

道路標識認識のための学習データ増強手法の検討 Study on Augmentation Technique of Training Data for Traffic Sign Recognition

堰澤 映[†] 中島 克人[†]
Akira Sekizawa Katsuto Nakajima

1. はじめに

機械学習を用いた画像認識システムを構築する場合には適切な学習アルゴリズムを選択するだけでなく、十分な量と質を持つ学習データを用意することが必要である。このとき十分な量の学習データを収集できない場合や収集したデータに偏りがある場合の対策として学習データを人工的に生成する手法が提案され[1,2]、道路標識認識[3]や人検出[4]、特定物体認識[5]などに幅広く応用されている。

一方で学習データ増強手法が多様化するとともに各手法が前提とする条件や目的が複雑化し、学習データ増強手法の提案者が意図する用途と手法の使用者が希望する用途の違いが起きやすい状況となっている。

そこで我々は学習データ増強手法の使用者が適切な手法を選ぶための統一的な指針を作成した。我々はまず学習データ増強手法の用途を 4 つに分類した。用途を明確に把握することが適切な学習データ増強手法を選択する最初のステップである。

次のステップではそれぞれの用途に適した具体的な学習データ増強手法を選択する必要がある。我々は個々の具体的な学習データ増強手法が 4 つの用途のどれを満たすのか判断するための 3 項目を提案した。そして道路標識認識における学習データ増強手法を題材に、3 項目を用いた適切な手法の選択がどのように機能するか示した。

以上の検討は道路標識認識のための学習データ増強を基に行ったが、それ以外の認識対象にも適用可能な一般性のある指針になると考えている。

2. 先行研究との関係

本稿と同様に学習データ増強手法の体系化を目指した研究として、村瀬による画像認識のための生成型学習[1]や Wong らによる研究[6]が存在する。これらの先行研究は学習データ増強手法の具体的なアルゴリズムに着目して分類している。

我々は学習データ増強の具体的なアルゴリズムではなく、生成された学習データが満たす要件に着目して量的要件と質的要件に分類した。さらに、目指す用途に対して適切な手法を選ぶ際に有用な 3 つの比較項目を提案した。

3. 学習データ増強手法の用途の分類

多くの場合に学習データ増強はデータ収集のコスト削減を目的に行われるが、実際にはそれ以外の目的が課されることがある。我々は学習データ増強手法に課される要件を量と質の観点から次の 4 つに整理した。

- ① (量的要件) 学習データの収集コストの削減
- ② (質的要件) レアデータのシミュレーション
- ③ (質的要件) 不均衡データのオーバーサンプリング

④ (質的要件) 学習効率の高いデータの生成 (能動学習)

上記の要件を全て満たす理想的な学習データ増強手法はまだ存在しないため、目的に応じて適切な手法を組み合わせる必要がある。

3.1 学習データの収集コストの削減

道路標識は種類ごとに設置場所や設置頻度に偏りがあるため、全ての種類の道路標識に対し十分な量の学習データを人手で集めることは高コストである。

このとき用いられる学習データ増強手法では新しい学習データの生成に必要なシードとなる実画像がなるべく少なく、理想的には実画像を全く必要としないことが望ましい。

道路標識の 2D モデルから学習データを生成する手法[3]は道路標識の実画像を全く必要とせずに学習データを生成することが可能である。

3.2 レアデータのシミュレーション

安全運転支援を目的とした道路標識認識では、発生率の低い特殊な状況やデータに対しても正しく認識できることが安全性のために求められる。例えば霧や雪などの特殊な天候条件における標識画像はそのイベントが起きない限りコストをかけても収集できないため、学習データ増強手法により特殊な状況やデータを積極的にシミュレーションして作り出すことが有効である。

