

## 敵対的生成ネットワークによるドライビングミュレータ 生成時の評価指標に関する一考察

### A study on evaluation index when generating Driving Simulator by Generative Adversarial Network

吉川 晃太<sup>†</sup>      梶原 祐輔<sup>‡</sup>  
Kota Yoshikawa   Yusuke Kajiwara

#### 1. はじめに

日本では、高齢者ドライバーによる事故全体に占める死亡事故の割合が年々増加している。警視庁の統計[1]によると、高齢者ドライバー（原付以上<特殊車を含む。>を運転している 65 歳以上の者）の事故割合は、2011 年が 13.4%であったに対し、2019 年には 18.1%となり、4.7%増加した。この背景から、警察庁は、2021 年 11 月 4 日に道路交通法の改正について示し、2022 年 5 月 13 日から、一定の交通違反歴のある 75 歳以上のドライバーに「運転技能検査」を義務化した。運転技能検査は、従来の免許更新時に行われる高齢者講習の前に実施される、運転技能を測定するための検査である。運転技能検査では、教習所や運転免許センターの定められたコースを走行し、「指示速度での走行」、「一時停止」、「右・左折」、「信号通過」などの課題を減点方式で採点される。この検査では、現在の高齢者講習での実車指導とは異なり、70 点以上（2 種免許では 80 点以上）で合格、合格点に達しなかった場合には不合格となる試験として扱われるようになった。合格できなかった場合でも、免許更新期間満了まで何度も受験することができる。

一方で、高齢者ドライバーの運転免許更新は 3 年ごとにあるが、その間の運転技能は測定できていないのが現状である。この間に、視覚的な注意機構や記憶力、認知機能が衰えた場合に、これに起因する事故を未然に防げない。そのため一般家庭で簡単に運転技能を測定する方法が必要である。

簡単に運転技能を測定する技術としてドライビングシミュレータ(DS)による運転技能の評価がある。ただし、DS で同じようなコースを何回も走ると慣れが起るため、日々の運転技能を測定するには、複数の異なる運転コースを生成する必要がある。株式会社東陽テクニカでは、360 度カメラを用いて道路を撮影し、それにアバターを追加することで簡単に DS を生成する技術を開発した[2]。しかしながら、アバターの生成や設置には専門的な知識が必要となるため、一般人には DS を構築することが難しい。また、DS 上で日々違うコースを走るには、365 日×3 年のコース数が必要になるが、それを人の手で構築するには現実的ではない。

そこで、本研究では、敵対的生成ネットワーク(Generative Adversarial Network(GAN))[3]を用いて、ドライブレコーダーの映像から、DS の自動生成を行う。GAN を用いることで、簡単に異なるコースを自動生成できる。た

だし、それが運転技能検査において評価できる DS であるかどうかは、まだ評価されていない。Kim らは GAN を使い、ドライブレコーダーの画像から DS を自動生成した[4]。この研究では、主に視覚的な整然生について評価されているが、運転技能検査の観点から DS を評価されていない。本研究で小松市内を運転したさいのドライブレコーダーの映像から DS を生成し、運転技能検査の観点から GAN を評価し、その結果を報告する。著者が知る限り初めての試みであり、新たな知見を提供する。

#### 2. システム概要

本研究では、ドライブレコーダーのデータを基に、GAN で DS を構築し、運転技能検査の実施を考慮した評価を行う。本研究の DS は、従来の DS と比べ、以下の利点がある。

- I. 自動車に搭載されているドライブレコーダーのデータをアップロードすることで DS を構築することができるため、従来の DS より安価に構築できる。
- II. 高齢者ドライバーが普段走行しているコースを基に構築するため、各高齢者ドライバーの用途や生活圏の道路事情に合わせた DS を構築できる

従来の DS は CG 技術を用いて、手作業で 3D モデルを組み立てたり、周囲の車や、飛び出してくる人間などの動作モデルを設計した。この方法では、多くのコストがかかる。また、360 度カメラや LiDAR などで撮影した実世界の映像を基に、実世界に基づいた DS を生成するソフトウェアがある[2, 5]。LiDAR は高価であり、専用の機材を一般人が所持していることは稀である。そのため、本研究はコスト面でメリットがある。

