

農作業支援ロボットの狭い畝間での 隠れ領域復元による作業者認識精度向上

Improving the accuracy of human recognition using hidden area recovery for an agricultural robot in a narrow furrow

○学 荒井 ルシア^{*1}, 正 萬 礼応^{*1}

正 大矢 晃久^{*1}, 正 坪内 孝司^{*1}

Lucia ARAI^{*1}, Ayanori YOROZU^{*1},
Akihisa OHYA^{*1}, and Takashi TSUBOUCHI^{*1}

^{*1} 筑波大学 University of Tsukuba

This study presents an agricultural robot that follows a worker for transportation support. We proposed a method to follow a human by obtaining the human's center position as a target point from an RGB-D camera using Deep Neural Network Model (PoseNet). But this method would be difficult to use in a place where plants hide human body. Therefore, we propose a method, that can recovery hidden area by plants in the RGB image where plants hide human body using a Generative Adversarial Network Model. In this method, it restores the hidden area of the human body by using the plants in front of the human as the shielded area. We verified using work data in the field and confirmed that the joint positions of a person can be detected on the RGB image even in a place with lush plants.

Key Words : Human recognition, Agricultural robot, Human following, Generative Adversarial Network, Image restoring

1. 緒 言

現在の日本の農業は農業従事者の高齢化，農家数の減少，食料自給率の低迷という3つの問題を抱えており⁽¹⁾，農業の小労力化が必要不可欠である。この解決策として近年 IoT，ロボット技術による農業の機械化・自動化が注目されており，労働力不足の解消の他に生産効率の向上，人件費の削減など期待されている。

大型の農機の導入が困難な日本特有の分散した中小規模の農場における農作業の支援を目的として，株式会社 DONKEY の小型多機能ロボットや，農作業管理システムによるスマート農業システムが提案されている⁽²⁾。Fig.1 に示すように，狭い畝間での収穫物の運搬支援などを行うためには，作業者を検出して，一定距離範囲を保ちながら追従する技術が必要である。本研究では，ロボットは前方に搭載した RGB-D カメラを用いて，作業者の検出を行う。

ロボットが作業者に対して追従を行うためには作業者との相対位置を取得する必要がある。そのための手法として，特定色の服を着用させて作業者の検出を行う手法⁽³⁾がある。この手法では，照明条件や遮蔽状況によっては検出が難しい場合がある。また，この手法⁽³⁾においては，検出した畝を回避しながら作業者への追従を実現しているが，Fig.1 に示すような狭い畝間では作業者に隠れて畝を検出することが難しく，特に作業者の体の中心から外れた部分を追従目標点と設定してしまった場合に，畝に乗り上げてしまう問題がある。

著者らは，特定色やマーカなどを使用せずに，作業者の関節情報を検出可能な深層学習器を用いた作業者中心部の検出手法を提案した。検出した関節情報（両腰，両膝）に基づいて，追従のための作業者の体の中心付近を追従目標点として算出する手法⁽⁴⁾を提案し，狭い畝間で畝に乗り上げずに追従できることを確認した。Fig. 2 に Fig.1 のような背の高く葉が茂るような植物による遮蔽が起こるような環境において，同手法⁽⁴⁾を適用した場合の関節位置検出結果の一例を示す。Fig.2 に示すように，腰や膝など追従目標点を算出するための関節が植物によって隠れてしまうため，追従が困難になる。



Fig.1 Human following for transportation support in narrow furrows

- Hip positions
- Knee positions
- Center position



Fig.2 Results of joint position detection ⁽⁴⁾

一方で、近年敵対的生成ネットワークを用いて、画像の指定した領域を復元する手法⁽⁵⁾が提案されている。そこで本研究では、RGB 画像に対して、作業者の体の植物によって遮蔽されている部分を敵対的生成ネットワークによる画像生成を用いて復元することで、手法⁽⁴⁾による作業者の関節情報の認識精度向上を目指す。手法⁽⁵⁾を活用して、植物に隠れた作業者の体の領域を復元するためには、作業者とロボット間にある植物を遮蔽領域として設定すること、Fig. 2 のような環境の画像に対して画像復元を行うネットワークの学習が必要である。本稿では、Depth データと一時刻前の追従目標点（ロボットから作業者中心）までの距離を用いて、奥行き方向の復元する距離の範囲を設定し、その範囲内の領域の復元を行うことで遮蔽がある状況に対応できるようなアプローチを提案する。2章で関連研究について述べ、3章で提案システムの概要を説明する。4章では、画像復元のために必要な圃場で収集したデータセットと学習について説明する。5章では、Depth データを用いた遮蔽領域の抽出方法について説明する。6章では、Fig.2 のような圃場で収集した植物による人体遮蔽が生じた場合に、画像生成を行うことによる関節の認識精度について検証する。

