

Abstract: 本研究計画では、果物栽培支援のための異種ロボットエージェント間の転移学習と物体状態推定について検討する。異なる形態やモダリティを持つロボットエージェント間でのスキル共有を可能にする転移学習と、花、果実、芽などの重要な対象物の正確な状態推定手法を提案する。これらの技術は、ロボットによる授粉、収穫、芽間引きなどの複雑なタスクの学習を加速させる。

Keywords: 転移学習, 異種ロボット, 状態推定, 果物栽培, マルチモーダル認識, 強化学習

1. 異種ロボットエージェントにおける果物栽培のための転移学習

転移学習により、異なる形態やモダリティを持つロボットエージェントが学習したスキルを共有し、果物栽培支援などの複雑なタスクの学習を加速することが可能になる。最近の研究では、異種ロボット間でのポリシー転移が探求されている。例えば、マニピュレータアームのポリシーをドローンにマッピングするなど、それらの行動空間と状態空間の間のマッピングを学習することで実現される [1][2]。主な課題は、ヒューマノイド、四足歩行ロボット、ドローン、アームなどのロボットが大きく異なる運動学と観測視点を持っているにもかかわらず、類似したタスク（花の授粉、果実の収穫など）を実行する必要があることである。本節では、ロボット間の転移のための理論的フレームワークについて論じる。これには、行動空間と観測空間の整合、マルチモーダル表現学習、強化学習と模倣学習のパラダイム、そしてフェデレーテッドマルチロボット学習が含まれる。

1.1. 形態間の行動空間マッピングとポリシー転移

異種ロボットは多くの場合、関節数、ダイナミクス、行動空間が異なるため、直接的なポリシー転移は自明ではない。行動空間マッピングは、あるロボットの行動を別のロボットの同等の行動に変換する関数を求める。一つのアプローチは接地行動変換であり、ロボット A のシミュレータで行動を取ることが、ロボット B によって実行された場合と同等の状態変化を生じさせるようなマッピングを学習する [3]。例えば、車輪型ロボットの前進動作は、ドローンの前進させるピッチ入力に対応する可能性がある。学習された変換によりこれらの行動が整合される。別のアプローチではグラフニューラルネットワーク (GNN) を用いてロボットの身体を表現する。各肢/関節をノードとしてモデル化することで、新しい形態にも一般化できるグラフ構造の形でポリシーを訓練できる [2]。Wang らの NerveNet は、異なる体構造を持つエージェントを制御できる GNN ベースのポリシーを導入した。各関節に対してモジュラーニューラルコントローラを共有することで実現される [4][5][6]。このモジュラーポリシーアプローチ（すべてを制御する一つのポリシーを学習する）は、各アクチュエータに同一のニューラルモジュールを割り当てることで、様々な平面エージェント（二足歩行、四足歩行など）での移動を単一のコントローラで実現することを示した [5][6]。図 1 はこ

の考え方を示している：異なる骨格形態を持つ複数のエージェントが共有モジュラーポリシーで共同訓練され、肢モジュール間のメッセージパッシングを通じて形態間のスキル転移を可能にする [5][6]。GNN を超えて、研究者たちは異なるロボットの行動を比較できる普遍的な潜在行動空間の学習を探索している。例えば、ポリシーは結果（エンドエフェクタをポイント A から B に移動するなど）を生成するように訓練され、その効果の中間表現を用いて異なる身体間で転移できる [1]。Zhu らは効果サイクル一貫性を強制するために効果空間マッピングを使用し、ソースドメインとターゲットドメイン間の行動シーケンスの結果を整合させた [1]。これにより、生の行動が異なる場合でも、両ドメインで誘発される状態（効果）のシーケンスが一致することを確保することで、あるロボットから別のロボットへのポリシー転移が可能になった。このような整合技術は、単純な転移と比較して、クロスドメインポリシーのパフォーマンスを大幅に向上させた [1]。

[5] 図 1：共有モジュラーポリシーは、各肢に同一のコントローラモジュールを割り当て、モジュール間の通信を可能にすることで、異なるロボット形態間で制御を一般化できる。これにより、単一のポリシー π_θ が一足歩行からヒューマノイドまでの様々なエージェントで一貫した行動（例：移動）を生成することが可能になる [5][6]。

別の研究方向は形態間の模倣学習に焦点を当てている。行動を明示的にマッピングするのではなく、ターゲットエージェントはソースエージェントの行動を観察し、同様の結果を達成するポリシーを推論することで学習する。三人称模倣学習は視点と身体の不マッチに対処する：Sharma らは身体に依存しない高レベルコントローラが別のエージェント（例：人間のデモンストレーション）を観察することからサブゴールを生成し、身体に特化した低レベルコントローラがこれらのサブゴールを実行するという階層的アプローチを提案した [7][8]。この分離により、例えば、ヒューマノイドロボットが四足歩行ロボットのデモンストレーションタスクを模倣することが可能になる。まずデモンストレーションを抽象的なゴール（「花に移動して授粉剤を適用する」など）に変換し、次に自身の行動空間を使用してそれらを達成する。このような階層的分離により、形態の違いにより直接模倣が失敗する場合でも転移が改善された [8]。同様に、観察からの模倣技術により、ロボットはそのエージェントの行動を見ることなく、異なるエージェントのビデオから学習できる。例えば、アームロボットは、コンテキスト変換ネットワークを使用して視覚的観察を自身の状態-

† is the presenter of this paper.

