

CNNとRNNの統合モデルによる車いす加速度データを用いた路面状態の評価 Evaluation of Road Surface Condition using Wheelchair Acceleration Data by Integrated Model of CNN and RNN

門田 流聖[†] 渡邊 拓実^{††} 佐藤 剛^{††} 岩澤 有祐^{†††} 矢入 郁子^{††}
Ryusei Kadota Takumi Watanabe Go Sato Yusuke Iwasawa Ikuko Eguchi Yairi

1. はじめに

歩道のアクセシビリティを確保し、高齢者や障がい者などの移動時の負担を軽減することは、重要な社会的課題の一つである。情報通信技術を用いた課題解決方法の一つに、歩道のアクセシビリティ情報を移動弱者に提供するための大規模な地理情報システムとしてのアクセシビリティマップを開発するという方法がある[1, 2]。アクセシビリティ情報を考慮した最適経路を推薦するパーソナライズドアクセシビリティマップ(PAM)の概念も提案されている[3]。アクセシビリティマップ開発には、歩道のアクセシビリティ情報をどのように大規模に収集するか、という課題がある。データ収集方法には現在、歩道が撮影された画像等から専門家が事例ごとに評価するもの[4]や、クラウドソーシングを利用して歩道の撮影や評価を一般人の協力を募って行うもの[5, 6]などがある。これらのマンパワーに基づいた手法では、時間的・金銭的成本を要することから大規模な情報収集が困難という問題があり、公開されているアクセシビリティ情報の多くは、大学や駅などの公共性の高い施設とその周辺に留まっているのが現状である。

スマートフォンやスマートウォッチのような多彩なセンサを搭載した電子デバイスの普及により近年、人間行動を計測し活用する動きが活発化している[7, 8]。車いすユーザーが持ち歩く電子デバイスが計測する人間行動データからそのユーザーが通った路面状態を推定することができれば、ローコストで大規模なアクセシビリティ情報の収集が可能になる。そこで筆者らはこれまで、車いすに設置した加速度センサの観測値が路面の凹凸のパターンや傾度などに影響されることに着目し、機械学習を用いて路面状態を推定するシステムを提案してきた。

身体に取り付けた加速度計により計測された生の人間行動データに機械学習を適用して人間行動認識を行う研究には、次のようなものがある。日常生活や工場での組み立て作業などのデータセットを用いて人間行動認識に対する各種機械学習方法の可能性の調査を行ったもの[9]、これらの認識精度の向上を目指したもの[10, 11]、時系列データを画像化して学習の効率化を目指したもの[12]、疾病による不随意的な体の振動と随意的な運動とを区別することを目的としたもの[13]、畳み込みニューラルネットワーク(CNN: Convolutional Neural Network)とリカレントニューラルネットワーク(RNN: Recurrent Neural Network)を統合して時間的依存関係を抽出したもの[14]などである。本稿では、CNN

