

歩行映像解析による体組成推定に関する一検討

廖若辰[†] 守脇幸佑[†] 榎原靖[†] 村松大吾[†] 武村紀子[†] 八木康史[†]

概要：体組成は健康状況を把握するための重要な指標である。体脂肪率や体水分率、筋肉量などを把握することにより、肥満や生活習慣病の予防や改善が可能になり、現代社会における健康維持のためにその必要が増しつつある。市販の体組成計の多くは、生体電気インピーダンス分析法を用いるものが多く、正確な結果を出せる一方、設備が高価という問題点がある。また一人ずつしか計測できないため、多人数を効率よく計測するには不向きである。そこで、本研究では、多人数を効率よく計測するための、歩行映像解析による体組成推定を試みる。具体的には、歩行映像から抽出するシルエットに基づく特徴表現である歩容エネルギー画像 (Gait energy image, GEI) を入力、各体組成の値を出力とする畳み込みニューラルネットワーク (Convolutional neural network, CNN) を構成し、被験者の歩行映像から抽出した GEI と市販の体組成計で計測した体組成の値の組を学習データとして、ネットワークパラメタを学習する。ここで、体組成を計測できる被験者数には限りがあることから、CNN を適切に学習することが困難となる。そこで、まず、大規模歩行映像データベースから抽出した GEI を入力、同データベースから抽出可能な、体組成と関連性のありそうな歩容個性 (腕振りの大きさや歩幅) を出力とする CNN を事前学習する。次に、事前学習されたパラメタを持つ中間層までのネットワークに対して、いくつかの層を追加した、即ち、構造的に成長させたネットワークの出力に体組成値を設定し、ネットワークのファインチューニングを行うことで、限られた体組成の学習データからでも効果的に学習可能なことを示す。実験では、体組成の学習データのみを用いた、サポートベクター回帰や CNN による推定手法と比較して、提案手法が高い精度を得られることを確認した。

キーワード：歩行映像解析、体組成、深層学習、健康管理

1. はじめに

体組成とは、体脂肪や筋肉量といった人間の身体を構成する各種成分の割合であり、それを把握することにより生活習慣病の予防や改善といった健康管理や、フィジカルトレーニングの成果確認などに活用されている[1]–[4]。

現在、体組成を計測する手法はいくつかあり、市販の体組成計では生体電流インピーダンス法[5]、[6]を使用するのが主流である。この手法では、対象の手や足を電極に接触させ、微弱な交流電流を流し、そのインピーダンスを解析することで体組成を推定している。この手法では正確な体組成計測ができるが、体組成計は高価であることから、一般への普及には困難が伴う。

一方で、体組成は人間の運動機能に強く関わっているため、運動機能を評価することで、体組成を推定できる可能性がある。そのような運動の中でも、歩行は、生活の中で最も一般的かつ頻繁に行う運動の一つである。そのため、歩行映像解析による性別や年齢、健康状況を推定する研究が多くなされている[7]–[11]。また、成人は身体の老化につれて体組成が変化し、筋肉量が減少する[12]と言われているが、それが原因で年齢別の歩行に違いが生じる。これらを踏まえると、体組成を推定するために観測する運動として、歩行動作が一つの有力な候補となり得る。

本研究では、歩行映像を解析することにより、体組成を推定する手法を提案する。具体的には、歩行映像から抽出するシルエットに基づく特徴表現である歩容エネルギー画像 (Gait energy image, GEI) [13]を入力、各体組成の値を出力とする畳み込みニューラルネットワーク (Convolutional

neural network, CNN) を構成し、被験者の歩行映像から抽出した GEI と市販の体組成計で計測した体組成の値の組を学習データとして、ネットワークパラメタを学習する。ここで、体組成を計測できる被験者数には限りがあることから、CNN を適切に学習することが困難となる。

そこで、まず、大規模歩行映像データベースから抽出した GEI を入力、同データベースから抽出可能な、体組成と関連性のありそうな歩容個性 (腕振りの大きさや歩幅) を出力とする CNN を事前学習する。次に、事前学習されたパラメタを持つ中間層までのネットワークに対して、いくつかの層を追加した、即ち、構造的に成長させたネットワークの出力に体組成値を設定し、ネットワークのファインチューニングを行うことで、限られた体組成の学習データからでも効果的に学習可能なことを示す。実験では、体組成の学習データのみを用いた、サポートベクター回帰や CNN による推定手法と比較して、提案手法が高い精度を得られることを確認した。

