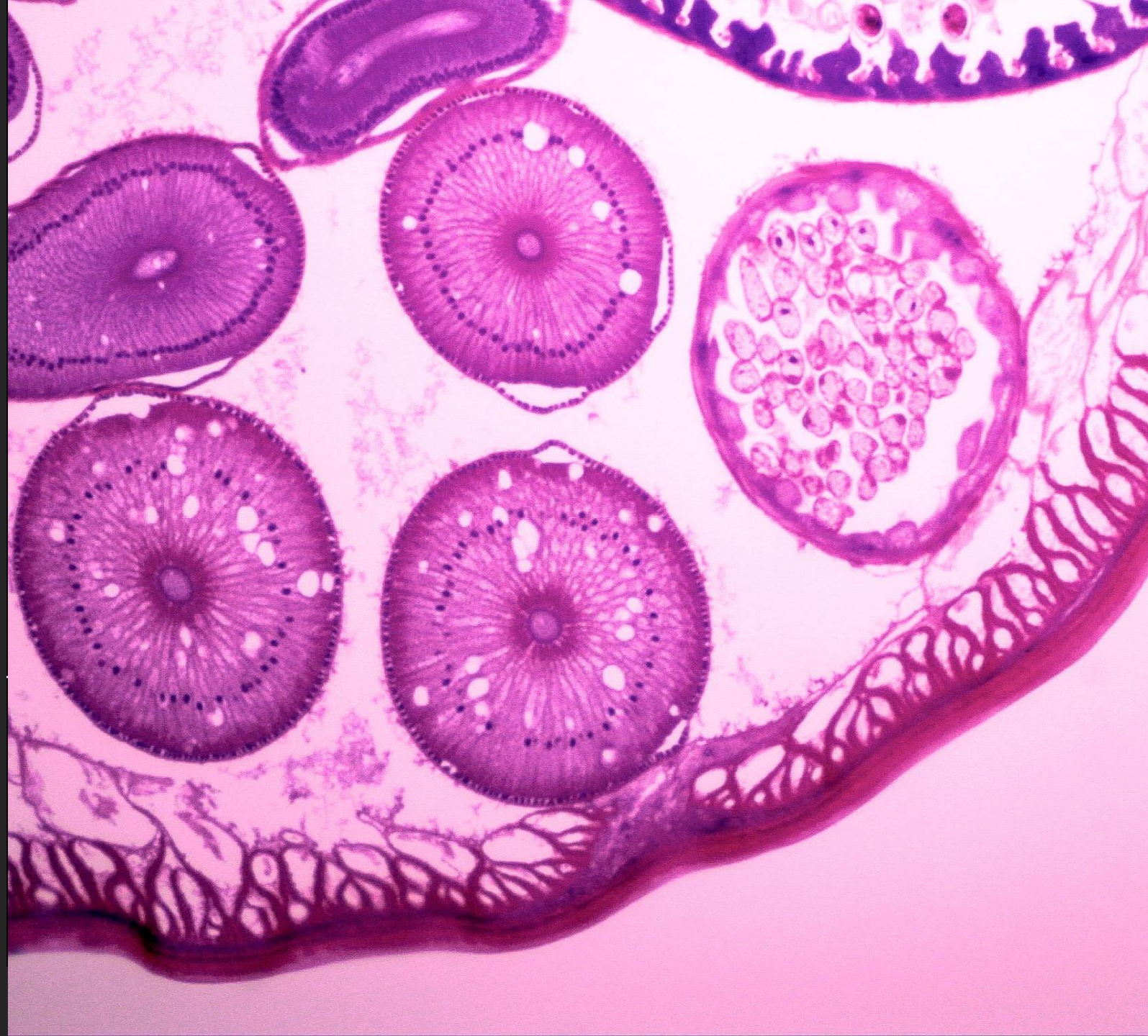
Combining Datasets with Different Label Sets for Improved Nucleus Segmentation and Classification

BIOIMAGING 2024

AMRUTA PARULEKAR, UTKARSH KANWAT, RAVI KANT GUPTA, MEDHA CHIPPA, THOMAS JACOB, TRIPTI BAMETA, SWAPNIL RANE, AMIT SETHI.

INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY, MUMBAI, INDIA

TATA MEMORIAL CENTRE-ACTREC (HBNI), MUMBAI, INDIA



Introduction

Histopathology images are images of stained, biopsied tissue, which are used to clearly view specific cells and tissue abnormalities and to diagnose diseases like cancer and eosinophilia.

Using deep learning techniques for nuclei instance segmentation and classification in histopathology images helps in the timely diagnosis, giving patients an improved prognosis.

Most annotated open-source histopathology datasets differ in the sets of nuclei class labels, magnification, source hospitals, scanning and staining equipment, organs, and diseases.

This proves as a challenge to clinical deployment of models trained on these datasets as they do not test well on custom hospital data.

Hence, for better domain generalization, we need to perform combined training over multiple datasets having different class labels, existing methods for which are unsatisfactory.

Aims and objectives

We propose a method to train DNNs for instance segmentation and classification over multiple related datasets for the same types of objects that have different class label sets.

- Address the issues of class imbalance, staining variability and magnification differences in different nuclei segmentation and classification datasets using various preprocessing techniques.
- Address the issue of different class labels in different nuclei segmentation and classification datasets using a custom loss function and hierarchical representation of available class labels.

- Evaluate the effectiveness of the technique on using test splits.
- Evaluate the effectiveness of the technique for domain generalization.

Datasets used

Multiple histopathology datasets of different organs and magnifications were surveyed

Dataset	Classes	Organs	Mag.	Nuclei	Images	Img. Size
PanNuke[1] (Gamper et al., 2020)	5: Inflammatory, Neoplastic, Dead, Connective, Non-neoplastic Epithelial	19: Bladder, Ovary, Pancreas, Thyroid, Liver, Testis, Prostrate, Stomach, Kidney, Adrenal gland, Skin, Head & Neck, Cervix, Lung, Uterus, Esophagus, Bile-duct, Colon, Breast	40x	216,345	481	224x224
MoNuSAC[2] (Verma et al., 2021)	4: Epithelial, lymphocytes, macrophages, neutrophils	4: Breast, Kidney, Liver, prostrate	40x	46,909	310	82x35 to 1422x2162
CoNSeP[3] (Graham et al., 2019)	7: Healthy Epithelial, Inflammatory, Muscle, Fibroblast, Malignant Epithelial, Endothelial,Other	1: Colon	40x	24,319	41	1000x1000

[1] Gamper, N. A. Koohbanani, K. Benes, S. Graham, M. Jahanifar, S. A. Khurram, A. Azam, K. Hewitt, and N. Rajpoot, "Pannuke dataset extension, insights and baselines," 2020.
[2] R. Verma, N. Kumar, et al, "Monusac2020: A multi-organ nuclei segmentation and classification challenge," IEEE Transactions on Medical Imaging, vol. 40, no. 12, pp. 3413–3423, 2021.
[3] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, and N. Rajpoot, "Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," 2019.

Data Preprocessing

Problem: Magnification differences in datasets

Solution: 256x256 patches, corresponding segmentation and classification masks extracted having similar magnification

Problem: Class imbalance of nuclei

Solution: Geometric augmentations (flips, rotations) and elastic augmentations to resample less frequent classes