行動ドメインにマッピングすることで、ドローンの成功した飛行を観察して授粉戦略を学習できる [9][10]。Desai らは、シミュレーションを観察して学習したポリシーが、異なる物理特性を持つ実際のロボットに転移できるようにするダイナミクスミスマッチに対応した観察からの模倣方法を実証した [11]。これらの方法は、ロボットが達成すること（結果）が、それをどのように達成するか（方法）よりも形態間で転移しやすいことを強調している。そのため、結果表現の整合が重要である。

1.2. 観測空間の整合と視点転移

明示的な 3D マッピング以外にも、表現学習が観測の整合を達成できる。時間対比ネットワーク (TCN) のような自己教師あり技術は、異なる角度から同じイベントの観測を正の対として扱うことで、視点に依存しない特徴を学習する [13]。Sermanet らは TCN を使用して、時間的に整合したフレーム（例：花が授粉される様子のドローンとアームの視点）が類似した表現を持つ埋め込みを学習した [13]。このような埋め込みにより、あるモダリティの観測で訓練されたポリシーを別のモダリティに適用できる。両方が共通の潜在空間にエンコードされるためである。観測表現の対比学習 (CPC, SimCLR の適応など) は、ドメインに依存しない特徴に効果的であることが示されている [14]。強化学習では、Zhang らはダイナミクスサイクル一貫性アプローチを導入した。エージェントが自身のドメインで別のドメインに対応する観測シーケンスを予測することを学習し、潜在空間でのクロスドメイン対応を強制する Zhang2020。これにより、ロボット間で観測-応答パターンが一致することを確保し、転移が改善される。

具体例として、ドローンから地上ローバーへの果実検査ポリシーの転移を考えてみよう。ドローンの空中画像とローバーの側面画像は大きく異なる。視点マッピング (3D ガウシアンスブラッキングモデルや対比視覚エンコードを通じて) を使用することで、ローバーはカメラ入力をドローンの視点特徴に対応する方法で解釈でき、ドローンのポリシーを使用して花に近づき、視界の中心に花を配置することができる。実験では、このような観測マッピングが、整合された行動効果と組み合わせることで、シミュレーションにおいて異なるカメラ視点とロボット形態に対してゼロショットポリシー転移を可能にすることが示されている [1]。残りのギャップはしばしば微調整を必要とするが、ゼロから訓練するよりもはるかに少ないサンプルで済む。

1.3. マルチモーダルセンサー抽象化と共有表現

視覚だけでなく、果物栽培におけるロボットは様々なモダリティを使用する: RGB カメラ、深度センサー、触覚センサー (果物を掴む際など)、力センサー、固有受容感覚 (関節エンコード)。マルチモーダル転移学習は、これらの感覚ストリームを異なるエージェントが理解できる抽象的な形で組み合わせた共有表現を見つけることを目指している。例えば、ヒューマノイドと四足歩行ロボットは両方とも視覚と触覚を持っているかもしれないが、生の読み取り値は異なる。抽象的な表現 (「物体が検出され、距離約 1m、接触力は穏

やか」といったどちらのエージェントのセンサーからも来る可能性のある埋め込み) を学習することで、ポリシーはより転移しやすくなる。Gupta らは、異なるロボットのペア間で不変特徴空間を学習することで、把持スキルをあるロボットから別のロボットに転移できることを実証した [16]。特徴空間は高レベルのアフォーダンス (「手の届く範囲に物体がある」や「手がターゲットを囲んでいる」など) をエンコードし、これらはロボットの特定のセンサー構成に依存しない。このような不変空間は、オートエンコーダやエージェント間で特徴の区別がつかないことを促進する敵対的訓練を通じて学習できる [1]。しかし、特徴が真に一般的であり、本質的な情報を失わないようにすることは課題である。不変性を過度に制約すると、パフォーマンスが損なわれる可能性がある [1]。

一つのアプローチは、エージェント間で一部の層やモジュールのみを共有し、他は各エージェント固有のままにすることである。例えば、視覚は一定レベルの抽象化まで共通の畳み込みネットワークで処理され、その後、別々のネットワークがアームとドローンのモーターコマンドを予測する。この部分的な重み共有により、共有知覚と異なる制御が実現される。別のアプローチであるモジュラーポリシー蒸留では、複数のエキスパート (エージェントごとに 1 つ) を訓練し、その知識をエージェント ID や形態パラメータを入力とする単一の学生ネットワークに蒸留する。この学生は、身体記述子 (アームとドローンをエンコードするベクトルなど) に条件付けられた一般化ポリシーを学習する [1]。これは効果的に行動間を補間し、記述子を調整することで未見の形態にも外挿できる (「形態学的転移」[1])。最近の研究である AnyMorph はこの方向を取り、潜在的な形態エンコーディングを推論しながら、様々なエージェント構造に適用可能な普遍的なポリシーを学習した [17]。

異種センサーモダリティを組み合わせるために、トランスフォーマーベースのアーキテクチャが注目を集めている。トランスフォーマーは複数の入力タイプ (視覚、触覚など) に注意を払い、クロスモーダル相関を学習できる。例えば、Visuo-Tactile Transformer (VTT) アーキテクチャは、視覚的および力/触覚シーケンスをセルフアテンションとクロスアテンションで並行して処理し、各モダリティの最も関連する特徴に焦点を当てることができる [18]。このようなモデルをマルチモーダルシミュレーション (例: ロボットグリッパーが花の蕾を感じて見る) で訓練することで、結果として得られる潜在表現は手がかりを統合する。これらの共有潜在表現は転移可能である: 同じタイプのモダリティを持つ異なるロボットがそのデータをトランスフォーマーに入力し、意味のある状態推定やアクションコマンドを得ることができる。主要な理論的洞察は、各ロボットの生のセンサー空間は異なるが、基礎となるタスク関連情報 (花の位置、接触達成など) は共有されているということである。抽象的なタスク固有の表現 (「ターゲットの花が授粉されたかどうか」など) を学習することで、知識が転移可能になる。マルチビューシミュレーション環境 (同じシナリオが異なるエージェントによって同時に観察される) は、モダリティを整合するための同時データを提供することで、このよう