2. 歩行映像解析による体組成推定

2.1 体組成と健康指標

本研究では、体組成計によって計測可能な、以下 26 の健康指標 (表 1) の推定を試みる。ここで、体組成計には、株式会社インボディ・ジャパンの体成分分析装置 InBody を用いた。

表 1 推定する体組成の健康指標とその単位

指標	単位	指標	単位	
体重	kg	部位別 体脂肪量	kg	
体水分量	L			左腕
タンパク質量	kg			右腕
ミネラル量	kg			体幹
体脂肪量	kg			左脚
				右脚
筋肉量	kg	BMI	kg/m ²	
除脂肪量	kg	体脂肪率	%	
骨格筋量	kg	InBody 点数	N/A	
部位別 筋肉量	kg	基礎代謝量	kcal	
		ウエストヒップ比	N/A	
		内臓脂肪レベル	レベル	
		肥満度	%	
		四肢の筋肉量	kg	

- 体重, 体水分量 (Total Body Water, TBW), タンパク質量, ミネラル量, 体脂肪量 (Body Fat Mass, BFM), 筋肉量 (Soft Lean Mass, SLM), 除脂肪量 (Fat Free Mass, FFM), 骨格筋量 (Skeletal Muscle Mass, SMM) の 8 項目は字面通り, 直接計測した体重と各体成分の質量を表す. 体水分量, タンパク質量, ミネラル量と体脂肪量の合計は体重に相当し, 体重から体脂肪量を除いた質量が除脂肪量となる. 骨格筋は筋肉全体の中に骨格を動かすための筋肉であり, 運動能力と直接かかわるものである.
- BMI (Body Mass Index, ボディマス指数) は身長 h と体重 w を用いて下記の式により計算される:

$$v_{\text{BMI}} = w/h^2$$
- 体脂肪率 (Percent Body Fat: PBF) は体脂肪量が体重に占める割合を示す.
- InBody 点数 (InBody Score: IBS) は体成分分析装置 InBody の開発会社が定めた指標で, 除脂肪量と体脂肪量の実測値を標準値と比較して算出される.
- 基礎代謝量 (Basal Metabolic Rate, BMR) は計測した除脂肪量 m_{FFM} に基づき, カニンガムの公式を用いて算出している:

$$v_{\text{BMR}} = 370 + 21.6 \times m_{\text{FFM}}$$
- 肥満度 (Obesity Degree, OD) は標準体重に対する現在の体重の割合である. 標準体重は身長により算出される.
- ウエストヒップ比 (Waist-Hip Ratio, WHR) は推定された腹囲と臀囲の比率であり, 内臓脂肪レベル (Visceral Fat Level, VFL) は臍まわりの腹部水平断面にある内臓脂肪の多さを評価している.
- 骨格筋指数 (Skeletal Muscle Index, SMI) は四肢の筋

肉量 m_{LBM_L} と身長 h に基づき, 以下の式を用いて計算できる:

$$v_{\text{SMI}} = m_{\text{LBM}_L}/h^2$$

ただし, 実験ではこの指標の出力が不安定であり, 計算できなかった被験者が多い. そのため, 部位別筋肉量から四肢の値を合計し, SMI を計算するための四肢の筋肉量として扱い, 推定した.

2.2 歩行特徴

推定に用いる特徴量としては, 歩行者のシルエット 1 歩行周期で平均することにより得られる GEI を使用する. GEI は動的特徴と静的特徴を組み合わせた特徴量であり, 対象の動作と体形を両方表現できる. これにより, 筋肉量のような運動能力とのかかわりが強い指標と, 体脂肪量のような体形への影響が強い指標を同時に評価できる.

2.3 事前学習

GEI を入力とした CNN の具体的な構造としては, GEINet [14]を用いる. GEINet は, CNN を用いた歩行映像解析のために設計されたネットワーク構造であり, 歩容認証に加えて, 年齢推定などにも使用されている.

