

不均衡データの取り扱いに着目した深層学習による運転者の眠気推定

Driver Drowsiness Estimation Using Deep Learning Considering Imbalanced Data

電気電子情報通信工学専攻 石田 直也

Department of Electrical, Electronic and Communication Engineering Naoya ISHIDA

1. 緒言

1.1 研究背景

近年、衝突被害軽減ブレーキや、車線逸脱警報装置などの機能を搭載した先進安全自動車 (ASV) の開発が進んでいる。そのため、自動車の性能は急速に向上しており、交通事故は年々減少傾向にある。しかし、脇見運転や漫然運転、居眠り運転といった、ドライバーのヒューマンエラーに起因する交通事故は減少傾向にあるものの、全体を占める割合は増加傾向にある。このような交通事故を減らすためには、このような交通事故を減らすためには、ドライバーの状態を無意識下でセンシングし、ドライバーの状態に応じたフィードバックを行う必要がある。運転中のドライバーに眠気が生じることが特に多く報告されている。眠気は環境の単調さ、病気や疲労、睡眠不足、薬の副作用など様々な要因によって生じる。自動車を運転している場合、ハンドルやペダル操作といった運転タスクによる負担が小さくなった時に、疲労や睡眠時間に伴わず意識水準が低下し眠気が生じると考える。この問題の対策として、運転中、ドライバーから取得される情報をもとに正確に眠気を推定する必要があると考える。

1.2 研究目的

本研究では、人間の状態を判断するために、眠気が現れやすい顔表情を用いて、近年、益々発展している手法である深層学習を用いることでドライバーの眠気推定を正確に行うことを目的とする。深層学習を用いた運転者の眠気推定において、不均衡データへの対処が課題の一つである。運転者の眠気推定において、顔表情から取得される情報やラベルに個人差が大きくある。つまり、ラベル間の比率が極端に偏っている場合もあり、それに対する何かしらの対応策をとらなければいけないことがある。したがって、本研究では不均衡データに対する解決策を調査し、その中で運転者の眠気推定に有効とされる手法について検討し、その効果を検証する。

2. 運転者の眠気検知システム手法

運転者の眠気検知システムは大きく3つのカテゴリに分けられる。生体情報を用いた手法、車両情報を用いた手法、運転者の行動指標を用いた手法である。生体情報を用いた手法では、皮膚に電子デバイスを付加してドライバーの状態をアクセスすることで計測するもので、脳電図 (EEG)、心電図 (ECG)、眼電図 (EOG) の3つの項目に大きく分けることが出来る。この手法では、心拍数、脈拍数、呼吸数、体温など、ドライバーの身体的状態から眠気を検知することが可能である。これらの生体パラメータは、ドライバーに物理的に何が起こっているかに関係するため、眠気検出においてより信頼性が高く正確である。車両情報を用いた手法では、車両制御システムがドライバーの眠気を分析するために使用されており、ステアリングホイールの動き、ブレーキパターン、車線逸脱計測のようなものが含まれる。運転者の行動指標を用いた手法では、車両ではなく運転者にフォーカスを当てているため、車両情報を用いた手法より信頼性が高いとされている。しかし、行動指標は閉眼率、瞬目、頭の位置、表情、あくびなどの運転者の行動指標を通じて運転者の眠気を検出する非侵襲的な手段である生体情報を用いた手法や車両情報を用いた手法を前項で述べたが、この手法が最も一般的になってきている。本研究でも、車内にカメラを搭載し、顔情報を取得することを想定して、それを基に運転者の眠気推定を行う。

3. 評価用データセットの取得

評価用のデータセットはドライビングシミュレータによる実験によって取得された。カメラによる顔情報を元に Facemesh と呼ばれるライブラリを用いて、EAR (Eye Aspect Ratio) を算出した。処理としては、目の横幅を