Problem: Staining variability of H&E dye

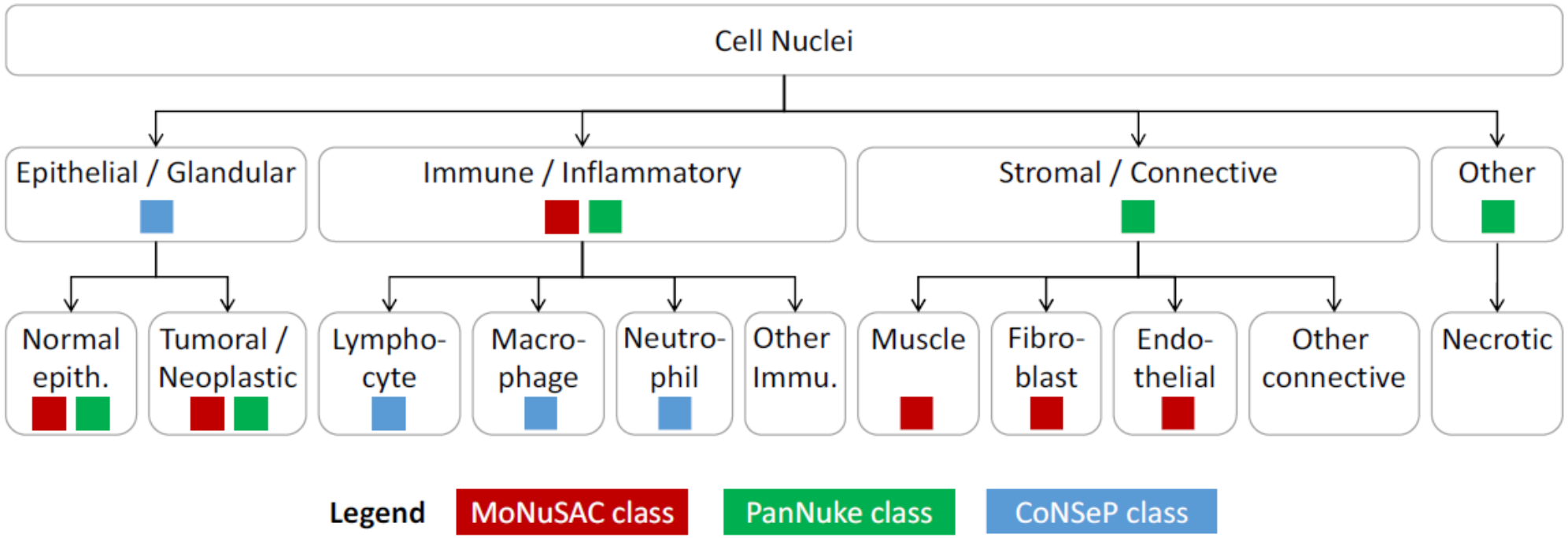
Solution: Random brightness, saturation and hue augmentations (Alternative – Colorjitter augmentation)

Class Hierarchy Tree

Fine-grained classification

Super-classes

Sub-classes



Technique - Custom loss function

IF ONLY SUPER-CLASS LABELS AVAILABLE, WE SUM THE PREDICTED SUB-CLASS PROBABILITIES, ELSE WE USE SUB-CASS PROBABILITIES

THE SET OF SUB-CLASSES TO BE COMBINED IN THE LOSS CAN DYNAMICALLY CHANGE OVER INSTANCES, BATCHES, EPOCHS OR DATASETS

Cross entropy loss $L_{CE} = -\sum_{i=1}^{n}\sum_{i=1}^{c} t_{ij}\log(y_{ij})$ modified to

$$L_{MCE} = -\sum_{i=1}^{n}\sum_{k=1}^{m} t_{ik}\log y_{ik} = -\sum_{i=1}^{n}\sum_{k=1}^{m} t_{ik}\log\left(\sum_{j\in S_k} y_{ij}\right)$$

Focal Tversky loss [4]

$$L_{FT} = \sum_{i=1}^{n}\left(1 - \frac{\sum_{j=1}^{c}(t_{ij}y_{ij}+\varepsilon)}{\alpha\sum_{j=1}^{c}t_{ij} + (1-\alpha)\sum_{j=1}^{c}y_{ij}+\varepsilon}\right)^{\gamma}$$ modified to

$$L_{MFT} = \sum_{i=1}^{n}\left(1 - \frac{\sum_{k=1}^{m}(t_{ik}\sum_{j\in S_k}y_{ij}+\varepsilon)}{\alpha\sum_{k=1}^{m}t_{ik} + (1-\alpha)\sum_{k=1}^{m}\sum_{j\in S_k}y_{ij}+\varepsilon}\right)^{\gamma}$$

Where tij are one-hot labels, yij are class probabilities, k are super classes – denoted by j∈Sk, n is number of training instances and c is number of classes, m is number of super classes, tik is a binary indicator for super class labels, ε is a small constant to prevent division by 0, $\alpha > 0$, $\gamma > 0$ are hyper-parameters

[4] Abraham, N. and Khan, N. M. (2018). A novel focal tversky loss function with improved attention u-net for lesion segmentation.

