

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

A Study on Driver's Facial Expression Recognition Based on GAN

関文鋭¹⁾

指導教員 亀田弘之²⁾, 研究協力者 カ ウンキョク³⁾

1) 東京工科大学大学 バイオ・情報メディア研究科 コンピュータサイエンス専攻 思考と言語研究室

2) 東京工科大学 コンピュータサイエンス学部

3) 東京工科大学大学 バイオ・情報メディア研究科 コンピュータサイエンス専攻 亀田研究室

キーワード : Driver's Emotional State; Facial Expression Recognition; Dataset; GAN Network

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。ドライバーは、自分自身のストレスや外部環境によって気分の浮き沈み時、交通安全を脅かすような危険な運転行動を起こしやすくなる。そのため、ドライバーの気分をリアルタイムにモニタリングし、感情状態に異常がある場合に警告を発したり、適時規制することで、交通事故の可能性を低減することができる。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。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

Lucey^[1]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^[2]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。

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

3.1 運転状態の表情データセットの構築

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。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。

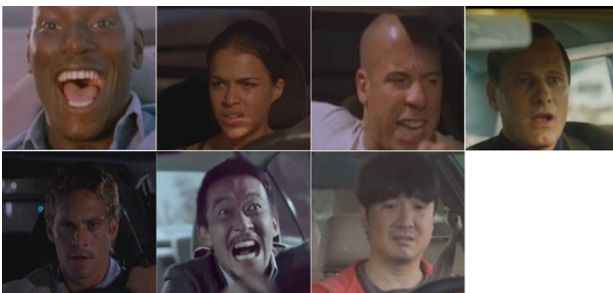


図 3.1 Example images of the seven emotion categories

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対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 3.2 に示す。

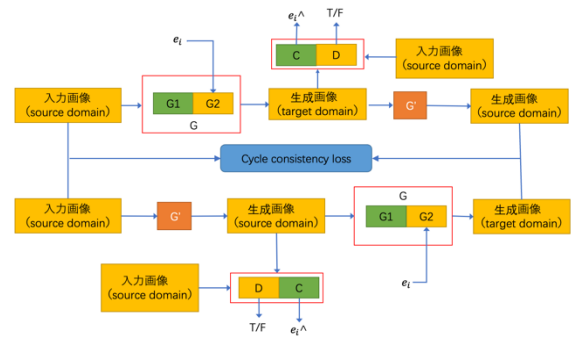


図 3.2 Network Architecture

4. 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る。ドライバーの表情のデータセットが公開されていないことを受け、映画やテレビ作品から顔画像を取り込んでドライバーの表情のデータセットを構築する芸術的複製法を用い、データセットのデータ量が少ないという問題を解決するために改良したデータ拡張法を提案しました。評価実験は、CK+データセットと本研究で確立したデータセットにそれぞれ基づいて行い、表情認識率を比較することでモデルの有効性を検証した。

参考文献

- [1] 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
- [2] 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.