

実践的な植物病自動診断のための画像生成技術

彌富 仁*

Image Generation Techniques for Practical Plant Disease Diagnosis

Hitoshi Iyatomi*

In recent years, data “generation” techniques using generative adversarial networks (GANs), in which multiple neural networks learn in an adversarial manner, have achieved remarkable results in many fields. The authors are participating in a project commissioned by the Ministry of Agriculture, Forestry and Fisheries (MAFF) called the “Artificial Intelligence Future Agriculture Creation Project” and are working on the development of automatic diagnosis of plants based on images and related technologies. In this paper, we introduce the trend of automatic diagnosis technology for plant diseases, the unique difficulties of this task, and the remarkable effect of GAN’s image generation technology on it.

Keywords: Generative adversarial networks (GAN), Image generation, Automated plant diagnosis, Deep learning.

1. 導入：深層学習技術の登場と画像認識

1.1 深層学習登場まで

2010年代初頭まで画像認識技術は、対象となる画像から(1)解析対象となる注目領域 (ROI: region of interest) の抽出や色補正といった前処理, (2) 識別のために重要な特徴量の設計と抽出, (3) Support vector machine (SVM) や, 人工的な神経回路網である Neural networks (NN), Random forest などの識別や回帰アルゴリズムによる識別の3段階からなっていた。これらの手法は限られた範囲の問題には、優れた性能を発揮できていたが、大規模な問題を扱う場合、特に(2)において致命的な問題があった。また(2)を実現するために必要となる(1)も、限定的な制約のある状況下では上手に働くが、そうでない場合には実現が困難であった。

1.2 深層学習技術の登場

2012年、ILSVRC (ImageNet Large Scale Visual Recognition Challenge) という1000種類のカテゴリの画像からなるデータセットの画像認識能を競う大会で、著者の名前にちなんで後に AlexNet[1]と呼ばれる実践的な構造の畳み込みニューラルネットワーク (Convolutional neural networks; CNN) モデルが、top5 error 率 (上位5つに正解が入っていれば正解) で15.8%を実現し、2位 (同26.8%) 以下に大差をつけて優勝を果たした。

CNNは現在の画像認識技術で中心を担う技術であり、その最大の特徴は識別対象から、識別に必要な特徴を自ら獲得して学習に用いることができるという点である。前述の古典的な手法で非常に困難であった(1)と(2)の手順がCNNでは(難しくない問題に対しては)必要ではなくなり、CNNでは各画像に対して、「りんご」「みかん」などといった答えの情報とともに学習を行うことで、高精度で画像を分類できる能力を獲得できる。CNNはその能力のおかげで、きわめて急激に広く人工知能分野の発展を後押しした。

深層学習以前のほとんどのneural networkモデルは、3層構造であった(それ以上深くするとうまくいかなかった)が、このAlexNetは、畳み込み層5層全結合層3層を含む全8層構造であり、それまでと比較して深い構造のものであった。以後、2014年に同競技で優勝したVGGモデルは16ないし19層 (top5 error = 7.4%)、2015年のResNetは152層 (同3.6%)と、どんどんと深く精度の高いモデルが提案された。現在では、問題に応じて適切なネットワークの構成を考慮した計算効率のよい(かつ深い)モデルが提案され広く用いられている[2]。こうした技術は、多層の深い人工ニューラルネットワークを用いる学習技術であるため、“deep learning”(深層学習)と呼ばれるようになった。

一般的な物体を対象とした認識問題では数年前より人間の識別能力をすでに超えており、身近なところではスマートフォンで撮影された写真の人物や内容などが自動的に認識され、登録者は便利に分類や検索できるようになっている。

1.3 物体検知+識別モデルの開発

CNNの発達に伴い、画像中に複数の種類の物体が複数

2022年1月25日受付

法政大学理工学部

(〒184-8584 東京都小金井市梶野町 3-7-2)

Hosei University, Faculty of Science and Engineering
(3-7-2 Kajino-cho, Koganei, Tokyo 184-8584, Japan)

