

センシング農業ローバーを用いたマルチモーダル観測による桑畑の病害観測手法の検討

寺田 憲司†††

藤波 香織†††

† 東京農工大学 大学院 生物システム応用科学府 生物機能システム科学専攻

†† 職業能力開発総合大学校 基盤ものづくり系 ††† 東京農工大学 大学院 工学研究院 先端情報科学部門

1 はじめに

世界の農業をめぐる状況は急峻な天候変化、農薬・肥料価格の高騰等により深刻な食糧生産危機に直面している。日本においても同様に農業従事者の高齢化、後継者不足と労働力不足による経営者あたりの平均耕地面積の規模拡大は喫緊の課題である [1]。そこで農場の省人化、データ化を目指したスマート農業、精密農業 [2] に大きな関心が寄せられている。営農者は航空ドローンに代表される IoT, AI, ロボット技術を利用し、農場内の「情報可視化」「作物毎への個別処置」が重要となる [3]。一例として多くの農家は決められた時期に農場全体へ農薬を散布するカレンダー-散布が一般的である。そこで観測箇所の病害分類が可能となればハウス栽培や果樹園でも低コストな検出箇所のスポット農薬散布が実現できる。

一方で作物を加害する様々な害虫として代表的な甲虫類、ハダニ、鱗翅目害虫等は葉の裏に寄生し食害を行う [4]。葉裏の解析は植物の良い育成のために重要である [5]。植物ウイルス性病害は一か所でも放置すれば、同じ農場にあるほかの作物に蔓延し 2 次被害を引き起こす。このような害虫による被害は食料全体の 40% にあたる [6]。病害の防除は主に農家自身の経験の蓄積に頼っており、それは必ずしも安定したものではなく、形成に長い時間を要する [7]。防除として農家の多くは農薬、除草剤を作物全体に散布しコストが高くなる。また、作物が化学物質を過剰に浴びてしまい、人体に悪影響を及ぼす可能性もある [8]。そこで病害予測のために画像による病害検出は有能であるが、病害の主原因はウイルスである。人の風邪でもウイルス感染から数日後に発症するように、植物も目視で確認できる状態の病害は数日前に感染していることが多く、病害の目視検出時には既に広範に広がっている可能性が高い。そこで作物をできる限り個別に病害や育成状況を把握し、農薬散布や振動処理などのスポット対処を実施することでコスト削減実現する高効率なシステムが求められている。

圃場の観測ではいくつかの手法に分類できる。ヒトの五感と記憶による観測、固定式センサーによる天候に左右されにくい堅牢な観測、航空用ドローンによる圃場上部の広域観測に分けられる。本研究では、これら観測から得られるデータを補完する形として圃場の不整地でも走破可能な小型農業用ローバーを用いて航空用ドローンでは観測が困難な作物の葉裏画像、地上

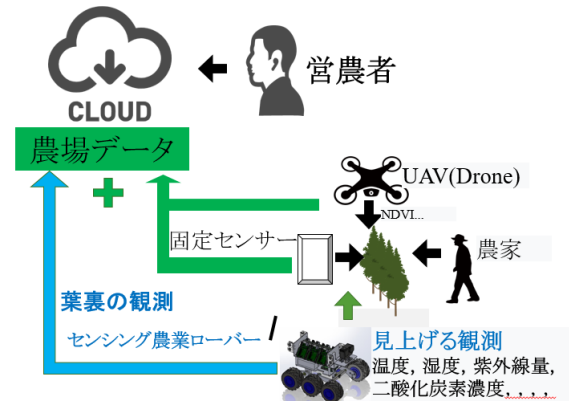


図 1: 営農者のための農場データ

の環境値のマルチモーダル観測及び分析を行い、営農者の支援を行う事を目的としている (図 1)。

2 章では関連研究について述べる。3 章では、関連研究から課題を整理し、課題解決のためのマルチモーダル観測が可能な農業用ローバーを提案し、データ収集と予測を行うための特徴量増加手法を提案する。4 章では機械学習アルゴリズム Gradient Boosting Decision Tree (GBDT) を適用するために不均衡なデータセットに対する対応手法について述べる。5 章では有効な特徴量について分析を行い、自然環境で計測を優先すべき特徴量を機械学習アルゴリズムの視点から考察を行う。6 章で本研究をまとめる。

2 関連研究

2.1 AI を使った営農システム

近年、センシング技術を利用した農業支援の研究の中で圃場観測に特化した研究は AI の登場により活発化している。日本の農林水産省は WAGRI と呼ばれる気象や農地、収量予測など農業に役立つデータやプログラムを提供する公的なクラウドサービスを提供している [9]。飼育予測や病虫害診断等の様々なオープンデータ、ビッグデータを WebAPI で提供し、農業現場の生産性を高めるサービスを実施している。病虫害診断は、キュウリ、ナス、イチゴ、モモなどの 12 品目の検出が可能となっている。若林ら [10] は AI, IoT のテクノロジーを採用した農業 ICT サービス「食・農クラウド Akisai」を提唱している。画像解析による茶葉の診断に AI を利用し、適切な時期に効率的に収穫作業を行えるよう研究を進めている。規格をオープンとすることで様々な自治体、大学、企業と連携しデータを集約している。課題は病虫害データベースの整備が足りない点である。病害予測の開発は 2020 年時点で有用な病虫害のデータベースが不十分であるため、連携団体の力を借りて、専用スマートフォンアプリケーションでデータ

Investigation of multi-modal observation methods for detecting diseases in mulberry fields using a sensing agricultural rover

††† Kenji TERADA ††† Kaori FUJINAMI

† Department of Bio-Functions and Systems Science, Tokyo University of Agriculture and Technology

†† Division of Core, Polytechnic University, Japan

††† Division of Advanced Information Technology and Computer Science, Tokyo University of Agriculture and Technology

