

深層学習を用いたスキーシミュレータ利用時における 初心者と熟練者の運動特徴抽出手法

渥美 亮祐¹ 金井 秀明¹

概要: スキーは多くの人が経験したことのあるスポーツである。しかし、スキーのトレーニングは他のスポーツに比べ行うことは難しい。なぜならば、スキーは特定の季節(冬)、特定の場所(雪と斜面)という制限があるためである。また、1人では技術を習得しにくいという点である。初めてスキーを滑る人や新たに技術を習得したい人はスキー学校のような場所でスキーの専門家から直接指導を受けることが多い。先行研究では、スキーシミュレータや情報技術を組み合わせた指導方法やトレーニングシステムに関する研究が行われている。しかし、スキー学校の指導からもわかるように、スキーの指導では専門家の経験に大きく依存する。そこで本研究では、スキーシミュレータを活用し、利用時に取得可能なセンシングデータから初心者と熟練者間の特徴を明らかにし、初心者への改善点教示を目指す。特徴調査にあたり1.VRスキー映像表示部分、2.姿勢検出部分、3.荷重取得部分、4.特徴抽出部分の4つからなるシステムを開発し被験者11名、スキー初心者からスキー熟練者まで幅広く集め実験を行なった。

キーワード: スキー、シミュレータ、VR、深層学習、センシングデータ、初心者と熟練者

1. はじめに

文部科学省が出している学習指導要領では「自然とのかかわりの深い雪遊び、氷上遊び、スキー、スケート、水辺活動等の指導については、学校や地域の実態に応じて積極的に行うことに留意すること。」と記載されていることから、多くの小学校もしくは中学校ではスキーの授業を実施しており、1度は経験したことのある人が多い馴染みのあるスポーツである。しかし、スキーのトレーニングは他のスポーツに比べて行うことは難しい。なぜならば、スキーは特定の季節(冬)、特定の場所(雪と斜面)という制限があるためである。また、1人では技術を習得しにくいという点である。初めてスキーを滑る人や新たに技術を習得したい人はスキー学校のような場所でスキーの専門家から直接指導を受けることが多いが、マンツーマンでの指導は少なく技術の習得に時間を要する。このような背景から先行研究では、スキーシミュレータや情報技術を組み合わせた指導方法 [1][2] やトレーニングシステムに関する研究が行われている [3][4]。また、多くのスポーツでは、正しいフォームはスキルの向上につながると考えられているため、熟練者のフォームをキャプチャーし、学習者に熟練者との違いを提示しトレーニングするシステムの開発も行われている

[5][6]。しかし、スキー学校の指導からもわかるように、スキーの指導では専門家の経験に大きく依存する。

そこで本研究では、スキーシミュレータを活用し、利用時に取得可能なセンシングデータから初心者と熟練者間の特徴を明らかにし、初心者への改善点教示を目指す。近年では、センシングデータと深層学習を活用した行動認識等の研究 [7][8] が行われていることから、人間が抽出しにくい特徴を深層学習は見つけてきていると考え、特徴を調査していくにあたり深層学習を活用する。

2. 開発したシステム

本研究で開発したシステムの構成図を図1に示す。本システムは主な機材として、PC(スペックは表1参照)、スキーシミュレータ(図2)、VIVE Pro、Kinect V2を用いる。

システムの利用の流れとしては、スキーヤーにはVIVE Proを着用した状態でスキーシミュレータに乗ってもらう。VIVE ProにはVRスキーを投影し、VRスキーで表示されているプロのスキーヤーの動きに合わせてスキーシミュレータを動かす。その時のスキーヤーの姿勢や荷重のデー

