

# 環境認識にAIを用いたトマト収穫ロボット

Tomato-Harvesting Robot using AI for Environment Recognition

上 垣 俊 平  
Shumpei Uegaki

新 崎 誠  
Makoto Shinzaki

荒 木 秀 和  
Hidekazu Araki

上 田 大 介  
Daisuke Ueta

戸 島 亮  
Ryo Toshima

山 崎 龍 次  
Ryuji Yamazaki

## 要 旨

国内農業において労働者不足が大きな課題となるなか、依然として野菜や果物といった農作物の栽培は人作業に頼っている。それらの社会課題に対して筆者らは作業者に大きな負担が発生する収穫作業を自動化するためのロボットを開発している。取り組みのなかで、画像処理により果実や環境認識を行って収穫条件を決定していたが、実際の農園で検証を行うと、果実検出率などで限界があった。そこで、人工知能（AI）を用いて、果実検出率の向上や、環境の詳細認識を試み、一定の成果を得た。それらの手法や結果について報告する。

## Abstract

In domestic agriculture, labor shortage is a major issue. However, cultivation of vegetables and fruits depends on work performed by humans. With regard to such social issues, we are developing robots to automate harvesting work that depends heavily on labor workers. In development, we used image processing to recognize both fruits and the environment. In this case, the fruit detection rate was limited. Therefore, we attempted to improve the fruit detection rate and the recognition of details in the environment using AI. Herein, we report on the methodology and results.

## 1. はじめに

国内の農業現場では、農業就業人口の平均年齢が66歳を超え、高齢化が急速に進み、労働力不足が深刻となっており、農作業における省力、軽労化をさらに進めるとともに、新規就業者への栽培技術の継承などが重要な課題となっている。他方、異業種では、ロボット技術や人工衛星を活用したりモータセンシング技術、クラウドシステムをはじめとしたICT（Information and Communication Technology）の活用が進展しており、農業分野でもそれらの活用が期待されている。

現在、実用化されているロボットの多くは産業用ロボットであり、農林水産分野におけるロボットの割合は少なく、(国研)農業・食品産業技術総合研究機構などにより農業用ロボットの研究開発が鋭意行われているが、実用化までに至る例は少ない[1]。農業用ロボットの実用化が遅れている原因として、コスト、安全性、屋外作業における環境の影響や人の経験による作業などの問題が考えられる[2][3]。特に単純ながらも長時間かつ重労働な収穫作業を自動化する取り組みは長く検討されているが、収穫率の低さが問題となり実用化されていない。これらの課題を克服し、農業用ロボットを普及させることができれば、新たなロボット市場が生まれる。

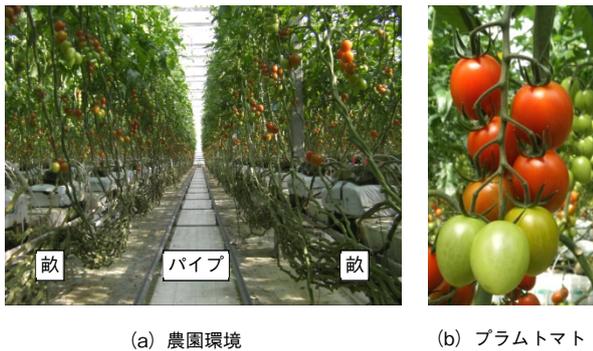
このようななか、筆者らはこれまで果実収穫作業の自動化・省人化の達成に向け、トマト収穫ロボットの試作機開発を行ってきた。開発初期から画像処理を駆使して

果実やその周囲環境を検出することで収穫対象とロボットとの相対位置や収穫姿勢を決定することを試みていたが、農園といった条件変動の激しい環境においては精度に限界があった。そこで、近年目覚ましく発展してきたディープラーニング（深層学習）による物体の検出や認識技術に着目し、それらを用いて果実の検出や環境認識を試みたので報告する。

