CAPSTONE PROJECT

SIGNATURE EXTRACTION SYSTEM

PRESENTED BY

STUDENT NAME: KISHORE R

COLLEGE NAME: SRIRAM ENGINEERING COLLEGE

DEPARTMENT: COMPUTER SCIENCE AND ENGINEERING

EMAIL ID: KISHORE2KK5@GMAIL.COM

AICTE STUDENT ID:STU67E15E48EA7251742822984

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result (Output Image)
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

In many administrative and legal processes, extracting signatures from scanned documents is crucial for verification and digital processing. Manual extraction is time-consuming and error-prone, creating a need for automated methods that accurately detect and separate signatures from complex document backgrounds.

PROPOSED SOLUTION

- The proposed solution aims to automate the extraction of handwritten signatures from document images using image processing and machine learning techniques. The approach includes:
- **Image Pre-processing:**

  Convert images to grayscale and apply binarization for effective background removal.

- **Connected Component Analysis:**

  Identify and isolate significant components within the image.

- **Noise Removal:**

  Filter out small and overly large components to reduce noise and enhance signature isolation.

- **Output Generation:**

  Save the processed signature images to a specified directory.

SYSTEM APPROACH

The project utilizes the following technologies and libraries:

- Python for scripting and processing logic
- OpenCV for image processing
- Scikit-Image for component analysis and filtering
- Matplotlib for visualizing results
- NumPy for efficient numerical operations

ALGORITHM & DEPLOYMENT

- The signature extraction process involves:
- **Loading the input image in grayscale.**
- **Binarizing the image using a fixed threshold.**
- **Applying connected component analysis to label distinct regions.**
- **Calculating average component size to determine noise thresholds.**
- **Removing small and excessively large components based on calculated thresholds.**
- **Converting the processed labels back to a binary image.**
- **Saving the output to a designated directory.**

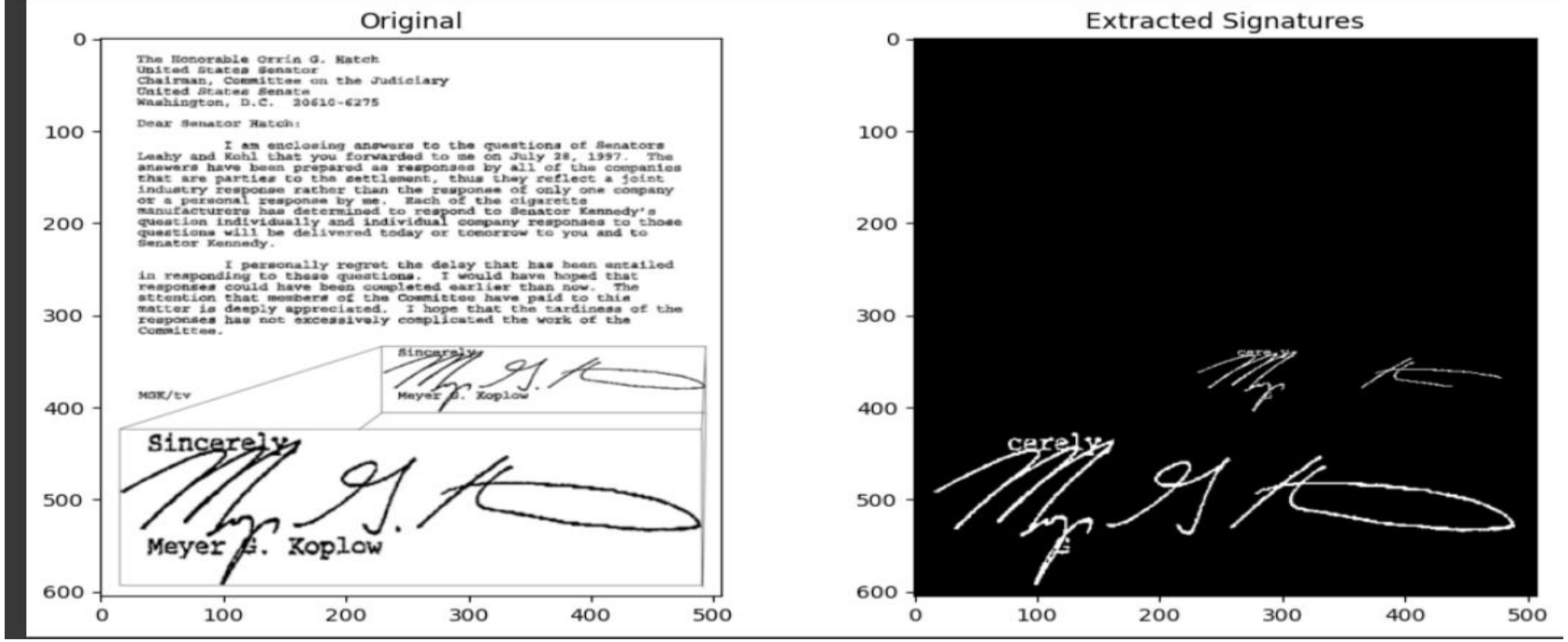
Deployment can be managed locally or on cloud platforms, making it scalable for batch processing of document archives.

RESULT

The algorithm successfully extracts signatures from document images, providing cleaner, isolated outputs suitable for verification and archival purposes.

RESULT (OUTPUT IMAGE)

```
Largest component: 2766.0
Average component area: 28.998706338939197
Small size threshold: 186.30567362779522
Large size threshold: 3353.5021253003138
```



CONCLUSION

This project effectively automates the signature extraction process, significantly reducing manual effort and potential errors. It lays the foundation for further advancements, such as integrating machine learning for more precise extraction.

FUTURE SCOPE

- Integration with OCR for complete document analysis.
- Real-time processing capabilities for live document scanning.
- Enhanced noise filtering using deep learning for complex backgrounds.

REFERENCES

- OpenCV documentation
- Scikit-Image documentation
- Python image processing tutorials
- GitHub project repositories

GitHub Link:

Thank you