

DNN を用いた画像分類のための画像強調型データ拡張法

張 偉航[†] 木下 裕磨[†] 貴家 仁志[†]

[†] 首都大学東京システムデザイン学部 〒191-0065 東京都日野市旭ヶ丘 6-6

E-mail: [†]q610714693@gmail.com, ^{††}kinoshita-yuma@ed.tmu.ac.jp, ^{†††}kiya@tmu.ac.jp

あらまし 本稿では DNN のモデルトレーニングのための新しいデータ拡張法として、画像強調型データ拡張法を提案する。DNN を用いた画像分類問題で、一つの重要な課題は十分な学習データセットの準備である。しかし、一般に天候や照明条件の異なる学習セットの準備は困難である、そのため様々なデータ拡張法が研究されている。本稿では、従来高品質な画像生成を目的としていた画像強調法を用いて、様々な条件で撮影された画像を生成し、DNN に適用する。実験の結果、提案法は画像データ拡張法としても有効であることが確認された。また、従来のデータ拡張法と併用できることも確認された。

キーワード DNN, 画像分類, 画像強調, データ拡張法

Image enhancement-based data augmentation for improving deep learning in image classification problem

Weihang ZHANG[†], Yuma KINOSHITA[†], and Hitoshi KIYA[†]

[†] Department of Information and Communication Systems, Tokyo Metropolitan University 6-6 Asahigaoka, Hino-shi, Tokyo, 191-0065 Japan

E-mail: [†]q610714693@gmail.com, ^{††}kinoshita-yuma@ed.tmu.ac.jp, ^{†††}kiya@tmu.ac.jp

Abstract In this paper, we propose a novel data augmentation method based on image enhancement. In DNN-based image classification, it is required to prepare sufficient training datasets taken in different weather conditions and lighting conditions for achieving a high performance. Hence, various data augmentation methods have been studied. In this paper, we propose applying an image enhancement method to data augmentation to generate images with different exposures. Experimental results show that the proposed method is effective as an image data augmentation method. The results also demonstrate that it can be used in combination with other existing data augmentation methods.

Key words Deep learning, Image classification, Image enhancement, Data augmentation

1. ま え が き

近年、DNN(Deep Neural Networks)の急速な発展に伴い、DNNは画像分析および画像分類の問題を解決するための重要な手法になった。しかし、DNNには対処すべき多くの課題もある。DNNは、画像分類タスクを処理するときに数多くのパラメータを学習する必要があるが、有用なトレーニングデータセットが取得できない場合、オーバーフィットなどの問題が発生する。最先端のDNNモデルは非常に深いため、さまざまな照明条件で撮影された大量のトレーニングデータが必要になったが、これらの画像を提供することは非常に困難である。したがって、データ拡張法の研究はますます注目されている。現在、代表的なデータ拡張法は画像に対してアフィン変換と弾性変換

二種類の操作を行う手法である。そこでは元の画像のローテーションまたは反射、ズームイン/ズームアウト、シフト、ディストーション、カラーパレットなどが施される [1], [2]。

本研究では、従来高品質な画像生成を目的としていた画像強調法を用いて、露出値の異なる画像を生成する新しいデータ拡張法を提案する。提案手法では、文献 [3], [4] に基づいた画像セグメンテーションに基づく単一画像ベース多重露出画像合成法を用いて、元画像から露出値が異なる画像を生成し、トレーニングデータセットを拡張する。提案法は画像データ拡張法としての有効性を評価するため、トレーニングデータセットの拡張なしと提案法を行った後のテストセットの分類正解率の比較実験を実施した。実験の結果、提案法は画像データ拡張法としては有効であることが確認された。さらに、提案法は他のデータ

拡張法と併用する可能性を確認するため、従来のデータ拡張法だけを施すと従来法と提案法を併用して施した後のテストセットの分類正解率の比較実験を実施した。実験の結果、提案法は他のデータ拡張法と併用可能であることが確認された。

