

修士論文

ロボットによる人工授粉のための 部分的に隠れた花の検出

佐藤 諒

2019 年 3 月 15 日

奈良先端科学技術大学院大学
情報科学研究科

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

佐藤 諒

審査委員：

小笠原 司 教授	(主指導教員)
向川 康博 教授	(副指導教員)
高松 淳 准教授	(副指導教員)
丁 明 助教	(副指導教員)
GARCIA RICARDEZ Gustavo Alfonso 助教	(副指導教員)

ロボットによる人工授粉のための 部分的に隠れた花の検出*

佐藤 諒

内容梗概

ロボットで人工授粉を行う場合，ある視点から観測された画像において部分的に隠れた花を検出することは，視点を移動したり，葉を払いのけたりするなど，ロボットの行動戦略の決定において重要である．部分的に隠れた花を検出する方法として，画像の各ピクセルを枝，葉，見えている花，隠れている花，背景の5クラスのいずれかに属するかを分類する領域分割に着目する．深層学習による画像の領域分割は認識精度の高さから注目されているが，大量の学習データが必要となる．しかし，現実環境において部分的に隠れている花の学習データを得ることは困難であるため，3DCGでレンダリングされた画像を学習データとすることで学習のために十分なデータ量を確保する手法を提案する．

本研究では，トマトの育成環境を模したRGB画像とそれらを5クラスに分類した画像のセットを用いて学習することで，RGB画像を領域分割するネットワークを構築し部分的に隠れた花を推定する．手法の有効性を示すためにトマトのCG画像および写真から隠れた花を検出した．現実の花の認識ができるネットワークを学習するためには，レンダリングの際に作物のテクスチャと背景・光源が重要な要素であることを確認した．

最後に，3DCGの画像のみで学習したネットワークを用いて，人工的な環境下において隠れたトマトの花の位置を推定し，ロボットの視点を移動する授粉実験を行なった．実験により現実環境において，3DCGの画像のみで学習したネットワークがロボットを用いた人工授粉に有効であることを示した．

*奈良先端科学技術大学院大学 情報科学研究科 修士論文, 2019年3月15日.

キーワード

自動化, 農業, ロボット授粉, セマンティックセグメンテーション, データ拡張

Detection of Partially Occluded Flowers for Robotic Pollination*

Ryo Sato

Abstract

In robotic pollination, detection of partially occluded flowers from a certain viewpoint is very important to create robot strategies such as changing viewpoint and removing leaves that occlude the flowers. In order to detect partially occluded flowers, we focus on region division which classifies each pixel of an image into five classes, namely branches, leaves, visible flowers, occluded flowers, and background. Recently, image recognition based on deep learning has drawn attention due to its high performance. Nevertheless, deep learning requires a large amount of training data, which is very difficult to prepare, especially in the case of real images. Therefore, we propose to use images generated using 3D computer graphics as the training data to overcome this difficulty. In this research, we learn pairs of 1) RGB images from scenes of tomato cultivating environments and 2) images classified into five classes. Then, we construct a neural network that segments RGB images and estimates partially occluded flowers. We showed that the texture of the plant, realistic background and illumination settings during image rendering are essential to successfully train an image recognition algorithm that can also be applied to real scenes. Finally, we estimated the position of hidden artificial tomato flowers using the trained network and created a demonstration of artificial pollination after finding the hidden flowers by changing the viewpoint

*Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, March 15, 2019.

of the robot, as a proof of concept. The experiments showed that the network trained only with images generated from 3D computer graphics is effective for artificial pollination using a robot in real environments.

Keywords:

Automation, Agriculture, Robotic Pollination, Semantic Segmentation, Data Augmentation

目 次

1. はじめに	1
1.1 研究背景	1
1.2 関連研究	3
1.2.1 トマトの授粉を行うロボット	3
1.2.2 収穫を行うロボットのビジョンユニット	3
1.2.3 3DCG 画像からの真値データ抽出	4
1.2.4 植物の隠れた部分の推定	4
1.3 研究目的	5
1.4 本論文の構成	6
2. 隠れた花を検出する手法	7
2.1 提案手法の概要	7
2.2 データセットの作成とネットワーク	9
2.2.1 トマトの 3DCG モデルの作成	9
2.2.2 レンダリング	12
2.2.3 DeepLabv3+	14
3. 隠れた花の検出実験	15
3.1 実行環境	15
3.2 CG 画像による隠れた花の検出性能評価	16
3.3 現実画像による隠れた花の検出性能評価	25
4. ロボットを用いた現実環境での応用	33
4.1 実験環境	33
4.2 現実の造花画像による評価	39
4.3 考察	41
5. おわりに	42
5.1 まとめ	42

5.2 今後の課題と展望	42
謝辞	43
参考文献	44

目 次

1	Target cutting point in tomato harvesting [6]	2
2	High-wire tomato	6
3	System flow	8
4	Rendered image for training	9
5	Ground truth image for training	9
6	Three views of 3D model of tomato flower and leaf	10
7	Plot of 3D model tomato	11
8	Picture of tomato leaf for texture	11
9	Picture of tomato flower for texture	12
10	Case divided and rendered images for training	13
11	Network structure of DeepLabv3+ [21]	14
12	Input image	17
13	Ground truth image	17
14	Comparison between rendered images	18
15	Input image - 2	19
16	Ground truth image - 2	19
17	Comparison between rendered images - 2	20
18	Input image - 3	21
19	Ground truth image - 3	21
20	Comparison between rendered images - 3	22
21	Hidden flowers by branches	25
22	Comparison of evaluations for hidden flowers by branches	26
23	Real tomato flower	27
24	Synthetic image	27
25	Comparison of evaluations for hidden real flowers	28
26	Real tomato flower - 2	29
27	Failure of flower detection	29
28	Manually annotated images of real tomato flowers	30

29	Artificial flower of tomato - front	34
30	Artificial flower of tomato - side	34
31	Artificial flower cluster of tomato	35
32	End effector and camera	35
33	Artificial plant	36
34	KUKA LBR iiwa 14 R820	37
35	Experiment for pollination of an artificial flower by the robot. . .	38
36	Occluded flowers by branches in artificial plants	39
37	Detection of artificial occluded flowers	39
38	Robot pollination with artificial flowers	40

表 目 次

1	Work time per task in tomato production (10a, Annual production average, 2007, Japan) [1]	2
2	Weight per class	15
3	Training parameters for DeepLabv3+	16
4	F1 score with rendered images in each class and condition	24
5	Recall with rendered images in each class and condition	24
6	Precision with rendered images in each class and condition	24
7	Accuracy with rendered images in each class and condition	24
8	F1 score with real images in each class and condition	32
9	Recall with real images in each class and condition	32
10	Precision with real images in each class and condition	32
11	Accuracy with real images in each class and condition	32
12	Basic specification of the robotic arm	38

1. はじめに

1.1 研究背景

世界人口の増加に伴う食糧の不足に対し、農作物の生産性の向上が求められている。日本国内トマト生産従事者における作業別労働時間の政府統計 [1] (Table 1) によると、授粉を含む管理と収穫の占める時間が長く、農作物の生産量を増加させるために、授粉 [2] [3] [4]、収穫 [5] [6] を行うロボットの開発が進められている。

枝や葉が複雑に重なった作物は収穫を行う際ロボットのみならず人によって果実の収穫をする際にも負担となっている。収穫を簡易にするためには複数の花の集まりである花房の全ての果実が同時に収穫できることが望まれる (Fig. 1) が、収穫時期は受粉時期に大きく依存するため、同一時期に授粉できることが望ましい。

一方受粉において蜂を利用した省力化が行われているが、米国において 2006 年から 2007 年にかけて蜂群崩壊現象とよばれる蜂の大量失踪により受粉用蜂の大規模な供給が不足が発生したことがあり、気候変動や農薬の散布、ウイルスによる再発が懸念される [7]。年間 240 億ドル相当の作物が天然の花粉媒介者に依存しており [8]、蜂の数が減少することでそれらを借りなければならない農家の費用が増加することになる。

受粉用蜂を使用する欠点として、蜂による受粉では気温や成長段階、花の形状といった商業的に適した条件の花に受粉する保証がなく、一つの花房中の各花に対する受粉時期が異なるため、果実の熟する時期に差が生じる点が挙げられる。

そのため、花房中の全ての花に対して一斉に授粉のできるロボットが求められている。

Table 1 Work time per task in tomato production (10a, Annual production average, 2007, Japan) [1]

	Working hours by work	Ratio
Raising seedlings	46.82	5%
Cultivate / Basal dressing	18.88	2%
Seeding / Seedlings	43.1	5%
Topdressing	9.51	1%
Weeding / Control of insect pests	35.31	4%
Irrigation and drainage / Air conditioning	37.19	4%
Management	325.57	34%
Harvest	308.49	33%
Processing	33.78	4%
Shipment	74.42	8%
Other	13.66	1%
Total	946.73	100%



Fig. 1 Target cutting point in tomato harvesting [6]

1.2 関連研究

はじめに授粉，収穫を行うロボットの先行研究について挙げる．本提案では花が周囲の枝や葉に隠れることの多い環境においても授粉，収穫を可能にするため，セマンティックセグメンテーションを用いる．深層学習によるセマンティックセグメンテーションは識別成功率の高さと処理速度から注目されているが，そのためには大量のアノテーションを行なった学習データが必要となる．しかし，植物の複雑な形状かつ隠れた部分のアノテーションは困難であるため，3DCGを用いたアノテーションデータ作成に関する研究を挙げる．最後に隠れた植物の形状推定についての研究について挙げる．

1.2.1 トマトの授粉を行うロボット

トマトは蜂に代表される花粉媒介者による接触や，風によって受粉する他に，ホルモン剤の噴霧によって授粉することができる．ホルモン剤の噴霧は蜂や風による受粉よりも結実率が高いとされるが，人による授粉が必要なためコストが高くなる，複数回噴霧してしまうと薬害が生じ出荷できないといった問題が生じることがある．先行研究では蜂が花に着地する際の接触を模倣し，リニアアクチュエータによって綿棒状のブラシを前後に動作させるエンドエフェクタを持つ形式 [2] と，ホルモン剤を人と同様に花に噴霧する形式が存在する [3] [4]．

