

人物姿勢推定アルゴリズムをあれこれ試す

ラズパイで エッジAI



新連載

第1回

開発環境の準備

笹尾 幸良

画像認識分野はアルゴリズム/計算機/ビッグ・データの進化によって急速に発展しています。また、広く知られているラズベリー・パイもさらに進化して高性能となり、さまざまなAIモデルを気軽に実行できるようになりました。

そこで、本連載では人物姿勢推定をラズベリー・パイ5上で動作させ、さらに転倒検出などを実現します。今回は本連載で使用するプログラムを実行するための環境構築をします。

● 人物姿勢推定とは特定の部位を検出すること

人物姿勢推定 (Human Pose Estimation) は、骨格検出/特徴点検出/キーポイント検出とも呼ばれ、画像から人物の顔だけでなく肩や足など特定部位の点位置を検出するタスク (さまざまな目標) の1つです。

2018年に論文とともに公開されたOpenPose⁽¹⁾は、論文タイトルに「Realtime Multi-Person ..」と付くように、複数人の姿勢推定をリアルタイムに実行できる点がブレイクスルーとなり、監視カメラやモーション・キャプチャ・システムに広く利用されました。OpenPoseでは、ニューラル・ネットワークが直接関節位置を出力せず、中間情報であるヒートマップ画像と関節の結び付きを意味する情報を出力することによって複数人の検出を実現していました。

AIモデル推論の枠組み

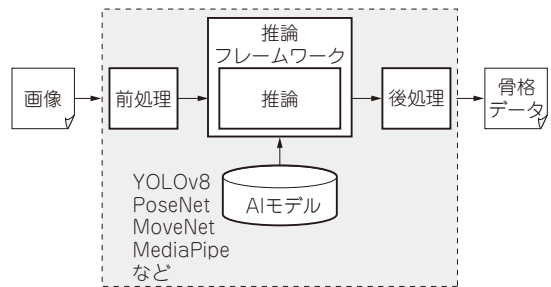
AIによる画像認識処理が一般的にどのような構成と処理フローで実行されるのかを図1に示します。

● 前処理

AIモデルの入力フォーマットに合わせて、

- ・リサイズ
- ・並び替え
- ・平均と分散による標準化
- ・量子化

といった処理が行われます。



推論ハードウェア (小型コンピュータ・ボード)

図1 姿勢推定の処理フロー

● 推論フレームワーク

TensorFlow Lite Runtime, PyTorchの他、TensorFlow, ONNX Runtimeといった一般的なPCで動作するフレームワークがラズベリー・パイ上でも動作します。

● 推論ハードウェア

今回はCPUのみを使用しますが、GPUや他の外付けハードウェアも使用できます。AIモデルの入出力および内部のテンソルと呼ばれるデータが浮動小数点型か、または量子化された整数型かを知っておく必要があります。型によっては推論ハードウェアが対応していない場合もあります。

● AIモデルと後処理

ユーザが欲しい骨格データを直接出力せず、例えばヒートマップのような、アルゴリズムにとって都合の良いデータを出力することがあります。

AIモデルの出力をユーザが欲しいデータに加工するのが後処理です。例えば検出系の画像認識ではしばしば、NMS (Non-Maximum Suppression) と呼ばれる後処理を実行します。検出系タスクにおいてNMSは、AIモデルが出力した近傍の検出候補をまとめるような処理を行います。

◆参考文献◆

(1) openpose, GitHub.

(2) H.A. Rowley, S. Baluja, T. Kanade : Neural Network-Based Face Detection, Appears in Computer Vision and Pattern Recognition, 1996.