

# キュウリウイルス病の画像による診断システムの開発

埼玉県農業技術研究センター 宇賀博之  
 法政大学理工学部 彌富仁・川崎雄介・藤田恵梨香  
 法政大学生命科学部 鍵和田聡

## はじめに

植物の栽培においては、数多くの障害が発生する。その要因としては、糸状菌や細菌、ウイルス病といった病害のほか、養分の欠乏・過剰、ガス害、高・低温障害等の生理障害がある。これらの障害の中には一目で判断がつくものも多いが、作物における障害は多種多様で、その原因を特定するためには、豊富な知識と経験が必要となる。また、その場で判断できないものもあり、専門機関における顕微鏡観察や遺伝子診断等によるところも大きい。防除の一步は診断からと言われるように、対策のためにはその原因究明が必要不可欠であり、時間的なロスを防ぐためにもできるだけ早く行いたい。

生産現場における簡易診断技術としては、多くの事例が報告されている。例えば、キュウリの3種ウイルスを判別できるRIPA法(大崎, 2012)や簡易な遺伝子診断法であるLAMP法を用いたウリ類退緑黄化ウイルス検出キットやトマト黄化葉巻病診断キット(いずれもニッポン・ジーン)が販売されている。また、黒田ら(2013)は、小型顕微鏡を用いて病原菌の胞子を観察する技術を開発している。前者は、診断にはキットが必要であるほか、特定のウイルス病に限られる。後者は、比較的簡単に行える技術であるが、やはりある程度の専門知識が必要である。Web上では、植物障害の診断方法や関連情報を提供してくれるサイト(<http://otuki3.org/plant/pdc8.html>; 植物の病気診断室関連リンク集)などがあり、手軽に病害虫や生理障害情報が得られる。

一方、画像を用いた機械学習による病害などの診断技術については、いくつか報告がある

(HUANG, 2007; ZHANG and ZHANG, 2010; BARBEDO, 2013; GAVHALE and GAWANDE, 2014)。いずれの方法も、識別器の構成に不可欠な特徴量をその目的に応じて人為的に設計

する必要があり、また、そのための前処理として特定領域の抽出が求められる。

近年、機械学習の一分野であるディープラーニング(深層学習)は研究が盛んに行われ、多方面に技術展開が可能であると注目されている。簡単に言えば、コンピュータに人間のような学習機能を持たせることである。この技術のうち、画像認識などで応用されるConvolutional Neural Networks(CNN)は、識別に必要な特徴量をコンピュータが獲得した情報をもとに自ら学習する。このため、これまで人手に頼っていた画像の前処理や特徴量の設計が省略できるようになったほか、識別精度が飛躍的に向上している。また、このような技術は、一般物体認識の分野で世界的に大変幅広く注目されており、ここ数年世界トップレベルの競争が行われている(ILSVRC2015; <http://image-net.org/challenges/LSVRC/2015/>)。この技術を活用した植物における障害の診断目的での研究は、筆者の知る限りこれまで国内外ではほとんどなされていない。本研究では、植物の障害をその画像によって判断することを目的としており、その第1段階として、キュウリのウイルス病を対象とした信頼性の高い診断システムが構築できたので紹介する。

## I 供試材料

まず訓練データ作成のため、キュウリのウイルス罹病株を栽培し、その病葉画像の撮影を行った。キュウリには主に8種類のウイルスが感染するが、メロン黄化えそウイルス(MYSV)、ウリ類退緑黄化ウイルス(CCYV)、キュウリモザイクウイルス(CMV)、ズッキーニ黄斑モザイクウイルス(ZYMV)、カボチャモザイクウイルス(WMV)、パパイヤ輪点ウイルス(PRSV)およびキュウリ緑斑モザイクウイルス(KGMMV)の7種を用いた。病徴がCCYVと極めて類似した近縁種であるピートシユードイエロースウイルスは、本試験では除いた。栽培期間中にウイルス症状を呈した葉をデジタルカメラで撮影した(口絵①)。カメラはソニーサイバーショットRX-100を使用し、撮影条件は、画像サイズ:S:5.0M、横縦比:4:3、画質:スタンダード、シングルオートフォーカス、マルチフォーカスエリア、ISO:200、測光モ

Development of a Diagnostic System by the Picture of Cucumber Virus Diseases. By Hiroyuki UGA, Hitoshi IYATOMI, Yusuke KAWASAKI, Erika FUJITA and Satoshi KAGIWADA

(キーワード: 診断, キュウリ, ウイルス病, 画像, 深層学習, CNN)

ード：マルチ、オートホワイトバランスとした。また、被写体については、画像中央領域に識別対象の葉が写っている状態とし、識別対象に強く日光が当たらないようにした。画像は様々なステージを網羅するために、幼植物、成植物のほか、病徴の形成過程も含めて撮影した。

