

修士論文概要書

Master's Thesis Summary

Date of submission: 01/24/2022

学科名 Department	情報理工・ 情報通信専攻	氏名 Name	奥野 彩佳	指導 教員 Advisor	渡辺 裕 
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	5120F012-8		
研究題目 Title	乳幼児の睡眠姿勢の分類と姿勢推定モデルによる骨格推定 Classification of Sleeping Postures in Infants and Skelton Estimation with Pose Estimation Model				

1. まえがき

近年、乳幼児のうつぶせが窒息死や乳幼児突然死症候群を引き起こすとして問題視されている。厚生労働省の調べによると[1]、不慮の事故のうち窒息による死亡率が大半であり、その原因として睡眠姿勢のうつぶせが挙げられている。また、乳幼児突然死症候群とは予兆や病気をもっていない乳幼児が睡眠中に突然に亡くなる病気で、正確な原因は解明されていないが、うつぶせに寝ること関連があるとされている。

また、近年スマートフォンなどの単眼カメラの映像から骨格情報をもとに人の動きを認識し分析する技術や予測値とターゲット値の関係をトレーニングによって見つける分類アルゴリズムが広く研究されている。

以上から、本論文では乳幼児の画像から睡眠姿勢の分類を行う。また、既存の姿勢推定モデルでは乳幼児の検出が難しいことを踏まえ、乳幼児の検出精度が高くなるモデルを作成する。本論文の乳幼児睡眠姿勢の分類処理過程を図1に示す。また、睡眠姿勢精度向上のための提案手法を図2に示す。



図1 乳幼児睡眠姿勢の分類処理過程

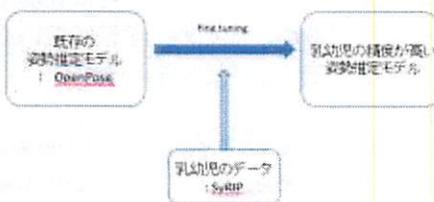


図2 睡眠姿勢精度向上のための提案手法

2. 関連研究

2.1 分類アルゴリズム

分類アルゴリズムとは予測値とターゲット値の関係をトレーニングによって見つける手法である。本研究では Support Vector Machine (SVM), 決定木, random forest, k近傍法 (k-NN), LightGBM の5種類の分類手法を用いる。

2.1 OpenPose

OpenPose はカーネギーメロン大学の Zhe Cao らが提案した深層学習を用いて人物から身体座標を取得する手法である[2]。画像平面上での二次元座標として、15, 18, 25 個の人物の関節位置を算出可能である。

2.2 fine tuning

fine tuning とは学習済みモデルをベースに出力層などを変更したモデルを構築し、自前のデータでニューラルネットワーク・モデルの結合パラメータを学習させる手法である。

3. 提案手法

3.1 乳幼児の睡眠姿勢の分類

画像中の乳幼児の関節座標から特徴量を抽出し、分類アルゴリズムを用いて乳幼児の睡眠姿勢の分類を行う。分類する姿勢は“仰向き”, “横向き”, “うつ伏せ”の3種類に大別する。乳幼児の関節情報から特徴量を算出する、使用する特徴量として、右肘, 右肩, 左肘, 左肩, 右尻, 左尻, 右膝, 左膝の8個の関節角度を用いる。各姿勢の特徴量を学習する手法として2.1で述べた5種類の分類アルゴリズムを用いる。

3.2 OpenPose による乳幼児の骨格推定

OpenPose での乳幼児の検出は成人への検出よりも精度が劣る。理由として、学習データに乳幼児が多く含まれていないためと考えられる。そこで既存の姿勢推定手法: OpenPose に乳幼児の画像をもとに fine tuning を行い、既存モデルから精度向上を図る。fine tuning を行う乳幼児のデータセットとして SyRIP (synthetic and real infant pose)[3]を使用する。このデータセットは、座っている、転がっている、這っているなどの様々なポーズをとった RGB の乳幼児画像であり、中には実画像と合成画像を含む。データセット内には 700 枚の実際の乳幼児の画像とそれに対するアノテーションと 1000 枚の乳幼児の合成画像が含まれている。

4. 実験

4.1 乳幼児の睡眠姿勢の分類

SyRIP の実画像のうち“仰向き”, “横向き”, “うつ伏せ”の姿勢をしている 229 枚を使用する。これらの画像を仰向き, 横向き, うつ伏せのポーズにラベル付けをする。これらのラベルと各画像の関節情報から算出した各関節の角度の特徴量を用いて、分類アルゴリズムに

よって分類を行う。

画像から乳幼児の姿勢を検知する手法で顔の向きを使用しているものがある。この手法と比較するため、顔全面が映っている場合(左右の目が見えている場合)を“仰向き”、顔の半分が映っている場合(左右の目どちらかが見えている場合)を“横向き”、顔全面が映っていない場合(目が映っていない場合)を“うつ伏せ”として分類し、その精度を算出した。

目から姿勢を判断した場合と分類アルゴリズムを用いたときの分類精度を表1に示す

分類手法	分類精度[%]
顔の向きからの判別	45.0
SVM	69.6
決定木	50.7
Random forest	58.1
k-NN	70.6
LightGBM	76.1

表1より、どの分類アルゴリズムでも顔の向きのみからの判別より高い精度で分類できていることが分かる。また分類アルゴリズムの中では LightGBM が最も高い精度で分類できることが分かる。

4.2 OpenPose による乳幼児の骨格推定

既存の OpenPose のネットワークモデルに対して、学習率、モーメントを変更した。学習データが 1200 枚(SyRIP)+8500 枚(COCO)の 9700 枚、検証データが 500 枚(SyRIP)+2550 枚(COCO)の 3050 枚に設定し、15epoch として学習を実行した。

従来モデルから fine tuning 後の検出結果の変化の例を図3で示す。左側が従来モデルの検出結果で右側が提案モデルの検出結果である。

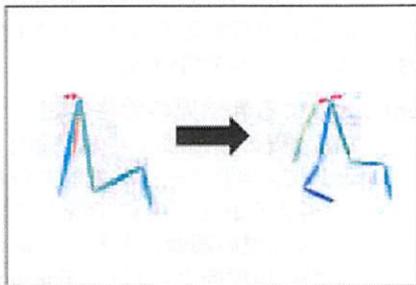


図3 従来モデルから fine tuning 後の検出結果の変化

図3に写る乳幼児はカメラと反対向きを向いてハイハイしており、目、鼻、右肘、右手首は見えていない姿勢である。図3では従来モデルより fine tuning 後のモデルでは右肩、左肩、左肘、左手首、左膝、左足首を正しく検出できている。他の画像においても、従来モデルより fine tuning 後のモデルの方がおおよそ正しく検出できていることが分かった。

本実験の評価に SyRIP の実画像 350 枚を使用する。この画像中に写る人は乳幼児 1 人になっており、乳幼児の姿勢は睡眠姿勢だけでなく立っているもの、ハイハイの姿勢のものも含まれる。

使用するデータセットの画像内の乳幼児の 18 点の関節を“検出”、“未検出”、“誤検出”、“検出不可”と判別し、総関節数からの“検出”の割合を算出した。“未

検出”とは関節が画像内に写っているが、検出できていないものを指す。“誤検出”とは関節が画像内に写っており、検出しているが、検出位置が正しい位置でなかったものを指す。“検出不可”は画像内に関節点が写っていない場合を指す。関節の検出数と成功率を表2に示す。

表2 モデルの関節検出数と検出成功率

	学習前のモデル	fine tuning 後のモデル
検出成功数[点]	3837	4293
未検出数[点]	415	124
誤検出数[点]	1039	874
検出不可数[点]	1009	1009
検出成功率[%]	72.5	81.1(+8.6[%])

SyRIP データセットに対するモデルの検出成功数と検出率が従来モデルより fine tuning 後のモデルの方が 8.6%向上していることが分かる。すなわち、従来手法を上回る精度を達成できた。

5. 結論

本研究では、乳幼児の睡眠時の姿勢分類に骨格座標から求めた特徴量を使用し、分類アルゴリズムによって判別する提案をした。従来手法である顔の向きのみから判別する手法よりも高い精度で分類できることが確認できた。

また、乳幼児のデータを使用し、姿勢推定モデルである OpenPose に対して fine tuning を適用する手法を提案した。乳幼児の画像に対し評価し、従来モデルよりも提案手法のモデルの方が、高い検出精度を与えることを確認できた。

実際のうつ伏せ検知に使用するにはより高精度の分類が必要である。また、乳幼児の姿勢推定を本実験での目的である睡眠姿勢の検出やその他の行動認識で使用するには、さらに高精度の検出精度が必要である。

分類精度向上には、本研究での提案手法に加えて時系列情報、体の向き情報などの特徴量を追加することで分類精度が高められると考える。また、検出精度向上にはデータ量を増やすことが必須である。本研究での実験において見られた傾向として、特に乳幼児の下半身部分の未検出が多かったため、学習に用いるパラメータを工夫することが必要だと考えられる。

参考文献

- [1] 消費者庁消費者安全課, “子どもの不慮の事故の発生傾向～厚生労働省「人口動態調査」より～,” 5 3 2021. Available: https://www.caa.go.jp/policies/policy/consumer_safety/meeting_materials/assets/consumer_safety_cms205_210305_03.pdf. [アクセス日: 21 1 2022].
- [2] Z. Cao, T. Simon, S-E Wei, Y. Sheikh, “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), No.121, pp.1302-1310., July 2017.
- [3] X. Huang, N. Fu, S. Liu, S. Ostadabbas, “Invariant Representation Learning for Infant Pose Estimation with Small Data,” 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021), , 15-18 Dec 2021.