表1 本研究に用いたPCスペック

CPU	i9-9900K
メモリ	DDR4 2666MHz 16GBx4
GPU	RTX2080Ti

¹ 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

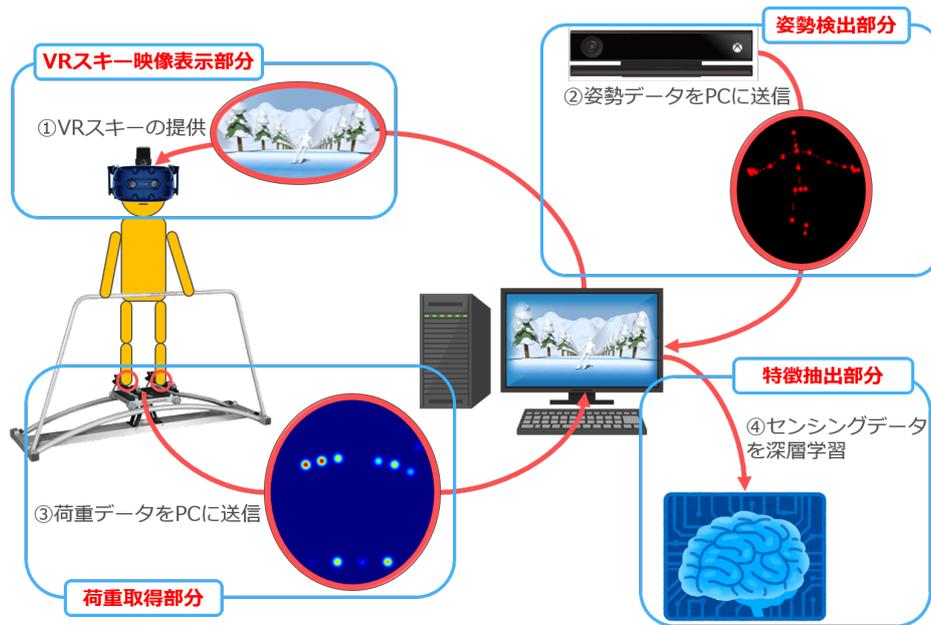


図 1 システム構成図



図 2 本研究に用いたスキーシミュレータ



図 3 圧力センサの配置

データを PC に送信する。その後、深層学習によって初心者と熟練者の特徴を探す。

本システムは、大きく分けて 4 つ要素で構成している。

- (1) VR スキー映像表示部分
- (2) 姿勢検出部分
- (3) 荷重取得部分
- (4) 特徴抽出部分

以上の 4 つである。VR スキー映像表示部分では先行研究で開発された VR スキー [9] を使わせていただいた。姿勢検出部分では Kinect V2 を用いて姿勢のキャプチャを行った。荷重取得部分、特徴抽出部分は以下の節で説明する。

2.1 荷重取得部分

荷重取得部分には、先行研究 [10] を参考に実装を行なった。この先行研究では、片足あたり 8 個ずつ計 16 個の圧力センサを取り付け、15 種類の足裏のジェスチャを識別す

ることに成功している。

先行研究では、靴下に対し圧力センサを取り付けていたが、靴下に対し圧力センサを取り付けるとスキー靴を履く際に取れてしまう可能性があったため、本研究ではスキー靴のインソールに対し圧力センサを取り付けた。センサの取り付け位置は先行研究と同様の位置 (図 3 参照) を採用している。作成した荷重取得装置を図 4 に示す。また、スキー靴を履いた際配線が足に違和感を与えないためにスキー靴のサイドには穴を開け配線を出すようにしている。荷重取得装置と PC 間はシリアルで通信を行い、送られてきた荷重データはヒートマップ化処理を行った (図 5)。

2.2 特徴抽出部分

本研究では深層学習を用いて特徴抽出を行うが、手順としては学習データで学習を行い Grad-CAM[12] という手法でネットワークがどの部分に注目して分類を行ったのかを確認する。Grad-CAM は、ネットワークから出力された特徴量を受け取り、ネットワークがどの部分に注目して分類を行ったのかをヒートマップで可視化する手法である。ネットワークには YOLOv3[11] のネットワークをベースに Feature Pyramid Network[13] の技術を活用し、出力の特



