ゲームエンジンを活用したデジタルツイン上における交通量計測AI学習データの生成

皐月　健悟[1]・柳沼　秀樹[2]・寺部　慎太郎[3]・鈴木　雄[4]・海野　遥香[5]

[1]学生非会員 東京理科大学 創域理工学研究科 社会基盤工学専攻（〒278-8510　千葉県野田市山崎 2641）
E-mail: 7623517@ed.tus.ac.jp
[2]正会員 東京理科大学 創域理工学部 社会基盤工学科 准教授（〒278-8510　千葉県野田市山崎 2641）
E-mail: yaginuma@rs.tus.ac.jp(CorrespondingAuthor)
[3]正会員 東京理科大学 創域理工学部 社会基盤工学科 教授（〒278-8510　千葉県野田市山崎 2641）
E-mail: terabe@rs.tus.ac.jp
[4]正会員 東京理科大学 創域理工学部 社会基盤工学科 助教（〒278-8510　千葉県野田市山崎 2641）
E-mail: yusuzuki@rs.tus.ac.jp
[5]正会員 茨城大学大学院 理工学研究科 助教（〒316-8511　茨城県日立市中成沢町 4-12-1）
E-mail: haruka.uno.uj38@vc.ibaraki.ac.jp

交通量は道路の維持管理，その他関連施策のために極めて基本的かつ重要な情報である．これまでは，調査員の常駐による計測が主であったが，近年の AI 技術の発展に伴い，交通量計測の現場への AI 導入の機運が高まっている．AI による高精度な交通量計測を実現するためには，AI に対して「学習データ」を与える必要がある．しかしながら，学習データ作成の作業は多大な労力とコストを要するものであり，AI のコスト面での利点を相殺しかねない．本研究は，ゲームエンジンを活用した効率的な学習データ作成手法の開発を目的として，当該手法の有効性の検証を行った．本論文においては，学習済みモデルが苦手とする夜間の映像に対する有効性の検証を行った．検証の結果，3D モデルにより学習を行ったモデルでは，従来の学習済みモデルと比較して精度の向上が確認された．

Key Words: *Digital Twin, Game Engine, Virtual Reality, Traffic Survey, Artificial Intelligence*

1. はじめに

交通量は道路の建設や維持管理において基本的かつ重要な情報であり，日々至る所で交通量調査が実施されている．今日までの長きにわたり，調査員が対象地に常駐する形で交通量調査は行われてきた．しかし，人がその場に居る必要があるという性質のために，コストは高く，精度も安定しないという問題点を抱えている．

近年の AI 技術の発展に伴い，様々な分野において，人手の必要であった作業が AI に取って代わられていることもあり，交通量計測の現場においても AI 導入の機運が高まっている．AI を用いて交通量を計測する方法はいくつか考えられるが，最も一般的な方法の一つとして，物体検出 AI を用いて車両を検出するものが挙げられる．交通量計測において，高精度である方が好ましいことはいうまでもないが，AI 計測において精度を追求するためには，AI に対して学習データを与える必要がある．

しかし，時間帯や天候などの組み合わせによっては，特定の状況の映像が手に入らないこともあり，学習データの作成が困難であるような状況も考えられる．その上，学習データの作成は人手により行われるため，多大な時間や費用を要する．このことは，AI による交通量計測の利点の一つである低コスト性を相殺してしまうことを意味している．

このような背景から，まれな状況に対応することができ，かつ効率的な学習データ作成のための手法が求められる．

よって，交通量計測 AI のための，ゲームエンジンを用いた効率的な学習データ生成手法の確立を本研究の目的とする．

2. 既往研究の整理と本研究の位置づけ

ディープラーニングは，近年多くの場面で活用されており，特に自動運転には欠かせない技術である．自動運転は非常に希なシナリオに遭遇することや，過酷な天候に遭遇することが想定される．そこで Bo-Yi Lin ら[1] は，非常に写実的な仮想空間を Unreal Engine というゲームエンジンを使って構築し，その中で物体検出技術のテストを行った．ゲームエンジンの中では，現実の様々な環境を模倣することが可能であり，現実ではほとんど起こりえない希な状況さえも再現可能である．テストの結果，小さい物体，一部が隠れた物体，および複数の物体に対する物体検出技術の精度は，改善の余地があることが示された．

