

状態分類による高齢者転倒危険の検知

Status Classification for Elderly Fall Detection

陳楽涵†, 孔祥博†, 富山宏之†, 孟林†

LeHan Chen, XiangBo Kong, Hiroyuki Tomiyama, Lin Meng

1. はじめに

現在、日本のみならず、先進国において、高齢者問題が重要な社会問題であるため、安全、安心な高齢化社会の実現は重要な研究テーマとなる。日常生活において、高齢者の安全を非常に脅かすのが転倒である。本研究は画像処理・機械学習・深層学習を用いて、高齢者の転倒などの危険をリアルタイムに検出する IoT システムを実現し、安全安心な高齢化社会の実現に貢献する。詳細について、Depth Camera で得られた深度情報に対して、画像処理・機械学習及び深層学習を用いて、高齢者の状態を分類し、危険な状態をリアルタイムに検知する。危険が検出された場合は直ちに家族や病院に通報することにより、高齢を見守る。本稿は、機械学習・深層学習を用いて、高齢者の状態を、高齢者の転倒、立ち及び他の姿勢との三クラスに分類し、最適な手法を議論する。

2. 状態の分類と画像の取得

2.1 状態の分類

図1に人体状態の分類を示す。本研究では文献[1]の理論に基づいて、六個の姿勢に対して、三つのグループ (Stand, Fall, Others) と分類する。詳細について、以下に示す。

- Stand : (W)人が立つ又は人が歩く姿勢
- Fall : (L)人が転倒する姿勢
- Others : (Si)人が椅子に座る姿勢
(Sq)人がしゃがむ姿勢
(B)人の腰をかかめる姿勢
(F)人がうつぶす姿勢

その中、Others に含まれた姿勢は通常の認識率に影響するだけでなく、体調又は健康状況を考察する有効な手段となる。従って、Others の検知も重要であると考えられる。

2.2 画像の取得

本研究では、深度画像(depth image)を撮る機能を搭載する Intel RealSense D435 で実験を行った。Intel RealSense D435 は 2018 年半導体会社 Intel で公開した深度カメラで、

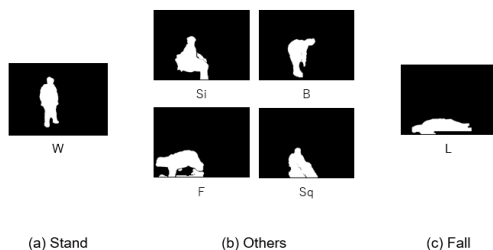


図1 状態の分類



図2 Intel RealSense D435 と実験環境

図2(左)に示す。二つの深度センサ、RGBセンサとIR投射器を搭載し、USB給電で動作する。

深度画像は物体とカメラの距離により生成された画像であり、RGB画像と比べると、深度画像のほうが環境(光線、ノイズ、位置など)の影響が少ないため、人体の認識精度率が高くなる。

今回、実験の原始データは2.1節に記載された六つの姿勢に対して、各状態を500の画像を取得した。実現環境は図2(右)で示し、RealSenseと地面の高度が1.3mを設定した。一つの姿勢について、図2(右)の矢印で示すように、人が八方向に向かって画像をランダムに取得した。六つの姿勢、八方向で全部3000枚の深度画像を取り、実験を行った。

3. 機械学習と深層学習を用いた転倒危険の検知

3.1 機械学習を用いた転倒の検知

機械学習での転倒検出は、まず、Histograms of Oriented Gradients(HOG)を用いて人の特徴点を抽出する。そして、Support Vector Machine(SVM)によりトレーニングする。最後、テスト画像を用いて、モデルの性能を評価する。

図3はHOG処理の流れを示す。まず、深度カメラから取得した画像を二値化する。そして、ピクセルの勾配(強度と方向)を計算し、選択セルのヒストグラムを作成する。最後は正規化の特徴量を抽出し、すべての特徴量を統合する。

SVMはパターン認識のモデルとして、幅広く使われている。本研究では、立つ姿勢(Stand)、転倒の姿勢(Fall)、他の姿勢(Others)を含める三つの状態を分類するため、one-versus-one(OVO SVM)法を用いる。

OVO SVMは、任意の二つ状態をSVMで分類し、その結果等を統合する方法である。k個の状態は、二分類のSVM過程を $k(k-1)/2$ 回で進行する。今回の実験は三分類から、 $k=3$ であり、二分類のSVMは三回を進行し、その結果は

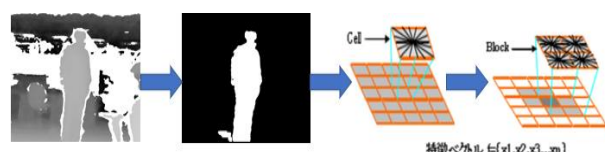


図3 HOG

† 立命館大学 大学院理工学研究科, Graduate School of Science and Engineering, Ritsumeikan University

投票の形式で確定する。アルゴリズムは以下のように示す：

```
mark stand=fall=others=0;
picture Stand, Fall, Others;
classifier (Stand, Fall) if Stand
stand=stand+1; else fall=fall+1;
classifier (Stand, Others) if Stand
stand=stand+1; else others=others+1;
classifier (Fall, Others) if Fall
fall=fall+1; else others=others+1 ;
result=max (stand, fall, others);
```