を集約し予測システムの開発を試みている。盛朝子ら [11] は AI を活用した病害感染のリスク予測を実現するハウス栽培向けのスマート農業ソリューション Plantect を紹介している。全国各地の試験・研究機関および生産者の協力のもと、延べ 1 万棟以上のハウスでデータ収集・調査を行い、病害の発生に関わりのある可能性が考えられる温度・湿度、病害名、発病が認められた日、薬剤散布記録、圃場および栽培に関する情報の情報を収集した。機械学習としてランダムフォレストを使用し、得られた 2 日後までのハウス内温湿度予測値を、病害感染リスク予測モデルに当てはめることによって、2 日後までの感染リスクの予測が可能となっている。病害予測に関する研究では、Fenu [12] らは、3 年間 (2016 年 - 2018 年) にわたってカリアリの場所から ARPAS 機関によって提供された過去の気象データを使用して行われ、地域の気象観測所からの気象データを用いてジャガイモ晩枯病のリスク指標を予測できるかどうかを検討した。Artificial Neural Network (ANN) の予測精度は 96 %、Support Vector Machine (SVM) 分類の予測精度は 98 % を達成した。Mallocci ら [13] は逆伝播ニューラルネットワークを使用して病気の発生を予測した。予測には、最高気温、最低気温、最高湿度、最低湿度、降雨量、および全インド共同研究プロジェクト (AICRP) から取得したジャガイモに関する作物データを考慮した。活性化関数には、Sigmoid, TanH, ReLu の 3 種類が使用された。テストの結果、Sigmoid 関数を使用することで、より良い結果が得られることが示され、0.909 の精度を達成した。

これらの課題は画像収集の困難さと予測の精密さである。画像収集は農家の方のスマートフォンなどの撮影で集めているため、多大な労力が必要となる。画像収集時の病害撮影時の温度、湿度等の詳しい観測条件の管理は不足かつ不十分である。また、固定式センサーや気象情報を利用した予測に関しても農場単位であり、作物毎もしくは 1m 四方のエリアなど精密な予測の取り組みは不十分である。そこで「作物毎」のより精密な観測のために、画像観測時の周囲の温湿度や二酸化炭素等の環境条件を自動で収集し、病害予測をより実現しやすいマルチモーダルな観測が実現できれば、営農に有益な情報を提供できると考えられる。

2.2 UAV(Unmanned Aerial Vehicle)

人が登場しない航空ドローンの登場により農業・肥料散布、及び農場センシングによる育成状況の把握の効率化がすすめられている [14]。Yang ら [15] は航空用ドローンで撮影した写真と衛星写真から多くの作物病害の検出とマップ作成に成功した。広範な画像から病気を他の交絡因子から区別するために高度な画像センサーや分析手法の評価が必要であると指摘し、早期検出が課題と述べている。杉浦ら [16] は日本の北海道にある大規模農場を航空用ドローンを用いて、上空よりマルチスペクトルカメラを利用して正規化植生指標を取得し、育成状況のバラツキを可視化している。課題として取得画像とグランドトゥルースのデータセットを揃える作業の必要性を指摘し、画像による病害検出に関しては、病害の種類によっては、地上での目視評価、特に熟練の目と同等の検出精度を得るのは難しい場合が多く万能なものとは望めないと指摘している。

UAV は雨天時、強風時や夜間、都市部での飛行が制

限される。夜行性の害虫であるヨトウムシの検出など夜間観測の有効性は想像しやすいが、各国の法規制も年々厳しくなっている UAV では難しく、精密農業を実現するために地上走破の観測が求められる。

2.3 UGV(Unmanned Ground Vehicle)

地上での目視評価の重要性に注目し、UGV と呼ばれる地上を走破するローバーに注目している研究がある [17]。

Lars ら [18] は、農場環境の多様化を課題としてとらえている。そこで温室、露地、傾斜、ポリトンネル等のあらゆる地形に対応できるように設計された全天候型移動ロボット Thorvald II を紹介している。悪路での走破性を確保するため、パッシブサスペンションを装備した四輪駆動であり、トラック幅は 1.5m、質量 200kg 以下のロボットである。ノルウェー国や英国内のキュウリやトマト農園のモニタリングや UVB 処理を実現する。Gonzalez-De-Santos [19] は、スマートファームを実現する無人走行車を提案しており、操舵機構を実現するための脚付き 4 輪ローバーを設計している。ただし、動作制御が難しいことが課題となっている。松本ら [20] は、小型ロボットは大型機械の導入が難しい中山間地域や傾斜地の多い果樹園、園芸ハウス内での運用の容易さが利点があると指摘している。更に、大型農機と比べ安全性が高く、他の農機では難しい夜間運用の可能性を示唆している。

日本国外の農業用ローバーの課題は大型のものが多くある。凹凸のある悪路を走破するためにはダンパーや操舵機構など多くの部品が必要となり大型化は避けられないと考えている。小型になれば段差の影響が大きくなりカメラ映像など、安定した観測が難しくなる。また、人が乗れるほどの大きさに近づけば現行のバイクやホビー用の乗用車部品を流用できるので、オーダーメイドでの部品制作となるケースが多い小型ローバーより低コストでの制作が可能である。しかし、乗用車の大きさとすれば山間部、住宅地で見られる密集栽培が多い環境では適用が難しい。

2.4 画像分類器

画像による病害検出はスマートフォンを利用したアプリケーションが世界中に存在する [21]。WANI ら [22] は、実際の農業環境下での作物病害の自動検出と詳細な病害予防の提案のための高性能な作業ソリューションを提案している。YOLOv3 CNN アーキテクチャによる葉の識別、画像処理による感染の広がり検出、InceptionV3 による病害検出を行い、自動農薬散布機を実現している。Lou ら [23] は、農場にいる幼虫や蛾、バッタなどを検出する画像分類器として YOLOv5 の有効性を示している。YOLOv5 の CIoUloss 関数を EIoUloss 関数に変更することにより精度向上を果たした。Shi ら [24] は、いくつかの農業害虫識別ソフトウェアが市場に存在するが、その識別精度は圃場によって、不安定であることを指摘している。EfficientNetV2 転移学習に基づき、病害によって学習データ数が違うデータセットの偏りを解決するために、各トマト病害の写真 300 枚をランダムに選択し、水平方向、垂直方向に鏡像変換を行った。結果として 9 つの一般的なトマトの病気が正確に識別され、mAP の値は 0.98 に達し、圃場によって学習が必要となることを示している。