但し, 体組成を計測できる被験者数には限りがあることから, GEINet を適切に学習することが困難となる. そこで, まず, 大規模歩行映像データベースから抽出した GEI を入力, 同データベースから抽出可能な, 体組成と関連性のあるような歩容個性 (腕振りの大きさや歩幅, 以下ではプリミティブと呼ぶ) を出力とする CNN を事前学習する

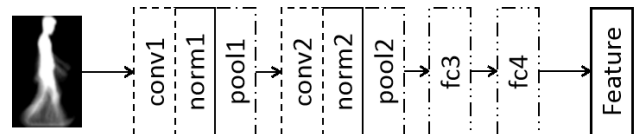


図 1 1in-GEI のネットワーク構造

表 2 各層の設定

Layer	#Kernels	Size/stride	Act.	Pooling
conv1	18	7 × 7 × 1/1	ReLU	
pool1		2 × 2/2		Max pooling
conv2	45	5 × 5 × 18/1	ReLU	
pool2		3 × 3/2		Max pooling

本研究では, GEINet の中でも, 一つの GEI を入力とする 1in-GEINet を使用する. ネットワークの構成と各層の設定は, 図 1 及び表 2 で示す通りである. 正規化層は局所応答正規化 (LRN: Local response normalization) を採用している. 全結合層では, fc3 層は 1024 次元であり, ReLU による活性化及びドロップアウトを適用している. 続く fc4 層は, 個人認証を目的として多次元ベクトルを出力していた元の GEINet とは異なり, スカラー値であるプリミティブの推定

(回帰問題)を解いているため、1次元の値を出力するものとする。以降、このようにして事前学習されたネットワークを、プリミティブネットワークと呼ぶ。

2.4 ファインチューニング

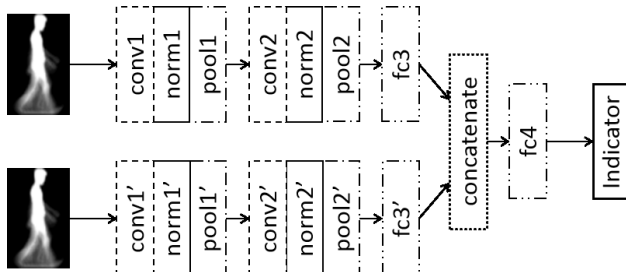


図 2 2in-GEINetのネットワーク構造

事前学習したプリミティブネットワークを用いて、体組成及び健康指標を推定するネットワークをファインチューニングする。ファインチューニングでは、2in-GEINetの構造を用いており、事前学習したプリミティブネットワークを2種類同時に使用することで推定精度の向上を図る。ネットワーク構造は図2で示す通りである。各層の設定について、各conv, pool, norm層及びfc3層の設定は事前学習時と同様であり、fc4層の直前に統合層を挿入して二つのプリミティブネットワークを統合する。

3. 実験

3.1 データ収集

体組成及び健康指標の情報を含むデータセットを構築するために、被験者実験を行い、データを収集した。実験では、被験者の通常歩行と体組成を計測した。通常歩行では、約10mのコースを被験者に通常で歩いてもらい、側面から約5m離れたところにカメラを設置し、その様子を撮影した。撮影した映像は背景差分を使い、患者1人あたり75枚のシルエット画像を抽出し、それを用いてGEIを生成した。また、被験者の直立姿勢の正面写真を撮影した。体組成と健康指標の計測は体成分分析装置InBody(型式:InBody270)を使用した。健康指標を計算するために必要な身長情報は被験者に自己報告してもらい、事前に入力している。

実験は、通常歩行の撮影を当日の午前と午後で2回行い、それぞれのGEIを抽出した。また、体組成計測を1回行った。参加者計203名の内、正しく計測と撮影ができ、かつデータの利用を許諾したのは175名であり、GEIは午前と午後の合計で337枚である。

3.2 学習

事前学習のデータセットは、大規模歩行映像データベースであるOULP-Age[15]を用いていた。OULP-Ageに含まれる40,000枚のGEIを入力とし、目標とする歩行特徴は既存

