

深層学習の利活用による植物表現型解析技術の展望

大倉史生^{1,2}・野下浩司^{1,3}・木下俊則^{4,5}・戸田陽介^{1,4,6}

¹JST さきがけ、〒332-0012 埼玉県川口市本町4丁目1-8

²大阪大学産業科学研究所複合知能メディア研究分野

〒567-0047 大阪府茨木市美穂が丘8-1

³九州大学理学研究院大学院理学研究院生物科学部門

〒819-0395 福岡県 福岡市西区 元岡774番地

⁴名古屋大学トランスフォーマティブ生命分子研究所

〒464-8602 愛知県名古屋市千種区不老町

⁵名古屋大学理学研究科 〒464-8601 愛知県名古屋市千種区不老町

⁶アグリエア株式会社 〒432-8011 静岡県浜松市中区城北3丁目5番1号

Fumio Okura^{1,2}, Koji Noshita^{1,3}, Toshinori Kinoshita^{4,5}, Yosuke Toda^{1,4,6}

Perspectives of Plant Phenotypic Analysis Utilizing Deep Learning

Keywords: Deep Learning, Machine Learning, Plant Phenotyping

¹Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

²Department of Intelligent Media, The Institute of Scientific and Industrial Research, Osaka

University, 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan

³Department of Biology, Faculty of Science, Kyushu University, 744 Motoooka, Nishi-ku,

Fukuoka 819-0396, Japan

⁴Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Chikusa-ku,

Nagoya 464-8602, Japan

⁵Division of Biological Science, Graduate School of Science, Nagoya University, Chikusa-ku

Nagoya 464-8602, Japan

⁶AgriHeir Co. Ltd, 3-5-1 Johoku Naka-ku Hamamatsu, Japan

DOI: 10.24480/bsj-review.10b7.00160

1. はじめに

近年、画像解析など様々な分野において深層学習の活用が進んでいる。深層学習のルーツであるニューラルネットワークは1940年代から研究が行われてきたが、様々な技術的困難から流行・衰退を繰り返してきた。2000年代後半以降、事前学習の有用性などのいくつかの技術的ブレイクスルーと計算資源の飛躍的進歩を背景に、大規模なニューラルネットワークの最適化が可能となったことから、様々な分野での応用が広がっている。特に、植物表現型解析の主要な技術要素である画像解析においては、畳み込み演算を行うフィルタの学習を伴う畳み込みニューラルネットワーク (convolutional neural network: CNN) の活用が進んでいる。CNNの初出は1980年代に遡る (LeCun et al. 1989; Fukushima and Miyake 1982) が、従来の (深

層学習を用いない)手法による画像分類と比較して精度を圧倒的に改善し CNN ブームの幕開けを告げたのは 2012 年に提案された AlexNet (Krizhevsky et al. 2012) である。コンピュータの計算能力の進歩と歩調を合わせるように、CNN の構造 (層の数や接続関係) も巨大化し続けている。AlexNet が 8 層のネットワークであったのに対し、2014 年に提案された GoogLeNet (Szegedy et al. 2015) は 22 層で分岐を含む構造をなし、翌年提案された ResNet (He et al. 2016) はループ構造を含む最大 152 層のネットワークを用いて人間による画像分類を超える精度を達成した。

大量の入出力データからネットワーク内のパラメータを学習することで、手作業での特徴設計なしに画像解析を実現できる手軽さを持つなどの利点から分野を問わず活用されつつある。植物表現型解析の分野においても、植物の葉の画像から 14 種の疾病を検出する手法が 99% の識別精度を達成する (Mohanty et al. 2016) など、深層学習を用いた植物表現型の解析への応用は数多く報告されている。しかし、層の増加や構造の複雑化は、CNN による判断根拠の理解を難しくしている。深層学習がある種の「ブラックボックス」である¹といわれ、特に基礎科学の分野において深層学習の活用を敬遠する一因となっている。

有効活用するにせよ、あえて使わないにせよ、多くの分野で深層学習と関わりが避けられなくなっている。その意味で、深層学習が本分野に何をもたらしたのか、その到達点を知っておくことは重要であろう。本稿では、特に画像を用いた植物表現型解析における応用事例を紹介し、深層学習の植物科学における応用可能性の一端を明らかにするとともに、その技術的限界と今後の展望を考察する。