図 4 荷重取得装置

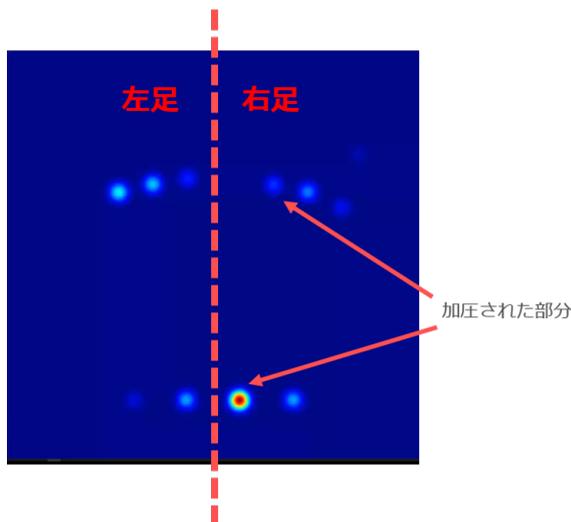


図 5 荷重データに対しヒートマップ化処理をした画像

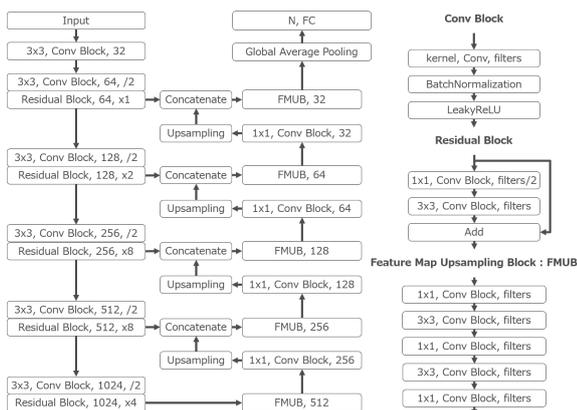


図 6 改良を行ったネットワーク図

微マップを入力した画像サイズと同等のサイズを出力できるように改良を行った(図6)。これにより、Grad-CAMを適用した際に入力画像に対しピクセル単位で着目した特徴部分を抽出できるようになったと考える。

被験者	滑走頻度・日数	滑り方
A	2-3年に1日	パラレル
B	2日	ボーゲン
C	1日	ボーゲン
D	10日	パラレル
E	年1日(10日程度)	パラレル
F	1日	ボーゲン
G	10日	ボーゲン
H	1000日	パラレル
I	3日	ボーゲン
J	10日	ボーゲン
K	年15日(600日程度)	パラレル

3. 実験

開発したシステムを用いて被験者11名を集め実験を行った。被験者11名の滑走頻度・日数や滑り方については表2に示す。ボーゲンはスキー板をハの字型にして方向転換を行うのに対し、パラレルは常にスキー板を平衡に保つことである。

3.1 実験の流れ

被験者は、VIVE Proを着用した状態でスキーシミュレータに乗った。なお、スキー靴にはインソールに対し圧力センサを取り付けたものを着用している。VIVE Proには目の前にお手本となる人が滑っており、被験者はそれに合わせてスキー動作を行う(実験の様子とその時の被験者が見ている映像を図7に示す)。実験中はスキー板に足が固定されているかつVIVE Proを着用しているため、安全を確保するためにスキーシミュレータに付属している手すりを掴んでもらったうえで行ったが、被験者にはあらかじめ手すりに力を入れず足の力でスキーシミュレータを動かすように促した。実験では、まずVRスキーの雰囲気を感じてもらったために1回目は計測をせずに行い、2回目で計測を行った。1回のVRスキーの時間は1分間である。実験では、荷重と姿勢のデータを取得し、図8に示す3種類の画像を生成した。それぞれの画像の右側には取得したセンシングデータ時のVRスキー内の映像を表示している。被験者11名から姿勢、荷重、姿勢+荷重の画像をそれぞれ10285枚ずつ取得することができた。深層学習の学習データには1分間の実験データのうち前半40秒の計6854枚の画像を入力し、テストデータには後半20秒の計3431枚の画像を用いた。

3.2 実験1(深層学習の出力を被験者ごとにした場合)