YOLOやEfficientNetによる病害分類は営農に適用可能な分類器の有効性を示している。ただし、関連研究からわかる事は画像は同一品種で揃える事が必要であり、加えてこれらの研究で学習に利用する画像は葉表が多い。これらのデータセットを転移学習を用いて本研究で利用したとしても有効となるとは言い難い。例えば、桑のようにキュウリやトマトと比べ生産量が少ない品種に適用する際は転用して使用できない。

3 小型農業ローバーによるデータ収集

3.1 課題

2章で散在した各技術の課題をまとめると以下の通りである。

1. 観測した病害画像取得時の温湿度等の環境情報を付加したマルチモーダル情報取得の多大な労力
2. 固定式センサーによる精密観測の困難さ
3. 飛行条件に制限がある UAV
4. 日本の密集栽培に適用した観測
5. 凹凸がある不整地での安定観測
6. 未知な作物での病害分類器の適用の低さ

これら課題を解決するためにセンシング農業ローバーを用いたマルチモーダル観測による病害観測手法が求められる。この手法により病害の症状が作物に見られる前に病害発生の予測をする狙いがある。作物の間近で観測できる小型農業ローバーの観測情報は作物毎の状態をヒートマップ等で可視化でき、作物毎の観測情報の把握、根元、葉裏情報の取得等利点が多い(図2)。例えば、航空用ドローンは葉裏自体の処置についてはプロペラの角度を変え、葉裏への農薬散布を行える製品[25]が存在する。しかし、観測となると上部のみとなるため圃場予測では不十分であることが指摘されている[26]。登場を走破するローバーであれば葉裏に対しても効果的な運用を行う事ができる。一方で課題となるのが以下2点となる。

1. 不整地での一定速度の走破性能の実現
2. 1日1回の巡回による観測データの少なさ

「1」の課題を解決するために本研究では図3に示す電動6輪駆動関連懸架構造を持った小型農業用ローバーを開発した[27](図4)。小型農業ローバーに設置したRaspberry Pi4B+に温度、湿度、照度、紫外線量、気圧、二酸化炭素濃度を計測できる環境センサーに加え、作物を撮影するためのカメラを搭載した。センサー値とカメラ映像によるマルチモーダル情報を可視化することにより自動農場監視が可能となる(図5)。「2」の課題は、仮にローバーが24時間稼働となり絶えず走破すると、水たまりの原因となる土壌の輻生成の促進や運用コストの増大等につながる。また海外営農に多く見られる50haを超える大規模農場での運用となると更に1日の観測回数は減ると考えられる。関連研究にあった病害予測は数年にわたる膨大なデータ数が必要となることが多い。1日1回の観測では作物毎の病害予測に必要な学習データを集めるには多くの年月が必要となってしまう。一方、小型農業用ローバーは1日に農場を巡回しながら多くの葉裏や作物の画像を温湿度等の環境情報とともに収集することが可能である。他の

観測点のデータを任意の観測点の病害予測にも使用できれば、短期間の観測でも予測に十分なデータを収集できると考えた。本研究は作物毎に得られる病害予測に必要なデータを農場巡回時に得られた環境センサー値、カメラによる観測画像を利用することで蓄積していく。更に、農場全体で得られた精密なデータを加工・分析を行う事で新たな病害予測を実現する狙いがある。

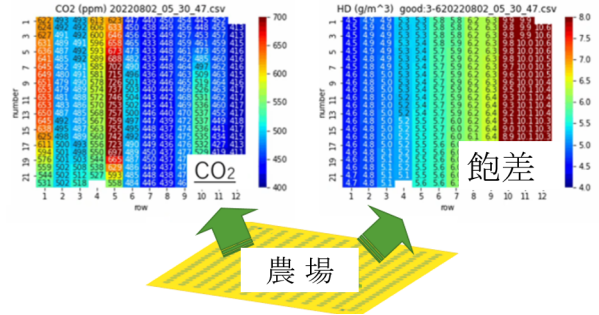


図2: 本研究の計測で可視化できる農場のヒートマップ例(二酸化炭素濃度, 飽差)

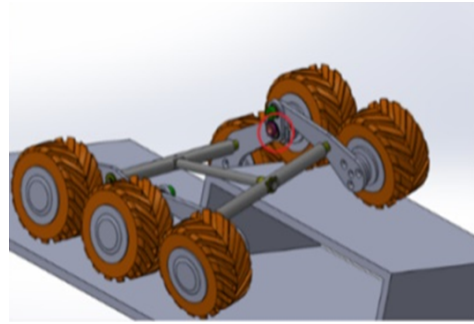


図3: 電動6輪駆動関連懸架構造



図4: 小型農業ローバー [27]