1.2.2 収穫を行うロボットのビジョンユニット

日本におけるトマトの収穫ロボットの研究はKawamuraら [9] の1984年から行われており，リンゴ，オレンジ，モモ，メロン，スイカ，トマト，ブドウ，キュウリ，イチゴまたはキノコなどの様々な種類の作物について，様々な収穫ロボットシステムが提案されている [10]．一般に果実の収穫ロボットはビジョンユニットによって農作物の規則通りでない構造条件下で果実を検出し収穫の判断を行う．しかし農業において果物の情報を正確に入手することは困難であるので，ビジョンユニットが農業ロボットの商業的応用のボトルネックであることが証明されている [11]．ビジョンユニットの方式として受動的方法と能動的方法に分類される．

受動的方法は画像を取得し分析するためにカメラのみ使用する方法である．受動的方法としてステレオカメラが知られているが特徴点マッチング問題が欠点として挙げられる．密集し隠れる果実に対して，この問題は位置の正確さを減少させそして収穫の失敗を引き起こす原因となる．能動的方法是，撮影において赤外線やレーザーの光を投射する装置および構造化照明を必要とし，主に受動的方式と組み合わせて用いられる [11] [12]．

1.2.3 3DCG 画像からの真値データ抽出

Stephan らは写実的なオープンワールドのコンピュータゲームによりレンダリングされた CG 画像からピクセル単位のセマンティックデータの生成を試みている．ゲームデータによるセマンティックデータとアノテーションされた実写データである CamVid [13] トレーニングセットの 1/3 で学習したモデルは，完全な CamVid トレーニングセットで学習したモデルよりも優れたセグメンテーション性能を示した [14]．また，Ros ら [15] の SYNTHIA データセットは仮想の街に対して，複数の視点から異なる季節，天候および照明条件をシミュレートしてピクセル単位でのセマンティックデータを含んだ自動運転のためのデータセットを生成している．SYNTHIA データセット単体および既存の現実環境のデータセットと併用し学習によって現実の環境に対するセマンティックセグメンテーション実験を行なっている．以上のセマンティックセグメンテーションデータセットに関する実験は，取得した仮想環境のデータを使用して実世界のセグメンテーションが可能であり，それらのデータによって実世界の画像を補完することで精度を大幅に向上できることを示唆している．

1.2.4 植物の隠れた部分の推定

植物の隠れた部分を推定する試みは Isokane ら [16] によって行われている．Isokane らは複雑な植物の形状を推定するために，多視点画像から見えない部分も含めた 3D モデルを GAN による生成モデルである pix2pix [17] を用いた画像変換によって再現を試みている．また，Isokane らの研究では植物形状の真値を得

るために 3DCG によるモデルを使用している。しかし，多視点からの画像を必要とする手法であるため，撮影角度が限定される作物が密集し並んだ農場における作物全体モデルの推定は困難である。

1.3 研究目的

本研究では，CG 画像を用いて現実環境の認識に利用できる深層学習に有効な学習データを得るために，学習データとして利用可能な 3DCG の条件について検討し，CG 画像と現実画像それぞれで隠れた花の検出の評価を行う。また，手法の応用として現実環境においてロボットによる人工授粉のための部分的に隠れた造花の検出デモを行う。

本論文では授粉対象として，日本で最も出荷量の多い果実であるトマトに着目する。日本のトマトの生産現場において，収穫量が多いオランダ式ハイワイヤー栽培 (Fig. 2) や低段密集栽培が大規模施設を中心に導入されている。ハイワイヤー栽培は長期間の栽培を行い誘引高さは 3 [m] 以上，主茎の総延長は 5 [m] 以上 [18] になることもあり，収穫花房は 30 段にいたることもある。低段密集栽培は苗の本数を増やすことで短期間のうちに集中して栽培する手法であり高品質かつ単位面積・時間当たりの収量が高くなる [19]。しかしながら，これらの作物を密集させた栽培法は，花房，果房の方向および高さが変動することや主茎が花房，果房に対し複雑に絡まることからロボットでの授粉，収穫が困難になると高市ら [20] は指摘している。Yuan ら [3] は花房の花が少ない場合や周囲の枝葉によって花が隠れている場合，花の認識率が低くなり授粉の成功率が低下すると述べている。

したがって周囲の枝や葉によって隠れた花を見つけることができれば，既に普及が進み収穫量の多い栽培方法においてロボットによる斉一授粉を効果的にできると考えられる。



Fig. 2 High-wire tomato

1.4 本論文の構成

以下、2節でRGB画像から隠れた花を検出する手法について述べ、3節で提案手法の有効性を示すための実験方法と実験結果および考察を述べる。5節で結びとする。

2. 隠れた花を検出する手法

2.1 提案手法の概要

本研究ではロボットアームの先端に取り付けたカメラから得た RGB 画像中の葉や枝の裏に隠れる花を検出し、視点移動および葉や枝の操作によって隠れた花の発見について提案する (Fig. 3)。花の検出にセマンティックセグメンテーションを用いるが、これは葉や枝の裏に隠れる花を検出するためだけでなく、枝や葉を傷つけずに操作するためである。セマンティックセグメンテーションのネットワークとして DeepLabv3+ [21] を用いる。トマトの育成環境を模した RGB 画像 (Fig. 4) とそれらを枝、葉、見えている花、隠れている花、背景の 5 クラスに分類する。5 つのクラスに対してピクセル単位のアノテーション処理をした画像 (Fig. 5) のセットを用いて学習することで、RGB 画像が与えられた時にセマンティックマップを出力するネットワークを構築し部分的に隠れた花を推定する。現実環境において、枝や葉によって隠れている花についての情報を得ることは困難であるため、本研究では、植物の 3DCG を用いてレンダリングした RGB 画像と、5 クラスにピクセル単位でアノテーション処理をした画像のみを使用して、学習用データセットを作成する。アノテーション画像の色は PASCAL VOC 2012 [22] のインデックス値に基づき背景は 0 (黒色)、葉は 1 (赤色)、枝は 2 (緑色)、見えている花は 3 (黄色)、隠れた花は 4 (青色) とする。