## 2. トマト収穫ロボット

### 2.1 動作環境と収穫対象

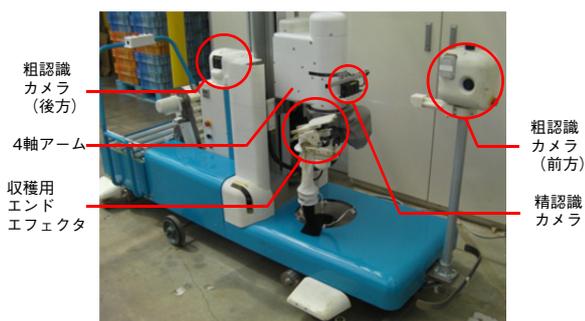
当ロボットは、オランダで開発され、近年日本でも導入が進められているフェンロー型施設でトマトを栽培している農園を対象としている。それらの農場では**第1図 (a)**のように、ハウス施設内の温度を制御するため畝と畝との間に温水用パイプを設置しており、当ロボットはそのパイプを走行レーンとして走行する方式を採用した。従って、フェンロー型の農園であれば使用可能な仕様となっている。収穫対象は他のミニトマトと比較して高価な品種であり、かつ密集してなることから人の手でも収穫作業の難易度が高く、ロボット化のメリットが大きいプラム種を採用した(**第1図 (b)**)。



第1図 農園環境とプラムトマト  
Fig. 1 Farm environment and plum tomato

## 2.2 ロボットの構成

ロボットの構成を第2図に示す。ロボットはレール上を走行するための車輪を4輪有し、前進/後進の両方で果実収穫を行うため、果実認識（粗認識）に使用するカメラを前後に搭載している。また、果実とロボットとの相対位置を認識するために、カラー画像、赤外線画像、深度画像を取得可能な当社グループ製精認識カメラを搭載している。アームは上下方向に1軸、水平方向に3軸の計4軸を有している。水平方向については本来2軸あれば手先位置を制御することが可能だが、障害物の回避や房背面側への回り込みなどを実現するため1軸冗長化させている。アーム先端に収穫用エンドエフェクタ（以下、EE）を取り付けている。駆動源に鉛蓄電池を使用し、充電は専用充電器を用いて手動で行う仕様となっている。



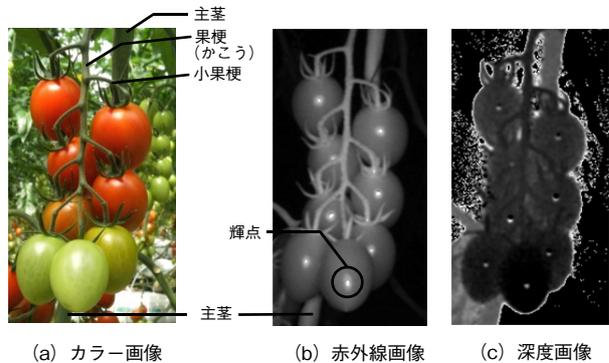
第2図 トマト収穫ロボット  
Fig. 2 Tomato-harvesting robot

## 3. 画像処理による果実環境認識

### 3.1 果実・環境画像の取得

果実検出や環境認識は精認識カメラを用いて取得した画像を基に実施する。精認識カメラは第3図に示した、(a) カラー画像、(b) 赤外線画像、(c) 深度画像を取得する。

その際、太陽光の影響を緩和するためと夜間対応を目的としてカメラ上部に搭載したLEDライトを発光させている。



第3図 精認識カメラによる取得画像  
Fig. 3 Image acquired by precision recognition camera

### 3.2 画像処理による果実検出

果実検出は次の手順で実施する。まず、赤外線画像により果実中央部付近に発生する輝点を検出する。この際、他の反射強度の強い光と果実中央部に発生する輝点とを混同しないよう、輝点の形状やサイズ、また近接画素の輝度勾配から導出されるHOG特徴（Histograms of Oriented Gradients）といった特徴量を抽出し、パターン認識モデルの1つであるSVM（Support Vector Machine）により判別を行う。そして、カラー画像から輝点位置の色相が熟した果実として設定した色味（HLS色空間のH値）であるかを判定したうえで、深度画像における輝点位置の距離データから奥行き方向で一定範囲内にある果実を収穫対象とする。

### 3.3 画像処理による環境認識

果実を収穫する際、果実を包み込む形でEEを進入させるが、果実周囲に主茎などがあるとEEと干渉することで収穫対象である果実の房が動いてしまい収穫することができない。また、EEは房中央部を上下に走る果梗（かこう）と、果実の中心位置とを結んだ線の延長上に配置することで収穫する仕組みとなっていることから、果実周囲の状況も詳細に把握する必要がある。