2. 関連研究

2・1 人検出に関する研究

橋本ら⁽⁶⁾は歩行者の移動方向予測することを目的とする、頭と肩の位置から顔の向きを推定する研究を行った。まずパーティクルフィルタを用いて基幹部を追跡し、体のどの部分なのかをある程度推測した。その後、頭部に関して、肩から頭にかけての輪郭と関節位置を点群で与え分析したモデルを用いて肩と頭の詳細な位置を推定した。両肩と頭を結んだ三角形の形から顔がどちらに向いているのかを推定可能にした。しかし、上半身か下半身の大部分が映っていないと人を検出することは困難である。

三井ら⁽⁷⁾は人の一部が隠れていても人の検出が出来ることを目的とする研究を行った。人の領域を6個のパーツに分け各パーツのラベルがついた一つのサンプルとして、マルチクラス検出器を構築し、パーツベースでの人検出方法を提案した。一部だけ遮蔽に隠れた場合でも人として検出が出来る。しかし、三井らの研究では遮蔽で隠された人の部位は一部でかつ狭い範囲であるため、Fig.2 のような状態で使用するの難しい。

Oved ら⁽⁸⁾ は人がいる場所と背景を畳み込みニューラルネットワーク(CNN : Convolutional Neural Network)を用いて判断し、さらに CNN を用いて人領域を 24 のパーツに分けることができる BodyPix を提案した。体全体が映っていないくても領域に分類することが可能であり、人間のどの部分であるかの判別が可能である。BodyPix の関節を検出する学習器として PoseNet⁽⁹⁾が利用されている。

著者ら⁽⁴⁾は、Oved ら⁽⁸⁾の PoseNet を利用して取得した画像上の作業者の関節情報のうち、左右の腰と膝の関節位置と Depth データを組み合わせることで、狭い畝間で農作業をする人の 3 次元的な中心位置を取得する手法を提案した。この手法では、背の低い植物が植えてある狭い畝間にて人の中心位置を追い続けることは可能だが、作業者の関節を隠してしまうような植物が植えてある畝間では使用が困難である。そのため、植物による遮蔽が起きるような Fig.1 のような場所でこの手法⁽⁴⁾を使うには、植物による遮蔽領域を人体領域に復元しなくてはならない。

2・2 画像復元に関する研究

Goodfellow ら⁽¹⁰⁾は、敵対的な過程を通して推定された生成モデルについてのフレームワークである敵対的生成ネットワーク (GAN: Generative Adversarial Networks) を提唱した。このフレームワークの中には 2 つの学習モデル、データ分布を得るための生成モデル Generator (以下 G) とサンプルデータが本物のデータか G が作ったデータかの確立を推定する識別器 Discriminator (以下 D) がある。この 2 つを敵対的に同時に学習させることで G は D に見破られないデータを生成し、D はデータの真偽判定の精度がより正確になる。GAN による画像生成を応用して、画像を復元するアプローチがいくつか提案されている。

Xiaosheng ら⁽¹¹⁾は GAN を応用し、画像上の人や物に遮蔽された自動車領域を復元させる手法を提案した。この手法では、自動車のセグメンテーションモデルを利用して入力画像上の遮蔽されていない自動車領域を抽出し、遮蔽部分を補完してマスクを作成、作成したマスクを GAN のモデルに組み込むことで、画像上で遮蔽された部分を含む自動車領域を生成している。Xiaosheng ら⁽¹¹⁾の手法では、復元したいものの形状がほとんど変わらないことを前提としているが、本研究で復元したい作業者は様々な姿勢を取るためこの手法は適していない。復元したいものの形状にとらわれずに復元が可能な手法が望ましい。

