



卒業研究報告書

平成28年度

研究題目

人感センサを用いた
要介護者の危険行動認識

指導教員 上野秀剛 講師

氏名 川上幹男

平成29年2月16日 提出

奈良工業高等専門学校 情報工学科

人感センサを用いた 要介護者の危険行動認識

上野研究室 川上 幹男

高齢や疾患のために1人で生活するのが困難な人（要介護者）が、転倒やベッドからの転落などの危険行動を起こした場合、介護者による補助が必要なことがある。危険行動の発見が遅れると要介護者は長時間1人で対処できない状態が続くため、疾患の悪化が起きる可能性がある。このため、介護者は危険行動が発生したことを迅速に認知しなければならない。しかし、介護者が要介護者に24時間付き添うことは、介護者にとって大きな負担となるため難しい。これに対し、カメラによる要介護者の監視や要介護者自身が自宅から心拍数や血圧の情報を担当医師に送信する手法が存在する。しかし、これらの手法は要介護者のプライバシーの侵害や介護者への負担、要介護者自身への負担の増大などの問題が挙げられる。

本研究の目的は、要介護者の行動をもとに危険行動が行われたか識別する手法を提案し、その識別精度を評価することである。提案手法では、要介護者の行動を識別するために人が動いた時、体温に反応して人を感じ取る人感センサを使用する。また、要介護者や介護者の負担を軽減するために、危険行動の識別には機械学習アルゴリズムの1つであるRandom Forestを使用し、自動で危険行動を識別する。本研究では、ベッド周りでの要介護者の行動を対象とし、危険行動を識別する。被験者にベッド周りで起こりうると考えた行動を行ってもらい、記録する。記録した行動を行動サンプルと呼び、これを使用してRandom Forestに学習させるデータセットを作成し、危険行動を識別する。データセットを作成する際、データセットに含まれる危険行動と非危険行動の割合が識別精度にどう影響するか各パラメータごとに比較した。

実験の結果、データセットに含まれる危険行動、非危険行動の割合は識別精度には影響しなかった。識別精度に影響する要因を分析した結果、同じパラメータで作成したデータセットでも識別精度は0%か35%以上に分かれた。35%以上の精度を示すデータセットは、2種類の行動サンプル(寝ている状態からベッドの右側に座る、ベッドの右側に立つ状態からベッドの右側に座る)から作られたデータセットであった。従って、特定の行動サンプルを含むデータセットがRandom Forestで危険行動を識別させた際、高い識別精度に影響すると期待できる。

目次

1	はじめに	2
2	関連研究	3
2.1	人感センサによる行動パターンの認識	3
2.2	カメラ画像を用いた異常動作の検出	3
3	原理	4
3.1	Random Forest	4
3.2	要介護者	4
3.3	危険行動	5
4	提案手法	7
4.1	人感センサによる行動の検知	7
4.2	危険行動の検出	7
5	実験	9
5.1	行動の記録	9
5.2	学習データセットの作成	11
5.3	Random Forest による予測	14
6	結果と考察	16
6.1	結果	16
6.2	考察	18
7	おわりに	21
	謝辞	22
	参考文献	23

1 はじめに

近年,高齢化や疾患のため日常生活に支援が必要であるとされた要介護者が増加の傾向にある [1]. 1人で生活する要介護者がベッドからの転落や転倒など,危険な行動を起こし自身で対処できない場合,介護者による補助が必要となる. 危険行動の発見が遅れると要介護者は長時間1人で対処できない状態が続くため,疾患の悪化が起きる可能性がある. このため,介護者は危険行動が発生したことを迅速に認知しなければならない. しかし,危険行動の発生を認知するために介護者が常に要介護者に付き添うことは介護者と要介護者への身体的,精神的負担が大きくなるため難しい. 危険行動の発生を介護者が付き添うことなく機械により自動で検知できるならば,検知した際に介護者が駆けつけることで介護者が常に付き添う必要がなく負担を減らすことができる.

本研究では,要介護者の行動をもとに危険行動が行われたか識別する手法を提案し,その識別精度を評価することを目的とする. 提案手法は要介護者の行動を知るために,人感センサを複数用いて人間の位置と体勢のデータを集める [2]. 得られたデータから危険行動を識別する学習器を作成する. 実際の介護現場では危険行動が発生する頻度は少ないと考えられる. そのため,学習器の作成には少ないデータから高い識別精度が期待できる機械学習アルゴリズムのRandom Forest[3]を用いる. 作成した学習器を用いて危険行動を識別し,実際の行動との比較によって提案手法の精度を検証する.

2 関連研究

2.1 人感センサによる行動パターンの認識

高齢者の危険行動を人感センサの情報から予測する際に Hidden Markov Model (HMM) を使用する手法が青木らによって提案されている [4]. 提案手法では, 人感センサを使用して得た高齢者の行動パターンから行動の種類を分類し, 各行動ごとに HMM を作成して危険行動を検出する. この提案手法では, 日常状態を HMM により表現し, 日常状態とは異なる結果が各行動ごとに作成した HMM 全てから出力されたら, 非日常状態として検出する. 非日常状態に着目することにより, 高齢者の行動によるセンサ反応だけではなく高齢者の体調変化からも危険行動を検出できる. しかし, 問題点として毎日の生活の中で行う行動の規則性から日常性を検出できる一方, トイレのように日常的に行う時間が決まっていない行動を行う場合には, 異常の検出が難しい. 本研究では危険行動を行った際のセンサ情報を機械学習で学習し, 日常性ではなく, センサ反応が危険行動と設定した動きに類似しているか判定し危険行動を検出する.

2.2 カメラ画像を用いた異常動作の検出

高齢者の異常を自動で検出する手法としてカメラ画像による検出手法が関らによって提案されている [5]. 提案手法では, 高齢者の日常をカメラにより画像として記録し, 記録した画像を学習データとして高齢者の異常を検出する. 検出の方法としては, 学習データとして得た画像と, モニタリングしている高齢者のカメラ画像の類似性に着目する. まず学習段階として, 学習データの画像を自己組織化マップ [6] により行動ごとに分類し, 結果を固有空間法を用いて固有空間上に点として抽出する. つぎに, モニタリングしている画像を固有空間法により点として抽出し, 学習データから得た固有空間上に投影する. 学習データの各点とモニタリング画像の点との距離に着目して相関を調べることで高齢者の行動が異常であるか検出する. この手法により, 携帯電話やテレビ電話のカメラなど, 安価で設置しやすい装置を用いて自動で異常を検知できる. しかし, 問題点としてカメラ画像を用いるので高齢者のプライバシーが守られていない点, 洋服や背景, 照明の条件によっては画像の抽出が上手くいかない点, 季節による服装の変化に対応しなければいけないため学習データを大量に用意しないと分類が上手くいかない点などが挙げられている. 本研究では, 要介護者のプライバシーに配慮してカメラではなく人感センサを使用し, 少ない学習データからでも精度が期待できる Random Forest を分類に用いる.