果梗の認識手順を説明する。果梗は重力方向に垂れ下がるため、赤外線画像から画像縦方向に細長い線成分を抽出し、2値化する。2値化した線分に対して一定間隔で横方向画素数と深度画像値を取得し、各位置の太さを算出、各位置の太さの中央値を線成分の太さとする。次に、太さが設定値（2.5 mm～9 mm）の範囲内に入っている線成分を果梗として認識する。

主茎の認識手順を説明する。主茎は赤外線画像で白色

が強い領域から、先に求めた果実と果梗以外を障害物とみなし、そこから色相が緑でかつ奥行き方向一定範囲内の物体を主茎として認識する。

以上の認識情報をまとめたうえで環境マップを作成し、EEが収穫する際の進入方向、姿勢を決定する。

#### 4. 画像処理による果実環境認識の課題

本章では、画像処理により果実・環境認識を実施するうえで発生する課題について説明する。

##### 4.1 果実検出における課題

果実検出は、果実が3次元的な楕円（だえん）形状である特性をいかして、果実中央部付近に現れる輝点をベースに検出する仕組みとなっている。従って、輝点が見つからないと検出候補にならないという問題がある。果実がロボットに対して整然と並んだ状態で結実する房であればすべての果実を検出することも可能だが、傾きが生じてしまうと手前側にある果実と奥側にある果実とで領域の干渉が発生し、輝点が見つからないことから検出できず、検出率を落とす要因となっている。

##### 4.2 果梗認識における課題

果梗の認識は、赤外線画像縦方向に見られる線成分を抽出することで行っている。その際、縦方向に長い線成分を中心に判定を行っているが、まれに果実表面に発生する輝点が高い線成分を形成する場合がある。また、果梗自体もロボットに対して全面が見える訳ではなく、手前に果実などが存在すると、それらが障害物となって果梗の一部しか確認できない場合がある。それらの条件が重なると、果梗が奥に存在する条件でも、手前に存在する果実表面の長い輝点を果梗として誤認識してしまうことがある。また、果梗頭頂部の萼（がく）が上方向に伸びた場合、果梗と同じような縦方向の線成分を表示し、誤認識してしまう場合がある。このように、赤外線画像を基として、果実や萼といった果梗以外の物体を果梗として誤検出すると、果梗位置と果実中心位置から求められるEEの姿勢に狂いが生じ、本来望ましい方向とは異なった角度からEEが進入することから、収穫率を落とす要因となっている。

##### 4.3 適収穫判断における課題

収穫すべき果実であるかは、果実検出時のカラー画像の色相値で判断している。しかしながら、農園の運営において収穫対象の色相レベルは常に変更されている。例えば、収量を増やしたい日などは多少赤みが弱い果実で

も収穫対象とすることがある。従って、日々の収穫対象を農園運営者側で制御する必要があるが、現在のシステムは作業者が直感的にそれを選択できるようには設計されていない。

#### 5. ディープラーニングによる果実環境認識

##### 5.1 ディープラーニングによる物体検出

近年、ディープラーニングを用いた物体検出は目覚ましい発展を遂げている。既存手法とディープラーニングを用いる手法との大きな違いは、画像から求める特徴量抽出を検出者自らが設計した手法で行うのではなく、ニューラルネットワークに委ねてしまう点にある。従って、既存手法でなされてきたような、検出対象に対して個別に行ってきた特徴量抽出手法の開発を行う必要がなく、非常に汎用的な検出手段を獲得することが可能である。また、検出精度は学習用画像を増やすことで高めることが可能であり、検出対象や周囲環境が徐々に変化していくような状況においても、随時取得した画像データから学習を重ねることで変化に対してロバストなシステムを構築することが可能である。