2021 年度

早稲田大学大学院基幹理工研究科情報理工・情報通信専攻 修士論文

乳幼児の睡眠姿勢の分類と姿勢推定モデルによる骨格推定

Classification of Sleeping Postures in Infants and

Skelton Estimation with Pose Estimation Model

奥野 彩佳

(5120F012-8)

提出日：2022 年 1 月 24 日

指導教員：渡辺 裕 教授

研究指導名：オーディオビジュアル情報処理研究

目次

第1章	序論.....	1
1.1	研究の背景.....	1
1.2	本研究の目的.....	1
1.3	本論文の構成.....	2
第2章	従来手法.....	3
2.1	まえがき.....	3
2.2	乳幼児の睡眠姿勢検知手法.....	3
2.3	行動認識手法.....	4
2.4	むすび.....	5
第3章	関連技術.....	6
3.1	まえがき.....	6
3.2	分類アルゴリズム.....	6
3.2.1	Support Vector Machine.....	6
3.2.2	決定木.....	7
3.2.3	random forest.....	7
3.2.4	k近傍法.....	8
3.2.5	LightGBM.....	9
3.3	OpenPose : Realtime Multi-Person Pose Estimation.....	9
3.4	fine tuning.....	12
3.5	むすび.....	12
第4章	提案手法.....	13
4.1	まえがき.....	13
4.2	睡眠姿勢の分類.....	13
4.2.1	睡眠姿勢の分類手法の概要.....	13
4.2.2	分類する姿勢.....	13
4.2.3	分類アルゴリズムによる分類.....	15
4.3	OpenPoseによる乳幼児の骨格推定.....	16
4.3.1	OpenPoseによる乳幼児の骨格推定手法の概要.....	16
4.3.2	OpenPoseによる乳幼児の骨格検出.....	17
4.3.3	fine tuningに用いる乳幼児のデータセット.....	17
4.4	むすび.....	19
第5章	実験.....	20
5.1	まえがき.....	20

5.2	睡眠姿勢の分類実験.....	20
5.2.1	データセット	20
5.2.2	実験結果及び考察	20
5.3	OpenPose による乳幼児の骨格推定実験.....	23
5.3.1	fine tuning 後のモデルの結果	23
5.3.2	評価の際に使用するデータセット	26
5.3.3	モデルの精度評価	26
5.4	むすび.....	27
第 6 章	結論と今後の課題.....	28
6.1	結論	28
6.2	今後の課題.....	28
	謝辞	29
	参考文献	30
	図一覧.....	31
	表一覧.....	32
	研究業績	33

第1章 序論

1.1 研究の背景

近年、乳幼児の不慮の事故による死亡者数は減少傾向にあるが、厚生労働省の調べによると [1]、不慮の事故が死因の上位にある。その中でも窒息による死亡率が大半である。

窒息死はうつぶせになることで乳幼児の顔が枕や敷布団・掛け布団などの寝具に埋まってしまうこと、掛け布団などが首に巻き付いてしまうこと、ベッドから転落してしまうことが原因としてあげられる。乳幼児はうつぶせによって窒息状態になっても寝返りを簡単に行えないため、自力で窒息状態を改善できない。よって、乳幼児期にうつぶせ状態にならないように保護者は見守る必要がある。

また、うつぶせによって乳幼児突然死症候群(SIDS)を引き起こすとされている。こちらも、厚生労働省の調べで0歳の死因第4位になっている。SIDSとは予兆や病気をもっていない乳幼児が睡眠中に突然に亡くなる病気であり、原因は解明されていないが、うつぶせに寝ることと関連があるとされている。

一方、近年スマートフォンなどの単眼カメラの映像から骨格情報をもとに人の動きを認識し分析する技術や予測値とターゲット値の関係をトレーニングによって見つける分類アルゴリズムが広く研究されている。

以上のような背景から、本研究では動画像から乳幼児の骨格情報を用いて姿勢の分類を目指す。

1.2 本研究の目的

乳幼児の保育者・監督者は乳幼児がどのような姿勢をしているかを常に見守ることが必要だが、常に見守ることは保育者・監督者の負担が大きくなる。そこで、人間が目視で確認するのではなく、スマートフォンなどの単眼カメラで撮影された映像をもとに姿勢を推定する手法を用いることで、家庭や保育園や病院などで、保育者の負担を軽減することができる。また、乳幼児の映像からわかる情報から姿勢を分類し、危険な姿勢をしているときにアラートすることができれば窒息事故やSIDSを防ぐことができる。

そこで本研究では、寝ている乳幼児の安全確保のため、動画像から乳幼児睡眠姿勢の分類を行う。

一方で、既存の姿勢推定モデルでは乳幼児の骨格検出が難しいことが問題点として挙げられる。理由としては学習データに乳幼児が含まれていないため、体の比率が成人と異なり、ポーズが多様である乳幼児の姿勢推定の精度が下がると考えられる。

以上より本論文では、乳幼児の姿勢検出精度を向上させる姿勢推定モデルの作成を目的としている。

1.3 本論文の構成

以下に本章以降の構成を示す。

- 第1章 本章であり、研究の背景及び目的について述べている。
- 第2章 本研究の従来手法について述べる。
- 第3章 本研究で使用する関連技術について述べる。
- 第4章 本研究で提案する手法について述べる。
- 第5章 提案手法の実験、結果及び考察について述べる。
- 第6章 本論文の結論と今後の課題について述べる。

第2章 従来手法

2.1 まえがき

本章では、本研究の従来手法、従来研究について述べる。まず、乳幼児の姿勢を推定する従来手法について述べる。次に、従来の人物の行動認識手法について述べる。

2.2 乳幼児の睡眠姿勢検知手法

乳幼児の姿勢を検知する手法として乳幼児にデバイスをつける手法、床にシート状のセンサをひく手法、ベビーモニターの普及もあり、画像や映像から動作を検知する手法があげられる。

おむつや衣類にデバイスを付ける手法の使用例は図 2.1 の図のようなものである。こちらの手法は正確な測定が期待できるが、センサ自体が外れてしまう危険性やセンサを取り付けられるような衣類の制限、また乳幼児に違和感や負担を与えてしまうことがデメリットとして挙げられる。図 2.1 は <https://plusstyle.jp/blog/item?id=168> から引用している [2]。



図 2.1 デバイスを付ける手法例

同様に、シート状のセンサもデバイス同様に正確な測定が期待できるが、乳幼児が移動しうる広範囲に敷く必要があり高価になるということがデメリットとして挙げられる。

画像や映像から動作を検知する手法では顔の向きを使用しているものがある。画像や動画からの検知はセンサと異なり外れてしまう危険性がないことや、衣類の制限がないこと、また乳幼児に違和感や負担を与えてしまう懸念点がないこと、一般的なカメラを用いる場合コストがかからないことがメリットとして挙げられる。図 2.2 は顔の向きのみをつかう手法であり、緑のフレームで囲われている場合は仰向け、黄色が横向き、青がうつぶせと検知されている。顔のみを検出に使用している場合、横向きやうつぶせであっても顔が見えている状態だと仰向けと検知してしまうことや、横向きやうつぶせでなくても顔の角度や顔が手でおおわれている場合にうつぶせと検知してしまう不具合が発生する。

図 2.2 は https://service.edgematrix.com/application/babydetection_neut/ から引用している [3]。



図 2.2 画像から姿勢推定している例

2.3 行動認識手法

人物行動認識の手法として、TOF カメラや Kinect などの深度カメラを用いたものが挙げられる。TOF カメラとは光の反射時間により奥行きなどの 3 次元情報を計測可能なカメラである。また、Kinect とは RGB カメラ・深度センサ・マルチアレイマイクロフォンからなるセンサである。これらの深度カメラは特殊なカメラなため、本論文が対象とする家庭での行動認識には適さない。

次に人物行動認識の手法として、ディープラーニングを用いて行動認識をする手法がある。株式会社アジラ [4] は人物の姿勢を推定する Convolutional Neural Network と Long short-term memory を組み合わせることで人物の行動を予測している。

寝ているときの体勢検知の手法として Kinect センサを使用したものがある。吉武ら [5] は Kinect センサから取得した 3D 骨格座標をもとに医療患者監視を行う実験を行っている。