な表現の訓練に大いに役立つ。

1.4. 転移のための強化学習、模倣学習、シミュレーション

ロボット間の行動転移は、強化学習 (RL) と模倣学習 (IL) のフレームワークと密接に関連している。深層強化学習 (RL) では、ポリシーはしばしばロボットごとにゼロから学習されるが、転移学習は関連するマルコフ決定過程に対する別のロボットの解決策からあるロボットのポリシーをブートストラップできる。理論的には、共通の報酬関数 (例: 1 分あたりに授粉された花の数) と共通の状態表現 (上記で議論したように) を定義すれば、その報酬を最大化しようとする異なるロボットは経験を共有できる。強力な概念の一つはフェデレーテッド強化学習であり、複数のロボットがそれぞれローカル RL 更新 (例: 自身の環境や農場で) を実行し、定期的にポリシーパラメータをグローバルモデルに集約する。ロボットが異なる場合でも (異種エージェント) フェデレーテッドスキームはパラメータの異なるサブセットやモデル構造を処理するように設計できる [19][20]。例えば、いくつかの農場ロボット (車輪型や飛行型) がそれぞれ少し異なる果樹園で花を見つけて授粉することを学習している状況を想像してみよう。フェデレーテッド学習を使用すると、各ロボットは生のセンサーデータをプライベートに保ちながら、グローバルポリシーを維持する中央サーバーに勾配更新を共有する。時間とともに、このグローバルポリシーは環境や身体全体で効果的な戦略をエンコードし、それがすべてのロボットに送り返される [19][20]。この協調学習はサンプル効率と一般化を向上させ、フェデレーテッドロボットが孤立した訓練よりも優れたパフォーマンスを示すシミュレーション研究で見られている。理論的に重要な側面は、異種性にもかかわらず収束を確保することである: 技術には信頼による更新の重み付け (特定のロボットのデータが非常に異なる場合、その更新の重みを低くする) や、グローバル更新後の迅速な特殊化を可能にするメタ学習の使用が含まれる。

模倣学習 (IL) 設定では、転移は共有デモンストレーションを通じてアプローチされる。タスク (芽の剪定など) のデモンストレーションがエキスパートエージェント (おそらく人間や参照ロボット) によって行われた場合、行動クローニングや逆 RL を通じてこれらを模倣するように様々なロボットエージェントを訓練できる。フェデレーテッド模倣学習の概念はこれを拡張する: Liu らは、複数のロボットがローカルでどのように模倣するかをモデルをアップロードし、クラウドサーバーがこれらを融合して各ロボットの学習を導くために送り返される優れたポリシーを作成するフレームワークを導入した [21]。これは異なる身体からのスキルをクラウドソーシングするようなものである。異種知識融合により、例えば、ロボットアームは四足歩行ロボットが授粉中にどのようにバランスを取ったかから恩恵を受けることができる。実行が異なっても、学習したモデルを補完的な方法で組み合わせることで可能になる [21]。

シミュレーションは転移学習研究において不可欠な役割を果たす。マルチエージェントシミュレーション環境は、異なるロボットモデルを同じ仮想果樹園で

ホストし、知識共有に関する制御された実験を可能にする。ヒューマノイドと四足歩行ロボットの両方がリングを掴むことを学習し、その後ヒューマノイドのポリシーを四足歩行ロボットに転移しようとすることができる。ドメインランダム化 (様々なテクスチャ、照明、物理特性でのトレーニング) のような技術は、学習したポリシーが高レベルの特徴に焦点を当て、実際の条件や他のロボットに転移された場合でも堅牢であることを確保するのに役立つ [1][22]。サイクルー貫性のある生成手法もシミュレーションから実際の転移に適用されている: 例えば、RL-CycleGAN はシミュレーションされたカメラ画像を基礎となる状態を保持しながら現実的に見えるように変換し、ポリシーが事実上「実際の」画像で訓練できるようにする [23]。同様に、RetinaGAN は知覚転移を改善するために物体固有のリアリズム (果物を実際のものと同一に見せるなど) に焦点を当てている [24]。これらは主にシミュリアル視覚ドメインシフトに対処するものですが、同じ考え方が異なるロボットの視点や形態間のシム-シムシフトにも適用できる。本質的に、中間ニューラルネットワーク (GAN) が二つのドメイン間の観測や行動を調整し、明示的なマッピングなしでポリシー転移を容易にする。

1.5. ロボット間の協調学習とフェデレーテッド学習

果樹園環境では、異なるタイプの複数のロボットと一緒に操作される可能性がある - ドローンが調査し、アームが摘み取り、脚付きロボットが物資を運ぶなど。協調学習により、これらのエージェントは共有経験を通じて互いのポリシーを改善できる。一つのアプローチは役割特化を伴うマルチエージェント RL であり、異種エージェントが同時に学習し、共通の報酬 (例: 全体的な収穫量) を共有する。訓練を通じて、彼らは暗黙的に戦略を交換するかもしれない: アームはドローンのアプローチ角度を観察することで、より良い摘み取り戦略を学ぶかもしれない。しかし、エージェントが同時に学習していない場合や異なる目標を持つ場合は、明示的な転移が必要になることが多い。前述のように、フェデレーテッド学習は単一のポリシーを共有するだけでなく、多様性から恩恵を受けるアンサンブルモデルを訓練するためにも適用できる。例えば、中央モデルは、どのエージェントタイプが特定の状況で処理すべきかを決定することを学習できる (高い枝の果物はドローンに任せるなど)。これは古典的な転移を超えてメタ学習に進み、システムはポリシーに対するポリシーを学習する。

もう一つの興味深いパラダイムはロボット間の教師-生徒フレームワークである。例えば、四足歩行ロボットが試行錯誤を通じて、木に到達するために厚い果樹園の草をどのようにナビゲートするか (車輪型ロボットが苦戦するスキル) を学習したとする。四足歩行ロボットは軌道を生成することで教師として機能し、車輪型ロボットはオフポリシー学習を通じてそれを (調整して) 模倣できる。車輪型ロボットはこのように、四足歩行ロボットの地形交渉の洞察を誘導された練習を通じて獲得する。このような形態間教育はシミュレーションで探求されている: あるエージェントのポリシー出力を、別のエージェントの学習プロセスのための行動提案や形成報酬として使用できる [10]。

