

**TITLE OF INVENTION: SAMPLE CONSTRUCTION METHOD AND DEVICE,
MODEL TRAINING METHOD AND DEVICE**

ABSTRACT

A sample construction method and device and a model training method and device are provided. The sample construction method includes: acquiring a document file and a text corpus, determining multiple document images and a document base corpus according to the document file, and determining a text base corpus according to the text corpus; extracting element style information and page layout information corresponding to each of the document images, constructing a document template according to the element style information and the page layout information, and determining a target base corpus according to the document base corpus and the text base corpus; fusing the document template and the target base corpus by using an element renderer and a layout manager according to a preset writing mode and a font resource library, to obtain a document recognition sample and a document recognition label; and constructing a document recognition sample pair based on the document recognition sample and the document recognition label, and storing the document recognition sample pair into a sample library.

(Figure 1)