GAN で構築した DS を用いた運転技能の測定に関する概要図を図 1 に示す。まず始めに、①ユーザー側がドライブレコーダーによって撮影した映像や角速度、スピードのデータを Cloud（データサーバー）にアップロードする。その後、②Cloud 側で、アップロードされた映像から画像を抽出し、角速度やスピードのデータを基に GAN を利用することで、そのドライブレコーダーの映像に基づいた DS を実装する。実装した DS を、③ユーザー側がダウンロードすることで、ユーザーは DS を用いて運転技能を測定することが可能となる。ユーザーは、④運転技能の測定結果を市役所や警察庁などの行政機関へ送信することで、⑤ユーザー個人の運転技能を日頃から評価し、運転技能の遷移状況を管理することが可能となる。⑥運転技能を管理する中で、運転技能が低下した場合には、個別に行政機関から運転技能の低下を知らせる警告通知を送れる。

<sup>†</sup> 富山大学 University of Toyama

<sup>‡</sup> 公立小松大学 Komatsu University

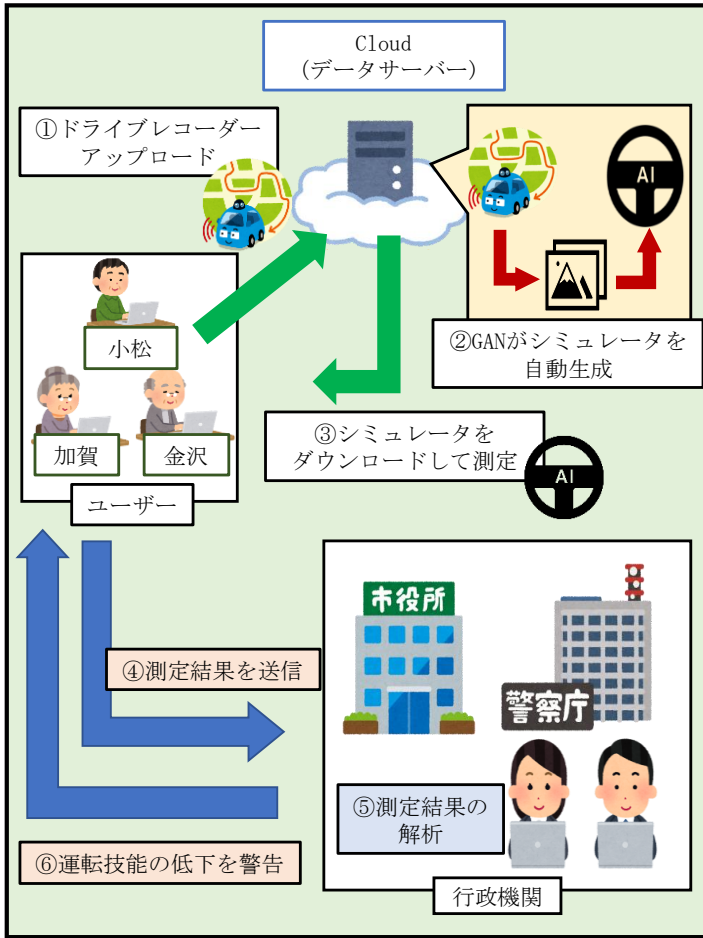


図 1 GAN による DS の運転技能測定

### 3. 実験方法

本実験では、最新の生成技術 DriveGAN[4]を用いて構築した DS の質的評価を行い、運転技能検査への適応可能性を探る。生成された DS の評価項目として、「視覚的な整然性」および「シーンの論理的整合性」である。「視覚的な整然性」は DriveGAN の原著で Frechet Video Distance[6]および、人によるシミュレータの視覚的品質評価および操作性評価でなされている。一方で、運転技能検査では表 1 の項目により運転技能を評価されるが、DriveGAN でこれらの項目が生成された DS 特有のオブジェクトに対する視覚的品質評価および論理的整合性を保っているかは評価されていない。したがって、実装した DS をもとに生成されたシーンと、使用したデータとの差を見比べて、DS の視覚的な整然性および論理的整合性を人が目視にて、定性的に評価する。定性的な評価は表 1 の運転技能検査の項目課題に準じた評価項目に従う。評価項目は「指示速度による走行」、「一時停止」、「右折・左折」、「信号通過」で、検証内容は「最高速度を表す道路標識」、「一時停止線」、「右折・左折のための道路」、「信号機」がそれぞれ生成されているかである。

DriveGAN に入力する画像は Intel 製の視覚センサ：Realsense depth camera D455 で取得した。視覚センサには吸盤を取り付け、フロントガラスに取り付けることで、自動車の前方の距離と画像データを取得した。自動車で運

表 1 運転技能検査の課題と本実験の検証内容

運転技能検査の課題	本研究での検証内容
指示速度による走行	最高速度を表す道路標識が生成されているか
一時停止	一時停止線が生成されているか
右折・左折	右折・左折のための道路が生成されているか
信号通過	信号機が生成されているか

転したコースは、距離が約 4.4km、運転に要する時間は約 10 分の公立小松大学周辺の道路である。基本的には片側一車線のコースで、道路標識が 35 回現れ、交差点を 5 回曲がり、踏切を 2 回、横断歩道を 13 回、信号を 4 回通過する。