また、株式会社富士通研究所 [6] では肺活量計を装着することで睡眠時の姿勢・寝返り回数・時間などを可視化するシステムも存在する。

2.4 むすび

本章では、本研究における従来手法について述べた。従来手法では、画像から乳幼児の姿勢推定を行うことが容易でないことを述べた。

第3章 関連技術

3.1 まえがき

本章では，本研究の関連技術について述べる．

本章の構成としては，まず，分類に用いる分類アルゴリズムの概要について述べる．次に，姿勢推定アルゴリズムの手法である OpenPose : REALTIME MULTI-PERSON POSE ESTIMATION 及び Cascaded Pyramid Network の概要について述べる．最後に，OpenPose での検出精度を改善するために使用する fine tuning について述べる．

3.2 分類アルゴリズム

3.2.1 Support Vector Machine

Support Vector Machine (SVM)は，教師あり学習を用いるパターン認識モデルの一つである．クラスの集合間の境界に位置するデータを基準とし，境界面に最も近いデータと境界との距離が最大化するように境界面を設定する手法である．SVM による分類を図 3.1 に示す．

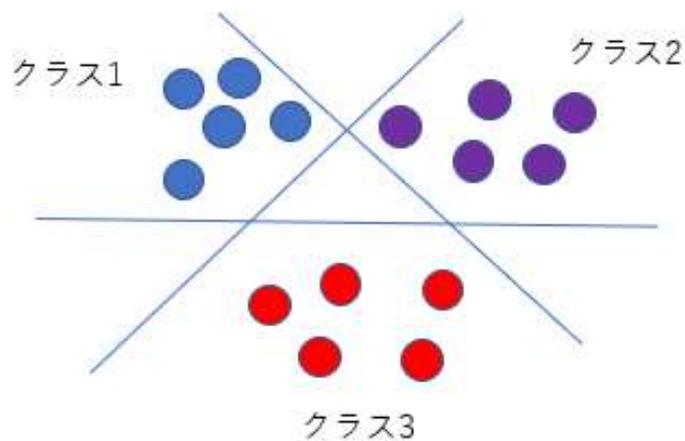


図 3.1 SVM による分類

3.2.2 決定木

決定木とは木構造で条件分岐によってクラスを分類する教師あり学習の手法である。与えられたデータに条件をもうけ、その条件に当てはまるか当てはまらないかで段階的にデータを分類する手法である。決定木による分類を図 3.2 に示す。

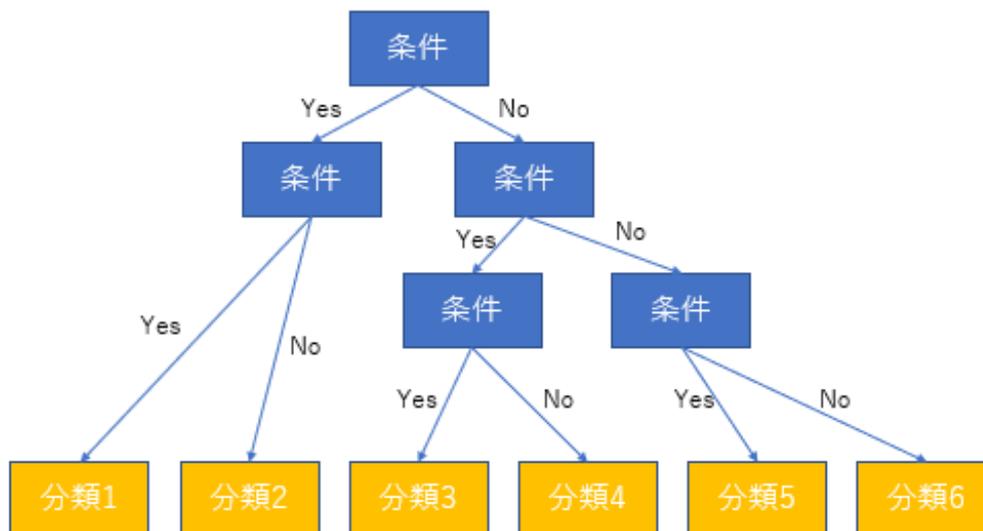


図 3.2 決定木による分類

3.2.3 random forest

random forest とは複数の決定木を作って多数決や平均を計算することができる手法である。条件分岐を何度もすることで、データが木のように構造化され、最小単位に分類されていく。ランダムフォレストは大量のデータを必要とするが、データの分類に重要な特徴量を判別できるため、精度の高い分類を行える特徴をもつ。random forest による分類を図 3.3 に示す。

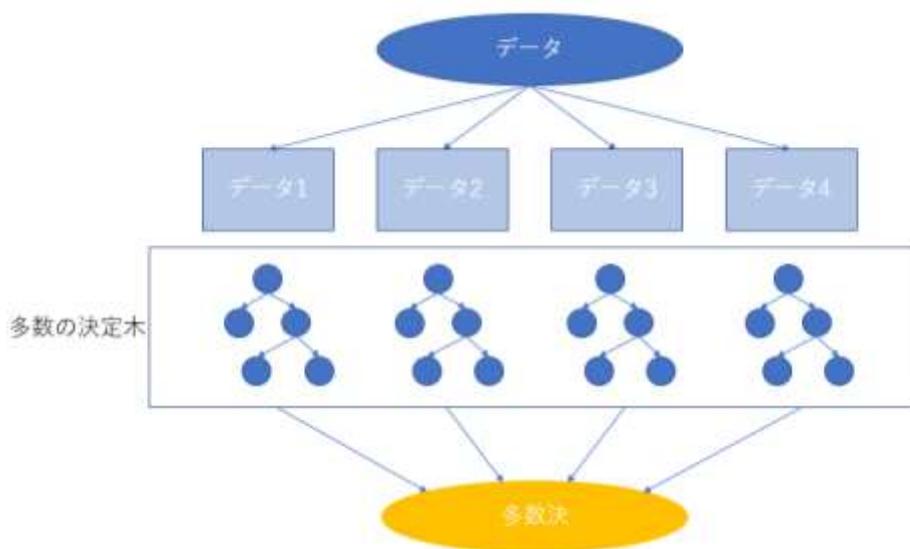


図 3.3 random forest による分類

3.2.4 k 近傍法

k 近傍法 (k-NN)とは、新たなデータが入力された際、それに最も距離の近い k 個のデータのラベルで多数決を取り、多い方にデータが属するクラスを分類するアルゴリズムである。

k-NN では k の数で結果が変わるため、データセットの特徴に適合している k を選ぶことが必要となる。k-NN による分類を図 3.4 に示す

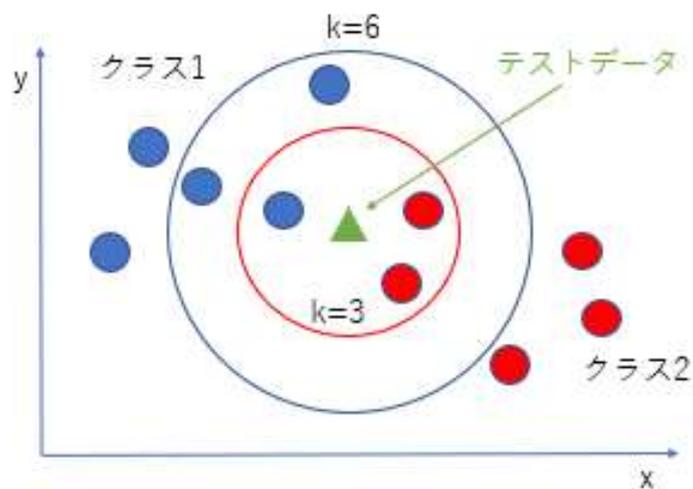


図 3.4 k-NN による分類

3.2.5 LightGBM

LightGBM は決定木アルゴリズムと勾配ブースティングに基づく機械学習手法である。勾配ブースティングとは前の弱分類器の予測値の誤差を新しい弱分類器が引継ぎ、段階的に誤差を小さくする手法である。random forest が決定木を並列に学習し、多数決をとる処理であるのに対し、勾配ブースティングは決定木を直列に学習する。勾配ブースティングの手法として XGboost もあるが、計算時間が長い特徴があるため、本実験では LightGBM のみを利用している。LightGBM による分類を図 3.5 に示す。

