Jiedong Hao

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

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 Vim/Neovim (daily dev tool); familiarity with PyCharm and Sublime Text
- Pattern Recognition and Machine Learning
  - OCR: text recognition, text image generation tool; document image reconstruction: object detection, document layout analysis and post-processing, PPT generation; image retrieval; video fingerprinting/deduplication algorithm
  - Familiarity with PyTorch
- Language
  - Languages: English - Fluent, Mandarin - Native speaker
  - Excellent reading and writing skills in English
  - CET-4: 650, CET-6: 613

Professional Experience

video fingerprinting/deduplication – vivo Jan. 2021 – present

Algorithm development and service deployment

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

- Technology involved: video key frame extraction; contrastive frame feature learning; large scale feature retrieval; video similarity via LCS matching; video duplication GID generation.
- (1) Participate in the design of the workflow of the whole Systems (2) Design and implement the video fingerprinting algorithm, the evaluation protocol, and the offline evaluation framework (3) Develop and deploy various video-deduplication related services, and video frame feature extraction service.
- The system processes over 30 million videos monthly. System performance metric, recall: 0.93, precision: 0.92

PPT reconstruction from a single image – vivo Oct. 2019 – Dec. 2020

Algorithm development and service deployment

We designed and implemented a system to turn PPT images to an editable PowerPoint document. The document background, text color, paragraph and layout are recovered as much as possible. Users can preview the reconstructed PPT and generate the PPT on the fly.

- Lead the whole project (data collection, quality goal, model development and bug fixes). I am responsible for all the work related to algorithms and partly responsible for deploying the whole system as 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 (by human inspector): 90%+.

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

Algorithm development

- Develop a tool to generate vertical text images for training and develop CRNN model to recognize vertical text. The image generation tool uses multiple text sources such Wikipedia and supports both simplified and traditional Chinese, with various background, color, noise, blur and cropping effects to simulate real world scenarios. 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 lines, accuracy: 0.97

Smart WiFi recognition – vivo Jun. 2019 – Dec. 2020

Algorithm development and service deployment

This project enables vivo smartphone users to take a photo of WiFi plate and connect to it via a single click.

- In charge of the whole project (data collection, quality goal, algorithm development, bug fixes). I also designed a smart WiFi info parsing module based on OCR result. The whole system is deployed 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 rotated images.
- Accuracy of the system, account: 0.87, password: 0.85

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

PUBLICATION

- 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**)
- 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

MISCELLANEOUS

- GitHub: https://github.com/jdhao
- Tech Blog: https://jdhao.github.io/ (total PV: 2.8 million, monthly PV: 30K)
- Stack Overflow: reputation 17k, top 0.09% among all users