Model – StarDist [5][6]

It segments objects accurately by approximating them with **star-convex polygons**

UNet backbone with additional heads that predict for each pixel:

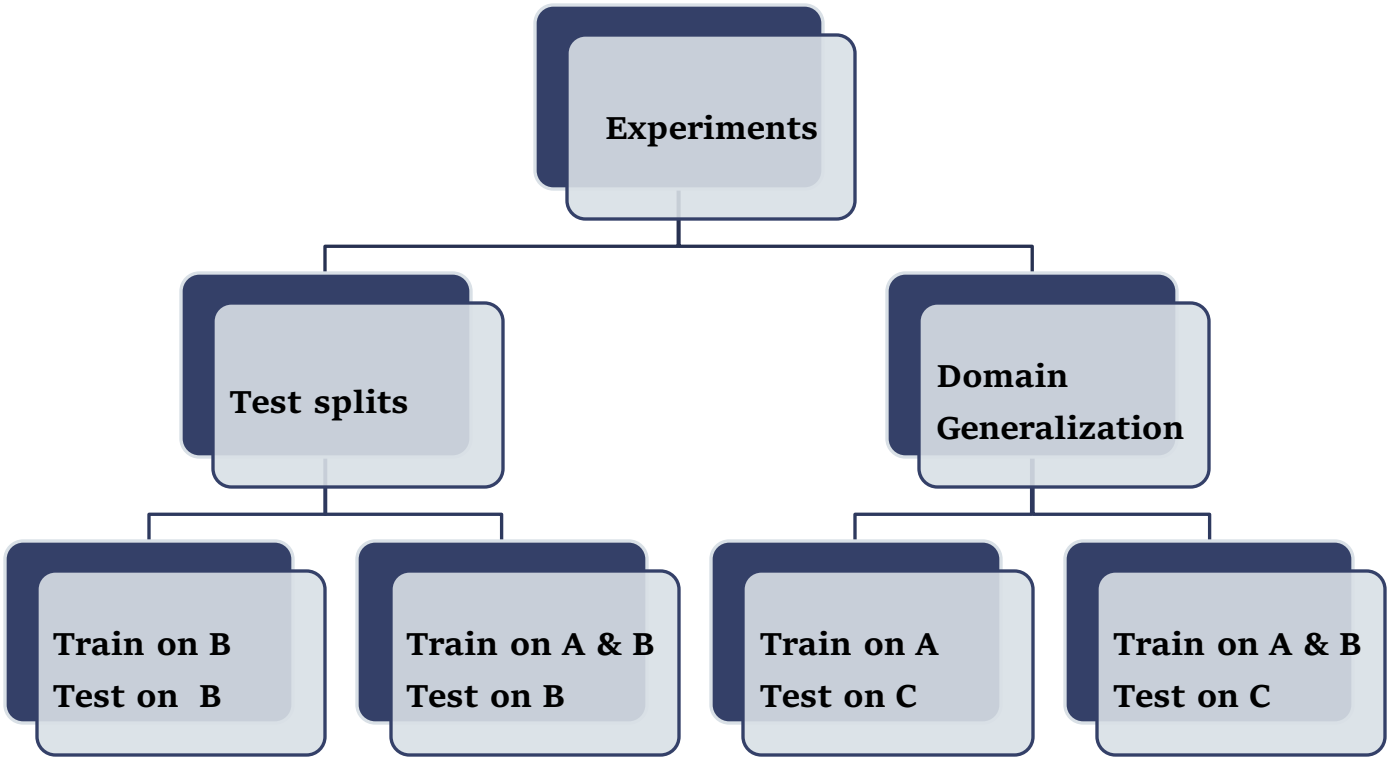
- Object probability
- Class probability
- Radial distance to object boundary

