Balance-batch: An Optimized Method for Semantic Segmentation Loss Functions

Yifeng Huang
Fuzhou University
No.2, Xueyuan Road, New District, Fuzhou University, Fuzhou, Fujian Province, China
+86 13015765772
N181120058@fzu.edu.cn

Zhirong Tang
Fuzhou University
No.2, Xueyuan Road, New District, Fuzhou University, Fuzhou, Fujian Province, China
+86 13107603356
111601136@fzu.edu.cn

Kaixiong Su*
Fuzhou University
No.2, Xueyuan Road, New District, Fuzhou University, Fuzhou, Fujian Province, China
+86 13600886028
skx@fzu.edu.cn

***ABSTRACT*-Class-imbalanced data easily generates under-fitting problems in deep neural networks, which seriously limits the performance of the network. Several schemes have been proposed to alleviate the class-imbalance, i.e., data augmentation and network structure optimization. Our work has two main contributions: First, we proposed an optimized method Balance-batch which invests equal consideration for each class in mini-batch and tries to balance the losses of classes. Optimizing loss functions is a simpler and more effective way to tackle the class-imbalance than others. Second, we extend the binary classification Dice loss which is employed in medical images segmentation to multiple classifications. Multi-classifications dice loss reflects the improvement of Balance-batch to unstable loss functions. Our various loss function experiments on Pascal VOC2012 show the effectiveness of Balance-batch, which is beyond the advanced level of these loss function methods.**

Keywords-Class-imbalanced data; deep neural networks; semantic segmentation; mini-batch; loss function.

I. INTRODUCTION

Deep learning has led to the development of pixel-level semantic segmentation [33]. In class-imbalance, the large number of samples focused on minority classes, while majority classes have few samples. As a result, high and low performance emerge in minority and majority classes, respectively. It reduces the average performance. Two methods are researched to tackle class-imbalance. One is data augmentation and the other is algorithm improvement.

A. Date Augmentation

Resampling: Resampling augment the samples to achieve a better balance. Random oversampling directly resamples the majority classes. However, it results in overfitting due to a large amount of duplicate sample [5]. SMOTE optimizes it by randomly selecting a sample x_{ij} from the nearest neighbors of the majority class x_i. The newly synthesized samples are arbitrary points on the line between them. Borderline-SMOTE makes a breakthrough on the sample overlap caused by SMOTE [15]. Under-sampling lessen the redundant samples in those minority classes, which is generally applicable to the case where the samples are extreme disequilibrium [14]. Due to a large amount of data are required in deep learning, oversampling is more appropriate than under-sampling. There are also some traditional extension methods for image processing, i.e., rotation, translation, scaling, and cropping [12, 16, 17]. The expansion of generative adversarial networks (GANs) has recently become a trend [1, 11, 32, 34].

Ensemble learning: Classic ensemble learning methods include Easy-Ensemble and Balance-Cascade [19]. Firstly, arbitrarily sample n times in the majority classes, and select the same number of samples on the minority classes each time to obtain n sample sets. Secondly, training n models in n sample sets. Finally, these models vote for final model results. However, ensemble learning requires more computational expense and time consumption than resampling methods.

B. Algorithm Improvement

Networks: At present, there are some outstanding semantic segmentation networks, i.e., FCN [20], Deeplab series [6-9], U-Net [29], V-Net [22], SegNet [3], etc. FCN combines a fully convolutional neural network with the jump structure. Rough but deep semantic information incorporates detailed but shallow representation information to generate accurate and fine segmentation results. The Deeplab series further improve the precision with atrous convolution, atrous spatial pyramid pooling (ASPP) and conditional random fields (CRF). U-Net and V-Net are employed in medical image segmentation. SegNet which has encoder-decoder architecture is more suitable for remote sensing image segmentation [2].