##### 5.2 ディープラーニングによる果実検出

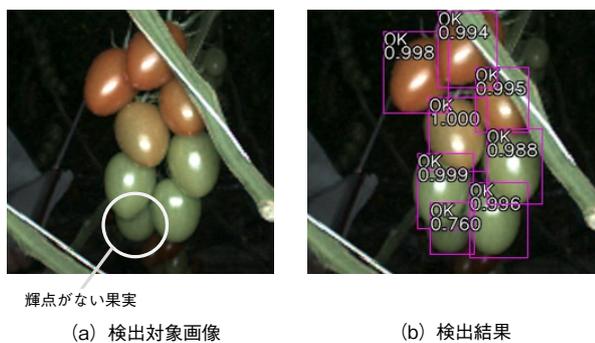
果実検出においては輝点が見られない果実を検出することが課題であることから、果実間で重なりが発生した房を対象に、後側に配置されている果実の検出を目的に検出モデルの開発を行った。また、一般的に、1枚の大きな画像のなかから対象となる物体を検出するためには、大きな画像の一部を大量に切り出してきて、それぞれの画像に対して判定を行う方法が考えられるが、切り出し画像の枚数が増えるほど計算量が増えてしまい、収穫タクトに影響を及ぼす。従って、高精度かつ高速な物体検出が行えるモデルが必要となる。本件では、検出モデルにPVANet[4]を用いて果実検出を試みた。

PVANetは大規模画像認識コンペティションの1つであるPASCAL VOC2012データセット（The PASCAL Visual Object Classes Challenge 2012）で検出率が1位のResNet-101[5]に続いて82.5%と高く、計算コストは12.3%しか必要としない非常に軽量で高精度な検出モデルであり、精度と速度を両立可能なモデルとして採用した。

本件では、農園で撮影した画像から果実画像を2000枚準備し、学習済みPVANetモデルをファインチューニングして検出モデルを作成した。PVANetモデルを用いて果実検出を行った際の検出結果を第4図に示す。結果の図は、果実を検出すると検出したことを示す矩形（くけい）窓

が追加され、矩形窓上部に分類時の確率が出力される。

**第4図 (a)** 検出対象画像にて左下にある果実に輝点がないことがわかる。既存手法ではこの果実は見落としてしまい検出対象とならない。PVANetモデルを用いた場合の検出結果 (**第4図 (b)**) では輝点のない果実も他果実と比べて若干分類確率は低いが検出されていることがわかる。また、他画像も用いて既存手法と比較した結果を**第1表**に示す。既存手法では、対象となる果実が157個あるうち、検出数が148個 (検出率95.5%)であったが、PVANetモデルでは155個 (検出率98.7%)の果実を検出し、検出率向上に有効なことを示した。また、本検出はVGA画像 (画素数: 640×480) に対して行った場合、GPU (NVIDIA製TitanX) を搭載する端末で約50 msで計算されており、計算時間も高速なことが確認できた。



第4図 果実検出結果

Fig. 4 Tomato detection result

第1表 果実検出結果

Table 1 Tomato detection result

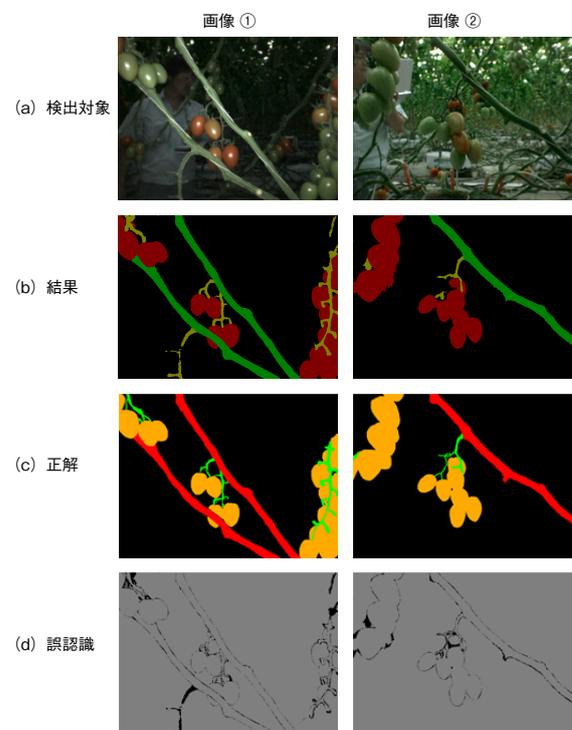
|         | 既存手法      | PVANetモデル |
|---------|-----------|-----------|
| 検出数/対象数 | 148 / 157 | 155 / 157 |
| 検出率     | 95.5 %    | 98.7 %    |
| 未検出数    | 9         | 2         |
| 誤検出数    | 10        | 0         |

