MODEL CARD

Human Detection

Model Details

This model has approximately 100,000 parameters and is trained using YOLOv5 for human detection. It is designed specifically for detecting individuals (Human) as the only class. The model is capable of real-time inference on webcam images, supporting over 5 different poses. When performing inference on the VA8801 platform, it achieves a speed of 50 frames per second (FPS) with an accuracy exceeding 90%.

Detecting the position and number of people

Model Specifications

Model Type : Convolutional Neural Network

Model Architecture : Yolov5n with customized backbone for VA8801

Input : 1*96*96

Output : [class,x0,y0,x1,y1]

Class : Human

Number of people : 5 (suggestion) or more

Angle of view : Suggested to be at eye level or within a 45-degree downward angle.

Posture :

stand, sit, walk, bend, crouch, raise hand, turn around, …

Application

Product for Human Detection : Surveillance Devices, People Counting

Application Example : To enhance overall performance, it is suggested to incorporate motion detection in preprocessing to capture moving objects. Subsequently, crop and resize the region of interest to match the input size of this model

Limitation

(1) Detection is not reliable beyond a distance of seven meters.

(2) Individuals too small (height less than 50 pixels) may not be detected.

(3) Not supported under lighting conditions with less than 30 lux.

(4) Special poses such as lying down, bending, or raising hands may result in decreased recall.

(5) People must have clear outlines and be distinguishable from the background

(6) Under backlight conditions, non-standing human forms may cause a decrease in recall.

Training Data

(1) Images which are processed by ISP ROI algorithms from CCTV footage dataset

(2) COCO-human

(3) Surveillance dataset

Total : about 600 images

Reference

https://cocodataset.org/#home