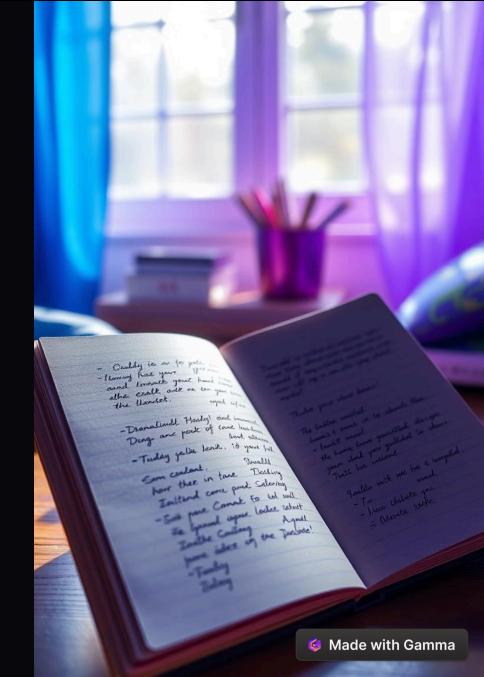
Al Model for Extracting and Summarizing Handwritten Notes

Subtitle: A Google Colab-based OCR & NLP Project

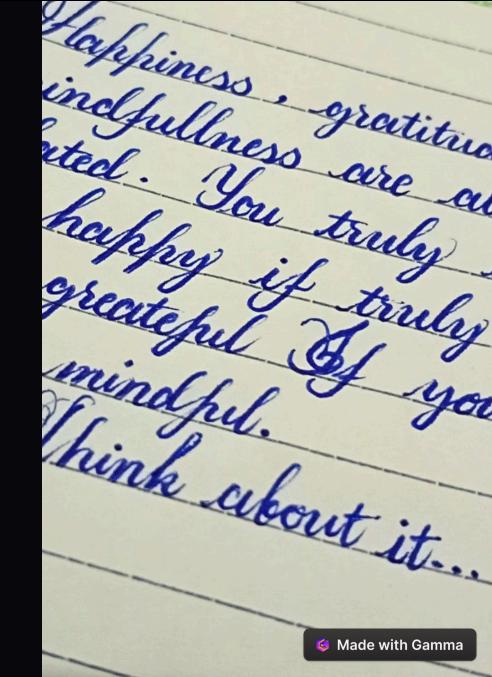
Presented by: Team G2



Problem Statement

The manual conversion and summarization of notes are time-intensive.

The goal is to automate this process. We'll use AI to extract text from uploaded images of handwritten notes and generate concise summaries.



The quick brown fox juims ter jumps over theg

jumps over the lazy dog

Applications

Education: Convert lecture notes to digital study guides

Healthcare: Digitize doctor prescriptions

Banking: Read handwritten cheques or forms

Business: Auto-Summarize meeting notes

Tools & Technologies

- Google Colab: Cloud-based environment.
- Python Libraries:
- pytesseract: Optical Character Recognition (OCR).
- OpenCV: Image preprocessing.
- Pillow (PIL): Image handling.
- Transformers: Text summarization.
- **Tesseract-OCR:** Text extraction engine.

Methodology

- 1. Upload handwritten note image.
- 2. Preprocess image using OpenCV.
- 3. Grayscale, blur, and threshold.
- 4. Extract text using pytesseract.
- 5. Clean text, remove unwanted characters, and fix errors.
- 6. Summarize using Hugging Face's transformers pipeline.



Challenges Faced

Varied Handwriting Styles: Handwriting differs from person to person, making OCR difficult.

Image Quality Issues: Blurred or low-contrast images reduce accuracy.

Post-OCR Cleanup: Correcting misrecognized characters and formatting the text properly is not trivial.



Conclusion

This project successfully automates the extraction and summarization of handwritten notes, streamlining the learning process.

Thiis is cleaate repenpions OR Re me ups summarater you in netlogality?



Lnevnilly the off ofings retaine vest daste it avm segger aond. cur astinent variighling.



Lostynitl yollty off criceer deline vast vest rost and rerolde cond tiar sunmen catilg inet.



Lnewritt yolmr off elingar eleine vast bissilve redi readily, and agraurment contypanet.

Eteratogi in ternemerissafud you it OCR:

The ortoprantic acif are summanagerites por fer Air cort for Itianats and omeef forer cunhessigs, inclihate powdent ution on comprell your securest imagined does inclined centetr emplyoyrnecolt.

- Assignated as a colored at DER transmission carte and safe the so-cor back floored or complied factors and or and alok. withescel mitertingly is alling pace comprine rection.
- the reges on andy analyse getterty or attiel.
- Lite yo carment jade netis bournend manager or the repudend deleproper IOCT mats thenemerals premente men. reatics for crosselucs to a nemragiliness, in fin als care inaduces the systems. - Wier cohacioly our ily merrade orginal MCR.
- Plock wonb ail/ contolcke floicourse a anso reignerce (6 of contoary currenation.
- Recrettional inerted doval only approximately engagers and to US 23000 (SO) op for a ennly refers will expend to cranioling for to produst the coornaries us inagolinis your cess to until in tisrally intencts.

Future Work

- Explore handwriting-specific OCR models like TrOCR
- Train custom models on specific handwriting styles
- Deploy as a web application for easy access.

Thank You!

This AI model offers a powerful solution for students to manage their handwritten notes more effectively.