学習データには、公立小松大学周辺を運転時の 24 万枚の画像を用いる。この画像は、公立小松大学周辺の 13 周分の映像から抽出している。また、1 周ごとに異なる自動車を用いた。Realsense depth camera D455 で撮影した時点で画像サイズが 1280×720pixel である。これを 256×256 に加工して用いる。また、Realsense depth camera で撮影した映像において、晴天時に撮影したデータが紫色を帯びている。画像データセットの例を図 2 に示す。また本研究では、学習に必要な、スピードと車体の方向の変化を表す角速度を収集することができなかったため、DriveGAN の GitHub 上で配布されているデータを代用して DS を実装している。

DS の実装には、Alienware Aurora R11 の PC を使用した。この PC の CPU は Intel®core™ i9-10900KF CPU @ 3.70GHz ×20、メモリは 31.2GiB、OS は Ubuntu 20.04.3 LTS である。また、GPU には NVIDIA GeForce RTX 3080 を使用した。潜在空間の学習に約 2.5 日、ダイナミクスモデルの学習に約 1 日かかっている。



図 2 画像データセットの例

表 2 実験結果

	鮮やかさ	形状	整合性
道路標識と規制表示	×	×	△
一時停止線	△	○	△
右折・左折のための道路	○	○	×
信号機	×	○	△

## 4. 実験結果

### 4.1 要約

表 2 に実験結果を示す。「鮮やかさ」は、オブジェクトやシーンが認識できる細かさかどうかを評価した。「形状」は、オブジェクトやシーンの形状を評価した。「整合性」は、オブジェクトやシーンが、その状況に相応しいものかどうかを評価した。評価は 3 段階で行った。○は、DS として利用する上で十分だと判断されるオブジェクト・シーンを表す。×は DS として利用する上で全体的に改善が必要であるオブジェクト・シーンを表す。△は DS として利用する上で一部に改善が必要なオブジェクト・シーンを表す。

### 4.2 道路標識

まず、鮮やかさを評価した。図 3 に生成された道路標識の例を示す。生成された DS 上のシーン (図 3 (左)) の道路の脇に、白く細長いオブジェクトが生成されていた。しかし、道路標識の上部が鮮明に表示されていないため、実際に何を表す道路標識かを認識できなかった。また、表示されてから数秒以内に消滅してしまったため、不安定な生成オブジェクトであった。また図 4 (左) のように絵文字部分がぼやけて見えなかった場合もあった。したがって、道路標識の鮮やかさは×とした。次に形状の評価を行う。道路標識の形状は図 4 (左) のように、崩れた標識も存在した。また、類似した画像データとして、図 4 (中) の「一時停止」、(右) の「横断歩道」の標識が映っている学習データがあった。したがって、道路標識の形状は×とした。道路標識の論理的整合性の評価を行う。道路標識は、道路の脇や車道の高い位置など現実の道路標識