2. 気孔開度定量自動化による表現型スクリーニングの実現

植物の気孔開口制御は個体の成長と環境ストレス応答において重要な役割を担う。光条件下において、植物は気孔を開口させることによって光合成に必要な二酸化炭素を取り込むと同時に、蒸散を通じて地下部から地上部への栄養転流を促進する。一方、乾燥条件下では気孔を閉鎖させることによって水分損失を抑制する。気孔開閉制御の分子機構を明らかにすると同時に、植物の光合成活性および乾燥ストレス耐性の人為制御を目指し、著者らは気孔開口を制御する化合物の探索およびその分子機能の解析を進めてきた (Toh et al. 2018)。前報においては、一般に表現型スクリーニングと呼ばれる手法を採用した。具体的には、候補化合物を葉に処理した後、一定時間後気孔の開口度 (気孔開度) を測定することによって化合物の生理活性を評価するという方法である。しかしながら、化合物処理した植物の葉の表面を顕微鏡で 1 枚 1 枚撮影し、気孔開度を手作業で測定せざるを得なかったため、大量の化合物評価は困難であるという問題点があった。

¹ これは程度問題であり、従来の機械学習も似たような意味で「ブラックボックス」である。

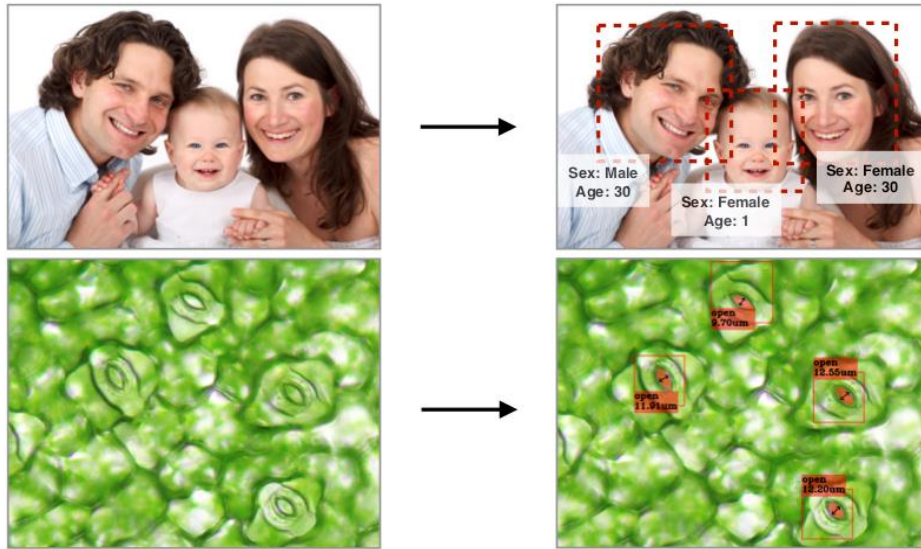


図1 顔認証技術をもとにした気孔開度測定技術²

そこで筆者らは、画像解析による気孔開度測定を自動化する手法を確立した。本節の詳細については (Toda et al. 2018) を参照されたい。本手法は、深層学習を活用した顔認識技術 (Taigman et al. 2014) などをベースにしている。顔認識技術の一例においては、まずは事前定義された特徴量を利用する機械学習モデルを用いて画像内における顔の場所を検出し、次にその顔から「年齢」や「表情」といった情報を深層学習によって識別する (図1 上段)。それと同様に本手法では、気孔の場所を特定した後、各々の開閉状態を識別する。閉鎖していると識別された気孔の開度を 0 と一律に扱う一方、開口していると識別された気孔を含む領域のみ、画像二値化を経た定量的計測プロセスが実行される (図1 下段)。

本技術の導入により、これまで手作業に依存していた定量解析の自動化が可能となったと同時に、「何割の気孔が開口していた、または閉鎖していた」といった、定性的な人の評価基準を模倣させることが可能となった。特筆すべきは、解析速度が手動に比べ 80 倍向上したことであり、日常的な定性・定量的表現型スクリーニングの実行が大幅に易化した。このように、他分野における深層学習を活用した技術を基にして、植物画像解析のモジュールに利用することで、人手が必要な解析作業を自動化が実現する一例として紹介した。