* Corresponding Author iyatomi@hosei.ac.jp

存在するような状況で、それらを検知し、かつその内容も判定するような物体検知モデルが2013年頃より多く提案されるようになってきた。当初のモデルは、注目領域の特定に時間がかかっていたが、こちらも高速、高精度なモデルが次々提案されており[3-5]、現在では多くの問題において実時間で十分な精度が実現できている。これらのモデルでは物体が存在する候補領域を事前に無数用意しておき、それぞれに対して識別し、位置的なずれと、その中身の内容の推定を行っている。身近なところでは、自動車の自動運転や、建物の入り口で顔を追跡して体温を検知するシステムなどで用いられており、すでに多くの分野で実用化されている。

2. 敵対的生成ネットワークとその発展

こうした深層学習技術の研究が進む中、2014年に深層学習技術を大きく加速する Generative adversarial networks (GAN) と呼ばれる、画像を“生成”する技術が発表された[6]。GANは、Generator (G:生成器)と Discriminator (D:識別器)の2つの学習器から構成される。観測されている(学習用)画像はある未知の確率分布からサンプリングされた結果として観察されるものと考え、Gは最尤推定の考え方にに基づき、その確率分布を獲得するように学習を行い画像を生成する。一方Dは入力された画像が、本物かGが生成した偽物の画像かを見抜くように設計、学習される。このGとDがお互いにだますように学習が行われるため、“敵対的”生成ネットワークと呼ばれ、Gは泥棒、Dは警察官に例えられることが多い。これまでの機械学習技術で見た目には自然な画像の生成は、きわめて難しい問題であり事実上実現できていなかったが、このGANは現実的な計算時間で成し遂げる画期的な成果となった。

GANの研究は深層学習技術の中できわめて盛んに行われ、生成できる画像の質は短時間で劇的に向上した。2021年現在、GANはもはや人間には本物と区別がつかないほど高精細な新しい画像を「生成」することが可能であり[7]、生成過程でDをだますように生成された画像であるため、コンピュータにとっても見抜くことは容易でない。

GANの高精度な画像生成能力は、「領域の検出」(pix2pix[8])、異なる画像特徴を持つ画像への「スタイル変換」(CycleGAN[9])、未知の高い解像度成分を推測し高精細画像を生成する「超解像」(ESRGAN[10])などさまざまな用途で利用されている。本稿では、こうしたGANによる画像生成技術を、植物病自動診断問題に適用した実施例を紹介する。

3. 画像を用いた植物病の自動診断技術

3.1 自動診断技術研究の概要と初期の研究

植物病虫害の診断には専門家の知識が必要であり、簡単で正確、低コストな植物病害診断システムの構築が求められている。深層学習技術をベースに、実際の圃場で撮影された画像を元に開発された最初のシステムの1つ

に、著者らが埼玉県協力を元にしたキュウリの自動診断システムがある[11]。このシステムは、今では当たり前になった深層学習技術を用いた初めての自動診断システムであり、800枚のキュウリの葉の圃場画像の3クラス(2種のウイルス病と健全)分類問題に対して94.9%の平均識別精度を実現した。

大規模な植物画像の open dataset である PlantVillage dataset が web で公開されて以降、深層学習を用いた自動診断の研究は急増した。Mohanty ら[12]は、このデータセットに含まれる14種の作物、病気と組み合わせ38ペア計54,306枚の大規模な葉の画像データをCNNで学習することで、平均99%以上の識別能を報告している。以降、同データセットを用いたいくつかの研究で、90~99%程度のきわめて良好な識別結果が報告されてきた。しかしこれらの研究で用いられた学習および評価用の画像データは、解析対象となる葉が事前に切り取られ、無地の背景に置かれた状態の画像であり、実際の圃場の写真とは大きく外観が異なる。Mohanty らは、自身の論文の中で、彼らのモデルが実践的環境で約31%まで精度が大幅に低下すること、同様のほかの研究でも学習とは異なるタイプの画像に対しては精度が99.5%から約33%まで低下したことの報告[13]があり、PlantVillageのような実験室環境の画像を用いて構築されたモデルは、実用面で深刻な課題を有している。また、実際の圃場で撮影された実際の植物の画像を多数学習したモデルであっても、ほかの圃場で撮影された画像で評価を行うと、一般的には大幅に精度が低下してしまう。

3.2 植物病自動診断における主要な問題点

前述のように深層学習を活用することによりきわめて高い診断精度が報告されてきたが、実践的な環境の下で評価した時に得られる本質的な診断能がこの数字に遠くおよばないことが多い。画像を用いた植物病自動診断技術のもっとも大きい問題は「過学習」と呼ばれる、識別器が学習データに対する過度な適合により汎化性(違う環境に対する適用性)が失われてしまうことにある。この過学習は、一般的な機械学習技術でも問題になるが、植物病を対象とする場合、特に影響が大きい。以下、植物病自動診断における問題点について簡単に紹介する。