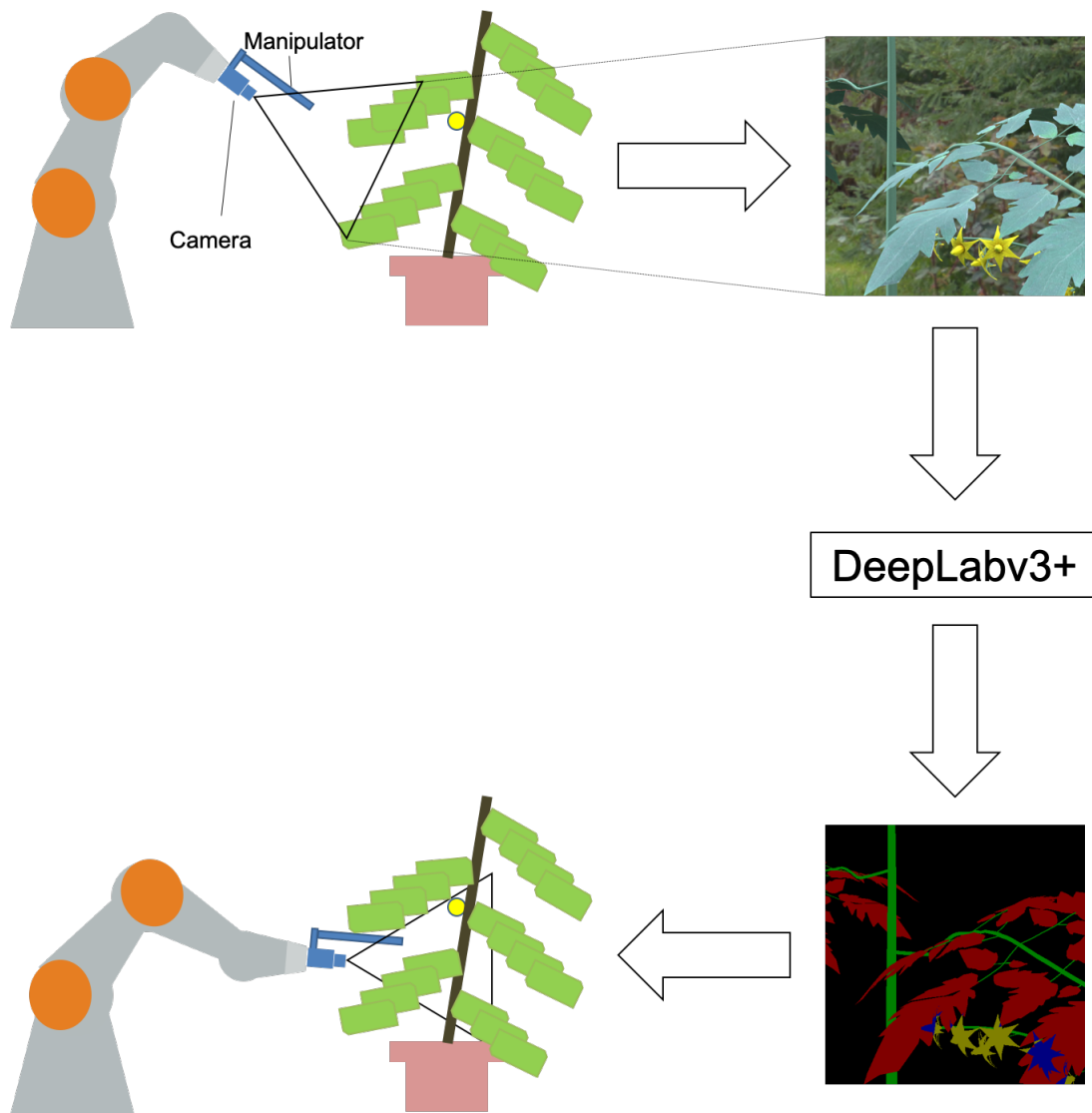


Fig. 3 System flow

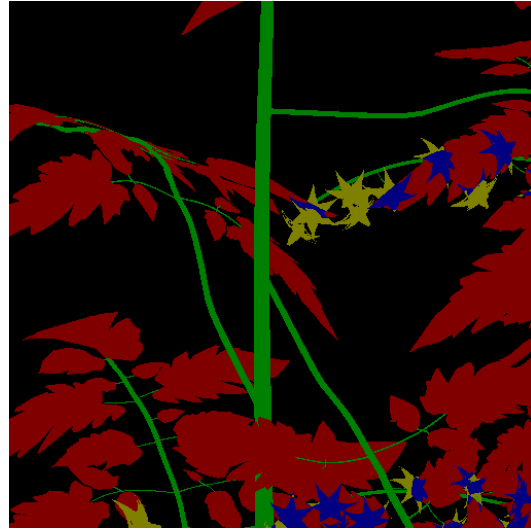


Fig. 4 Rendered image for training Fig. 5 Ground truth image for training

2.2 データセットの作成とネットワーク

2.2.1 トマトの 3DCG モデルの作成

トマトの着花規則はトマトの種類，栽培方法によって異なるが，今回は3葉ごとに花房を付け，主茎に対して一方向にのみ花がつく枝が生じるといった特徴を持つとされる非心止まり系の品種 [23] を参考に 3DCG モデルを作成する (Fig. 6)．実際の栽培では，成長に従い茎のねじれや，花のつく枝の生じる方向がずれる段跳びとよばれる現象のため，花が近接した枝や葉に隠れることがある．



Fig. 6 Three views of 3D model of tomato flower and leaf

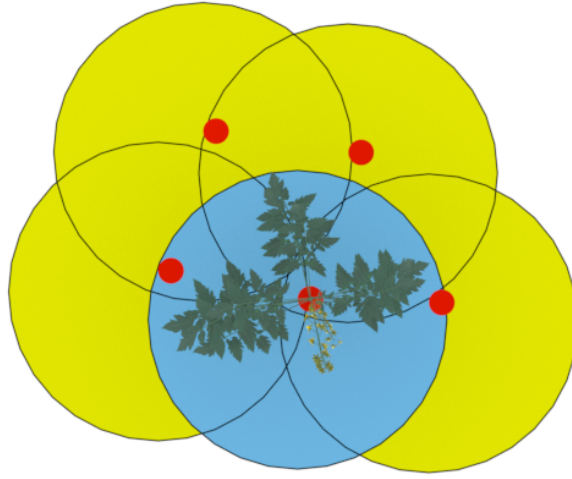


Fig. 7 Plot of 3D model tomato

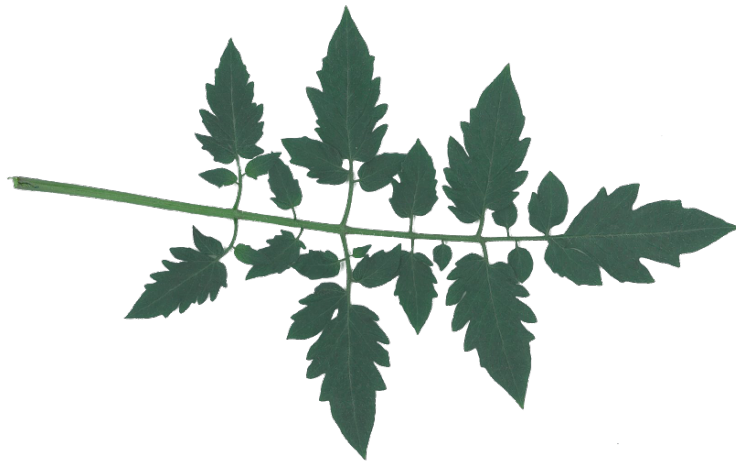


Fig. 8 Picture of tomato leaf for texture



Fig. 9 Picture of tomato flower for texture

本実験はモデルの作成を簡単にするため、段跳びや大きくねじれの生じていない状態のモデルを作成し、5つそれぞれのモデルの葉が互いに干渉する配置にすることで、ハイワイヤー栽培における隣のトマトの枝や葉が、互いに重なり合うことによって生じる花のオクルージョンが発生しやすい状況に近づけた。Fig. 7の赤点が各トマトの主茎座標であり、円が枝や葉が重なり合う範囲である。また、トマトの葉や花のテクスチャを、トマトの葉 (Fig. 8) とトマトの花の写真 (Fig. 9) それぞれ1枚から作成する。ただし Fig. 9 はモデル作成上の都合から上部を加工している。

2.2.2 レンダリング

作成したトマトのモデルを、撮影地点、撮影方向の2つの要素についてモデルを中心とした範囲内においてランダムに変更し、画像をレンダリングする。画像サイズは縦512ピクセル、横512ピクセルとする。また背景を切り替えるため、背景画像として全天球画像を17種類用意する。レンダリングにおいて全天球画像が



(A) Texture and background

(B) No texture and background



(C) Texture and no background

(D) No texture and no background

Fig. 10 Case divided and rendered images for training

ら光源情報を計算する IBL (Image Based Lighting) を用いるため光源は全天球画像 (背景光源) に依存する。ランダムにレンダリングした画像内において、トマト作物が写っていない画像を除くために、トマト作物全体の画面占有率を 0.015% 以上と限定する。

CG の質と認識精度の関係を評価するために、(A) テクスチャ有しかつ背景光源有り、(B) テクスチャ無しかつ背景光源有り、(C) テクスチャ有しかつ背景光源無し、(D) テクスチャ無しかつ背景光源無しの 4 つの条件に分け、A, B, C, D はそれぞれ同一の撮影地点、撮影方向とし、A, B の背景光源は同一としてレンダリングする (Fig. 10)。

2.2.3 DeepLabv3+

本研究では DeepLabv3+ を用いて隠れた花の位置の推定を行う。DeepLabv3+ は、後処理なしで PASCAL VOC 2012 と Cityscapes [24] において最高精度を達成したセマンティックセグメンテーション手法であり、Fig. ?? のネットワークによって構成されている。

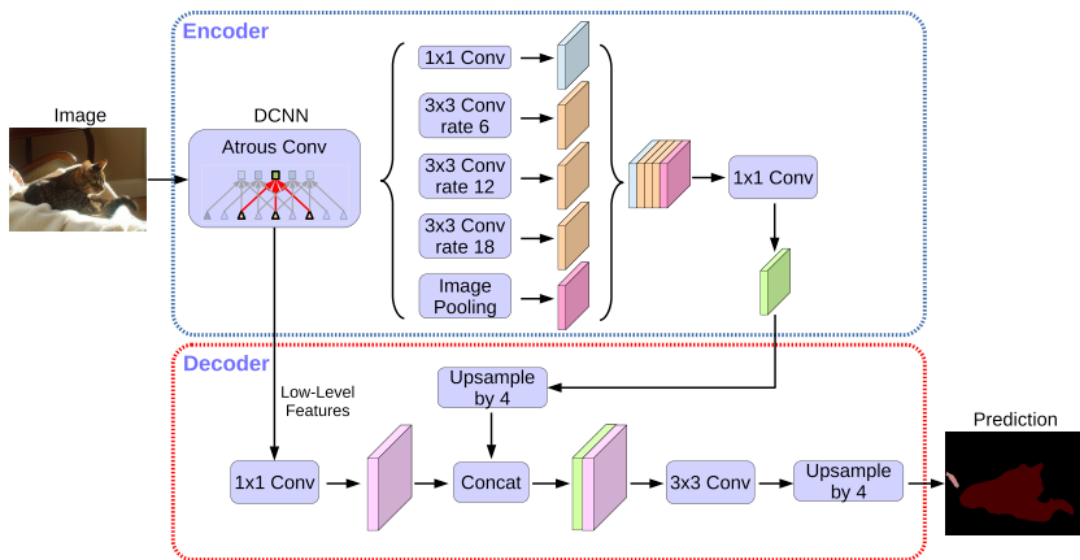


Fig. 11 Network structure of DeepLabv3+ [21]

3. 隠れた花の検出実験

3.1 実行環境

本実験では DeepLabv3+ を 3DCG でレンダリングされた画像を用いて学習し，認識・評価することによって，現実の画像での隠れた花の位置推定に必要な要素を調査する．

条件 A, B, C, D のそれぞれについて，RGB 画像 2995 枚と，対応するアノテーション画像 2995 枚をレンダリングし，データセット A, B, C, D とした．各データセットのうち 2696 枚を学習データ，299 枚を学習時に用いる検証データとした．

TenforFlow で実装された DeepLabv3+ によって NVIDIA GeForce 1070, Intel Core i7-6700 を用いて Table 2, Table 3 の条件で学習と比較評価を行なった．ただし，重み (Table 2) は学習データにおけるクラス毎の出現頻度に応じた損失を設定するハイパーパラメータであり，出現頻度が少ないほど値は大きくなる．

Table 2 Weight per class

background	1
branch	10
leaf	10
flower	100
occluded flowers	100

Table 3 Training parameters for DeepLabv3+

model	xception_65
train step	40000
batch normalization	false
batch size	8
crop size	257×257
atrous rate	[6 12 18]
output stride	16
decoder output stride	4

3.2 CG 画像による隠れた花の検出性能評価

CG 画像により学習したモデルにおける隠れた花の検出性能評価について述べる．よりリアリティのある画像（条件 A）を認識するための学習用データに求められる条件を調べるために，A の条件で評価用データセットとして新たに RGB 画像（Fig. 12, Fig. 15, Fig. 18）と真値となるピクセル単位のアノテーション画像（Fig. 13, Fig. 16, Fig. 19）を 378 組レンダリングした．