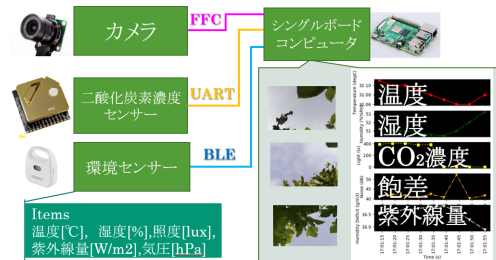


図5: 観測システム部の構成

3.2 データ収集方法

データ収集は農場規模が約10[m]×7[m]である屋外の桑畑を利用した。桑はある一定の間隔で直線上に植えられており、ローバーは12樹列を走破し3.1節に示

すセンサー、カメラを利用し観測を実施した。ただし旋回については農場隅が公道道路のため旋回スペースが無く手作業で行った。観測は晴天時に気孔が安定して開くとされる夜明けから1時間以内に農場を巡回し計測した。農業用ローバーは土壌での轍抑制のため1日に1巡回の走破が望ましい。加えて雨天時は土壌を痛めるため未観測とした。環境値、画像はセンサー精度の限界である3秒毎に計測した。1列当たり1分で走破できるよう速度調整を行い両端のデータを除いた18点の観測を得た。農場は小さい枝や鳥糞等がつくった段差がある不整地である。土壌は水はけが良いよう設計されているが、日によって含水率の違いから走行にスリップが生じ、1樹列当たりの観測点数は日によって1個の差が生じた。

3.3 病害分類のためのラベル付与

葉裏画像から病害情報を得て機械学習の説明変数で利用するためにYOLOv5を利用した分類器を構築した。農場で2022年7月2日から7月22日に収集した葉裏画像を利用し、目視で変色がみられる葉裏を病害とした。収集できた330枚の葉裏画像データに矩形アノテーション処理を行い、YOLOv5を利用した結果、適合率0.90、再現率0.91となった(図6)。この分類器の分類スコアを0.5を閾値として目視に頼らず病害検出を1、病害未検出を0とする病害情報をセンサーによる観測点に付加することとした。少量の画像でも高い精度を検出できた主な理由としては葉裏画像をローバーで収集する際、見上げる撮影をしたためと考えられる。どの観測点でも背景は上空の単色画像であり、観測時間も明け方であったため逆光による影響が少なく検出画像のパターンが少ない結果となった。



図 6: YOLOv5 による病害分類

4 病害予測手法

4.1 検討手順

予測手法の検討の流れを以下に示す。

1. 機械学習アルゴリズムの選択
2. 特徴量の定義
3. 不均衡データ対策
4. 予測日を固定した学習データ期間の比較
5. 予測日をスライドさせたラグ特徴量による病害予測

4.2 機械学習アルゴリズムの選択

本研究の観測環境を考慮し、機械学習アルゴリズムはGBDT手法であるLighGBM[29][30]とXGBoost[31]を利用した。GBDTの特徴は以下の通りである。

1. 特徴量重要度の把握がしやすい
2. 標準化・正規化の影響を受けにくい
3. 外れ値の影響を受けにくい
4. 特徴量をヒストグラム化することでメモリの計算負荷を抑制

「1」の特徴量重要度とは学習させた数々のデータのうち、どれが予測に貢献しているのかを定量的に比較することで、どの特徴量が重要なのか推測する際に有効である。「2」について決定木は分類境界を引くときにある次元に垂直な線を引くため、特徴量の大小関係のみに着目している。この特徴は観測する温度、紫外線量等単位が違う観測値を多く扱う際に有効であり、本研究との親和性が高い。「3」について決定木はデータ分割によって特徴を顕在化させるためにデータの中に稀に表れる異常に高い値や低い値があったとしても外れ値として含むデータブロックとして分割し、外れ値のある領域だけが除外されるように分割されるため、外れ値の影響が少ない手法である。この特徴は自然環境センシングに見られる誤検出に有効である。「4」は処理を行う計算機の小型化につながる。これは設置スペースが限られる小型ローバー搭載に有効である。

評価手法は再現率を用いる。理由は本研究を利用する営農を考慮しているためである。営農では病害検出箇所への対応の優先順位が高い。例えば農薬散布の場合、病害が無い箇所への農薬散布より病害の検出失敗による未検出を避けたい。

4.3 特徴量の定義

本研究では教師つき機械学習により病害分類器の構築及び病害予測を行う。特徴量は3.1節で示したセンサーにより観測した温度、湿度、照度、紫外線量、気圧、二酸化炭素濃度の時系列データから算出する。加えて農場座標、定点の風速、風向情報を利用する。病害は観測点周辺から影響を与える事を考慮し式(1)を用い周囲 v か所の近傍群 9 ヶ所の約1.5メートル四方の平均値を算出した。当日を含む過去 d 日間の平均、分散、標準偏差、最大値、最大値と最小値の差、尖度、歪度を用い合計86個の特徴量を使用した(図7)。

$$\bar{M} = \frac{1}{(v+2)^2} \sum_{i=-v}^v \sum_{j=-v}^v M_{i,j} \quad (1)$$

- x 農場の横座標
- y 農場の縦座標
- $M_{x,y}$ 温度、湿度、照度、紫外線量、気圧、二酸化炭素濃度の観測値
- v 近傍相対範囲(本研究では1)、奇数
- \bar{M} 近傍群 $(v+2)^2$ 点の平均

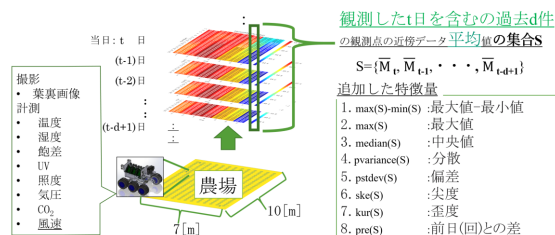


図 7: 観測データと特徴量の生成方法

4.4 不均衡データ対策

病害検出箇所は未検出場所と比べて少ないため観測データが不均衡であると考えられる。そこで不均衡データ対策手法として、多数のクラスのデータを減らす Under Sampling, 少数のクラスのデータを増やす Over Sampling, 決定境界の近くにあるサンプルのウェイトを下げる Sample Weight, 少ない方のデータをただ単にコピーするのではなく検出した近接データとの内挿を使って増やす SMOTE[32], SMOTE 手法にクラス分類をしやすくするために分類境界線近辺のデータを増やす Boderline SMOTE[33], Over Sampling の一手法である SMOTE と Under Sampling の一手法である Edited Nearest Neighbor (ENN) の 2 つを組み合わせさせた手法である SMOTE + ENN[34], 少数クラスのデータ付近に多数クラスがどの程度あるかの重みを動的に増やす ADASYN[35] を利用する。