3 原理

3.1 Random Forest

Random Forest は,2001年にLeo Breimanによって提案された機械学習のアルゴリズムである[3]. Random Forestは,木構造の条件分岐で予測を行う決定木を弱学習器として作成した決定木の集団学習により予測を行い,分類の精度を向上させる. 集団学習とは,精度の高くない分類器である弱学習器を複数組み合わせることで,高精度な分類器を構築する方法である. 機械学習は,多種多様なデータの中から,規則性・パターンを発見し,予測に役立てるために考案された.

図1にRandom Forestの概要を示す. まず,学習データを用いて決定木を複数個作成する. 学習データ内の説明変数をランダムに抽出し,その説明変数を元に1つの決定木を作成する. 作成された決定木それぞれに予測対象のデータを入力する. 各決定木により算出された予測結果を平均して出力結果とする. Random Forestは少ないデータ数からでも高い精度が期待できる.

この理由として,機械学習の予測で学習データの違いに由来する誤差をランダムに抽出した説明変数から作成した決定木による集団学習を行うことで下げられるからである.

本研究では,R言語を使用してRandom Forestを実行する.

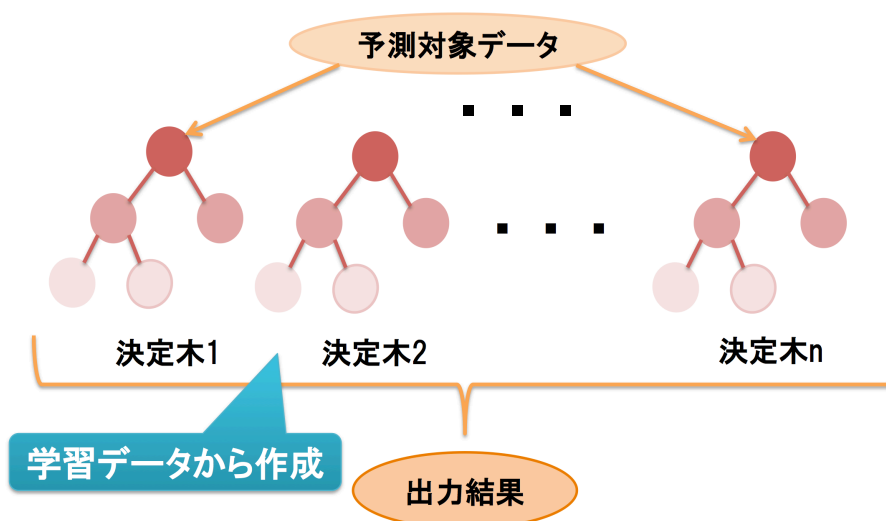


図1 Random Forestの概要

3.2 要介護者

要介護者は,身体上の障害により介護なしでは日常生活に支障があると見込まれている状態にある人を指す. 要介護者の例としては,転倒した場合,足腰が衰えているため自力では立ち上がれない人が挙げられる. 身体上の障害の原因として

は主に、高齢化による身体機能の低下が考えられる。要介護者が転倒すると自力では立ち上がれないため、転倒による怪我の治療ができなくなったり、疾患を持っている場合はその症状が悪化する恐れがあると考えられる。

そのため介護者によって要介護者が転倒から起き上がるための補助が必要となる。介護者は、要介護者の入浴補助や食事の補助に加えて転倒や転落などの危険行動が発生するごとに補助する必要がある。しかし、近年の介護現場の問題として、要介護者の人数に対し介護者の人数が圧倒的に足りていないことが挙げられる [7]。介護者が不足していることで、居住型の介護施設など複数の要介護者が生活している中で介護者が要介護者 1 人 1 人に配慮しながら介護することが困難となり要介護者が危険行動を起こしてもすぐに気づくことができない。

この問題を解消する方法として、要介護者の徘徊や転倒などの危険行動をビデオカメラで検出し、介護者に知らせるシステムが導入されている [8]。このシステムには、介護者が別室でモニターすることで複数の要介護者を同時に見ることができるというメリットがある。しかし、このシステムの問題点としてプライバシーの問題が挙げられる。介護者は要介護者をカメラで監視するため要介護者に常に監視されている不快感を与える可能性がある。また、カメラで介護者を常に監視するのは介護者側にも負担となる。本研究で提案する手法は、人感センサを使用することで介護者の監視されているという不快感を緩和することができると考えられる。加えて、提案手法は Random Forest によって自動で危険行動を検知するので介護者は常に要介護者をモニタリングする必要がなく負担が減ると考えられる。

3.3 危険行動

多くの要介護者はベッド上で寝たきり生活を行っており、要介護者は1日の大部分をベッド周りで過ごす。それに伴って、危険行動のほとんどがベッド周りで発生すると考えられる。そこで本研究では、要介護者のベッド周りの行動を対象として危険行動を検知する。

要介護者が危険な状態にあると考えられる行動は以下の2つが挙げられる。1つ目はベッドから転倒し、そのまま動かなくなっている状態である。これは、転倒によって身体を強打し要介護者が意識を失っていたり、身体の一部を損傷(骨折や脱臼)したため身動きが取れない状態が考えられる。この状態で発見が遅れると意識の回復が遅れたり怪我の重症化が起きてしまう。2つ目はベッド上で座っている状態である。ベッド上で座っている状態は要介護者が寝たきり状態から起き上がって次の行動を起こそうとしていると考えられる。しかし、要介護者は、介護なしでは日常生活を行うのが難しいため、ベッド上以外で行動することが危険行動につながりやすい。ベッドから起き上がってトイレに移動しようとした時に転倒することや起き上がってベッド上で身体の向きを変えようとした時に転落するといった例が挙げられる。本研究では、以上で述べた2つを危険行動とする。

また,内閣府の調査によると1年間寝室で転倒を経験した高齢者は,アンケートを取った高齢者のうち0.5%であった[9]. この調査より,要介護者の生活の中で,危険な行動の発生する頻度と発生しない頻度を比較した時,危険行動の発生する頻度の方が圧倒的に少ない. このため,人感センサを使用して要介護者の行動を検知した時,危険行動を検知する回数は少ないと考えられる. そこで,少ない学習データからでも高い予測精度が期待できる Random Forest を予測に用いる.