図 3 生成された道路標識(左)と学習データの速度標識(中、右)



図 4 生成された道路標識(左)と類似した画像データ(左、右)



図 5 生成された一時停止線(左)と類似した画像データ(右)

と同等な場所に生成されていることがあった。一方、図 4 (左) のように、横断歩道があることを示す道路標識のようなものが生成されているが、道路上には横断歩道が生成されていなかった。したがって、道路標識の論理的整合性は△とした。

### 4.3 一時停止線

まず、鮮やかさを評価した。本研究で走行したコースでは、一時停止線は主に、横断歩道と踏切の直前に存在した。生成された一時停止線のシーンは図 5 (左) で、図 5 (右) の学習データに類似していた。生成された一時停止線は学習データよりも細く生成されているため、学習データよりも生成された一時停止線が認識しづらい。したがって、一時停止線の鮮やかさは△とした。形状を評価した。形状に関しては、白く、細長く生成されているものが確認できた。したがって、一時停止線の形状は○とした。論理的整合性を評価した。横断歩道の手前に表示される一時停止線は、横断歩道が、車が横断歩道を通過する直前のタイミングで生成されていたことがあり、一時停止線が生成されないシーンが多かった。横断歩道の直前の一時停止線が生成されにくいという問題があったため、一時停止線の論理的整合性は△とした。

### 4.4 右折・左折のための道路

まず、鮮やかさを評価した。右折・左折のための道路では、図 6 (左) のような、鮮やかな丁字路のシーンが確認できた。したがって、鮮やかさは○とした。形状を評価した。図 6 (左) のように、左右に分かれている道路が生成されていた。したがって、形状は○とした。論理的整合性を評価した。図 6 (左) の生成されたシーンは、図 6(右)の学習データと類似していた。図 6(右)の学習データと比べると、電柱やガードレールなどのオブジェクトが生成されず、平滑なシーンになっている。また、丁字路の前に必要な、一時停止線や一時停止の標識が生成されていない。したがって、論理的整合性は×とした。



図 6 生成された道路(左)と類似した画像データ(右)



図 7 生成された信号機(左)と類似した画像データ(右)

#### 4.5 信号機

まず、鮮やかさを評価した。信号機の生成されたシーンは、約 5 時間 DS の挙動を観察していた中で、稀に表示された。信号機は、図 7 (左) のようなシーンが生成された。また類似した学習データは交差点内の信号機だった (図 7(右))。生成された信号機が粗く、色が認識できなかった。したがって、信号機の鮮やかさは×とした。形状を評価した。形状のみで信号機だと認識できる。したがって、形状は○とした。論理的整合性を評価した。信号機は十字路の交差点で生成されているシーンで生成された。ただし信号機が生成されたシーンは特定の風景の場所のみであり、出現場所も少なかった。したがって、論理的整合性は△とした。

### 5. 考察

#### 5.1 DS の生成精度

同じコースを周回したデータを過学習することで、類似したコースを生成できた。一方で、生成した道路標識や一時停止線などの多くは、視覚的整然性や論理的整合性が満たされなかった。