5 実験

5.1 実験データの取得

実験は自然環境の影響を受ける屋外の桑畑を利用する。台風や数日続く雨天による精度の影響を検証するため、図 9 に示す日毎の訓練データを増やしながらか評価を行った。学習データは 2022 年 8 月 6 日から 2022 年 9 月 26 日の雨天時を除く観測できた 33 日を利用した。予測する日を 2022 年 9 月 27 日として観測結果の再現率評価を行った。更に観測点の近傍データを利用するために農場の端の観測点は除外した。そのため 1 日の観測で利用できるデータは 10 列×16=160 点の観測点となり、これら観測点に 3.3 節で説明した特徴量を計算したデータを利用した。目的変数は病害情報の 2 値 (検出: 1, 未検出: 0) とした (式 (2))。あわせて、病害情報の例として予測日で利用した 9 月 27 日の病害農場マップを図 8 に示す。

$$y(t) = A(\vec{F}_t) \quad (2)$$

t 予測日

A() GBDT アルゴリズム

$y(t)$ 目的変数: 病害の有 (1) 無 (0)

\vec{F}_t 特徴量 (学習データ: 観測点数×86×日数)

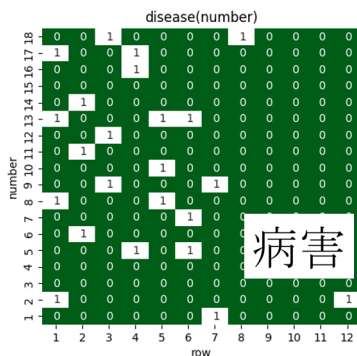


図 8: 観測農場の病害ヒートマップ例 (図中の数字は 0 は病害なし, 1 は病害ありの検出結果を表す)

5.2 不均衡データ対策の比較

4.4 節で示した手法を LightGBM で試し再現率の比較を行った。表 1 に結果を示す。

表 1: 不均衡データ対策における再現率の比較

対策手法	再現率
なし	0.089
Under Sampling	0.649
Over Sampling	0.659
Sample Weight	0.637
SMOTE	0.593
Boder line SMOTE	0.659
SMOTE+ENN	0.819
ADASYN	0.666

不均衡データのため再現率は他の項目と比べて極端に低い。最もよい結果は SMOTE + ENN となった。原因は不均衡データの割合であると考えている。学習データ期間において病害の不均衡データの平均割合は約 30.1 % である。多くの不均衡対策手法は数%程しか発生しないデータに対して適用することを目的とした手法である。

それら手法の中で SMOTE + ENN はオーバーサンプリングによって少数派データを増やした後に、分類をする上で不要なデータを除去することで、クラス間の境界をより明確にしている。病害が発生しているデータが多ければ不要なデータ除去が減り、過学習を避けられるとともに再現率の高いデータセットを生成できる。よって、本研究では不均衡データ対策として SMOTE + ENN を利用する事とした。

5.3 GDBT アルゴリズムの比較

GDBT アルゴリズムの比較を行った。表 2 に示すように、SMOTE + ENN を適用した LightGBM の再現率が最も高い値となった。原因として考えられるのは Leaf-wise により決定木が形成される LightGBM は Level-wise により決定木が形成される XGBoost に比べて決定木が複雑になりやすい。複雑になるということは効率的に分類が進んでいると考えられる。一方、データセットに対して適合し過ぎて汎用性が低くなる過学習をしている可能性がある。この点を検証するには、観測システムの時間当たりの観測数を増加する必要がある。しかし、現状の手法を農場観測に当てはめた場合、有効な面も存在する。農業の周囲を含めた環境は統一されていない。特に日本は広大な平野部から山間部、住宅街の中心部等、農場の周囲条件は多種多様である。その中から複雑な発生条件である病害を定める事は難しい。そこで、本研究手法のように 1ヶ月等、例えば一定期間の観測で観測した農場に適合した学習モデルを利用する方が、より良い運用となると考える。そこで、GBDT アルゴリズムの利点をいかし、有効な特徴量を分析する必要がある。

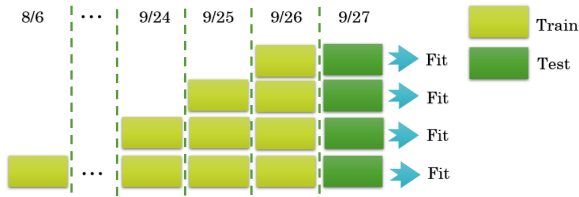


図9: 訓練データとテストデータの作成方法

表2: 各分析アルゴリズムの平均再現率

	XGboost	LightGBM
特徴量増加無	0.013	0.021
特徴量増加有 SMOTE+ENN 無	0.034	0.028
特徴量増加有 SMOTE+ENN 有	0.332	0.819

5.4 特徴量重要度の分析

SMOTE + ENN を適用した LightGBM について特徴量重要度に注目した。特徴量重要度は各特徴量が分類時にモデルの性能向上に寄与した情報利得を計算する。情報利得とは分割前後の不純度の減少量を示す。計算結果から特徴量がモデルの予測精度にどれだけ寄与したかを重要度として評価する。特に影響が出ている値は図10に示す通り当日の風速、過去3日の観測点近傍エリア9点の温度の平均値を用いた標準偏差、風速の平均値を用いた最大値、湿度の平均値を用いた最大値及び紫外線量である。風と紫外線はビニールハウスの防除目的と一致している。また、病害は数日の多湿の影響が大きいため湿度上昇に影響がある温度のバラツキが特徴量の重要度が大きい原因と考えられる。更に、観測期間の9月18日、22日の台風遭遇により連続した未観測日が発生している。台風後はどの営農でも農場を巡回するのでシステム上での予測ニーズは低い、風速は今回の特徴量重要度でも上位であるため台風後の観測データの適用について熟考する必要があると考えられる。