3. 深層学習と三次元復元による枝の三次元構造推定

詳細な植物表現型解析、特に枝の一本一本レベルで植物の三次元形状・構造を解析することが可能になれば、分枝過程の定量化や剪定の最適化など、育種や栽培管理における新たな応用が見えてくる。しかし、植物は細い枝、同色の葉、遮蔽の多さなどから、三次元形状の復元が非常に難しい対象である。特に、枝構造を再現するには、葉などによる枝の遮蔽を適切

² 本図で使用されている人物像にはクリエイティブ・コモンズ 1.0 ライセンスが付与されており、<http://publicdomainpictures.net> から援用した。なお、各顔に割り振られた性別や年齢は、本来ならば深層学習による推論で得られる値であるが、ここでは筆者が定めたランダムな値で意味はない。

に補完する必要がある。いかにして「見えない部分」を推定するかは、植物科学分野のみならず、情報科学、特に画像解析・コンピュータビジョン分野においても非常に挑戦的な課題である。

提案手法は、複数の視点から撮影された植物画像（多視点植物画像）を入力とした三次元枝構造の推定手法である。深層学習による画像変換（image-to-image translation）と呼ばれる手法を用いて、各画像で葉などに隠された枝の存在確率を推定し、推定された枝確率を用いて三次元復元することにより、見えない枝も含めた枝の構造を再構成する。なお、本節で述べる手法の詳細については (Isokane et al. 2018) を参照されたい。

3-1. 画像変換：画像をつくる深層学習

本手法で活用するのは、画像変換と呼ばれる技術である。画像変換は、入力画像に含まれる物体情報などを保持したまま、その「ドメイン」を変換する（例えば線画から写真など）手法である。画像変換に類する技術は深層学習の登場以前から活発な研究が行われてきた (Efros and Freeman 2001; Darabi et al. 2012)。画像のテクスチャに着目した変換や色、スタイルの変換がその例であり、画像の矩形小領域（パッチ）ごとに変換画像を生成し合成する手法が多く試みられ、Adobe Photoshop 等にも画像修復の手法として搭載されている。

他方、本研究で用いるのは Pix2Pix (Isola et al. 2017) と呼ばれる深層学習を利用した画像変換手法である。本手法は、conditional generative adversarial networks (cGAN) と呼ばれるネットワークを活用する。入出力画像の正解ペア群を事前に学習データとして与えると、生成器が画像生成を行い、生成された画像と入力画像が正しいペアであるかどうか判別する CNN である識別器と「競い合う」ことにより学習が進む。識別器は、一般的な CNN が画像全体を畳み込むのとは異なり、画像のパッチを入力とする CNN を用い、与えられた入力・出力パッチペアが本物かどうかを判定する。この方法は、「深層学習以前」の画像変換が活用していた考え方に近い手法である。深層学習はあくまで優秀な最適化器であり、良いシステムの構築には「深層学習以前」の知識の蓄積も重要であることを示す例として興味深い。