取得する。そのあと目じりと上下瞼の中心座標より目の中心座標を算出。その座標から目じり間の距離の二倍の範囲内の区画を切り出す。また、目の向きが通常とは大きく異なる可能性があるため、両方の目じりの点での角度を算出する。通常両目じりの角度は0であるため、角度が0になるように回転する。今回使用するのは、眠気の実値は被験者の顔表情動画を第三者が観察して眠気段階を評価する顔表情評定により得る。顔表情による客観的眠気評価の妥当性については関連研究においても十分述べられており、VASなどの主観的眠気評価に比べて有用な評価指標である [1].

$$EAR = \frac{|P_2 - P_6| + |P_3 - P_5|}{2|P_1 - P_4|} \quad (1)$$

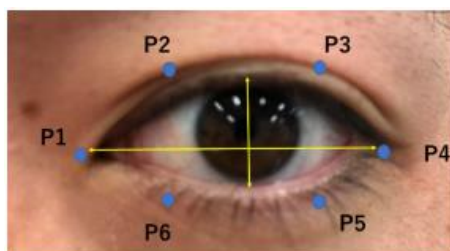


Figure. 1 目の輪郭上のランドマーク

4. 不均衡データの解析

不均衡データとは、正例と負例の比率が「0.01 : 0.99」のような比率の偏ったデータのことを指す。このようなデータセットにおいては精度の高さのみ追求する場合、問題ではないが、2つのラベルを分類する必要がある場合には大きな問題となり、タスクによっては負例を検出することが重要な場合がある。眠気推定に関して、起きている状態より眠っている状態の検出の方が本来重要である。不均衡問題に対処するために、データに基づくアプローチとして、オーバーサンプリングとアンダーサンプリングがある。オーバーサンプリングとは、少数派クラスのデータを多数派クラスのデータに合わせて補完するというものである。オーバーサンプリングのメリットは、クラス間の不均衡の問題を解消することが出来ることである。デメリットとしては、少数派クラスのデータを水増しすることになるため、過学習が起りうる可能性がある。オーバーサンプリングの手法として、SMOTEやADASYNなどがある。アンダーサンプリングとは、多数派クラスのデータを少数派クラスのデータに合わせてランダムに抽出するというものである。アンダーサンプリングのメリットは、クラス間の不均衡の問題を解消することが出来ることである以外に、データを減らすため、計算コストの低下もすることができる。デメリットとしては、多数派クラスのデータを削減するため、データの情報損失が起りうる可能性がある。

5. Mixupを用いた運転者の眠気推定

5.1 Mixupを用いた提案手法

本研究では、Mixupを用いることによる提案手法によって、前述した手法の問題を解消し、より多様なデータを作成し、不均衡データにおける分類問題での精度を向上させる。Mixup[2]は、当初深層学習における過学習を減少させる簡単な正則化技術として紹介されたものである。2つの訓練サンプルのペアを混合して新たな訓練サンプルを作成するデータ拡張方法であり、データとラベルのペア (x_i, y_i) , (x_j, y_j) から、下記の式により新たな訓練サンプル (\tilde{x}, \tilde{y}) を作成する。 λ は0から1の間の値を取り、これはハイパーパラメータ α である β 分布から得られる値となっている。以下のように合成データを作成する。

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j \quad (2)$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j \quad (3)$$

Mixup を適用する提案手法として、以下のような手順を示す。

- I. バッチサイズごとに Mixup を適用する前に行われるサンプリングをランダムではなく、各クラスのデータが均等にサンプリングされるようにする。
- II. β 分布から得られたハイパーパラメータ α を使い、2つの訓練サンプルを混合し、新たなサンプルを作成
- III. 決められたバッチサイズで作成されたデータを元に学習を行う。

5.2 結果と考察

本実験では、学習用のデータに目の動きのデータを用いるが、異なる特性を持つ被験者のデータがノイズとなり、良好な学習の妨げになることが知られているため、被験者個人ごとにモデルを作成することにした。また、一つのデータは 20 秒区間であり、層化抽出法を用いた K 分割交差検証法 (K=5) でデータセットを分割し、学習を行った。各モデルの学習率やドロップアウトの比率などのパラメータは、事前に Optuna というハイパーパラメータの最適化を自動化するためのツールを使用して決定した。学習回数は最大で 500 回とし、バッチサイズは 32 で固定することにした。また使用したモデルは CNN をベースにしたモデルと LSTM Fully Convolutional Network[3]というモデルである。また、実験のパターンは以下の場合である。

