

Lightweight and High Performance Model for Visual Object Tracking based on Deep Learning

黎, 一丁

<https://hdl.handle.net/2324/7363817>

出版情報 : Kyushu University, 2024, 博士 (工学), 課程博士
バージョン :
権利関係 :



氏 名 : 黎 一丁 (レイ イチチョウ)

論 文 名 : Lightweight and High Performance Model for Visual Object Tracking
based on Deep Learning
(深層学習に基づく物体追跡のための軽量かつ高性能モデル
に関する研究)

区 分 : 甲

論 文 内 容 の 要 旨

近年、視覚的な物体追跡は、カメラドローンや運転支援などの実世界のアプリケーションにおいて重要性が増している。Siamese トラッカーは、ニューラルネットワークを活用した高精度な追跡技術として注目されているが、複雑な設計がリソース制約のあるデバイスへの展開を困難にしている。また、対象物体の遮蔽や高速移動などの実環境特有の課題による性能低下も克服すべき重要な問題である。そこで本論文では、これらの課題を解決するために、軽量かつ高性能な物体追跡モデルの開発を目指す。

まず本論文では、小型で軽量の Siamese トラッカー SDST (Specialized teachers Distilled Siamese Tracker) を提案する。本手法では、基本的な追跡性能に優れたトラッカーと、厳しい環境下で高性能を発揮するトラッカーの 2 種類を教師モデルとして利用し、知識蒸留により小規模な生徒モデルを訓練する。これにより、モデルの軽量化と高性能化の両立を実現する。ベンチマークデータセット (LaSOT Dataset) を用いた実験では、最大 8 倍のモデル圧縮率と 252FPS の高速処理を達成しつつ、全体的に高い追跡性能を維持できることを確認した。

次に、遮蔽に対して頑健なトラッカー AMKD (Adaptive Multi-teacher Knowledge Distillation) を提案する。物体の可視比率が低い状況は追跡精度に大きな影響を与えるため、本手法では低可視比率の状況下で訓練された複数の教師モデルを事前に用意し、それらを生徒モデルへの知識蒸留のアシスタントに利用する。その際に、適応選択機構 (ASM) を組み込むことで効率的な知識蒸留を実現する。GOT-10K データセットを利用した実験では、9 倍の圧縮率と 6 倍の速度向上を達成し、低可視比率シナリオにおける追跡精度も向上することを確認した。

さらに、より高速かつ低メモリコストのトラッカー RST (RepSiamese Tracker) を提案する。本手法は、精度を犠牲にすることなくモデルの圧縮を行える RepVGG に着想を得て、Siamese トラッカーの圧縮に応用する。Re-parametrization 手法により、モデル訓練時に使用される複雑なモジュールと等価な単純構造に変換し、モデルサイズを圧縮する。複数の RST モデルを設計し、CPU や GPU、利用可能メモリサイズなど異なるハードウェア要件に対応する。VOT2018 と LaSOT のデータセットを用いた実験では、CPU 上で 69FPS のリアルタイム処理を 16.4MB という極めて低いメモリコストで物体追跡を行えることを確認した。

以上、本研究は物体追跡問題におけるリソース制約や実環境での課題を克服しつつ、高速かつ高性能な物体追跡モデルを低コストで実現する新たな方向性を示すものである。