

## Black Press + AI

### Automated Image Stitching and Page Segmentation for Historical Newspaper Digitization

Kingston Davies

Howard University, [kingston.davies@bison.howard.edu](mailto:kingston.davies@bison.howard.edu)

Saurav Keshari Aryal

Howard University, [saurav.aryal@Howard.edu](mailto:saurav.aryal@Howard.edu)

Gloria Washington

Howard University, [gloria.washington@Howard.edu](mailto:gloria.washington@Howard.edu)

Historical newspaper archives are critical for preserving Black press history, yet many digitized newspapers exist as fragmented image strips that are difficult to access and study. This fragmentation hinders researchers' ability to efficiently navigate and analyze these invaluable historical documents. Working with the Black Press Archive at Howard University, this project addresses this challenge by developing an automated pipeline to reconstruct complete newspaper pages from image strips. Using OpenCV computer vision library, the approach vertically concatenates sequential newspaper strips into complete documents. The system then employs brightness analysis to detect dark horizontal gaps between pages, automatically identifying page boundaries and splitting the stitched image into individual pages. This eliminates manual segmentation work that would otherwise be time-intensive and error-prone. The implemented pipeline successfully processes multiple newspaper strips, generates intermediate stitched outputs for quality verification, and produces organized individual page files. This automation significantly improves the accessibility and usability of the Black Press Archive's digitized newspaper collection for historical research and preservation.

**CCS CONCEPTS** • Computing methodologies~Artificial intelligence~Computer vision~Computer vision tasks~Scene understanding • Information systems~Information systems applications~Digital libraries and archives

**Additional Keywords and Phrases:** Computer Vision, Document Image Analysis, Historical Newspaper Digitization, Image Strip Stitching, Page Segmentation, Brightness-Based Gap Detection, OpenCV, Digital Archives

## REFERENCES

- [1] OpenCV. n.d. *Image file reading and writing*. OpenCV 3.4.20-dev Documentation (imgcodecs module). [https://docs.opencv.org/3.4/d4/da8/group\\_imgcodecs.html](https://docs.opencv.org/3.4/d4/da8/group_imgcodecs.html)