Loss functions and Evaluation metrics: The Dice loss is mainly employed for medical binary image segmentation [22]. Log loss [27] gives the same weight to each class in back-propagation which is susceptible to class-imbalance. A loss function employed in instance segmentation incorporates Soft IoU loss, Soft box IoU loss, and score cross-entropy (CE) loss to achieve higher accuracy [26]. In allusion to class-imbalance, [10, 18, 31] add the weight inversely proportional to the number of samples to achieve balanced loss. The differences between [18] and [31] is that the former employs soft weight (between [0,1]), which is currently only applicable to the binary classification tasks, while the latter is able to multi-classification tasks by employing weighted CE loss. Semantic segmentation often employs IoU/IoU variants [28], F-score [13, 30], Pixel Accuracy (PA) and Kappa [24] as evaluation metrics. In recent years, it is proved that the loss function based on IoU is a direct way to deal with class-imbalance [4, 23, 25].

978-1-7281-9481-3/20/$31.00 ©2020 IEEE
DOI 10.1109/CVIDL51233.2020.00-74

Compared with the other two methods, it is a trend to directly optimize the loss function to tackle class-imbalance. In this paper, we propose a Balance-batch which can balance distribution in each mini-batch. It ensures that the majority classes and the minority classes have as equal attention as possible. The contributions of our work can be summarized:

(1) Balance-batch ensures the diversity of samples in each mini-batch from the perspective of the loss function.

(2) Expanding dice loss to multi-classifications field, highlight the improvement of balance-batch on unstable loss function.

To our knowledge, Balance-batch is one of the most direct and simple ways to tackle class-imbalance. Experimental results prove that Balance-batch makes a special promotion in unstable and common loss functions, namely, Soft Dice loss, Soft IoU loss, CE loss.

II. METHOD

A. CE Loss

CE loss replaces the mean square error (MSE) loss with the advantages of fast weight update and convergence. CE loss is a commonly used loss function in deep learning. We assume that the dataset contains C categories, $c = \{0, 1, \ldots, C - 1\}$ where $c \in C$. Each batch contains m samples. The set of pixels in the training set is V, where $i \in V$. Let P represent the predicted probability distribution output by the network through Softmax and G represent the real distribution of the data, where $p \in P$, $g \in G$. The CE loss in each batch can be defined as Eq. (1):

$$L_{CE} = -\frac{1}{m}\sum_{c\in C}\sum_{i\in V} g_{ic}\cdot \log_a\left(p_{ic}\right) \tag{1}$$

where p_{ic} represents the probability of i being predicted as c. The value of g_{ic} can be 0 or 1. Section 3.1 adopts Eq. (1) to construct the multi-class CE loss. The gradient of CE loss can be calculated by Eq. (2) to update the parameters:

$$\frac{\partial L_{CE}}{\partial P_{ic}} = -\frac{1}{m}\sum_{j=1}^{m}\sum_{c}\left(\frac{g_{ic}}{a\cdot p_{ic}}\right)_j \tag{2}$$

B. Multi-classifications Soft Dice Loss

Dice similarity coefficient (DSC) represents the degree of similarity. For given set X and Y, the definition of DSC is shown in Eq. (3):

$$DSC\left(X,Y\right) = 2\frac{\left|X\bigcap Y\right|}{\left|X+Y\right|} \tag{3}$$

Dice loss is constructed based on DSC. Combined with Eq. (3), DSC can be rewritten as Eq. (4):

$$DSC = \sum_{c\in C}\frac{\sum_{i\in V} 2TP_{ic}}{\sum_{i\in V} 2TP_{ic} + FP_{ic} + FN_{ic}} \tag{4}$$

We define *TP*, *FP*, and *FN* as the number of true-positive samples, false-positive samples, and false-negative samples. TP_{ic}, FP_{ic}, and FN_{ic} represent the number of true positive-samples, false positive-samples, and false-negative samples of i in the corresponding c, respectively.

According to Eq. (4), DSC has a characteristic of discrete counting, which results in the non-differential of Dice loss function. It deviates from the principle of the loss function. In recent years, a lot of work has been carried out to design it. In this paper, we employ multi-classifications Soft Dice loss as the baseline to accomplish our experiments. Multi-classifications Soft Dice loss in each batch is shown in Eq. (5):

$$L_{Soft-Dice} = -\frac{1}{m}\sum_{c\in C}\frac{\sum_{i\in V} 2g_{ic}\cdot p_{ic}}{\sum_{i\in V} {g_{ic}}^2 + {p_{ic}}^2} \tag{5}$$