実験1では、被験者11名から取得したセンシングデータから生成した3種類の画像(姿勢、荷重、姿勢+荷重)を深層学習に入力し被験者11名の分類を行う。本実験の目



図 7 実験の様子

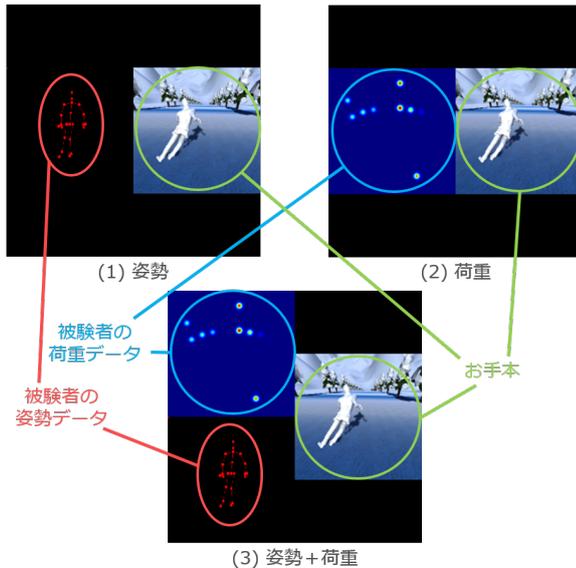


図 8 センシングデータから生成した画像

的は以下の 2 点である。

- 姿勢、荷重、姿勢+荷重の 3 種類の画像のうちどの入力が最も分類精度が高いのかを確認
- 分類結果をもとに Grad-CAM を用いて視覚化を行い、被験者同士で似た傾向があるのかどうかを調査

3.2.1 分類精度結果

姿勢、荷重、姿勢+荷重データに対する分類精度結果を表 3 に示す。姿勢のみ、荷重のみ、姿勢+荷重の組み合わせそれぞれにおいて分類精度は 90% 以上と高い結果となった。平均の精度および標準偏差の値から一番高い精度の入力方法は姿勢+荷重の組み合わせの画像に対してということがわかる。また、姿勢+荷重データの分類結果で唯一 9 割を下回った結果となった被験者 K の誤分類先は被験者 H で、両名ともスキー経験が豊富な方同士で熟練者同士似た傾向があったのではないかと考えられる。

3.2.2 Grad-CAM で視覚化した結果

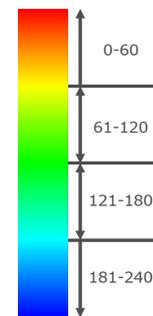
最も分類精度の高かった姿勢+荷重データの画像に対し Grad-CAM を適用した結果が図 9 である。図の見方としては、被験者 B を例に挙げると被験者 B の姿勢データ部分のヒートマップは青色で荷重データ部分には青色以外の色が付いていることから、深層学習が被験者 B と分類した

表 3 姿勢、荷重、姿勢+荷重データに対する分類精度結果 [%]

被験者	姿勢	荷重	姿勢+荷重
A	91.9	93.5	99.4
B	100	94.9	98.1
C	99.3	98.6	100
D	100	97.7	97.4
E	91.7	86.9	96.8
F	100	94.5	99.7
G	96.5	96.5	100
H	93.5	94.8	92.2
I	99.4	80.5	96.9
J	66.0	85.6	98.7
K	73.3	91.9	81.1
平均	92.0(± 11.1)	92.3(± 5.39)	96.4(± 5.28)

表 4 被験者ごとの姿勢、荷重部分の色相値

被験者	姿勢	荷重
A	40	137
B	240	69
C	230	50
D	86	60
E	180	240
F	190	232
G	60	236
H	183	55
I	121	236
J	159	42
K	205	40



