

Advancements in Historical Document Image Processing using CNN

ID: 2017-054

Executive Statement:

This technology enhances the classification accuracy of historical document images through semantic segmentation using convolutional neural networks.

Technology Overview:

The document discusses a novel deep learning architecture for accurately segmenting and classifying diverse content in historical documents. By employing convolutional neural networks (CNNs) for pixel-level labeling, the technology can differentiate between handwriting, machine print, form lines, and stamps, even when these elements overlap. It introduces a unique combination of downsampling and upsampling layers, enabling precise semantic parsing of document images and facilitating the processing of large-scale historical records.

Key Advantages:

- High classification accuracy with minimal training data through data augmentation and balanced training approaches
- Ability to assign multiple class labels per pixel, enhancing detail in content differentiation
- Improved understanding of spatial context without aggressive downsampling, allowing for efficient processing of large document images
- Capability to approximate human-level performance in per-pixel classification tasks

Problems Addressed:

- Challenges in segmenting and classifying densely packed and diverse content types in historical documents
- Difficulties in processing large-scale document images with existing methodologies
- Limited accuracy in automated indexing and transcription of mixed-content document images

Market Applications:

- Automated indexing and transcription services for libraries, archives, and museums with historical document collections
- Enhanced document processing solutions for companies involved in digital document management, such as Adobe and Google
- Intellectual property management, including patent filing and technology transfer in

document processing technologies