2. 準備

ここでは、画像強調法の一つである画像セグメンテーションに基づく単一画像ベース多重露出画像合成法について説明する。この方法を用いることによって、画像の輝度分布を M 個の領域に区別して露出補正とトーンマッピングを行い、露出値が異なる M 個の画像を生成することができる [4]。

画像セグメンテーションに基づく単一画像ベース多重露出画像合成法の処理は以下5つのステップで実行される (図1 参照)。

- (1) 入力画像 x から輝度値 l を計算する。
- (2) 局所コントラスト強調

入力画像 x の局所コントラストを強調するために、*dodging and burning* アルゴリズム [5] が使用される。強調された輝度値 l' は次のアルゴリズムから計算する。

$$l'(p) = \frac{l(p)^2}{\bar{l}(p)}. \quad (1)$$

ここで、 $l(p)$ は画像 x の輝度値、 $\bar{l}(p)$ はピクセル p 周辺の輝度の局所平均値である。 $\bar{l}(p)$ は、 l にパイラテラルフィルターを適用することによって計算する。

$$\bar{l}(p) = \frac{\sum_{p' \in P} l(p') g_{\sigma_1}(\|p' - p\|) g_{\sigma_2}(l(p') - l(p))}{\sum_{p' \in P} g_{\sigma_1}(\|p' - p\|) g_{\sigma_2}(l(p') - l(p))}. \quad (2)$$

ここで、 $g_{\sigma}(t)$ はガウシアン関数を用いて計算する。

$$g_{\sigma}(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right), t \in \mathbb{R}. \quad (3)$$

- (3) 露出補正

- (a) 輝度分布に基づく画像セグメンテーション

画像 x を M 個のエリアに分割するために、ガウシアン混合モデルを用いて、輝度分布 $l'(p)$ を計算する。

$$p(l'(p)) = \sum_{k=1}^K \pi_k N(l'(p) | \mu_k, \sigma_k^2), \quad (4)$$

ここで、 K は混合成分、 π_k は k 番目の混合係数、 $N(l'(p) | \mu_k, \sigma_k^2)$ は一次元ガウシアン分布である。GMM を $l'(p)$ に適用させるには、変分ベイジアンアルゴリズムが使用される [6]。このアプローチと大きな K ($K \geq M$) を適用すると、不要な混合成分が自動的に削除される。本稿では、 M の最大値として $K = 10$ を使用する。

ここで、特定の要素 z_k が 1 に等しく、他のすべての要素が 0 に等しい 1-of- K 表現を持つ K 次元バイナリ確率変数を z とする。 z の周辺分布は、以下のような混合係数 π_k によって決定する。

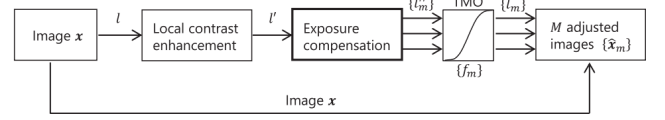


図1 画像セグメンテーションに基づく単一画像ベース多重露出画像合成法

$$p(z_k = 1) = \pi_k, (0 \leq \pi_k \leq 1, \sum_{k=1}^K \pi_k = 1). \quad (5)$$

ピクセル p のクラスター m は、条件付き確率として与えられる γ によって決定される。

$$\begin{aligned} \gamma(z_k = 1 | l'(p)) &= p(z_k = 1 | l'(p)) \\ &= \frac{\pi_k N(l'(p) | \mu_k, \sigma_k)}{\sum_{j=1}^K \pi_j N(l'(p) | \mu_j, \sigma_j)}. \end{aligned} \quad (6)$$