3.3 不均衡データのオーバーサンプリング

道路標識は種類ごとの設置数に偏りがあり、撮影時間帯や経年劣化の度合いにも偏りが生まれやすいことから、道路標識のデータセットは不均衡データ[7]になりやすい。

このような不均衡を改善するためのオーバーサンプリング手法としても学習データ増強は使用できる。この場合のデータ増強手法は単にデータの数を増やすだけでなく、不足した性質のデータを選択的に生成する能力が求められる。

3.4 学習効率の高いデータの生成 (能動学習)

レアデータのシミュレーションによる生成や不均衡データの改善が間接的に予測誤差を改善するのに対して、予測誤差を定量的に推定し、予測誤差の推定値を直接的に減少させる学習データを人工的に生成するクエリ生成型能動学習[8]が能動学習[9]の一手法として提案されている。

クエリ生成型能動学習を画像認識に適用する際は人間が見ても識別困難でラベル付けできないサンプルが生成されることが問題であった[10]が、近年では深層生成モデルを用いて問題の解決に取り組んだ手法[11]も提案されている。

4. 用途に適した学習データ増強手法の選択法

前章で述べた 4 つの用途を全て満たす万能な学習データ増強手法はまだ存在しないため、使用者は用途に応じて適切な学習データ増強手法を選択する必要がある。個々の手

[†] 東京電機大学未来科学部
School of Science and Technology for Future Life,
Tokyo Denki University

法が前章で述べたどの用途に適合するかを判断するには次の 3 項目に当てはめて比較することが有用である。

- ① 学習データ生成器が事前の学習を必要とするか
- ② 新しい学習データを生成するための入力値の形式
- ③ 出力画像にラベルやアノテーションは付与されるか
道路標識認識に適用可能な学習データ増強手法の中から選んだ[3,12,13,14]を項目別に比較した結果を表 1 に示す。

表 1 既存の学習データ増強手法の比較

		データ増強手法			
		[3]	[12]	[13]	[14]
事前の学習	不要	○	○	○	
	教師なし学習が必要				○
入力値の形式	3Dや2Dモデル	○			
	1枚の実画像		○		
	複数の実画像			○	
	ランダムノイズ				○
出力画像のラベル	付与される	○	○	○	
	付与されない				○

4.1 事前の学習の要否

事前の学習の要否は量的要件に影響を与える。一部の学習データ増強手法はそれ自身が新しい学習データを生成するために学習を必要とする。例えば Generative Adversarial Networks (GAN)[14]を学習データ増強手法として用いる場合は GAN の学習に大量の教師なしデータが必要となるためデータ収集のコスト削減の目的には適さない。

4.2 入力値の形式

入力値の形式は量的要件と質的要件に影響を与える。実画像に射影変換や明度調整を施して新しい学習用画像を生成する伝統的なデータ増強手法[12]や、複数の実画像をブレンドして新しい学習用画像を生成する手法[13]はシードとなる実画像を予め用意する必要がある。これは実画像の収集コストが高い場合には問題となる。また生成される学習用画像のアピアランスは入力値の実画像に強く依存するため、アピアランスの自由な制御は難しい。

3D や 2D のモデルデータを基にレンダリングして学習用画像を生成する手法[3]は認識対象物の実画像を全く必要とせずに学習用画像を生成できる。またレンダリングのパラメータを調整することで学習用画像のアピアランスを自由に制御できる。ただしこの方法は認識対象物の 3D や 2D モデルが用意できることが前提となる。道路標識は剛体かつ平面的で形状の種類が少ないためモデルデータを用意しやすい。複雑な対象物への応用例として人体の 3D モデルを用いた人検出の学習データ増強手法も提案されている[4]。

ランダムノイズをシードとして画像を生成する GAN はランダムノイズと生成される画像のアピアランスの対応が自明ではないため、レアデータ生成や不均衡データ改善の用途には使用しにくい。一方でアピアランスを直接的には制御する必要のない能動学習には適用できる可能性がある。

