https://doi.org/10.2352/EI.2022.34.10.IPAS-191

Class Specific Biased Extrapolation of Images in Latent Space for Imbalanced Image Classification

Suhyeon Jeong and Seungkyu Lee
Department of Computer Science and Engineering, Kyung Hee University, Republic of KOREA

Abstract

Image classification performance using a deep neural network is based on the quality of training images. Well-designed and selected training set representing truth distribution of the classes enable the network to achieve improved accuracy. On the other hand in real applications, class training data imbalance problem limits training performance. Minor classes of relatively smaller training instances suffer from under-training and networks are over-trained on major classes. In this work, we study the effectiveness of prior re-sampling approaches for imbalanced image classification. We propose to investigate inter-class and within-class characteristics and conduct class specific extrapolation re-sampling for optimal imbalanced learning. The proposed algorithm is evaluated on CIFAR-10 data set using a biased extrapolation method.

Introduction

Deep neural network based image classification requires training images of related classes that are well designed and selected to represent truth distribution of the classes. In many real applications, given training instances of a class are not enough to represent the truth distribution. Furthermore, if the number of training images is relatively much smaller (minor classes) than other classes (major classes), a minor class with relatively smaller training samples will be overwhelmed by other major classes in the training. Because deep neural networks prefer to reflect the supervision from major class to optimize given objective function. Addressing such class imbalance problem corresponds to the generalization of minor class. In order to avoid class imbalance problem in the training, re-weighting assigns higher importance on the minor samples so that they can be considered with equal importance in the training. On the other hand, re-sampling methods interpolate new samples of minor classes. For example, Chawla et al. [2] create new samples from existing minor samples using K-nearest neighbors. Methods are referring to the distribution of classes for re-sampling. RAMOBoost [12] determines the ranking of minor classes following respective data distribution. GAMO [10] performs oversampling of minority classes to handle the class imbalance problem. Xu et al. [5] propose an automatic data augmentation method using stochastic natural gradient even though it is a time-consuming complex process. Lin et al. [4] propose a deep reinforcement learning method for imbalanced classification.

However, the class imbalance problem is not just about the imbalance in the number of training samples. It is more about the quality difference of the sample distribution of major and minor classes. With the limited number of training samples, describing every aspect of a class may be limited. Simple re-weighting or oversampling by interpolation may not resolve the problem because they only replicate the description of minor class samples. For example, two classes with an equal number of training samples may have different complexity of truth distribution. One of the classes with complex truth distribution with within-class diversity may require more training samples outside of the current training sample distribution. Sample extrapolation methods [15, 6, 7, 8, 9, 13] are the solution for the problem. Han et al. [6] create new minor samples on the borderline with other classes rather than within minor class samples. Lee et al. [7, 8] propose a feature space extrapolation with deep convolutional neural networks. Decision boundary re-sampling (DBR) finds new minor samples on the decision boundary of latent space. Li et al. [9] propose a margin tuning scheme by using asymmetric large margin loss to move their activation distribution towards under-represented classes across the decision boundary. Similarly, label-distribution-aware margin (LDAM) [13] assigns a larger margin to minor class in the classification. Jeong and Lee [15] propose biased extrapolation method in latent space to concentrate on challenging cases for improved classification.

In this work, we study the effectiveness of prior re-sampling approaches with data sets of diverse inter-class and within-class characteristics. We claim that the uniform application of a re-sampling scheme on a data set is not efficient and effective in many imbalanced learning applications. We propose class specific extrapolation algorithms for the optimal and efficient extrapolation re-sampling. The proposed algorithm is evaluated on CIFAR-10 data set using a biased extrapolation method.

Biased Extrapolation in Latent Space

One of recent approaches for class imbalanced learning is re-sampling new training data between major and minor classes. Usual re-sampling approaches choose confident existing instances and conduct sample interpolation or extrapolation synthesizing new samples. On the other hand, Jeong and Lee [15] claim that re-sampling with confident instances add new but less effective samples. Instead, they synthesize new samples around struggling locations of feature space such as decision boundary between major and minor classes so that new samples investigate and improve the separation of the classes. Sampling by extrapolation enables a deep learning network to be induced to improve decision boundary separation with the new samples. In other words, adding challenging new samples improves the classification providing fine-grained descriptions.