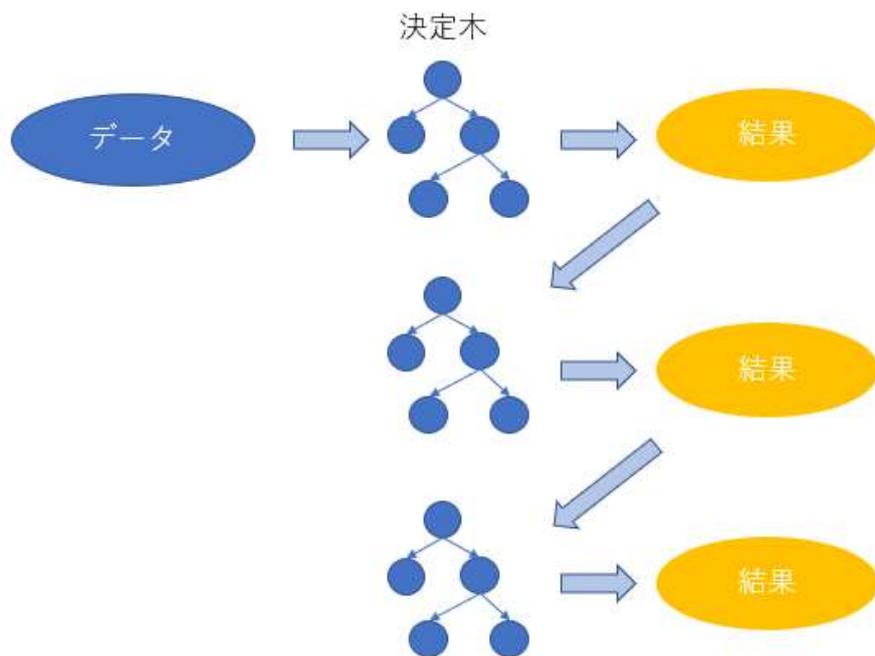


図 3.5 LightGBM による分類

3.3 OpenPose : Realtime Multi-Person Pose Estimation

OpenPose [7]とは、画像からキーポイントを見つけ出し、その後にキーポイントを人ごとに繋ぎ合わせていくことで姿勢推定を行うボトムアップ型の姿勢推定処理である。深層学習を用いて単一の画像から身体や頭などの人間のキーポイントの検出をリアルタイムにすることが可能な手法である。画像を入力すると画像中の人物の肩や肘などの 15, 18 もしくは 25 点の位置を推定し、各人物ごとの関節点を線で結んで人体モデルとして出力できる。動画を解析するオプションもあり、その場合は、動画のフレームごとに姿勢推定をする。OpenPose から 15, 18, 25 のキーポイントの座標は json ファイルで出力でき、その際、x

座標・y座標・信頼度が出力される。動画を解析する際もフレームごとにキーポイントの座標を json ファイルで出力できる。検出できないキーポイントの座標は x 座標も y 座標も 0 で表示される。図 3.6 に OpenPose によって検出される関節位置を示す。OpenPose を乳幼児に使用したとき、図 3.7 を入力画像とした場合、図 3.8 のような出力画像になる。

図 3.7, 3.8 は <https://www.photo-ac.com/main/detail/22369512> から引用しており、非商用の利用が可能である。

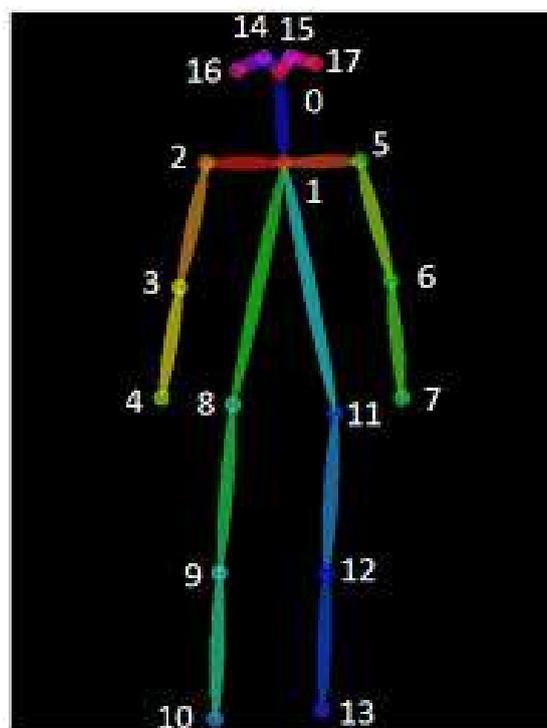


図 3.6 OpenPose によって検出される関節位置



図 3.7 OpenPose の検出例(入力画像)

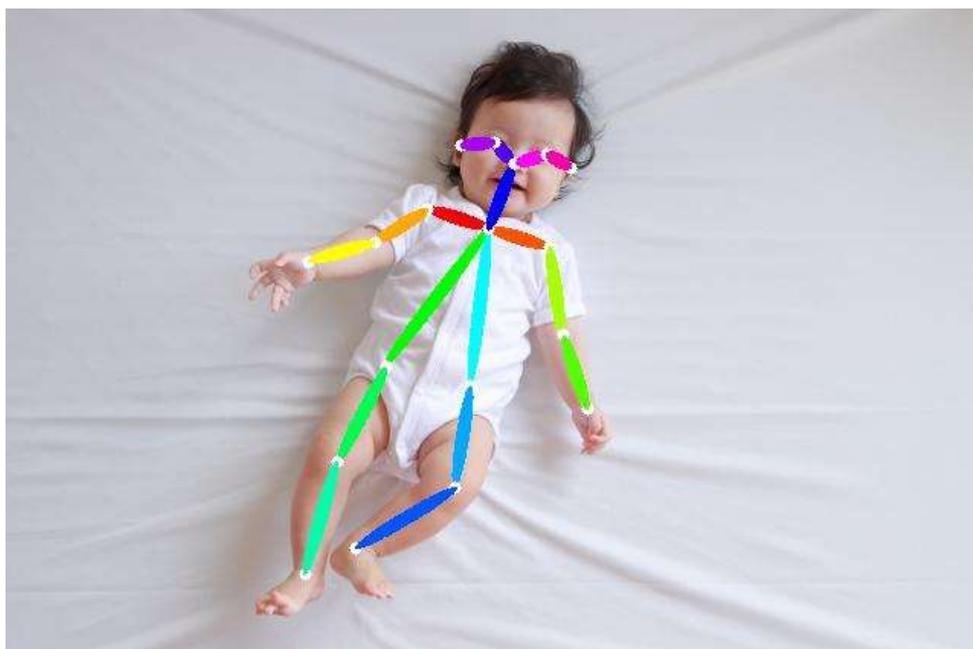


図 3.8 OpenPose の検出例(出力画像)

3.4 fine tuning

fine tuning とは学習済みモデルをベースに出力層などを変更したモデルを構築し、自前のデータでニューラルネットワーク・モデルの結合パラメータを学習させる手法である。

fine tuning では学習済みのネットワークの重みを初期値として、全層のパラメータを再学習させる。その際、入力層に近い部分のパラメータは学習率を小さく設定し、出力層に近い部分のパラメータは大きく設定することが多い。

fine tuning は学習済みのモデルをベースとしているため、用意したデータセットが少量でも性能の良いモデルを学習できるメリットがある。fine tuning の仕組みを図 3.9 に示す。

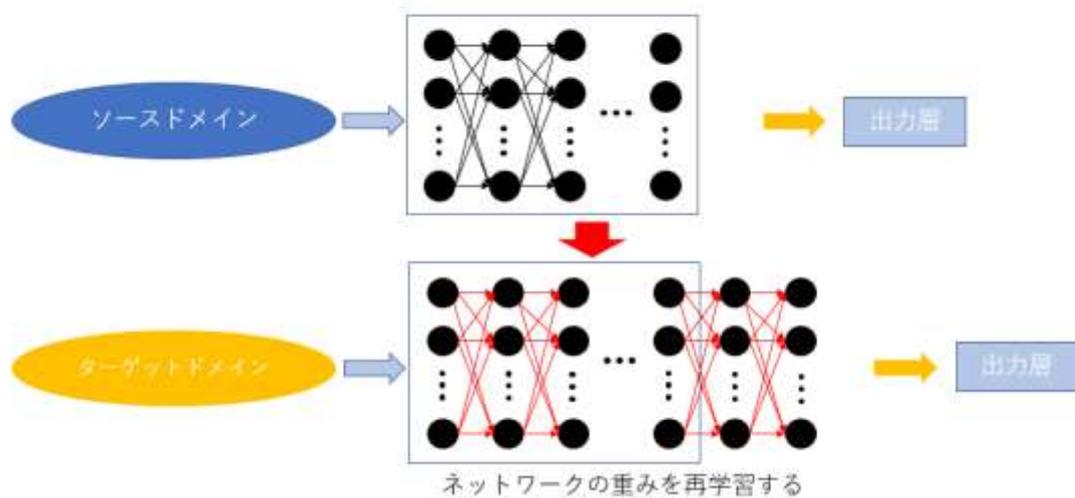


図 3.9 fine tuning の仕組み

3.5 むすび

本章では、本研究の関連技術について述べた。本研究ではこれらの関連技術を用いて画像中の乳幼児の関節座標から乳幼児の睡眠姿勢の分類を行う。また、fine tuning を適用することにより OpenPose の乳幼児検出の精度向上を図る。

第4章 提案手法

4.1 まえがき

本章では、はじめに乳幼児が寝ている画像から乳幼児の睡眠姿勢を分類する手法を提案する。次に、OpenPose に fine tuning を適用し乳幼児検出の精度を向上させる手法を提案する。最後に、本章のまとめを行う。