4 提案手法

本研究では,要介護者の危険行動を識別するため,センサから得た人の行動を Random Forest で学習し危険行動を識別する手法を提案する.以下,人感センサによる行動の検知方法について説明し,その後,危険行動を検出する手法について述べる.

4.1 人感センサによる行動の検知

本研究では,プライバシー侵害の問題を考慮して人感センサを用いて人間の行動を検知する.本研究で使用する人感センサは,ワンボードマイコン Arduino YUN を用いて制御する. Arduino は AVR マイコン,入出力ポートを備えた基盤,Arduino 独自の言語「Arduino 言語」とその統合開発環境から構成されるシステム全体を指す.なお,本研究においては Arduino はそのワンボードマイコンのことを言う.その Arduino の中でも本研究では,ワイヤレス通信によるデータの送受信を可能とした Arduino YUN を使用する.この理由として,有線接続の Arduino を使用すると実際の介護現場はセンサの数だけコードが必要となり要介護者がコードに引っかかって転倒する恐れがあると考えたからである. Arduino に焦電型赤外線センサーモジュール (SB412A) を接続して人感センサを作成する. SB412A は,検知距離が約 3m である.このためベッド周辺から離れた場所にいる人間を誤検知する可能性が低く,ベッド周辺のみ人間の行動を検知できる.使用する人感センサは,センシング結果を 0~255 の値で出力する.人を検知していないときは 0 を出力し,人を検知したときはセンサと人との距離や温度に応じた値が 1~255 で出力される.本研究では,得られたセンサ情報をシリアルポートを通じて PC 側で受信し,学習に用いる要介護者のベッド周りでの行動データとして記録する.使用した Arduino と接続したセンサを図 2 に写真で示す.

4.2 危険行動の検出

高齢者の危険行動は日常で頻繁に起こらない.従って危険行動の検出には少ない事例の学習からでも高い精度が期待できる Random Forest を用いる.

Random Forest の学習に使用するデータをデータセットと呼ぶ.データセットは,人感センサより取得した要介護者の行動をもとに作成する.人感センサは動く熱源に反応するため,一定時間要介護者が動かないとセンサは人間を感知しなくなる.そのため,人間はセンサの反応範囲内に存在するのに作成されるデータセットにはセンサ反応がない部分が出てしまい,これでは識別を正確に行えない.

そこで,作成するデータセットには過去のセンサ情報を含める.また,データセットに含まれる要介護者の行動データを数秒間ごとにまとめる.これは,データセッ

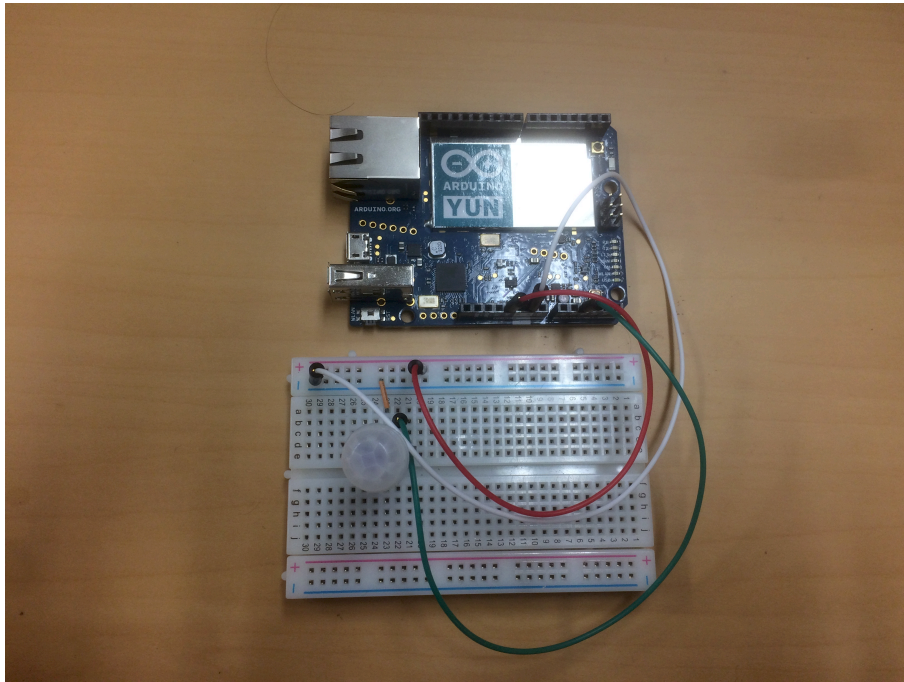


図2 作成した人感センサ

トを小さくすることで識別にかかる時間を減らし,(要介護者の危険行動をできるだけはやく発見する)ために行う。以下,本研究ではこの処理を圧縮と呼ぶ。また,作成したデータセットには,ラベルとして危険行動か非危険行動かを手動で入力し与えるものとする。

最後に,作成したデータセットを2等分し,片方を学習データとしてRandom Forestによる学習器を作成する。そして,作成した学習器によって残りのデータセットの行動を予測し,クロス集計によって危険行動の検出精度を考察する。クロス集計とは,異なる2つの項目に注目して同時に集計する集計方法である。本研究では,実際の要介護者の行動とRandom Forestによる予測結果を組み合わせるクロス集計を行う。

5 実験

1名の被験者に危険行動と非危険行動を行ってもらい,学習器を作成し,危険行動の検出精度を測った.

5.1 行動の記録

情報工学科5年生の1名に対して,所属研究室にあるベッド周りで指定した行動をそれぞれ行ってもらう.指定した行動を人感センサにより記録し学習に用いるデータセットを各行動ごとに作成する.

被験者に行ってもらう行動は,以下の11種類(危険行動:5種類,非危険行動:6種類)とする.この11種類の行動は,実際に要介護者がベッド周りで行うと考えられる行動を想定している.以下に示す"→"は,状態の変化を表す.例として,「ベッド上で寝ている→ベッド右側に座る」という表記は,ベッド上に寝ている状態からベッド右側に座る状態に移るまでの行動を表し,この一連の行動のセンサ反応をProcessingによってcsvファイルとして記録する.