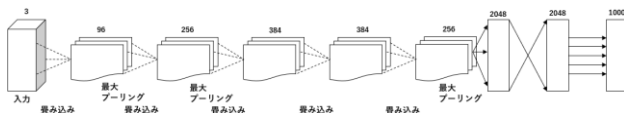


図 4 AlexNet

3.2 深層学習を用いた転倒の検知

Convolutional Neural Networks (CNN)はよく使われている深層学習の手法である。本研究は AlexNet と GoogLeNet を用いて、体の三つの状態を認識し、その有効性を示す。AlexNet は 5 層の畳み込み、3 層のプーリング、及び 3 層の全結合で構成させる [2]。AlexNet のモデルは、図 7 に示し、画像の核心を変わず、元画像にランダムで一定ピクセルサイズのブロックを選択し、それぞれのピクセルの R,G,B 数値を公式で変換し、データの増強を行う。それにより、認識率を向上させる。

GoogLeNet はもう一つ CNN に基づいてニューラルネットワークである。GoogLeNet の一番の特徴は、複数の畳み込み層や pooling 層から構成される Inception モジュールと呼ばれる小さなネットワーク (micro networks) を定義し、これを通常の畳み込み層のように重ねていくことで 1 つの大きな CNN を作り上げている点である [3]。

4. 実験

4.1 実験内容

訓練データは、3.2 を述べたデータ集で Stand, Fall, Others (Si, B, F, Sq は各 100 枚) の画像がそれぞれ 400 枚を取り、全 1200 枚画像を二値化処理した。実験データは、データ集に訓練データ以外で、上記の三分類それぞれ 100 枚を取り、二値化後全部で 300 枚の画像を組む。

本実験は、機械学習 HOG+SVM、深層学習 (AlexNet、GoogLeNet) を用いて訓練データをトレーニングし、生成のモデルを実験データでテストする。その結果を評価し、有効性を議論する。

HOG+SVM は、Microsoft に基づく OpenCV を使用した。AlexNet、GoogLeNet は NVIDIA を公開した Deep Learning Box で実験を行った。入力画像は、サイズ 640*480、グレースケールの深度画像である。CPU: Xeon E5-1620v3, GPU は GTX1080ti x4 で、メモリは 64GB である。

4.2 実験結果、考察

表 1 に機械学習と深層学習の認識結果を示す。

表 1 実験結果

	元の状態/ 認識状態	Fall	Others	Stand
HOG+ SVM	Fall	100%	0%	0%
	Others	59%	38%	3%
	Stand	0%	0%	100%
Alex Net	Fall	100%	0%	0%
	Others	13%	51%	36%
	Stand	0%	0%	100%
GoogLe Net	Fall	49%	51%	0%
	Others	0%	98%	2%
	Stand	0%	0%	100%

表 2 比較結果

	SE	SP	AC	ER	train time	test time
(1)	100%	70.5%	80.3%	19.7%	33.0s	31min5s
(2)	100%	93.5%	95.7%	4.3%	51.6s	5s
(3)	49.0%	100%	83.0%	17.0%	1min3s	5.85s

HOG+SVM で Fall と Stand の認識率が高いが、Others で 59% の画像は Fall として認識した。その中で、Fall 姿勢の g 誤認識率は 100% であり、Sq の姿勢の誤認識率は 96% である。その二つの部分が Others の認識率を下降させる原因であると考えられる。

AlexNet は Fall と Stand の認識率が高いが、Others の部分は認識率が 51% になった。36% の Stand として間違っただ部分の中で、B 姿勢は 10%、F 姿勢は 9%、Si 姿勢は 8%、Sq は 9% である。

GoogLeNet は Others と Stand の認識率が高いが、Fall の認識率がただ半分未満であり、転倒検知として認識率が低いと考えられる。

表 1 は、SVM+HOG(1)、AlexNet(2) と GoogLeNet(3) の結果比較である。ただモデルの転倒検知の感度を考え、Stand と Others は全部で体調良好となっていた。(Fall はポジティブの結果としてみて、Stand と Others は同時にネガティブとして結果を出した。)

5. おわりに

本研究は、機械学習・深層学習を用いて、高齢者の状態を三つのグループ (Stand, Fall, Others) に分類することを目指し、高齢者の転倒などの危険を検出することにより、高齢者の見守りに貢献する。さらに、その中から最適な手法を探す。その実験の結果は、AlexNet は認識率が高く、有効性を示した。しかし、Others の認識率が高くないため、特に B と Sq 姿勢の認識率がほとんど 0% となり、特徴量の抽出方法又はモデルを改良する必要があり、今後の課題となる。

参考文献

- [1] Erdem Akagündüz et al., "Silhouette Orientation Volumes for Efficient Fall Detection in Depth Videos", IEEE J. of Biomedical and Health Informatics, Vol. 21, No. 3, pp756-763, 2017
- [2] 紙徳直生, 孟林, 山崎勝弘, "深層学習を用いた二段階甲骨文字認識", 第 18 回情報科学技術フォーラム, 2017.
- [3] 内田祐介, 山下隆義, "物体認識のための畳み込みニューラルネットワークの研究動向", 電子情報通信学会論文誌 D, Vol.J 102-D, No.3, pp.203-225, 2019.