4.2 睡眠姿勢の分類

4.2.1 睡眠姿勢の分類手法の概要

画像中の乳幼児の関節座標から特徴量を抽出し、分類アルゴリズムを用いて乳幼児の睡眠姿勢の分類を行う。睡眠姿勢の分類処理過程を図 4.1 に示す。

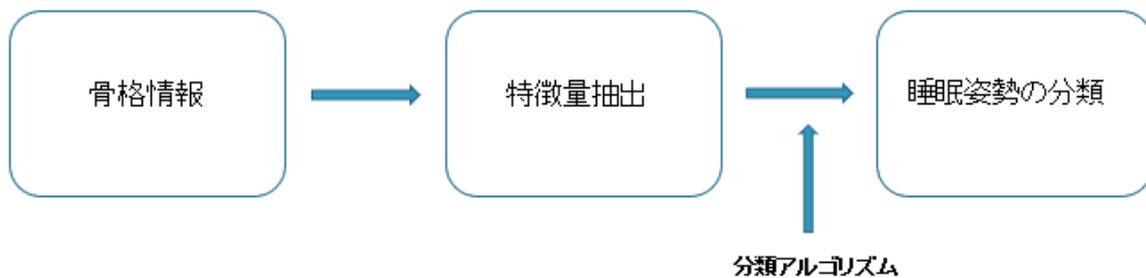


図 4.1 睡眠姿勢の分類の処理過程

4.2.2 分類する姿勢

本論文では、乳幼児の睡眠姿勢を“仰向け”、“横向き”、“うつぶせ”の 3 種類に分類する。それぞれの姿勢の例を図 4.2～4.4 で示す。図 4.2～4.4 は

<https://www.photo-ac.com/main/detail/22512255>

<https://www.photo-ac.com/main/detail/22511750>

<https://www.photo-ac.com/main/detail/5111381>

から引用しており、非商用の利用が可能である。



図 4.2 仰向け姿勢例



図 4.3 横向き姿勢例



図 4.4 うつぶせ姿勢例

4.2.3 分類アルゴリズムによる分類

乳幼児の関節情報から特徴量を算出する。使用する特徴量として、右肘、右肩、左肘、左肩、右尻、左尻、右膝、左膝の 8 個の関節角度を用いる。右肘、右肩、左肘、左肩、右膝、左膝は隣接する関節との角度を特徴量とする。右尻、左尻は隣接する関節座標の 3 個のうち、右尻の場合は右肩と右膝、左尻の場合は左肩と左尻を用いて関節角度を算出し特徴量とする。

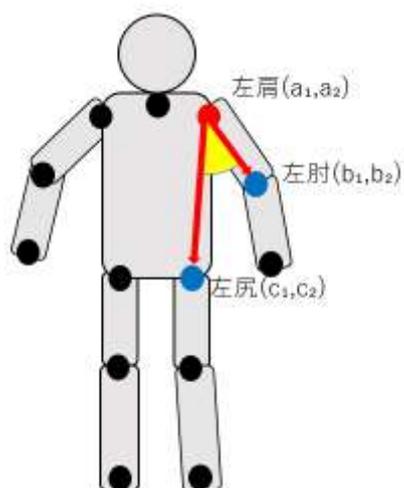


図 4.5 関節角度算出例

図 4.5 のような姿勢のとき，左肩の各節角度は以下のように求める。
左肩 (a_1, a_2) ，左肘 (b_1, b_2) ，左尻 (c_1, c_2) のとき，左肘と左肩からなるベクトルを \vec{b} ，左尻と左肩からなるベクトルを \vec{c} と表せる。

$$\vec{b} = (b_1 - a_1, b_2 - a_2) \quad (4.1)$$

$$\vec{c} = (c_1 - a_1, c_2 - a_2) \quad (4.2)$$

2つのベクトルのなす角を θ としたとき，式(4.3)から $\cos \theta$ が分かる。

$$\cos \theta = \frac{\vec{b} \cdot \vec{c}}{|\vec{b}| |\vec{c}|} \quad (4.3)$$

ここからアークコサインを算出することで関節角度 θ を求められる。

同様に 8 個の関節角度を算出し，特微量とする。各姿勢の特微量を学習する手法として 3.4 に述べた 5 種類の分類アルゴリズムを用いる。

4.3 OpenPose による乳幼児の骨格推定

4.3.1 OpenPose による乳幼児の骨格推定手法の概要

OpenPose での乳幼児の検出は成人への検出よりも精度がおとる。理由として，学習データに乳幼児が多く含まれていないためと考えられる。幼児の体の比率は成人と異なることや，乳幼児のポーズが成人よりも多様であることが，検出を困難にしていると考えられる。そこで図 4.5 に示すように，既存の姿勢推定手法：OpenPose に乳幼児の画像をもとに fine tuning を適用し，既存モデルから精度向上を図る。

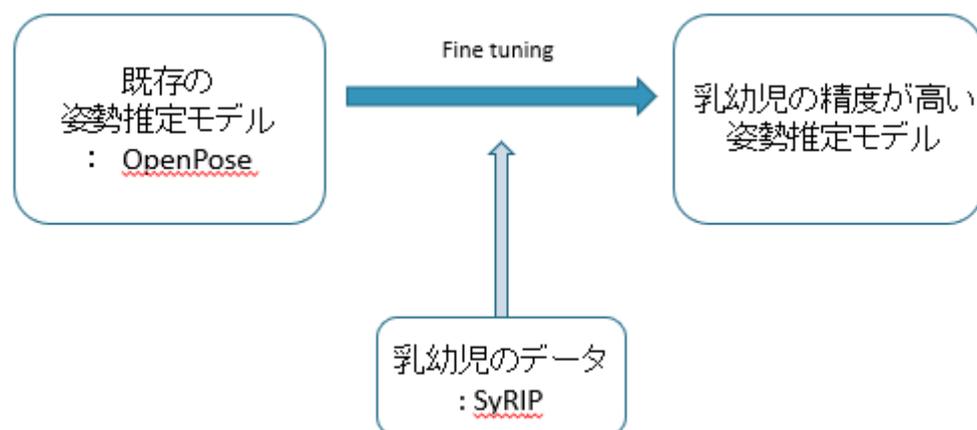


図 4.6 姿勢推定精度向上のための提案手法

4.3.2 OpenPoseによる乳幼児の骨格検出

OpenPose はその実装プログラムコードを公開している [8]. COCO 2016 keypoints challenge dataset による学習済みモデルを使い, この実装モデルによって 18 点の関節が推定される. OpenPose によって推定される関節点を表 4.1 に示す.

表 4.1 OpenPose によって推定される関節点

番号	OpenPoseで推定される関節点
1	Nose(鼻)
2	Neck(首)
3	RShoulder(右肩)
4	RElbow(右肘)
5	RWrist(右手首)
6	LShoulder(左肩)
7	LElbow(左肘)
8	LWrist(左手首)
9	RHip(右尻)
10	RKnee(右膝)
11	RAnkle(右足首)
12	LHip(左尻)
13	LKnee(左膝)
14	LAnkle(左足首)
15	REye(右目)
16	LEye(左目)
17	REar(右耳)
18	LEar(左耳)

4.3.3 fine tuning に用いる乳幼児のデータセット

fine tuning には乳幼児のデータセットとして SyRIP (synthetic and real infant pose) [9] を使用する. このデータセットは Xiaofei Huang, Nihang Fu, Shuangjun Liu, Sarah Ostadabbas, らが “Invariant Representation Learning for Infant Pose Estimation with Small Data,” [10]

内で作成した乳幼児のデータセットである。このデータセットは、座っている、転がっている、這っているなどの様々なポーズをとった RGB の乳幼児画像であり、実画像と合成画像を含む。実画像は Google 画像や YouTube 動画から収集した 100 種類以上の幼児画像から構成されている。また、合成画像は 12 種類以上の被写体の画像を用いて、独自に生成されている。データセット内には 700 枚の実際の乳幼児の画像とそれに対するアノテーションと 1000 枚の乳幼児の合成画像が含まれている。

SyRIP のアノテーションを表 4.2 に示す。SyRIP のアノテーションは fine tuning を行う際に使用する [Microsoft](#) Common Object in Context (MS COCO) のデータセットのアノテーションと異なっているため、マッピングを行う。MS COCO とは画像データを使用するディープラーニング向けのデータセットである。様々な画像とアノテーションが含まれており、本研究では姿勢推定用の画像とアノテーションを用いる。SyRIP にない首のアノテーションはアノテーションにも画像にも存在しないラベルづけを行う。また、SyRIP にない対象人物の中心を示すアノテーションは左肩、右肩、左尻、右尻の中点を用いる。

表 4.2 SyRIP のアノテーション

番号	SyRIP のアノテーション
1	nose(鼻)
2	left_eye(左目)
3	right_rye(右目)
4	left_ear(左耳)
5	right_ear(右耳)
6	left_sholder(左肩)
7	right_shoulder(右肩)
8	left_elbow(左肘)
9	right_elbow(右肘)
10	left_wrist(左手首)
11	right_wrist(右手首)
12	left_hip(左尻)
13	right_hip(右尻)
14	left_knee(左膝)
15	right_knee(右膝)
16	left_ankle(左足首)
17	right_ankle(右足首)