Optimizer used: **Adam**

Non-maximum suppression done to remove multiple polygons representing the same object

[5] U. Schmidt, M. Weigert, C. Broaddus, and G. Myers, "Cell detection with star-convex polygons," in Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer International Publishing, 2018, pp. 265–273. [Online]. Available: https://doi.org/10.1007%2F978-3-030-00934-2_30
[6] M. Weigert and U. Schmidt, "Nuclei instance segmentation and classification in histopathology images with stardist". in 2022 IEEE International Symposium on Biomedical Imaging Challenges (ISBIC). IEEE, mar 2022. [Online]. Available: https://doi.org/10.1109%2Fisbic56247.2022.9854534

Experiments conducted



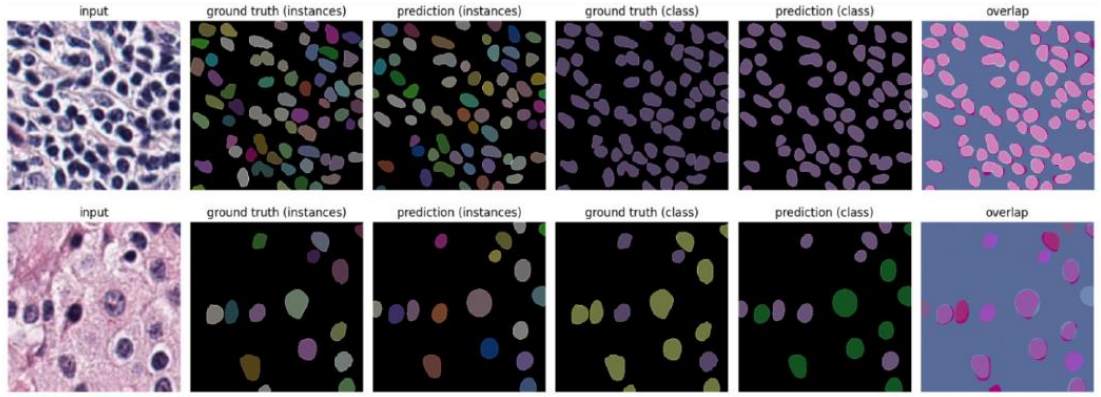
Metric used: Panoptic Quality[7]

$$\mathrm{PQ} = \underbrace{\frac{\sum_{(p,g)\in TP} \mathrm{IoU}(p,g)}{|TP|}}_{\text{segmentation quality (SQ)}} \times \underbrace{\frac{|TP|}{|TP| + \frac{1}{2}|FP| + \frac{1}{2}|FN|}}_{\text{recognition quality (RQ)}}$$

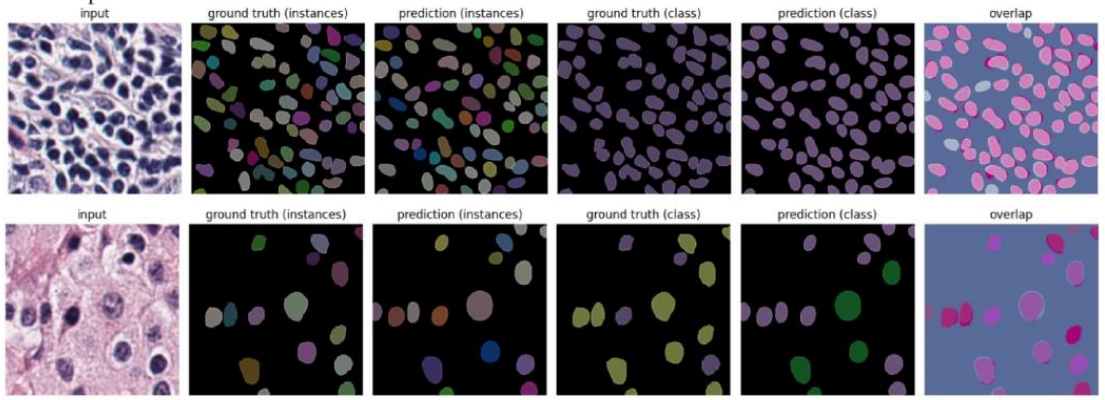
Where,
TP = True positive predicted pixels
FN = False negative predicted pixels
FP = False positive predicted pixels
IoU= Intersection of prediction and ground truth pixels divided by their union (TP)/(TP + FP + FN)
p=Predicted pixels
g=Ground truth pixels