復元するものの形状にとらわれない画像復元の手法として Jiahui ら⁽⁶⁾の手法がある。Jiahui ら⁽⁶⁾は、動的に特徴量を選択して畳み込みを行う Gated Convolution や各特徴マップに対して、本物か偽物かの判定を行うことができる SN-PatchGAN のアーキテクチャを採用することで、画像上の指定した自由領域のみをその周囲に合わせて復元する手法を提案した。学習時にランダムに生成した自由形状領域を遮蔽部分として入力することで、生成器 G は遮蔽された領域がある場合にはその領域のみを周囲に合わせて生成することができる。しかし、本研究で使用するには作業者の体を遮蔽している植物の領域を抽出する方法が必要である。

本研究では、Fig. 2 のような状況に対して、一時刻前の作業者の距離情報と RGB-D カメラの Depth データを用いて、人体を遮蔽している植物領域部分を抽出し、その領域のみを Jiahui ら⁽⁶⁾の手法を活用して復元することで、作業者の関節情報の認識精度の向上を目指す。

3. 遮蔽領域復元による追従目標点算出の概要

Fig.4 に本研究で提案する遮蔽領域復元による作業者の中心部を検出する手法の概要を示す。Fig.4-a のように植物による遮蔽が生じている場合に、Fig.4-b ロボットと作業者間にある植物を遮蔽が起きている領域と設定し、Fig.4-c のように事前に学習した復元モデルによって復元する。その後、Fig.4-d のように復元された画像に対して、左右の腰と膝の関節位置の検出および作業者の画像上の中心位置(Center position)を決定する。Fig.4-e では、Fig.4-d で検出した関節位置と、遮蔽領域を比較して、遮蔽領域にない関節 (人体上にある関節) の Depth データを取得する。Fig.4-f では、Fig.4-e で取得した関節の Depth データの平均値を作業者の位置の Depth データとし、Fig.4-d で求めた Center position と合わせて作業者のロボット座標系での追従目標点 (Target position)とする。

本研究では、全身が写っていないくても人の関節情報を検出可能な深層学習器である PoseNet⁽⁹⁾を使用する。PoseNet を用いて左右の腰と膝の関節を検出し、その観測状況に応じて、ロボット座標系での作業者の中心付近の位置を追従目標点として決定する⁽⁴⁾。一時刻前に追従目標点すなわち、作業者までの距離情報が取得できた場

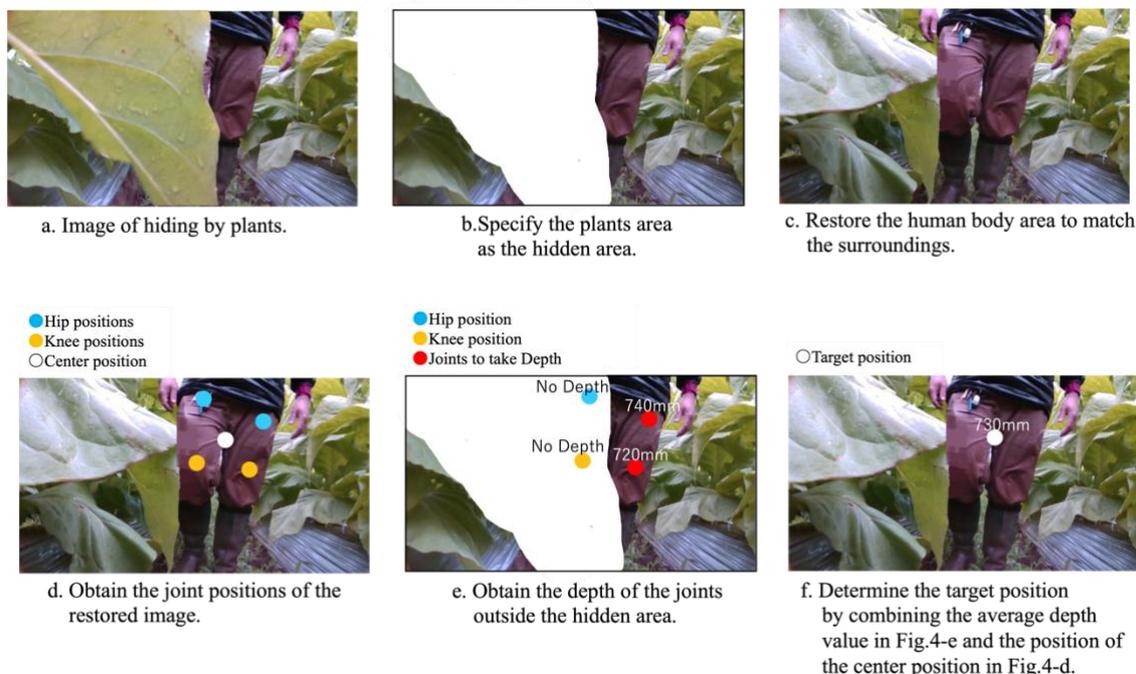


Fig.4 Overview of the method for calculating the target position when plants shielding