- 危険行動

1. ベッド下に落ちる
2. ベッド上で寝ている→ベッド右側に座る
3. ベッド上で寝ている→ベッド左側に座る
4. ベッド右側に立つ→ベッド右側に座る
5. ベッド左側に立つ→ベッド左側に座る

- 非危険行動

1. ベッド右側に座る→ベッド上で寝る
2. ベッド左側に座る→ベッド上で寝る
3. ベッド右側に座る→ベッド右側に立つ
4. ベッド左側に座る→ベッド左側に立つ
5. ベッド下で倒れている→ベッド右側に立つ
6. ベッド下で倒れている→ベッド左側に立つ

この11種類の行動を検知できるように人感センサをベッド周りに設置する.設置場所は研究室内のベッド周りとし人感センサを9つ設置する.設置方法は要介護認定率が急激に上がる日本人65~70歳の平均身長である165cmに合わせる[10].身長がおよそ165cmの情報工学科学生1名にベッド上で寝る,座る,ベッド周りで立つなど

の行動を行ってもらい,センサの反応をPC上で表示し,センサが行動の種類ごとに異なる反応を示すか確認しながら設置する. 設置したセンサの配置を図3に示す. 図3でセンサが書かれている面はベッドに対して垂直に立っている壁を表す. センサ1~4はベッド上で人間が寝ているか起き上がっているか分かるよう記録するのに使用する. センサ5~9はベッド周りで人間が座っているか立っているのか分かるよう記録するのに使用する.

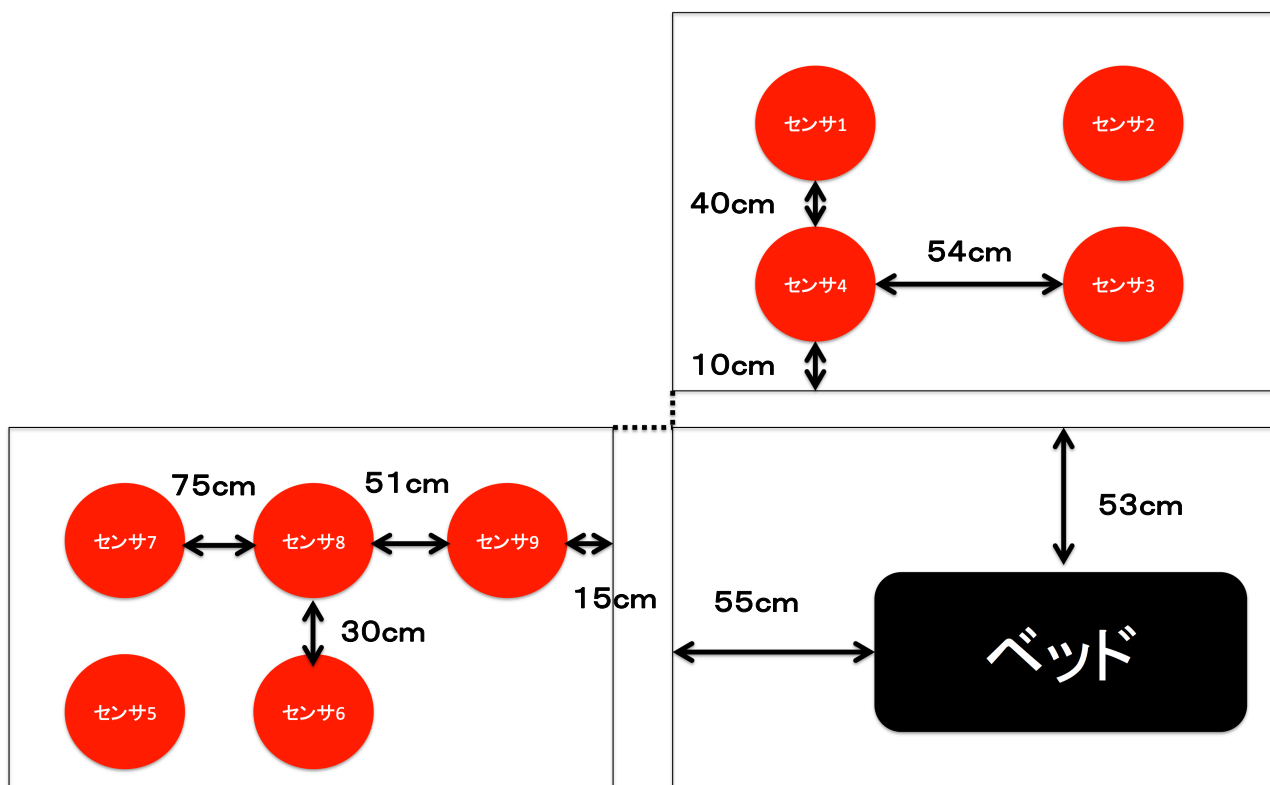


図3 人感センサの設置場所

被験者には,指定した11種類の行動を行ってもらい,設置したセンサでセンシングする. 本研究では,指定した行動をデータセット内でランダムに配置して機械学習に用いるデータを作成する. そのため,各行動ごとにデータを分ける必要があるので11種類の行動を1種類ごとに記録する. 人感センサは反応があった場合,その後,数秒間人間に動きがなくても反応が続く. そのため,計測時には,数秒間安静にしてもらい全てのセンサの反応が消えてから指定した行動を行ってもらい,行動後全てのセンサ反応が消えるまで安静にしてもらい行動の記録を行った. 人感センサにより取得した行動のデータは,Arduinoを介してPCに送る. 送られた行動データは,プログラミング言語Processing[11]によりcsvに書き込み記録する. 9種類のセンサによる行動データはPC上で別々のcsvファイルに記録する.

Random Forest による学習には,9種類のセンサの情報を使うため,各センサが出力

した値を1つのcsvファイルにまとめる必要がある。本実験では、各センサの値を記録する際に、時間(日付/時/分/秒)を併せて出力することで、9つのセンサの出力を1つのファイルにまとめる。こうして作成された1つのcsvファイルを行動サンプルと呼ぶ。実際に作成されるcsv形式でEXCEL上に表示した行動サンプルを表1に示す。表1は、被験者に指定した11種類の行動の内、「ベッド右側に立つ→ベッド右側に座る」行動を行った場合である。表中の最上行は、9個のセンサを示し、S1は図3のセンサ1、S2はセンサ2のようにSの後に続く数字が設置したセンサの番号にそれぞれ対応する。表1のようなcsvファイルを11種類の行動ごとに作成し、行動検出に使用する。

表1 作成される行動サンプル例

S1	S2	S3	S4	S5	S6	S7	S8	S9
0	0	151	0		0	0	0	144
0	154	152	153	0	0	151	150	144
149	154	151	153	0	153	152	150	144
149	153	151	153	152	154	152	150	145
149	153	151	152	152	153	152	150	0
149	154	0	153	152	0	152	150	0
149	153	0	153	152	0	152	150	0
.
.
.
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