A, B, C, D の条件で学習した各ネットワークモデルで出力を行った結果を Fig. 14, Fig. 17, Fig. 20 に示す．

評価用データセットについて，学習に背景光源を含まない画像を用いたネットワーク C, D に比べ学習に背景光源を含んだ画像を用いたネットワーク A, B は隠れている花の位置を推定できている．しかし葉がカメラの近くにレンダリングされている場合，テクスチャ情報を学習していないネットワーク B（Fig. 20(B)）は葉を認識できず隠れた花の位置を推定できない傾向が見られた．



Fig. 12 Input image



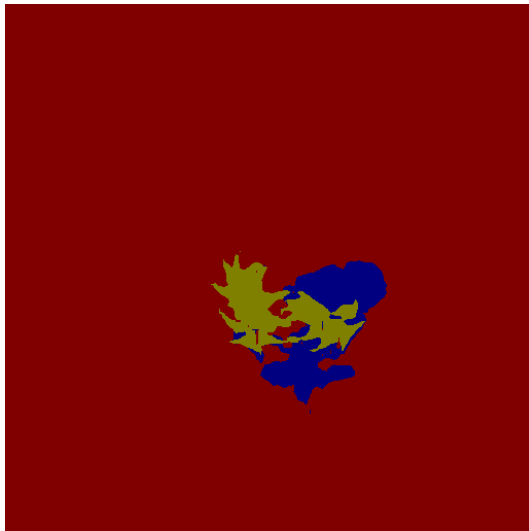
Fig. 13 Ground truth image



(A) Texture and background



(B) No texture and background



(C) Texture and no background



(D) No texture and no background

Fig. 14 Comparison between rendered images



Fig. 15 Input image - 2

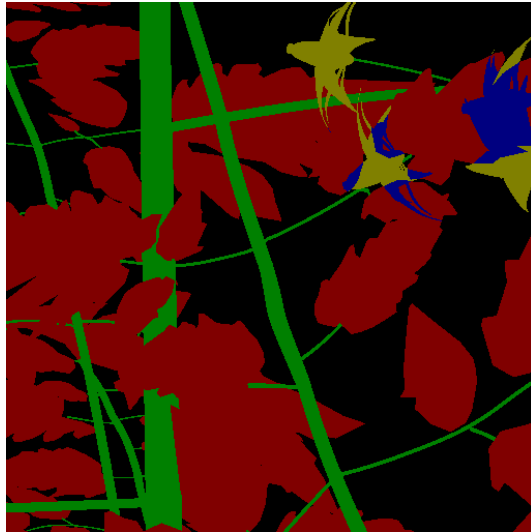
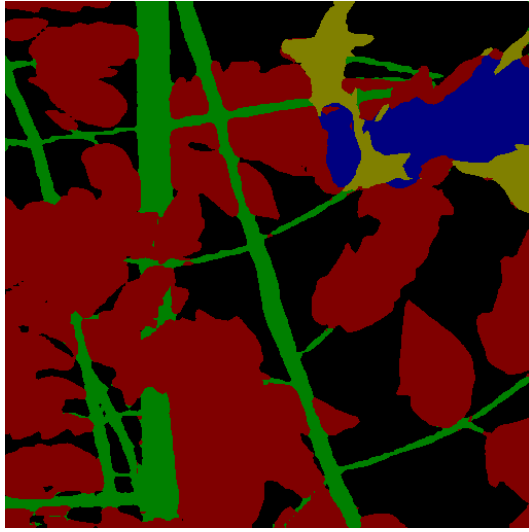
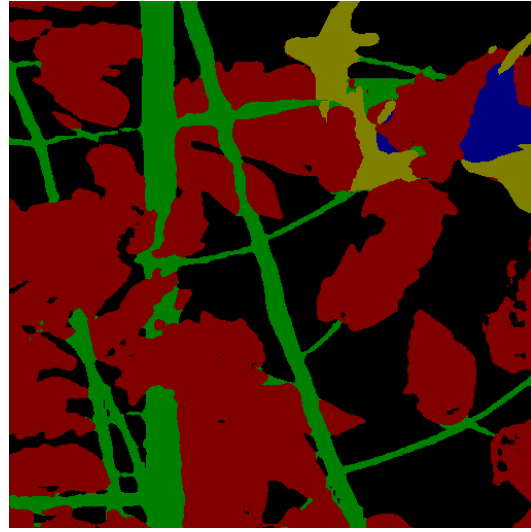


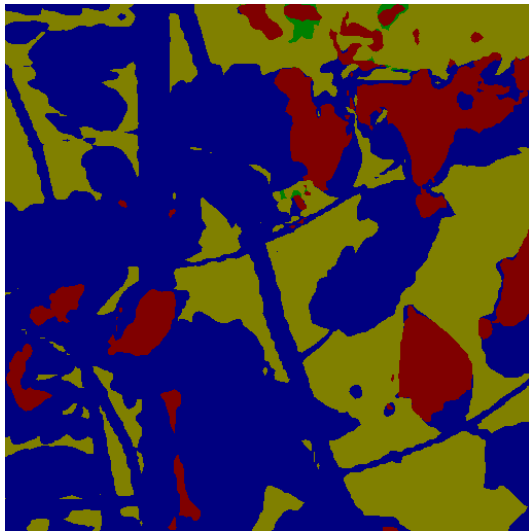
Fig. 16 Ground truth image - 2



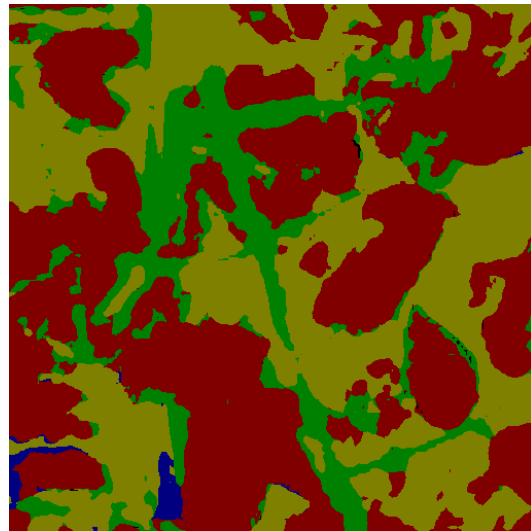
(A) Texture and background



(B) No texture and background



(C) Texture and no background



(D) No texture and no background

Fig. 17 Comparison between rendered images - 2



Fig. 18 Input image - 3

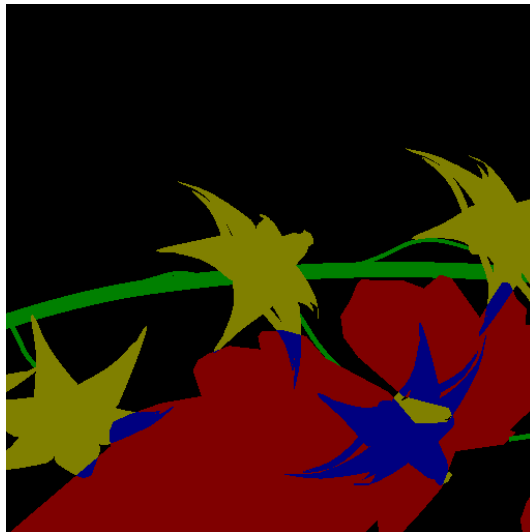


Fig. 19 Ground truth image - 3

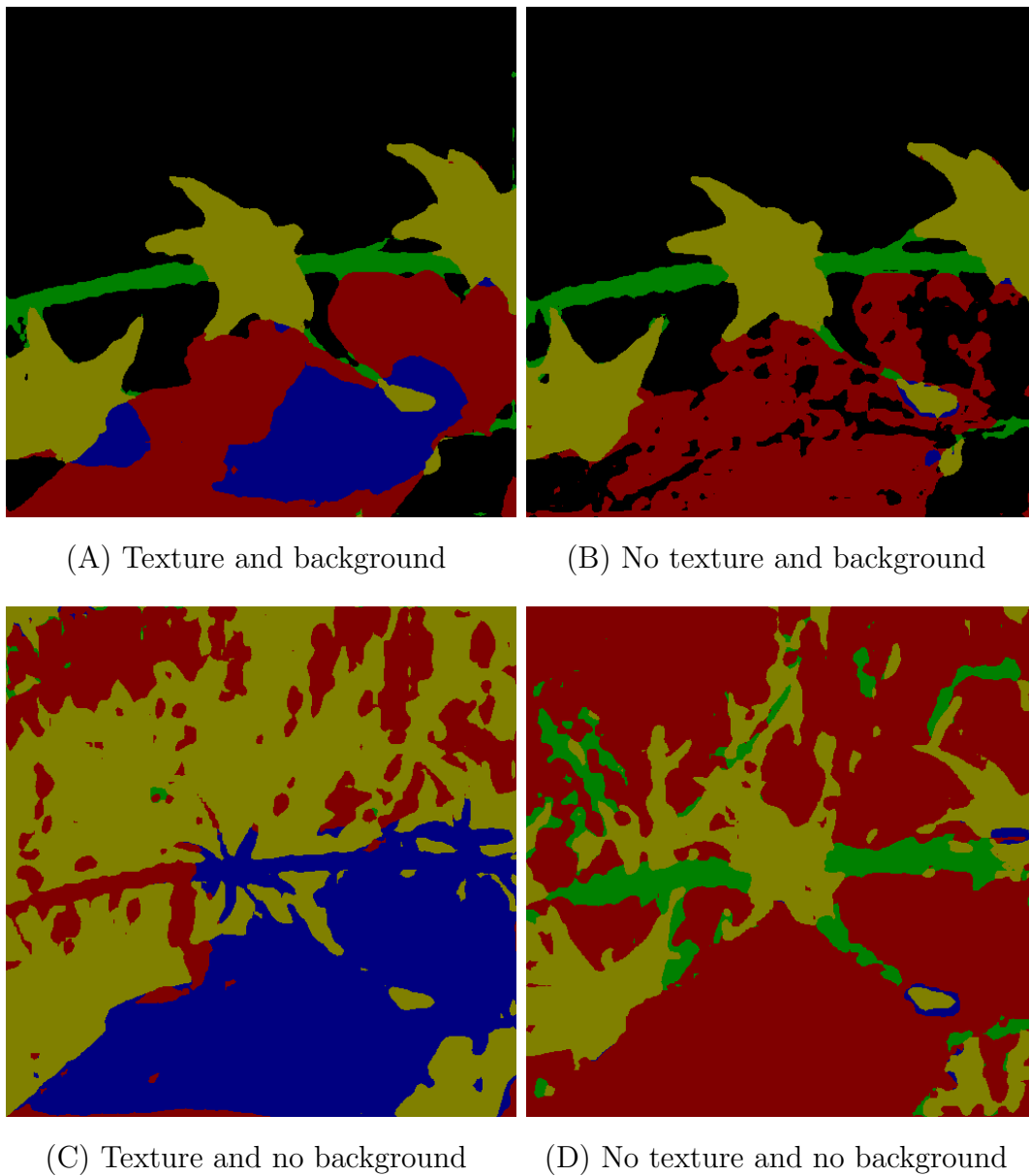


Fig. 20 Comparison between rendered images - 3

セグメンテーション結果の評価指標であるF1値, Recall, Precision, AccuracyをTable 4, Table 5, Table 6, Table 7に示す. F1値はRecallとPrecisionの調和平均で1に近いほどRecallとPrecisionのバランスのとれた正確性が高い指標