判断基準は荷重データ部分にあるということが言える。しかし、このままでは主観で判断していくしかないため、姿勢、荷重データ部分に対し色相値を用いてどれだけヒートマップがかかったのかどうかをまとめた(表 4 参照)。先ほどの被験者 K のセンシングデータを被験者 H と誤分類することが多いことからわかるように被験者 H と K の姿勢および荷重の色相値はほぼ同様の値を示している。姿勢の色相値は高く荷重の色相値は低いことから被験者 H と K は荷重に対し深層学習は特徴として抽出したと言える。姿勢の色相値が高く荷重の色相値が低い被験者は他に 2 名おり、被験者 B と C である。この 2 名はともにスキー初心者であった。色相値だけではそれぞれの被験者の荷重に対しどのような特徴があるのかわからなかったため、計測した荷重マップを確認したところ、被験者 B と C の荷重のかかり方は移動方向と逆側にかかっていたことに対し、被験者 H と K の荷重のかかり方は移動方向に合わせて荷重がかかっているという特徴があった。

3.3 実験 2(深層学習の出力をスキー経験に応じ 4 段階に振り分けた場合)

実験 1 より分類精度は荷重+姿勢データが一番高いことがわかり、初心者および熟練者においてもそれぞれ似た特

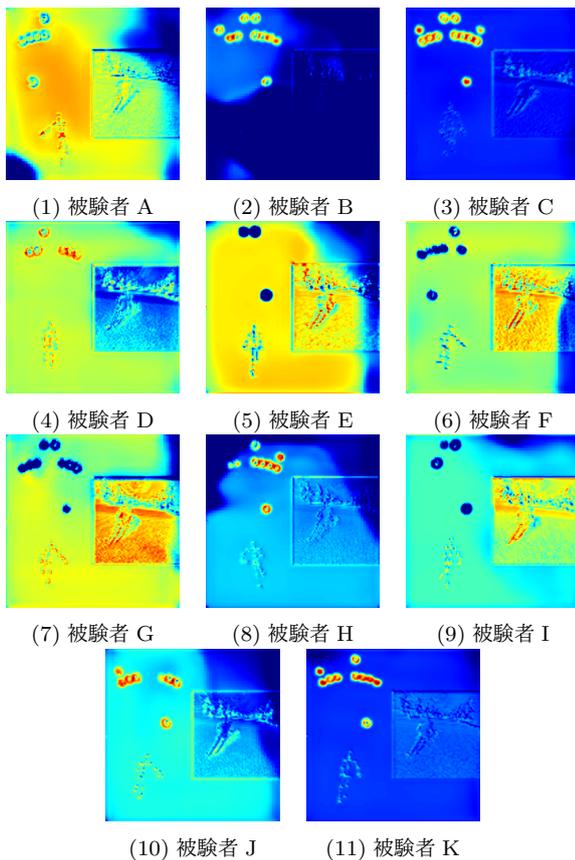


図 9 姿勢+荷重データに対し Grad-CAM を適用した画像

表 5 被験者のスキルレベル

被験者	スキルレベル	
A	ミドル	ビギナー
B	ビギナー	スキー経験：3日以下
C	ビギナー	滑走方法：ボーゲン
D	ミドル	ノービス
E	ミドル	スキー経験：4日以上
F	ビギナー	滑走方法：ボーゲン
G	ノービス	ミドル
H	エキスパート	スキー経験：4日以上
I	ビギナー	滑走方法：パラレル
J	ノービス	エキスパート
K	エキスパート	スキー経験：100日以上
		滑走方法：パラレル

徴があったのではないかと考える。また、VR スキーの映像にも少なからずヒートマップがかかってしまった。そこで、実験2ではVR スキーの映像を取り除いた画像を入力、また被験者のスキー経験をもとに表5のように4つの段階に振り分け、姿勢+荷重データを用いて深層学習に学習させ分類を行う。本実験の目的は以下の点である。

- スキー経験に応じて4段階に振り分けた結果、それぞれの段階で姿勢や荷重に対し特徴の傾向があるのか調査

3.3.1 分類精度結果

実験1では、被験者11名のテストデータをどの被験者のデータなのかを予測し分類を行ったが、実験2では被験

者11名のテストデータに対し、入力した被験者のデータがビギナー、ノービス、ミドル、エキスパートのどの段階のデータかを予測し分類を行った。スキー経験に応じた4段階の分類精度結果を表6に示す。実験2においてもそれぞれの被験者に対し95%以上の精度で分類することができた。