これは形態間のポリシー蒸留の一形態である。

理論的観点からは、これらすべての転移技術は、異なるエージェント間の違いを考慮しながら情報の再利用を最大化することを目指している。これにより、サンプル複雑性が減少する - 実世界の試行（木に登る、繊細な花の近くを飛ぶなど）が高コストでリスクを伴う農業ロボティクスでは重要である。事前知識（他のロボットやシミュレーションから）を活用することで、新しいロボットは果樹園での試行をはるかに少なくして能力を獲得できる。文献によれば、ロボットがタスクの共通表現（整合された状態/行動/効果空間を通じて）を共有し、更新を通信するマルチモーダル、マルチエージェントアプローチが、迅速な学習に最も効果的である [19][20]。このような手法の成功は、互いから、そして人間の専門家から継続的に学習する果樹園での異種ロボットチームを予見させる。時間とともに剪定、授粉、収穫などのタスクでのパフォーマンスを向上させる。

2. 果物栽培ロボティクスのための物体状態推定

ロボットが授粉、収穫、芽間引きなどを果樹園で実行するためには、重要な対象物（花、果実、芽）の正確な状態推定が不可欠である。ロボットはこれらの対象物の存在だけでなく、その正確な姿勢、状態（例：授粉されたかどうか）、操作後の変化（例：芽が正常に除去されたか、果実が取り外されたか）を認識する必要がある。本章では、マルチモーダル設定 - 視覚（RGB、深度）と触覚の組み合わせ - における先進的な状態認識技術と、トランスフォーマーや生成モデルのような現代的アプローチがオクルージョンなどの実世界の課題下での状態推定をどのように改善するかを探求する。また、花が授粉されたかどうか、芽が間引き中に適切に押しつぶされたかどうかなどの特殊な状態分類タスクと、これらがロボットの学習プロセス（例：報酬やサブタスクトリガーとして模倣学習で）にどのように統合できるかについても議論する。

2.1. オクルージョン下での把持物体の姿勢推定（視触覚融合）

ロボットによる収穫や授粉では、エンドエフェクタ（グリッパーやツール）が相互作用の重要な瞬間にターゲット物体を遮蔽することがよくある - 例えば、グリッパーが果実や花を掴む際にそれを覆い、ロボットのカメラビューをブロックする。このような場合に物体の6自由度姿勢（位置と向き）を確実に推定するために、ロボットはRGB-Dと触覚センサーの融合を活用する。視覚は接触前のグローバル情報を提供し、触覚は把持中の局所的な形状と接触データを提供する。Bimboらによる古典的アプローチは、視覚追跡と触覚フィードバックを融合して操作中の物体の姿勢追跡を改善した [25]。視覚センシングは手の中の物体を位置特定できるが、物体が滑ったり部分的に見えなくなったりした場合、触覚センサーの読み取り値（圧力分布、接触点）が姿勢推定を更新する。YuとRodriguez（2018）はこれをSLAMのような状態推定問題として定式化した：彼らは増分平滑化フレームワーク（iSAM）を使用し、触覚接触をランドマーク観測に類似したものとして扱い、カメラと触覚データを継続的にブレンドして物体

が視界から消えても姿勢を追跡した [26]。触覚読み取り値はグリッパーに対して物体を固定し、最後に知られた視覚姿勢はそれをグローバルに固定し、特にオクルージョン後は視覚のみよりも優れた融合推定を生み出す。

現代の学習ベースの手法はこれをさらに進める。Dikhaleら（2022）は、視覚と触覚の入力を組み合わせから直接手の中の物体の姿勢を回帰するディープネットワークを訓練した [27]。彼らの視触覚姿勢推定器は、RGB画像（物体の一部とグリッパーのみを示すかもしれない）とグリッパーからの高次元触覚センサー読み取り値を取り、物体の6D姿勢を出力する。このエンドツーエンドアプローチは、触覚パターンと物体の向きの間の暗黙的な相関関係を学習することで、視覚のみのCNNよりも大幅なオクルージョン（ロボットハンドが物体のほとんどを覆うケース）をより適切に処理することが示された [28]。しかし、純粋に学習されたアプローチは訓練物体を超えて一般化することが難しく、既知の幾何学を明示的に使用しないという制限が指摘されている。姿勢をブラックボックス回帰問題として扱う [28]。より多くの構造を組み込むために、最近のアプローチでは中間ステップとして明示的な形状再構成を実行する。例えば、Liら（2023）はViHOPEを提案した。これは最初に視触覚入力を使用して部分的に見える物体の3D形状を完成させ（条件付き生成モデルを介して）、次にこの再構成された形状から姿勢を推定する [29][30]。物体の遮蔽された部分を「想像」することを学習することで、システムは完全な形状モデル上に姿勢を固定でき、精度が大幅に向上する - 本質的に、ロボットが片側の触覚と他側の視覚に基づいて果実の隠れた側面を推測できれば、空間内でより適切に配置できる。ViHOPEの形状完成のためのGANの使用は、直接回帰と比較して姿勢推定を大幅に改善した [29][30]。これは状態推定における生成モデリングの統合の利点を強調している。

もう一つの問題は操作中の推定一貫性である。ロボットハンドが把持した物体を動かすとき、接触点が変わると従来の姿勢フィルターは追跡を失う可能性がある。これに対処するために、Muraliら（2022）は能動的視触覚知覚戦略を導入した：ロボットは物体の観察可能性を意図的に改善するために、わずかにグリッパや視点を調整する [31]。探索的行動（より多くの物体表面を感じるために指をスライドさせたり、遮蔽された部分をカメラビューに覗かせるために手首を動かしたりする）を計画することで、ロボットは姿勢推定を能動的推論問題として扱う。密集した状況での彼らの方法は、より良い組み合わせの視覚-触覚データを得るために物体を突いたり移動させたりすることを含み、密集したシナリオでより正確な姿勢をもたらした [31]。このインタラクティブ知覚の概念は、状態が不確かな場合、ロボットが物体をより多く露出させようとすることを保証する（例えば、カメラが茎の付着を見ることができるよう果実の塊をわずかに持ち上げる）。これは強化学習に結びつく：ロボットはタスク実行だけでなく、情報収集のためのポリシーも持つことができる。