The Soft Dice loss converted into a continuous domain has a differentiable property. Dice loss was first proposed to tackle the uneven sample distribution. Most biomedical image segmentation employs the binary classification Dice loss as a loss function. It is commonly to eliminate certain instabilities by adding a smoothing term to both the numerator and denominator. In this paper, to further seek the effectiveness of Balance-batch, the Soft Dice loss adopted in our experiment is in the form of Eq. (5), which adds no smoothing term. We update the parameters according to Eq. (6):

$$\frac{\partial L_{Soft-Dice}}{\partial P_{ic}} = -\frac{1}{m}\sum_{j=1}^{m}\sum_{c\in C}\left[\frac{2g_{ic}\left(1-2p_{ic}^2\right)}{\left(g_{ic}^2+p_{ic}^2\right)^2}\right]_j \tag{6}$$

C. Balance-batch

Traditional semantic segmentation employs mini-batch GD as the training strategy. It is proved that the training effect of small batches is better than that of large batches [21]. Mini-batch GD costs less time and computational expense to obtain higher performance. Mini-batch GD divides all samples M into every single small batch. After training each small batch, loss and gradient can be obtained to back-propagate and update the parameters. It is worth noting that the data in the mini-batch is randomly selected, and the number of samples in each batch is fixed.

We define the maximum batch size *batch_max*, and *batch_cl* is the number of non-repeating categories in each batch. In data loading, we alter the random combination of mini-batch GD to the regular distribution of the classes. Assume that the number of non-repeating categories in the current batch is *n_cl*, and the number of samples is *N*. Firstly, while *n_cl* is not greater than *batch_cl*, it is determined whether *N* has reached *batch_max* in advance. When *N* is less than *batch_max*, the samples and labels which contain the non-repeating categories can be continuously loaded. Once the upper limit *batch_max* is reached and *n_cl* is less than the *batch_cl*, the currently loaded sample and label pairs are

discarded and reselected. Balance-batch ends and small-batch training starts, until the *batch_cl* is satisfied.

We strictly limit the total number of categories in each batch to *batch_cl*. Training with non-repeating categories in the batch can help eliminate the inherent impact of imbalance. Algorithm. 1 shows Balance-batch with multi-classifications Soft Dice loss. It is similar to transfer Balance-batch to other loss functions, except that the replacement of the loss functions is required. The algorithm is easily adopted by various semantic segmentation networks.

Algorithm 1: Soft Dice loss with Balance-batch

Input: The upper limit of batch size *batch_max*;

The number of non-repeating categories in batch *batch_cl*;

Output: $I_c \leftarrow \sum_{i \in V} 2 p_{ic} \cdot g_{ic}$

$U_c \leftarrow \sum_{i \in V} {g_{ic}}^2 + {p_{ic}}^2$

$\left(L_{Soft-Dice}\right)_{batch} \leftarrow -\frac{1}{N} \sum_{c \in C} \frac{I_c}{U_c}$

Initialize *batch_max* and *batch_cl*;

if $n\_cl \le batch\_cl$ **then**

  if $N \le batch\_max$ **then**

    load images and labels;

  else

    reload images and labels;

  end

else

  balance-batch over;

end

return I_c, U_c, $\left(L_{Soft-Dice}\right)_{batch}$

III. EXPERIMENTAL RESULTS

In this section, we employ augmented PASCAL VOC2012 dataset which involves 20 foreground classes and one background class. The number of training images is increased to 10,582, which expanded 9,118 training images. We adopt ResNet-101 as the baseline network and the model is pre-trained on MS-COCO. Combining multi-scale (MSC) and maximum fusion, the input data is scaled at any scale to complete data augmentation. We set the initial learning rate to 0.00025, momentum to 0.9, and weight decay to 0.0005. The models are trained on NVIDIA GTX P100 by iterates for 20,000 times and is saved per 5,000.

A. CE Loss

This section compares the results of CE loss before and after Balance-batch processing. To compare with the experimental results in [7], we employ batch size as a variable, batch size in {1, 2, …, 12}. We find that when batch size = 9, the maximum mIoU = 76.353%, FWIoU = 90.419%, MA = 86.087%, PA = 94.635%. Consistent with the experimental results obtained in [7].