と考えられるため、ネットワーク A が全体を通して最も精度が良い結果となっている。本研究目的において重要である花、隠れた花の推定においてネットワーク C, D に比べネットワーク A, B の F1 値が高く、わずかながらネットワーク B に比べネットワーク A のモデルの F1 値が高い結果となっている。これは CG 画像の評価においてはテクスチャの影響より背景光源の影響が大きいことを示している。ネットワーク C, D において部分的にネットワーク D の F1 値が高い箇所があるが、これは今回評価に用いたデータ数に偏りがあったためであると考えられる。また、隠れている花の検出タスクでは、隠れた花の Recall の値が高いことが有効とされるが Recall においてもネットワーク A のスコアが最も高いことが分かる。

Table 4 F1 score with rendered images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.9494	0.9263	0.7331	0.7082	0.3890
(B)	0.9315	0.9166	0.7244	0.7064	0.3559
(C)	0.3050	0.4093	0.0805	0.2743	0.0912
(D)	0.2417	0.6445	0.3630	0.3242	0.1094

Table 5 Recall with rendered images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.9099	0.9606	0.7600	0.9740	0.5870
(B)	0.9225	0.9337	0.7635	0.9683	0.4393
(C)	0.2634	0.4953	0.0713	0.8208	0.6260
(D)	0.2108	0.8026	0.4796	0.8914	0.2576

Table 6 Precision with rendered images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.9957	0.8976	0.7191	0.5748	0.3352
(B)	0.9558	0.9072	0.7001	0.5739	0.3892
(C)	0.9558	0.5697	0.2734	0.2278	0.0724
(D)	0.7542	0.6541	0.3747	0.2639	0.1434

Table 7 Accuracy with rendered images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.9520	0.9554	0.9779	0.9717	0.9775
(B)	0.9415	0.9452	0.9766	0.9712	0.9850
(C)	0.4311	0.4253	0.8113	0.4291	0.5250
(D)	0.4645	0.6042	0.8665	0.5632	0.8893

3.3 現実画像による隠れた花の検出性能評価

現実の枝に隠れたトマトの花の写真 (Fig. 21) に対して比較評価を行なった結果を Fig. 22 に示す. また, 隠れの効果を検証するために, 現実のトマトの花の写真 (Fig. 23) に対して意図的に葉を合成し作成したオクルージョンデータ (Fig. 24) に対して比較評価を行った結果を Fig. 25 に示す. 現実の画像はインターネット [25] から取得した. 人が意図して撮影した画像では, 花は隠れのない状況で撮影されることが多く, 仮想的に CG で隠れを再現することとした.



Fig. 21 Hidden flowers by branches

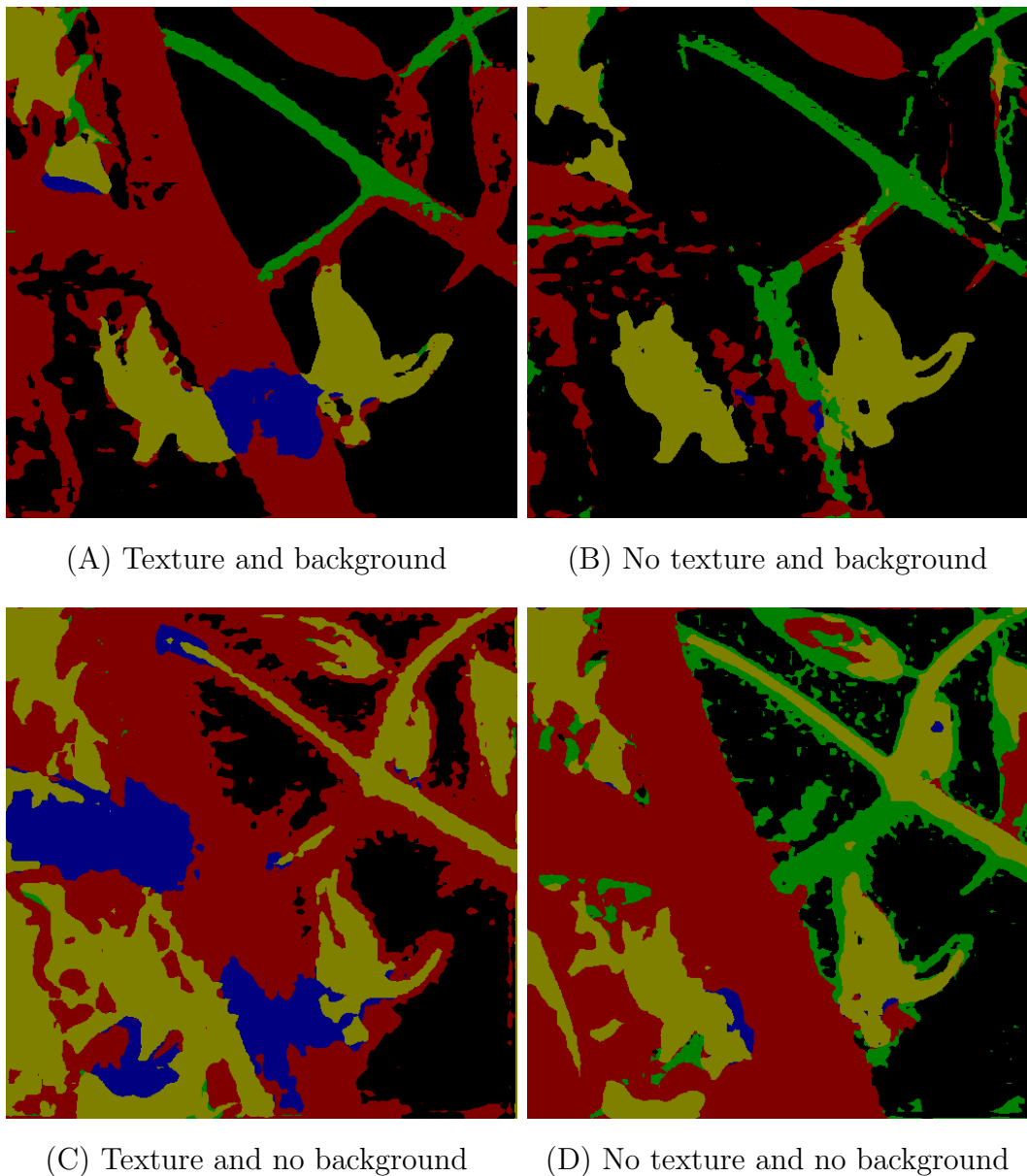


Fig. 22 Comparison of evaluations for hidden flowers by branches

現実の花においても学習に背景光源を含まない画像を用いたネットワーク C, D に比べて学習に背景光源とテクスチャ情報を含んだ画像を用いたネットワーク A は隠れている花の位置を推定できていることが分かる. ネットワーク B は隠れ

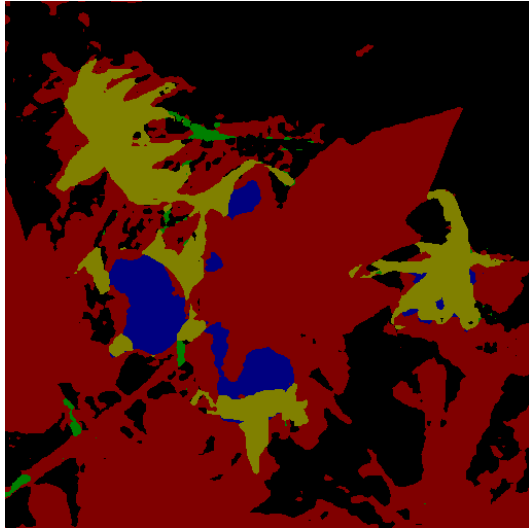
ている花を部分的に推定できているが、レンダリングした画像と同様に葉がカメラの近くに存在している場合、テクスチャ情報を学習していないため、葉を認識できず隠れた花の位置を推定できない傾向が見られた。ネットワーク B に対し、葉のテクスチャ情報を学習しているネットワーク A は、現実の写真においても認識できていることから、単一のテクスチャであっても認識精度の向上のために有効であることを示している。また、1 枚の画像の評価につきおよそ 0.2 秒の時間を要した。これはロボットでの授粉を行う際、隠れた花の推定を行う画像処理として実用可能な処理速度だと考えられる。