4.3 出力画像の正解ラベルやアノテーションの有無

学習データ増強手法により出力される画像に正解ラベルや物体位置のアノテーションが付与されるか否かは、量的

要件に影響を与える。正解ラベルやアノテーションの付与に専門家の協力が必要で高コストになる場合は、生成画像にそれらが付与される手法が望ましい。

5. まとめ

本稿では画像認識における学習データ増強手法の用途を 4 つに分類し、現在提案されている具体的な手法がそれらのどの用途を満たすか判断するための比較項目を提案した。また道路標識に適用できる学習データ増強手法の候補として[3,12,13,14]を例に挙げ、実画像のサンプルが全く得られない状況では[3]の 2D モデルを用いる手法が有望であるとともに、実画像のサンプルが得られる状況では[3,12,13,14]のいずれの手法も使用できることを示した。

今後は、本稿の検討結果を基に複数の学習データ増強手法を組み合わせる効果的な学習データ増強を行い、道路標識の認識率の改善を示したい。

参考文献

- [1] 村瀬洋, “画像認識のための生成型学習”, 情報処理学会論文誌コンピュータビジョンとイメージメディア (CVIM), Vol.46, No.SIG15(CVIM12), pp.35-42 (2005).
- [2] Annika Kuhl, Lars Krüger, Christian Wöhler, Ulrich Kreßel, “Training of Classifiers Using Virtual Samples Only”, Proceedings of the 17th International Conference on Pattern Recognition (ICPR), Vol.3, pp.418-421 (2004).
- [3] 道満恵介, 出口大輔, 高橋友和, 目加田慶人, 井手一郎, 村瀬洋, “色変動を考慮した生成型学習法による道路標識検出器の構築”, 電子情報通信学会論文誌 D, Vol.J93-D, No.8, pp.1375-1385 (2010).
- [4] 土屋成光, 山内悠嗣, 藤吉弘亘, “人検出のための生成型学習と Negative-Bag MILBoost による学習の効率化”, 電気学会論文誌 C, Vol.134, No.3, pp.450-458 (2014).
- [5] 黄瀬浩一, 野口和人, 岩村雅一, “参照特徴ベクトルの増加による低品質画像の高速・高精度認識”, 電子情報通信学会論文誌 D, Vol.J93-D, No.8, pp.1353-1363 (2010).
- [6] Sebastian C. Wong, Adam Gatt, Victor Stamatescu, Mark D. McDonnell, “Understanding Data Augmentation for Classification: When to Warp?”, Proceedings of the 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pp.59-64 (2016).
- [7] Haibo He, Edwardo A. Garcia, “Learning from Imbalanced Data”, IEEE Transactions on Knowledge and Data Engineering, Vol.21, No.9, pp.1263-1284 (2009).
- [8] Dana Angluin, “Queries and Concept Learning”, Machine Learning, Vol.2, No.4, pp.319-342 (1988).
- [9] 福水健次, “能動学習: 3. ニューラルネットワークの能動学習”, 情報処理, Vol.38, No.7, pp.569-574 (1997).
- [10] Kevin J. Lang, Eric B. Baum, “Query Learning Can Work Poorly when a Human Oracle is Used”, Proceedings of the 1992 International Joint Conference on Neural Networks (IJCNN), Vol.2, pp.335-340 (1992).
- [11] Jia-Jie Zhu, José Bento, “Generative Adversarial Active Learning”, arXiv:1702.07956 [cs.LG], (2017).
- [12] Patrice Y. Simard, Dave Steinkraus, John C. Platt, “Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis”, Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR), Vol.2, pp.958-963 (2003).
- [13] Joseph Lemley, Shabab Bazrafkan, Peter Corcoran, “Smart Augmentation Learning an Optimal Data Augmentation Strategy”, IEEE Access, Vol.5, pp. 5858-5869 (2017).
- [14] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, “Generative Adversarial Networks”, arXiv:1406.2661 [stat.ML], (2014).