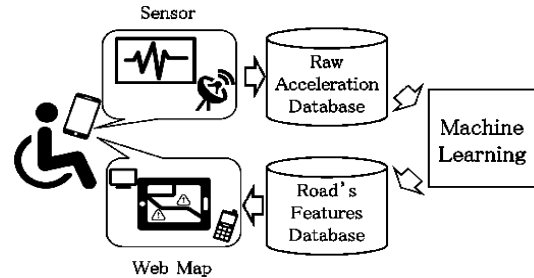


図1 車いすセンシングと機械学習による路面状態の自動評価システム

とRNNの統合モデルを用いて、加速度センサデータから時間的依存関係も考慮した路面状態を表現している特徴量を抽出することで路面状態分類タスクの精度向上を試みる。

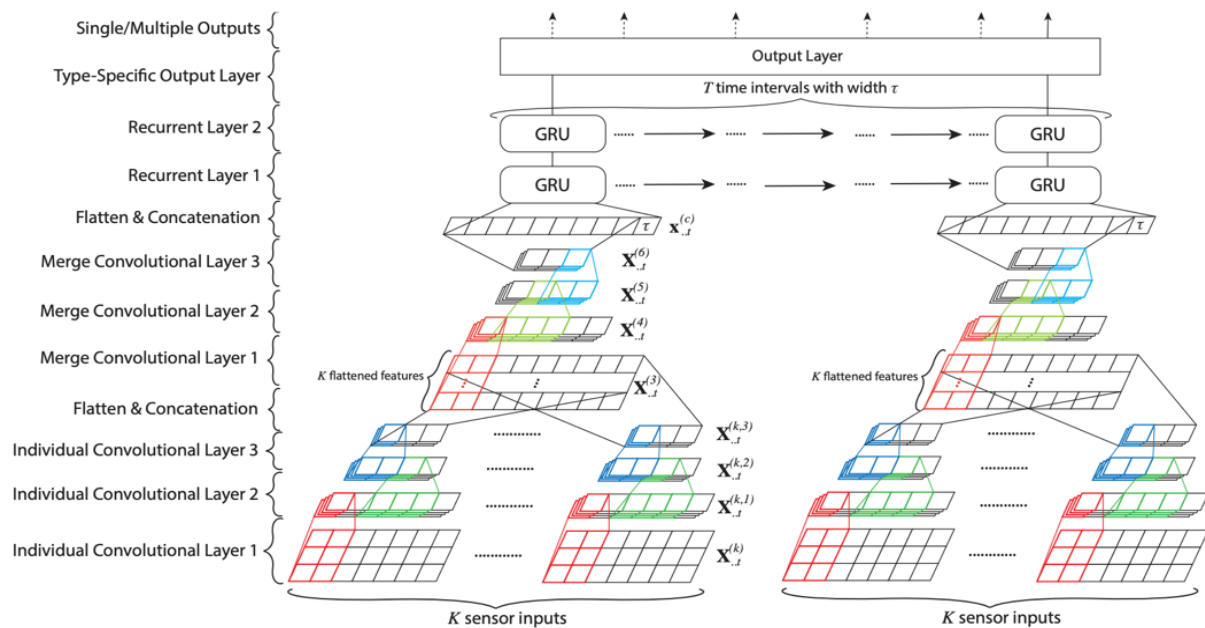
2. 提案手法の概要

2.1 システム構成

図1にシステムの概要図を示す。提案システムは、車いすに設置した加速度センサ(iPod touch)により走行時の加速度センサデータを収集し、機械学習を用いて情報を抽出した後、抽出した路面情報を蓄積し地図上に可視化するものである。

車いすに設置した加速度センサによる加速度センサデータから路面の情報を抽出するにあたり、生データから路面状態による影響を推定するのは容易ではない[15, 16]。そのため、観測された加速度センサデータから路面状態を表す指標に変換することが重要となる。福島らは、加速度センサデータを振動加速度レベル(VAL: Vibration Acceleration Label)に変換することで路面状態に対するユーザーの主観を抽出することを試みている[17]。しかし、加速度センサデータは系列データであるので生データのままで解釈がしにくく、ユーザーの主観を推定する指標を人間が設計するのは極めて困難である。長峯らは、手動車いすの漕ぎ行動に注目し、加速度センサデータのP-P値と心拍データの相関から車いす利用者の負担度合いを評価している[18]。岩澤らは、機械学習を用いて加速度センサデータの分類器を作成することで、路面状態をいくつかの離散的なクラスで表現する方法を提案している[19]~[21]。高橋らは、路面状態を表すいくつかの離散的なクラスのラベル付けが施された加速度センサデータにより学習されたDCNNを利用して、路面状態を表現していると期待される特徴量を取り出し、事前に施されたラベル以上の特徴を取得することを試みている[23]。

[†] 上智大学理工学部情報理工学科
Faculty of Science and Technology, Sophia University
^{††} 上智大学大学院理工学研究科理工学専攻情報学領域
Graduate School of Science and Engineering, Sophia University
^{†††} 東京大学大学院工学系研究科技術経営戦略学専攻
Graduate School of Technology Management for Innovation

