

CycleGAN による教師なしドメイン適応を用いた運転者視点画像に対する物体検出

Object Detection for Driver's Field of View Using Unsupervised Domain Adaptation with CycleGAN

高橋 暖人[†]猿田 和樹[‡]陳 国躍[‡]寺田 裕樹[‡]

Haruto Takahashi

Kazuki Saruta

Guoyue Chen

Yuki Terata

1. はじめに

交通事故の発生要因のうちドライバーの認知ミスが多くを占めていることから、ドライバーの認知能力向上を目的とした運転者教育システムの研究が行われている。先行研究では、視線計測器により運転者視点映像と視線座標を取得し、映像に対して物体検出を行い、視線座標と重畳することで運転者の注視している物体を判定することで注視行動を分析する仕組みを構築した[1, 2, 3]。これにより、ドライバーが何をどの程度注視していたかを明らかにし、見逃し等のフィードバックを与えることが可能となった。

これらの先行研究の物体検出タスクにおいて、モデルの訓練・検証に用いたデータセットはすべて昼間に撮影された運転者視点画像で構成されている。交通事故は 17 時台～19 時台の薄暮の時間帯に最も多く発生することから、この時間帯の運転時の注視行動のフィードバックが与えられることがドライバーの認知能力向上に寄与すると考えられる。しかし、薄暮および夜間のデータセットを作成する際、物体の境目や詳細が分かりにくいいため、アノテーションが困難である。そこで、本研究では薄暮時の運転者視点画像における物体検出を目的とし、CycleGAN[4]を用いたドメイン適用を行い、その有効性を明らかにする。

2. 提案手法

関連研究として Arruda ら[5] は CycleGAN を用いてドライブレコーダ映像データセットの昼間画像を偽夜間画像に変換し、その画像を使用して Faster R-CNN を用いて実夜間画像に対する車両検出を行い、昼間画像のみで学習させた場合よりも検出性能を向上させた。この研究では検出対象は車両のみである。また、頭部運動を伴う運転者視点画像に対する検証は行われていない。

本研究では、画像生成モデルである CycleGAN を用いてラベル付きの昼間運転者視点画像データセットを偽夜間画像に変換し、アノテーション情報を利用して偽夜間画像データセットを作成する。そして、作成したデータセットを用いて物体検出モデルを訓練させる。これによりアノテーションコストを削減しつつ薄暮画像に対する検出性能の向上を図る。図 1 に昼間画像から偽夜間画像への変換を通じたデータセット作成の流れを示す。



図 1 データセット作成の流れ

[†] 秋田県立大学大学院 Graduate School of Akita Prefectural University

[‡] 秋田県立大学 Akita Prefectural University

3. CycleGAN を用いた昼から夜への画像変換

3.1 CycleGAN の概要

CycleGAN は生成モデル GAN をベースとした画像変換モデルで「変換してから逆変換すると元に戻る」というアイデアのもとに提案された。異なるドメイン間の画像変換を行う 2 つの Generator と、入力された画像がそのドメインであるか判定する 2 つの Discriminator で構成されている。従来の GAN ベースの手法と異なり、訓練の際にペアデータを必要としないという特徴を持つ。

3.2 実験概要

本実験では、運転者視点画像データセットの訓練データ画像を、CycleGAN を用いて偽夜間画像に変換することでドメイン適応を図り、変換後の画像を定性的に評価する。CycleGAN の学習は物体検出の訓練に使用しない昼間画像と夜間画像をそれぞれ 1395 枚ずつ使用した。また、計算量の都合上、変換前の画像は 1920×1080 に対し、変換後の画像は 640×360 とする。

3.3 実験結果

学習済みの Generator で運転者視点画像を昼間画像から偽夜間画像に変換した画像の一部を図 2、図 3 に示す。画像変換を行った結果、車両等の大きなオブジェクトの特徴を残したまま夜間画像に変換される傾向にあることが確認できた。しかし、図 2 で示した画像変換では、画像中央の車両の左上に存在する標識が変換後に消失していた。また、図 3 の例のように変換後の画像に存在しないはずの場所に信号や照明のような光(以下、アーティファクト)が多く生じていた。



図 2 オブジェクトの消失が見られた例



図 3 アーティファクトが見られた例

4. 変換後の画像を用いた物体検出

4.1 実験概要

本実験では、変換後の画像を学習に用いた場合の物体検出性能を評価する。昼間画像 2045 枚で構成されたデータセットと、このデータセットのうち、1000 枚を CycleGAN で偽夜間画像に変換したものを比較する。物体検出モデルは YOLOX[6]および co-dino[7]で、事前学習済みモデルを前述のデータセットでファインチューニングする。その後、昼間画像 395 枚、薄暮画像 100 枚のテストデータに対し物体検出を行い、各データセット間で mAP を比較する。

4.2 実験結果

各データセットを YOLOX で学習させた結果を示す。表 1 に昼間画像に対する物体検出の検出性能を、表 2 に薄暮画像に対する検出性能を、図 4 に各ラベルの薄暮画像に対する mAP を示す。