We fix the batch size to 9 and *batch_cl* varies in {2, 3, …, 12}. The results are shown in Figure 1. When we set *batch_cl* to 2, mIoU has reached 73.232%. While *batch_cl* = 8, mIoU = 76.658%, exceeding the highest value of 76.35% in [7]. As *batch_cl* continues to increase, the mIoU processed by Balance-batch is always higher than without processing. When *batch_cl* =12, mIoU reaches the highest value of 76.788%, an increase of 0.438%. FWIoU, MA, and PA are improved accordingly.

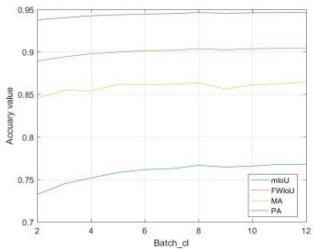

Figure 1. Cross-entropy loss with Balance-batch.

We provide the segmentation results of the best performance models before and after the Balance-batch processing in Figure 2. As we can see from Figure 2, although a preferable segmentation effect has been achieved before Balance-batch processing, some edges are still coarse. After Balance-batch, the edge of the object is refined and the model accuracy is improved.

Figure 2. Result of CE loss on Pascal VOC2012 validation set. (a) ground-truth segmentations, (b) segmentations before Balance-batch, (c) segmentations after Balance-batch.

B. Multi-classifications Soft IoU loss

IoU loss is constructed based on Jaccard index, which is used to measure the degree of similarity between samples. Since the Jaccard index and DSC are similar, we modify the formula to obtain multi-classifications Soft IoU loss for training. Similarly, we vary the batch size to find the best performance before balance-batch. The results are shown in Figure 3(a). We can find that when batch size = 4, the maximum mIoU = 61.945%, FWIoU = 87.287%, MA = 75.418%, and PA = 92.365%.

In the Balance-batch process, we fix the batch size to 4 and let the value of *batch_cl* varies in {2, 3, …, 6}. As shown in Figure 3(b), when *batch_cl* = 3, mIoU has exceeded 61.945% and reached 65.389%. As *batch_cl* increases to 4, mIoU = 65.700%, FWIoU = 87.443%, MA = 81.407%, PA = 92.318%. Comparing with the best result before Balance-batch, our method makes an improvement of 3.755% on mIoU.

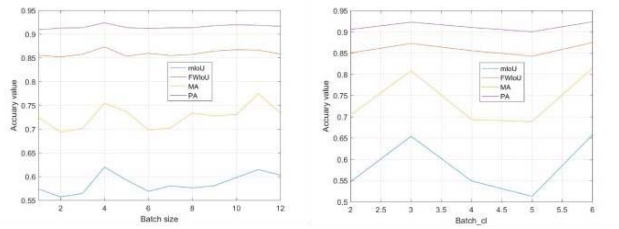

Figure 3. Soft IoU Loss before/after Balance-batch.
(a) before Balance-batch, (b) after Balance-batch.

Figure 4 shows a comparison of the before and after Balance-batch segmentation results. Due to the instability, IoU loss without Balance-batch is prone to misjudgment. In addition, the boundary of the object is over-segmented. In contrast, the segmentation effect of Balance-batch is better than before, especially on small object classes.

Figure 4. Result of Soft IoU loss on Pascal VOC2012 validation set. (a) ground-truth segmentations, (b) segmentations before Balance-batch, (c) segmentations after Balance-batch.

C. *Multi-classifications Soft Dice Loss*

This section employs the multi-classification Soft Dice loss of Eq. (5) for training. As can be seen from Figure 5(a), although the principle of IoU loss is similar to that of Dice loss, the square term of Dice loss eliminates the instability. When batch size = 6, the maximum mIoU = 73.062%, FWIoU = 88.775%, MA = 87.959%, PA = 93.351%.

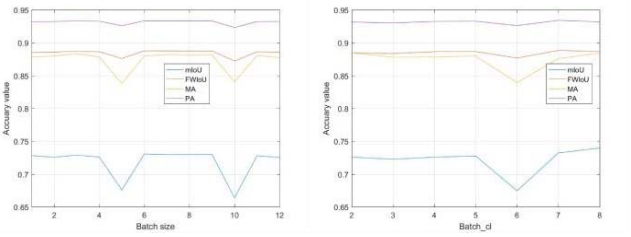

Figure 5. Dice loss before/after Balance-batch. (a) before Balance-batch, (b) after Balance-batch.