5.2 学習データセットの作成

危険行動の検出に用いるデータをデータセットと呼ぶ。データセットは、5.1節で作成した行動サンプルを用いて作成する。データセットは、行動を記録した人物のみの結果に偏らないようにランダムに行動サンプルを配置して作成する。学習に用いるデータセットは6000秒に含む行動サンプルの割合と行動サンプル内の危険行動と非危険行動の割合をパラメータとして作成する。このように作成する理由としては、データセットのパラメータが危険行動検出の精度にどう影響するのか検証するためであり、高い精度が得られるパラメータがわかれば実際の介護現場で危険行動を検出する際に作成するデータセットのパラメータ調整に役立つと考えたためである。行動サンプル内の割合は、細かく変化させることは可能であるが全ての割合を試すには検証に膨大な時間がかかるため本研究では、10%~50%間を5%間隔で変化させる。50%よりも大きい割合は高齢者が6000秒のうち半分以上動き続ける状態が非日常的であると考えたため作成しない。また、行動サンプル内の危険行動と非危険行動の割合も、検証時間が大きくなることを考慮して危険行動:非危険行動を1:9~5:5の中を1割ごとに変化させる。5:5で止めた理由として

は,6:4以上の割合は,すでに試した4:6以下の割合から危険行動と非危険行動の割合を入れ替えたものとなるため,結果も危険行動と非危険行動が入れ替わった精度が出るものと考えられる.理由としては,危険行動か非危険行動であるかはラベルとして手動であたえるからである.そのため,本研究では危険行動:非危険行動を1:9~5:5の範囲内のみで変化させてデータセットを作成する.作成するデータセット内を図4に示す.図4は,データセットが含む行動サンプルが10%で行動サンプルの割合は危険行動1:非危険行動9を表す.データセット全体が6000秒で,そのうち10%(600秒)が行動サンプルとなる.行動サンプル内の危険行動と非危険行動の割合は1:9なので,危険行動60秒,非危険行動540秒となる.

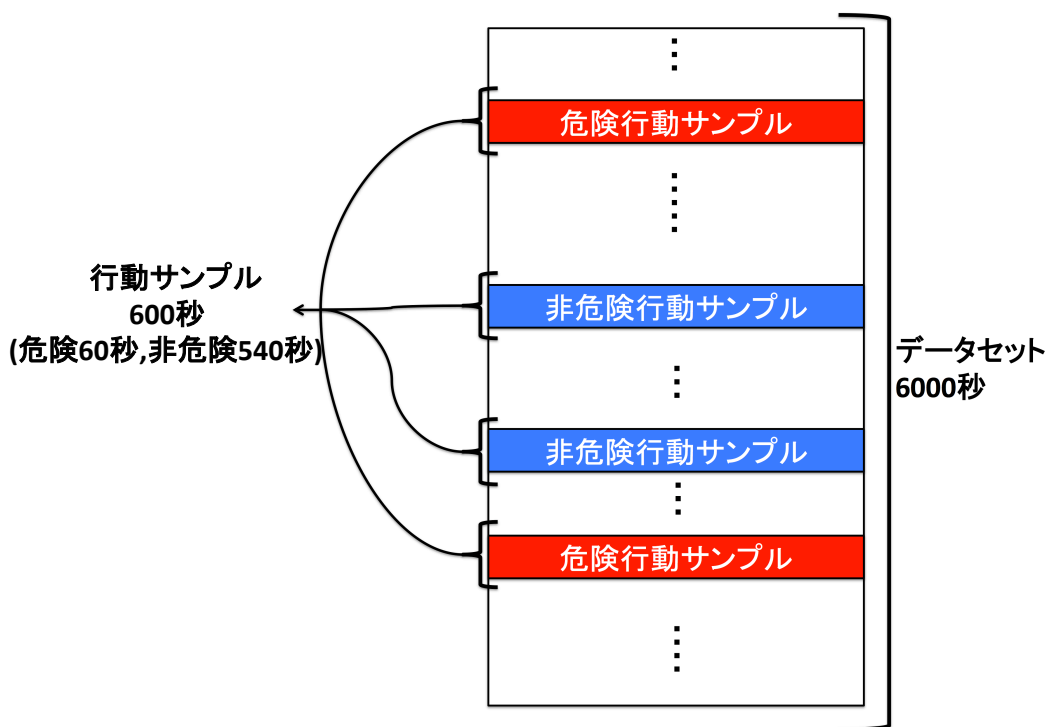


図4 データセット例(行動サンプル10%で割合が危険行動1:非危険行動9)

作成するデータセットは,5分前までの過去のセンサ情報を参照する.5分前まで使用する理由としては,5分前まで遡れば過去に人間が行動を起こしており危険行動か非危険行動かを識別する精度を向上するのに利用できるセンサ情報が得られると考えたためである.また,作成したデータセットは6秒ごとに圧縮を行う.6秒で区切る理由は,要介護者が1つの行動を行うのに約6秒かかると考え,圧縮してもデータに実際には行動していないのに,圧縮することで行動していると記録される影響が起こりにくいと考えたからである.圧縮の方法は,データセットを6秒ごとに区切り,区切った6秒間の中でセンサの反応があれば1,なければ0として記録する.このように5分前までの過去データを参照し6秒で圧縮したデータセットは表2のようなものになる.表2で,最上行がデータセットの各列の説明変数である.

S1,S9-49などのSの後につく数字は,センサの種類を表す. S1はセンサ1,S9-49はセンサ9の反応を指す. -の後の数字は,遡った過去の時間を表す. 本研究では,5分前まで過去を参照するが,6秒で圧縮するため後の数字が1つ大きくなるごとに6秒前を参照することを表す. 例として,S1-1はセンサ1の6秒前,S9-49はセンサ9の4分56秒(49×6秒)前を指す. また,最上行の1番右にあるstatusはあらかじめ入力し与える危険か非危険かを表すラベルである. Yが危険行動,Nが非危険行動を表す. Random Forestでは,このstatusの列をYかNか予測し精度を検証する.

