MARKS TO CSV v2

Mini Project Presentation : Zeroth Review

Guided by: Ms. JABIN MATHEW
Assistant Professor, Department of AD

Presented By :
Ms. ANEETA JOSE, SJC21AD013
Mr. DON SABU, SJC21AD024
Mr. JACOB PHILIP, SJC21AD029
Mr. PRANAV P, SJC21AD049

February 29, 2024

Outline

- Introduction
- Where they left of
- Objectives
- Our plan
- Technical Glimpse
- Current status
- Project Timeline
- References
- Conclusion

Introduction

- Marks2CSV is an initiative of a team from our preceding batch consisting of Ajay, Justin, Emil, Vishnuprasad under the mentorship of Dr. Deepa V
- It aimed to streamline post-evaluation documentation efforts.
- Version 1 achieved significant results, highlighting future potential and motivating continued development to support our faculty.

Where they left of

Marks To CSV v1 achieved the following:

- Transforms handwritten marks into structured CSV files.
- Interprets marks ranging from 0 – 7, including instances with no awarded marks.
- Model capable of analysing 5 papers within 13 seconds.
- Model delivered an accuracy of 99.2%
- Outputs are in machine as well as human friendly CSV format, ensuring easy readability and editability.

Objectives

1. Scale up the model's capability to recognize digits up to 14 including fractional marks

2. A validation system to cross check the marks providing real time outputs

3. Develop a more seamless method to interact with the system and provide efficient processing of data.

Our Plan

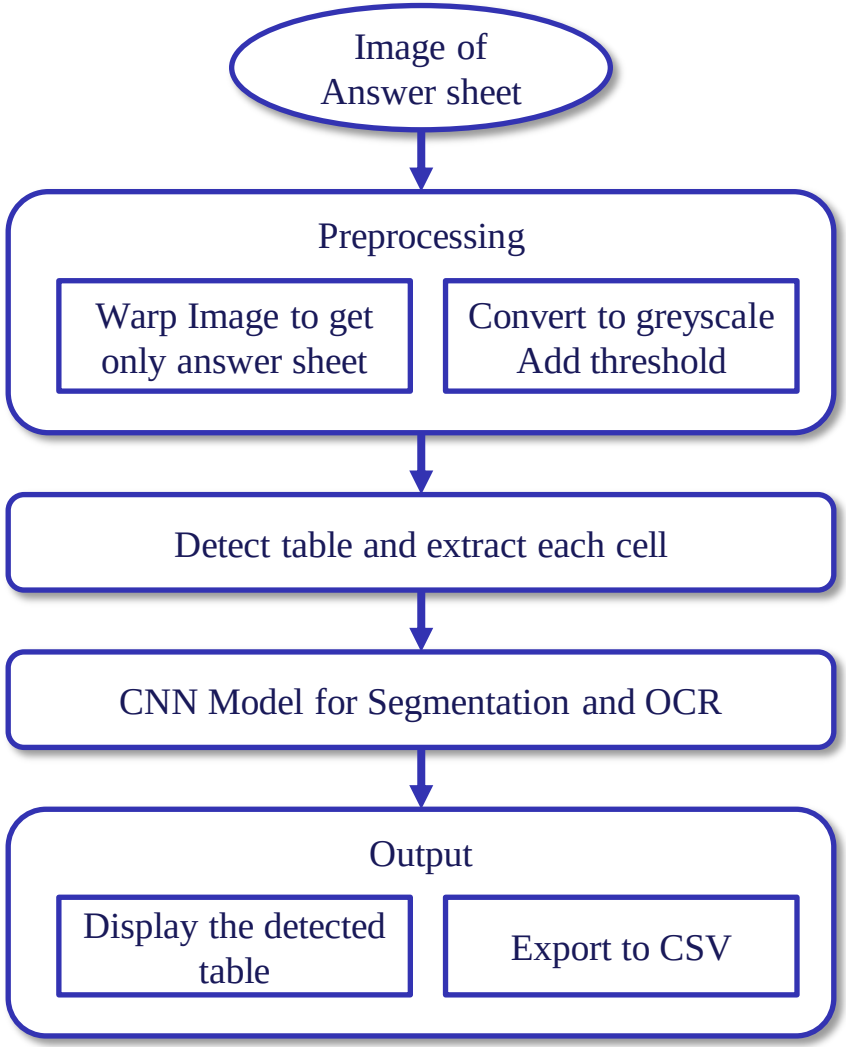
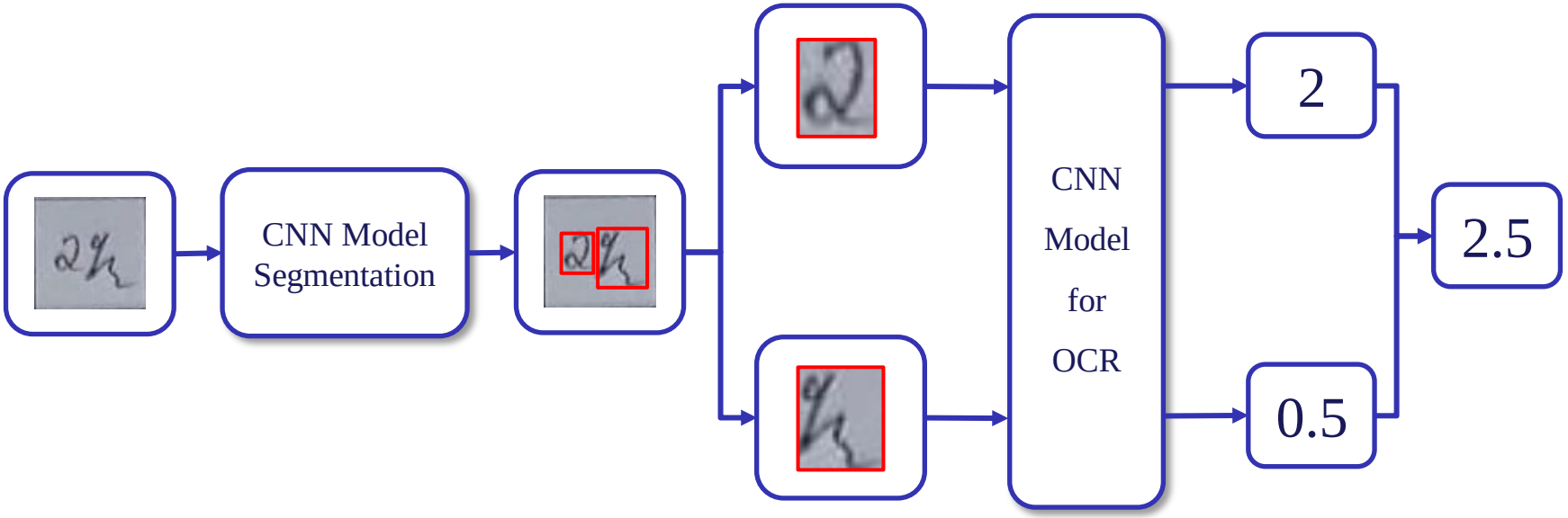