YOLO などの物体検出技術の訓練のためには，現実の画像に対して人手によりラベル付けをする方法が一般的である．この方法は，精度が高い物体検出モデルの構築を可能にするが，大変な労力を要するものである．そこで asmussen, I. ら[2] は，短時間で大量の学習データを生成すべく，Unreal Engine から生成した画像により，YOLO などの物体検出技術の訓練が可能であるかどうかを検証した．検証の結果，3D レンダリング画像により，現実世界を対象とした物体検出技術の訓練が可能であることが示された．しかし，モデルの安定性と信頼性を高めるためには今後のさらなる調整が必要であることが同時に示された．

上記のとおり，ゲームエンジンにより生成した画像を，物体検出 AI の推論や訓練に活用するという研究はすでに行われている．しかし，CCTV により撮影された映像に対する検出精度向上を目的とした，ゲームエンジンを活用した研究成果は未だもたらされていない．

前述のとおり，ゲームエンジン内では，物体検知精度に影響を与える要因である，地形や天候，太陽光などの視覚的要素を容易に変更し，任意の状況を再現することが可能である．すなわち，ゲームエンジンを活用した手法は，自動化による効率化だけではなく，様々な状況に対応したモデルの作成をも可能にする．本論文においては，3D モデルの作成から AI モデルの訓練までを行うことにより，上記のような利点を活かしながら，ゲームエンジンを交通量計測 AI の訓練に活かすことができるのか否かを検証する．

3. ゲームエンジンの概要

本研究においては，前述の目的を達成するために，米 Epic Games 社によって開発されているゲームエンジンである”Unreal Engine”を用いる．ゲームエンジンとは，ゲーム開発を行うためのソフトウェアフレームワークである．有名なゲームエンジンはいくつか挙げられるが，Unreal Engine は競合のソフトウェアよりも映像品質に優れていることから，実際の映像に対して適用する AI の学習という，本研究の目的に対して最も適していると判断した．

また，交通量調査は，様々な時間や天候において行われるものであり，AI を導入する場合も例外ではない．様々な状況下で，AI に精度良く車両を検出させるためには，学習データに同様の状況の画像が含まれていることが好ましい．Unreal Engine では，**図-1**, **図-2** に示すように，時間や天候について，任意の状態を作り出すことが可能である．

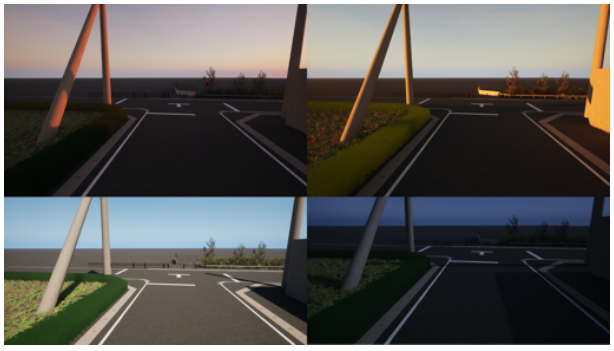


図-1 時間変化の一例

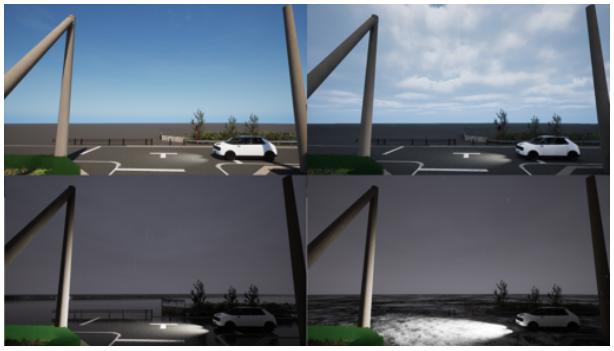


図-2 天候変化の一例



図-3　検証に用いる地点の CCTV 画像

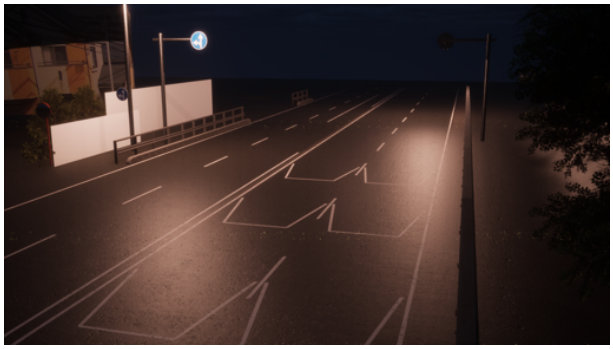


図-5　夜間の CCTV 画像を 3D で再現したもの



図-4　夜間の CCTV 画像に対する検出結果



図-6　車両を配置し光源を調整したもの

4.　有効性の検証とその結果，および考察

(1)　事前学習済みモデルの問題点

事前学習済みモデルは，大量の学習データにより学習がされているため，非常に多くの状況において正確に物体を検出することができる．しかしながら，事前学習済みモデルは，車両の検出，とりわけ交通量調査用途に対して必ずしも適しているとはいえない．