図2 DeepSenseの主な構造¹

2.2 DeepSenseの構造

本稿では速度センサデータから路面の情報を抽出するためにDeepSenseを用いる。DeepSenseはShuochaoらによって提案されたCNNとRNNを統合したモデルであり、人間行動認識のタスクにおいて高い分類性能を持っている[14]。図2にDeepSenseの主な構造を示す。ただし、ベクトルを太字の小文字(\mathbf{x} , \mathbf{y} など)、行列とテンソルは太字の大文字(\mathbf{X} , \mathbf{Y} など)で表し、テンソル $\mathbf{X}_{\dots t}$ は3番目の軸に沿った t 番目の行列を表す。また、DeepSenseでは K 個の異なるセンサがあると仮定しているが、本研究では加速度センサは1つだけ用いているため $K=1$ である。測定値の行列、タイムスタンプのベクトルをそれぞれ、 $\mathbf{V}_t^{(k)} \in \mathbb{R}^{d^{(k)} \times n^{(k)}}$ 、 $\mathbf{u}_t^{(k)} \in \mathbb{R}^{d^{(k)} \times n^{(k)}}$ と表す。ただし、 $d^{(k)}$ は測定値の次元数、 $n^{(k)}$ は測定数を表す。入力測定値 \mathbf{V} と時間 \mathbf{u} を時間に沿って重複しないように分割し、幅 τ の $\mathbf{W} = \{(\mathbf{V}_t^{(k)}, \mathbf{u}_t^{(k)})\}$ を生成する。 \mathbf{W} の各要素にフーリエ変換を適用し $\mathbf{X}^{(k)} \in \mathbb{R}^{d^{(k)} \times 2f \times T}$ を生成する。ただし、 f は振幅位相からなる周波数領域の次元数、 T は時間間隔の数、つまり $T=|\mathbf{W}|$ とする。以上の前処理を行った $\mathbf{X} = \{\mathbf{X}^{(k)}\}$ をDeepSenseの入力とする。

DeepSenseは入力データをCNNとRNNで処理する。CNNは、個々のセンサの特徴量を畳み込むパートとそれをマージして畳み込みをするパートに分けられる。最初にセンサの測定値の局所的な特徴と測定値の次元との間の特徴をとらえるために($d^{(k)}$, conv1)の2dフィルタをかけ、 $\mathbf{X}_{\dots t}^{(k,1)}$ を得る。次に(1, conv2)の1dフィルタをかけ、 $\mathbf{X}_{\dots t}^{(k,2)}$ を得た後、(1, conv3)の1dフィルタをかけ、 $\mathbf{X}_{\dots t}^{(k,3)}$ を得ることでより高度な特徴を獲得する。その後 $\mathbf{X}_{\dots t}^{(k,3)}$ を平らにしてベクトル $\mathbf{x}_{\dots t}^{(k,3)}$ を得る。 K 個の $\mathbf{x}_{\dots t}^{(k,3)}$ を結合して行列 $\mathbf{X}_{\dots t}^{(3)}$ を作り、

これを次の入力とする。マージして畳み込みをするパートでは、(K , conv4)の2dフィルタをかけ $\mathbf{x}_{\dots t}^{(4)}$ で K 個のセンサ間の相互作用を学習し、次に(1, conv5)をかけ、 $\mathbf{x}_{\dots t}^{(5)}$ を得た後に(1, conv6)をかけ $\mathbf{x}_{\dots t}^{(6)}$ を得てより高度な特徴量を得る。それを平らにしてベクトル $\mathbf{x}_{\dots t}^{(f)}$ を得る。 $\mathbf{x}_{\dots t}^{(f)}$ と τ を連結して $\mathbf{x}_{\dots t}^{(c)}$ を作り、これをRNNの入力とする。RNNでは、 T 個の $\mathbf{x}_{\dots t}^{(c)}$ を2層のstacked GRUに入力し、 $\mathbf{x}_{\dots t}^{(r)}$ を得て、これをもとに予測確率を得る。

3. DeepSenseによる分類結果の評価

検証に用いられた車いすセンサデータは、6名の手動車いすユーザ、3名の電動車いすユーザ、計9名の手動車いすユーザが東京都四ツ谷駅周辺の約1.4kmの経路を走行した際の行動を、車いすの座席下側に設置した加速度センサ(iPod touch)により計測したものである。位置情報として、実験時の2012年11~12月時に1機の試験的運用が行われていた準天頂衛星による測位データを取得した。利用した加速度センサではx軸、y軸、z軸の3軸方向の加速度値が50Hzでサンプリングされ合計1,341,602サンプル(約7.5時間)が得られた。また、加速度サンプルを取得した場所がどのような状況であったかを確認できるように、実験中は被験者の走行している様子とその時の走行路面が動画撮影された。経路全体の傾向として平坦な歩道が多く、建物等に設置された車いす用スロープの上り下りに問題のないユーザであれば身体への過度な負担や事故の危険性は発生しにくいと考えられる経路となっている。