の手法[16]を用いてシルエット系列から抽出した腕の前振り、腕の後振り、背筋の伸び及び歩幅の4つの特徴についてそれぞれ学習した。学習時はミニバッチサイズを128、学習率の初期値を0.001、エポック数を200と設定している。

ファインチューニングのデータセットは、被験者実験で収集したデータを使用し、20分割交差検証を用いてネットワークを学習した。2in-GEINetとして使用する2つのプリミティブネットワークを選出するため、事前学習した4つのネットワークの全組み合わせを試し、誤差が最も小さい腕の前振りと腕の後振りの組み合わせを選出した。一部の被験者は実験時、スカートもしくは裾の広いズボンを着用しており、図3のようにGEI画像の前景面積を著しく増大させる。服装の違いが推定に与える影響を検証するために、全員のデータを使用するデータセット(Allと称する)と、服装により前景面積が増大したGEIを除いたデータセット(Omittedと称する)両方で推定を行った。また、体組成の量は対象の身長と強く関係していると思われるが、GEIはシルエットを正規化しているため身長を含まれていない。そのため、被験者の身長を正面写真から計測し、体組成と健康指標の中に質量(単位がkgのもの、例えば、体重、筋肉量)または体積(単位がLのもの、例えば、体水分量)で表現するものは身長を三乗を用いて正規化した。学習時のミニバッチサイズと学習率の初期値は事前学習と同じ数値に設定したが、エポック数は250に設定した。

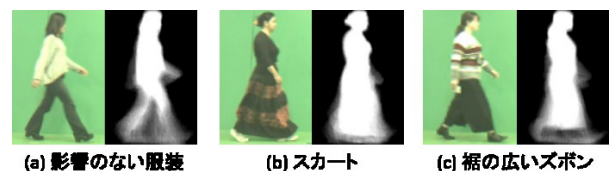


図 3 スカートや裾の広いズボンの GEI に対する影響

手法の有効性を検証するために、同じデータセットを用いてサポートベクター回帰(Support Vector Regression, SVR)及び1in-GEINetによるスクラッチからの学習も行った。1in-GEINetはファインチューニングと同じパラメータ設定をした。SVRの入力として、GEIに主成分分析(Principal Component Analysis, PCA)を適用して、情報量保持率90%になるように次元圧縮したデータを入力データとして使用した。

3.3 実験結果

実験結果の評価には、被験者実験での計測値と学習したネットワークによる推定値の平均絶対誤差(Mean Absolute Error, MAE)を用いた。MAEを計算する前には、各値の身長による正規化した値から元の値に戻した。

表 3 推定手法ごとの各指標の MAE. 部位別 LBM/BFM における部位名は RA: 右腕, LA: 左腕, T: 体幹, RL: 右脚, LL: 左脚, L: 四肢.

Indicator	GEINet, Fine-tuning		GEINet, Scratch		SVR	
	All	Omitted	All	Omitted	All	Omitted
Weight	3.84	3.95	8.49	7.71	6.96	6.67
TBW	1.93	1.95	6.05	4.05	3.37	3.37
Protein	0.54	0.54	1.56	1.15	0.89	0.94
Minerals	0.19	0.19	0.61	0.48	0.30	0.30
BFM	3.03	2.93	3.41	3.39	6.14	5.58
SLM	2.41	2.49	8.07	4.90	4.39	4.00
FFM	2.64	2.55	7.82	5.44	4.62	4.52
SMM	1.64	1.71	4.15	3.21	2.76	2.96
LBM_RA	0.22	0.22	0.48	0.49	0.38	0.43
LBM_LA	0.22	0.22	0.46	0.42	0.38	0.38
LBM_T	1.24	1.20	2.61	2.75	2.23	2.13
LBM_RL	0.49	0.47	1.51	1.32	0.76	0.73
LBM_LL	0.46	0.47	1.40	1.25	0.73	0.79
BFM_RA	0.27	0.26	0.31	0.29	0.56	0.52
BFM_LA	0.27	0.26	0.30	0.30	0.52	0.51
BFM_T	1.66	1.59	1.87	1.88	3.01	2.62
BFM_RL	0.49	0.47	0.61	0.58	0.95	0.98
BFM_LL	0.48	0.45	0.59	0.59	0.95	0.89
BMI	1.67	1.58	3.76	2.84	2.94	2.64
PBF	3.80	3.69	4.85	4.70	7.55	6.74
IBS	4.05	3.80	15.53	12.57	7.00	6.59
BMR	269.86	212.63	111.99	114.47	129.00	144.47
WHR	0.04	0.04	0.22	0.18	0.15	0.06
VFL	1.71	1.57	1.81	1.62	3.05	2.94
OD	7.48	6.74	13.93	11.29	13.19	10.64
LBM_L	1.27	1.32	3.23	2.85	2.37	2.03