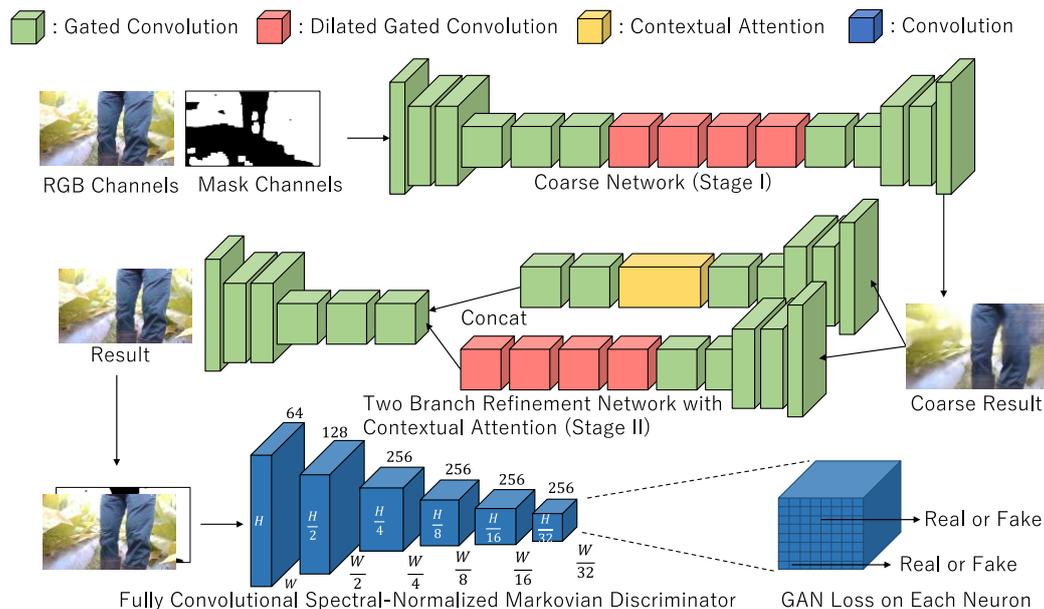


Fig.5 Image of the network configuration

合に、遮蔽領域の抽出および復元処理を行う。ロボットと作業者の間にある植物を復元する遮蔽領域として、抽出できることが望ましい。そこで、一時刻前の追従目標点（ロボットから作業者中心）までの距離を用いて、奥行き方向の復元する距離の範囲を設定し、現在の Depth データの値がその範囲にある領域を復元する遮蔽領域として抽出する。復元する距離の範囲設定については6章で述べる。遮蔽領域の取得後は、RGB 画像と遮蔽領域から次章で述べる復元モデルを通して植物遮蔽領域を復元し、追従目標点算出までの処理を行う。



Fig.6 Example of a Mask area image.

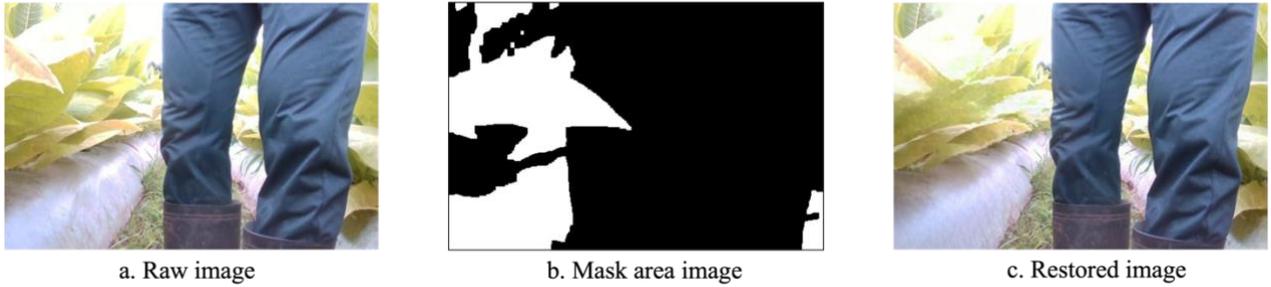


Fig.7 Restoration results using the trained image.