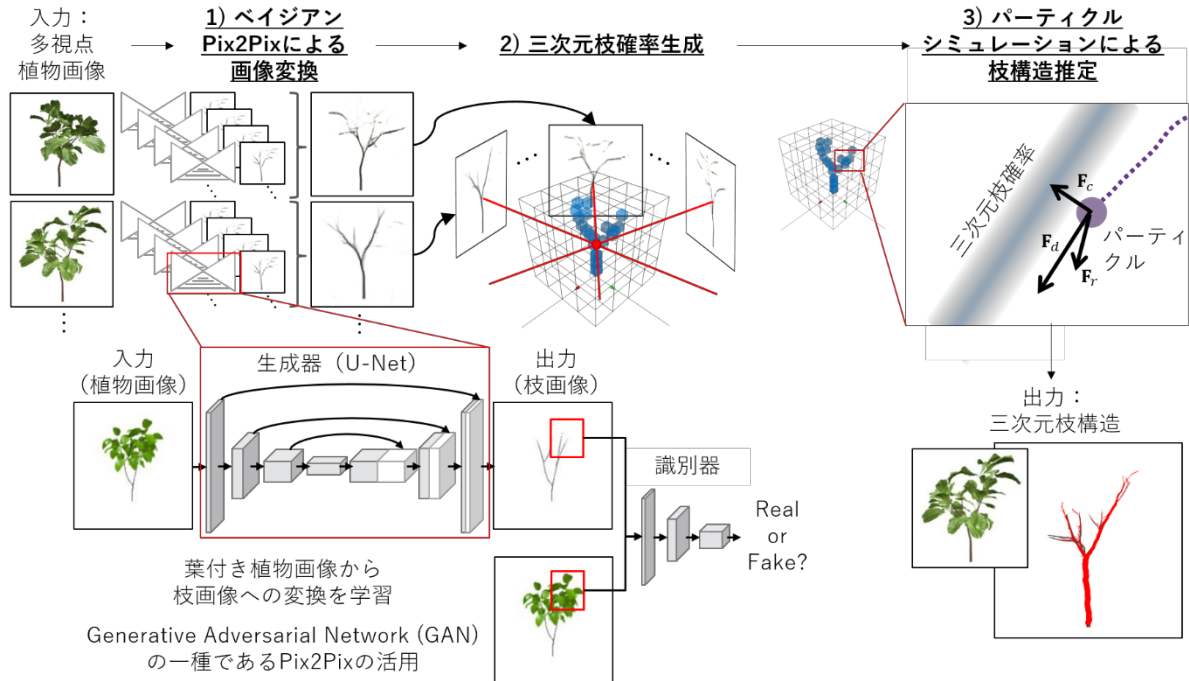


図2 深層学習による画像変換を用いた枝の三次元構造復元 (Isokane et al. 2018)

3-2. 枝の三次元構造復元への画像変換の活用

では、いかにして画像変換を枝の三次元構造復元に活用するのか。手法の流れを図2に示す。本手法は、複数の視点から撮影された植物画像（多視点植物画像）を入力とし、葉に隠れた部分も含め、枝の三次元構造を復元する。そのために、各画像上で深層学習を用いた画像変換 (Isola et al. 2017) を使い、葉付きの植物画像から枝の存在確率を推定する。各画像上での枝の存在確率は、従来の多視点画像を用いた三次元復元 (structure from motion: SfM (Wu 2013)) によって推定されたカメラ位置・姿勢を用いて3次元空間上における枝存在確率に変換できる。本研究では、この三次元枝確率に基づき、従来、CG分野で研究されてきた樹木の自動モデリングと同様の枠組みで枝構造を推定する。

提案手法の精度および有効性を検証するための三次元復元実験を行った。画像変換ネットワークの学習のために、CG植物生成ツール (Palubicki et al. 2009) を用いて10本のCG植物を作成し、葉付き・枝のみの状態で各72視点からの撮影を行った。ネットワークの学習に用いた画像とは異なるCG画像を入力として枝形状の推定を行うと、提案手法の枝位置の誤差は、植物の高さに対しおよそ1%~2%であり、良好な精度を示した。また、復元に用いる視点数を減らした場合や、葉の色・形が異なる植物種を入力画像とした場合にも有効であった。

本節では、「見えない枝」を復元する植物の三次元枝構造復元手法を紹介してきた。見た目・形状・構造が複雑（葉による遮蔽、枝の細さ、構造のバリエーションの多さ）な植物は、植物科学のみならず、情報科学分野においても非常に挑戦的かつ興味深い研究対象である。詳細な構造推定手法の開発は未だ途上であり、初期評価の段階である。今後は、植物表現型解析の実応用に即した手法として改良することが重要である。

4. 深層学習の病害虫診断への応用と判断根拠の可視化

前節までに述べたように、深層学習の活用は、植物表現型解析の分野において「これまでできなかった」応用を実現し、画像解析手法の可能性を大幅に広げた。一方で、深層学習の広まりは、植物科学も含めた基礎科学分野において「これまで簡単にできたことができない」状況を生み出している。「深層学習はブラックボックスであり、判断根拠が人の目で判断できない」という問題であり、これは基礎科学において致命的な問題となり得る。

