

敵対的生成ネットワークに基づくドライバーの表情認識に関する研究

G2121039 関文鋭

1. はじめに

近年、世界的な自動車の爆発的増加や高齢化社会の深化に伴い、多くの交通事故が発生している。人的要因による交通事故を減らすために、各国政府は一連の施策を策定し、関連法規を施行し、一定の成果を上げていますが、その効果は限定的であり、特に感情や病気の問題による交通事故をこれらの規制で防ぐことは困難である。ドライバーの心身の状態は不安定で、運転中に事故を起こす可能性がある[1]。怒り、疲労、興奮、悲しみなどの不安定な感情が一般的である。安全運転のためにAI技術による運転支援を提案する研究者もいますが、現段階での自動運転技術は、想定されるすべての道路状況を解決できるわけではなく、現在の不完全な規制と高いコストでは、当面は普及させることができず、現段階ではまだ手動運転が自動車の運転方法のメインとなっている。多くの学者や研究チームが、運転者の感情状態が運転に与える影響について研究している。運転者が現在、恐怖や悲しみの状態にある場合、運転行動はより安全で、事故のリスクも低くなる。そのため、ドライバーの感情状態をモニタリングすることは、交通安全向上分野への応用が期待されている。

交通事故を減らすために、大学や研究機関では、運転中のドライバーの生理状態や認知状態について多くの研究が行われている。第一のタイプは接触型で、主に物理デバイスを用いてドライバーの生理信号（ECG信号、体温、呼吸など）を取得し、ドライバーの情動状態を判断するもので[2]、信頼性が高くリアルタイム性が高いが、センサーなどのハードウェアデバイスはドライバーに接触する必要が

あり、運転にある程度の影響を与え安全性が低い、第二は非接触型である。Mehravian[3]の研究によると、顔の表情は全体の55%と半分以上を伝えており、顔の表情が最も感情情報を伝えることができることを示している。一方、音声は7%と低い割合で情報を伝えており、音声では十分に表現できないことを示している。音声による情報伝達の割合が7%と低いのは、音声では個人の感情を十分に伝えられないことを表しています。そのため、ドライバーの表情をモニタリングすることで、ドライバーの感情状態を効果的に判断することができ、安全に運転できると判断することで、ドライバーが良い状態で運転できるような対策を促すことができるのである。今後、自動運転の時代には、ドライバーの状態が悪くないと自動的にシステムが引き継ぎ、無人で運転するようになる。

そこで、本研究の目的は以下の通りである。

- ① 運転中の表情データセットを構築し、今後の関連研究のために運転者の表情認識問題を研究する。
- ② Cycle GANを用いたドライバーの表情認識アルゴリズムの改良版を提案する。データセットが少なく、様々な表情画像のサンプルに偏りがあるという問題に対して、データで補強できる敵対的生成ネットワークを導入している。

2. 関連研究

研究者たちは、以前から表情の重要性を認識し、それに応じた研究を行ってきました。著名生物学者 Darwin[4]は、表情は性別や人種を越えて機能し、同じ感情を伝えることができるという結論を出し、科学書『Expressions of Affection in Man and Animals』に取り入れた。コンピュータビジョンでは、1970年代に Ekman と Friesen[5]が制作した Facial Action Coding System (FACS) が最初の表情認識の実験と研究を開始し、まず顔の動作データを取得し、それを6つの基本的な感情タイプに分類していた。しかし、現在の研究成果は、適切な顔表情データセットがないことや、表情認識手法に新しいブレイクスルーがなく、従来の認識アルゴリズムが中心となっていることもあり、非常に限定的なものにとどまっている。そこで、本章では、表情データセットとドライバー表情認識の関連研究をそれぞれ紹介する。

2.1 表情データセット

JAFFE データセットは1998年に Lyons[6]らによって作成され、10人の日本人女性に髪を結んでもらい、同じ背景で嫌悪、幸福、恐怖、怒り、驚き、悲しみ、冷静の7つの情動状態を表現してもらった213枚の画像から構成されている。五感の位置関係や画像の大きさによる影響を排除することを目的としている。このデータセットはオープンソースであり、高度な表現校正がなされている。