さらに、トランスフォーマーベースの融合が視触覚姿勢追跡に適用されている。Chenら（2022）は、視覚

トークンと触覚読み取り値が互いに注意を払い、無関係な特徴をフィルタリングし、接触点と視覚的エッジに焦点を当てる Visuo-Tactile Transformer (VTT) を開発した [18]。トランスフォーマーのセルフアテンションとクロスアテンションメカニズムは、異なるセンサーモダリティを自然に処理し、下流の状態推定や計画に使用される潜在表現を生成する。果樹園ロボティクスの文脈では、グリッパーが果実を閉じる際のカメラ画像のシーケンスと、グリッパーの指からの対応する触覚読み取り値のシーケンスを受け取るトランスフォーマーを想像できる。特定の触覚センサーがトリガーされると、指が果実の反対側にあるはずであり、したがって果実はカメラがもはや見ることができなくても、手に対して特定の方向に向いているはずであることを学習できる。このような学習された相関関係は、物体の一部が視覚で観察されない場合の姿勢推定の堅牢性を向上させる。

2.2. 花の授粉状態認識と分類

授粉は繊細なタスクである：ロボットはどの花が授粉されたか（自身によるか自然に）、どの花がまだ注意を必要とするかを判断する必要がある。視覚の手がかりは微妙である - 授粉された花は色のわずかな変化、しおれた花弁、または単に柱頭上の花粉の存在を示すかもしれない。授粉完了の検出はオープンな課題である。一部の研究ではこれを花の画像に対する二値分類として扱っている：花の接写が与えられたとき、「授粉された」対「授粉されていない」と分類する。これは柱頭が暗くなったか花粉粒が見えるかを確認するために色分割を使用するかもしれない。例えば、マクロ写真と画像分析を組み合わせて、柱頭上の花粉粒を定量化できる [32]。原則として、ロボットは授粉試行後に花の高解像度画像を撮り、エコロジー研究で ImageJ で実証されているように、画像処理を使用して花粉沈着をカウントできる [32]。しかし、照明や花粉の微小なスケールのため、フィールドでリアルタイムでこれを行うことは困難である。

代わりに、実用的なアプローチはしばしばプロセス知識に依存している：ロボットの授粉ツールが花の生殖器官と適切に接触した場合、授粉を想定する。最近の授粉ロボット設計では、カメラがロッドを花の中心に触れるように誘導する。接触が行われると（閾値距離内で）、システムはその花を授粉されたとマークする [33][34]。しかし、二重授粉を避けたり一部を見逃したりしないためには、より多くの知能が必要である。ディープラーニングは、新鮮な花と授粉された花の間の微妙な違いを検出するのに役立つ。Singh ら（2024）は支援授粉のためのトマトの花と蕾を検出するための CNN アプローチを開発した。これは、開いた（おそらく授粉された）対閉じた花を区別することで、授粉状態を識別するように拡張できる可能性がある [35]。彼らのシステムは YOLOv5 検出器を使用して、温室画像内の開いた花と未開花の蕾の両方を見つけ、これらのクラスを識別する際に約 82 % の mAP を達成した [35]。このような検出器が授粉直後の花の画像（おそらく花弁がわずかに変色したり雄しべが乱れたりする）で訓練されれば、それらを別のカテゴリとして分類できる。リンゴやキウイのような果樹園では、授粉の成功は果実の形成が始まる数日後にのみ確認できる

かもしれない - ロボットが行動するには遅すぎる。そこで、授粉を多段階状態として扱う代替案がある：（1）花の準備完了、（2）授粉中（ロボットが行動中）、（3）授粉完了（完了としてマーク）。ロボットの視覚システムは、その花を訪問したかどうかによって、状態（1）対（3）を検出できる。おそらく物理的なマーカーを使用する（一部のシステムでは授粉時に食品安全染料の小さな点を噴霧してマークし、コンピュータビジョンが容易に見ることができる）。

それでも、研究の完全性のために、授粉の直接的な視覚検出を考慮する：マルチスペクトルイメージングを使用すると、授粉後の花の化学的变化（例：花粉の UV 蛍光）を明らかにできる。機械学習モデルは、授粉された花と授粉されていない花のスペクトル署名で訓練できる。これは推測的であり、授粉状態検出に直接取り組む出版物はほとんどない。関連する問題は、キウイフルーツのような雌雄異株作物における雄花と雌花の区別である - Tang ら（2024）によるディープラーニングモデルは、色と形状の手がかりを使用して、（授粉が必要な）アクチニディア雌花を識別した [36]。ロボットはこのような分類器を活用して、雌花に努力を集中させ、柱頭がもはや生存可能でない（過熟または既に受精した）場合を検出できる可能性がある。

要約すると、現在のロボティクスソリューションは授粉のための間接的な状態推定に依存している：授粉ツールによって物理的に接触された花を追跡する。視覚は花の検出と位置特定により多く使用される。例えば、Hulens ら（2022）は花の位置を検出するために Inception-V3 CNN を使用し、その後授粉ブラシを各花に誘導する授粉ロボットを説明している [33]。彼らは成功の検証が制限であることを指摘しているが、適切な接触が授粉された花に等しいという仮定がある。将来の研究では、実際に柱頭上の花粉を見るために、エンドエフェクタに微小な顕微鏡カメラや分光法を統合するかもしれない。それまでは、ロボットは保守的な戦略を採用できる：検出されたすべての花に授粉を試み、環境センサー（ミツバチの活動のモニタリングや以前に記録された授粉された花のデータなど）を使用して冗長な作業を避ける。