本稿では、作業者とロボット間にある植物を遮蔽領域の抽出と、圃場のロボット走行時のデータを基に学習した復元モデルを用いて画像を復元することで、植物による遮蔽が生じた場合に、追従目標点の算出に使用する両腰、両膝の関節検出精度について検証する。

4. 植物による遮蔽領域復元モデルの学習

本研究で用いる Jiahui ら⁽⁶⁾の手法のネットワークの構造を Fig.5 に示す. それぞれ緑の部分は動的に特徴量を選択して畳み込みを行い、赤の部分は穴が空いているフィルターを用いて畳み込みを行う. 黄の部分は不完全な画像に対して、バックグラウンドのどこに注目をして画像の補完をするかを決め、青の部分では入力されたものの真偽を判別するための畳み込みを行うことを示している.

生成器Gの目的関数を式(1)に、識別器 D_{sn} の目的関数を式(2)に示す. ここで、 D_{sn} はスペクトル正規化された識別器、Gは不完全な画像 z を受け取る画像インペインティングネットワークである. $E_{z-P_{data}(z)}[LeRU(1 - D^{sn}(x))]$ は本物のデータ集合 $data$ からランダムに選択された x を入力とした識別器 D_{sn} の結果のスコア、 $E_{z-P_z(z)}[D^{sn}(G(z))]$, $E_{z-P_z(z)}[LeRU(1 - D^{sn}(G(x)))]$ は生成器 G によって生成された画像を入力された識別器 D_{sn} の結果のスコアである.

$$L_G = -E_{z-P_z(z)}[D^{sn}(G(z))] \quad (1)$$

$$L_{D^{sn}} = E_{z-P_{data}(z)}[LeRU(1 - D^{sn}(x))] + E_{z-P_z(z)}[LeRU(1 - D^{sn}(G(x)))] \quad (2)$$

Fig.5 より、モデル学習の際には、植物による遮蔽が起きていない状態の作業者が写っている RGB 画像と遮蔽領域画像が必要である. ロボットによる作業支援を実施した際に取得した RGB-D カメラのデータから、約 9000 枚の植物による遮蔽がない農作業中の作業者の RGB 画像を用意した. 学習を行う際に RGB 画像とともに入力する遮蔽領域画像 (以下 Mask 画像とする) は、なるべく実際の環境と似ているものが望ましい. そこで Fig.2 のよ

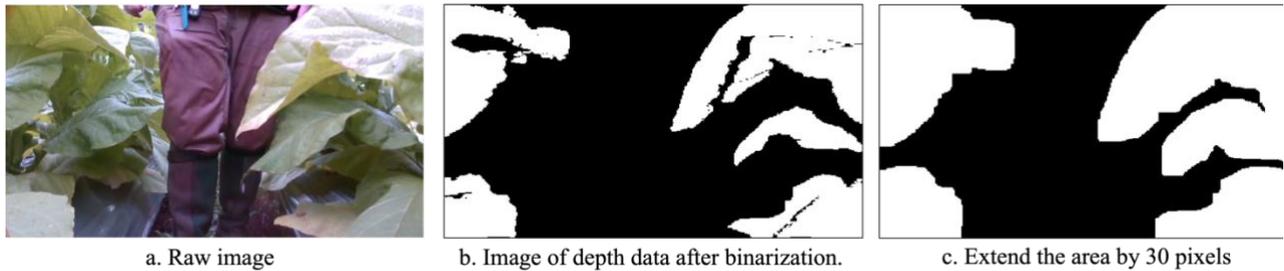


Fig.8 Example of binarization and dilation

うな、実際に圃場でロボットが作業者を追従している際に取得した RGB 画像を基に、植物の色領域を抽出して、210 パターンの Mask 画像を作成した。Fig.6 に Mask 画像の一例を示す。白の領域が復元したい領域である。学習時には、RGB 画像に対して、210 パターンの Mask 画像をランダムに組み合わせて入力させた。約 9000 枚の RGB 画像に対して、学習ループは 500 回とした。学習データを用いた復元結果の一例を Fig.7 に示す。Fig.7-a の画像と Fig.7-b の遮蔽領域を入力すると、Fig.7-c のように復元される。