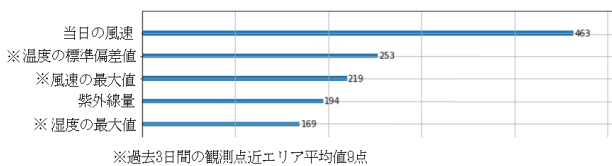


図10: 特徴量重要度の上位

5.5 病害発現遅延を考慮した特徴量生成

数日後の病害予測が可能か検証を行った。まず、再現率の変化を図11に示す。9月27日の病害の検証で利用する学習データについて台風発生時のデータセットを適用した際に再現率低下がみられる。加えて降雨による未観測期間の8月30日から9月4日より前の再現率について顕著な低下がみられる。このことから急激な天候変化、一定期間の未観測期間が生じるような1か月以上の学習データを利用した場合、再現率の低下がみられることがわかる。そこで3から6日間の学習データを利用した LightGBM による予測を検証する(図12)。数日後の検証は(2)式を以下の(3)式に変更し

たラグ特徴量による検証を行う。病害の誘因となる環境が続くと遅れて葉の病害が発生する[36]。計測した環境値に x 日後の病害結果をつなげたデータセットを利用する事で学習を行う。検証データセットにも数日ずれたラグ特徴量を使うことでずれた分の病害予測が可能であると考えた。

$$y(t+x) = A(\vec{F}_t) \quad (3)$$

x ラグ日数

t 予測日

A() GBDT アルゴリズム

$y(t)$ 目的変数: 病害の有(1)無(0)

\vec{F}_t 特徴量 (学習データ: 160点×日数)

ラグ特徴量を利用した再現率の結果を表3に示す。0は当日の推定を表す。表中の括弧の中の数値は偏差である。全センサーの結果より、学習利用期間 S を増やすことにより再現率向上を確認できた。更にラグ特徴量とすることで、2日目までの精度向上が明らかとなった。偏差まで見てみるとばらつきがある中、再現率0.995を達成した学習期間6日ラグ2日は偏差0.370となった。これは日によって極端に低い再現率がある事を示しており、よりバラツキの少ない再現率達成手法が求められる。全センサーの特徴量を利用した場合と比べ再現率が高い結果となったのは風速、紫外線量である。紫外線量はフィルム等で量を減らすことでいくつかの病害が防除できることが明らかとなっている[37][38]。農場内の不均一な紫外線量から機械学習による病害予測ができる可能性がある事がわかった。標準偏差からデータのバラツキも少ない。本研究では変色が見られる葉裏画像を病害と位置付けた。分類器でより緻密な害虫分類モデルを適用すれば紫外線量が原因となる起因も特定できる。植物病原菌の促進に風が関係している事はすでに明らかになっている[39]。病原菌の感染行動には水と風が必要であり、高湿度条件下で菌が形成され、風や水滴化により周囲に伝搬する。故に、風の変化が特徴量として高い再現率を示したと考えられる。ただし、風速の再現率1.000のスコアを達成したのは全て病害と検出したためである。よって風速変化が最善とは結論付ける事は難しい。温度、湿度は他のセンサーと比べて再現率のスコアは低い。しかし、主にハウス栽培で作物の気孔の開閉状態を把握するための育成支援情報である飽差についてはラグ2日について、どの学習利用期間でも再現率向上が確認できる。二酸化炭素濃度は光合成に影響があるため、飽差と同じ指標となりえたが、ラグ日数が多くなるにつれ再現率は低下した。

一方で作物を育成する際に重要となる光に関係がある照度の再現率が低い。この原因について日の光が届きづらい地上近辺による観測のためと考えるが、これは人間が感じる物理量である照度を光の測定で利用したためと考えられる。光量子を測定する必要があると考える。

あくまで、表3の結果は研究で利用した桑畑での特徴に過ぎない。本研究で測定した農場は道路に面しているため、郊外に多くある農場とは違う特徴が表れている可能性がある。環境によって優先される環境測定値が変わる可能性があるため、大きく環境が異なる農

場での検証を重ねていく。この結果によりラグ日数に関わらず学習利用期間を増やせば再現率のスコア向上が確認できた。一方でセンサーの組み合わせはより多く検証する必要がある。

以上、ラグ日数についても日数を増やすことで精度に上下が見られるものも学習期間を増やせば0.95の高い精度が望める事がわかった。

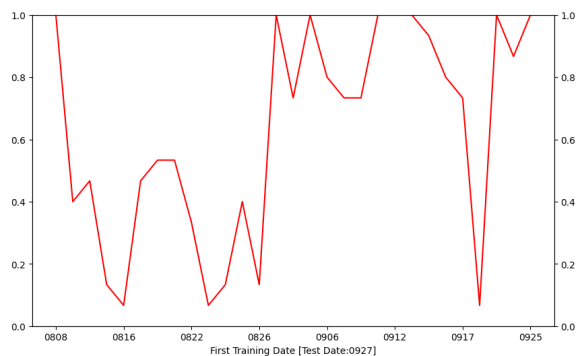


図 11: 学習データ開始日別の再現率の変化 (評価日: 2022年9月27日)