[7] Kirillov, A. et al. (2019). Panoptic segmentation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 9404-9413

Labels splits results



(a) Predictions on MoNuSAC of the model pretrained on PanNuke for 250 epochs followed by fine-tuning on MoNuSAC for 130 epochs

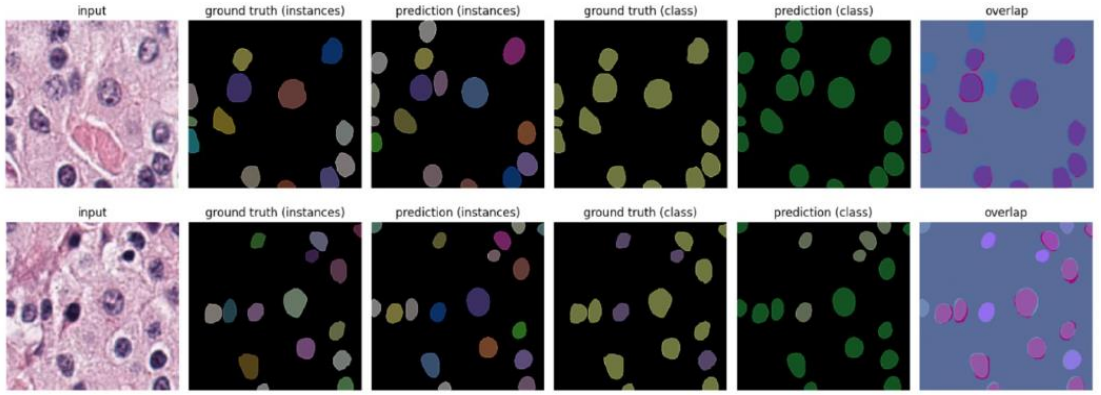


(b) Predictions on MoNuSAC of the model trained on MoNuSAC for 175 epochs before overfitting starts to occur

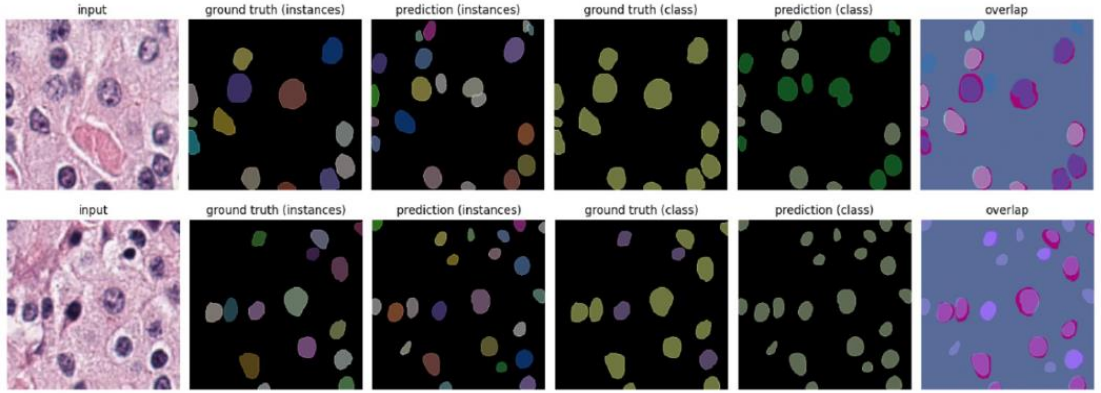
Pre-Train	Epochs	Fine-tune	Epochs	Test	PQ
CoNSeP	100	-	-	CoNSeP	0.540
MoNuSAC	175	CoNSeP	75	CoNSeP	**0.555**
PanNuke	250	CoNSeP	75	CoNSeP	**0.571**
MoNuSAC	175	-	-	MoNuSAC	0.579
CoNSeP	100	MoNuSAC	130	MoNuSAC	**0.587**
PanNuke	250	MoNuSAC	130	MoNuSAC	**0.602**
PanNuke	250	-	-	PanNuke	**0.610**
CoNSeP	100	PanNuke	187	PanNuke	0.605
MoNuSAC	175	PanNuke	187	PanNuke	**0.610**