5. 植物による遮蔽領域の抽出

植物による人体遮蔽が起きている時に、復元する領域は必要な部分のみであることが望ましい。本研究では、人体が隠れている作業者よりも手前の植物領域のみを復元する。方針としては、一時刻前の追従目標点(作業者の中心位置)から、作業者とロボットの移動を考慮した値を引いたものを取得範囲の上限として、その値よりもロボット側にある Depth データを遮蔽領域とすることにする。

Fig.8 に遮蔽領域抽出の一例を示す。提案システムでは、遮蔽領域を取得する範囲の上限を一時刻前の作業者までの距離- α として、Depth データが範囲内に含まれる場合に復元する領域を白として二値化した画像を生成する。Fig.8-b に Fig.8-a の状況において、設定した範囲に対して、Depth データから生成した二値化画像の例を示す。二値化後の領域の境界付近は植物領域の境界付近である可能性が高く、このままでは境界部分に微小な植物領域が残り、残った領域から植物領域が復元される場合がある。そこで、Fig.8-c に示すように、Fig.8-b の白領域を膨張させ、これを遮蔽領域とした。

6. 遮蔽領域復元による作業者認識の検証

6.1 概要

Fig.1 に示すように、植物が茂った狭い畝間での不要な葉を取り除く作業時に手法⁽⁴⁾で追従を行うロボットから取得した RGB-D データに対して、4 章で学習した画像復元モデルを用いて、画像復元とそれにより関節位置の認識精度が向上するのかどうかの検証を行った。また、遮蔽領域を抽出する際に、一時刻前に検出した作業者までの距離に対して、 α の値を 0~250mm まで 50mm 区切りにした場合を比較し、領域を復元する距離の範囲の検証を行った。

評価には一時刻前に追従目標点が取得できた 195 フレームを使用した。本検証での認識成功 (○) は、左右の腰と膝の関節位置が画像上の人体領域内に検出できており、かつ、片方の脚に両方分の腰や膝の関節が検出されていないものとし、目視による判定を行った。

6.2 結果・考察

Table.1 に α の値による関節位置認識結果を示す。復元をしなかった場合の認識成功フレーム数は 195 フレーム中 117 フレーム(60%)であった。これに対し、復元を行なった際の認識成功フレーム数が最大のものは $\alpha=200$ の場合であり、そのフレーム数は 147 フレーム(75%)であった。 $\alpha=200$ では、復元しない場合には失敗したが、復元により認識成功した($\times \rightarrow \circ$)ものは 39 フレーム増加した。復元をしない場合では認識に成功したが、復元をし

Table.1 Consideration of α to be the upper limit of the range (total number of frames: 195)

α	No restoration \rightarrow With restoration				Successful recognition frames
	$\bigcirc \rightarrow \bigcirc$	$\bigcirc \rightarrow \times$	$\times \rightarrow \bigcirc$	$\times \rightarrow \times$	
0 mm	74	43	19	59	93 (48%)
50 mm	84	33	28	50	112 (57%)
100 mm	96	21	44	34	140 (71%)
150 mm	100	17	42	36	142 (73%)
200 mm	108	9	39	39	147 (75%)
250 mm	105	12	34	44	139 (71%)