Similarly, we fix batch size = 6 and let *batch_cl* changes between {2, 3, …, 8}. As shown in the experiment results in Figure 5(b), when *batch_cl* = 7, mIoU exceeds the original highest 73.062% to 73.211%. When *batch_cl* = 8, the maximum mIoU = 73.964%, an increase of 0.902%. The comparison of the before and after Balance-batch segmentation results are shown in Figure 6.

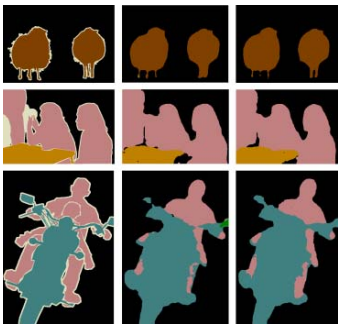

Figure 6. Result of Dice loss on Pascal VOC2012 validation set. (a) ground-truth segmentations, (b) segmentations before Balance-batch, (c) segmentations after Balance-batch.

IV. CONCLUSION

In this paper, we propose Balance-batch, a training strategy that can be adapted to the semantic segmentation loss functions. Balance-batch tackles the class-imbalance from the perspective of each mini-batch. The proposed method enables each mini-batch to contain non-redundant categories to restore balance. Meanwhile, we have expanded the binary classification Dice loss to multi-classifications to highlight the effect of Balance-batch on unstable loss functions. Contrast experimental results on CE loss, Soft Dice loss, and Soft IoU loss demonstrate the effectiveness of the algorithm in Pascal VOC2012. Balance-batch has a certain guiding significance for the research of the class-imbalance in the algorithm. In future work, we will continue the research of class-imbalance with the concern of the network in semantic segmentation.

ACKNOWLEDGMENTS

This research was supported by the National Basic Research Program of China (Item ID: 2017YFB0504202).

REFERENCES

[1] ALHAIJA, H. A., MUSTIKOVELA, S. K., MESCHEDER, L., GEIGER, A., AND ROTHER, C. Augmented reality meets computer vision: Efficient data generation for urban driving scenes. *International Journal of Computer Vision 126*, 9(2018), 961–972.

[2] AUDEBERT, N., LE SAUX, B., AND LEFÈVRE, S. Segment-before-detect: Vehicle detection and classification through semantic segmentation of aerial images. *Remote Sensing 9*, 4 (2017), 368.

[3] BADRINARAYANAN, V., KENDALL, A., AND CIPOLLA, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence 39*, 12 (2017), 2481–2495.

[4] BERMAN, M., RANNEN TRIKI, A., AND BLASCHKO, M. B. The lovász-softmax loss: a tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 4413–4421.

[5] CHAWLA, N. V., BOWYER, K. W., HALL, L. O., AND KEGELMEYER, W. P. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research 16* (2002), 321–357.

[6] CHEN, L.-C., PAPANDREOU, G., KOKKINOS, I., MURPHY, K., AND YUILLE, A. L. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv preprint arXiv:1412.7062* (2014).

[7] CHEN, L.-C., PAPANDREOU, G., KOKKINOS, I., MURPHY, K., AND YUILLE, A. L. Deeplab: Semantic image segmentation with deep

convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence 40*, 4 (2017), 834–848.

[8] CHEN, L.-C., PAPANDREOU, G., SCHROFF, F., AND ADAM, H. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* (2017).

[9] CHEN, L.-C., ZHU, Y., PAPANDREOU, G., SCHROFF, F., AND ADAM, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (2018), pp. 801–818.

[10] CUI, Y., JIA, M., LIN, T.-Y., SONG, Y., AND BELONGIE, S. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2019), pp. 9268–9277.

[11] EATON-ROSEN, Z., BRAGMAN, F., OURSELIN, S., AND CARDOSO, M. J. Improving data augmentation for medical image segmentation.