他方、深層学習の「ブラックボックス」の中身を明らかにしようとする研究も活発に行われている。特に、画像解析に広く用いられる CNN の計算過程を、人が理解できる形に可視化する手法が多く提案されている (Selvaraju et al. 2017; Ghosal et al. 2018)。しかし、多くの手法は ImageNet データセット (Russakovsky et al. 2015) (一般的な物体を集めた 1000 カテゴリからなる画像データセット) を実験対象として開発されており、特定の物体 (植物) を対象として適用した場合にどのような特性を示すのかは明らかにされていない。そこで本節では、葉の病害虫診断を例として、CNN 可視化手法の特性および活用可能性を議論する。

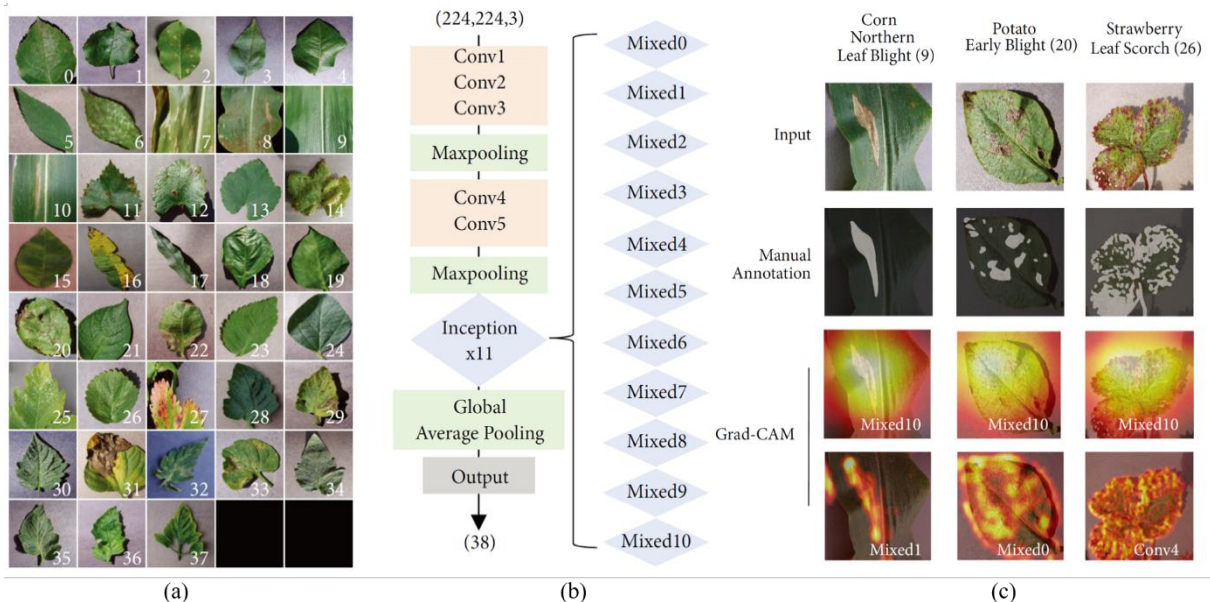


図3 CNNによる葉の病害虫診断と判断根拠の可視化 (Toda and Okura 2019)

本節では、InceptionV3 (Szegedy et al. 2016) と呼ばれる CNN を用いた葉の病害虫診断ネットワークを学習し可視化する。学習には、葉の画像データセットである PlantVillage (Hughes and Salathé 2015) を用いる。本データセットは、各画像に健康および様々な疾病に対応する 38 のラベル (図 3(a)参照) が付与されている。InceptionV3 は、図 3(b)に示すようなネットワークであり、各 Inception モジュール (Mixed) は分岐構造を含む畳み込み層である。

図 3(c)に、CNN の可視化手法でも特に広く用いられる、ヒートマップによる可視化手法の一つである Grad-CAM (Selvaraju et al. 2017) による可視化結果を示す。他の手法群も含めた網羅的な比較は (Toda and Okura 2019) を参照されたい。Grad-CAM は、各ニューロンの重要度を可視化する手法の一つであり、CNN の畳み込み層に適用した場合、画像中の位置に対応づいた重要度のヒートマップとして可視化ができる。図 3(c)より、Grad-CAM による可視化は