データセットはx軸、y軸、z軸の3軸の加速度値が計測された地点がどのような路面状態であったかを実験動画を目視することによって、坂/縁石/点字ブロック/その他、の4種類のラベルを付した。この四つを選択した理由は、それぞれ順に、継続する傾斜/突発的な段差/継続する凹凸、という路面の代表的な特徴を表していると考えられること、分類学習に十分な地点数が確保できる(全ラベル付きデータ数に対して、どの種類のラベルのデータも最低1%程度

¹ Shuochao Yao et al. 著, ACM DL 出版, p. 3 の Figure 1 より転載 [14]

表 1 既存手法と DeepSense の比較

Method	F-score	Accuracy (%)
Raw+k-NN [22]	28.2	75.5
MV+k-NN [22]	45.0	69.9
Heuristic+SVM [22]	51.9	80.3
Heuristic+MLP [22]	56.5	78.2
DCNN [22]	66.0	83.9
DeepSense	55.9	81.6

は確保されている場合に分類精度が高かった)ためである。また、3 軸の加速度データを入力するにあたり、ラベル付きデータを、ウィンドウ幅 450 サンプル (約 9 秒)、重なり率 0.5 としたスライディングウィンドウ方式により 5949 個切り出した。この幅と重なり率は、同じデータを用いた研究によって適した値を探索して決定した[21]。

次に、4 種類の路面分類タスクの分類精度について既存手法と DeepSense を比較し、獲得された特徴量が路面状態とどの程度対応しているのかを説明する。表 1 は、既存手法と DeepSense の、坂/縁石/点字ブロック/その他、の分類タスクにおける認識精度を比較したものである。ここで、各表の値はそれぞれ Leave-One-Subject-Out 法により得られた 9 回の試行の平均値である。Leave-One-Subject-Out 法とは、8 名分のデータセットで訓練したモデルを残りの 1 名分のデータセットで評価することを繰り返す方法であり、学習データに含まれないユーザのデータで評価を行うことで、未知のユーザに対する認識精度を検証することができる評価方法である。評価指標には各クラスのマクロ平均 F 値(F-score)及び正解率(Accuracy)を用いた、今回の分類問題はクラスインバランスであり、平均 F 値が高いことはデータセット中に出現する地点が少ないバリアについてもよく認識できていることを意味する。なお、既存手法では前処理としてウィンドウ幅 450 サンプル、重なり率 0.1 としたスライディングウィンドウ方式により 29,727 個切り出した[22]。比較に用いた手法の詳細は次のとおりである。

- **Raw+k-NN**: 9 人分のデータセットを入力に、k-近傍法 (KNN: K-Nearest-Neighbor)で分類した。なお入力には、データごとに 450 ウィンドウ分の 3 軸加速度値を単純に結合したベクトルを用いた。K の値としては K=1, 5, 10, 15, 20, 25, 30 を検証し、検証データにおいて最も精度の高かった K=1 を利用した。
- **MV+k-NN**: Raw+k-NN と同様に、K=1 の KNN を利用した。入力には、450 ウィンドウ分の 3 軸加速度の xyz 各軸の平均値及び標準偏差の 6 つを利用した。

表 2 クラス毎の F 値による評価

Method	Per-Class F-score			
	Slope	Curb	TI	Oths
Raw+k-NN [22]	15.8	7.93	3.10	86.1
MV+k-NN [22]	35.8	42.9	18.8	82.5
Heuristic+SVM [22]	19.6	63.2	35.8	89.1
Heuristic+MLP [22]	26.7	65.6	46.4	87.3
DCNN [22]	47.4	77.3	48.5	90.7
DeepSense	42.3	62.4	29.0	89.8

表 3 クラス毎の適合率による評価

Method	Per-Class Precision (%)			
	Slope	Curb	TI	Oths
Raw+k-NN [22]	31.2	59.8	14.2	76.2
MV+k-NN [22]	40.9	42.5	18.8	82.9
Heuristic+SVM [22]	49.1	70.2	46.9	85.1
Heuristic+MLP [22]	43.3	66.6	48.9	88.7
DCNN [22]	55.6	81.6	65.4	89.3
DeepSense	56.3	76.0	64.7	84.3