表 3 は各手法, 各指標で得られる MAE を示す. 各指標に対する推定精度は図 4 が示したように, 指標の大半は誤差が 5~7% であるが, BFM や PBF, VFL などのような体脂肪を表している指標は誤差が大きく, 18~30% に達している. 他にも, 質量で表す FFM のみ使用して定数との四則演算で計算しながら, 身長を用いた正規化が行われていない BMR も誤差が大きい. さらに, 推定手法の精度は図 5 が示したように, ほとんどの指標はファインチューニングした GEINet の結果が最もよく, スクラッチから学習したネットワークは SVR による推定よりも誤差が大きいが, 前述の体脂肪関係の指標のみスクラッチから学習してもファインチューニングしたものとは大きな差はなく, SVR の結果よりよかった. それから, 服装の影響に関しては図 6 が示したように, いずれの手法に対しても, GEI の前景面積に影響するような服装を着用した被験者を除いたデータセットの方が, わずかではあるが精度がより高い.

この結果に対して, 我々は以下のように考察する. 事前学習は体組成や健康指標の推定精度を向上させる効果がある. 体脂肪関係の指標の推定において, 事前学習が精度向上の効果薄い原因は, 事前学習する際に使用した 4 つの歩行特徴のいずれに対しても体の厚みは重要ではないためであると考えられる. 特に, ファインチューニングでは腕の前振り-腕の後振りの組み合わせを使用したが, この 2 つの情報は GEI の動的特徴のみ評価している. 体脂肪は運動に対する負担になると考えられるが, その変化は緩やかであるため, 人間は普段の生活中でその負担に慣れており, 歩容への影響は少ないと考えられる. また, GEI の前景面積への影響を取り除いても精度向上が小さいことについて, 事前学習において GEI の静的特徴への評価についての重みが小さいからだと考えられる.

4. おわりに

本研究は, 歩行映像解析に基づく体組成推定の手法を提案した. GEINet を歩行特徴に対して事前学習させることで, 推定精度の向上を成功した. 推定結果について, 体脂肪関係以外の指標は比較的に高い精度で推定できた. また, 適切な服装を着用することで推定精度を少ないながらも向上できることを証明した.

今後の課題として, 体脂肪を表す指標の推定精度を上げるが挙げられる. そのために, GEI の前景面積を評価するような特徴を事前学習させることや, それに応じてファインチューニングにおけるプリミティブネットワークの使用方法を見直したい. また, データ数を増やすため, 今回でわかったデータセットの問題点を解消しつつ, データ収集を続行する.