- Hip positions
- Knee positions

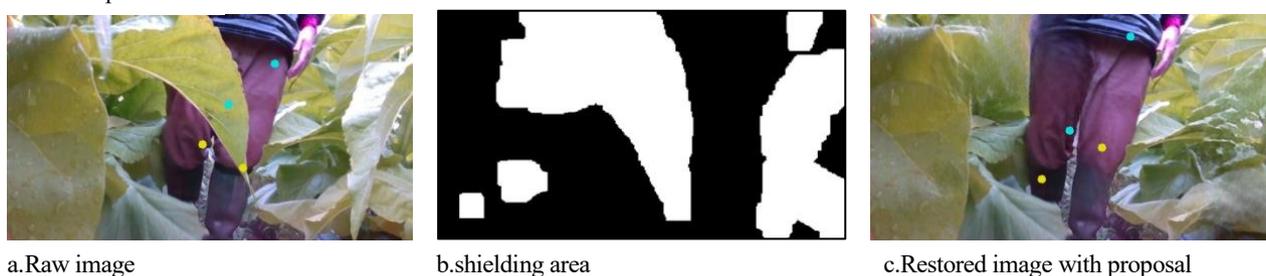


Fig.9 Example of joint detection results: $\times \rightarrow \bigcirc$ case

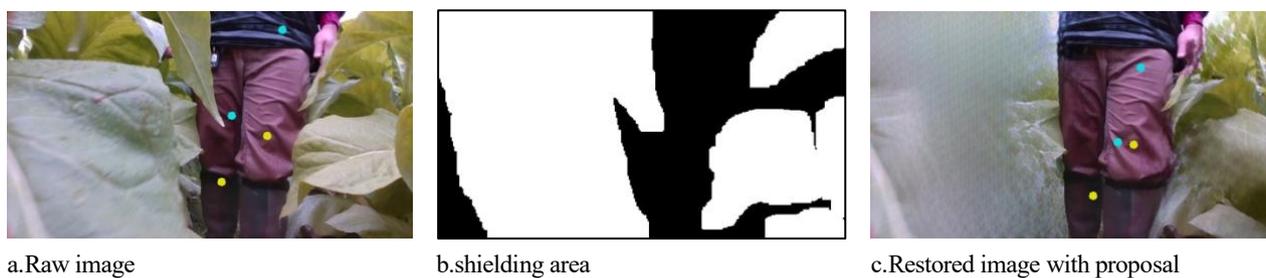


Fig.10 Example of joint detection results: $\bigcirc \rightarrow \times$ case

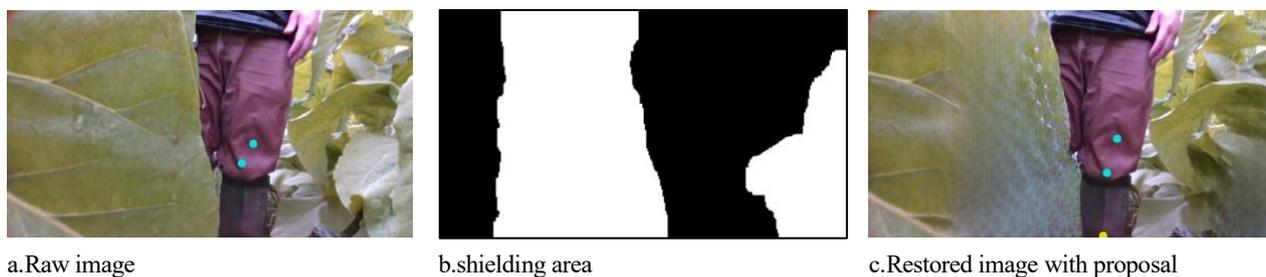


Fig.11 Example of joint detection results: $\times \rightarrow \times$ case

た場合に認識失敗となった($\bigcirc \rightarrow \times$)ものは9フレーム存在した。また、復元をしてもしなくても認識に失敗したものは39フレーム存在した。

$\alpha = 200$ の時のそれぞれの場合の結果画像を Fig.9 から Fig.11 に示す。Fig.9-a から Fig.11-a は復元していない画像に対して関節検出を行なった結果を、Fig.9-b から Fig.11-b は遮蔽領域を、Fig.9-c から Fig.11-c は復元後の画像に対して関節検出を行なった結果を示す。

Fig.9-a では関節位置が植物常に出てしまっているが、Fig.9-b のような人の手前にある植物があった領域を復元することで、Fig.9-c のような画像になり、関節位置が作業者の人体上に出るようになった。Fig.9-a から Fig.9-c を比べると、植物による人体への遮蔽が起きている領域を遮蔽領域として抽出することができている。一部遮蔽領域に含まれていない植物もあるが、関節認識に大きな影響を与えるほどではなく、復元された画像の関節は深層学習機によりある程度検出可能であると考えられる。

Fig.10-a では認識が成功しており、Fig.10-b の修復領域も人の手前にある植物をほとんど全て含むことができる。しかし、Fig.10-c では復元した領域を含む右大腿部に関節が出ておらず、左足に 2 箇所腰の関節位置が描写された。このように、復元されていない領域に左右どちらの関節も出ることや、復元した領域に人体関節があると認識されない場合があり、このような場合にはロボットが畝に乗り上げてしまう可能性があると考えられる。