We employ decision boundary re-sampling [8] and biased extrapolation[15] schemes in our study and evaluation. Biased extrapolation between two classes can reveal the characteristics of the classes such as diversity and complexity. Following the training scheme of biased extrapolation[15], the backbone network is trained with class imbalance data, and latent space projection of all training data is conducted. Biased extrapolation creates new minor samples in the latent space. We choose pairs of major and minor classes for the extrapolation. And then, closest minor-major instance pairs are decided for all instances. Biased extrapolation re-sampling is performed in each minor-major instance pair. Decision boundary re-sampling [8] adds the new sample on the minor-major class boundary. On the other hand, biased extrapolation finds the optimal location of new samples based on diverse bias condition. We only use sample probability and asymmetric biases. Sample probability bias selects major instances that are used for the decision boundary re-sampling. Extrapolation is conducted only with the minor-major pairs including selected major instances.

Asymmetric sample probability bias is a type of extrapolation bias that adjusts sampling locations of minor instances based on existing instances of low probability using the following formulation.

$$P_{Asym} = 0.5 + \gamma \cdot P_{max} \cdot (max(1, (d_{normal})/(d_{abnormal})) - 1) \quad (1)$$

This asymmetry bias shifts new minor samples toward the major class regions of feature space.

Class Specific Extrapolation

Due to the difference of class characteristics in a data set, the aspect of imbalance vary along with the pairs of minor-major classes.

First, within-class characteristics of given data set are diverse. Just by the number of training instances, there exist relatively large or small classes. On the other hand, regardless of the number of training instances, there exist classes with given training instances that are sufficient or deficient to describe truth distribution. Large and sufficient class is good enough to be applied as it is. Small and deficient class definitely has to be re-sampled to resolve the imbalance problem. Large but deficient class may not be ignored in a training thanks to the number of training instances, however given set of instances are not enough to describe the class correctly. In this case, similar instances can be sub-sampled and then the class has to be considered as a minor class. On the contrary, small but sufficient class does not suffer from the description problem. It only needs to assign increased importance to the given instances by re-weighting or interpolation re-sampling.

Secondly, a pair of minor-major classes can be either well separated or significantly overlapped with the given data set. Well-separated minor and major classes already form good decision boundary for classification. Therefore re-sampling around the given training samples may not change the decision boundary much showing a limited gain. On the contrary, significantly overlapped minor and major classes form distorted decision boundary favoring major class. In this case, extrapolation re-sampling may function extending the distribution of minor class and shifting original decision boundary toward the major class region.

Therefore uniformly applying extrapolation re-sampling is not efficient and effective in many imbalanced learning applications. Investigating within-class and inter-class characteristics of the given data set enables class specific extrapolation. Unfortunately, it is not possible to explicitly obtain truth distribution of a class or estimate if given training instances are sufficient enough to represent the truth distribution. Only the number of training instances is counted for re-weighting or re-sampling approaches. On the other hand, an inter-class characteristics can be estimated in pair-wise classification performance.

Proposed class specific extrapolation is as follows. First, estimate the number of given training instances of N classes if they are relatively small or large in the data set and assign minor or major classes. Now we have M major classes and m minor classes ($N = M + m$). And then, find closest minor-major pairs in the latent space of trained deep neural network for extrapolation. At each pair, check the binary classification performance of the trained network to decide inter-class separation. Based on a classification performance threshold τ, each pair is decided to be separable or overlapped. Finally, extrapolation is conducted only on the overlapped pair. An alternative method in class specific extrapolation without any fixed threshold, iteration number of extrapolation is decided inversely proportional to the binary classification performance. Algorithms 1 and 2 summarize proposed class specific extrapolation with and without classification performance threshold τ.