3.3.2 Grad-CAM で視覚化した結果

VR スキー映像を取り除いた姿勢+荷重データの画像に対し Grad-CAM を適用した結果が図10である。また、色相値を用いてどれだけヒートマップがかかったのかわかるとまとめた表を表7に示す。結果は、エキスパートの被験者に対しては荷重データの特徴として抽出してきたことがわかる。被験者Jはエキスパートの被験者HおよびKと同じように荷重を特徴として抽出してきている。この傾向は実験1と同様の結果である。分類精度結果を見ても、被験者Jはノービスと割り振ったがエキスパートと誤分類する確率が15%程度あったことから、被験者Jはエキスパートの被験者と同じような荷重のかけ方であったと言えるのではないかと考える。しかし、姿勢データに関してはビギナーからエキスパートまで特徴として抽出してきていないことがわかる。原因としては、今回の実験ではVIVE Proを着用したうえで行ったため、被験者は安全確保のためスキーシミュレータの手すりを持った状態で実験を行った。そのため、それぞれの被験者の姿勢に対しあまり変化が起これなかったと考えられる。

3.4 システムの改善点

今回の実験から、システムの改善点が2つ挙げられる。

3.4.1 荷重取得装置の改善

スキー熟練者から、今回計測していない部分(土踏まず等)も計測したほうが良いとアドバイスをいただいたため、足全体をカバーできるように圧力センサの数を増やす必要がある。そのため、図11に改善案を示す。今回の実験からはスキー初心者と熟練者では重心のかかり方が主に違っていたが、このように荷重の取得部分を増やすことで重心のかかり方以外にも初心者と熟練者で荷重をかける部分の違い等を見つけることが可能と考える。

3.4.2 映像表示部分の改善

VIVE Pro を着用状態では被験者の姿勢に対しあまり変化が起きないことがわかった。そのため、VIVE Pro を使用しない形での映像表示部分の改善を行う必要がある。改善案を図12で示す。このように、複数のディスプレイを使用することで、VIVE Pro で映像表示を行っていた理由である没入感をある程度維持したまま代替可能であると考ええる。また、ヘッドマウントディスプレイを使用しないため、スキーヤーの視線の計測等も可能となりより多くの情報を取得することが可能となる。

表 6 スキー経験に応じた 4 段階の分類精度結果 [%]

被験者	スキルレベル	ビギナー	ノービス	ミドル	エキスパート	精度
A	ミドル	0.3	0.0	99.4	0.3	99.4
B	ビギナー	99.0	1.0	0.0	0.0	99.0
C	ビギナー	100	0.0	0.0	0.0	100
D	ミドル	0.0	0.0	100	0.0	100
E	ミドル	2.2	0.0	97.8	0.0	97.8
F	ビギナー	99.4	0.0	0.6	0.0	99.4
G	ノービス	0.3	99.4	0.3	0.0	99.4
H	エキスパート	0.0	4.6	4.2	91.2	91.2
I	ビギナー	93.5	0.0	6.5	0.0	93.5
J	ノービス	1.6	77.9	3.8	16.7	77.9
K	エキスパート	0.0	0.0	7.2	92.8	92.7
平均		—	—	—	—	95.5(± 6.35)