また、Fig.11-a のように左右の半身しか写っていない場合と、Fig.11-b のような人の手前にある大きな植物領域のうちの一部が遮蔽領域に設定できていない場合が組み合わさると、Fig.11-c のように修復した領域が全て植物領域となり復元する前の画像とほとんど変わらない結果となることがあることがわかった。

Fig.10-c と Fig.11-c のように、復元をしても認識に成功しない場合もあるが、75%と大半のフレーム数で Fig.9-c のように認識に成功することから、本研究の手法は有効だと考えられる。

7. 結 言

本研究では、植物による遮蔽があるような環境で関節位置から人の中心部を検出する手法⁽⁴⁾の認識精度を向上させることを目的に、敵対的生成ネットワークを利用した植物による遮蔽領域を復元するアプローチを提案した。Depth データと一時刻前の追従目標点（ロボットから作業者中心）までの距離を用いて、奥行き方向の復元する距離の範囲を設定し、復元する遮蔽領域を抽出する方法を提案した。また、ロボットによる作業支援を実施した際に取得した RGB-D カメラから学習データセットを構築し、圃場での遮蔽領域復元モデルを学習した。

植物が茂った狭い畝間で農作業を行う作業者に追従した際に、ロボットで取得した RGB-D データに対して、提案手法を適用し、遮蔽領域抽出の際の復元する奥行き方向の距離の適切な範囲を検証した。提案手法により、植物による遮蔽領域を復元することで、両腰、両膝の関節位置の認識精度が大きく向上することを確認した。

謝 辞

本研究は、株式会社 DONKEY との共同研究であり、ロボットの利用および実験環境の提供に協力いただいたことを記し、謝意を表す。

文 献

- (1) 農林水産省: “農林水産基本データ集”, <https://www.maff.go.jp/j/tokei/sihyo/>, (2019).
- (2) 三輪泰史, 井熊均, 木通秀樹: アグリカルチャー4.0 の時代農村 DX 革命, (2019), 日刊工業新聞社.
- (3) 萬礼応, 石上玄也, 高橋正樹: 農作業支援ロボットの畝間人追従走行, (2019), 第 62 回自動制御連合講演会.
- (4) 荒井ルシア, 萬礼応, 大矢晃久, 坪内孝司. 農作業支援ロボットの狭い畝間での追従のための作業者認識—RGB-D カメラと PoseNet を用いた人の重心部の認識—, ロボティクス・メカトロニクス講演会 2021 (ROBOMECH2021), (2021).
- (5) Yu Jiahui, Lin Zhe, Yang Jimei, et al. Free-form image inpainting with gated convolution. Proceedings of the IEEE International Conference on Computer Vision, (2019), pp. 4470-4479.
- (6) 橋本潔, 加賀屋智之, 片岡裕雄 et al. 姿勢変動を考慮した基幹部位追跡と肩位置推定-単眼カメラからの 3 次元姿勢推定に向けて-. 画像の認識・理解シンポジウム, (2011), pp. 618-674.
- (7) 三井相和, 藤吉弘亘. Randomized Trees を用いたパーツベースによる人検出法. 動的画像処理実利用ワークショップ (DIA), Vo.12 (2011), pp. 1-6.

- (8) Dan Oved . Introducing BodyPix: Real-time Person Segmentation in the Browser with TensorFlow.js . <https://medium.com/tensorflow/introducingbodypix-real-time-person-segmentation-in-the-browser-with-tensorflowjs-f1948126c2a0>, (2019), アクセス:2020
- (9) Dan Oved. Real-time Human Pose Estimation in the Browser with TensorFlow.js. <https://medium.com/tensorflow/real-time-human-poseestimation-in-the-browser-with-tensorflow-js-7dd0bc881cd5>, (2018), アクセス:2020
- (10) Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza. Generative Adversarial Networks. arXiv:1406.2661, 2014, アクセス 2021.
- (11) Yan Xiaosheng, Wang Feigege, Liu Wenxi, et al. Visualizing the invisible:Occluded vehicle segmentation and recovery. Proceedings of the IEEE International Conference on Computer Vision, (2019), pp. 7617-7626.