ここで、 m は以下の式を満たす。

$$m = \arg_k \max \gamma(z_k = 1 | l'(p)). \quad (7)$$

- (b) セグメンテーションに基づく露出補正
次式を用いて、 α_m を計算する。

$$G(l | P_m) = \exp\left(\frac{1}{|P_m|} \sum_{p \in P_m} \log(\max(l(p), \epsilon))\right), \quad (8)$$

$$\alpha_m = \frac{0.18}{G(l' | P_m)}. \quad (9)$$

ここで、 ϵ は $l(p) = 0$ での特異点を避けるために小さな値に設定される。

- (c) 露出補正

α_m によって、 l''_m を計算する。

$$l''_m(p) = \alpha_m l'(p). \quad (10)$$

- (4) トーンマッピング

スケールされた輝度値 $l''_m(p)$ は一般的な画像形式の最大値を超えることが多いため、値の切り捨てによりピクセル値が失われる可能性がある。この問題を解決するために、トーンマッピング操作を使用して、輝度値の範囲を $[0, 1]$ に制限させる。

擬似多重露出画像の輝度 $\hat{l}_m(p)$ は、トーンマッピングオペレーター f_m を適用することにより取得される。

$$\hat{l}_m(p) = f_m(l''_m(p)). \quad (11)$$

ここでトーンマッピングオペレーター f_m として Reinhard のグローバルオペレーターを使用する [7]。

$$f_m(t) = \frac{t}{1+t} \left(1 + \frac{t}{L_m^2}\right) (t \in [0, \infty)). \quad (12)$$

パラメータ $L_m > 0$ は、 t の値を $f_m(t) = 1$ として決定する。Reinhard のグローバルオペレーター f_m は単調増加関数であ

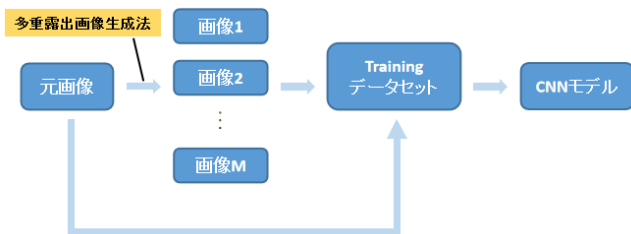


図 2 提案法の概要

るため、 $\max l'_m(p)$ に L_m を設定することにより、輝度値の切り捨てを防ぐことができる。

(5) 多重露出画像を生成

\hat{l}_m , 入力画像 x , およびその輝度 l を組み合わせて、多重露出画像 \hat{x}_m を生成する。

$$\hat{x}_m(p) = \frac{\hat{l}_m(p)}{l(p)} x(p). \quad (13)$$

3. 画像強調型データ拡張法の提案

提案法では、多重露出画像合成法を用いて、一枚の画像から露出値が異なる M 枚の画像を生成し、学習データセットを増加する。さらに、提案法は他のデータ拡張法と併用することができる。提案法の概要を図 2 に示す。

3.1 従来のデータ拡張法

現在、DNN に用いたデータ拡張の最も一般的な方法は、アフィン変換と色変換の組み合わせを実行することである。アフィン変換はローテーション、反射、スケーリング (スムーズイン/スムーズアウト)、せん断の手法を定義される。DNN モデルの学習データ数を増やし [8], データセットの各クラスのバランスをとるため [9], 及びモデルの学習効率を向上するために [10], このような幾何学的な変形は広く使用される。色変換でよく使用されている方法はヒストグラム均等化, コントラスト/明るさの増強, ホワイトバランス, シャープ化及びブラー [11].