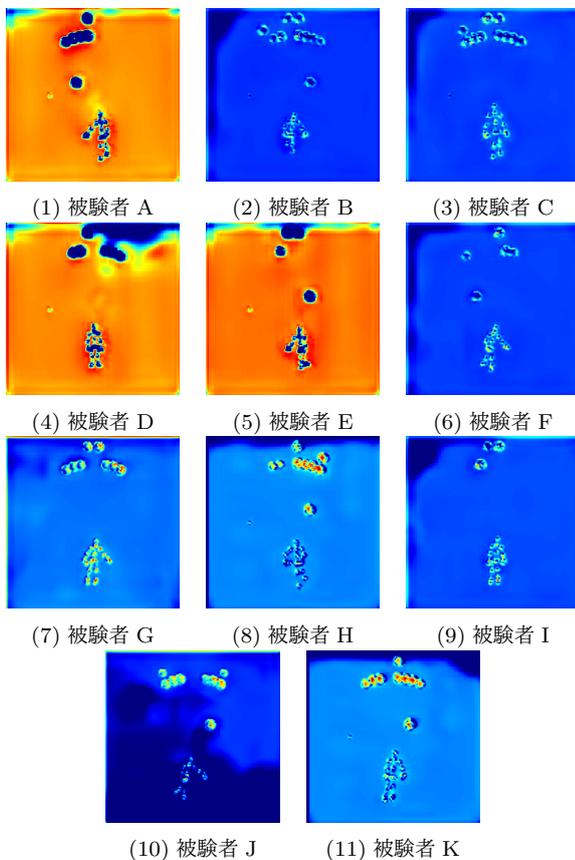


図 10 VR スキーの映像なしでの姿勢+荷重データに対しスキー経験ごとで深層学習したモデルの Grad-CAM の画像

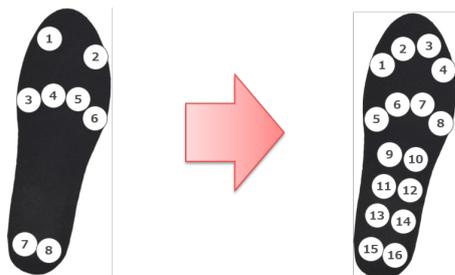


図 11 荷重取得装置の改善

表 7 スキー経験に応じ 4 段階に振り分けた姿勢, 荷重部分の色相値

被験者	姿勢	荷重
A	224	240
B	212	220
C	211	215
D	226	240
E	222	240
F	205	217
G	103	179
H	196	73
I	198	208
J	212	93
K	201	52

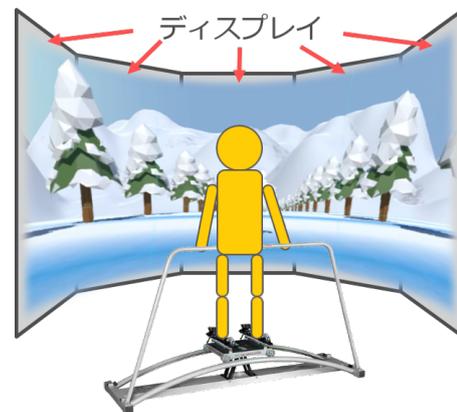
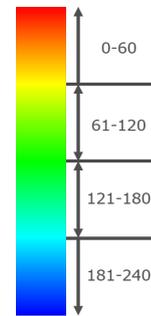


図 12 映像表示部分の改善

3.5 まとめ

開発したシステムを用いて被験者 11 名を集め実験を行った結果、深層学習を用いて被験者ごとに分類する場合もスキー経験に応じ 4 段階に振り分けた場合どちらの場合においても未学習のデータに対して 95%以上の精度で分類することができた。本研究では、Grad-CAM を用いて深層学習のネットワークがどの部分に注目して分類を行ったのか視覚化した結果、荷重データを特徴として抽出してきている

ことがわかった。しかし、特徴として抽出してきているとだけしかわからず、具体的に初心者と熟練者の特徴の違いはこの手法のみではわからなかった。そのため計測した荷重マップを確認したところ、以下の点がわかった。

- 初心者から経験者
 - 荷重が移動方向と逆側にかかっている
 - 片側に荷重がかかりっぱなしである
 - 荷重がかかっていない
- 熟練者
 - 移動方向に合わせて荷重がかかっている