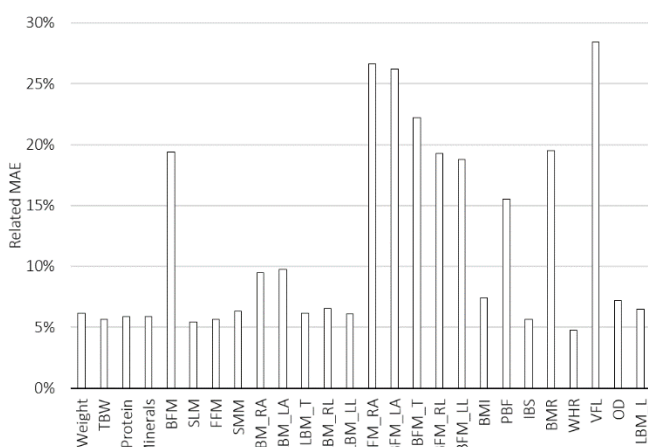


図 4 GEINet (ファインチューニング, Omitted) の相対誤差: 各指標の MAE が計測値平均に対する割合

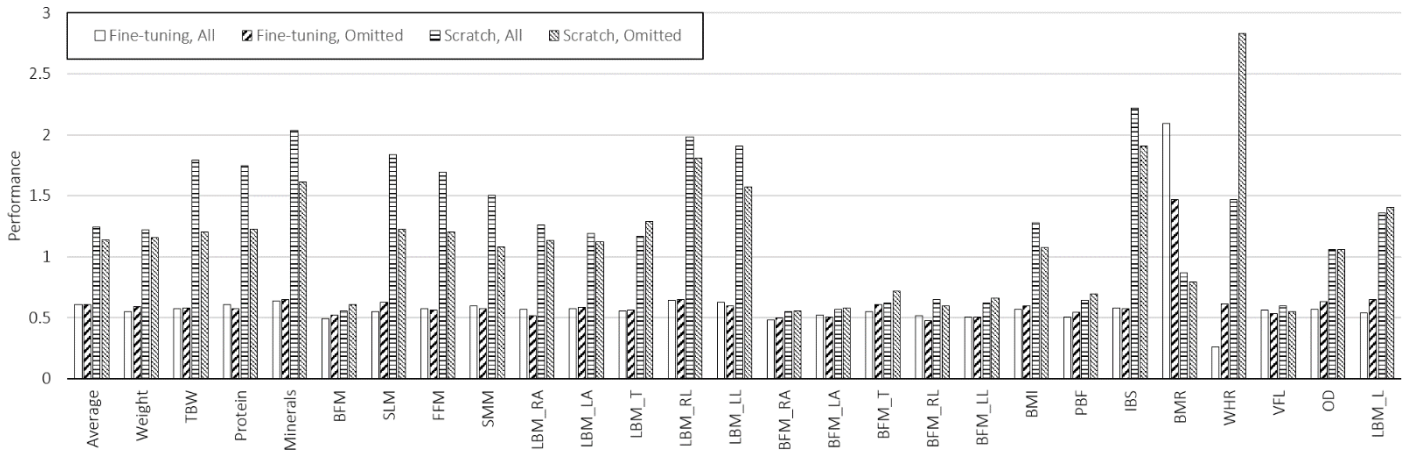


図 5 推定手法ごとの性能：GEINet で得られた MAE をそれぞれ
同じデータセット，同じ指標を SVR における推定で得られた MAE に対する割合

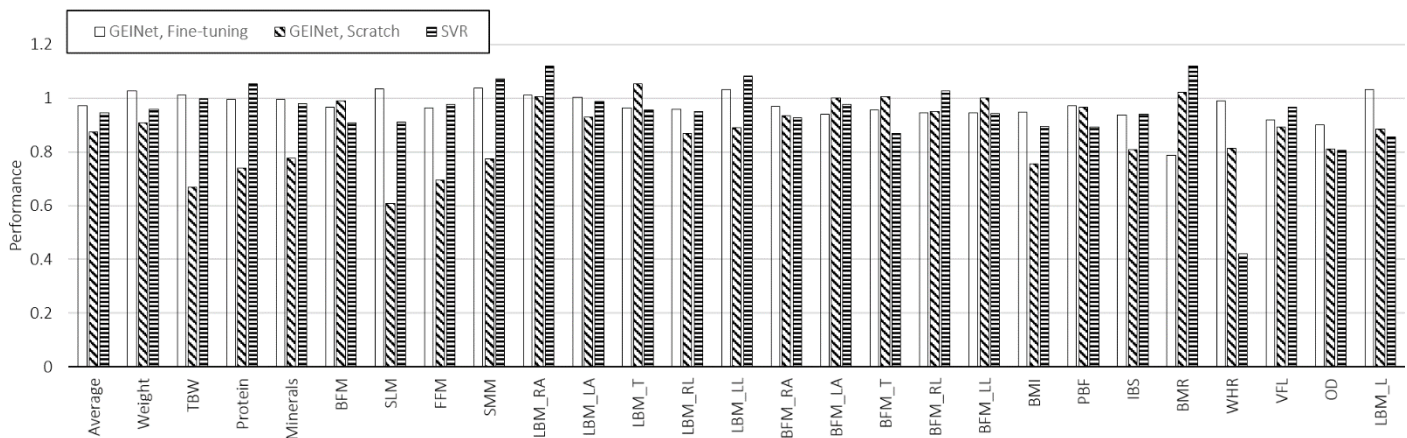


図 6 服装が精度に対する影響：Omitted の MAE を All の MAE に対する割合

参考文献

- [1] A. S. Jackson and M. L. Pollock, "Practical Assessment of Body Composition," *The Physician and Sportsmedicine*, vol. 13, no. 5, pp. 76–90, May 1985.
- [2] R. Baumgartner, "Body composition in healthy aging.," *Annals of the New York Academy of Sciences*, vol. 904, pp. 437–448, 2000.
- [3] K. Sanada *et al.*, "Differences in body composition and risk of lifestyle-related diseases between young and older male rowers and sedentary controls," *Journal of Sports Sciences*, vol. 27, no. 10, pp. 1027–1034, Aug. 2009.
- [4] J. S. Brach, E. M. Simonsick, S. Kritchevsky, K. Yaffe, and A. B. Newman, "The Association Between Physical Function and Lifestyle Activity and Exercise in the Health, Aging and Body Composition Study," *Journal of the American Geriatrics Society*, vol. 52, no. 4, pp. 502–509, 2004.
- [5] R. F. Kushner and D. A. Schoeller, "Estimation of total body water by bioelectrical impedance analysis," *Am J Clin Nutr*, vol. 44, no. 3, pp. 417–424, Sep. 1986.
- [6] H. C. Lukaski, W. W. Bolonchuk, C. B. Hall, and W. A. Siders, "Validation of tetrapolar bioelectrical impedance method to assess human body composition," *Journal of Applied Physiology*, vol. 60, no. 4, pp. 1327–1332, Apr. 1986.
- [7] S. Yu, T. Tan, K. Huang, K. Jia, and X. Wu, "A Study on Gait-Based Gender Classification," *IEEE Transactions on Image Processing*, vol. 18, no. 8, pp. 1905–1910, Aug. 2009.