## II 診断システムの構築

### 1 識別器の構造

CNNは脳の視覚野の処理モデルに基づき、畳み込み演算処理ユニットをネットワーク状に多段接続したものである。その識別器の構造を図-1に示した。本研究では、畳み込み層とプーリング層を交互に3層持つ構造とした。また、フレームワークとしてはCaffe (Jia, 2014)を使用した。画像を入力することによって、それがどのクラスに分類されるかの事後確率を出力することになる。

### 2 識別予備試験

キュウリのウイルス病を葉の画像のみで診断できるかの可能性を検証するため、予備試験として、MYSVの感染による発病葉ならびに健全葉をCNN識別器に機械学習させた。まず、口絵①に示す通り、各画像は葉が1枚入るように撮影し、キュウリ健全葉画像を797枚、MYSVによる発病葉画像を235枚使用した。識別能力の評価は4-fold cross-validationによって行った。データ量の増大と様々な撮影角度に対する汎用性を確保するために、学習時のデータ画像を10度刻みで360度回転を行い、学習画像量を36倍とした。その結果、識別率(罹病葉と健全葉を正しく判別した割合)は86.5%、感度(罹病葉を正しく判別した割合)は79.1%、特異度(健全葉を正しく判別した割合)は88.7%であった。正しく識別された画像の一部を口絵②に示した。一方、正しく識別できなかった画像の一部を口絵③に示す。これらは、見た目にも判断が難しく、今回得られた精度は、MYSVの感染病徴に見慣れた人と同等以上と思われた。

### 3 CNNを用いた複数ウイルス病徴の診断

予備試験において、CNNを用いることでキュウリの葉の画像のみからウイルス病の診断が可能なが示唆されたため、複数のウイルス種の識別が可能かどうか検証した。使用した画像数は、以下の両試験ともCMVのみ320枚で、その他はすべて1,000枚とした。CNNの構成や学習画像を増加させる手法は予備試験と同様とした。

試験1：4種のウイルス(MYSV, CCYV, CMV, ZYMV)と健全株の5分類を識別させた結果を表-1に示した。MYSVの感度(ウイルスに対する識別能)がやや低く87.8%となったものの、その他ウイルスに対する感度、および健全葉に対する識別能である特異度、またそれらの平均値である正解率のすべてが90%以上と高精度であった。

試験2：試験1の4種にWMV, RSVおよびKGMMVの3種ウイルスと健全株を合わせた合計8分類を識別させた。その結果、試験1と同様に畳み込み層を3層とし、学習画像に対して回転処理のみを行った場合の正解率は75.5%となり、5分類の場合と比較してかなり精度が低下した。特異度は86.9%と比較的安定していたが、各ウイルスの感度は軒並み60%台にまで落ち込んだ(表-2)。

表-1 4種類のウイルス感染葉および健全葉の5分類時における診断精度(試験1)

評価指標	精度[%]
正解率	90.6
感度(MYSV)	87.8
感度(ZYMV)	91.8
感度(CCYV)	90.8
感度(CMV)	90.3
特異度	92.0

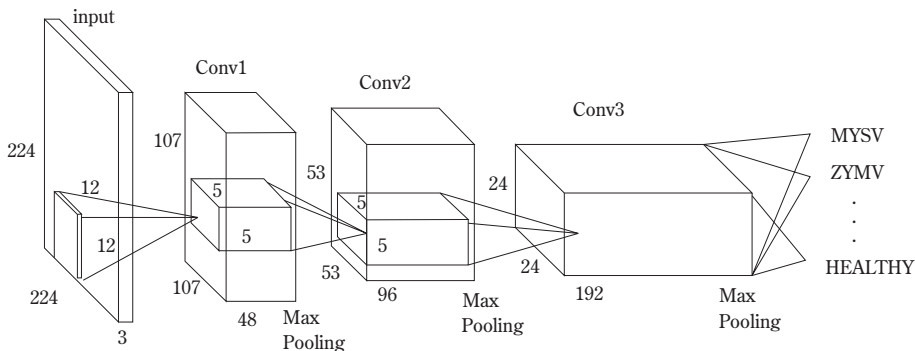


図-1 CNN識別器の構造

表-2 7種類のウイルス感染葉および健全葉の8分類時における診断精度(試験2)

評価指標	精度[%]
正解率	75.5
感度 (MYSV)	80.6
感度 (ZYMV)	67.6
感度 (CCYV)	86.3
感度 (CMV)	60.4
感度 (PRSV)	69.3
感度 (WMV)	69.7
感度 (KGMMV)	69.1
特異度	86.9

表-3 7種類のウイルス感染葉および健全葉の8分類時における診断精度(試験2:精度向上版)

評価指標	精度[%]
正解率	83.2
感度 (MYSV)	85.4
感度 (ZYMV)	79.3
感度 (CCYV)	92.5
感度 (CMV)	67.8
感度 (PRSV)	79.9
感度 (WMV)	81.5
感度 (KGMMV)	77.5
特異度	91.0