Training on dataset A followed by training on dataset B gives better results on dataset B as compared to training on dataset B alone.

Domain Generalization results



(a) Predictions on MoNuSAC of the model pre-trained on CoNSeP for 100 epochs followed by fine-tuning on PanNuke for 62 epochs

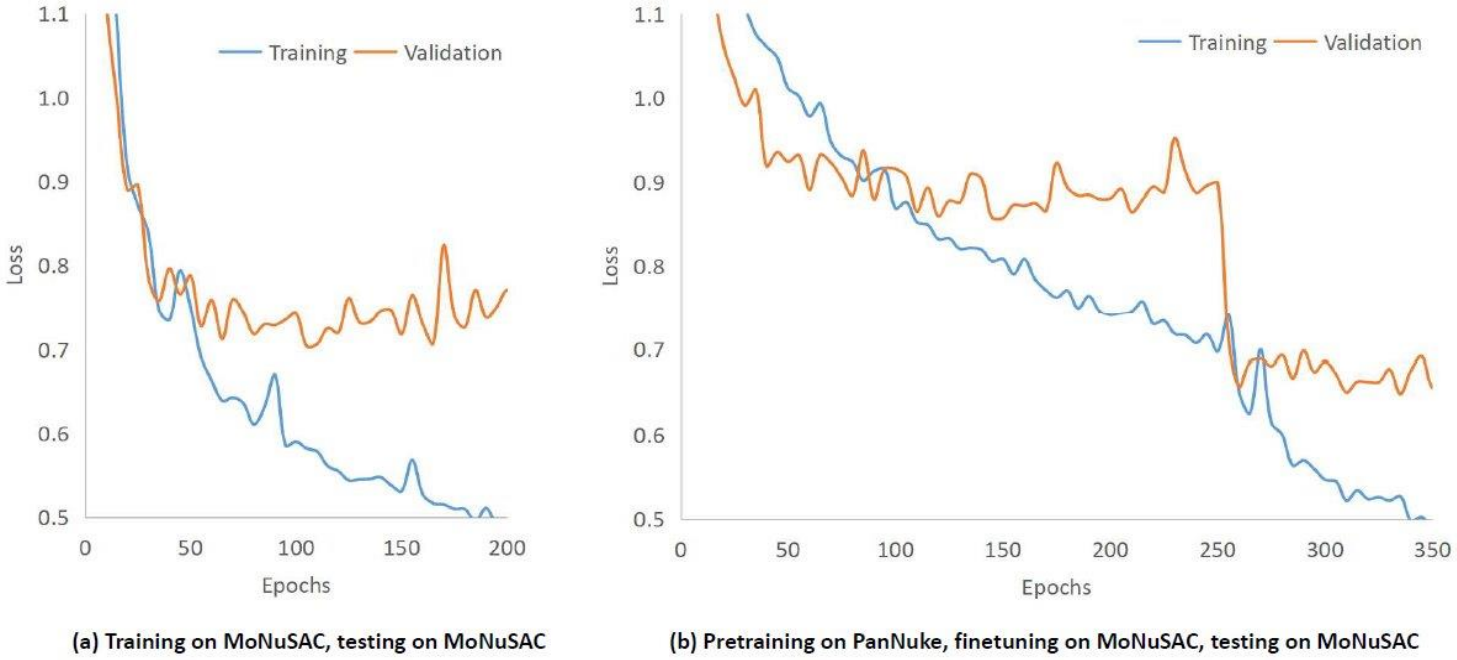


(b) Predictions on MoNuSAC of the model trained on CoNSeP for 100 epochs, before overfitting starts to occur