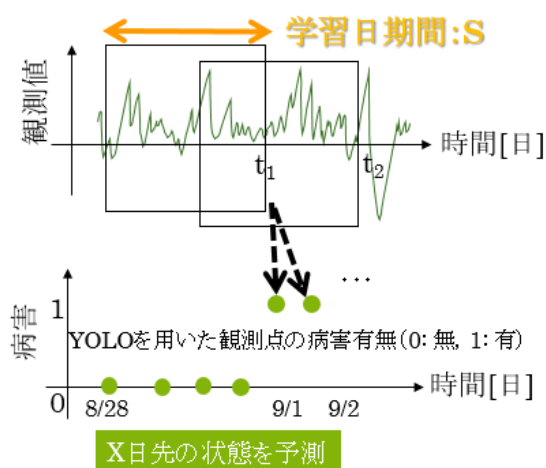


図 12: ラグを利用した予測手法

6 おわりに

本研究では観測が困難な作物の葉裏画像及び地上の環境値のマルチモーダル観測及び分析を行った。営農者の支援を実現する病害予測手法として農場観測データセットに SMOTE + ENN を適用し再現率向上を確認し、LightGBM を用いて再現率 0.819 を達成した。加えてラグ特徴量を用いる事で再現率向上、1,2 日後の病害予測の可能性を確認した。今後は光量子センサーの追加や様々なセンサーの測定結果の組み合わせ検証を重ね、病害予測精度の向上を目指す。更に病害分類の細分化を行い、営農を支援する手法の確立を目指す。

謝辞

小型農業ローバーの開発・支援はペンタリンク株式会社により遂行された。桑畑のデータ収集は小暮農場の提供により実現した。

参考文献

- [1] 吉田真悟, 松久勉. 労働力不足下にある農業法人の特徴と課題—日本農業法人協会アンケート調査に基づいて—. 農業経済研究 93.3, pp.325-330, 2022.
- [2] 澁澤 栄. 精密農業. 朝倉書店, 2006.
- [3] Julien Champ, Adan Mora-Fallas, Hervé Goëau, Erick Mata-Montero, Pierre Bonnet, Alexis Joly. Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots. Applications in plant sciences Vol8-7, pp.e11373, 2020.
- [4] 日本国 農林水産省. 病虫害防除. 秋田県野菜栽培技術指針, pp.23-25, 2007.
- [5] 彌富仁. 植物病害自動診断技術の動向と課題. 日本神経回路学会誌 Vol.26 No.4, pp.123-124, 2019.
- [6] David Pimentel. Agricultural science. Eolss Publishers/UNESCO Singapore Vol.2, pp.272-293, 2009.
- [7] Hu Wan, Zheng Lu, Wang Qi, Yuanyuan Chen. Plant disease classification using deep learning methods. Proceedings of the 4th international conference on machine learning and soft computing, pp.5-9, 2022.
- [8] Keith Kirkpatrick. Technologizing agriculture. Communications of the ACM Vol.62 No.2, pp.14-16, 2019.
- [9] Takehiro Shiomi, Takeshi Saito. Overview and System Operations of Agricultural Data Collaboration Platform W AGRI. Journal of the Japanese Society of Agricultural Machinery and Food Engineers Vol.82 no.3, pp.229-233, 2020.
- [10] 若林毅. 食・農分野におけるデータ活用を支援する ICT プラットフォームの開発. JATAFF journal, pp.34-38, 2019.
- [11] Asako Mori. 日植防シンポジウムからセンサーデータと AI を活用した病害感染リスク予測. Plant protection Vol.73 No.6, pp.381-386, 2019.
- [12] Gianni Fenu, Francesca Maridina Mallocci. An application of machine learning technique in forecasting crop disease. Proceedings of the 3rd International Conference on Big Data Research, pp.76-82, 2019.
- [13] Mallocci, FRANCESCA MARIDINA. Predictive Analytics Models and Tools for Decision Support of Stakeholders in Digital Agriculture. Università degli Studi di Cagliari, 2022.
- [14] Haruki Murofushi, Chong Nak-Young. Report on Current Situation and Issues of Drone Utilization in Smart Agriculture. IPSJ SIG Technical Report Vol.2017-CE-140 No.3, pp.1-5, 2017.
- [15] Yang Chenghai. Remote sensing and precision agriculture technologies for crop disease detection and management with a practical application example. Engineering 6.5, pp.528-532, 2020.
- [16] 杉浦綾. ドローン空撮画像による大規模圃場のリモートセンシング. 日本農業学会誌 45.2, pp.146-149, 2020.
- [17] Hugo Rafacho Fernandes, Edna Carolina Moriones Polania, Angel Pontin Garcia, Oscar Barrero Mendonza, Daniel Albiero. Agricultural unmanned ground vehicles: A review from the stability point of view. Revista Ciência Agronômica Volume 51, pp.1-12, 2020.
- [18] Lars Grimstad, Pål Johan From. Thorvald II - a Modular and Re-configurable Agricultural Robot. IFAC-PapersOnLine Volume 50 No.1, pp.4588-4593, 2017.
- [19] Gonzalez-De-Santos, Pablo and Fernández, Roemi and Sepúlveda, Delia and Navas, Eduardo and Armada, Manuel. Unmanned ground vehicles for smart farms. Agron.-Clim. Chang. Food Secur Vol.6, pp.73-86, 2020.
- [20] 松本 将大. 過酷な農業環境下で安全に使用するために小型農業用ロボットへの期待と留意点. Farming mechanization No.3260, pp.20-23, 2023.
- [21] Nikos Petrellis. Mobile application for plant disease classification based on symptom signatures. Proceedings of the 21st Pan-Hellenic Conference on Informatics, pp.1-6, 2017.
- [22] Buchammagari Avinash Reddy, Ghanta Sai Krishna, Kani-ganti Priyanka Saraswathi, Indala Sathvika, Venkanna Udu-talappally, Debanjan Das. dScout: Unmanned Ground Vehicle for Automatic Disease Detection and Pesticide Atomizer. 2022 IEEE 7th International conference for Convergence in Technology (I2CT) IEEE, pp.1-6, 2022.