そこで、識別器の畳み込み層を1層追加して4層にし、学習させる画像に反転や抽出位置に変化をつけるなどの工夫をした結果、正解率は83.2%、特異度は91%にまで向上した(表-3)。また、感度についてもおおむね80%程度以上に精度が向上したが、CMVでは67.8%とあまり向上が見られなかった。これは、CMVのみ使用した画像数が少なかったことと近年のキュウリはCMVの耐病性品種がほとんどで、葉における病徴が非常に軽微であることが起因していると考えられた。

### III 診断システムの検証

ウイルス病徴は常に典型的な症状を示すとは限らないため、どのようなサンプルを訓練画像として用いるべきか検討が必要になる。また、紹介した技術開発に使用したウイルスの病徴画像は、1機関で取得したものである。栽培した作型は複数ではあるが、品種やウイルス分離株の違いによって病徴に多様性があるため、さらに訓練データを追加していく必要がある。一方、画像の取得方法について、撮影するカメラの性能や解像度、日射条件、背景への写り込み等、あらゆる場面を想定する必要があるが、カメラに関する部分においては、これまでの経験によりさほど影響はないと考えられる。また、現時点では、葉における診断に限っているが、花や果実、茎等、葉以外の部位に奇形やえそ症状等の特徴的な障害も見られるため、植物を全体的に見て判断する場合も考慮するなど、随時、診断精度の向上に努めなければならない。

同じ作物で新たな障害を診断対象に加えた場合、診断精度が維持できるかどうかについては大変重要な事項である。前述の試験では、判別対象が健全を含めて5分類から8分類に増えた場合、識別精度の低下が見られた。しかし、学習する画像数が多くなっていることを考慮し、CNN構造の大規模化やチューニング、訓練画像を増やすことにより、8~9割程度の識別精度が期待される。

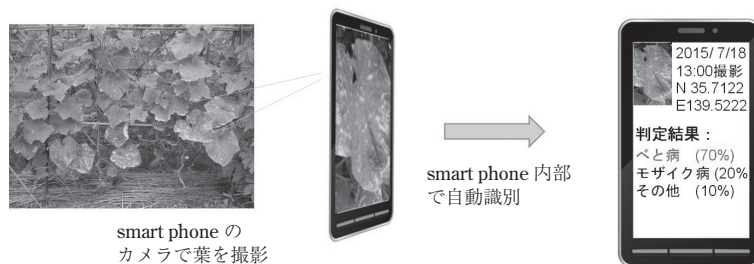


図-2 診断アプリの将来像

デジタルカメラや携帯電話、スマートフォン等で障害を撮影し、診断アプリに画像を入力すると、その原因と様々な対応策が出力される。

## おわりに

本研究ではキュウリのウイルス病をモデルとして、診断システムの構築を行った。その結果、CNNを利用することで、特徴の設計や抽出、前処理を人為的、機械的に行うことなく8分類の識別が可能であることが示された。しかし、冒頭でも述べたが、作物の障害はウイルス病に限ったものではなく多種多様である。今後は他の病害や生理障害、あるいは、虫害による食害痕等にも対応させる予定である。当然ながら、作物もキュウリに限ったものではないが、この技術の鍵は何と言っても正しい訓練データが如何に数多く収集できるかにかかっている。一つの障害について最低千枚程度は必要であり、多く学習させるほど精度は高まる。また、一つの作物で判別したい障害の数が多いほど各障害の必要画像数も増える。現在、この画像データの取得方法については検討を進めており、順次、このようなシステムを作物ごとに作製し、スマートフォンのアプリなどでの公開を検討して

いる(図-2)。アウトプットされる情報については、障害の原因や対策等を想定しているが、生産者や指導者、企業等に対して行ったアンケートも参考にしながら、よりよいシステムを構築していく予定である。

**謝辞** 本研究の一部は、JST研究成果展開事業 A-STEP FS ステージ探索タイプ「農作物における病気の画像診断システムの構築」および同マッチングプランナープログラム探索試験「キュウリ病害の画像診断システムの開発」により実施した。

## 引用文献

- 1) BARBEDO, J. G. A. (2013): Springer Plus, Vol. 2: 660.
- 2) GAVHALE, K. R. and U. GAWANDE (2014): IOSR Journal of Computer Engineering, Vol. 16: 10 ~ 16.
- 3) HUANG, K.Y. (2007): Computers and Electronics in Agriculture 57: 3 ~ 11.
- 4) JIA, Y. et al. (2014): Caffe: Convolutional architecture for fast feature embedding, arXiv preprint arXiv:1408.5093.
- 5) 黒田克利ら (2013): 植物防疫 67: 304 ~ 306.
- 6) 大崎秀樹 (2012): 同上 66: 224 ~ 227.
- 7) ZHANG, J. and W. ZHANG (2010): Proc. Advanced Computer Control (ICACC) 5: 264 ~ 266.