

防犯システムのための CNN を用いた不審動作認識手法の提案

1. はじめに

近年, コロナの影響により女性や高齢者を狙った強盗が多発し, 防犯システムの開発が進められている. 一原ら[1]は 3DCG による画像データに Openpose で取得した関節の 2 次元の座標データを加えて機械学習を行い, 不審動作を認識している. また, 梅津[2]は Kinect を用いて人物領域を抽出し, 転倒人物の姿勢推定を行っている. しかし, これらの手法では十分な認識精度は得られていない. そこで, 本研究では関連研究から人物領域とスケルトンを表示させた画像データを CNN で学習させ不審動作認識をする方法を提案する.

2. 不審動作認識

2.1 CNN について

本研究では畳み込みニューラルネットワーク (CNN) を用いて不審動作の学習, 判別を行う. CNN は局所的に特徴量を抽出することができ, 画像処理分野では優れた性能を発揮している. CNN を構成する層の一つである中間層(隠れ層)は決まりがなく, 扱う情報に合わせて設定可能である. 本研究では, 畳み込み層や ReLU, プーリング層で特徴を学習し, ソフトマックス関数を用いて分類出力を行う.

2.2 認識する動作と学習データ

本研究では, 暴力行為やそれに伴う行動を不審動作として取り扱い, さらに一般的な動作も識別する. そのため, 不審動作として「殴る」, 「蹴る」, 「蹴る構え」, 「うづくまる」を認識する. また, 一般的な動作として「座る」, 「立っている」を認識し, 計 6 つの姿勢に分類する.

関連研究[2]では図 1 左のような人物領域を表示した画像のみで学習を行うのに対し, 本研究では図 1 中央のように画像にスケルトン (骨格情報) を追加する. さらに, 図 1 右のようにスケルトンに色を付け, 関節の意味情報を加える. 画像に関節の特徴を追加することで服装の変化や未学習の人物への汎用性を高め, 認識精度の向上を図る.

2.3 拡張方法と画像データ量

本研究では, 認識精度の向上と画像収集の負担軽減のため Keras の ImageDataGenerator を用いて次のようなデータ拡張を行う. まず, 画像中心を軸にランダムな左右反転を行う.



図 1 殴る動作の学習データ (抜粋)

表 1 各動作の認識精度

動作	画像	画像+ スケルトン	画像+色付き スケルトン
殴る	0.97	0.81	0.21
蹴る	0.58	0.74	0.21
蹴る構え	0.27	0.73	0.53
うづくまる	0.73	0.87	1.00
座る	0.83	0.98	0.60
立っている	0.71	0.73	1.00
精度	0.72	0.81	0.70

さらに, 画像を左右に平行移動する. このようなデータ拡張を各画像に対して行う. 今回は画像を各動作 1500 枚に拡張し, 6 つの動作で合計 9000 枚の学習データを用いた.

3. 精度評価

不審動作の認識実験を行った. 被験者は学習済みの人物 1 名, 未学習の人物計 5 名の計 6 名とした. テストデータは各動作 60 枚ずつの画像を用意し, 合計で 360 枚を用いた. 各動作の認識精度を表 1 に示す. 画像にスケルトンを追加することで精度が向上していることが確認できた. また, 画像のみやスケルトンを追加した場合では, 蹴る構えの精度が低くなっており, 立つや蹴る動作と似ているためだと考えられる. 色付きスケルトンでは全体的な精度は下がってしまったが, うづくまるや立つの精度が向上しており, 関節の意味情報を追加した効果だと考えられる.

4. おわりに

本研究では CNN を用いて不審動作を認識する手法を提案した. 人物領域にスケルトンを追加した学習データを CNN で学習させることで, 未学習の人物に対しても高い認識精度が得られた. しかし, 学習データが少なく, 似た画像に対する精度が低くなっていた. 今後は, 学習データを増やし, CNN のパラメータを変更することで認識精度の向上を図る.

【参考文献】

- [1] 一原賢吾, 竹内建, 金井謙治, 甲藤二郎: 姿勢推定技術に基づく機械学習を活用した人物の不審動作認識手法の精度評価, 電子情報通信学会信学技報, 2018 年
- [2] 梅津啓太: 見守りシステムのための CNN を用いた転倒人物の姿勢推定, 福岡工業大学情報工学部情報工学科卒業論文, 2020 年度