vinetur

ブドウ樹の枝を画像上で特定するAIモデルを開発、IoUは88.25％

冬季のブドウ畑で既存の画像セグメンテーションツールを上回ったAIモデル。安定したロボット剪定に向けた一歩となる

2026年11月14日（土）



農業工学の研究チームが、複雑なブドウ畑の画像からブドウ樹の幹と剪定対象の枝を識別する人工知能システムを開発した。将来的にロボットが冬季剪定をより安定して行う助けとなる可能性がある。11月13日金曜日に発表された研究成果は学術誌「Smart Agricultural Technology」に掲載されており、完全自律型の剪定機ではなく、コンピュータービジョンに焦点を当てている。

剪定はブドウ畑で日常的に行われる作業のなかでも、とりわけ重要なものの一つだ。樹の成長や果実の生産、樹勢、ひいてはブドウの品質を調整する役割を担う。多くのワイン産地では、作業の大部分を今も熟練した作業員に頼っている。そのため費用と時間がかかり、特に人手不足のなか、大規模なブドウ畑全体で作業を標準化することも難しい。剪定はブドウ生産者にとって大きなコスト

となるため、技術が商業用の畑で信頼性を実証できれば、自動化の進展は将来、ワイン産業にとって重要な意味を持つ可能性がある。

新たなモデルの名称は、Dual-branch Transformer-Mamba Fusion Network（DTMF）。休眠期にあたり、ブドウ樹が葉を落として冬季剪定に適した状態になった時期に撮影した屋外画像から、木質化した樹体構造を識別することを目的としている。一見すると対象は限定的だが、ブドウ畑のロボット化における基本的な課題に取り組むものだ。機械はまず、ブドウ樹と剪定すべき枝、周囲の背景を見分けられなければ、枝を切ることはできない。

背景はソフトウェアにとって判別が難しいことが多い。細い枝は、棚線や土壌、雑草など、周囲のさまざまな視覚情報に紛れやすい。研究チームによると、一般的な画像セグメンテーションシステムでは、ブドウ樹全体の大まかな形状と、細い枝の輪郭にある微細な情報の両方を捉えるのが難しい場合がある。この課題に対処するため、研究チームは二つの経路からなるシステムを設計した。一方は画像全体の場面や長距離にわたる要素間の関係を分析し、もう一方は細い木質部の追跡に役立つ微細なテクスチャーや輪郭、周波数パターンに着目する。

技術的には、このモデルはビジョントランスフォーマー、離散ウェーブレット変換、Mambaの状態空間モジュールを組み合わせている。トランスフォーマー側は画像全体の文脈を捉えることを狙い、ウェーブレットとMambaの側は細部の構造や細い枝に沿った連続性を保つことを目指す。二つの経路を早い段階で無理に統合するのではなく、処理の大半で分けておき、最終段階でのみ予測を統合する。研究チームは二段階の学習手法も採用した。学習初期には基本的なピクセル分類を行い、後半には境界の精緻化を強化する。これには、予測を物体の自然な輪郭に沿わせるよう設計された、スーパーピクセルベースの損失関数も含まれる。

研究チームは、中国・新疆ウイグル自治区で冬季休眠期に収集した高解像度RGB画像2,150枚を使ってシステムを検証した。データセットには、石河子大学の実験用ブドウ畑で2024年11月に撮影した1,250枚と、トルファンの商業生産用果樹園で2024年12月に撮影した900枚が含まれる。全画像について、幹と枝を示すピクセル単位のラベルを手作業で付与した。データセットは、学習用1,505枚、検証用215枚、テスト用430枚に分割された。データ拡張後、実質的な学習用データは4,515枚に増えた。

研究によると、DTMFの完全モデルはテストデータで、平均ピクセル精度94.24％、一般的なセグメンテーション指標である平均IoU（交差/和）88.25％を達成した。既存の複数モデルとの比較では、DeepLabV3、PSPNet、FCN、Segformer、U-Net、トランスフォーマーベースのU-Netを上回った。枝の検出で差が特に大きかった。枝は細く、背景に紛れやすいため、この課題のなかでも難しい部分だ。DTMFの枝のIoUは85.12％で、このカテゴリーで最も高いベースラインモデルの70.13％を14.99ポイント上回った。幹のセグメンテーションも87.68％と高かったが、この分類では他の最新モデルも差を縮めた。

研究チームは、システムの各構成要素が性能にどれほど寄与するかを調べるため、一部を取り除くアブレーションテストも実施した。その結果、デュアルエンコーダー・デュアルデコーダー構造、ウェーブレットとMambaの融合モジュール、二段階の損失関数戦略はいずれも性能向上に寄与したことが示された。トランスフォーマー経路のみを使った場合、平均IoUは71.98％だったが、完全な設計では88.25％に上昇した。学習手法を個別に検証したテストでは、後半段階の教師信号とスーパーピクセル制約を加えることで、枝のIoUが74.33％から85.12％に上がった。

論文に掲載された視覚例では、比較対象のシステムよりも枝の途切れや欠落を抑えながら、棚線や起伏のある土壌など判別の難しい背景のなかで枝を追跡する様子が示された。研究チームは、他の木本植物を写した外部画像でもモデルを試した。これらの実演から、手法がより幅広い構造物の認識に役立つ可能性が示唆された一方、論文では、ブドウ畑のデータ以外への汎化性能を正式に検証したものではなく、あくまで例示にとどまると説明している。

研究の著者らは、この成果が完全なロボット剪定ソリューションを意味するものではないと強調した。正確な画像セグメンテーションは、より大きな課題の一部分にすぎない。実用段階の剪定ロボットには、ブドウ樹の構造を三次元で再構築し、切断位置を特定し、移動経路を計画したうえで、実際のブドウ畑の環境下で切断工具を制御する能力がなお必要となる。また研究チームは、モデルの並列アーキテクチャーによって計算コストが増し、農業用ロボットに搭載する組み込みハードウェアへの実装が難しくなる可能性も指摘した。

論文では、ほかにもいくつかの限界が挙げられている。画像は新疆の2地域で冬季休眠期にのみ収集されており、ほかの栽培地域や棚仕立て方式、葉が茂る夏季の条件では、まだモデルの有効性が検証されていない。データセットは個々の株単位ではなく画像単位で分割されており、異なる乱数初期値によるばらつきを測るための反復学習の結果も報告されていない。こうした限界を踏まえると、良好なベンチマーク結果は認識技術の進展を示すものとして捉えるべきで、広範な商業利用の準備が整った証拠ではない。

それでも、この研究はブドウ畑の自動化に向けた実践的な方向性を示している。今後、処理速度が向上し、より多様な実環境で検証されれば、ブドウ生産におけるロボット剪定の主な技術的障壁の一つを減らす助けとなる可能性がある。ワインやその他のブドウ由来飲料の原料を供給するブドウ畑では、将来的に労働力計画の安定化や、より均一な樹体管理につながる可能性もある。ただし、それには画像セグメンテーションにとどまらない、現場への統合と検証が欠かせない。