3.2.1 診断そのものの困難さ

植物病害の診断は、一般物体認識問題(たとえば車と人、道路、植栽、標識など)に比べて、識別したいカテゴリ(病害)を表す特徴が多くの場合小さく曖昧である場合が多く、また同じ病害であってもその見た目の多様性がきわめて大きい。このため、学習モデルが画像の背景や、構図、明るさ、色合いなどといった、捉えるべき病徴とは無関係なものに適合し学習してしまう問題が無視できない。こうした過学習の問題は、多様な学習データを無数に用意することで理論上、経験上解決に向かうが、その数は膨大であり現実的ではない。

著者らの経験では、別する分類のクラス数にもよるが、数千~数万枚程度の画像を、自動認識分野で新しく実績のあるCNNモデル(たとえば2019年の最高記録更新モ

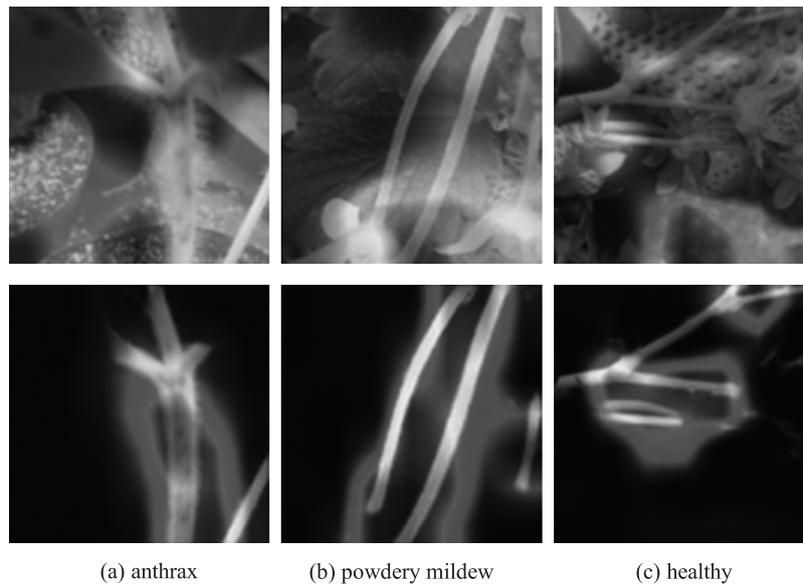


Fig. 1 Examples of diseases that occur on strawberry vines (top row), the results of extracting vine regions using a GAN-based region extraction method[16] (bottom row), and the results of visualizing the diagnostic basis for each. From left to right: anthracnose, powdery mildew, and healthy case. The discriminator focuses on the red region to make a diagnosis

デル EfficientNet[14]) に、性能向上の効果が強く近年広く用いられる data augmentation 手法である RandAugment [15]を組み合わせ、fine-tuning をした識別器の本質的な診断性能は、病害ごとに 30%~85%程度にとどまることが多い。

また定点観測カメラなどの利用を想定した広範囲の自動診断を行う際、カメラから解析対象(例:葉や実)までの距離が遠い場合、光学的な制約から十分な解像度が得られず十分な診断精度が得られないことも問題となる。

3.2.2 教師データ入手の困難さ

植物病の診断器構築のためには正しい病害ラベルのついた大量の画像データの収集が必要であるが、病害の特定は困難な場合も多くきわめてコストが高い。また健全症例は入手しやすい一方、発症例が少なく入手しにくい病害種もある。学習データ、あるいは評価用データの数の偏りは、一般的に精度の低下を招くため、理想的には各クラスとも多量かつ同程度のデータが準備できることが望ましいが、学習用データと評価用データが異なる環境(圃場)で撮影されていることを担保したうえで、これを実現することはきわめて難しい。