Lukey[7]らが2010年に作成したCK+データセットは、123人が撮影し、研究者が自由に利用できる593枚の画像シーケンスを含んでいる。集められたボランティアの年齢分布は18歳から30歳で、60%以上が女性、アジア系とヒスパニック系はわずか3%でした。他のデータベースとは異なり、保有者はデータの一部にセンチメントのラベルを付けており、このデータセットは多くの研究で利用されている。

2.2 ドライバーの表情認識関連研究

Moriyama[8]らは、時間的変化をマークする方法を用いて、運転者が穏やかなとき、楽しいとき、怒っているときの表情をフィルタリングし、分離顔情報空間法を用いて運転者の複数の顔特徴を抽出し、感情を識別している。

Azman[9]らは、ドライバーの表情をリアルタイムで検出し、3秒間怒っていると検出された時点でドライバーに警告を発する。

そのため、自動車ドライバーのための表情認識技術の研究は非常に重要である。関連研究から明らかのように、ほとんどの研究がデータベースに焦点を当てたものではありません。既存の顔表情データセットは、顔の表情を通して人間の感情状態を認識する研究に利用できますが、運転はドライバーの認知資源を大量に必要とするプロセスであり、運転がドライバーの感情表現に影響することを意味し、通常的生活シーンとは異なる。通常表情データセットを運転シーンに適用した場合、信頼性の高い認識結果が得られない可能性がある。

3. 本研究の概要

本章では、運転状態表情データセットの作成と改良型 GAN を用いた運転者表情認識アルゴリズムから、それぞれ本研究の概要を紹介する。

3.1 運転状態の表情データセットの構築

1つ目は、エモーションカテゴリーの選択に関するセッションである。ニーズに応じて離散的な感情が選ばれ、「嬉しい」「悲しい」「怒り」「驚き」「嫌悪」「恐怖」「冷静」の7つの基本感情が選ばれる。次に、表情画像の撮影を行いました。まずドライバーの表情画像の命名基準を作成し、その後、映画やテレビ作品から適切なリソースを選定し、約400枚の表情画像の画像撮影を行った。続いて、感情のカテゴリーを決定した。まず WEN[10]を参考に、適切な感情判定ルールと判定方法を開発した。画像の感情判定を手動で行い、データセットを作成した。

最後に、データセットに含まれる発現画像を処理した。その画像をMTCNNで検出し、整列させることで、サイズ128*128の均一な画像を生成した。

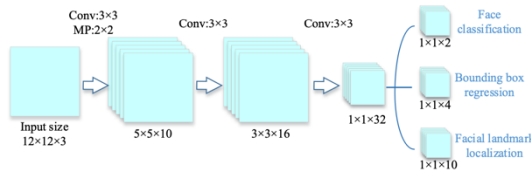


図 3.1 Structure diagram of P-Net

3.2 改良型 GAN を用いた運転者表情認識アルゴリズム

一般的な表情データセットでは、穏やかな状態の表情が最もデータ量が多いため、穏やかな表情画像から他の種類の表情画像への変換がデータ拡張の主な作業となる。ネットワーク内に一对多の対応関係がある場合、すなわち、穏やかな表情画像から多カテゴリーの表情画像への変換を同時に扱う場合、モデルの学習を数回行う必要があり、アルゴリズムの時間コストが非常に大きくなってしまふ。そこで、本研究では、顔画像合成と表情認識を1つのフレームワークで行うことで、CycleGAN ネットワークの改良を提案する。本手法は、元の CycleGAN ネットワークにカテゴリ制約を加えることで、多カテゴリーの画像サンプルの変換を一つのモデルで実現するものである。同時に、識別器に補助表現分類器 C を追加し、ネットワーク内の2つの識別器を新たな識別分類器 DC に置き換える。データセットのサンプル数が少なく、各カテゴリーの画像サンプルのバランスが悪いという問題を解決し、時間とコストを節約し、ネットワークの表情認識率を向上させた。識別器の損失関数に勾配ペナルティ機構を追加し、モデルの崩壊しやすいという問題を解決した。ネットワーク全体のアーキテクチャを図 3.1 に示す。