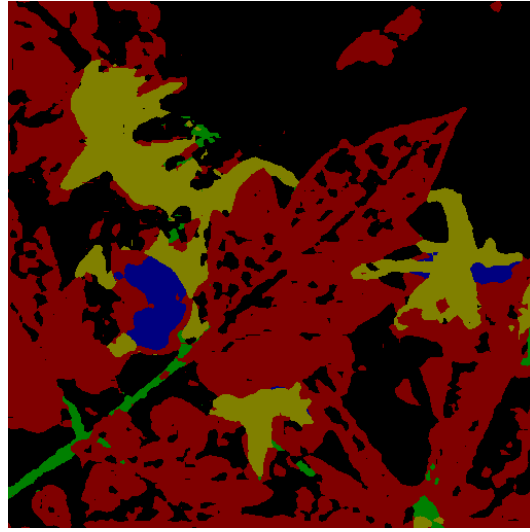
Fig. 23 Real tomato flower



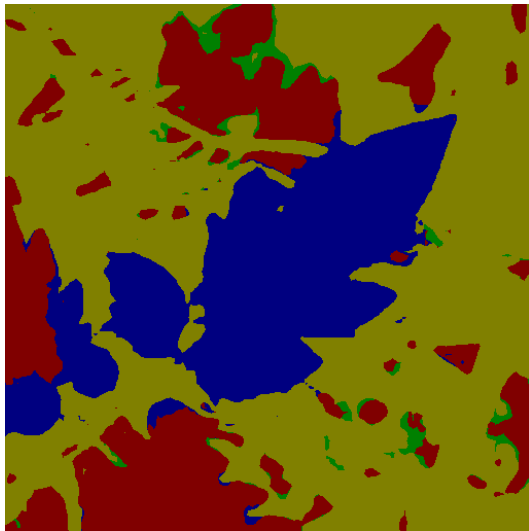
Fig. 24 Synthetic image



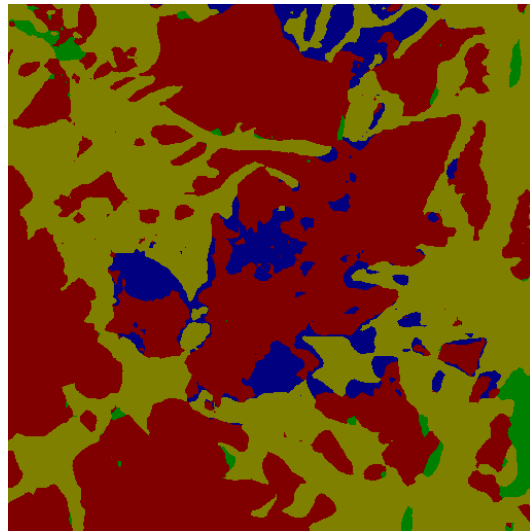
(A) Texture and background



(B) No texture and background



(C) Texture and no background



(D) No texture and no background

Fig. 25 Comparison of evaluations for hidden real flowers

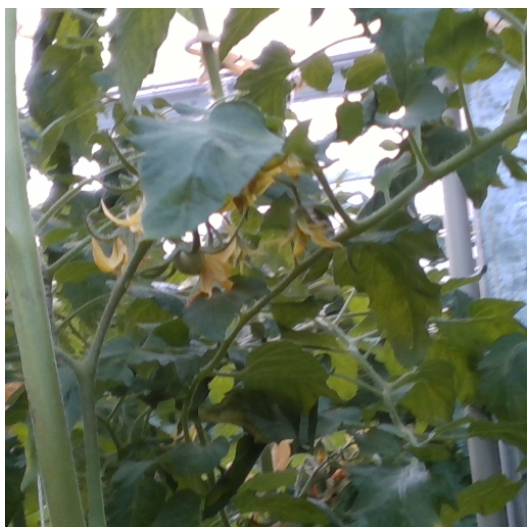


Fig. 26 Real tomato flower - 2

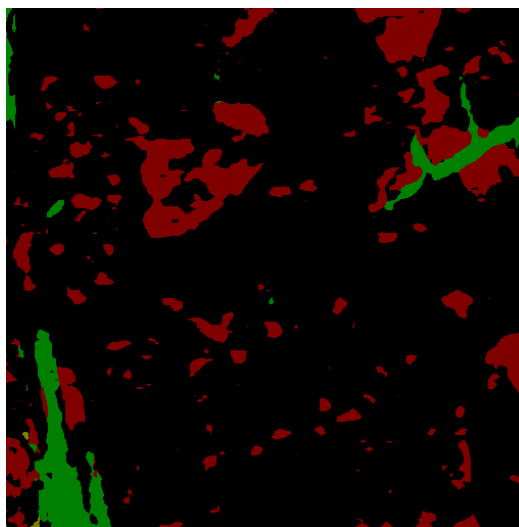
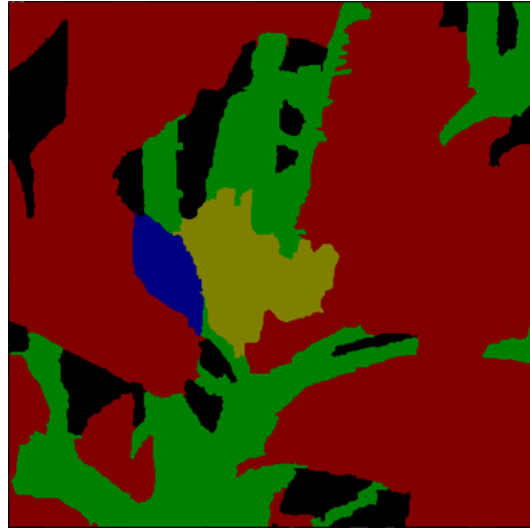


Fig. 27 Failure of flower detection

Intel RealSense D415 により撮影した隠れたトマトの RGB 画像（Fig. 26）に対して，ネットワーク A で検出を行なった結果を Fig. 27 に示す．学習データに含まれていない品種のトマトであったことや，カメラの露出とホワイトバランス，撮影時のブレといった学習データに含んでいない要素によってネットワーク A の認識が失敗することを示唆している．



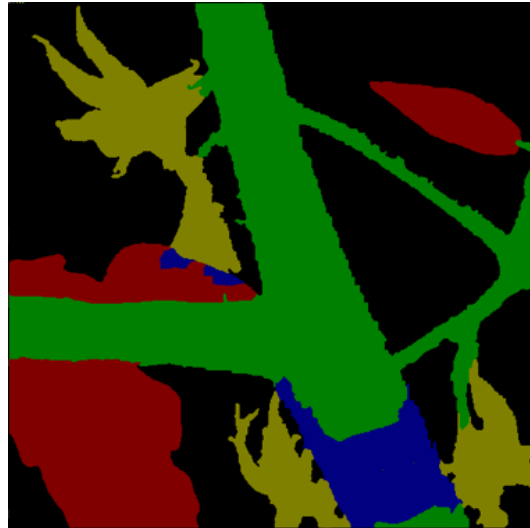
Real tomato flower image - 1



Manually annotated image - 1



Real tomato flower image - 2



Manually annotated image - 2

Fig. 28 Manually annotated images of real tomato flowers

手動でアノテーションを行なった画像（Fig. 28）を正解の値として、レンダリングした画像と同様に、F1 値, Recall, Precision, Accuracy を Table 8, Table 9, Table 10, Table 11 に示す。F1 値と Accuracy においてネットワーク A が全体を

通して最も精度が良い結果となっている。また，Recall においてネットワーク C がネットワーク A より良い値となっているが，これは Fig. 25(C) のように正解クラスの領域に対して多くの領域に対して花，隠れた花として推定しているためであると考えられる。そのため Precision においてはネットワーク C の花と隠れた花クラスが最も低い値となりネットワーク A が最も高い値となっている。

Table 8 F1 score with real images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.6683	0.4293	0.2384	0.6389	0.3141
(B)	0.6244	0.2921	0.1505	0.5848	0.1601
(C)	0.1845	0.2860	0.0313	0.2159	0.2988
(D)	0.2111	0.3422	0.1084	0.2425	0.1242

Table 9 Recall with real images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.8647	0.4436	0.1536	0.7082	0.2628
(B)	0.8728	0.2628	0.0910	0.7256	0.1009
(C)	0.1240	0.2590	0.0191	0.9375	0.7464
(D)	0.1524	0.4014	0.0785	0.9335	0.1061

Table 10 Precision with real images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.5956	0.6297	0.7219	0.6687	0.6921
(B)	0.5287	0.6784	0.6321	0.5169	0.5930
(C)	0.7834	0.4711	0.1540	0.1306	0.1996
(D)	0.7400	0.5127	0.2348	0.1567	0.2644

Table 11 Accuracy with real images in each class and condition

	background	leaf	branch	flower	occluded flowers
(A)	0.6189	0.6343	0.8407	0.9165	0.9399
(B)	0.5467	0.5949	0.8096	0.8888	0.9227
(C)	0.1247	0.3702	0.0445	0.8161	0.4696
(D)	0.3975	0.3904	0.6356	0.3160	0.8452

4. ロボットを用いた現実環境での応用

ロボットを用いた現実環境での実験について述べる．

本実験はロボットアーム先端に装着したカメラにより取得した画像から，ネットワーク A を用いて隠れた花の位置を推定する．推定した花の位置を元に手動によるロボット操作によって視点を変更することで花を発見し，Fig. 1 のように花房における枝の根元をエンドエフェクタによって接触することで，本研究における隠れた花の位置推定手法の妥当性を示すことを目的とする．

4.1 実験環境

実験のため，トマトの花と花の咲いているトマトの作物全体を模したモデルを作成した．正面から見た一つの花の直径は 30 [mm] であり実際のトマトの花の大きさと形状 (Fig. 9) を参考に設計した造花を Fig. 29, Fig. 30 に示す．3D プリンタで作成した物を塗装しており素材はアクリル系樹脂である．トマトの造花と針金を接続し枝全体（花房）を模したものを Fig. 31 に示す．操作対象である作物の全体図を Fig. 33 に示す．