[12] GARCIA-GARCIA, A., ORTS-ESCOLANO, S., OPREA, S., VILLENA-MARTINEZ, V., AND GARCIA-RODRIGUEZ, J. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857* (2017).

[13] GOUTTE, C., AND GAUSSIER, E. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *European Conference on Information Retrieval* (2005), Springer, pp. 345–359.

[14] HAIXIANG, G., YIJING, L., SHANG, J., MINGYUN, G., YUANYUE, H., AND BING, G. Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications 73* (2017), 220–239.

[15] HAN, H., WANG, W.-Y., AND MAO, B.-H. Borderline-smote: a new over-sampling method in imbalanced data sets learning. In *International conference on intelligent computing* (2005), Springer, pp. 878–887.

[16] HOU, Y., LI, L., LI, B., AND LIU, J. An anti-noise ensemble algorithm for imbalance classification. *Intelligent Data Analysis 23*, 6 (2019), 1205–1217.

[17] KANG, D., AND OH, S. Balanced training/test set sampling for proper evaluation of classification models. *Intelligent Data Analysis 24*, 1 (2020), 5–18.

[18] LIN, T.-Y., GOYAL, P., GIRSHICK, R., HE, K., AND DOLLÁR, P. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (2017), pp. 2980–2988.

[19] LIU, X.-Y., WU, J., AND ZHOU, Z.-H. Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 39*, 2 (2008), 539–550.

[20] LONG, J., SHELHAMER, E., AND DARRELL, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 3431–3440.

[21] MASTERS, D., AND LUSCHI, C. Revisiting small batch training for deep neural networks. *arXiv preprint arXiv:1804.07612* (2018).

.

[22] MILLETARI, F., NAVAB, N., AND AHMADI, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)* (2016), IEEE, pp. 565–571.

[23] NAGENDAR, G., SINGH, D., BALASUBRAMANIAN, V. N., AND JAWAHAR, C. Neuro-iou: Learning a surrogate loss for semantic segmentation. In *BMVC* (2018), p. 278.

[24] NOGUEIRA, K., DALLA MURA, M., CHANUSSOT, J., SCHWARTZ, W. R., AND DOS SANTOS, J. A. Learning to semantically segment high-resolution remote sensing images. In *2016 23rd International Conference on Pattern Recognition (ICPR)* (2016), IEEE, pp. 3566–3571.

[25] RAHMAN, M. A., AND WANG, Y. Optimizing intersection-over-union in deep neural networks for image segmentation. In *International symposium on visual computing* (2016), Springer, pp. 234–244.

[26] REN, M., AND ZEMEL, R. S. End-to-end instance segmentation with recurrent attention. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 6656–6664.

[27] RESHEFF, Y. S., MANDELBAUM, A., AND WEINSHALL, D. Every untrue label is untrue in its own way: Controlling error type with the log bilinear loss. arXiv preprint arXiv:1704.06062 (2017).

[28] REZATOFIGHI, H., TSOI, N., GWAK, J., SADEGHIAN, A., REID, I., AND SAVARESE, S. Generalized intersection over union: A metric and a loss for bounding box regression. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2019), pp. 658–666.

[29] RONNEBERGER, O., FISCHER, P., AND BROX, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (2015), Springer, pp. 234–241.

[30] SOKOLOVA, M., JAPKOWICZ, N., AND SZPAKOWICZ, S. Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation. In *Australasian joint conference on artificial intelligence* (2006), Springer, pp. 1015–1021.

[31] SUDRE, C. H., LI, W., VERCAUTEREN, T., OURSELIN, S., AND CARDOSO, M. J. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, 2017, pp. 240–248.

[32] VOLPI, R., NAMKOONG, H., SENER, O., DUCHI, J. C., MURINO, V., AND SAVARESE, S. Generalizing to unseen domains via adversarial data augmentation. In *Advances in Neural Information Processing Systems* (2018), pp. 5334–5344.

[33] ZHANG, D., HE, F., TU, Z., ZOU, L., AND CHEN, Y. Pointwise geometric and semantic learning network on 3d point clouds. *Integrated Computer-Aided Engineering*, Preprint (2020), 1–19.

[34] ZHANG, H., CISSE, M., DAUPHIN, Y. N., AND LOPEZ-PAZ, D. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412* (2017)