このことから、初心者への改善点として、スキーシミュレータを利用した状態での重心移動の習得が挙げられる。

4. おわりに

本研究では、スキーシミュレータを活用し、利用時に取得可能であった姿勢、荷重データを対象に初心者と熟練者間の特徴を明らかにし、初心者への改善点教示を目指した。

開発したシステムは大きく分けて、1.VR スキー映像表示部分、2. 姿勢検出部分、3. 荷重取得部分、4. 特徴抽出部分の4つの要素で構成した。

開発したシステムを用いて被験者11名を集め実験を行った結果、深層学習は高い精度で被験者のデータを分類した。そして、荷重データを特徴として抽出してきていることがわかった。計測した荷重マップを確認したところ、スキー熟練者は荷重のかかり方が移動方向に合わせて荷重がかかっているのに対し、スキー初心者や経験者に対しては、荷重が移動方向と逆側にかかっている、片側に荷重がかかりっぱなしである、荷重がかかっていない等の傾向があった。このことから、初心者への改善点として、スキーシミュレータを利用した状態での重心移動の習得が挙げられる。

今後の展望としては、本研究ではスキーの初心者や熟練者の特徴を深層学習は抽出することができたと考える。そこで、新しく提案するのがスキーの初心者や熟練者の被験者数を増やし、初心者と熟練者の滑り方の予測モデルを深層学習によって作成することである。技術的にはFuturePose[14]を応用することで実現可能なのではと考える。初心者と熟練者の滑り方の予測モデルを作成することができれば、今後のスキー指導やトレーニングに大いに役立つほか、他のスポーツにも応用が利くと考えられる。

謝辞 本研究を進めるにあたり、東京工業大学小池研究室で開発されたVRスキーの利用を快諾して下さった小池英樹先生をはじめ、小池研究室の方に深く感謝いたします。

参考文献

[1] 多田憲孝; シミュレータを用いたスキー指導システムの開発, スポーツ産業学研究, Vol.21, No.1, pp.19-26, 2011.

[2] 多田憲孝; 拡張現実感技術を用いたスキー指導システムの開発, スポーツ産業学研究, Vol.22, No.2, pp.263-270, 2012.

[3] 田中幹大, 根本啓次; VRを利用したスキートレーニングシステムの実現, テレビジョン学会技術報告, Vol.19, No.21, pp.13-18, 1995.

[4] Erwin Wu, Takayuki Nozawa, Florian Pertenede, Hideki Koike; How to VizSki: Visualizing Captured Skier Motion in a VR Ski Training Simulator, VRCAI '19, No.5, pp.1-9, 2019.

[5] 浦脇 浩二, 増田 泰, 眞鍋 佳嗣, 千原 國宏; スポーツフォーム練習支援のための生体情報可視化, 第3回情報科学技術フォーラム, pp.471-472, 2004.

[6] 本莊 直樹, 伊坂 忠夫, 満田 隆, 川村 貞夫; HMDを用いたスポーツスキルの学習方法の提案, 日本バーチャルリアリティ学会論文誌, Vol.10, No.1, pp.63-69, 2005.

[7] 園田 謙人, 岡田 滉太郎, 平井 重行; シート型圧力センサを用いた洗面台前生活行動識別とその応用, 情報処理学会インタラクティブセッション, pp. 359-363, 2019.

[8] Ossama Abdel-Hamid, Abdel-rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu; Convolutional Neural Networks for Speech Recognition, IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL.22, NO.10, pp.1533-1545, 2014.

[9] Takayuki Nozawa, Erwin Wu, Hideki Koike; VR Ski Coach: Indoor Ski Training System Visualizing Difference from Leading Skier, 2019 IEEE Conference on Virtual Reality and 3D User Interfaces, pp.1341-1342, 2019.

[10] 深堀 孔明, 坂本 大介, 五十嵐 健夫; 靴下型圧力センサを用いた足裏ジェスチャの設計と実装, WISS, pp.1-6, 2014.

[11] Joseph Redmon, Ali Farhadi; YOLOv3: An Incremental Improvement, arXiv.org, pp.1-6, 2018.

[12] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra; Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, The IEEE International Conference on Computer Vision (ICCV), pp.618-626, 2017.

[13] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie; Feature Pyramid Networks for Object Detection, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2117-2125, 2017.

[14] Erwin Wu, Hideki Koike; FuturePose - Mixed Reality Martial Arts Training Using Real-Time 3D Human Pose Forecasting With a RGB Camera, 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp.1384-1392, 2019.