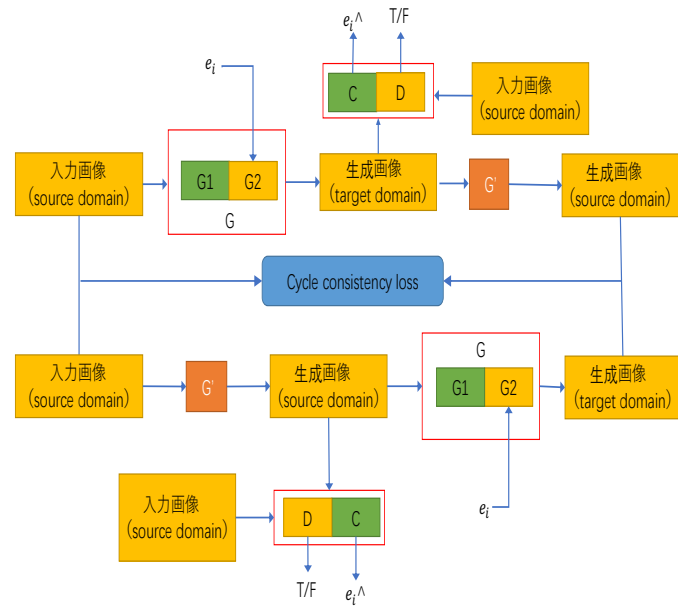


図 3.1 改良型 GAN の構造

4. 進捗状況と得られた成果

現在、表情画像のインターセプトとデータの预处理を実施中。今後の研究スケジュールは以下の通りである。

	8月	9月	10月	11月	12月	1月	2月
表情画像を加工してデータセットを作成する							
Tensorflowを用いた評価実験							
データの集計と論文の完成							
学外発表予定					大学コンソーシアム八王子		電子情報通信学会

図 4.1 今後研究スケジュール

5. おわりに

本研究では、ドライバーの表情認識手法に着目しています。公開されているドライバー表情データセットが存在しない現状に対し、映画やテレビ作品から芸術的複製手法により顔画像を取り込み、ドライバー表情データセットを確立し、データセットのデータ量が少ない問題に対し、改良型データ拡張手法を提案し、確立した運転中の表情データセットとCK+データセットで評価実験を行っている。

6. 参考文献

- [1] Klara S, Felix L, Kathrin M, et al. Effects of emotions on driving behavior[J]. Transportation Research Part F: Traffic Psychology and Behaviour, 2018, 59(A):150-163.
- [2] Matsuo H, Sato T, Yokoya N. Vehicle Driver Face Detection in Various Sunlight Environments Using Composed Face Images[C]. 2014 22nd International Conference on Pattern Recognition. Stockholm, Sweden, IEEE, 2014:1687-1691.
- [3] Mehrabian A. Communication without words[J]. Psychology Today, 1968, 2:53-55.
- [4] Darwin C. The Expression of the Emotions in Man and Animals[M]. Chicago: Oxford University Press, 1872: 10-11.
- [5] Ekman P, Friesen W. Facial Action Coding System: A Technique for the Measurement of Facial Movement[M]. Palo Alto: Consulting Psychologists Press, 1978: 26-27.
- [6] Lyons M, Akamatsu S, Kamachi M, et al. Coding facial expressions with Gabor wavelets[C]. Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. Nara, Japan, IEEE, 1998:200-205.
- [7] Lucey P, Cohn JF, Kanade T, et al. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression[C]. 2010 IEEE International

Workshop on CVPR for Human Communicative Behavior Analysis. CA, USA, IEEE, 2010:94-101.

- [8] Moriyama T, Abdelaziz K, Shimomura N. Face analysis of aggressive moods in automobile driving using mutual subspace method[C]. Proceedings of the 21st international Conference on Pattern Recognition (ICPR2012). Tsukuba, Japan, IEEE, 2012:2898-2901.
- [9] Azman A, Raman K, Mhlanga I, et al. Real Time Driver Anger Detection[M]. Singapore: Springer Singapore, 2019:157-167.
- [10] 温万惠. 基于生理信号的情感识别方法研究[D]. 西南大学, 2010.