また、広域診断において物体検知モデルを活用する場合、学習データの準備には、注目する領域ごとにラベルを付与する必要があるが、これも大変コストが高い。

4. 画像生成技術を用いた実用的な対応策

ここでは、上記の課題について GAN をベースとしたいくつかの対応策について紹介する。

4.1 診断注目領域の抽出

注目する領域の抽出技術は、その多様性からこれまできわめて難易度の高い問題であったが、GAN により劇的

に進化した。機械学習技術そのものの進歩および、学習データの解像度化などにより、病害識別器の背景などへの過学習は一定量抑えることが可能となったが、病害がきわめて小さい領域に発生する場合(茎など)や、虫害の診断場合、注目領域の事前抽出は必須となる。

Fig. 1 は、いちごの茎部(葉柄、果柄部など)に発生する病徴がきわめて小さい病害に対して、CNN 識別器が診断の根拠とした部分を(強)赤~黄~緑~青(弱)で表している。上段は画像をそのままの状態に CNN 識別器に入力しており、下段は我々が開発した GAN ベースの領域抽出手法[16]を適用し茎部を事前抽出してから学習ならびに診断を行った例である。上段では、過学習の影響を受け識別器は病変以外の部位に着目しており、その結果 3 種の病変識別において 6 割弱の平均精度しか得られない。一方下段のように、茎領域を特定できると背景への不適切な過学習が抑制され 8 割程度の実用に耐える識別精度を実現できるようになった。学習画像の数が限られている中(数千枚程度)で小さい病徴を捉える必要がある場合、注目領域の事前抽出の効果は小さくない。

4.2 超解像技術

超解像技術は、失われてしまっている高精細な情報を何らかの方法で実現する技術で、これまでは複数の画像を用いて理論に基づく超解像が主流であった。GAN を利用した超解像技術は、無数の学習データから低解像度画像と高解像度画像の関係を学ぶことで、未知の画像に対しても高精細な画像を生成する。

病害の高精度の識別には、一定の画像の解像度が必要であるが、常に必要な解像度の学習画像が入手できるとは限らない。特に物体検出のモデルを用いて広域の画像から診断を行う場合、カメラから遠い領域の画像の解像

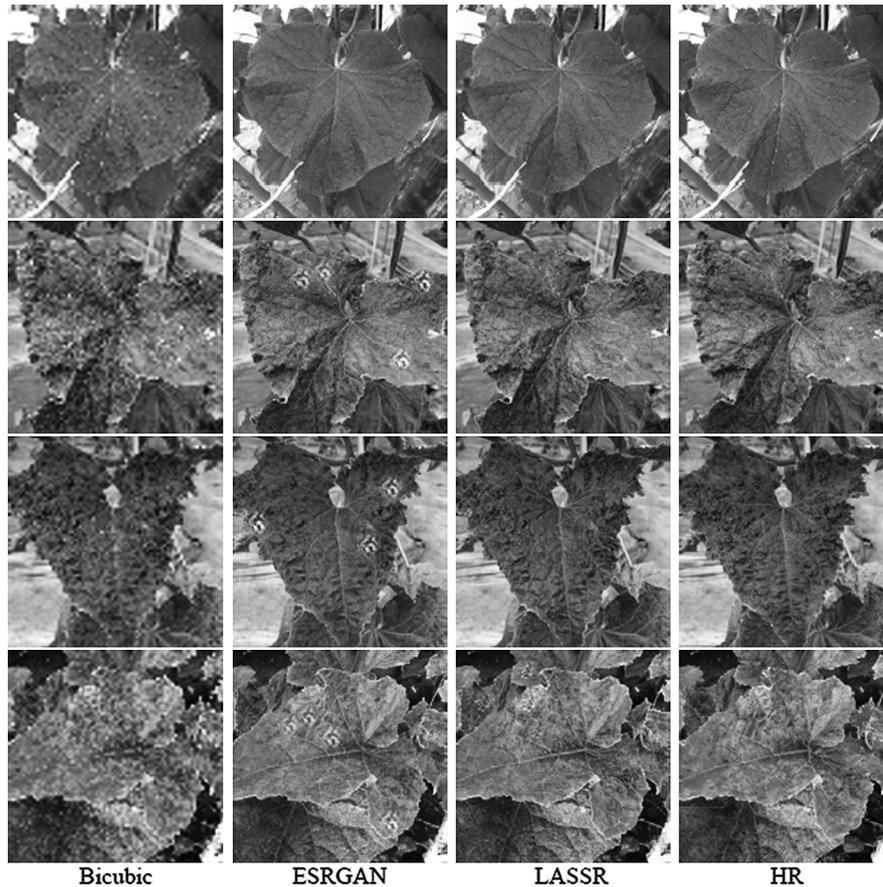


Fig. 2 Example of super-resolution of cucumber leaves. From left to right: low-resolution image (Bicubic interpolation), ESRGAN[10], LASSR (proposed by the authors)[17], (unknown) high-resolution image (HR)