夜間の CCTV 映像では，ヘッドライトの光に車両の大部分が隠れてしまい，車両全体の像も進行方向にぶれてしまうということがある．事前学習済みモデルは多くの状況において精度を確保することができているが，上記のような夜間の映像において，その精度は著しく低下してしまう．**図-4** は夜間の映像に対して，学習済みモデルを適用した一例である．停止中の車両は検出することができている一方，走行中の車両は検出できていないことが分かる．

(2)　検証の方法

本論文では，事前学習済みモデルの問題点である，夜間の映像における車両の検出精度を，3D による学習前後で比較することにより，本手法の有効性を検証する．本検証において対象とするのは，**図-3** に示す地点である．すなわち，この地点での実測値に対して，AI による計測数がどの程度正確であるのか，学習前後でそれがどのように変化したのかを検証する．

また，精度の検証は，実際の AI による交通量計測においても使用される，カウントラインを用いた計測台数により行う．カウントラインとは，画面上に仮想的に設定する直線で，その直線上を車両が通過した場合に計測台数として数え上げるというものである．

(3)　検証の準備

前述の方法で検証を行うために，夜間の CCTV 映像を Unreal Engine で再現したものが**図-5** である．なお，夜間に撮影された映像の場合，画面内の光源が検出精度に対して非常に大きな影響を与えるため，光の表現については現実に忠実になるように特に注意を払った．

図-5 内に車両を配置，光源を調整したものが**図-6** である．この中で車両の種類，位置，向きを変化させ，様々な組み合わせの画像を出力する．

夜間の CCTV 映像は，シャッター速度の低下等により，移動中の車両は進行方向に像がぶれている．しかしながら，3D モデルから生成した画像中の車両は，完全に静止している状態であり，夜間を対象とした学習データとして理想的ではない．そこで，夜間を再現した画像すべてに対して，画像処理により



図-7　モーションブラーの適用例

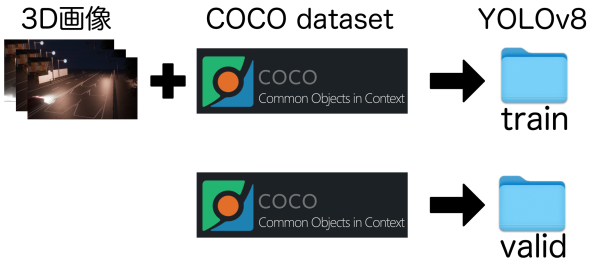


図-8　データセットの構成

モーションブラーを適用した．モーションブラーを適用した画像の例が**図-7** であり，同様の画像を 245 枚作成した．

物体検出モデルは，高速処理を特徴とする YOLO シリーズの YOLOv8m を用いる．YOLOv8m に対して，訓練用データとして，COCO dataset の訓練用データと，上記の 3D データを与えた．また，検証用データについては，COCO dataset の検証用データを与えた．COCO dataset とは，最も有名なオープンデータセットの一つであり，YOLOv8 の事前学習済みモデルも同データセットにより訓練されている．

(4)　検証の結果

事前学習済みモデルと，3D 画像により作成したモデルを用いて，検証地点の映像に対して推論処理を施した．**図-9** および**図-10** は，それぞれ事前学習済みモデルおよび 3D 学習モデルによる推論結果である．事前学習済みモデルでは検出ができていない車両に対して，3D 学習モデルでは検出ができるようになっていることが見て取れる．

また，これらのモデルを使用して車両を追跡し，交通量を計測した結果が**表-1** である．なお，こちらに近づいてくる向きが”coming”，こちらから離れていく向きが”going”である．事前学習済みモデルが，交通量を大幅に過小評価してしまっている一方，3D 学習モデルは非常に精度よく計測することができている．