2.3. 芽間引き：サブタスク状態推定と結果検出

芽間引きは、残りの果実がより大きく成長できるように、過剰な芽（または若い花）を除去することを含む。芽間引きを実行するロボットは、ターゲットの芽を識別し、各ターゲットが正常に「間引かれた」（通常は押しつぶすか除去することによって）ことを確認する必要がある。ここでの状態推定は、物体検出（芽を見つける）と操作後の状態認識（芽が除去されたか）の両方を含む。視覚的には、無傷の芽と押しつぶされた芽は形状/テクスチャによって区別できる - 健康な芽は丸々として無傷であり、正常に間引かれた芽は壊れているか部分的に欠けているかもしれない。研究者たちはディープラーニングを使用して堅牢な芽と花の検出器を開発した。Sahu と He（2023）は、YOLOv4 が果樹園画像内のリンゴの花芽を非常に高い精度（mAP₉₅ %）で検出でき、そのタスクでより新しい YOLOv5/7 よりも優れていることを示した [37]。このような検出器は、ターゲットとする芽の位置をリアルタイムで提供できる。実際、Khanal ら（2023）は

YOLOv5を使用した初期段階のリンゴの花検出（間引きと授粉のための）視覚システムを提案し、間引き決定のために検出を集めて密集した花の塊を識別している[38]。これらの進歩により、ロボットは芽がどこにあるか、そして塊にいくつあるかを確実に認識できる。

芽が識別され、ロボットがそれを挟む（例：グリッパーの指のペアやカッターで）と、芽が適切に除去されたかどうかを判断することが重要になる。これは異常検出の一形態である：期待される結果は芽が消えているか目に見えて損傷していること。まだ無傷であれば、間引き行動は失敗し、繰り返す必要があるかもしれない。単純な視覚ソリューションは「前と後」の画像を撮って比較することである。前に存在していた芽が後の画像で欠けていれば、成功である。これには芽の位置の正確な整合または追跡が必要である。現代的なアプローチでは、間引き前後のエリア内の芽をカウントするために物体検出器を使用できる - カウントの減少は成功を示す。しかし、芽がその場で押しつぶされた場合（落下しない）、まだ見えるかもしれないが、外観が変わっている。様々な状態（健康対押しつぶされた）の芽の画像で訓練された分類器が結果にラベルを付けることができる。これは品質検査に類似している：例えば、CNNはロボットの行動後の芽サイトの接写から「芽無傷」または「芽押しつぶし」を出力するように訓練できる。特徴には色の変化（押しつぶされた芽は内部組織や液体を露出させ、色を変える可能性がある）、形状（無傷の芽は丸みを帯びており、押しつぶされた芽は不規則または平らに見える可能性がある）、サイズ（破片が散らばる可能性がある）が含まれる。グリッパーの力センサーも貢献する：挟む間の力の特定のスパイクまたは低下は、芽が正常に折れたことを示す可能性がある。それを統合することで、システムは視覚のみに頼ることなく、成功の高い信頼性を持つことができる。

模倣または強化学習フレームワークでは、これらのサブタスク状態（例：「芽を把握した」、「芽を除去した」）はしばしばチェックポイントや報酬として機能する。ロボットは各ステップがラベル付けされたデモンストレーション（芽に近づく、把握する、除去する）で訓練され、そのシーケンスを複製する必要がある。ロボットが状態（芽がまだ存在するかどうか）を推定できれば、次の芽に進むか行動を再試行するかを決定できる。強化学習では、サブタスク（芽の除去）の達成に対する明示的な報酬を含めることができ、これにはロボットの知覚が達成を正しく識別する必要がある。

注目すべきアプローチは、操作タスクにおける一連の状態をモデル化するためのシーケンスモデルやトランスフォーマーの使用である。芽間引きの場合、トランスフォーマーは視覚フレームと固有受容データのシーケンスを取り込み、行動に条件付けられた次の状態（例：芽が消える）を予測することを学習できる。結果の状態が予測と一致しない場合（芽が予想外にまだそこにある）、モデルは何か間違っていることを知る。この種の予測モニタリングは、状態遷移の深層生成モデル（変分再帰ネットワークやトランスフォーマーなど）を使用して、暗黙的に成功を測定する。

同時に、マルチモーダル学習もここでの状態検出を

改善できる。ロボットの芽に対するカメラビューが挟む行為中に一時的に自身の腕によって遮られるシナリオを考えてみよう（再びオクルージョン！）。グリッパーの触覚センサーは芽のテクスチャや破裂イベントを感知できる。視覚と触覚を組み合わせることで、ロボットはカメラがそれを明確に見ていなくても、芽が確かに押しつぶされたことをより確信できる。VTTに類似したトランスフォーマーモデルを採用できる：接触前の芽を記述する視覚トークンと、接触中の触覚トークンが共同で「成功」または「失敗」を出力する分類器に情報を提供する。触覚センサーが特徴的なクランチを感じ、その後芽の抵抗がなくなった場合、それは成功の強い指標である。

研究では、「芽押しつぶし状態」に関する明示的な研究は少ないが、類似の問題は存在する：例えば、ケーブルが適切に差し込まれたか（力/視覚を通じて）、またはボトルのキャップが正常に外れたかの確認。これらのドメインからの技術は転用できる。深層生成視覚モデルを使用することもできる：拡散モデルは「間引き前」画像と行動が与えられた「間引き後」画像を想像するように訓練できる。実際の後画像がこれから大きく逸脱する場合（例えば芽がまだ存在する）、モデルはそれにフラグを立てる。