さらに、GAN (Generative Adversarial Network) ベースデータ拡張法, ランダム消去法 [12], テクスチャトランスファー [13] などの新しい手法も研究された。本稿では、画像強調法を用いた新しい画像データ拡張法を提案する。

3.2 提案法

従来、画像強調法の目的は、画像の全体的または局所的な特徴を強調し、画像の視覚効果を改善し、高品質の画像を生成することである。前の章で説明した単一画像ベース多重露出画像合成法は、多重露出画像と呼ばれる一連の異なる露出画像を生成する。そして、これらの多重露出画像を合成し、高品質な画像を生成することができる。

本稿では、画像強調法を画質の改善ではなく、一枚の画像から種々の画質の画像を複数枚生成するために使用することを提案する。



図 3 画像例 (GTSRB データベース)

4. 実験

提案手法がデータ拡張法としての有効性を確認するため、また他のデータ拡張法と併用する可能性を確認するため、二つのデータベースでいくつの実験を行った。

4.1 実験条件

本実験では、表 1 のように GTSRB (German Traffic Sign Recognition Benchmark) および Cifar-10 のデータベースを用い、ResNet-18 モデルで各実験のテストセットの分類正解率の比較を行った。

GTSRB はドイツの道路標識画像のデータベース (図 3) であり、制限速度, 進入禁止, でこぼこ道など 43 個のクラスがある。Cifar-10 は日常生活から獲得された画像のデータベースであり、飛行機, 車, 鳥など合計 10 個のクラスがある。二つのデータベースは全部サイズ 32x32 のカラー画像で構成される。

本実験では、提案されたデータ拡張法の有効性を確認するために、二つのデータベースのトレーニングセットの各クラスの画像をランダムに 1/4 まで減少し、新しいデータセット「1/4 トレーニングセット」とする。さらに、1/4 トレーニングデータセットの各クラスの画像をランダムに 1/2 減少し、「1/8 トレーニングセット」とする。また、提案法と従来のデータ拡張法を併用可能性を確認するために、従来法を実施した結果と従来法+提案法を実施した結果も計算した。

本実験で用いた従来法は、トレーニングセットの画像枚数を増加せずに、トレーニングデータが DNN モデルに渡す直前に ± 30 度の範囲内ランダムにアフィン変換と ± 1 の範囲内ランダムにコントラスト増減する方法の組み合わせ (RandomAffine+RandomContrast) である。また、すべてのデータ拡張法はトレーニングデータセットのみに適用され、テストセットには適用されていない。

4.2 実験結果

(1) データ拡張例

図 4 は提案法でトレーニングデータを拡張した画像例を示す。ここで、(a)(e) はそれぞれ Cifar-10 データベースと GTSRB データベースの原画像であり、(b-d)(f-h) は提案法用いて生成された画像である。

(2) テストセット分類の正解率

表 1 実験条件

| データベース | | |
|------------|----------------------|----------|
| データベース名 | GTSRB | Cifar-10 |
| クラス | 43 | 10 |
| トレーニングデータ数 | 39209 | 50000 |
| テストデータ数 | 12630 | 10000 |
| 画像サイズ | 32x32 | 32x32 |
| モデル | | |
| DNN モデル | ResNet-18 | |
| バッチサイズ | 128 | |
| 学習率 | 初期値 0.1(40epoch*0.1) | |
| エポック | 160 | |
| オプティマイザ | SGD | |



図 4 提案法によるデータ拡張例

二つのデータベースでの各実験条件で 160 エポック内の一番高いテストセットの分類正解率 (%) を表 2, 表 3 に示す。

表 2 では, 学習データを拡張なし (表 2 の None 列) の正解率は, それぞれ 97.70, 95.11, 94.70 である。ここで, 提案法を適用し, オリジナルのトレーニングセットを 4 倍に拡張し (一枚の画像から三枚の多重露出画像を生成する), 1/4 と 1/8 トレーニングセットをオリジナルトレーニングセットの枚数に拡張する (1/4 セットを 4 倍に拡張, 1/8 セットを 8 倍に拡張する)。提案法で学習データを拡張した場合 (表 2 の提案法列) の分類正解率は 98.12, 96.43, 95.85 とすべてのトレーニングセット "None" を上回った。さらに, "従来法 + 提案法" では, 従来法を上回った。表 3 も同様の傾向であり, 提案法を適用した場合の正解率は 74.97, 65.51 と "None" を上回った。提案法の有効性及び従来法と併用する有用性が確認できた。