4.4 むすび

本章では、睡眠姿勢の分類手法、OpenPose による乳幼児の骨格推定の手法について述べた。

第5章 実験

5.1 まえがき

本章では、前章で示した提案手法をもとに睡眠姿勢の分類, OpenPose による乳幼児の骨格推定の実験を行う。睡眠姿勢の分類実験ではデータセット, 実験結果及び考察について述べる。また, OpenPose による乳幼児の骨格推定実験では fine tuning 後のモデルの結果, 評価の際に使用するデータセット, モデルの精度評価について述べる。

5.2 睡眠姿勢の分類実験

5.2.1 データセット

本実験では SyRIP の実画像のうち“仰向け”, “横向き”, “うつぶせ”の姿勢をしている 229 枚を使用する。各画像中に写る人物は乳幼児 1 人のみである。“仰向け”, “横向き”, “うつぶせ”の画像数を表 5.1 に示す。これらの画像を仰向け, 横向き, うつぶせのポーズにラベル付けした。

表 5.1 姿勢ごとの画像枚数(実験 1)

画像の種類	枚数[枚]
仰向け	99
横向き	24
うつぶせ	106

5.2.2 実験結果及び考察

従来手法で述べたように, 画像から乳幼児の姿勢を検知する手法で顔の向きを使用してあるものがある。この手法と比較するため, 顔全面が映っている場合(左右の目が見えている場合)を“仰向け”, 顔の半分が映っている場合(左右の目どちらかが見えている場合)を“横向き”, 顔全面が映っていない場合(目が映っていない場合)を“うつぶせ”として分類し, その精度を算出した。