表2 作成するデータセット例

S1	...	S9	S1-1	...	S9-1	S1-49	...	S9-49	S1-50	...	S9-50	status
0		0	0		0		0		0	0		0	N
0		0	0		0		0		0	0		0	N
0		0	0		0		0		0	0		0	N
1		1	0		0		0		0	0		0	Y
1		1	1		1		0		0	0		0	Y
0		0	1		1		0		0	0		0	N
0		0	0		0		0		0	0		0	N
■ ■ ■													
0		0	0		0		0		0	0		0	N
0		0	0		0		0		0	0		0	N
0		0	0		0		0		0	0		0	N
1		1	0		0		0		0	0		0	N
1		1	1		1		0		0	0		0	N
0		0	1		1		0		0	0		0	N
0		0	0		0		0		0	0		0	N

本研究では,6000行のデータを6秒で圧縮して作成するので1000行のデータセットが表2の形式で作られる. 同じパラメータで複数のデータセットから学習し出した精度を平均してそのパラメータでの危険行動予測精度を決定する. これは,1回の学習による予測精度だけでは,精度に有意性があるか判別できないからである. 本研究では,同じパラメータで100回ずつ学習し求めた危険行動の予測精度を平均し精度を求める. しかし,同じパラメータで100個のデータセットを作成するのは膨大な時間がかかるため同じパラメータのデータセットを10個ずつ作成し,それぞれ10回ずつ予測に使用する. 従って,本実験では機械学習の予測用に計450個のデータセットを作成する. 図5に作成するデータセットを図示する. 縦の並びは異なるパラメータのデータセットを表し,横の並びは各パラメータで作成したデータセット10個を表す. データセットの下にある%表記がデータセットに含まれる行動サンプルの割合を示す. その横にある1:9や2:8はデータセット内の行動サンプルに含まれる危険行動と非危険行動の割合を表す. データセット名の一番下に描く丸付き数字は同じパラメータで作成したデータセットを判別するためのIDである.

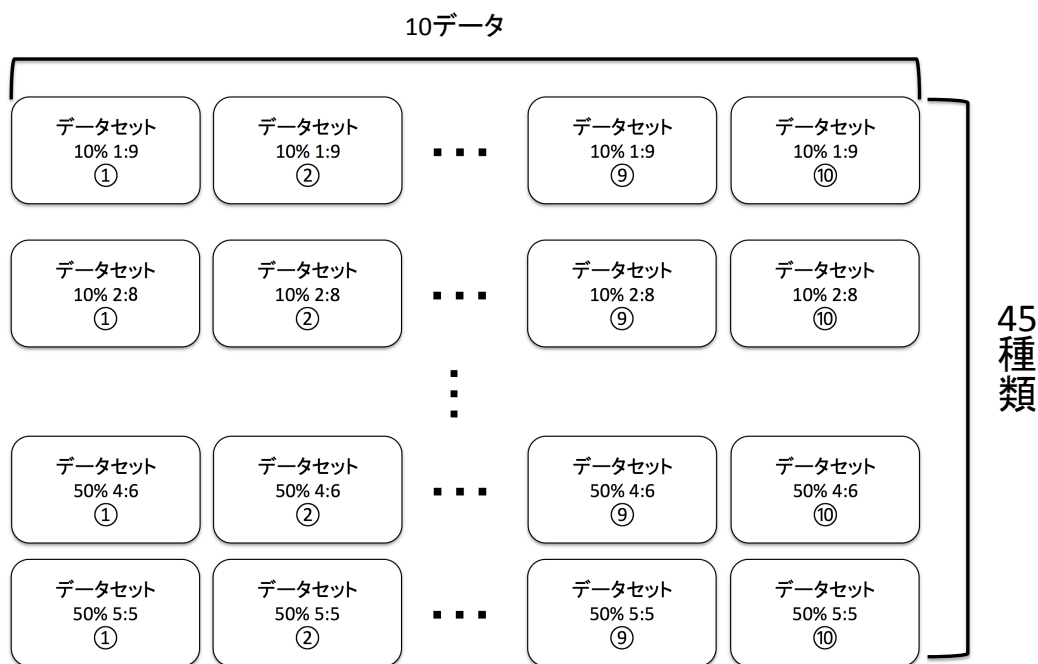


図5 データセット

5.3 Random Forest による予測

作成したデータセットをRandom Forestによって学習し危険行動の検出を行う。データセットの半分を学習用データとし、残りを予測対象とし予測する。多くの機械学習法では、乱数を使用しているため、同一のデータセットに対して予測を繰り返しても精度が同じになるとは限らない。本研究では、100回分（10データセット×10回）の精度の平均をそのパラメータでの精度とする。学習器による予測結果は、表3のようになる。

表3 予測精度の算出法

		実際の行動	
		非危険行動	危険行動
予測結果	非危険行動	A	B
	危険行動	C	D

表3を用いて予測精度の求めかたについて説明する。表3中の項目が実際の行動と予測結果でのクロス集計の結果を表す。例として、実際の行動が非危険行動でRandom Forestによる予測結果が危険行動と出力された数はCとなる。予測精度は、全体の行動の正答率、危険行動のみの正答率、非危険行動のみの正答率の3種類を求める。ここでは、各正答率の評価には、評価尺度 *accuracy*(正確度), *recall*(再現率), *fallout*(フォールアウト)を用いる。

まず,*accuracy* は,危険行動と非危険行動両方の正答率を表し,A,B,C,Dにより式(1)のように求める.

$$accuracy = \frac{A+D}{A+B+C+D} \quad (1)$$

次に,*recall* は,実際の危険行動の中でどれだけ正しく予測したかを表し,式(2)のように求める.

$$recall = \frac{D}{B+D} \quad (2)$$

最後に,*fallout* は,実際の非危険行動の中でどれだけ正しく予測したかを表し,式(3)のように求める.

$$fallout = \frac{A}{A+C} \quad (3)$$

accuracy,*recall*,*fallout*をそれぞれ求めることで,全体の正答率と危険行動,非危険行動の正答率がどう関係しているかを考察するとともに各パラメータ(例えば,データセット内の行動サンプルが10%で行動サンプルの割合比は危険1:非危険9)ごとの*accuracy*,*recall*,*fallout*を比較することで危険行動の検出に最適なデータセットを考察する.

6 結果と考察

5.3節で述べた精度の求めかたに基づいて、各パラメータでの学習器の精度を評価する。

6.1 結果

各パラメータで得られた行動検出の精度を表4に示す。各行はデータセットに入れた行動サンプルの割合を示し、各列は行動サンプルの危険、非危険の割合を表している。精度の単位はパーセントで表し、全ての予測結果のうち実際の行動と一致している予測の割合を示す。行動サンプルの割合に着目すると、割合が変化しても行動の正答率には影響していないことが分かる。また、使用する行動サンプルの内、非危険行動の割合が高いものほど予測精度は高くなっている。次に、危険、非危険それぞれのみでの正答率を表5、表6に示す。表4、表5、表6において、最も左の列はデータセットに含んだ行動サンプルの割合を示す。

表4 行動の予測精度

	危険1:非危険9	危険2:非危険8	危険3:非危険7	危険4:非危険6	危険5:非危険5
10%	88.1	73.8	64.8	59.1	57.2
15%	89.5	79.0	67.3	59.8	57.1
20%	89.2	77.7	65.1	58.3	56.3
25%	90.4	79.0	71.4	62.2	56.5
30%	90.5	79.7	68.5	60.7	58.5
35%	90.1	79.5	69.7	60.4	58.5
40%	90.0	78.7	70.0	62.9	59.4
45%	88.8	78.6	70.4	62.5	58.0
50%	89.2	78.9	68.3	59.9	56.8