- [8] Y. Makihara, H. Mannami, and Y. Yagi, "Gait Analysis of Gender and Age Using a Large-Scale Multi-view Gait Database," in *Computer Vision – ACCV 2010*, 2011, pp. 440–451.
- [9] A. Sakata, Y. Makihara, N. Takemura, D. Muramatsu, and Y. Yagi, "Gait-Based Age Estimation Using a DenseNet," in *Computer Vision – ACCV 2018 Workshops*, 2019, pp. 55–63.
- [10] 松浦拓, 阪下和弘, 満上育久, 大倉史生, 八木康史, "早期認知障害検出に向けた高齢者デュアルタスクデータセットの統計分析," *研究報告コンピュータビジョンとイメージメディア (CVIM)*, vol. 2018-CVIM-212, no. 34, pp. 1–7, May 2018.
- [11] R. Chang, L. Guan, and J. A. Burne, "An automated form of video image analysis applied to classification of movement disorders," *Disability and Rehabilitation*, vol. 22, no. 1–2, pp. 97–108, Jan. 2000.
- [12] G. B. Forbes and J. C. Reina, "Adult lean body mass declines with

- age: Some longitudinal observations,” *Metabolism*, vol. 19, no. 9, pp. 653–663, Sep. 1970.
- [13] Ju Han and Bir Bhanu, “Individual recognition using gait energy image,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316–322, Feb. 2006.
- [14] K. Shiraga, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, “GEINet: View-invariant gait recognition using a convolutional neural network,” in *2016 International Conference on Biometrics (ICB)*, 2016, pp. 1–8.
- [15] C. Xu, Y. Makihara, G. Ogi, X. Li, Y. Yagi, and J. Lu, “The OU-ISIR Gait Database comprising the Large Population Dataset with Age and performance evaluation of age estimation,” *IPSI Transactions on Computer Vision and Applications*, vol. 9, no. 1, p. 24, Dec. 2017.
- [16] Y. Makihara, M. Okumura, Y. Yagi, and S. Morishima, “The Online Gait Measurement for Characteristic Gait Animation Synthesis,” in *Virtual and Mixed Reality - New Trends*, 2011, pp. 325–334.