Cairo University Image Processing and

Faculty of Engineering Introduction to Computer Vision

Computer Engineering Department CMP3020

Phase 2 Report

Grade Auto Filler

Team #7

**The Used Algorithms:**

* Module 1 : Grade Sheet
  1. Image Preprocessing Techniques Like:
     + Morphological Operations (Erosion , Dilation)
     + Thresholding
     + Inversion and Grayscale Conversion
  2. Contour Detection:
     + Find Contours
     + Bounding Box Calculation
  3. OCR
  4. Feature Extraction
     + Training and Extract Features By HOG
  5. Excel File Generation
     + Pandas: Used for creating and manipulating data frames.
     + OpenPyXL: Used for writing data to Excel files and applying styles.
* Module 2 : Bubble Sheet

1. Image Preprocessing Techniques Like:
   * + Morphological Operations (Erosion, Dilation)
     + Thresholding
     + Inversion and Grayscale Conversion
   1. Contour Detection:
      * Find Contours
      * Bounding Box Calculation
   2. OCR
   3. Feature Extraction
      * Training and Extract Features By HOG
   4. Excel File Generation
      * Pandas: Used for creating and manipulating data frames.
      * OpenPyXL: Used for writing data to Excel files and applying styles.

**Test Cases Description:**

Images captured from different angles, distances, and with varied background colors present a range of diverse conditions for text extraction. Each image might have unique perspectives, offering a variety of alignment and orientation, which adds flexibility to handle different viewing situations. The distance from which the images are captured can provide different levels of detail, allowing for both close-ups and wider views. Background colors may vary, creating unique contrasts that could highlight the text differently, making it suitable for various environments and lighting conditions. This diversity ensures adaptability to real-world scenarios where documents may not always be uniform.

| Image | Accuracy | Recall | Precision |
| --- | --- | --- | --- |
| Image 1 | 0.627450980392157 | 0.790123456790123 | 0.752941176470588 |
| Image 2 | 0.656862745098039 | 0.933333333333333 | 0.571428571428571 |
| Image 3 | 0.57843137254902 | 0.9428571428571431 | 0.680412371134021 |
| Image 4 | 0.745098039215686 | 0.987012987012987 | 0.57 |
| Image 5 | 0.53921568627451 | 0.555555555555556 | 0.752475247524752 |
| Image 6 | 0.725490196078431 | 0.973684210526316 | 0.948275862068966 |
| Image 7 | 0.627450980392157 | 0.790123456790123 | 0.74 |
| Image 8 | 0.549019607843137 | 0.933333333333333 | 0.752941176470588 |
| Image 9 | 0.549019607843137 | 0.931034482758621 | 0.571428571428571 |
| Image 10 | 0.57843137254902 | 1 | 0.5625057 |
| Image 11 | 0.745098039215686 | 0.987012987012987 | 0.752475247524752 |
| Image 12 | 0.5 | 0.961538461538462 | 0.510204081632653 |
| Image 13 | 0.666666666666667 | 0.942028985507246 | 0.68421052631579 |
| Image 14 | 0.647058823529412 | 0.984375 | 0.642857142857143 |
| Image 15 | 0.7528823529412 | 0.914134226201082 | 0.53027142857143 |
| Average | 0.618954248366013 | 0.914134226201082 | 0.914134226201082 |

**Conclusion**

The project successfully automates the grading process for both handwritten grade sheets and MCQ bubble sheets. It leverages advanced image processing techniques and machine learning models to ensure accuracy and efficiency. The integration with Gradio provides a user-friendly interface for uploading images and receiving processed results.

**Proposed Enhancements**

1. AI-Powered Handwriting Recognition: Incorporate deep learning models for more accurate handwriting recognition.
2. LMS Integration: Integrate with Learning Management Systems (LMS) for seamless grade submission.
3. Real-Time Processing: Enhance processing speed to handle large datasets in real-time.
4. Enhanced Error Handling: Improve error handling and logging for better debugging and user feedback.

**Cons of the Used Methodology**

1. Dependency on Image Quality: The accuracy of the OCR and image processing algorithms heavily depends on the quality of the input images.
2. Limited Handwriting Styles: The custom classifiers may not generalize well to all handwriting styles.
3. Processing Time: The current implementation may be slow for large datasets or high-resolution images.
4. Complex Setup: The setup and dependencies might be complex for non-technical users.

**Work Division**

| Name | Work Load |
| --- | --- |
| Ahmed Mostafa | The Block of Processing The Grade Sheet and Extracting The Data of Each Column at the Grades table and Exporting the Output of it into Excel |
| Omar Ibrahim | Bubble sheet correction module, digits model, main |
| Mazen Adel | Paper extraction module, code extraction module |
| Ahmed Kamal | Making the GUI and Preprocessing The Grade Sheet paper and Make the Perspective Correction of Grade Sheet |

## **Data:**

* + (module2) Our dataset exhibits high diversity, with IDs ranging from 2 to 7 digits. Each entry contains between 12 and 45 questions, with 2 to 4 answer choices per question. The questions are distributed across multiple columns. The images are captured at various angles, distances, and orientations relative to the paper. Additionally, the background color varies across the images

