Jiedong Hao

jdhao@hotmail.com · (+86) 132-6152-3068 · https://jdhao.github.io ·

Education

Institute of Automation, the Chinese Academy of Sciences, Beijing, China 2014 – 2018

Master in Pattern Recognition and Intelligent Systems

Central South University, Changsha, China 2010 – 2014

B.S. in Automation

Professional Skills

- Programming
  - Proficient with Python; know how to use Flask, uwsgi and docker to deploy online web services; know C++ and Lua; familiarity with common algorithms and data structures
  - Proficient with Linux, shell (Bash, Zsh), shell scripting; Git (branch, merge, rebase, stash etc.); make; Docker (building using Dockerfile and push and daily usage)
  - Proficient in using Vim/Neovim (daily dev tool); familiarity with PyCharm and Sublime Text
- Pattern Recognition and Machine Learning
  - OCR (model training, training sample generation) and evaluation method; techniques related to document layout analysis and reconstruction: object detection, document layout analysis and post-processing, PPT generation; image retrieval; video fingerprinting/deduplication algorithm
  - Familiarity with PyTorch framework
- language
  - Languages: English - Fluent, Mandarin - Native speaker
  - Excellent reading and writing skills in English
  - CET-4: 650, CET-4: 613

Professional Experience

vivo business card OCR and general OCR Oct. 2018 – Jun. 2019

Algorithm development

- In charge of developing model to recognize vertical text and developing tools to generate images of vertical text. The image generation tool uses multiple text sources such Wikipedia and supports both simplified and traditional text, with background, color, noise, blur and cropping effects. Accuracy, (1) business card OCR:, CER: 0.06, line accuracy: 0.85 (2) general OCR, CER: 0.14
- Design and train document orientation classification model, accuracy: 0.99
- Design a model to classify the orientation of text line images, accuracy: 0.97

vivo smart vision – WiFi recognition Jun. 2019 – Feb. 2020

Algorithm development and service deployment

This feature enables the user to take a photo of WiFi info and recognize the account and password in it. Then the user can connect to WiFi with a single click.

- In charge of the whole project (data collection, quality goal, developing algorithm, fixing bugs). I also design a smart WiFi info parsing module based on OCR result, and deploy the whole system as a web service using Flask, docker and uwsgi.
- The system can parse complex real-world WiFi images, including one-line, multiple layout, print/hand-written images and images with different angle of orientations.
- Accuracy of the system, for account 0.87, for password: 0.85

vivo smart vision – PPT reconstruction Jan. 2020 – Dec. 2020

Algorithm development and service deployment

We designed and implemented a system to transform the user-taken PPT image to an editable PowerPoint document. The document background, text color, paragraph and layout are reconstructed as much as possible. Users can preview the reconstructed doc and generate the document on the fly.

- In charge of the whole project (data collection, quality goal, developing the model and fixing bugs). I am responsible for all the work related to algorithms and partly responsible for engineering deployment to make the whole system a web service.
- Technology involved: document element detection via maskRCNN; document layout analysis and post-processing, which involves post-processing detection results, OCR results (forming lines and paragraphs), deciding the right text size, line width, and correct layout analysis for images, text, tables etc; generate an editable PowerPoint document using python-pptx, based on the document image analysis result.
- Reconstruction acceptance rate (via human inspection): 87%+.

vivo content platform – video deduplication Jan. 2021 – present

Algorithm development and service deployment

In order to reduce the content duplication rate in video feeds to promote user experience, we designed and implemented a system to find duplicated videos provided by different content providers.

- Design and implement the video fingerprinting algorithm; design the evaluation metric; design and implement the offline evaluation framework
- Technology involved: video frame and audio extraction; video frame feature extraction; large scale feature retrieval; video and audio deduplication; detecting black edge from videos; video gid generation.
- Participate in the design of the workflow of the whole system; in charge of developing and deploying various video-deduplication related services, and video frame feature extraction service.
- The system processes over 30 million videos monthly. Evaluation metric, recall: 0.98, FP rate: 15 ppm (previous value, recall: 0.7, FP rate: 40 ppm)

MISCELLANEOUS

- GitHub: https://github.com/jdhao
- Tech Blog: https://jdhao.github.io/ (total PV: 2.6 million, monthly PV: 30000)
- Stack Overflow reputation 16000+, top 0.09% among all users

PUBLICATION

- Jiedong Hao, Yafei Wen, Jie Deng, Jun Gan, Shuai Ren, Hui Tan and Xiaoxin Chen, EEM: An End-to-end Evaluation Metric for Scene Text Detection and Recognition. ICDAR 2021
- Jiedong Hao, Jing Dong, Wei Wang and Tieniu Tan, DeepFirearm: Learning Discriminative Feature Representation for Fine-Grained Firearm Retrieval. ICPR 2018 (**Best Scientific Paper Award**)