Fig. 29 Artificial flower of tomato - front

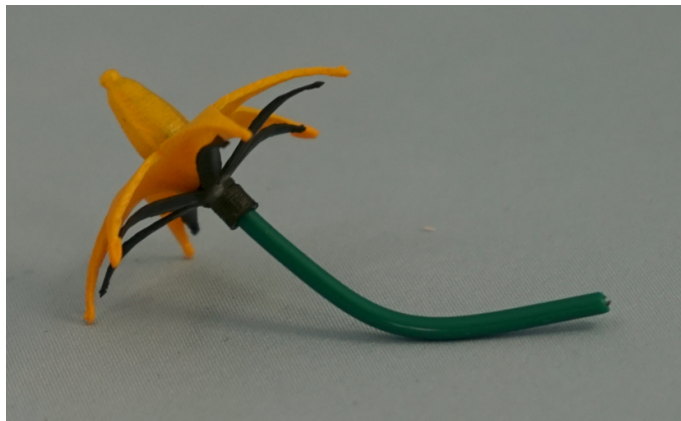


Fig. 30 Artificial flower of tomato - side



Fig. 31 Artificial flower cluster of tomato

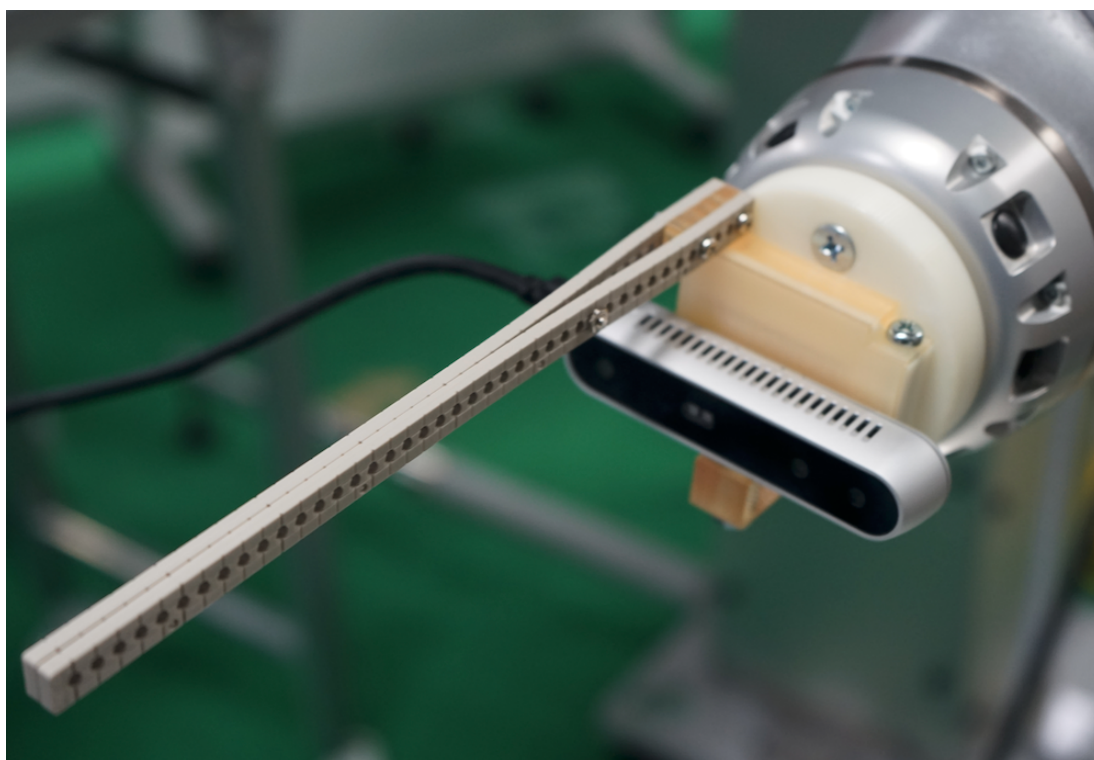


Fig. 32 End effector and camera

本研究ではロボットのエンドエフェクタの接触用の棒が造花根元に接触するこ

とで授粉成功と定義する．カメラは Intel RealSense D415 を使用する．RGB 画像の取得条件は fps30, 縦 720 ピクセル, 横 1280 ピクセルとし DeepLabv3+ 上で学習したモデルを動作させるために取得画像の中心から縦 512 ピクセル, 横 512 ピクセルのサイズにトリミングを行なった上で, 隠れた花の位置の推定を行なった．エンドエフェクタとカメラを Fig. 32 に示す．



Fig. 33 Artificial plant



Fig. 34 KUKA LBR iiwa 14 R820

枝葉の操作を行うロボットとして KUKA 社の LBR iiwa 14 R820 (Fig.34) [26] を使用する．基本的な仕様を Table 12 に示す．このロボットは 7 自由度で本実験において十分な精度を有している．

ロボットの操作とカメラからの情報，DeepLabv3+を統合してモニタリングするために ROS Kinetic [27] を用いた．ロボットの制御を行うために ROS パッケージの iiwa_stack [28] を用いた．ロボットと造花を配置した実験環境全体像を Fig. 35 に示す．

Table 12 Basic specification of the robotic arm

Specification	Value
Model	LBR iiwa 14 R820
Number of controlled axes	7
Volume of working envelope	1.8 [m ³]
Pose repeatability	± 0.15 [mm]
Related payload	14 [kg]
Maximum reach	820 [mm]



Fig. 35 Experiment for pollination of an artificial flower by the robot.

4.2 現実の造花画像による評価



Fig. 36 Occluded flowers by branches in artificial plants

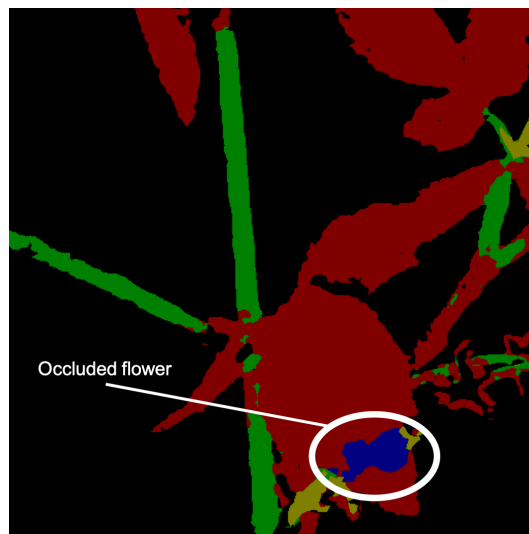
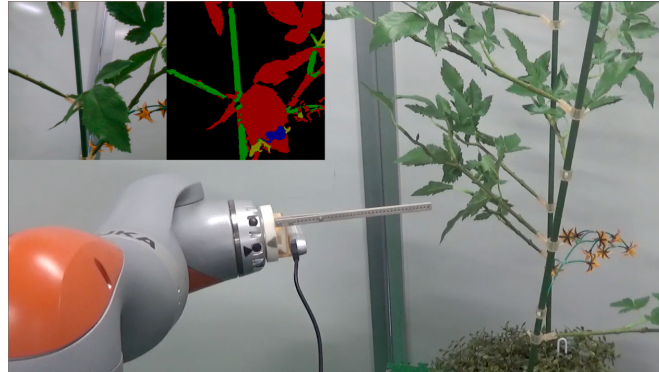
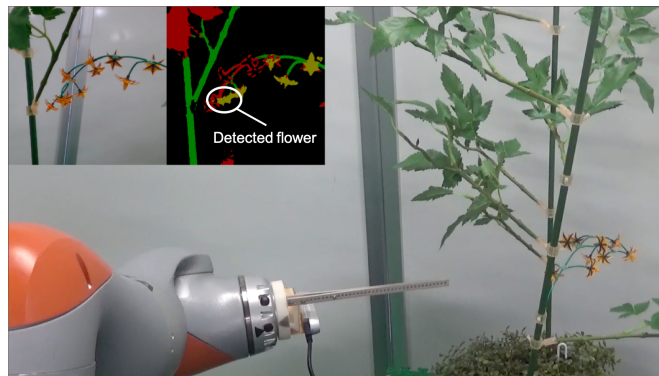


Fig. 37 Detection of artificial occluded flowers



(a) Detection of occluded flowers



(b) Exploring flowers



(c) Contact with flower branch

Fig. 38 Robot pollination with artificial flowers

実験の結果を Fig.38 に示す。画像左上がエンドエフェクタについてカメラから取得した画像であり、右がセマンティックセグメンテーション結果である。Fig.38(a)におけるカメラから取得した画像を Fig.36, 隠れた花の位置推定の結果を Fig.37 に示す。Fig.37 の青く示された領域が隠れた花が存在すると推定された領域であり、画像下部に存在しているため視点を下方に平行移動させた。下方に視点を移動した結果 (Fig.38(b)) 推定された隠れた花の位置に花が存在することを確認し、Fig.38(c) において花房根元に接触することに成功した。以上の結果より、本研究の提案手法である CG のみの学習によって現実環境における部分的に隠れた花の検出手法の妥当性を示した。また花の検出の過程において花の一部が検出できない例を確認した。

4.3 考察

一部の花を検出できなかった理由については学習に用いた CG データには含まれていない露出やブレなどのカメラの特性、実験環境における光源の差であると考えられる。これらの課題をの解決する方法として CG 環境においてカメラ特性として生じる露出やブレの表現と、花の色、形に微小な個体差を発生させること、撮影時の条件を限定して撮影することが挙げられる。

5. おわりに

5.1 まとめ

本研究では、3DCGの単純なモデルと単一のテクスチャでレンダリングした画像のみをDeepLabv3+の学習に利用することで、現実の画像に対しても部分的に隠れた花の位置推定を実現した。それによって、3DCGでレンダリングされた画像のみの学習によって現実の画像での隠れた花の位置推定のために作物のテクスチャと背景（光源）が重要な要素であることを示した。

3DCGの画像のみで学習したネットワークを用いて、人工的な環境下において隠れたトマトの花の位置を推定し、ロボットの視点を移動する授粉実験を行なった。実験により現実環境において、3DCGの画像のみで学習したネットワークがロボットを用いた人工授粉に有効であることを示した。