表3: ラグ特徴量による平均再現率 (標準偏差)

			学習利用期間 [S 日]			
			3	4	5	6
全センサー	ラグ X 日数	0	0.781(0.330)	0.788(0.139)	0.806(0.174)	0.813(0.393)
		1	0.895(0.395)	0.976(0.334)	0.753(0.423)	0.957(0.126)
		2	0.734(0.470)	0.896(0.365)	0.689(0.355)	0.995(0.370)
温度のみ	ラグ X 日数	0	0.497(0.440)	0.613(0.470)	0.646(0.471)	0.777(0.483)
		1	0.668(0.364)	0.763(0.320)	0.535(0.268)	0.581(0.068)
		2	0.306(0.490)	0.370(0.489)	0.321(0.480)	0.235(0.456)
湿度のみ	ラグ X 日数	0	0.764(0.515)	0.749(0.458)	0.698(0.363)	0.859(0.425)
		1	0.574(0.479)	0.600(0.375)	0.601(0.478)	0.644(0.354)
		2	0.739(0.435)	0.667(0.449)	0.673(0.448)	0.577(0.315)
飽差のみ	ラグ X 日数	0	0.683(0.412)	0.737(0.301)	0.653(0.443)	0.742(0.406)
		1	0.558(0.441)	0.570(0.394)	0.572(0.383)	0.447(0.461)
		2	0.768(0.380)	0.884(0.431)	0.832(0.447)	0.922(0.466)
二酸化炭素のみ	ラグ X 日数	0	0.686(0.423)	0.703(0.430)	0.684(0.343)	0.764(0.477)
		1	0.735(0.448)	0.724(0.373)	0.814(0.385)	0.644(0.304)
		2	0.519(0.372)	0.546(0.374)	0.624(0.399)	0.415(0.477)
紫外線量のみ	ラグ X 日数	0	0.750(0.150)	0.480(0.020)	0.698(0.165)	0.578(0.000)
		1	0.915(0.149)	0.912(0.000)	0.896(0.076)	0.878(0.047)
		2	0.993(0.355)	0.677(0.327)	0.673(0.417)	0.977(0.263)
風速のみ	ラグ X 日数	0	0.930(0.437)	0.889(0.441)	0.700(0.453)	1.000(0.468)
		1	0.923(0.269)	0.928(0.378)	0.922(0.292)	0.956(0.333)
		2	0.832(0.023)	0.788(0.511)	0.819(0.509)	1.000(0.437)
照度のみ	ラグ X 日数	0	0.880(0.279)	0.651(0.380)	0.608(0.444)	0.626(0.421)
		1	0.703(0.359)	0.645(0.379)	0.702(0.351)	0.742(0.200)
		2	0.660(0.456)	0.638(0.434)	0.522(0.315)	0.676(0.302)

- [23] Lijun Lou, Junya Liu, Zhen Yang, Xin Zhou, Zhijian Yin. Agricultural Pest Detection based on Improved Yolov5. Proceedings of the 2022 6th International Conference on Computer Science and Artificial Intelligence, pp.7-12, 2022.
- [24] Zhanhao Shi, Cui Wang, Lin Zhao. Tomato Disease Identification Application Based on EfficientNetV2. Proceedings of the 6th International Conference on Computer Science and Application Engineering, pp.1-6, 2022
- [25] Oleksandr Kholodiuk, Oleksii Tokarchuk. The efficiency of using agras drones for spraying, their design, technical and technological features. Engineering, Energy, Transport AIC No.2 (117), pp.63-74, 2022.
- [26] 中島満, 五月女淳, 前沢嘉彰, 柳真一, 植松勉. 航空機(無人ヘリを含む)を利用した農薬散布の現状と研究課題. 植物防疫 Vol.56 No.2, 2002.
- [27] 寺田憲司, 藤波香織, 木下研吾. 精密農業のための観測可視化ローバー ～不整地対応台車を利用した圃場走破小型6輪農業用ロボットの開発～. Farming mechanization No.3260, pp.14-19, 2023.
- [28] Angelina Gill, Tarandeep Kaur, Yendrembam K Devi. Application of Machine Learning Techniques in Modern Agriculture: A Review. AProceedings of the 2022 Fourteenth International Conference on Contemporary Computing, pp.263-270, 2020.
- [29] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems Vol.30, 2017.
- [30] Chunzhi Wu, Xiaofei Xue, Yongtao Song. Research on Cancer Diagnosis Method Based on LightGBM-Gridsearchcv. Proceedings of the 4th International Conference on Big Data Engineering, pp.122-126, 2022.
- [31] Tianqi Chen, Carlos Guestrin. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp.785-794, 2016.
- [32] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, W Philip Kegelmeyer. SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research Vol.16, pp.321-357, 2002.
- [33] Hui Han, Wen-Yuan Wang, Bing-Huan Mao. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In International conference on intelligent computing, pp.878-887, 2005.
- [34] Gustavo Eapa Batista, Ronaldo C Prati, Maria Carolina Monard. A study of the behavior of several methods for balancing machine learning training data. SIGKDD Explor. Newsl. 6(1), pp.20-29, 2004.
- [35] Haibo He, Yang Bai, Eduardo A Garcia, Shutao Li. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In 2008 IEEE International Joint Conference on Neural Networks, pp.1322-1328, 2008.
- [36] 桑病診断技術研究会. クワ病害診断の手引き. 桑病診断技術研究会, pp.2-3, 2007.
- [37] 小豆澤斉, 安田雄治, 山本孝司, 梶野康行. ブドウの紫外線除去栽培が樹体生育と果実収量及び品質に及ぼす影響. Bulletin of the Shimane Agricultural Experiment Station Vol.35, pp.33-42, 2004.
- [38] 岡久美子, 山田真, 石渡正紀, 岡田清嗣. 紫外光 (UV-B) 照射による施設野菜の病害防除. Kinki Chugoku Shikoku agricultural research Vol.16, pp.9-14, 2010.
- [39] Toshiyuki Usami. Disease Prediction in Greenhouses. HortResearch Vol.74 No.2, pp.8-10, 2020.