- (a) オーバーサンプリングを行わない場合 (提案手法も同様)、クラス毎のデータ量に基づいた重みを与える。
- (b) SMOTE によるオーバーサンプリング
- (c) ADASYN によるオーバーサンプリング
- (d) 各バッチで Mixup ($\alpha=0.1,0.3,0.5$) の適用

評価指標として、混同行列、正解率 (Acc) またマシューズ相関係数 (MCC) を使用する。混同行列とは、あるデータを分類したとき、その正解・不正解の数を整理しておく表である。正解率とは、予測結果全体のうち正しく予測されたデータの割合を示す指標である。マシューズ相関係数とは、予測値と真値の間の相関係数であり、以下のような式で示される。

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4)$$

結果は交差検証法で学習された 5 つのモデルのうち、中央値をとるモデルの結果を示している。Subject1 に関して、CNN, LSTM_FCIN とともに $\alpha=0.1$ での提案手法が従来手法と比較して、最も精度が高い結果となっている。また、混同行列では、ラベル 0, 3 に関しては比較的正しく推定できているものの、ラベル 1, 2 の推定はあまりできていない。ラベル 1 はラベル 0, 3 と、ラベル 2 はラベル 1, 3 と間違えて推定している。Subject2 に関して、CNN の $\alpha=0.3$ での提案手法が従来手法と比較して、精度が高い結果となっている。しかし、LSTM_FCIN では、従来手法の ADASYN による結果が最も良くなっている。また、混同行列では、ラベル 0, 1 に関しては比較的正確に推定できているものの、ラベル 2, 3 の推定は 1 件も推定することが出来なかった。これらの結果から、CNN モデルのような軽量モデルの方が提案手法の効果が現れやすいことが分かった。また、Mixup における適切な α 値は、使用するデータセットによって変わってくるのが考えられる。クラス間の不均衡具合によっては、提案手法では効果が期待できないこともあり、クラス毎のデータの絶対数はある程度必要である。全体の結果を見ると、従来手法と比較して提案手法は多様なデータを作成できたのではないかと考えられる。

Table. 1 Acc, MCC (CNNID (Subject1))

α	Acc \uparrow	MCC \uparrow
Class Weight \uparrow	62.96 \uparrow	50.21 \uparrow
SMOTE \uparrow	60.56 \uparrow	48.70 \uparrow
ADASYN \uparrow	61.85 \uparrow	48.63 \uparrow
Mixup ($\alpha=0.1$) \uparrow	65.19\uparrow	52.45\uparrow
Mixup ($\alpha=0.3$) \uparrow	64.63 \uparrow	51.75 \uparrow
Mixup ($\alpha=0.5$) \uparrow	64.63 \uparrow	52.00 \uparrow

Table. 2 Acc, MCC (LSTM_FCIN (Subject1))

α	Acc \uparrow	MCC \uparrow
Class Weight \uparrow	57.41 \uparrow	42.57 \uparrow
SMOTE \uparrow	56.85 \uparrow	41.30 \uparrow
ADASYN \uparrow	55.93 \uparrow	40.08 \uparrow
Mixup ($\alpha=0.1$) \uparrow	59.44\uparrow	44.19\uparrow
Mixup ($\alpha=0.3$) \uparrow	57.96 \uparrow	42.31 \uparrow
Mixup ($\alpha=0.5$) \uparrow	58.89 \uparrow	43.47 \uparrow

Table. 3 Acc, MCC (CNN1D (Subject2))

	Acc [±]	MCC [±]
Class Weight [±]	66.48 [±]	52.43 [±]
SMOTE [±]	84.81 [±]	71.89 [±]
ADASYN [±]	85.19 [±]	72.64 [±]
Mixup ($\alpha=0.1$) [±]	86.30 [±]	74.71 [±]
Mixup ($\alpha=0.3$) [±]	86.30[±]	74.90[±]
Mixup ($\alpha=0.5$) [±]	85.74 [±]	74.07 [±]