本実験のデータの特徴量の可視化を行った結果を図 5.1 に示す。

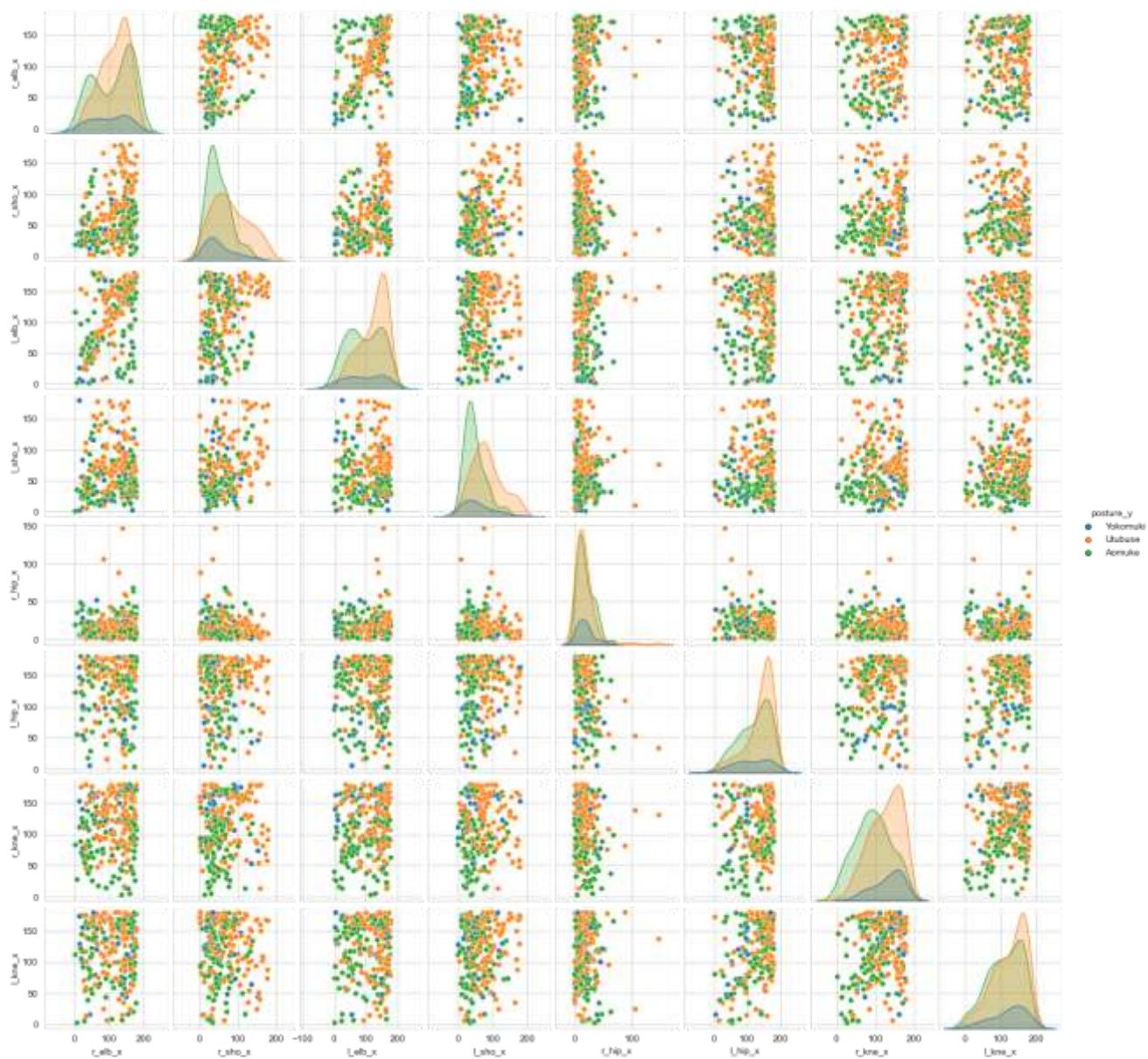


図 5.1 SyRIP データと特徴量の可視化

決定木で分類を行った際の木構造を図 5.2 で示す。

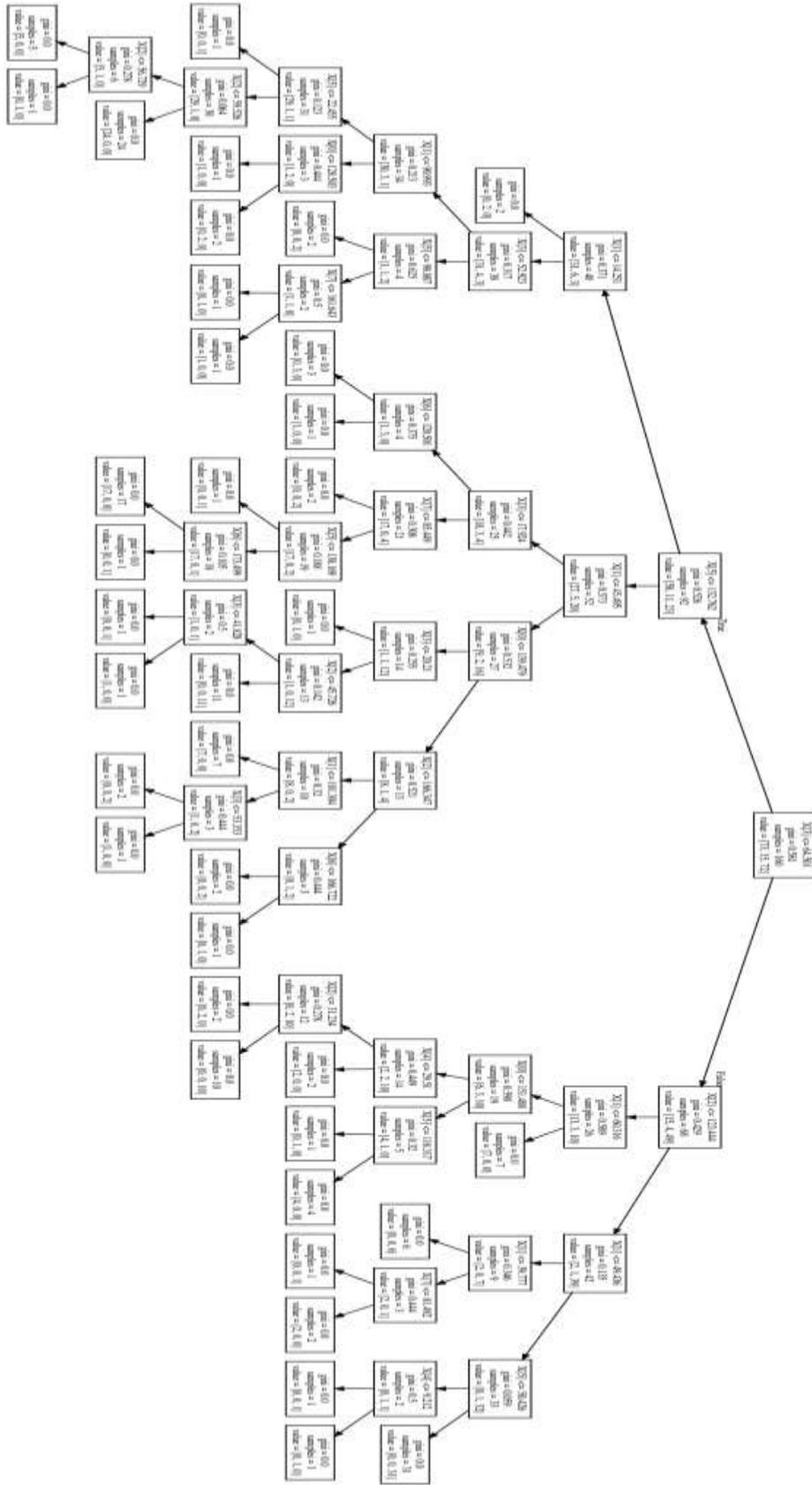


図 5.2 決定木の木構造

本実験で k-NN の k の値は図 5.3 から分かるように正解率が最大となった k=7 を使用した。

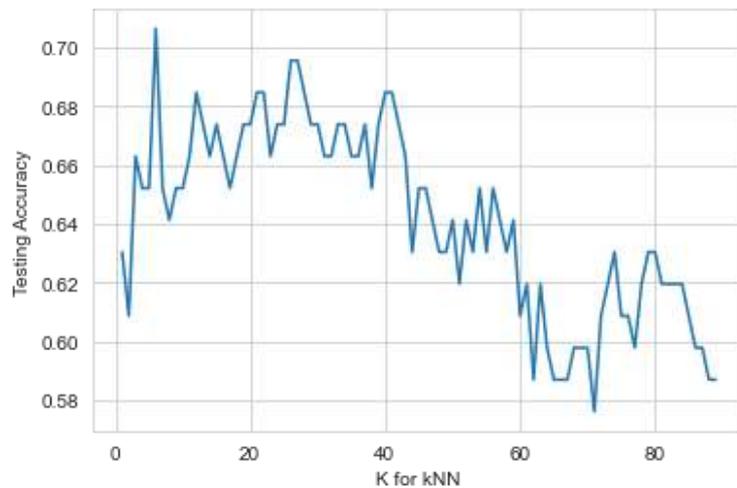


図 5.3 k-NN の正解率

目から姿勢を判断した場合と分類アルゴリズムを用いたときの分類精度を表 5.2 に示す。

表 5.2 睡眠姿勢の分類結果

分類手法	分類精度[%]
顔の向きからの判別	45.0
SVM	69.6
決定木	50.7
Random forest	58.1
k-NN	70.6
LightGBM	76.1

表 5.2 より、どの分類アルゴリズムでも顔の向きのみからの判別より高い精度で分類できていることが分かる。また、分類アルゴリズムの中では LightGBM が最も高い精度で分類できることが分かる。

5.3 OpenPose による乳幼児の骨格推定実験

5.3.1 fine tuning 後のモデルの結果

既存の OpenPose のネットワークモデルに対して、学習率、モーメントを変更した。学習データが 1200 枚(SyRIP)+8500 枚(COCO)の 9700 枚, 検証データが 500 枚(SyRIP)+2550 枚(COCO)の 3050 枚に設定し、15epoch として学習を実行した。

各 epoch の Loss を図 5.4 に示す。

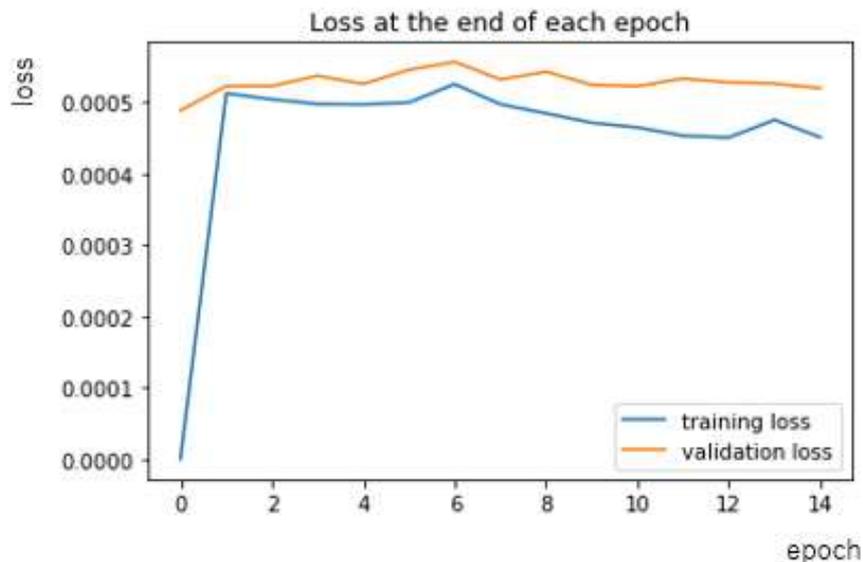


図 5.4 fine tuning の Loss

従来モデルから fine tuning 後の検出結果の変化を図 5.5 と図 5.6 に示す。左側が従来モデルの検出結果で右側が提案モデルの検出結果である。図 5.5 に写る乳幼児はカメラと反対向きを向いてハイハイしており、目、鼻、右肘、右手首は見えていない姿勢である。図 5.6 にうつる乳幼児はうつぶせの姿勢をしており、目、鼻、右肘、右手首は見えていない姿勢である。

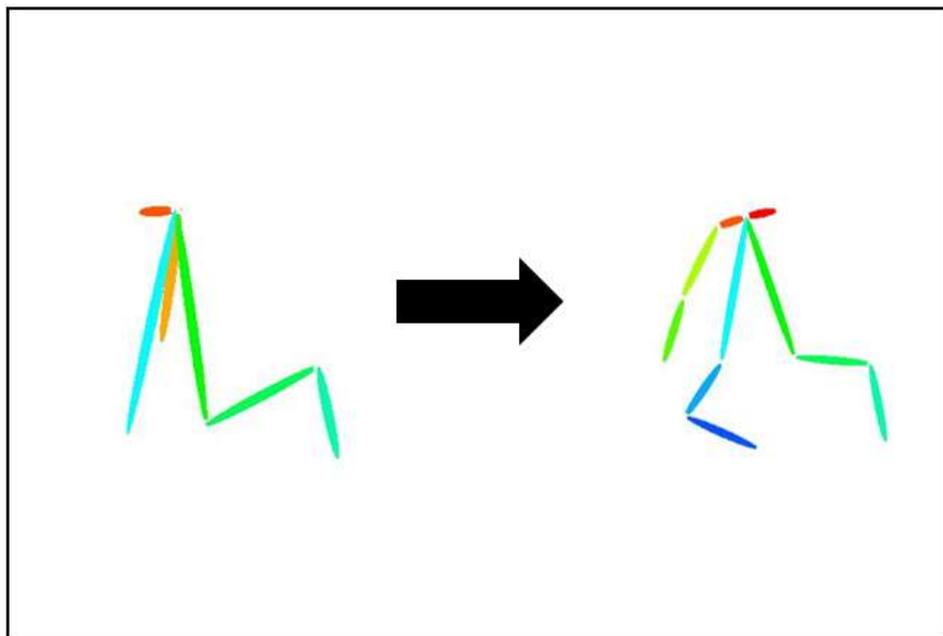


図 5.5 従来モデルから fine tuning 後の検出結果の変化 1

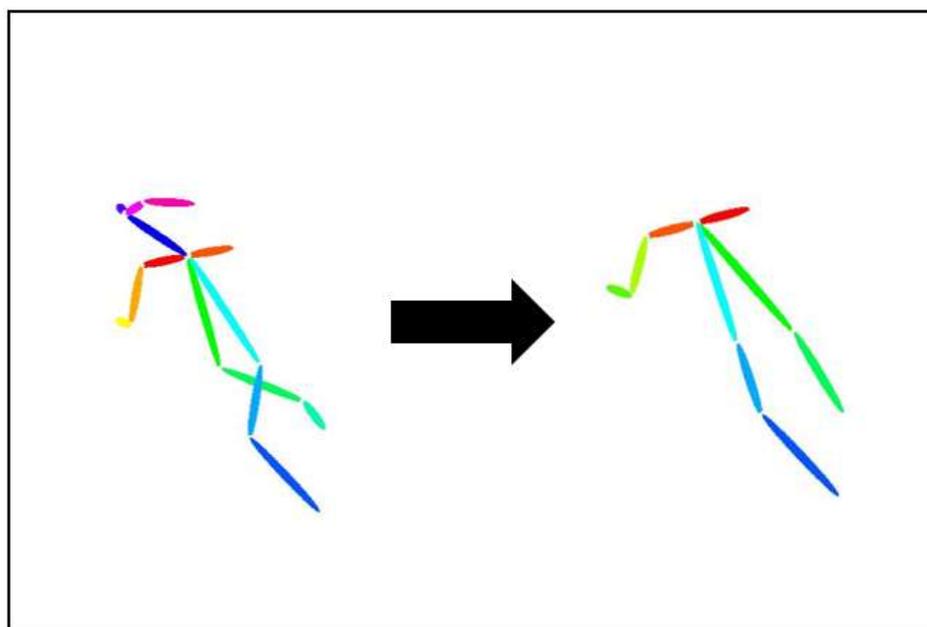


図 5.6 従来モデルから fine tuning 後の検出結果の変化 2

図 5.5 に示した従来モデルから fine tuning 後の検出結果の変化から、fine tuning 後のモデルのほうが右肩、左肩、左肘、左手首、左膝、左足首を正しく検知できていることが分かる。また、図 5.6 に示した従来モデルから fine tuning 後の検出結果の変化から、従来モデルによる検出結果では、後ろを向いているのに頭を顔と認識してしまっており、左右が反転