表5 危険行動のみの予測精度

	危険1:非危険9	危険2:非危険8	危険3:非危険7	危険4:非危険6	危険5:非危険5
10%	5.20	7.00	22.9	35.7	60.4
15%	0.36	4.40	22.9	29.7	59.9
20%	0.32	11.6	28.0	40.1	61.0
25%	0.31	10.6	17.9	46.4	56.4
30%	0.95	4.84	22.5	35.1	56.0
35%	0.24	7.04	12.8	35.5	54.8
40%	0.44	2.48	12.8	36.1	60.4
45%	3.24	1.08	10.4	31.2	57.2
50%	0.00	0.77	8.19	31.1	55.7

表6 非危険行動のみの予測精度

	危険1:非危険9	危険2:非危険8	危険3:非危険7	危険4:非危険6	危険5:非危険5
10%	97.5	94.5	80.9	69.1	47.3
15%	99.7	96.9	85.8	75.8	49.5
20%	99.6	95.2	83.3	69.1	48.7
25%	99.6	96.3	91.5	73.0	54.9
30%	99.7	97.5	88.2	76.3	60.6
35%	99.9	97.7	93.1	76.1	59.1
40%	99.9	98.5	93.8	80.6	57.1
45%	99.7	99.4	95.7	83.1	57.6
50%	100	99.6	95.5	80.4	57.3

結果より、全体の正答率が高いもの程、非危険行動の正答率は高く、危険行動の正答率は低い。このことから、行動サンプルとして含む危険行動の割合が低いほど全体の予測精度は高くなるが、危険行動の予測精度は低くなる。しかし、これは危険行動の割合が低いと予測対象に含まれる危険行動の割合も低くなるため精度が低くなっているといえる。つまり、データセットに含まれる危険行動と非危険行動の割合が学習結果にそのまま反映されていて、学習データとして与えた情報しか予測できていない可能性がある。この結果より、行動サンプルを使用して作成したデータセットによるRandom Forestの学習は過学習を行っていると考えられる。過学習とは、機械学習において学習データに対しては識別されるがそれ以外の未知データは識別できないような学習器の学習のことを指す。従って、本実験で作成したRandom Forestでは行動サンプルの割合の変化が予測精度に影響していないといえる。

6.2 考察

実験結果より,各パラメータに含まれる行動サンプルの割合は精度に影響しないことが分かった.そこで,本考察では予測精度に影響を与えているパラメータについて考察する.

まず,1つのパラメータ内での学習器の出力にばらつきが存在するか検証する.表7に例として,データセットに含む行動サンプルの割合が10%,危険行動1:非危険行動9のパラメータでの学習精度を示す.表7で,データのIDは1つのパラメータで作成した10個のデータセットのそれぞれを示し,精度は各データセットを用いて危険行動,非危険行動の両方合わせた予測の精度を10回分の平均で示す.

表7 学習パラメータ (10%, 危険行動1:非危険行動9)での予測精度

データのID	10回予測の平均精度 (%)
学習データ (1)	71.6
学習データ (2)	92.3
学習データ (3)	82.5
学習データ (4)	95.3
学習データ (5)	87.5
学習データ (6)	91.6
学習データ (7)	98.4
学習データ (8)	80.8
学習データ (9)	94.3
学習データ (10)	86.1
平均	88.04

表7より,精度は70%台~90%台でばらつきがあることが分かる.このばらつきがデータセットに含まれる説明変数によるものなのか探るために寄与率に着目する.寄与率とは,学習に用いたデータセットの各構成要素(本研究では,S1やS4-49のように表されている)が予測精度にどれだけ貢献しているか示す指標である.ここでは,70%台の精度がある学習データ(1)と90%台の精度がある学習データ(2)を比較する.