表-1　実測交通量および AI 計測値の比較

	going	coming	合計
実測値	29	26	55
事前学習済みモデル	21	10	31
3D 学習モデル	29	26	55

結果を詳細にみると，事前学習済みモデルは coming において著しい精度低下がみられる．これは，前述のとおり，ヘッドライトによるハレーションに車体が隠れてしまっていることが原因であると考えられる．しかしながら，3D 学習モデルについては，ハレーションやぶれを 3D により再現した画像を用いて訓練されているため，そのような条件下での検出精度が改善したものと考えられる．

以上の結果から，3D レンダリング画像を用いてモデルを訓練することにより，悪条件下での物体検出 AI の検出精度が改善することが確認された．また，悪条件下での追跡を用いた交通量計測の精度も，同手法により訓練されたモデルを用いることにより改善することが確認された．

5.　まとめと今後の課題

本論文における結果から，3D モデルによる学習が物体検出モデルの精度向上に寄与することが確認された．しかしながら，本論文における検証のために作成したデータは，多大な労力を要して作成したものである．本研究が扱う 3D による学習という手法の，目的の一つとして省力化を掲げているが，これは未だ達成できているとはいえない．この手法は，ゲームエンジンを活用していることが特徴であり，これは車両の配置や画像の出力を自動化できるということを意味している．本論文において取り扱ったのは夜間の CCTV 映像であるが，物体検出 AI の交通量調査利用において，事前学習済みモデルが不得意とする状況は他にも数多くある．学習データ生成の自動化に取り組みつつ，そういった他の状況においても本手法が有効であるのか否かを検証するために，引き続き同様の取り組みを行っていきたい．

謝辞：本研究は，国土交通省新道路技術会議「カメラ画像および複数の観測データを融合した次世代交通計測手法に関する研究開発」の助成を受けたものです．ここに記して感謝いたします．



図-9　事前学習済みモデルによる推論結果



図-10　3D 学習モデルによる推論結果

Notes

REFERENCES

1) B. -Y. Lin, C. -S. Huang, J. -M. Lin, P. -H. Liu and K. -T. Lai, ”Traffic Object Detection in Virtual Environments,” 2023 International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan), PingTung, Taiwan, 2023, pp. 245-246, doi: 10.1109/ICCE-Taiwan58799.2023.10226661.
2) Rasmussen, I.; Kvalsvik, S.; Andersen, P.-A.; Aune, T.N.; Hagen, D. Development of a Novel Object Detection System Based on Synthetic Data Generated from Unreal Game Engine. Appl. Sci. 2022, 12, 8534. https://doi.org/10.3390/app12178534
3) Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2015. cite arxiv:1506.02640.
4) Yolo: Real-time object detection. https://pjreddie.com/darknet/yolov1/. (Accessed on 01/30/2023).
5) 今井龍一, 神谷大介, 山本雄平, 田中成典, 中原匡哉, 中畑光貴. 汎用的な深層学習器を用いた 交通量調査手法に関する基礎的研究. 土木学会論文集 F3(土木情報学), Vol. 75, No. 2, pp. I150–I159, 2019.

(Received October 13, 2023)
(Accepted ??? ?? 202?)

CREATING TRAINING DATA FOR TRAFFIC SURVEY AI WITH GAME ENGINE

Kengo Satsuki, Hideki Yaginuma, Shintaro Terabe, Yu Suzuki and Haruka Uno

Traffic volume is a fundamental and important information for a road construction and maintenance. In many fields, AI has been utilized to replace the old methods to reduce labor. In contrast, a traffic survey is still mainly conducted by manpower, whose cost is high and accuracy is not enough. Given the above, many people has been working hard to introduce AI into a traffic survey. There are some types of AI, but what is mainly utilized for a traffic survey should be object detection AI. Object detection AI must be trained by training data consisting of pictures and labels. To make training data, you must draw rectangles one by one around cars in pictures and add labels to each, which takes much time and effort. It can result in that making training data could spoil the advantage of AI that it does not cost so much. Considering what has been mentioned above, the old method of making training data should be replaced somehow. By the way, game engine like Unreal Engine can simulate visual factors like time, terrain, weather, reflection, and so on. It seems that the game engine is appropriate for training AI. That means, in the game engine, you can easily construct a dataset of every situation expected at the location where CCTV is installed. The reseach is aimed at establishing the mothod to build dataset in the game engine and verifying whether AI can be trained with it or not. As a result, it was confirmed that 3D trained AI model could detect cars more precisely than the pretrained model.