### 5.3 ディープラーニングによる果梗認識

本ロボットが果実収穫を行ううえで必要な分類は、果実、主茎、果梗が主たる対象となるが、果実以外は長尺な物体であり、かつ他物体との干渉が多く、果実検出で用いたような物体検出に不向きである。そこで、画像の各画素がどの物体に属するかを推定するセマンティック・セグメンテーションという手法を用いてそれらの分類を行った。

検出モデルにはResNet-38[6]を使用した。ResNet-38はPASCAL VOC2012データセットにおいて精度84.9% (8位) の高精度のモデルである。スクリプト言語を使用しない形で実装されており、ロボットへの搭載が比較的ス

ムーズに実施可能なため本モデルを採用した。特徴として、ネットワークの階層を深くすることを目的としてネットワークにバイパスを設けたResidual Networksの層数を減らし、代わりに特徴マップ数 (チャンネル数) を増やすことで構造の最適化を図っている。また、セマンティック・セグメンテーション向けに、一般的に画像認識に用いられる畳み込みニューラルネットワークをすべての層に用いたFully Convolutional Networksに類似した構造へカスタマイズされている。学習は、画素ごとに果実、主茎、果梗、その他背景の分類にラベル付けを行った1050枚の画像を用いて行った。本検出モデルを用いて環境認識を行った結果を**第5図**、**第2表**に示す。なお、**第2表**の結果は評価画像35枚の結果の平均値を示す。



第5図 環境認識結果

Fig. 5 Environment recognition result

第2表 環境認識結果

Table 2 Environment recognition result

|         | 果実   | 主茎   | 果梗   | クラス平均 |
|---------|------|------|------|-------|
| 認識率 [%] | 97.9 | 94.3 | 83.7 | 93.2  |

**第5図**、**第2表**より、検出対象と背景とはよく分離できており、果実の個体や主茎本体はすべて検出に成功していることが確認できた。また、物体境界の画素で誤差が多く認識率を若干落とす結果となったが、果実、主茎については問題ない範囲と言える。果梗では他物体と比較

して認識率が低くなった。セマンティック・セグメンテーションの特性上、細い（画素数が少ない）物体は認識誤差が起きやすいため精度向上の対策が必要となる。今後の課題として検討を進めていく。なお、計算時間は果実検出と同じ条件で330 ms程度であった。

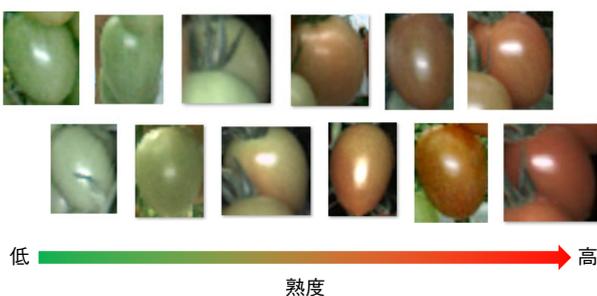
このように果梗位置を画素レベルで正確に認識可能となったことから、従来手法において発生していたような果梗の誤検出によるEEの姿勢の間違い発生は減少すると考えるが、今後の検討のなかで詳細な評価を実施していく。

#### 5.4 ディープラーニングによる収穫対象の決定

農園における収穫対象の判定は人間が目視によって行っているため、ロボットが実施する場合でもカメラで取得したカラー画像によって実施可能と判断して検討を行った。

収穫すべきか否かの判定は果実画像から熟度を判定し、設定された熟度以上の果実を収穫するようなシステムが構築できれば人作業の代替として運用することが可能となる。従って、ロボットが果実の熟度判定をする必要があることから、本ロボットでは適収穫判定に果実熟度をランキング学習させて判定することを試みた。

学習は、果実画像2枚を1ペアとし、どちらの画像が熟しているかを人が判断した結果を教師データとした。例えば、入力画像A、Bがあり、熟度が画像A>画像Bの場合は判定結果が1、画像A<画像Bの場合は判定結果が0、判別が難しい場合は0.5とラベリングする。それらのペアを1000組準備し、ペアとなる画像A、Bを別々の特徴抽出ネットワークに入れ熟度を出力させ、その結果が異なる結果となった場合は誤差値を逆伝播（でんぱ）することで学習を進めた。なお、特徴抽出にはLeNet[7]を使用した。推定は、これらの学習済み特徴抽出ネットワークの片側のみを使用して行っている。この適収穫判定モデルを用いて果実熟度を判定した結果を第6図に示す。