重要なことに、正確な状態推定により、ロボットはこれらのタスクで閉ループ制御を実行できる。収穫では、果実の茎を見て切れたかどうかを推定できるロボットは、もっと引っ張るべきか止めるべきかを知る。同様に、芽間引きでは、芽の状態を知ること、ロボットは自律的に再試行するか次に進むかを決定できる。この自律性は堅牢な状態知覚に基づいている。フィールド試験では、Bhattaraiら（2024）は花の塊を識別するために視覚を使用し、機械的スピナーでいくつかの花を除去する花間引きロボットをテストした[39,?].彼らのシステムは各花の除去を完全に検証していなかったが、後で全体的な間引き率を測定した。視覚と学習技術が向上するにつれて、将来のシステムはおそらくオンライン検証を組み込むだろう：ターゲットにされた各芽は知覚システムから成功/失敗のラベルを受け取り、即時の修正を可能にする。これにより、繊細な農業タスクでの効率性と信頼性が向上する。

2.4. 先進技術：状態推定のためのトランスフォーマーと生成モデル

精度を最大化するために、研究者たちは先進的な機械学習モデルに目を向けている。マルチモーダルトランスフォーマー（前述のように）は異なるセンサーストリームの統一処理を可能にする。また、状態追跡のための時間データも組み込むことができる。例えば、トランスフォーマーはロボットが果実に近づく一連の画像、把握の瞬間（おそらく突然の加速や触覚イベントによって識別される）、そして分離後の後続画像を取ることができる。時間にわたって注意を払うことで、最終状態（果実が除去されたかどうか）を出力することを学習できる。これは効果的に状態オブザーバーを学習することであり、時間を考慮した学習されたカルマンフィルタに類似しているが、非線形の注意ベースのマッピングを持つ。トランスフォーマーは長距離依存関係をエンコードする能力を持っているため、結果が早期と後期の手がかりの組み合わせによって示さ

れる（例えば、「フレーム 10 でグリッパーが閉じるのを見て、フレーム 15 までに果実があった場所の色が変わっていれば、果実は消えている」）シナリオで有用である。

もう一つの有望な道は、変分オートエンコーダー（VAE）や拡散モデルのような深層生成モデルを使用して、物体の状態とその不確実性を表現することである。VAE は物体の外観の潜在空間を学習できる。例えば、様々な数の花を持つ花の塊の多くの画像で訓練できる。画像が与えられると、VAE はそれを潜在空間に投影でき、おそらく「完全に間引かれた」対「間引かれていない」塊に対応する潜在変数の間に明確な分離が見られるだろう。芽が除去された場合のシーンの外観を生成し、実際の画像と比較して芽の存在を推測することができる。画像を反復的に洗練する拡散モデルは、初期状態（芽あり）を取り、芽をインペイントアウトすることで除去をシミュレートできる。ロボットの行動が成功した場合、行動後の観察は拡散モデルの予測と一致するはずであり、そうでなければ不一致がある。これらのアイデアはこのドメインではまだ大部分が理論的だが、タスクが正しく実行されたときに環境がどのように変化するかの良い生成モデルがあれば、現実がそのモデルから逸脱したときにエラーを捉えることができるという概念に基づいている。

最後に、膨大なデータで事前訓練されたトランスフォーマーベースの大規模モデル（CLIP のような視覚-言語モデルや、ビデオとテキストを組み合わせた最近のロボティクスの研究に類似）は、微妙な状態を認識するために微調整できる。多様な画像で事前訓練されたモデルは、少しの集中訓練を与えられれば、「花粉のある花」対「花粉のない花」をすでに理解しているかもしれない。また、物体を表現するためのニューラルフィールド（NeRF）に関する研究もあり、これが触覚入力と組み合わせられると（例：2024 年の NeuralFeel システム）、操作中に形状と姿勢を同時に推定できる。これらのニューラルフィールドは、継続的に更新される内部状態表現として効果的に機能する。芽が除去された場合、シーンのニューラルフィールド表現は適切に更新されるはずである（芽の密度が消える）。その変化を検出することで成功が確認される。

要約すると、果物栽培ロボティクスにおける正確な状態推定は、複数のモダリティ（視覚、深度、触覚、力）の融合、現代的な深層学習（検出のための CNN、融合のためのトランスフォーマー、完成のための GAN/VAE）の活用、そしてオクルージョンを軽減するための能動的戦略によって達成されている。これらの進歩により、ロボットは単に「行う」だけでなく、「行ったことを知る」ことも確保され、自律操作のためのループが閉じられる。

REFERENCES

- [1] R. Zhu, et al., “Cross Domain Policy Transfer with Effect Cycle-Consistency,” *arXiv:2403.02018*, 2024. <https://arxiv.org/html/2403.02018v1>
- [2] R. Zhu, et al., “To enable policy transfer across robots with different embodi-

ments,” *arXiv:2403.02018*, 2024. <https://arxiv.org/html/2403.02018v1>

- [3] J. Hanna and P. Stone, “Grounded action transformation for robot learning in simulation,” *AAAI*, 2017.
- [4] T. Wang, et al., “NerveNet: Learning structured policy with graph neural networks,” *U. Toronto Tech Report*, 2017.
- [5] W. Huang, et al., “One Policy to Control Them All: Shared Modular Policies for Agent-Agnostic Control,” *ICML*, 2020. <https://wenlong.page/modular-rl/>
- [6] W. Huang, et al., “This work takes a step towards learning a single policy that can generalize across agents,” *ICML*, 2020. <https://wenlong.page/modular-rl/>
- [7] P. Sharma, et al., “Third-person visual imitation learning via decoupled hierarchical controller,” *NeurIPS*, 2019.
- [8] H. Niu, et al., “A Comprehensive Survey of Cross-Domain Policy Transfer for Embodied Agents,” *IJCAI Survey Track*, 2024. <https://www.ijcai.org/proceedings/2024/0906.pdf>
- [9] Y. Liu, et al., “Imitation from observation: Learning to imitate behaviors from raw video via context translation,” *ICRA*, 2018.
- [10] M. Taylor and P. Stone, “Transfer learning for reinforcement learning domains: A survey,” *JMLR*, 2009.
- [11] S. Desai, et al., “Imitation from observation for transfer learning with dynamics mismatch,” *NeurIPS*, 2020.
- [12] J. Abou-Chakra, et al., “Physically Embodied Gaussian Splatting: Visually learnt and physically grounded 3D representation for robotics,” *Under review*, 2024. <https://embodied-gaussians.github.io/>
- [13] P. Sermanet, et al., “Time-Contrastive Networks: Self-supervised learning from video,” *ICRA*, 2018.
- [14] Y. Tian, et al., “What makes for good views for contrastive learning?” *NeurIPS*, 2020.
- [15] Q. Zhang, et al., “Learning cross-domain correspondence for control with dynamics cycle-consistency,” *arXiv:2012.09811*, 2020.
- [16] A. Gupta, et al., “Learning invariant feature spaces to transfer skills with reinforcement learning,” *arXiv:1703.02949*, 2017.
- [17] B. Trabucco, et al., “AnyMorph: Learning transferable policies by inferring agent morphology,” *ICML*, 2022.
- [18] Y. Chen, et al., “Visuo-Tactile Transformers for Manipulation,” *CoRL*, 2022. <https://arxiv.org/abs/2210.00121>
- [19] Y. Jiang, et al., “FedRobo: Federated Learning Driven Autonomous Inter Robots Communication For Optimal Chemical