度の低下は問題となる。GANを用いた従来の最高峰の手法の超解像手法であるESRGAN[10]は、優れた超解像画像を提供するが、植物病の画像に適用する場合、斑点状のartifactが発生してしまう場合がある。近年我々が提案したLASSR[17]はESRGANをベースにした超解像手法で、こうした欠点を克服し良質な超解像画像を生成できる。Fig. 2に低解像度画像をESRGAN, LASSRでそれぞれ4×4倍に拡大した画像例を示す。(注：この図では超解像の妥当性を見るために、未学習の高解像度画像(HR)を1/4にダウンサンプルした低解像画像(LR)を作成し、それに対して超解像処理を施し、元の高解像(HR)画像と比較をしている。LR画像はbi-cubic補間により大きさをそろえてある。)

LASSRを用いてきゅうりの葉画像の解像度を元の128×128から512×512に超解像処理をし、EfficientNet識別器にて6病害の識別器を構築した場合、未知の圃場の画像の平均識別率は64.9%から86.0%まで向上した。この数字は最先端手法であるESRGANを用いた場合よりも約2~2.5%高い値であり、提案する超解像技術の有効性ならびに、学習用画像の高解像度化の重要性が確認できる。

4.3 学習用画像生成

前述のように信頼できる病害ラベル情報のついた学習データの準備はたやすいことではない。一方、学習デー

タの不足や、カテゴリ間の偏りは診断精度に悪影響をおよぼす。我々は、不足する学習データを補うため、スタイル変換技術を応用した病害画像生成技術を構築した。Fig. 3に生成画像の例を示す。従来のGANに基づくスタイル変換手法(CycleGAN[9])は、注目していない背景部分などに意図しない変化が起こり不自然な画像が生成されることが多い。一方我々のLeafGAN[18]は、注目する葉領域のみに病徴を付与する独自の学習機構を持たせることで、自然な病害画像を生成することができるため、生成された画像を識別器の学習用画像として活用、つまり効果的なdata augmentationの一種として活用することができる。既存のCycleGANで病害画像を生成し、識別器の学習画像を増加させた場合には、診断精度は0.7%しか向上しなかった(71.3→72.0%)のに対し、LeafGANは、平均7.4%の性能向上が得られた(71.3→78.7%)。これはLeafGANが学習により獲得する画像生成のための確率分布が、元の学習データとは異なるように学習され獲得されるためであり、識別器から見ると新しい画像データの学習ととらえられるようになる。

この手法のさらに発展版となる我々のPPIG(Productive and Pathogenic Image Generation)[19]では、Fig. 4のように無数の健全画像を生成したうえで、それらに病徴を付与することで、さらに多様な画像を生成できる。この図

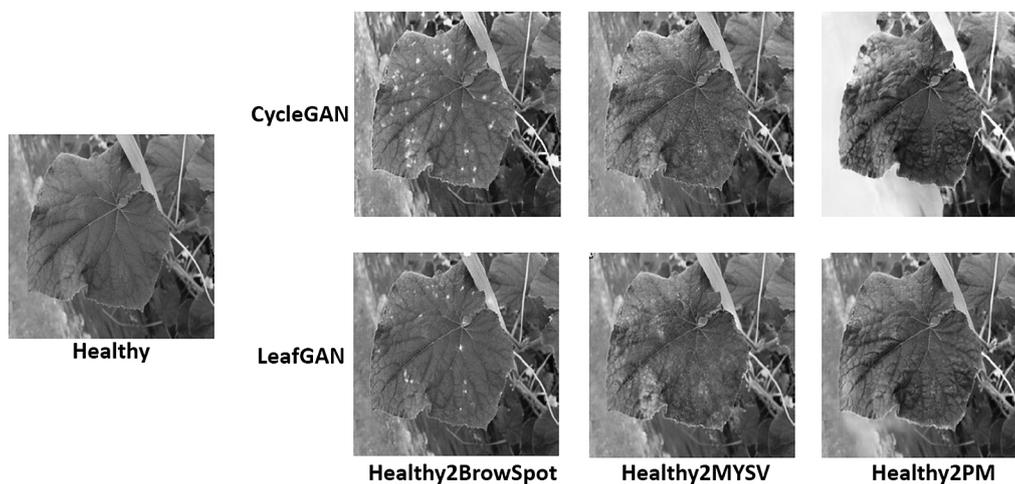


Fig. 3 Healthy image of cucumber (left) and examples of images generated by CycleGAN (conventional method)[9], LeafGAN[18] (our method) for Brown Spot, MYSV (viral disease name), and Powdery mildew (PM) diseases