本研究で得られたアノテーション情報と、深度情報を組み合わせることで、現実環境において枝や葉の裏に隠れた花を推定を行い、枝や葉の操作や視点の移動といったアプローチによってハイワイヤー栽培等の複雑な環境下においてもロボットによる人工授粉精度の向上が期待できる。また、ロボットにより隠れた花を検出し人工授粉を行う可能性を示した。

5.2 今後の課題と展望

今後の課題として、ロボットと連動したリアルタイムでの隠れた花の推定システムの構築と農場への設置を前提とした移動を制限を考慮した動作開発により実際に授粉させることが挙げられる。また、本手法は単純な植物モデルと少数のテクスチャを用意するだけで有効なデータセットの作成が可能であり、トマトの授粉だけでなく他の作物への応用、摘花や測定、収穫といった別の作業分野への応用も考えられる。

謝辞

本論文は著者が奈良先端科学技術大学院大学情報科学研究科博士前期課程在学中にロボティクス研究室で行った研究をまとめたものである。

本研究を遂行するにあたり，貴重なご指導とご助言と賜りました先端科学技術研究科 情報科学領域の小笠原司教授に深く感謝致します。

本研究を遂行するにあたり，貴重なご指導とご助言と賜りました先端科学技術研究科 情報科学領域の向川康博教授に深く感謝致します。

本研究の方針決定，遂行，論文執筆などにおいて多くのご助言，ご指導を頂きました先端科学技術研究科 情報科学領域の高松淳准教授に深く感謝致します。

本研究の方針決定，遂行，論文執筆など研究全般に渡り幾度となく丁寧にご指導，ご助言を頂きました先端科学技術研究科 情報科学領域の丁明助教に深く感謝致します。

本研究の方針決定，遂行，論文執筆など研究全般に渡り幾度となく丁寧にご指導，ご助言を頂きました先端科学技術研究科 情報科学領域の GARCIA RICARDEZ Gustavo Alfonso 助教に深く感謝致します。

大学生活や学会発表における事務処理の面で様々なサポートをして頂きました本研究室秘書の大脇美千代氏，池田美輝氏に深く感謝致します。

本研究の遂行にあたり，撮影の協力やご助言を頂いた三重県農業研究所の磯山陽介氏に感謝致します。

本研究の遂行にあたり，様々な議論やご助言を頂いた本研究室博士後期課程の山崎亘氏，趙崇貴氏に感謝致します。

先輩・同期ならびに後輩の皆様には研究に対してのご助言や様々な形での支援をいただきました。感謝致します。

最後に，長年の学生生活を支えて頂いた家族，友人達に深く感謝致します。

参考文献

- [1] 政府統計の総合窓口 (e Stat): “農業経営統計調査-平成 19 年産品目別
経営統計-施設野菜作経営-大玉トマト”, [https://www.w3.org/TR/2004/
REC-rdfconcepts-20040210/](https://www.w3.org/TR/2004/REC-rdfconcepts-20040210/) (2019 年 1 月 31 日参照) .
- [2] Nicholas Ohi, Kyle Lassak, Ryan M. Watson, Jared Strader, Yixin Du,
Chizhao Yang, Gabrielle Hedrick, Jennifer Nguyen, Scott Harper, Dylan
Reynolds, Cagri Kilic, Jacob Hikes, Sarah Mills, Conner Castle, Benjamin
Buzzo, Nicole Waterland, Jason N. Gross, Yong-Lak Park, Xin Li, and
Yu Gu: “Design of an autonomous precision pollination robot,” in Proc.
of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), pp.
7711–7718, 2018.
- [3] Ting Yuan, Shunlu Zhang, Xiyu Sheng, Dashuai Wang, Yue Gong, and Wei
Li: “An autonomous pollination robot for hormone treatment of tomato
flower in greenhouse,” Proc. of 3rd Int. Conf. on Systems and Informatics
(ICSAI), pp. 108–113, 2016.
- [4] 黒崎秀仁, 大森弘美, 高市益行, 岩崎泰永: “トマト低段密植栽培に対応した
着果処理ロボットの開発 (第 1 報)”, 農業機械学会誌, vol. 74, no. 6, pp.
490–497, 2012.
- [5] Naoshi Kondo, Koki Yata, Michihisa Iida, Tomoo Shiigi, Mitsuji Monta,
Mitsutaka Kurita, and Hiromi Omori: “Development of an end-effector for
a tomato cluster harvesting robot,” Engineering in Agriculture, Environment
and Food, vol. 3, no. 1, pp. 20–24, 2010.
- [6] Takeshi Yoshida, Takanori Fukao, and Takaomi Hasegawa: “Fast detection
of tomato peduncle using point cloud with a harvesting robot,” Journal of
Robotics and Mechatronics, vol. 30, no. 2, pp. 180–186, 2018.

- [7] Kathleen V. Lee, Nathalie Steinhauer, Karen Rennich, Michael E. Wilson, David R. Tarpy, Dewey M. Caron, Robyn Rose, Keith S. Delaplane, Kathy Baylis, Eugene J. Lengerich, Jeff Pettis, John A. Skinner, James T. Wilkes, Ramesh Sagili, Dennis vanEngelsdorp, and for the Bee Informed Partnership: “A national survey of managed honey bee 2013-2014 annual colony losses in the usa,” *Apidologie*, vol. 46, no. 3, pp. 292–305, 2015.
- [8] The White House Office of the Press Secretary: “Fact Sheet: The economic Challenge Posed by Declining Pollinator Populations,” <https://obamawhitehouse.archives.gov/the-press-office/2014/06/20/fact-sheet-economic-challenge-posed-declining-pollinator-populations,2014> (2019 年 1 月 31 日参照) .
- [9] Noboru Kawamura, Kiyoshi Namikawa, Tateshi Fujiura, and Motonobu Ura: “Study on agricultural robot (part 1),” *Journal of the Japanese Society of Agricultural Machinery*, vol. 46, no. 3, pp. 353–358, 1984.
- [10] Mitsuji Monta: “Robots for bioproduction systems,” *IFAC Proceedings Volumes*, vol. 33, no. 29, pp. 1–11, 2000. 2nd IFAC/CIGR Int. Workshop on Bio-Robotics, Information Technology and Intelligent Control for Bioproduction Systems (BIO-ROBOTICS II), Sakai, Osaka, Japan.
- [11] Qing Chun Feng, Wei Cheng, Jian Jun Zhou, and Xiu Wang: “Design of structured-light vision system for tomato harvesting robot,” *International Journal of Agricultural and Biological Engineering*, vol. 7, no. 2, pp. 19–6, 2014.
- [12] Paula Tarrío, Ana M. Bernardos, and José R. Casar and Juan A. Besada: “A Harvesting Robot for Small Fruit in Bunches Based on 3-D Stereoscopic Vision,” in *Proc. of the 4th World Congress Conf. on Computers in Agriculture and Natural Resources*, pp. 270–275, 2006.

- [13] Gabriel J. Brostow, Julien Fauqueur, and Roberto Cipolla: “Semantic object classes in video: A high-definition ground truth database,” *Pattern Recogn. Lett.*, vol. 30, no. 2, pp. 88–97, 2009.
- [14] Stephan R. Richter, Vibhav Vineet, Stefan Roth, and Vladlen Koltun: “Playing for data: Ground truth from computer games,” in *Proc. of the 14th European Conf. on Computer Vision (ECCV)*, pp. 102–118, 2016.
- [15] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio M. Lopez: “The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 3234–3243, 2016.
- [16] Isokane Takahiro, Okura Fumio, Ide Ayaka, Matsushita Yasuyuki, and Yagi Yasushi: “Probabilistic plant modeling via multi-view image-to-image translation,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros: “Image-to-image translation with conditional adversarial networks,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967–5976, 2017.
- [18] 佐々木英和, 河崎靖, 安場健一郎, 鈴木克己, 高市益行: “ハイワイヤー誘引栽培したトマトの主茎基部側枝が果実の糖度と収量に及ぼす影響”, *野菜茶業研究所研究報告*, no. 12, pp. 1–6, 2013.
- [19] 渡辺慎一: “低段密植栽培による新たなトマト生産”, *野菜茶業研究集報*, no. 3, pp. 91–98, 2006.
- [20] 高市益行, 田中和夫, 中島規子: “高度機械化・省力化作業に適合したトマトの新栽培様式の開発ハイワイヤー整枝における生育・収量特性と機械収穫に対する適性”, *野菜・茶業試験場久留米支場研究年報*, no. 12, pp. 122–126, 1999.

- [21] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam: “Encoder-decoder with atrous separable convolution for semantic image segmentation,” in Proc. of the 15th European Conf. on Computer Vision (ECCV), 2018.
- [22] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman: “The pascal visual object classes (voc) challenge,” International Journal of Computer Vision, vol. 88, no. 2, pp. 303–338, 2010.
- [23] 薄上秀男: “トマトの着果習性と出葉体系との関係”, 農及園, vol. 39, pp. 1281–1282, 1964.
- [24] Cordts Marius, Omran Mohamed, Ramos Sebastian, Rehfeld Timo, Enzweiler Markus, Benenson Rodrigo, Franke Uwe, Roth Stefan, and Schiele Bernt: “The cityscapes dataset for semantic urban scene understanding,” in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 3213–3223, 2016.
- [25] “flickr,” <https://www.flickr.com/photos/48224214@N02/6115505139/> (2019 年 1 月 31 日参照) .
- [26] “Industrial-robots/LBR-iiwa,” <https://www.kuka.com/en-us/products/robotics-systems/industrial-robots/lbr-iiwa> (2019 年 1 月 31 日参照) .
- [27] “ROS Kinetic Kame,” <http://wiki.ros.org/kinetic> (2019 年 1 月 31 日参照) .
- [28] “iiwa_stack,” https://github.com/IFL-CAMP/iiwa_stack (2019 年 1 月 31 日参照) .