- Sprays,” *arXiv:2408.06382*, 2023. <https://arxiv.org/html/2408.06382v2>
- [20] Y. Jiang, et al., “Federated Learning enables robots to learn collaboratively without sharing raw data,” *arXiv:2408.06382*, 2023. <https://arxiv.org/html/2408.06382v2>
- [21] B. Liu, et al., “Federated Imitation Learning: Heterogeneous knowledge fusion for cloud robotics,” *arXiv:1912.12204*, 2020. <https://arxiv.org/abs/1912.12204>
- [22] R. Zhu, et al., “Domain randomization for robust transfer learning,” *arXiv:2403.02018*, 2024. <https://arxiv.org/html/2403.02018v1>
- [23] K. Rao, et al., “RL-CycleGAN: Reinforcement Learning aware Simulation-to-Real,” *CVPR*, 2020.
- [24] D. Ho, et al., “RetinaGAN: An object-aware approach to sim-to-real transfer,” *ICRA*, 2021.
- [25] J. Bimbo, et al., “Object pose estimation and tracking by fusing visual and tactile information,” *IEEE MFI*, 2012.
- [26] K.T. Yu and A. Rodriguez, “Realtime State Estimation with Tactile and Visual Sensing for Planar Manipulation,” *IEEE RA-L/ROS*, 2018. <https://arxiv.org/abs/1709.09694>
- [27] S. Dikhale, et al., “VisuoTactile 6D Pose Estimation of an In-Hand Object Using Vision and Tactile Sensor Data,” *IEEE RA-L*, 2022.
- [28] A. Villalonga, et al., “Visuotactile policy learning for generalizable object reorientation,” *CoRL*, 2021.
- [29] H. Li, et al., “ViHOPE: Visuotactile In-Hand Object 6D Pose Estimation with Shape Completion,” *NeurIPS Workshop*, 2023. https://www.touchprocessing.org/2023/camera_ready/camera_ready_11.pdf
- [30] H. Li, et al., “In this paper, we present ViHOPE, a novel approach for in-hand object 6D pose estimation,” *NeurIPS Workshop*, 2023. https://www.touchprocessing.org/2023/camera_ready/camera_ready_11.pdf
- [31] P.K. Murali, et al., “Active visuo-tactile interactive robotic perception for accurate object pose estimation in dense clutter,” *IEEE RA-L*, 2022.
- [32] C. MacGregor and A. Scott-Brown, “Quantifying pollen deposition with macro photography and image analysis,” *J. Pollination Ecol.*, 2018. <https://pollinationecology.org/index.php/jpe/article/view/410>
- [33] D. Hulens, et al., “Autonomous Visual Navigation for a Flower Pollination Drone,” *Machines*, vol. 10, no. 5, p. 364, 2022. <https://www.mdpi.com/2075-1702/10/5/364>
- [34] D. Hulens, et al., “pollination of crops and plants is its size and maneuverability,” *Machines*, vol. 10, no. 5, p. 364, 2022. <https://www.mdpi.com/2075-1702/10/5/364>
- [35] R. Singh, et al., “Deep learning approach for detecting tomato flowers and buds in greenhouses on a gantry robot,” *Sci. Rep.*, vol. 14, p. 20552, 2024.
- [36] Y. Tang, et al., “Deep learning based approach for Actinidia flower detection and pollen analysis,” *Sci. Rep.*, 2024. https://www.researchgate.net/publication/369381897_Artificial_pollination_of_kiwifruit_Actinidia_chinensis_Planch_var_chinensis_Ericales_Actinidiaceae_results_in_greater_fruit_set_compared_to_flowers_pollinated_by_managed_bees_Apis_mellifera_L_Hymenop
- [37] R. Sahu and L. He, “Real-Time Bud Detection Using YOLOv4 for Automatic Apple Flower Bud Thinning,” *ASABE Annual Intl. Meeting*, 2023. <https://elibrary.asabe.org/azdez.asp?JID=5&AID=54133&CID=oma2023&T=2>
- [38] S.R. Khanal, et al., “Machine Vision System for Early-stage Apple Flowers and Flower Clusters Detection for Precision Thinning and Pollination,” *arXiv:2304.09351*, 2023. <https://arxiv.org/abs/2304.09351>
- [39] M. Bhattarai, et al., “Robotic System for Precision Blossom Thinning in Apples,” *WSU Tree Fruit*, 2024. <https://treefruit.wsu.edu/article/robotic-system-for-precision-blossom-thinning>
- [40] M. Bhattarai, et al., “similar to a weed cutter, triggered and performs the thinning,” *WSU Tree Fruit*, 2024. <https://treefruit.wsu.edu/article/robotic-system-for-precision-blossom-thinning>