Algorithm 1 Class Specific Extrapolation with τ

procedure EXTRAPOLATION($C_1, C_2, ..., C_N$)
    $Minors, Majors \leftarrow EstimateSize(C_k)$
    $Pairs \leftarrow FindPairs(Minors, Majors)$
    $Accuracy \leftarrow Classification(Pairs)$
    for all *Pairs* **do**
        if $Accuracy(Pair(C_i, C_j)) \leq \tau$ **then**
            for Iter **do**
                $\hat{C}_i \leftarrow Extrapolation(Pair(C_i, C_j))$
                $C_i \leftarrow \hat{C}_i$
            end for
        end if
    end for
    return $C_1, C_2, ..., C_N$
end procedure

Algorithm 2 Class Specific Extrapolation without τ

procedure EXTRAPOLATION($C_1, C_2, ..., C_N$)
    $Minors, Majors \leftarrow EstimateSize(C_k)$
    $Pairs \leftarrow FindPairs(Minors, Majors)$
    $Accuracy \leftarrow Classification(Pairs)$
    for all *Pairs* **do**
        for $Iter \times$ @@@@@@$Accuracy(Pair(C_i, C_j))$@@@@@@ **do**
            $\hat{C}_i \leftarrow Extrapolation(Pair(C_i, C_j))$
            $C_i \leftarrow \hat{C}_i$
        end for
    end for
    return $C_1, C_2, ..., C_N$
end procedure

Experimental Evaluation

We evaluate our class specific biased extrapolation algorithm on CIFAR-10 data set. The number of training instances is adjusted to assign Ship, Automobile, Deer, Cat, Bird as minor