Image of
Answer sheet
Preprocessing
Warp Image to get only answer sheet
Convert to greyscale
Add threshold
Detect table and extract each cell
CNN Model for Segmentation and OCR
Output
Display the detected table
Export to CSV

Technical Glimpse

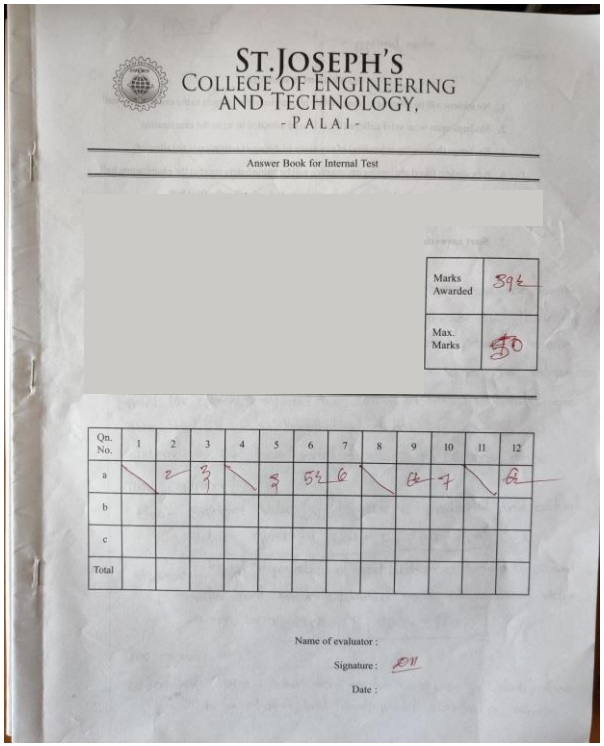
- Segmentation and Detection of Double Digits



Technical Glimpse

- Validation System

Input



Output to be Displayed

- Display the detected marks in the same table format.
- Show the cells that has less confidence while detection

Qn No	1	2	3	4	5	6	7	8	9	10	11	12
a		2	3		3	5.5	6		-	7		-
b												
c												
Tot												

Current Status

- We've consulted with the former batch and gathered their inputs and feedbacks.
- Trained a CNN model based on the inputs gathered and scaled up the scope to digits upto 9, including fractional marks. However, model is overfit due to lack of data.
- We're in the process of collecting more data for half marks. Currently we have around 200 instances for each, but we would need significantly more.

Project Timeline

- Achieved Milestones

First Team Meeting	Consulted former batch	Abstract Submitted	Experimented on CNN Models
15/01/2024	24/01/2024	29/01/2024	20/02/2024

- Future Deadlines

Complete Data Collection and Pre - Processing	Train new model for digits up to 14	Final CNN Model
By 11/03/2024	By 20/03/2024	By 29/03/2024
Implement to a usable form	**Testing Phase**	**Final Release**
By 12/04/2024	By 22/04/2024	By 26/04/2024

References

[1] Ian J. Goodfellow, Yaroslav Bulatov, Julian Ibartz, Sacha Arnoud, Vinay Shet, "Mutli-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks", 14 April 2014.

[2] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.

[3] R. Dixit, R. Kushwah, & S. Pashine, "Handwritten digit recognition using machine and deep learning algorithms", International Journal of Computer Applications, vol. 176, no. 42, p. 27-33, 2020.

[4] hristos N.E.Anagnostopoulos, Ioannis E. Anagnostopoulos, Vassili Loumos, Eleftherios Kayafas, “A License Plate-Recognition Algorithm for Intelligent Transportation System Applications”, IEEE Transactions On Intelligent Transportation Systems, Vol. 7, No. 3, pp: 377-392, September 2006

Questions ?

Thank You