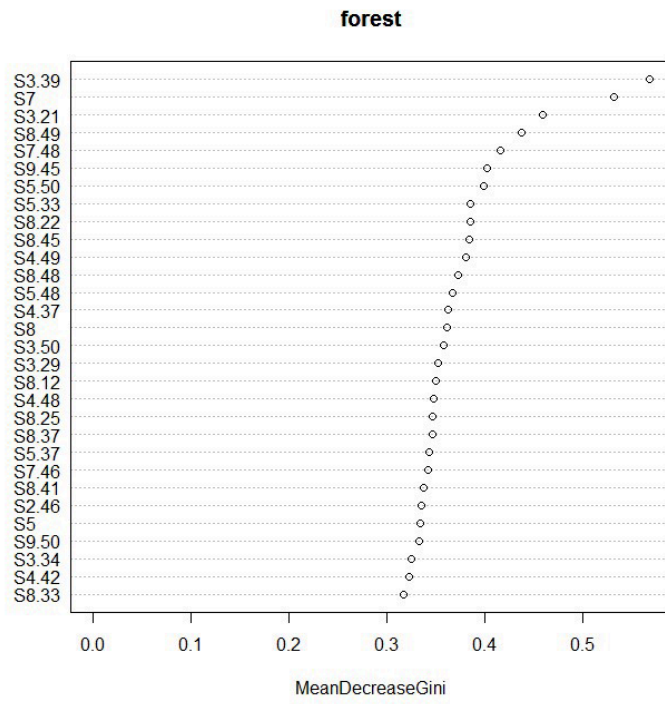


図6 学習データ(1)の寄与率

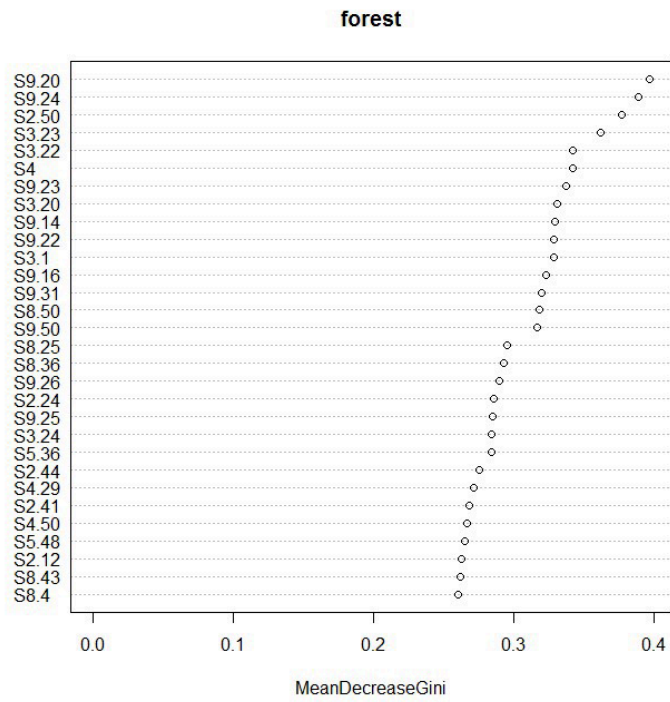


図7 学習データ(2)の寄与率

学習データ(1)と学習データ(2)の寄与率をそれぞれ図6,図7に示す. 図の縦軸は, 学習に用いた説明変数を寄与率が高い順に表す. 変数名Sの直後の数字は, センサの番号を示し1~9で表す. "."の後の数字は, さかのぼった過去分の数を表す. また, 横軸は寄与率を表す. 図6,図7で高い寄与率を示す説明変数は, どちらもセンサ, 過去の時間ともに傾向がみられない. このことより, 特定の説明変数が精度に関係している可能性は低い.

説明変数が精度に関係していないことがわかったので次に危険行動と非危険行動の予測精度にばらつきがあるのか調べる. 学習データ(1)と学習データ(2)の危険行動と非危険行動ごとに求めた精度を表8にまとめる.

表8 危険,非危険行動別の予測精度

	学習データ(1)	学習データ(2)
危険行動	59.4	0.00
非危険行動	76.3	100
平均	67.9	50.0

表8より全体の精度が70%台の学習データ(1)は危険行動を60%の精度で予測しているのに対し,90%台の学習データ(2)は危険行動の予測精度が0%で予測できていないことが分かる. 寄与率に共通点がないことは分かっているので,残っている学習データ(1)と(2)の異なる点は学習サンプルに用いた行動サンプルである.

そこで,学習に用いた行動サンプルに着目する. 学習データ(1)が含んでいる危険行動のサンプルは,「ベッド上で寝ている→ベッド右側に座る」,「ベッド右側に立つ→ベッド右側に座る」の2種類であった. 一方で,この2種類の行動サンプルは学習データ(2)に含まれていない. 以上のことから,学習データ(1)に含まれる2種類の行動サンプルが危険行動の予測精度を上げる要因になっていると考えられる. 従って,特定の行動サンプルがRandom Forestによる危険行動の予測精度の向上に関連していると考察できる.

7 おわりに

本研究では,要介護者の行動をもとに危険行動が行われたかをRandom Forestを用いて識別する手法を提案し,その識別精度を評価することを目的とした.

提案手法は,Arduino YUNと人感センサを用いて要介護者の行動を記録し,記録した行動からデータセットを作成しRandom Forestに学習させ危険行動を検出する.

実験では,情報工学科生1名に危険行動と非危険行動を行ってもらう.その行動を人感センサで記録して作成したデータセットからRandom Forestによって行動の正答率を求める.その結果,全体の正答率が高いものほど危険行動の正答率は低く,全体の正答率が低いものほど危険行動の正答率は高くなっていた.

考察として,正答率に影響を及ぼしているパラメータを分析した結果,特定の行動サンプルが学習の精度を高めているであろうということが考えられる.

今後の課題としては,特定の行動サンプルの影響の立証や要介護者の行動を記録するための人感センサの設置場所の変更,データセットを作成する際のパラメータの変更による予測精度への影響の評価が挙げられる.また,他の機械学習を用いた際の予測精度とRandom Forestの予測精度の比較,ベッド周り以外の場所での危険行動予測などの課題も挙げられる.本研究では,情報工学科生1名に対してのみ行動を記録し機械学習により行動の識別を行った.しかし,年齢が20歳前後の人間と高齢者の行動パターンには差があるものと考えられる.そこで,今後,本研究で提案した手法を使用して実際の要介護者の危険行動識別を行い有用性を検証することも今後の課題として挙げられる.

謝辞

本論文の執筆および研究を進めるに当たり、多くの方々に協力して頂きました。この場をお借りしてお礼を申し上げます。ありがとうございました。

指導教員である上野秀剛講師には、この1年間を通して研究に関する知識のご教授から、アドバイス、論文のチェックまで多くの面でご多忙の中、ご指導頂きました。また、大学編入における推薦書の執筆や面接練習も受けて下さり、アドバイスを頂くなど1年間支えていただきました。ここに深謝の意を表明させていただきます。ありがとうございました。

査読教員である松村寿枝准教授には、ご意見・ご指摘を頂きました。ここに深謝の意を表明させていただきます。ありがとうございました。

同じ上野研究室であった皆さまには、発表での至らぬ点の指摘や研究に対するアドバイスなどの面で支えていただきました。また、新入生歓迎会や研究室旅行などの様々な行事も主催して頂き楽しませてくださいました。ここに深謝の意を表明させていただきます。ありがとうございました。

同級生の皆様には、研究や課題に追われる中、被験者として研究実験に協力していただきました。ここに深謝の意を表明させていただきます。ありがとうございました。

参考文献

- [1] 厚生労働省 介護保険事業状況報告,<http://www.mhlw.go.jp/topics/0103/tp0329-1.html>(2017-1-12 参照)
- [2] Arduino-Home,<https://www.arduino.cc/>(2017-1-12 参照)
- [3] L. Breiman, "Random Forest", Machine Learning, Vol.45, No.1, pp.5-32, (2001).
- [4] 青木茂樹, 大西正輝, 小島篤博, 福永邦雄, "独居高齢者の行動パターンに注目した非日常状態の検出", 電気学会 E 部門, Vol.125, No.6, (2005).
- [5] 関弘和, 堀洋一, "高齢者モニタリングのためのカメラ画像を用いた異常動作検出", 電気学会産業応用部門, Vol.122, No.2, (2002).
- [6] T. Kohonen, "The Self-Organizing Map", Proc. IEEE, Vol.78, No.9, pp.1464-1480, (1990).
- [7] 介護人材の確保について-厚生労働省,<http://www.mhlw.go.jp/file/0000062879.pdf>(2017-1-12 参照)
- [8] 小規模介護施設向け介護システム ラムロックシステム,<http://www.ramrock.co.jp/ramrocksys/ram-mini/>(2017-1-12 参照)
- [9] 平成17年版高齢社会白書-内閣府,<http://www8.cao.go.jp/kourei/ishiki/h17sougou/19html/2syou-2.html>(2017-1-12 参照)
- [10] 平成 27 年版高齢社会白書-内閣府,<http://www8.cao.go.jp/kourei/whitepaper/w-2015/html/zenbun>(2017-1-12 参照)
- [11] Processing,<https://processing.org/>(2017-1-12 参照)