## **Experiment Results and Analysis**

Results and Accuracy (module2)

| Test case number | Student answers accuracy | Student ID bubble accuracy |
| --- | --- | --- |
| 1 | 16/16 | 4/4 |
| 2 | 16/16 | 4/4 |
| 3 | 11/13 | 2/2 |
| 4 | 13/13 | 2/2 |
| 5 | 13/13 | 2/2 |
| 6 | 16/16 | 4/4 |
| 7 | 16/16 | 4/4 |
| 8 | 13/13 | 1/2 |
| 9 | 12/13 | 2/2 |
| 10 | 12/16 | 3/4 |
| 11 | 13/13 | 2/2 |
| 12 | 16/16 | 4/4 |
| 13 | 16/16 | 4/4 |
| 14 | 16/16 | 3/4 |
| 15 | 16/16 | 2/4 |
| 16 | 13/13 | 2/2 |
| 17 | 13/13 | 2/2 |
| 18 | 13/13 | 2/2 |
| 19 | 13/13 | 2/2 |
| 20 | 40/40 | 4/4 |
| 21 | 13/13 | 2/2 |
| 22 | 13/13 | 2/2 |
| 23 | 9/13 | 1/2 |
| 24 | 10/13 | 1/2 |
| 25 | 11/13 | 1/2 |
| 26 | 12/13 | 1/2 |
| Average accuracy | 363/392=92.6% | 63/72=87.5% |

## **Weakness**

From the test cases, one notable issue arises from poorly shaded bubbles, which impacts accuracy. Achieving significantly higher accuracy under these conditions may not be feasible. Therefore, it is recommended that students use at least a 2B pen and ensure the bubbles are fully shaded for optimal results.

* Motivation: Manual grading of exams is time-consuming, prone to human error, and can be a tedious task for educators. An automated OMR system can significantly improve grading efficiency, reduce human error, and provide faster feedback to students.
* Module Description: This project aims to develop an automated system for grading multiple-choice exam papers. The system will utilize image processing techniques to extract relevant information from scanned images of the answer sheets, including student IDs, and marked answers.
* Methodology:
  + Image preprocessing: Image enhancement, noise reduction, and binarization.
  + Paper extraction: Isolate the student's paper from the background using contour detection and perspective transformation.
  + Bubble code extraction: Extract the student's ID/code using contour analysis and OCR.
  + Answer sheet extraction: Locate and extract the answer sheet region from the paper.
  + Bubble detection: Detect individual bubbles using contour analysis and shape analysis.
  + Answer marking: Determine whether a bubble is marked or not using pixel intensity analysis.
  + Grading: Compare student answers with the model answer key and assign grades.
* Expected Outcomes: A functional prototype of an OMR system capable of accurately grading multiple-choice exams from scanned images.

2. Used Algorithms

* Image Processing:
  + Paper Extraction: Contour detection, perspective transformation.
  + Bubble Code Extraction: Contour analysis, region of interest extraction, HOG + SVM for OCR.
  + Answer Sheet Extraction: Region of interest extraction based on image dimensions.
  + Bubble Detection: Contour analysis, shape analysis, and area filtering.
  + Answer Marking: Pixel intensity analysis within bubble regions.

3. Experiment Results and Analysis

* Data:
* Metrics: Accuracy
* Results:
  + [Present results in tables or graphs, showing accuracy, precision, recall, and F1-score for different components (bubble detection, OCR, overall grading).]
  + [Analyze the results, discussing factors that may have influenced performance, such as image quality, lighting conditions, and variations in handwriting.]
* Analysis:
  + [Discuss the overall performance of the system. Analyze the strengths and weaknesses of each component.]
  + [Identify potential sources of errors, such as incorrect bubble detection, misclassification of digits, or incorrect marking of answers.]

4. Performance & Accuracy

* Overall System Accuracy: [Report the overall accuracy of the system in grading the answer sheets.]
* Component-Wise Accuracy: [Report the accuracy of individual components, such as bubble detection, answer marking, and OCR.]
* Comparison with Pretrained DNN:
  + [Specify the chosen pretrained DNN model (e.g., a CNN-based OCR model).]
  + [Compare the performance of your system with the pretrained DNN model in terms of accuracy, speed, and robustness to noise and variations.]
  + [Discuss the strengths and weaknesses of each approach. For example, the custom OCR model might be more robust to specific types of noise or variations in handwriting, while the pretrained DNN might have higher accuracy on clean images.]

5. Conclusion

* Summary: Summarize the key findings and accomplishments of the project.
* Limitations: Discuss the limitations of the current system, such as sensitivity to extreme variations in image quality, difficulties in handling complex background patterns, and potential limitations in handling irregular shapes of bubbles.
* Future Work:
  + Explore more advanced image processing techniques for improved robustness to noise and variations.
  + Investigate the use of deep learning models for bubble detection and answer marking.
  + Develop a user-friendly interface for easy integration into educational settings.

6. References

* [List all the sources cited in the report]

7. Work Division Between Team Members (if applicable)

* [Describe how the work was divided among team members]

8. Additional Comments

* [Include any additional observations, insights, or challenges encountered during the project.]

9. Level of Variety for Test Cases

* Test Cases:
  + Variations in image quality (blur, noise, low resolution)
  + Different handwriting styles
  + Variations in lighting conditions
  + Simulated errors (e.g., misaligned answer sheets, stray marks)

10. Choice of Comparison Metric

* Metrics: Accuracy, precision, recall, and F1-score were chosen as they are widely used metrics for evaluating the performance of classification systems.
  + Accuracy provides an overall measure of correct classifications.
  + Precision measures the proportion of true positives among all predicted positives.
  + Recall measures the proportion of true positives among all1 actual positives.
  + F1-score provides a balance between precision and recall.

11. Complete Analysis of System Strengths and Weaknesses

* Strengths:
  + Robustness to moderate levels of noise and variations in handwriting.
  + Effective handling of common image quality issues.
  + Efficient processing of multiple answer sheets.
* Weaknesses:
  + May struggle with severely degraded images or highly irregular bubble shapes.
  + Performance might degrade significantly with extremely low-quality images or unusual handwriting styles.