比較的浅い層 (Conv4 から Mixed1) で実際の病害領域に対応した領域がハイライトされていることがわかる。深い層 (Mixed10) においては、畳み込みの繰り返しにより解像度が低下し、詳細な位置の特定は難しい。

一般に、ImageNet データセット等の広く様々な物体カテゴリを含むようなデータセットを用いた場合の Grad-CAM の可視化能力は浅い層で下がる (つまり、浅い層では画像の判別に有効な高レベル特徴が抽出できていない) ことが知られているが、従来の知見とは異なり、葉の病害虫画像を対象とすると、浅い層での特徴抽出が非常に有効であった。これは、葉の病害虫などの見た目が色の明暗やテクスチャなどの低レベル画像特徴によく現れており、比較的少ない回数の畳み込みで抽出できたためであると考えられる。

本可視化結果を用いると、ネットワークの浅い層のみを用いて再学習し、同程度の分類精度を維持したままパラメータ数 (学習にかかる時間・メモリの量に関わる) を削減することができる。本実験条件においては、Mixed6 以降の畳み込みを削除した場合、精度を下げないまま 75% のパラメータを削減できた。従来、深層学習の研究においてパラメータ削減を自動で行う方法は数多く提案されている (Han et al. 2015) が、可視化に基づく直感的な削減手法は、植物科学等の専門的知見をネットワークの設計に活用し得る、新たな指針を与えるかもしれない。

5. おわりに

本稿では、深層学習の植物表現型解析、特に植物画像解析への応用例を紹介してきた。各節で述べたように、情報科学においてこれまで広く扱われてきた一般的な対象と比較して、植物は特異な部分が多い。第 3 節で述べたように、植物の形状・テクスチャ・遮蔽の特性は情報科学分野においても非常に挑戦的である。また、第 4 節における葉の病害診断においては、病害虫の見た目の特性が異なることにより、一般的な物体認識タスクと深層学習による判断過程に相違が見えてきた。

深層学習ライブラリのコモディティ化や、可視化技術の進展により、植物科学分野においても深層学習の参入障壁は飛躍的に低くなってきている。しかし、各タスクに機械学習を活用するにあたり、どのようなモデルを選択するのか、そして、出てきた結果をどのように解釈するのかは、深層学習登場以前と比較し重要になっている。例えば、本稿で取り上げた CNN や GAN 以外にも文章や動画といった系列データを扱う再帰型ニューラルネットワーク (RNN) や 試行錯誤から自律的に学習することを目指す強化学習に用いられる Deep Q-Network (DQN) など異なるネットワーク構造も提案されている。目的に応じて適切なネットワークや学習アルゴリズム (場合によってはハードウェアも) を選択する必要がある、そもそも深層学習的な手法が必要かも検討の余地はあるだろう。また、深層学習は優秀な最適化器にすぎず、大量の学習データを与え当該タスクで高い精度を達成したとしても、その結果を解釈するのは (可視化手法などの手助けを借りるにしても) 人類である。人工知能が人間の仕事を奪う、などと盛んに叫ばれているが、物事の解釈こそが重要になる基礎科学分野においては、深層学習の登場によって、それを解釈する役目—人類が持つ専門的知識の重要性—は以前より高まっているともいえる。植物科学者が、強力な深層学習の力を借りつつ、その過程や結果と

自身の専門性を組み合わせ、新たな知見を創出することが、植物科学における新たなワークフローの一つとなるかもしれない。

2019年3月、日本における植物表現型解析研究の進展を加速するため、本稿著者らを含む有志が立ち上げ人となり、2019年3月に日本植物フェノタイピングネットワーク (Japanese Plant Phenotyping Network: JPPN)³を設立した。植物表現型解析は、植物科学において特に情報科学や画像解析が応用されやすい分野である。今後は植物表現型解析に関する情報交換やワークショップ、チュートリアルなどを行なっていく予定であるが、実際に解析をおこなっている（もしくは困難に直面している）植物学会員の皆様からの提案・フィードバックは、こうしたコミュニティの健全な発展に有益であろう。

謝辞

本総説で紹介した研究は、JST さきがけ「情報科学との協働による革新的な農産物栽培手法を実現するための技術基盤の創出」JPMJPR17O5 (戸田), JPMJPR17O3 (大倉), JPMJPR16O5 (野下) および、文部科学省 科学研究費補助金 新学術領域研究 15H05956 (木下), JSTALCA JPMJAL1011 (木下) の支援を得て遂行した。