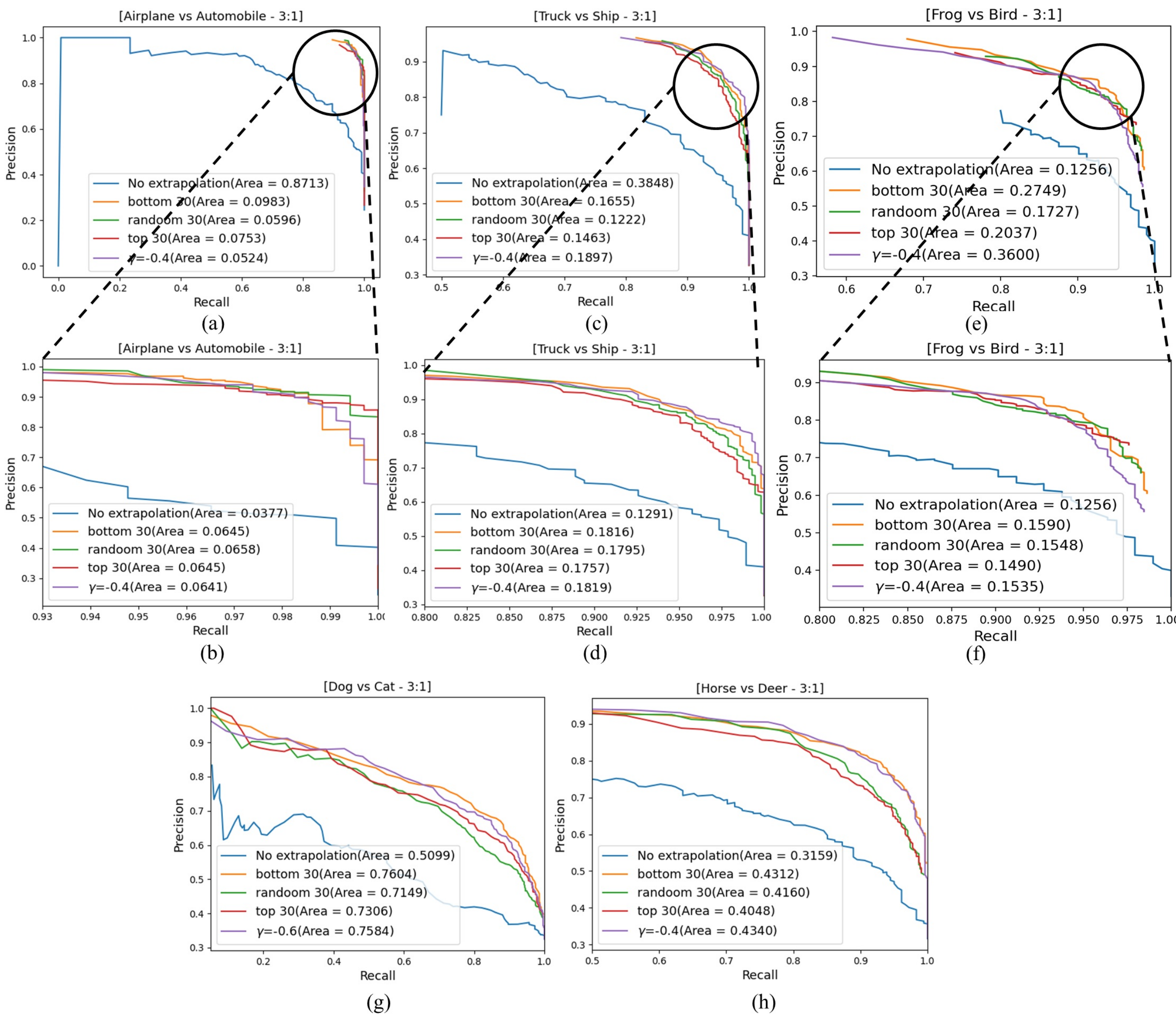

Figure 1. *Precision-Recall curves of five minor-major pairs in Cifar10 data set*

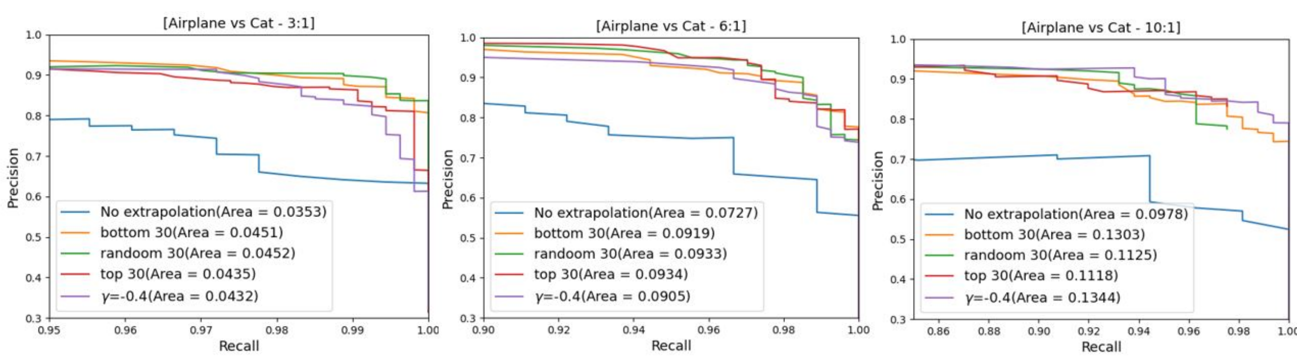

Figure 2. *Precision-Recall curves of Airplane and Cat pair with varying size of minor class*

classes. The total number of major and minor samples is 3000 and 1000. Decided Minor-Major pairs are shown in figure 3. In all experiments, VGG19 is used as our backbone network that is initially trained (epoch = 10). Extrapolation is performed 19 iterations (Iter=19) and 10% of new minor samples are generated and trained 10 epochs more for each iteration. P_{max} and γ are obtained