まず表 1 の昼間画像に対する検出性能を見ると、mAP は顕著な差が見られなかったが、mAP_M はドメイン適応後のデータセットの方が 9.2% 向上している。表 2 の薄暮画像に対する検出性能を見ると、mAP はドメイン適応後のデータセットの方が高くなっており、その差は 4.9% であった。しかし、mAP_M は 4.5% 低下していた。次に、図 4 を見ると、車両、ルームミラー等の運転者視界において大きく映る物体に対して検出性能が向上しており、一方で、標識や看板等の小さい物体に対して検出性能が低下していることがわかる。

表 1 昼間画像に対する検出性能

train_dataset	mAP	mAP _S	mAP _M	mAP _L
昼間画像	42.6	8.3	26.9	57.7
昼間+偽夜画像	43.3	7.0	36.1	57.9

表 2 薄暮画像に対する検出性能

train_dataset	mAP	mAP _S	mAP _M	mAP _L
昼間画像	24.5	1.1	14.4	29.9
昼間+偽夜画像	29.4	1.4	9.9	36.4

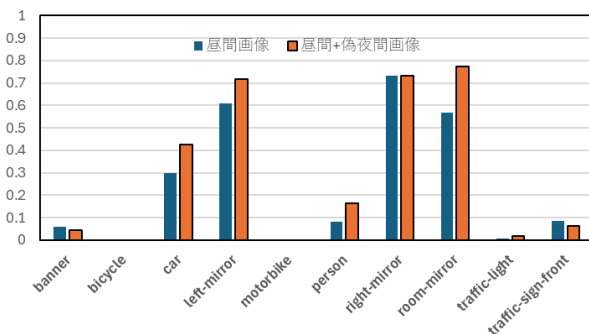


図 4 薄暮画像に対する各ラベルの mAP

5. 考察

CycleGAN による昼から夜への画像変換の結果について、変換後の画像にはアーティファクトやオブジェクトの消失などの GAN ベースの画像生成で生じる欠点が見られた。これに関して、学習の進行と並行して適切に学習率を減少させることである程度軽減できると考えられる。データセット間の物体検出結果の比較では、ドメイン適用後のデータセットを用いたほうが薄暮画像の mAP は高いことから、昼から夜のドメイン適応が薄暮画像の物体検出に一定の効果があったと考えられる。また、BBox のサイズが medium の物体に対して検出性能が低下していたことに関して、標識や看板等の検出性能が低下していたことから、オブジェクトが画像変換により消失し、該当箇所が物体検出タスクにおいて学習されてしまうため、検出性能に大きく影響を及ぼしたと考えられる。また、変換後の画像のサイズが小さくなっており、画質の低下に伴い情報が減少していたことも検出性能の低下の要因となったと考えられる。

6. おわりに

本研究では、CycleGAN を用いて昼間の運転者視点画像を夜間画像に変換し、物体検出の訓練データとして適用することで、薄暮時の物体検出性能を検証した。実験の結果、CycleGAN によるドメイン適用が、薄暮時の運転者視点映像中の物体検出性能の向上に有効であることを明らかにした。

今後の課題として、他の物体検出モデルを導入することでさらに検出性能を向上させることが挙げられる。また、昼間画像から夜間画像に変換した際の画質の低下やオブジェクトの消失の問題を解決するために、計算効率の向上や他のモデルでの画像変換を検討したい。

参考文献

- [1] 池田光汰, 猿田和樹, 陳国躍, 寺田裕樹, "Deep Learning を用いたドライバーの注視対象物の評価", 情報処理学会第 81 回全国大会講演論文集, pp.217-218 (2019)
- [2] 和田健太郎, 猿田和樹, 陳国躍, 寺田裕樹, "深層学習を用いた仮想空間提示映像に対する注視物体の判定と評価", 第 23 回情報科学技術フォーラム講演論文集, 第 3 分冊, p.225-226 (2024).
- [3] 小林陽, 猿田和樹, 陳国躍, 寺田裕樹, "Transformer-based モデルによる運転者視界に対する物体検出性能の検討", 第 23 回情報科学技術フォーラム講演論文集, 第 3 分冊, p.227-228 (2024).
- [4] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros., "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2223-2232 (2017).
- [5] Vinicius F. Arruda, Thiago M. Paixao, Rodrigo F. Berriel, Alberto F. De Souza, Senior Member, IEEE, Claudine Badue, Nicu Sebe, Thiago Oliveira-Santos, "Cross-Domain Car Detection Using Unsupervised Image-to-Image Translation: From Day to Night", 2019 IEEE International Joint Conference on Neural Networks(IJCNN) (2019).
- [6] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, Jian Sun, "YOLOX: Exceeding YOLO Series in 2021", arXiv:2107.08430 (2021).
- [7] Zhuofan Zong, Guanglu Song, YuLiu, "DETRs with Collaborative Hybrid Assignments Training", Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.6748-6758 (2023).