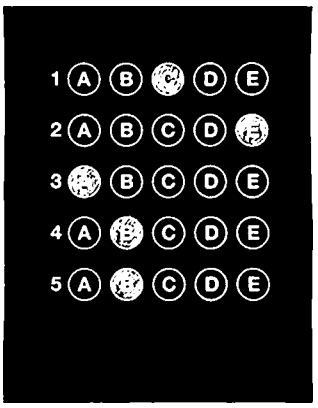


3. Extracting and Detecting User answers :

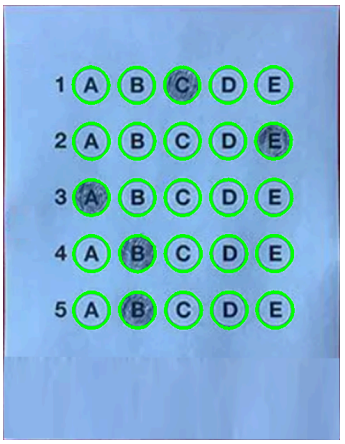
- **Contours Detected :** can adjust the size and aspect ratio conditions**.**



- **Threshold Image :** make the contours more distinct and accurate.



- **Question Contours :** make sure the question contours are properly sorted top-to-bottom, especially when the image contains noise or unwanted elements.



- A **pattern recognition** method was applied to identify and extract the questions automatically from the text

6. Displaying Results

- The answers, along with their correctness, were displayed in an organized table, making it easy for the user to review their responses.
- A **recommendation** message was shown based on the score, offering feedback like "Great job!" or "Needs improvement."

7. Testing and Validation

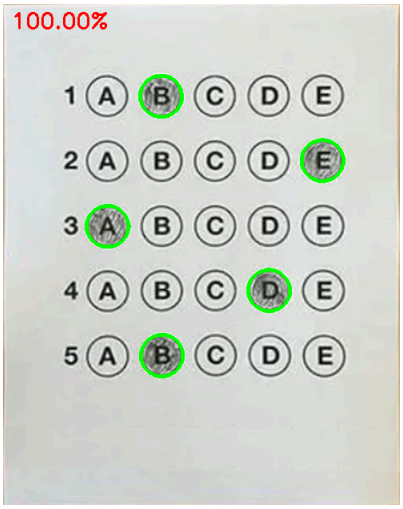
- The system was tested on various scanned questionnaires to ensure that the answer detection and evaluation were accurate.

➢ Results:

The system accurately detects marked answers, automates grading, and provides visual feedback. It ensures efficient, fast, and reliable performance for large-scale questionnaire evaluations.

1- Test case 1:

The system accurately detected all marked answers, resulting in a perfect score for the user.
output :



Question Number	Question Text	User Answers	Correction
1	What is the capital of Palestine ?	B	True
2	Which sea borders Palestine to the west ?	E	True
3	What is the official currency of Palestine ?	A	True
4	What is the name of the holiest mosque in Palestine ?	D	True
5	Which Palestinian territory is located on the coast ?	B	True

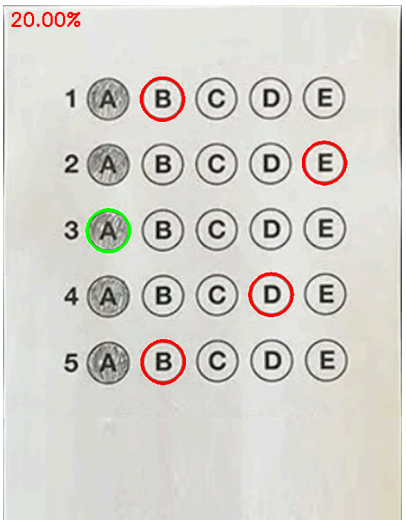
Grade:	100

Note : **Excellent! Keep up the good work!**

2- Test case 2:

The system detected a few marked answers correctly, resulting in a score of 20%.

output :



Question Number	Question Text	User Answers	Correction
1	What is the capital of Palestine ?	A	False
2	Which sea borders Palestine to the west ?	A	False
3	What is the official currency of Palestine ?	A	True
4	What is the name of the holiest mosque in Palestine ?	A	False
5	Which Palestinian territory is located on the coast ?	A	False

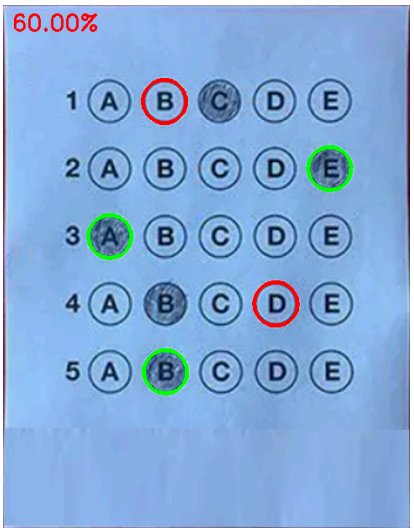
Grade:	20

Note : **Needs improvement. Consider seeking help or focus more on your studies.**

3- Test case 3:

The system detected most of the marked answers correctly, resulting in a score of 60%.

output :



Question Number	Question Text	User Answers	Correction
1	What is the capital of Palestine ?	C	False
2	Which sea borders Palestine to the west ?	E	True
3	What is the official currency of Palestine ?	A	True
4	What is the name of the holiest mosque in Palestine ?	B	False
5	Which Palestinian territory is located on the coast ?	B	True

Grade:	60

Note : **Good job! Keep practicing to improve your score.**

➢ Comments:

- **Pre-processing**: Enhances and corrects scanned images for better answer detection.
- **Contour Detection**: Identifies marked answers by detecting contours.
- **Answer Extraction**: Compares extracted answers with a predefined key.
- **Output**: Visualizes correct/incorrect answers for feedback.
- **Efficiency**: Automates grading and improves accuracy.
- **Future Work**: Improve error handling and support more layouts.

➢ Conclusion :

The **Questionnaire Analysis System** efficiently automates the detection, extraction, and evaluation of answers from scanned questionnaire forms. Using image processing techniques like edge detection and OCR, it accurately identifies marked responses, compares them to an answer key, and calculates scores.

The system pre-processes scanned images for quality enhancement, detects answers through contour analysis, and outputs the results in an organized table with feedback. This reduces manual data entry and improves accuracy, making it a valuable tool for educational assessments and surveys.

In conclusion, this solution streamlines the questionnaire analysis process, providing a fast and reliable method for data collection and evaluation.

➢ References:

- Python Code for OMR Answer Sheet Evaluation Using Deep Learning
- optical mark recognition
- Bubble sheet multiple choice scanner and test grader using OMR, Python, and OpenCV
- Use pytesseract OCR to recognize text from an image
- Tables — Rich 13.6.0 documentation

----------Thank you !----------