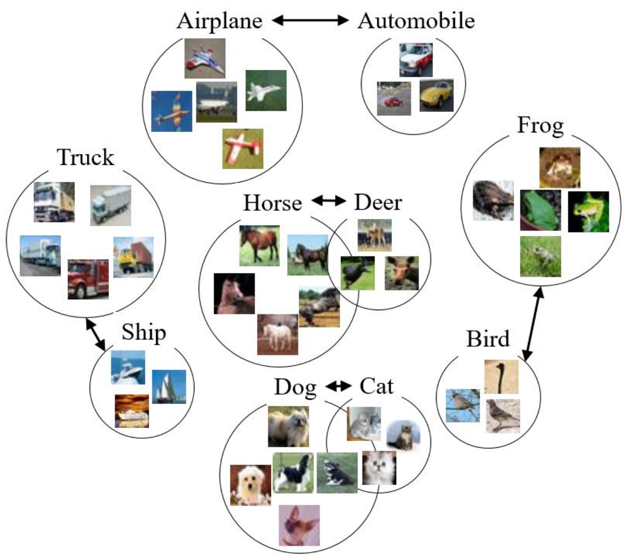

Figure 3. *Minor-Major pairs of Cifar10 data set*

empirically and the probability bias is 30% in all tests. For classification performance evaluation, we employ the precision-recall curve and the value of area under the curve(PR-AUC). Figure 1 shows precision-recall curves of all minor-major pairs before and after biased extrapolation. Each graph compares precision-recall curves of five cases: No extrapolation, Extrapolation with 30% randomly chosen normal samples (Random) [8], Extrapolation with bottom 30% of normal sample probability (Bottom), Extrapolation with top 30% of normal sample probability (Top), Asymmetry extrapolation based on the equation (1).

Separable Classes

If the binary classification accuracy of a pair is higher, extracted features of the classes are more separated. Since airplanes and automobiles are man-made objects, their characteristics are clear and classification accuracy is relatively higher than other pairs. Figure 1-(a) shows classification results of airplanes and automobiles. Figure 1-(b) is an enlarged graph of 1-(a). In this pair, extrapolation itself improves original classification performance significantly. However, all types of biased extrapolations do not show significant improvement as expected. On the other hand in 1-(c) and (d), the classification of trucks and ships show relatively worse original classification performance compared to the pair of airplanes and automobiles. Therefore, biased extrapolations show expected improvement. Asymmetry extrapolation with $\gamma = -0.4$ shows the best AUC and the extrapolation with the bottom 30% of normal sample probability shows the second best improvement in the performance.

Overlapped Classes

Figure 1-(g) shows classification results of dogs and cats classes. Dogs and cats are natural objects and show very similar characteristics across the classes. Therefore major and the minor classes overlap with each other in the latent space distribution. As a result, classification performance is relatively low (in 1-(g)) compared to other pairs indicating that extrapolation re-sampling is required to correct the current decision boundary due to imbalanced training data. In the classification of similar classes, asymmetry extrapolation with $\gamma = -0.4$ and the extrapolation with bottom 30% of normal sample probability show clear improvement compared to random [8] and top extrapolations. Similarly, Horse-Deer pair in figure 1-(h) also shows the characteristic of overlapped class pair. Horses and deer have sharing features, though not as much as dogs and cats do. For this reason, this pair also shows a higher effect of biased extrapolation compared to separated class pairs.

Class Specific Extrapolation

Based on the evaluation above on the CIFAR-10 data set, we apply our class specific extrapolation algorithm with threshold $\tau = 0.7$ on the data set. As a result, out of chosen 5 pairs of minor-major classes (Airplane-Automobile, Truck-Ship, Horse-Deer, Frog-Bird, and Dog-Cat), two overlapped pairs (Horse-Deer, Dog-Cat) are biased extrapolated and remaining pairs are randomly extrapolated (Note that extrapolation itself gives additional training to the networks, and therefore we perform random extrapolation with separable pairs rather than doing nothing with them) keeping overall classification accuracy saving 20.7% of computational cost.