引用文献

- Darabi, S., Shechtman, E., Barnes, C., Goldman, D.B. & Sen, P. 2012. Image melding: Combining inconsistent images using patch-based synthesis. *ACM Trans. Graph.* 31(4): Article No. 82.
- Efros, A.A. & Freeman, W.T. 2001. Image quilting for texture synthesis and transfer. In: *Proc. SIGGRAPH2001*. pp. 341–346. ACM.
- Fukushima, K. & Miyake, S. 1982. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In: Amari, S. & Arbib, M. A. (eds.) *Competition and Cooperation in Neural Nets*. pp. 267–285. Springer, Berlin Heidelberg.
- Ghosal, S., Blystone, D., Singh, A.K., Ganapathysubramanian, B., Singh, A. & Sarkar, S. 2018. An explainable deep machine vision framework for plant stress phenotyping. *Proc. Natl. Acad. Sci. U.S.A.* 115(18): 4613–4618.
- Han, S., Pool, J., Tran, J. & Dally, W. 2015. Learning both weights and connections for efficient neural network. In: *Proc. 2015 Conf. Neural Information Processing Systems (NIPS)*. pp. 1135–1143.
- He, K., Zhang, X., Ren, S. & Sun, J. 2016. Deep residual learning for image recognition. In: *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*. pp. 770–778. IEEE.
- Hughes, D. & Salathé, M. 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*.
- Isokane, T., Okura, F., Ide, A., Matsushita, Y. & Yagi, Y. 2018. Probabilistic plant modeling via multi-view image-to-image translation. In: *Proc. 2018 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*. pp. 2906–2915. IEEE.
- Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A.A. 2017. Image-to-Image Translation with Conditional

³ <https://www.plant-phenotyping.jp/>

F. Okura, K. Noshita, T. Kinoshita & Y. Toda-8

- Adversarial Networks. In: *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*. pp. 5967–5976. IEEE.
- Krizhevsky, A., Sutskever, I. & Hinton, G.E. 2012. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60(6): 84–90.
- LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W. & Jackel, L.D. 1989. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* 1(4): 541–551.
- Mohanty, S.P., Hughes, D.P. & Salathé, M. 2016. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 7:1419.
- Palubicki, W., Horel, K., Longay, S., Runions, A., Lane, B., Měch, R. & Prusinkiewicz, P. 2009. Self-organizing tree models for image synthesis. *ACM Trans. Graph.* 28(3): Article No. 58.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C. & Fei-Fei, L. 2015. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vision* 115(3): 211–252.
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. & Batra, D. 2017. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: *Proc. 2017 IEEE Int. Conf. Computer Vision (ICCV)*. pp. 618–626. IEEE.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. 2016. Rethinking the inception architecture for computer vision. In: *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*. pp. 2818–2826. IEEE.
- Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. 2015. Going deeper with convolutions. In: *Proc. 2015 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*. pp. 1–9. IEEE.
- Taigman, Y., Yang, M., Ranzato, M. & Wolf, L. 2014. DeepFace: Closing the gap to human-level performance in face verification. In: *Proc. 2014 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*. pp. 1701–1708. IEEE.
- Toda, Y. & Okura, F. 2019. How convolutional neural networks diagnose plant disease. *Plant Phenomics* Article ID 9237136, 14 pages.
- Toda, Y., Toh, S., Bourdais, G., Robatzek, S., Maclean, D. & Kinoshita, T. 2018. DeepStomata: Facial recognition technology for automated stomatal aperture measurement. *BioRxiv* 10.1101/365098.
- Toh, S., Inoue, S., Toda, Y., Yuki, T., Suzuki, K., Hamamoto, S., Fukatsu, K., Aoki, S., Uchida, M., Asai, E., Uozumi, N., Sato, A. & Kinoshita, T. 2018. Identification and characterization of compounds that affect stomatal movements. *Plant Cell Physiol.* 59(8): 1568–1580.
- Wu, C. 2013. Towards linear-time incremental structure from motion. In: *Proc. 2013 Int. Conf. 3D Vision (3DV)*. pp. 127–134.