LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis

Zejiang Shen[1] (✉), Ruochen Zhang[2], Melissa Dell[3], Benjamin Charles Germain Lee[4], Jacob Carlson[3], and Weining Li[5]

[1] Allen Institute for AI
shannons@allenai.org
[2] Brown University
ruochen_zhang@brown.edu
[3] Harvard University
{melissadell,jacob_carlson}@fas.harvard.edu
[4] University of Washington
bcgl@cs.washington.edu
[5] University of Waterloo
w422li@uwaterloo.ca

Abstract. Recent advances in document image analysis (DIA) have been primarily driven by the application of neural networks. Ideally, research outcomes could be easily deployed in production and extended for further investigation. However, various factors like loosely organized codebases and sophisticated model configurations complicate the easy reuse of important innovations by a wide audience. Though there have been on-going efforts to improve reusability and simplify deep learning (DL) model development in disciplines like natural language processing and computer vision, none of them are optimized for challenges in the domain of DIA. This represents a major gap in the existing toolkit, as DIA is central to academic research across a wide range of disciplines in the social sciences and humanities. This paper introduces LayoutParser, an open-source library for streamlining the usage of DL in DIA research and applications. The core LayoutParser library comes with a set of simple and intuitive interfaces for applying and customizing DL models for layout detection, character recognition, and many other document processing tasks. To promote extensibility, LayoutParser also incorporates a community platform for sharing both pre-trained models and full document digitization pipelines. We demonstrate that LayoutParser is helpful for both lightweight and large-scale digitization pipelines in real-word use cases. The library is publicly available at https://layout-parser.github.io.

Keywords: Document Image Analysis · Deep Learning · Layout Analysis · Character Recognition · Open Source library · Toolkit.

1 Introduction

Deep Learning(DL)-based approaches are the state-of-the-art for a wide range of document image analysis (DIA) tasks including document image classification [11,