Dinh Cong Vu

AI ENGINEER

CONTACT ME

[redacted]

[redacted]
om

[redacted]
m/dinhcongvu/

Alley 20 Le Thanh Nghi,
Hai Ba Trung District,
Hanoi

EDUCATION

Data Science and Artificial Intelligence at HUST

9/2020 - Now

CPA: 2.9

SKILLS

- Program Language: Python, C
- Read, speaking and write fluently

CERTIFICATE

- B2 - VSTEP Certificate (3/2024)

WORK EXPERIENCE

AI Engineer Intern — 3/2023 - 5/2024

ICOMM Media & Tech JSC

PROJECTS

LICENSE PLATE RECOGNITION

- Description: License plates are recognized as one of the important problems in CCTV. Includes detected and recognized license plates (image Super Resolution can also be used). Detection model used: Retina-Net combined with Perspective Transform to give the most linear figure possible. Using GFP-GAN to recover counts. Get Yolov4 interface for detection and sorting.
- Experiments: The results are quite good. The combination of alignment and resolution methods makes the LPR task more efficient.

GENDER CLASSIFICATION

- Description: Gender classification is simply a classification problem with two outputs: Male or Female. The requirement is that the model must be fast and highly accurate. Tested on many light Vit models such as: DeIT-Tiny, Pit-Tiny, TinyVit, ..., small models Dino-S, Dinov2-S,.. as well as large Vit models such as Base version: Unicom-Base , OpenCLIP-base.
- Experiments: Base models have high accuracy as expected (highest Unicom-Base : 94) Small (Dinov2 :92.5) and Tiny (Pit-Tiny :88.2). Use the Dinov2 model as the model for the problem because it is highly accurate and also quite fast.

FACE ATTRIBUTE RECOGNITION

- Description: Face attribute recognition is a technology that leverages computer vision algorithms to analyze facial images and videos. It focuses on identifying and extracting specific attributes such as age, gender, facial expressions, ethnicity, and facial features.
- Experiments: By using Retina-Face combining Model mutli-task (backbone is Iresnet160) for recogntion attributes, the results are quite good. Person' age is recognized quite outperformace.

DOCUMENTS INFORMATION EXTRACTION

- Description: Automatic Information Extraction of Documents are very popular in nowadays. In this project, I extracted some specific documents and general documents. In specific documents, including Citizen Information Card, High School Diploma, Residence Information and table extraction (even multi-tables, merged cells and long line table). Using Yolov5 for information detection, corner detection for rotating (table and diploma). In general documents, using rule to algin words in texts and filtering texts in Stamps. All results will return file docx (excel in table). Finally, Newspapers are exported by using Layout model and Sort algorithm to extract.
- Experiments: CIC extraction is very good (almost correct information in back and front card). Residence Information and High School Diploma are quite accurately. Table recognition is very excellent with high accuracy recognition. General Documents are quite accurately, difficult to recognize in complex layout documents. Newspapers are exported as forms in their original forms.

VIDEO SYNOPSIS

- Description: Video Synopsis is a technique that condenses videos into shorter summaries by extracting key frames and highlighting important events, enhancing video understanding and review efficiency. With using NvDCF tracking and Yolor as detection model to extract sequential frames. The combination of space and time relationship and greedy algorithm to group tubes and arrangement tubes. Finally, video synopsis will be created by rendering frames to background, which is created by temporal median technique.
- Experiments: Synopsis video follows the laws of time and ensures interaction between objects. Videos range in length from 30 minutes to 3 minutes with reasonable timing.