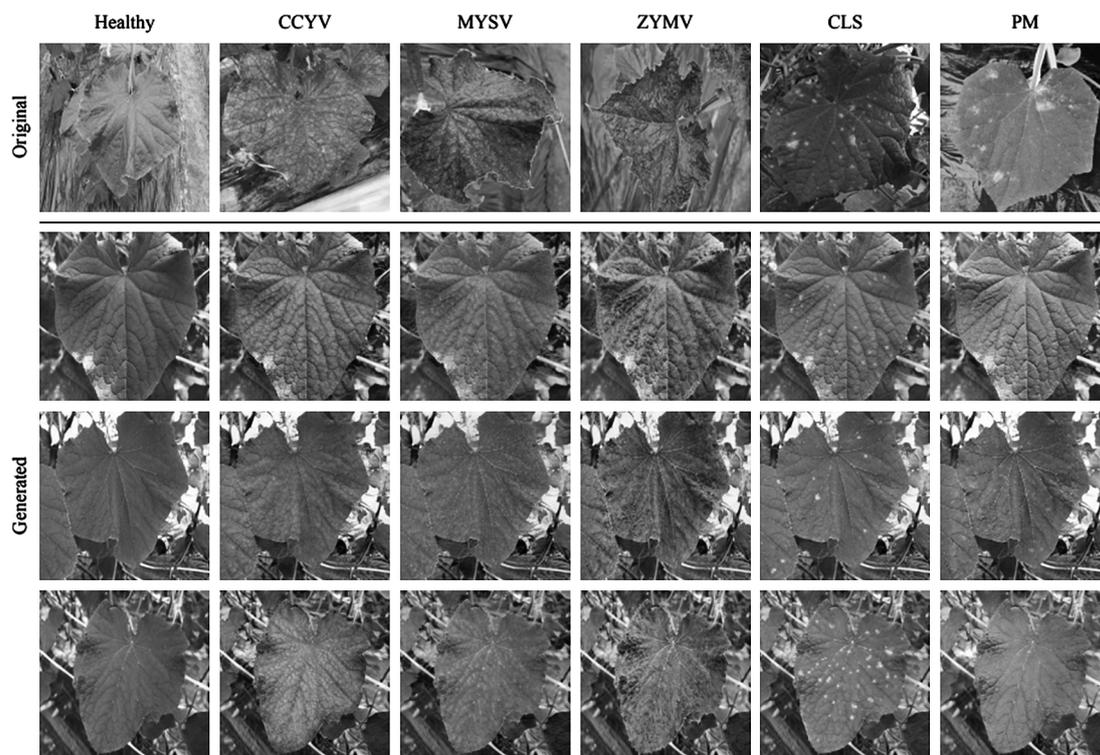


Fig. 4 Sample of generated image using PPIG[19]

で本物の葉画像は最上段のみで、左の列の2段目以降の画像が生成された偽の健全画像、それより右の列の画像が、その疑似健全画像に病徴を付与して生成した画像である。PPIGにおいて特筆すべきは、同じ構図で、病気のみ異なる画像を生成することが可能であるため、そうして生成された画像を学習する識別器は、背景などへの過学習が低減でき、平均識別率を71.5%から81.2%まで高めている。

このように data augmentation (学習用画像を疑似的に増やす手法)として活用できる画像生成技術は、植物病

害タスクに限らず広く有用な技術といえる。我々はこれをさらに発展させ、より頑健な識別器の構築を目指している。

5. 結 言

データ生成技術を含む、機械学習手法の劇的な進歩は、さまざまな分野で我々の生活を変えつつある。植物病自動診断技術もここ5年間で大幅に発展し、それに伴う企業や国、研究者など新たなプレーヤーが急増している。我々の自動診断システムは来年度より農水省より公開さ