第6図 熟度のランク結果  
Fig. 6 Ranking maturity results

第6図より、果実は緑から赤くなるにつれて熟度が高くなるよう判定されており、人の感覚ともよく合致した結果が得られた。従って、農園でロボットを稼働させる場合、農園作業者が事前に適収穫の閾（しきい）値を果実画像で選択すれば自動的にロボットが収穫対象を判別するシステムの構築が可能となった。また、例えば、収穫対象となるトマトの品種が変わり、熟度判定の色味が変わった場合でも、その品種の画像を用いて熟度判定の学習をすることが可能なことから、栽培している品種が異なる農園においても収穫ロボットの適用が容易になる。なお、本検討ではLeNetとは別に特徴抽出にGoogLeNet[8]を用いて精度比較を実施したが、両ネットワーク間で精度の優位差が見られなかったため、より軽量なLeNetを採用している。適収穫判定は色による判定となり、複雑なテクスチャを扱うモデルは必要ではないと考えられる。

## 6. まとめ

トマト農園においてロボットが果実を収穫するうえで必要となる果実検出および周囲環境を認識する技術を開発した。特定の画像処理手法では検出できない果実の検出や、主茎、果梗といった果実以外の環境についての詳細把握、また果実熟度の判定に対して、ディープラーニングによって学習したモデルが有効であることを示した。果実以外の環境把握については細い果梗などへの対応が今後の課題である。これらのモデルは、トマトに特化して開発を進めた画像処理手法と比較して汎用性が高く、例えば、果実検出モデルでは検出したい果実の画像を新たに学習すればそれを検出することが可能である。トマト以外のイチゴ、リンゴといった他の果実向け収穫ロボットへの寄与が期待される。

## 参考文献

- [1] 経済産業省, “2012年ロボット産業の市場動向,” [http://www.meti.go.jp/policy/mono\\_info\\_service/mono/robot/pdf/20130718002-3.pdf](http://www.meti.go.jp/policy/mono_info_service/mono/robot/pdf/20130718002-3.pdf), 参照 Apr. 20, 2018.
- [2] 野口伸, “車両系農業ロボットの現状と展望,” 建設の施工規格, no. 728, pp.58-61, Oct. 2010.
- [3] 野村融典, “未来農業を拓く農作業の自動化・ロボット化,” 豊穰, no. 29, pp.81-84, 1991.
- [4] K. Kim et al., “PVANET: Deep but Lightweight Neural Networks for Real-time Object Detection,” arXiv:1608.08021, Aug. 2016.
- [5] Kaiming He et al., “Deep Residual Learning for Image Recognition,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.770-778, 2016.
- [6] Zifeng Wu et al., “Wider or Deeper: Revisiting the ResNet Model for Visual Recognition,” arXiv:1611.10080, Nov. 2016.
- [7] Y. Lecun et al. “Gradient-based learning applied to document

recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp.2278-2324, 1998.

- [8] Christian Szegedy et al., “Going deeper with convolutions,” 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, Jun. 2015.

## 執筆者紹介



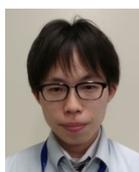
上垣 俊平 Shumpei Uegaki  
生産技術本部 ロボティクス推進室  
Robotics Promotion Office,  
Manufacturing Technology and Engineering Div.



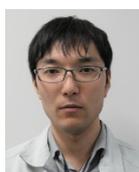
荒木 秀和 Hidekazu Araki  
生産技術本部 ロボティクス推進室  
Robotics Promotion Office,  
Manufacturing Technology and Engineering Div.



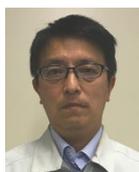
戸島 亮 Ryo Toshima  
生産技術本部 ロボティクス推進室  
Robotics Promotion Office,  
Manufacturing Technology and Engineering Div.



新崎 誠 Makoto Shinzaki  
コネクテッドソリューションズ社  
イノベーションセンター  
Innovation Center, Connected Solutions Company



上田 大介 Daisuke Ueta  
コネクテッドソリューションズ社  
イノベーションセンター  
Innovation Center, Connected Solutions Company



山崎 龍次 Ryuji Yamazaki  
コネクテッドソリューションズ社  
イノベーションセンター  
Innovation Center, Connected Solutions Company