ProMatch: Semi-Supervised Learning with Prototype Consistency

Subjects: Computer Science, Artificial Intelligence Contributor: Ziyu Cheng , Xianmin Wang , Jing Li

Semi-supervised learning (SSL) methods have made significant advancements by combining consistencyregularization and pseudo-labeling in a joint learning paradigm. The core concept of these methods is to identify consistency targets (pseudo-labels) by selecting predicted distributions with high confidence from weakly augmented unlabeled samples.

semi-supervised pseudo-label prototype consistency

1. Introduction

In the past few decades, machine leaning has demonstrated remarkable success across various visual tasks ^{[1][2][3]} ^{[4][5][6][7][8]}. This success can be attributed to advancements in learning algorithms and the availability of extensive labeled datasets. However, in real-world scenarios, the construction of large labeled datasets can be costly and often impractical. Therefore, finding ways to effectively learn from a limited number of labeled data points has become a major concern. This is where semi-supervised learning (SSL) ^{[9][10][11]} comes into play. SSL is an important branch of the machine learning theory and its algorithms, which has emerged as a promising solution to address this challenge by leveraging the abundance of unlabeled data. It has proven to be a remarkable achievement in the field.

The goal of SSL is to enhance the generalization performance by leveraging the potential of unlabeled data. One widely accepted assumption, known as the Low-density Separation Assumption ^[12], posits that the decision boundary should typically reside in low-density regions in order to improve generalization. Building upon this assumption, two prominent paradigms have emerged: pseudo-labeling ^[11] and consistency regularization ^[13]. These approaches have gained significant popularity in the field as effective methods for leveraging unlabeled data in the pursuit of a better generalization performance. Consistency-regularization based methods have become widely adopted methods in SSL. These methods aim to maintain the stability of network outputs when presented with noisy inputs ^{[13][14]}. However, one limitation of consistency-regularization based methods is their heavy reliance on extensive data augmentations, which may restrict their effectiveness in certain domains such as videos and medical images. On the other hand, pseudo-labeling based methods are alternative approaches that have gained popularity in SSL. These methods select unlabeled samples with high confidence as training targets (pseudo-labels) ^[11]. One notable advantage of pseudo-labeling based methods is their simplicity, as they do not require multiple data augmentations and can be easily applied to various domains.

In recent trends, a combination of pseudo-labeling and consistency regularization has shown promising results ^[15] ^{[16][17][18]}. The underlying idea of these methods is to train a classifier using labeled samples and use the predicted distribution as pseudo-labels for unlabeled samples. These pseudo-labels are typically generated by weakly augmented views ^{[16][19]}, or by averaging predictions from multiple strongly augmented views ^[9]. The objective is then constructed by applying the cross-entropy loss between the pseudo-labels and the predictions obtained from different strongly augmented views. It is worth noting that the pseudo-labels are often sharpened or processed using argmax, and each instance is assigned to a specific category to further refine the learning process.

2. Consistency Regularization

Consistency regularization is a commonly used technique in machine learning to improve the generalization ability and stability of models. Specifically, it often employs input perturbation techniques [10][20]. For example, in image classification, it is common to elastically deform or add noise to an input image, which can dramatically change the pixel content of an image without altering its label. In other words, it can artificially expand the size of a training set by generating a near-infinite stream of new, modified data. Up to now, many methods based on the pseudo-label have been proposed. For instance, in ^[20], it increases the variability of the data by incorporating stochastic transformations and perturbations in deep semi-supervised learning. Simultaneously, it minimizes the discrepancy between the predictions of unlabeled samples and their true labels. Temporal Ensembling [21] meets the consistency requirements by minimizing the mean square difference between the predicted probability distributions of the two data-augmented views. Mean Teacher ^[10] further extends this concept by replacing the aggregated predictions with the output of an exponential moving average (EMA) model. In VAT ^[22], consistency regularization is implemented through the introduction of virtual adversarial loss. It perturbs samples in the input space with a small noise and maximizes the adversarial loss, forcing the model to generate consistent predictions for these perturbed samples. To encapsulate, in semi-supervised learning, a classifier should output the same class distribution for an unlabeled example, whether it was augmented or not. For unlabeled points x, in the simplest case, it is achieved by adding a regularization term to the loss function as follows:

 $\| \mathrm{p}_{\mathrm{model}} \left(y \mid \mathrm{Augment} \left(x
ight); heta
ight) - \mathrm{p}_{\mathrm{model}} \left(y \mid \mathrm{Augment} \left(x
ight); heta
ight) \|_{2}^{2}.$

Note that Augment(*x*) is a stochastic transformation, so the two terms in Equation (1) are not identical. VAT ^[22] computes an additive perturbation to apply to the input, which maximally changes the output class distribution. MixMatch ^[9] utilizes a form of consistency regularization through the use of standard data augmentation for images (random horizontal flips and crops). FixMatch ^[19] distinguishes between two degrees of data augmentation (weak and strong): weak augmentation uses standard data augmentation, and strong augmentation may include greater random clipping, rotation, scaling, affine transformation, etc.

