

PageOCR API

Specifications

Swagger URL	https://bhasha.iiit.ac.in/pageocr/docs
URL	https://bhasha.iiit.ac.in/pageocr/api
Method	POST
Headers	{“Content-Type”: “multipart/form-data”} (Auto-generated)
Files	image : The page level image to be processed is uploaded.
Payload	language : specify the language of the page. Can take values of: [hindi, marathi, tamil, telugu, kannada, gujarati, punjabi, bengali, malayalam, assamese,manipuri, oriya, urdu]
Response	Json response object containing the word level bounding boxes and Page OCR text
Sample Code (Python)	<pre>import requests url = "https://bhasha.iiit.ac.in/pageocr/api" payload={'language': 'hindi'} files=[('image', ('sample.jpg', open('<path-to-image>.jpg', 'rb'), 'image/jpeg'))] headers = {} response = requests.post(url, headers=headers, data=payload, files=files) print(response.text)</pre>
Sample Code (cURL)	<pre>curl --location --request POST 'https://bhasha.iiit.ac.in/pageocr/api' \ --form 'image=@<path-to-image>.jpg' \ --form 'language="hindi"'</pre>