Pre-Train	Epochs	Fine-tune	Epochs	Test	PQ
CoNSeP	100	-	-	MoNuSAC	0.433
CoNSeP	100	PanNuke	62	MoNuSAC	**0.563**
CoNSeP	100	-	-	PanNuke	0.433
CoNSeP	100	MoNuSAC	43	PanNuke	**0.434**
MoNuSAC	175	-	-	CoNSeP	0.344
MoNuSAC	175	PanNuke	62	CoNSeP	**0.449**
MoNuSAC	175	-	-	PanNuke	0.396
MoNuSAC	175	CoNSeP	25	PanNuke	**0.405**

Training on dataset A followed by training on dataset B gives better results on dataset C as compared to training on dataset A alone.

Discussion – Training strategies



(a) Training on MoNuSAC, testing on MoNuSAC

(b) Pretraining on PanNuke, finetuning on MoNuSAC, testing on MoNuSAC

For test split results, we can observe that the improvement is more pronounced when the pretraining dataset is more generalized and has a super-set of classes and organs as compared to the target dataset.

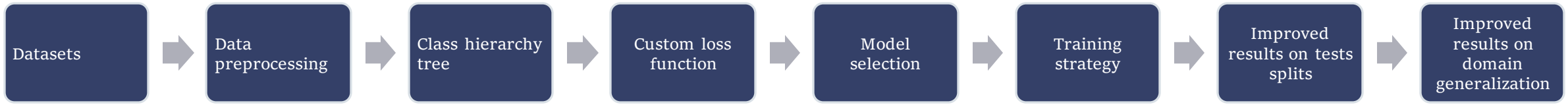
For domain generalization, we can observe that a more pronounced improvement occurs when the fine-tuning dataset is more generalized and has a super-set of classes and organs as compared to the other datasets.

Scope - Generalization

Datasets	Loss functions	Model architectures
• This technique can be used with datasets having distinct as well as overlapping class labels which can be structured as a hierarchy.	• Loss functions must be additive, i.e. involving sums over predicted and ground truth class probabilities while summing over instances or pixels.	• This technique can be applied to any deep neural network architectures trainable using these loss functions.

Conclusion

Novel method for combining datasets with different label sets



Achieved state-of-the-art panoptic quality on the CoNSeP dataset

Model	PQ		Model	PQ
UNet++ [29]	0.405		Hovernet [1]	0.532
DiffMix [30]	0.505		GradMix [31]	0.504
DSCANet [32]	0.426		DRCANet [33]	0.546
SMILE [34]	0.530		DeepLabV3+ [27]	0.373
Stardist [2]	0.540		Our Method	**0.571**

References

[1] Gamper, N. A. Koohbanani, K. Benes, S. Graham, M. Jahanifar, S. A. Khurram, A. Azam, K. Hewitt, and N. Rajpoot, "Pannuke dataset extension, insights and baselines," 2020.

[2] R. Verma, N. Kumar, et al, "Monusac2020: A multi-organ nuclei segmentation and classification challenge," IEEE Transactions on Medical Imaging, vol. 40, no. 12, pp. 3413–3423, 2021.

[3] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, and N. Rajpoot, "Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," 2019.

[4] Abraham, N. and Khan, N. M. (2018). A novel focal tversky loss function with improved attention u-net for lesion segmentation.

[5] U. Schmidt, M. Weigert, C. Broaddus, and G. Myers, "Cell detection with star-convex polygons," in Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Springer International Publishing, 2018, pp. 265–273. [Online]. Available: https://doi.org/10.1007%2F978-3-030-00934-2 30

[6] M. Weigert and U. Schmidt, "Nuclei instance segmentation and classification in histopathology images with stardist". in 2022 IEEE International Symposium on Biomedical Imaging Challenges (ISBIC). IEEE, mar 2022. [Online]. Available: https://doi.org/10.1109%2Fisbic56247.2022.9854534

[7] Kirillov, A. et al. (2019). Panoptic segmentation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 9404–9413