3. Pseudo-Labeling

Pseudo-labels are artificial labels generated by the model itself; they aid the model in learning more robust and generalized representations. However, one should be cautious when using pseudo-labels, as prediction results

may include errors or uncertainty, potentially introducing noise. Among pseudo-labeling-based approaches, such as ^[11], they conduct entropy minimization implicitly by constructing hard (one-hot) labels from high-confidence predictions on unlabeled samples. TSSDL ^[23] introduces confidence scores which are determined based on the density of a local neighborhood surrounding each unlabeled sample to measure the reliability of pseudo-labels. In ^[24], it involves training a teacher model on labeled data to generate pseudo-labels for unlabeled data. Then, a noisy student model is trained using the unlabeled data with pseudo-labels. R2-D2 ^[25] attempts to update the pseudo-labels through an optimization framework. It generates pseudo-labels by the decipher model during the repetitive prediction process on unlabeled data. In summation, appropriate measures can ensure the reliability of the pseudo-labeled data. It encourages us to focus on the sample with a high confidence (low entropy) that is away from the decision boundaries.

4. The Combination of Consistency Regularization and Pseudo-Labeling

Some methods ^{[26][27][28][29]} propose to integrate both approaches in a unified framework, which is often called the holistic approach. As one of the pioneering works, FixMatch ^[19] first generates a pseudo-label from the model's prediction on the weakly-augmented instance and then encourages the prediction from the strongly-augmented instance to follow the pseudo-label. Their success inspired many variants that use, e.g., curriculum learning ^{[28][29]}. FlexMatch ^[28] dynamically adjusts the pre-defined threshold in a class-specific manner based on the estimated learning status of each class, which is determined by the number of confident unlabeled data samples. Dash ^[29] dynamically selects the unlabeled data whose loss value does not exceed a dynamic threshold at each optimization step to train learning models. These methods show high accuracy, comparable to supervised learning in a fully-labeled setting.

References

- Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
- 2. Everingham, M.; Van Gool, L.; Williams, C.K.; Winn, J.; Zisserman, A. The pascal visual object classes (voc) challenge. Int. J. Comput. Vis. 2010, 88, 303–338.
- 3. Girshick, R. Fast r-cnn. In Proceedings of the IEEE international Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
- 4. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.

- Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In Proceedings of the Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Proceedings, Part V 13. Springer: Berlin/Heidelberg, Germany, 2014; pp. 740–755.
- Su, X.; Huang, T.; Li, Y.; You, S.; Wang, F.; Qian, C.; Zhang, C.; Xu, C. Prioritized architecture sampling with monto-carlo tree search. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 10968–10977.
- 7. Tang, K.; Ma, Y.; Miao, D.; Song, P.; Gu, Z.; Tian, Z.; Wang, W. Decision fusion networks for image classification. IEEE Trans. Neural Netw. Learn. Syst. 2022, 2022, 1–14.
- 8. Zhu, P.; Hong, J.; Li, X.; Tang, K.; Wang, Z. SGMA: A novel adversarial attack approach with improved transferability. Complex Intell. Syst. 2023, 1–13.
- 9. Berthelot, D.; Carlini, N.; Goodfellow, I.; Papernot, N.; Oliver, A.; Raffel, C.A. Mixmatch: A holistic approach to semi-supervised learning. Adv. Neural Inf. Process. Syst. 2019, 32, 5049–5059.
- Tarvainen, A.; Valpola, H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Adv. Neural Inf. Process. Syst. 2017, 30, 1195–1204.
- 11. Lee, D.H. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In Proceedings of the Workshop on challenges in representation learning, ICML, Atlanta, GA, USA, 20–21 June 2013; Volume 3, p. 896.
- Chapelle, O.; Zien, A. Semi-supervised classification by low density separation. In Proceedings of the International Workshop on Artificial Intelligence and Statistics, PMLR, Bridgetown, Barbados, 6–8 January 2005; pp. 57–64.
- 13. Verma, V.; Kawaguchi, K.; Lamb, A.; Kannala, J.; Solin, A.; Bengio, Y.; Lopez-Paz, D. Interpolation consistency training for semi-supervised learning. Neural Netw. 2022, 145, 90–106.
- Tang, K.; Shi, Y.; Lou, T.; Peng, W.; He, X.; Zhu, P.; Gu, Z.; Tian, Z. Rethinking perturbation directions for imperceptible adversarial attacks on point clouds. IEEE Internet Things J. 2022, 10, 5158–5169.
- 15. Zhang, H.; Cisse, M.; Dauphin, Y.N.; Lopez-Paz, D. mixup: Beyond empirical risk minimization. arXiv 2017, arXiv:1710.09412.
- Berthelot, D.; Carlini, N.; Cubuk, E.D.; Kurakin, A.; Sohn, K.; Zhang, H.; Raffel, C. Remixmatch: Semi-supervised learning with distribution alignment and augmentation anchoring. arXiv 2019, arXiv:1911.09785.
- 17. Kim, D.J.; Choi, J.; Oh, T.H.; Yoon, Y.; Kweon, I.S. Disjoint multi-task learning between heterogeneous human-centric tasks. In Proceedings of the 2018 IEEE Winter Conference on