れる。我々は逐次新しい技術の開発を行いながら実践的な識別器の開発を継続していく。

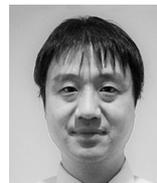
[謝辞] 本研究は農林水産省「人工知能未来農業創造プロジェクト」のうち、「AIを活用した病害虫診断技術の開

発」(JP17935051)、官民研究開発投資拡大プログラム(PRISM)、科学研究補助金(17K08033)の支援によって行われ、本稿の多くはその研究成果をもとに執筆した。関係各位ならびに共同研究でお世話になった方々にこの場を借りて深くお礼を申し上げたい。

References

- [1] A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet classification with deep convolutional neural networks, NIPS 2012 (2012) 1097–1105.
- [2] M. Tan, Q. V. Le, EfficientNet: Rethinking model scaling for convolutional neural networks, Proc. ICML (2019) 6105–6114.
- [3] S. Ren, K. He, G. Kaiming, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, Adv. Neural Inf. Process. Syst. (2015) 91–99.
- [4] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, A. C. Berg, SSD: Single shot multibox detector, Proc. European Conf. on Computer Vision (ECCV) (2016) 21–37.
- [5] C-Y. Wang, A. Bochkovskiy, H-Y. M. Liao, Scaled-YOLOv4: Scaling cross stage partial network, arXiv preprint (2020) 2011.08036.
- [6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Adv. Neural Inf. Process. Syst. (2014) 2672–2680.
- [7] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. Aila, Analyzing improving the image quality of StyleGAN, IEEE Proc. CVPR (2020) 8110–8119.
- [8] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, IEEE Proc. CVPR (2017) 1125–1134.
- [9] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cyclic-consistent adversarial networks, IEEE Proc. CVPR (2017) 2223–2232.
- [10] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, X. Tang, ESRGAN: Enhanced super-resolution generative adversarial networks, ECCV Workshop (2018) 63–79.
- [11] Y. Kawasaki, H. Uga, S. Kagiwada, H. Iyatomi, Basic study of automated diagnosis of viral plant diseases using convolutional neural networks, Lect. Notes Comput. Sci. 9475 (2015) 638–645.
- [12] S. P. Mohanty, D. P. Hughes, M. Salathe, Using deep learning for image-based plant disease detection, Front. Plant Sci. 7 (2016) 1419.
- [13] K. P. Ferentinos, Deep learning models for plant disease detection diagnosis, Comput. Electron. Agric. 145 (2018) 311–318.
- [14] M. Tan, Q. V. Le, EfficientNet: Rethinking model scaling for convolutional neural networks, Proc. ICML (2019) 6105–6114.
- [15] E. D. Cubuk, J. Shlens, Q. V. Le, RandAugment: Practical automated data augmentation with a reduced search space, IEEE Proc. CVPR Workshops (2020).
- [16] T. Saikawa, Q. H. Cap, S. Kagiwada, H. Uga, H. Iyatomi, AOP: An anti-overfitting pretreatment for practical image-based plant diagnosis, IEEE Proc. BigData2019 (2019) 5177–5182.
- [17] Q. H. Cap, H. Tani, S. Kagiwada, H. Uga, H. Iyatomi, LASSR: Effective super-resolution method for plant disease diagnosis, Comput. Electron. Agric. 187 (2021) 106271.
- [18] Q. H. Cap, H. Uga, S. Kagiwada, H. Iyatomi, LeafGAN: An effective data augmentation method for practical plant disease diagnosis, IEEE Trans. Autom. Sci. Eng. (2021) 3041499.
- [19] S. Kanno, S. Nagasawa, Q. H. Cap, H. Uga, S. Kagiwada, H. Iyatomi, PPIG: Productive pathogenic image generation for plant disease diagnosis, IEEE Proc. IECBES2020 (2021) 554–559.

〈著者紹介〉



彌富 仁

平成10年慶大・理工・電気卒。平成16年同大院博士課程了。博士(工学)。平成12~16年日本HP(株)勤務,平成16年法大・工・助手,平成21年同大・理工・専任講師,平成23年同准教授,博士(医学)東京女子医大,平成30年より法大・理工・教授(現職)。専門:機械学習および,それを応用した画像認識,画像処理,文書解析などの研究に従事