しているが、fine tuning 後のモデルによる検出結果では左右が正しく検知できていることが分かる。他の画像においても、従来モデルより fine tuning 後のモデルの方がおおよそ正しく検出できていることが分かった。

5.3.2 評価の際に使用するデータセット

本実験の評価に SyRIP の実画像 350 枚を使用する。この画像中に写る人は乳幼児 1 人になっており、乳幼児の姿勢は睡眠姿勢だけでなく立っているもの、ハイハイの姿勢のものも含まれる。使用した画像内の姿勢の分類を表 5.1 に示す。

表 5.3 姿勢ごとの画像枚数(実験 2)

画像の種類	枚数[枚]
仰向け	90
横向き	18
うつぶせ	100
ハイハイ	72
座っている	55
立っている	5
上記以外の姿勢	10

5.3.3 モデルの精度評価

学習前の従来モデルと今回 SyRIP によって従来モデルに fine tuning を適用し作成したモデルの精度を比較する。使用するデータセットの画像内の乳幼児の 18 点の関節を“検出”，“未検出”，“誤検出”，“検出不可”と判別し、総関節数からの“検出”の割合を算出した。“未検出”とは関節が画像内に写っているが、検出できていないものを指す。“誤検出”とは関節が画像内に写っており、検出しているが、検出位置が正しい位置でなかったものを指す。“検出不可”は画像内に関節点が写っていない場合を指す。関節の検出数と成功率を表 5.2 に示す。

表 5.4 モデルの関節検出数と検出成功率

	学習前のモデル	fine tuning 後のモデル
検出成功数[点]	3837	4293
未検出数[点]	415	124
誤検出数[点]	1039	874
検出不可数[点]	1009	1009
検出成功率[%]	72.5	81.1(+8.6[%])

SyRIP データセットに対するモデルの検出率が従来モデルより fine tuning 後のモデルのほうが 8.6%向上しているのが分かる。すなわち、従来手法を上回る精度を達成できた。

5.4 むすび

本章では、睡眠姿勢の分類と OpenPose による乳幼児の骨格推定と提案手法に基づいた実験を行った。睡眠姿勢の分類では従来の検出手法より高い精度で分類できることが分かった。また、OpenPose による乳幼児の骨格推定では従来モデルより fine tuning を適用した後のモデルで上回る精度を達成できた。

第6章 結論と今後の課題

6.1 結論

本研究では、乳幼児の睡眠時の姿勢分類に骨格座標から求めた特徴量を使用し、分類アルゴリズムによって判別する提案を行った。従来手法である顔の向きのみから判別する手法よりも高い精度で分類できることが確認できた。

また、乳幼児のデータを使用し、姿勢推定モデルである OpenPose に対して fine tuning を適用する手法を提案した。乳幼児の画像に対し関節検出数と検出成功率で性能の評価を行った。その結果、従来モデルよりも提案手法のモデルの方が、高い検出精度を与えることを確認できた。

6.2 今後の課題

実際に「乳幼児のうつぶせアラート」を出すにはより高精度の分類が必要である。また、乳幼児の姿勢推定を本実験での目的である睡眠姿勢の検出やその他の行動認識で使用するには、さらに高精度の検出精度が必要である。

本研究での提案手法に加えて時系列情報、体の向き情報などの特徴量を追加することで分類精度が高められると考える。また、検出精度向上にはデータ量を増やすことが必須である。本研究での実験において見られた傾向として、特に乳幼児の下半身部分の未検出が多かったため、学習に用いるパラメータを工夫することが必要だと考えられる。

謝辞

本研究の実験環境を与えてくださり、研究の方向性を示して下さいなど丁寧な指導を賜った渡辺裕教授に深く感謝申し上げます。

また、日頃からご意見やアドバイスを下さった研究室の皆様に御礼申し上げます。

最後に、私が学業に専念できる環境を与え、これまで温かく見守ってくれた家族に感謝いたします。

参考文献

- [1] 消費者庁消費者安全課, “子どもの不慮の事故の発生傾向～厚生労働省「人口動態調査」より～,” 5 3 2021. Available:
https://www.caa.go.jp/policies/policy/consumer_safety/meeting_materials/assets/consumer_safety_cms205_210305_03.pdf. [アクセス日: 21 1 2022].
- [2] S. MAGAZINE, “赤ちゃんの寝返りや呼吸など小さな動きを見逃さない。育児に役立つ IoT グッズ,” +Style, 13 6 2019. Available: <https://plusstyle.jp/blog/item?id=168>. [アクセス日: 21 1 2022].
- [3] EDGEMATRIX 株式会社, “赤ちゃんうつぶせ寝検知,” EDGEMATRIX 株式会社, Available: https://service.edgematrix.com/application/babydetection_neut/. [アクセス日: 21 1 2022].
- [4] 株式会社アジラ, “人工知能/画像認識技術を提供するアジラ、『行動認識』に関する 3 つの特許取得のお知らせ,” 25 6 2019. Available: <https://prtimes.jp/main/html/rd/p/000000004.000043312.html>. [アクセス日: 20 1 2020].
- [5] 吉武 伸泰, 恒田 晃完, 灘口 尚大, 田中 康一郎, “Kinect センサを用いた医療用患者監視システム,” 2013. Available: <https://www.ipsj-kyushu.jp/page/ronbun/hinokuni/1002/B-5/B-5-3.pdf>. [アクセス日: 20 1 2020].
- [6] 西川株式会社, “東京西川 快適な睡眠環境をアドバイスする新サービス「ねむりの相談所®」を 3 月 23 日より全国規模で展開!,” 西川株式会社, 23 3 2017. Available: <https://prtimes.jp/main/html/rd/p/000000049.000010201.html>. [アクセス日: 30 1 2020].
- [7] Z. Cao, T. Simon, S-E Wei, Y. Sheikh, “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), No.121, pp.1302-1310., July 2017.
- [8] Z. Cao, T. Simon, S.-E. Wei and Y. Sheikh, “GitHub:Realtime Multi-Person Pose Estimation,” Available: https://github.com/CMU-Perceptual-ComputingLab/caffe_rtpose. [アクセス日: 29 12 2019].
- [9] X. Huang, “Synthetic and Real Infant Pose (SyRIP) Dataset,” northeastern university, 9 10 2020. Available: <https://coe.northeastern.edu/Research/AClab/SyRIP/>. [アクセス日: 21 1 2022].
- [10] X. Huang, N. Fu, S. Liu, S. Ostadabbas, “Invariant Representation Learning for Infant Pose Estimation with Small Data,” 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021), , 15-18 Dec 2021.

図一覧

図 2.1	デバイスを付ける手法例	3
図 2.2	画像から姿勢推定している例	4
図 3.1	SVM による分類.....	6
図 3.2	決定木による分類.....	7
図 3.3	random forest による分類.....	8
図 3.4	k-NN による分類.....	8
図 3.5	LightGBM による分類	9
図 3.6	OpenPose によって検出される関節位置	10
図 3.7	OpenPose の検出例(入力画像).....	11
図 3.8	OpenPose の検出例(出力画像).....	11
図 3.9	fine tuning の仕組み	12
図 4.1	睡眠姿勢の分類の処理過程.....	13
図 4.2	仰向け姿勢例.....	14
図 4.3	横向き姿勢例.....	14
図 4.4	うつぶせ姿勢例	15
図 4.5	関節角度算出例	15
図 4.6	姿勢推定精度向上のための提案手法.....	16
図 5.1	SyRIP データと特徴量の可視化.....	21
図 5.2	決定木の木構造	22
図 5.3	k-NN の正解率	23
図 5.4	fine tuning の Loss.....	24
図 5.5	従来モデルから fine tuning 後の検出結果の変化 1	25
図 5.6	従来モデルから fine tuning 後の検出結果の変化 2	25

表一覧

表 4.1	OpenPose によって推定される関節点.....	17
表 4.2	SyRIP のアノテーション	18
表 5.1	姿勢ごとの画像枚数(実験 1)	20
表 5.2	睡眠姿勢の分類結果	23
表 5.3	姿勢ごとの画像枚数(実験 2)	26
表 5.4	モデルの関節検出数と検出成功率	27

研究業績

- [1] A. Okuno, T. Ishikawa, H. Watanabe, “Rollover Detection of Infants Using Posture Estimation Model,” 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), pp.490-493, October. 2020.