5. おわりに／まとめ

本稿では, 新しいデータ拡張法として, 画像強調型データ拡張法を提案した。画像強調法を用いて, 一枚の学習画像から何枚の露出値が異なる画像を生成して, 学習データセットを拡張する。実験結果より, 提案された画像強調型データ拡張法が有効であることが確認された。さらに, 従来のデータ拡張法と併用可能であることが確認された。

文 献

[1] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," 2018 International Interdisciplinary PhD Workshop

表 2 実験結果 (%) (GTSRB)

| GTSRB | | | | |
|----------------|-------|-------|-------|-----------|
| トレーニングセット (枚数) | None | 従来法 | 提案法 | 従来法 + 提案法 |
| オリジナル (39209) | 97.70 | 97.93 | 98.12 | なし |
| 1/4(9793) | 95.11 | 96.15 | 96.43 | 97.24 |
| 1/8(4886) | 94.70 | 95.14 | 95.85 | 96.70 |

表 3 実験結果 (%) (Cifar-10)

| Cifar-10 | | | | |
|----------------|-------|-------|-------|-----------|
| トレーニングセット (枚数) | None | 従来法 | 提案法 | 従来法 + 提案法 |
| オリジナル (50000) | 85.73 | なし | なし | なし |
| 1/4(12500) | 73.28 | 75.70 | 74.97 | 77.61 |
| 1/8(6250) | 62.82 | 66.57 | 65.51 | 68.99 |

(IIPhDW), pp.117–122, 2018.

- [2] C. Shorten and T.M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol.6, no.1, p.60, 2019.
- [3] Y. Kinoshita and H. Kiya, "Scene Segmentation-Based Luminance Adjustment for Multi-Exposure Image Fusion," IEEE Transactions on Image Processing, vol.28, no.8, pp.4101–4116, 2019.
- [4] Y. Kinoshita and H. Kiya, "Automatic exposure compensation using an image segmentation method for single-image-based multi-exposure fusion," APSIPA Transactions on Signal and Information Processing, vol.7, 2018.
- [5] Y. Huo, F. Yang, and V. Brost, "Dodging and burning inspired inverse tone mapping algorithm," Journal of Computational Information Systems, vol.9, pp.3461–3468, 2013.
- [6] C.M. Bishop, "Pattern recognition and machine learning," Information science and statistics, New York, 2006.
- [7] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," ACM Trans. Graph., vol.21, no.3, pp.267–276, 2002.
- [8] A. Kwasiroch, A. Mikolajczyk, and M. Grochowski, "Deep convolutional neural networks as a decision support tool in medical problems – malignant melanoma case study," Trends in Advanced Intelligent Control, Optimization and Automation, pp.848–856, Cham, 2017.
- [9] A. Kwasiroch, A. Mikolajczyk, and M. Grochowski, "Deep neural networks approach to skin lesions classification – a comparative analysis," 2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR), pp.1069–1074, 2017.
- [10] M. Wasowicz, M. Grochowski, M. Kulka, A. Mikolajczyk, M. Ficek, K. Karpieniko, and M. Cičkiewicz, "Computed aided system for separation and classification of the abnormal erythrocytes in human blood," Biophotonics—Riga 2017, International Society for Optics and Photonics, pp.49–55, 2017.
- [11] A. Galdran, A. Alvarez-Gila, M.I. Meyer, C.L. Saratxaga, T. Araujo, E. Garrote, G. Aresta, P. Costa, A.M. Mendonça, and A.J.C. Campilho, "Data-driven color augmentation techniques for deep skin image analysis," CoRR, 2017.
- [12] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random erasing data augmentation," CoRR, 2017.
- [13] A.A. Efros and W.T. Freeman, "Image quilting for texture synthesis and transfer," Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, pp.341–346, New York, 2001.