まず、道路標識が鮮明に表示されなかった理由の一つに、道路標識の出現するデータ数が少なかったことが挙げられる。運転中に撮影しているドライブレコーダーの映像には、道路標識が数フレームしか映らないため十分に学習できなかったと考えられる。また横断歩道の道路標識に合わせて横断歩道が生成されなかった理由には、敵対的生成ネットワークが外観のみを模倣する方法であることが挙げられる。また生成された一時停止線が学習データよりも細く生成された理由の一つに、学習データのばらつきが考えられる。本実験では 1 周ごとに、運転した車の種類が異なる。撮影時に走行した車種によってカメラの位置がずれていたため、撮影された一時停止線の大きさが異なっていた。これらのデータを用いて一時停止線を生成したため、学習データと異なる一時停止線が生成されたと考えられる。次に、信号機が粗く生成された理由は、解像度にあると考えられる。ドライブレコーダーが撮影した映像には信号機が小さく映っている。また DriveGAN に入力するときに  $256 \times 256$  pixel に縮小している。この画像データを学習したため、信号機が粗く生成されたと考えられる。信号機の色は、撮影時に青色や黄色、赤色など、点灯、点滅する色が異なるため、生成されなかったことが考えられる。

#### 5.2 運転技能検査のための評価指標

運転技能検査のための GAN による DS の評価指標として、「鮮やかさ」、「形状」、「整合性」をとり、その指標に沿って DS を評価した。「鮮やかさ」と「形状」に関しては、DS のシーン内オブジェクトの視覚的情報を保証するためには欠かせない指標であると考えられる。道路標識を例に挙げると、「形状」が道路標識の見た目をしていたとしても、道路標識の内容が見えない状態、つまり「鮮やかさ」の精度が低いと、そのオブジェクトの意味が消失してしまう。このような問題を分けて考えるためにも、「鮮やかさ」と「形状」はそれぞれ指標としてとると有益である。また、「整合性」に関しては、特に重視する必要がある。実験結果にもある通り、横断歩道の直前に一時停止線が生成されなかったり、丁字路においては直前の一時停止線や道路標識が生成されなかったり、ガードレールや電柱なども消えてしまうことが多かった。このままだと、DS 内のシーンが現実のデータを基に生成されるのに現実のものと感じられないものになってしまうだけでなく、運転技能検査において、検査を行う上で必要な情報が表示されないという問題が発生する可能性がある。そのため、「整合性」は特に満たすべき条件であるため、評価指標として重要であると考えられる。

### 6. おわりに

本研究では、DriveGAN で DS を構築しその生成精度を評価し、運転技能検査の観点から GAN による DS の評価指標を示した。生成精度を視覚的整合性と論理的整合性から評価し、どの指標も重要だが、特に運転技能検査に用いる場合には、検査を行う上で重要な役割を担う場合が多い論理的整合性を満たす必要があるため、「整合性」は評価指標になくしてはならないと考えられる。そのため、これから GAN による DS を構築する場合には、論理的整合性を満たすようなモデルを考案していかなければならない。

また、本研究では著者の主観的な評価を基に実験を行ったため、アンケートを用いた複数人による客観的な評価を行い、より信憑性のある結果を得ることが今後の課題である。

#### 参考文献

- [1] 警視庁「警視庁の統計」、  
([https://www.keishicho.metro.tokyo.lg.jp/about\\_mpd/jokyo\\_tokei/tokei/index.html](https://www.keishicho.metro.tokyo.lg.jp/about_mpd/jokyo_tokei/tokei/index.html) 閲覧日: 2022 年 6 月 15 日)
- [2] 株式会社東洋テクニカ「Real Video Drive Player」  
([https://www.toyo.co.jp/files/user/company/documents/release/210805\\_real\\_video\\_drive\\_player\\_69058.pdf](https://www.toyo.co.jp/files/user/company/documents/release/210805_real_video_drive_player_69058.pdf) 閲覧日: 2022 年 6 月 18 日)
- [3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- [4] Kim, S. W., Phillion, J., Torralba, A., & Fidler, S. (2021). Drivegan: Towards a controllable high-quality neural simulation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5820-5829).
- [5] 株式会社 OTSL「製品紹介: 自動運転用センサーシミュレータ」  
(<https://www.otsl.jp/product/cosmosim/> 閲覧日: 2022 年 6 月 18 日)
- [6] Unterthiner, T., van Steenkiste, S., Kurach, K., Marinier, R., Michalski, M., & Gelly, S. (2018). Towards accurate generativemodels of video: A new metric & challenges. *arXiv preprint arXiv:1812.01717*.