- **Heuristic**: 入力には、3 軸加速度の xyz 各軸に対して平均、標準偏差、最大、最小、ゼロクロッシング、階差の平均、階差の標準偏差、階差の最大、階差の最小、FFT 周波数成分(サンプリング周波数は 50Hz としている)、FFT 周波数成分のエネルギー、エントロピーを算出し、利用した。分類には、サポートベクターマシン(SVM: Support Vector Machine)、多層パーセプトロン(MLP: Multilayer Perceptron)を用いた。
- **DCNN**: 1. 入力層(3 軸加速度データ)、2~4. 畳み込み層、5. 全結合層、6. 出力層の全 6 層から構成される。ニューラルネットワークの学習には Adam を利用し、学習率は 0.0001 とした。活性化関数には出力層で Softmax 関数を利用したものを除き、ReLU を利用した。

4. 結果

表 1 が示すように DeepSense を利用した手法は DCNN よりも高い精度を得ることができなかった。DCNN よりも F 値で 10.1 ポイント、正解率で 2.3 ポイント低い値となった。正解率は DCNN と近いが、F 値で大きく下回っていることが分かる。また、Heuristic+MLP と比較すると、正解率は DeepSense の方が高いが、F 値は低い値になっていることがわかる。それ以外の手法と比較すると F 値も正解率も DeepSense が上回っている。表 2 に 4 クラス (坂 Slope, 縁石 Curb, 点字ブロック TI, その他 Oths) 毎の F 値、表 3 にそれぞれの適合率を示す。F 値と適合率でほとんどのクラスで DCNN の精度には及ばなかった。また、点字ブロックでの F 値が顕著に低いことも分かる。これは今回用いたデータのインバランス性に起因すると思われる。以上のように、本研究では既存手法より詳細な路面状態を捉えることはできなかった。

5. おわりに

本稿は車いす加速度データを用いた路面状態分類タスクの向上を目的として、CNN と RNN の統合モデルである DeepSense を適用した。DeepSense は筆者らがこれまで適用してきた DCNN を用いた手法の分類精度には及ばなかったが、DeepSense は複数のセンサデータを扱うことを想定しており、今回用いたデータは加速度センサ 1 つのみであったため、DeepSense 自体の構造をデータに合うように変更する必要があると思われる。また、今回用いたデータはクラスによってデータ数が大きく異なる。このようなインバランスデータの多クラス分類問題は、高い精度での分類が

難しいとされるタスクである。本稿ではインバランス性を考慮した CNN と RNN の統合モデル改良について未着手である。今後はデータインバランス性を考慮することや、モデルの構造を工夫するなどして、DCNN だけでは抽出することができない時間的依存関係を考慮できる教師あり学習のより高精度な手法を提案する必要があると考える。