Table. 4 Acc, MCC (LSTM_FCN (Subject2))

	Acc [±]	MCC [±]
Class Weight [±]	81.85 [±]	67.74 [±]
SMOTE [±]	85.19 [±]	72.89 [±]
ADASYN [±]	85.37[±]	73.20[±]
Mixup ($\alpha=0.1$) [±]	84.07 [±]	70.79 [±]
Mixup ($\alpha=0.3$) [±]	83.89 [±]	70.64 [±]
Mixup ($\alpha=0.5$) [±]	83.70 [±]	70.01 [±]

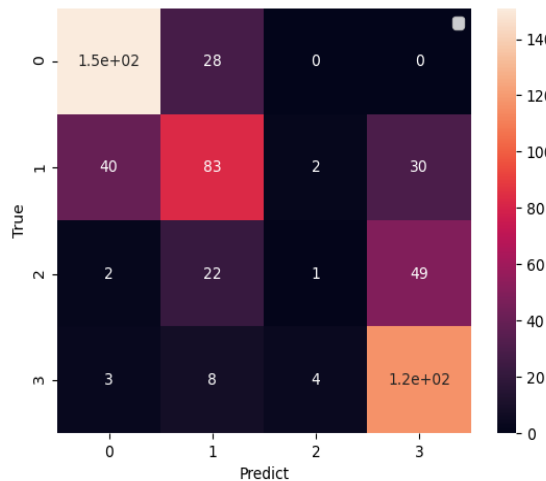


Figure. 2 混同行列 (CNN1D (Subject1))

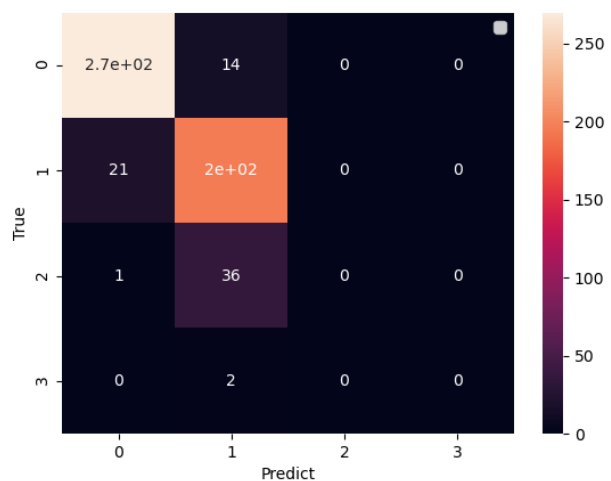


Figure. 3 混同行列 (CNN1D (Subject2))

6. 結言

本研究では、運転者の眠気推定タスクにおいて、サンプリングに工夫をし、そこに対して Mixup を使用することで、不均衡データに対応した手法を提案した。提案手法は多くの場合で高い精度を実現することが出来た。Mixup におけるハイパーパラメータである α 値による結果の検証をすることが必要であり、モデルによらない、不均衡度合いによらない改善策を検討する余地がある。また、ドライビングシミュレータによる実験は、他の被験者に対しても行っているため、被験者ごとに同様の実験を行うことができる。その他にも、評価指標として扱ったものがあるが、運転者の眠気推定タスクにおいては、起きている状態より眠りに近い状態を正確に推定することが非常に重要である。そのため、各ラベルの重みを考慮した上で眠気推定することも重要な課題といえる。本研究では、運転者の目に関する情報を使用した。入力情報に口の情報や生体情報を増やした際の提案手法の効果を確認することも試しうるところであろう。

参考文献

- [1] 大須賀 美恵子, 鎌倉 快之, 井上 裕美子, 野口 祥宏, ノプスワンチャイ ルーングロシユ: “ドライバの覚醒状態推定”, 自動車技術 62(12), pp46-52, 2018
- [2] H. Z. C. N. D. Lopez-Paz, “Mixup: Beyond Empirical Risk Minimization”, arXiv:1710.09412.
- [3] F. K. M. D. Chen, “LSTM Fully Convolutional Networks for Time Series Classification”, arXiv:1709.05206.