Applications of Computer Vision (WACV), IEEE, Lake Tahoe, NV, USA, 12–15 March 2018; pp. 1699–1708.

- Kuo, C.W.; Ma, C.Y.; Huang, J.B.; Kira, Z. Featmatch: Feature-based augmentation for semisupervised learning. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part XVIII 16. Springer: Berlin/Heidelberg, Germany, 2020; pp. 479–495.
- Sohn, K.; Berthelot, D.; Carlini, N.; Zhang, Z.; Zhang, H.; Raffel, C.A.; Cubuk, E.D.; Kurakin, A.; Li, C.L. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. Adv. Neural Inf. Process. Syst. 2020, 33, 596–608.
- 20. Sajjadi, M.; Javanmardi, M.; Tasdizen, T. Regularization with stochastic transformations and perturbations for deep semi-supervised learning. Adv. Neural Inf. Process. Syst. 2016, 29, 1171–1179.
- 21. Laine, S.; Aila, T. Temporal ensembling for semi-supervised learning. arXiv 2016, arXiv:1610.02242.
- Miyato, T.; Maeda, S.i.; Koyama, M.; Ishii, S. Virtual adversarial training: A regularization method for supervised and semi-supervised learning. IEEE Trans. Pattern Anal. Mach. Intell. 2018, 41, 1979–1993.
- Shi, W.; Gong, Y.; Ding, C.; Ma, Z.; Tao, X.; Zheng, N. Transductive Semi-Supervised Deep Learning Using Min-Max Features. In Proceedings of the Computer Vision—ECCV 2018, Munich, Germany, 8–14 September 2018; pp. 311–327.
- 24. Xie, Q.; Luong, M.T.; Hovy, E.; Le, Q.V. Self-training with noisy student improves imagenet classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10687–10698.
- 25. Wang, G.H.; Wu, J. Repetitive Reprediction Deep Decipher for Semi-Supervised Learning. In Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020.
- 26. Rizve, M.N.; Duarte, K.; Rawat, Y.S.; Shah, M. In defense of pseudo-labeling: An uncertaintyaware pseudo-label selection framework for semi-supervised learning. arXiv 2021, arXiv:2101.06329.
- 27. Xie, Q.; Dai, Z.; Hovy, E.; Luong, T.; Le, Q. Unsupervised data augmentation for consistency training. Adv. Neural Inf. Process. Syst. 2020, 33, 6256–6268.
- Zhang, B.; Wang, Y.; Hou, W.; Wu, H.; Wang, J.; Okumura, M.; Shinozaki, T. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling. Adv. Neural Inf. Process. Syst. 2021, 34, 18408–18419.

29. Xu, Y.; Shang, L.; Ye, J.; Qian, Q.; Li, Y.F.; Sun, B.; Li, H.; Jin, R. Dash: Semi-supervised learning with dynamic thresholding. In Proceedings of the International Conference on Machine Learning, PMLR, Virtual, 18–24 July 2021; pp. 11525–11536.

Retrieved from https://encyclopedia.pub/entry/history/show/110244