謝辞

実験に参加してくださった方々に感謝します。本研究は科学研究費補助金、基盤研究(B)17H01946 と基盤研究(B)20H04476 のもとで実施されました。

参考文献

- [1] M. Laakso, T. Sarjakoski, and L. T. Sarjakoski, "Improving accessibility information in pedestrian maps and databases," *Cartographica: Int. J. Geographic Information and Geovisualization*, Vol. 46, No. 2, pp. 101–108, 2011.
- [2] H. Matthews, L. Beale, P. Picton, and D. Briggs, "Modelling access with GIS in urban systems (MAGUS): Capturing the experiences of wheelchair users," *Area*, Vol. 35, No.1, pp.34–45, 2003.
- [3] H. A. Karimi, L. Zhang, and J.G. Benner, "Personalized accessibility map (PAM): A novel assisted wayfinding approach for people with disabilities," *Annals of GIS*, Vol.20, No.2, pp.99–108, 2014.
- [4] C. Ponsard and V. Snoeck, "Objective accessibility assessment of public infrastructures," in *Computers Helping People with Special Needs*, pp.314–321, Springer, 2006.
- [5] K. Hara, "Scalable methods to collect and visualize sidewalk accessibility data for people with mobility impairments," *Proc. Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology*, pp.1–4, 2014.
- [6] C. Cardonha, D. Gallo, P. Avegliano, R. Herrmann, F. Koch, and S. Borger, "A crowdsourcing platform for the construction of accessibility maps," *Proc. 10th International Cross-Disciplinary Conference on Web Accessibility*, p.26, 2013.
- [7] M. Swan, "The quantified self: Fundamental disruption in big data science and biological discovery," *Big Data*, Vol.1, No.2, pp.85–99, June 2013.
- [8] K. Nagamine, Y. Iwasawa, Y. Matsuo, and I.E. Yairi, "An estimation of wheelchair user's muscle fatigue by accelerometers on smart devices," *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, pp.57–60, UbiComp/ISWC'15 Adjunct, ACM, New York, NY, USA, 2015.
- [9] T. Plötz, N.Y., Hammerla, and P. Olivier, "Feature learning for activity recognition in ubiquitous computing," *Proc. International Joint Conference on Artificial Intelligence*, Vol.22, p.1729, 2011.
- [10] M. Zeng, L.T. Nguyen, B. Yu, O.J. Mengshoel, J. Zhu, P. Wu, and J. Zhang, "Convolutional neural networks for human activity recognition using mobile sensors," *6th Int. Conf. on Mobile Computing, Applications and Services*, pp.197–205, Austin, TX, USA, Nov. 2014.
- [11] J.B. Yang, M.N. Nguyen, P.P. San, X.L. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition," *Proc. 24th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, pp.25–31, Buenos Aires, Argentina, July 2015.
- [12] W. Jiang and Z. Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks," *Proc. 23rd ACM Int. Conf. on Multimedia*, pp.1307–1310, Brisbane, Australia, Oct. 2015.
- [13] N.M. Rad and C. Furlanello, "Applying deep learning to stereotypical motor movement detection in autism spectrum disorders," *2016 IEEE 16th Int. Conf. on Data Mining Workshops (ICDMW)*, pp.1235–1242, Barcelona, Spain, Dec. 2016.
- [14] Y. Shuochao, et al., "Deepsense: A unified deep learning framework for time-series mobile sensing data processing," *Proceedings of the 26th International Conference on World Wide Web*, 2017.
- [15] O.D. Lara and M.A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tutorials*, Vol.15, No.3, pp.1192–1209, Nov. 2012.
- [16] H. Liu, T. Taniguchi, Y. Tanaka, K. Takenaka, and T. Bando, "Visualization of driving behavior based on hidden feature extraction by using deep learning," *IEEE Trans. Intelligent Transportation Systems*, Vol.18, No.9, pp.2477–2489, Feb. 2017.
- [17] Y. Fukushima, H. Uematsu, R. Mitsuhashi, H. Suzuki, and I.E. Yairi, "Sensing human movement of mobility and visually impaired people," *Proc. 13th Int. ACM SIGACCESS Conf. on Computers and Accessibility*, pp.279–280, Dundee, Scotland UK, Oct. 2011.
- [18] 長峯 洗弥, 岩澤 有祐, 松尾 豊, 矢入 郁子, "加速度センサを用いた手動車いす走行時の身体的負担評価手法の提案", *信学論(D)*, Vol.J100-D, No.8, pp.773–782, Aug. 2017.
- [19] Y. Iwasawa and I.E. Yairi, "Life-logging of wheelchair driving on web maps for visualizing potential accidents and incidents," *12th Pacific Rim Int. Conf. on Artificial Intelligence*, pp.157–169, Kuching, Malaysia, Sept. 2012.
- [20] Y. Iwasawa, K. Nagamine, Y. Matsuo, and I. Eguchi Yairi, "Road sensing: Personal sensing and machine learning for development of large scale accessibility map," *Proc. 17th Int. ACM SIGACCESS Conf. on Computers & Accessibility*, pp.335–336, Lisbon, Portugal, Oct. 2015.
- [21] Y. Iwasawa, I.E. Yairi, and Y. Matsuo, "Combining human action sensing of wheelchair users and machine learning for autonomous accessibility data collection," *IEICE Trans. Inf. & Syst.*, Vol.E99-D, No.4, pp.115–124, Apr. 2016.
- [22] T. Watanabe, et al., "Weakly Supervised Learning for Evaluating Road Surface Condition from Wheelchair Driving Data." *Information*, No.11, Vol.1, 2020.
- [23] 高橋 宏紀, 長峯 洗弥, 岩澤 有祐, 松尾 豊, 矢入 郁子, "深層畳み込みニューラルネットワークを利用した車いすセンシングデータからの路面状態の定量化", *信学論(D)*, Vol.J100-D, No.6, pp.1009–1021, Aug. 2018.