Ablation Study

To see the effect of different ratios in original training instances, we vary the size of the minor classes. In figure 2, the number of major class instances is fixed to 3000. In 3:1, 6:1, and 10:1 cases, the number of minor class instances is set to 1000, 500, and 300, respectively with Airplane and Cat pair. As the number of minor classes decreases, classification performance without extrapolation decreases. In all cases, biased extrapolations increase the classification performance. In the 3:1 case, the sample probability of bottom 30 and asymmetric bias with $\gamma = -0.4$ show the best performance over others as reported in [15]. However, such results cannot be observed with the smaller minor classes such as the 10:1 case where the number of the minor classes is too small to show the meaningful effect of biased extrapolation.

Conclusion

In this work, we have studied the effectiveness of prior re-sampling approaches including biased extrapolation for imbalanced image classification. We proposed to consider inter-class and within-class characteristics and suggested class specific extrapolation re-sampling algorithm for optimal imbalanced learning. The proposed algorithm is evaluated on CIFAR-10 data set using a biased extrapolation method showing the effectiveness of imbalanced learning.

References

[1] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, Ronald Summers, ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases, IEEE CVPR, pp. 3462-3471, 2017

[2] N.V. Chawla., K.W. Bowyer., L.O. Hall., W.P. Kegelmeyer.: Smote: Synthetic minority over-sampling technique. In: J. Artificial Intelligence Research, vol. 16, pp. 321-357. (2002)

[3] S. Bae., K. Yoon.: Polyp detection via imbalanced learning and dis-

criminative feature learning. In: IEEE Transactions on Medical Imaging, pp. 2379-2393. (2015)

[4] E. Lin., Q. Chen., X. Qi.: Deep reinforcement learning for imbalanced classification. In: CoRR (2019)

[5] J. Xu., M. Li., Z. Zhu.: Automatic data augmentation for 3d medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp.378-387. (2020)

[6] H. Han., W.Y. Wang., B.H. Mao.: Borderline-smote : A new oversampling method in imbalanced data sets learning. In: Proc. Int'l Conf. Intelligent Computing. pp.878-887. (2005)

[7] C. Lee., J. Min., J. Cha., S. Lee.: Feature space extrapolation for ulcer classification in wireless capsule endoscopy images. In: The IEEE International Symposium on Biomedical Imaging (ISBI). pp.100-103. (2019)

[8] C. Lee., J. Min., J. Cha., S. Lee.: Decision boundary re-sampling in imbalanced learning for ulcer detection. In: The IEEE Access. pp.186274-186278. (2020)

[9] Z. Li., K. Kamnitsas., B. Glocker.: Overfitting of neural nets under class imbalance: Analysis and improvements for segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp.402-410. (2019)

[10] S. Mullick., S. Data., S. Das.: Generative Adversarial Minority Oversampling. In: IProceedings of the IEEE/CVF International Conference on Computer Vision. pp.1695-1704. (2019)

[11] S. Barua., M. M. Islam., X. Yao., K. Murase.: MWMOTE–majority weighted minority oversampling technique for imbalanced data set learning. In: IEEE Transactions on knowledge and data engineering. pp.405-425. (2012)

[12] S. Chen., H. He., E. A. Garcia.: RAMOBoost: Ranked Minority Oversampling in Boosting. In: IEEE Transactions on Neural Networks. pp.1624-1642. (2010)

[13] K. Cao., C. Wei., A. Gaidon., N. Arechiga., T. Ma.: Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss. In: Conference on Neural Information Processing Systems(NeurIPS) (2019)

[14] K. He, X. Zhang, S. Ren and J. Sun,: Deep Residual Learning for Image Recognition, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

[15] S. Jeong, S. Lee. : Biased Extrapolation in Latent Space for Imbalanced Deep Learning, International Workshop on Machine Learning in Medical Imaging, 2021

Author Biography

Suhyeon Jeong received the B.S. degree in computer engineering from Kyung Hee University, Yongin, South Korea, in 2020. Her research interests include image classification, computer vision, and machine learning.

Seungkyu Lee received the B.S. and M.S. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 1997 and 1999, respectively, and the Ph.D. degree in computer science and engineering from The Pennsylvania State University in 2009. His research interests include time-of-flight depth camera, color/depth image processing, symmetry-based computer vision, and 3-D modeling and reconstruction.