# CLASS-M: Adaptive stain separation-based contrastive learning with pseudo-labeling for histopathological image classification

Bodong Zhang, Hamid Manoochehri, Man Minh Ho, Fahimeh Fooladgar, Yosep Chong, Beatrice Knudsen, Deepika Sirohi, Tolga Tasdizen

## Introduction

- Histopathological image classification is an important task in medical image analysis.
- Achieving a highly accurate model requires a substantial amount of training labels, demanding significant time and effort from human experts.
- We propose a semi-supervised patch-level histopathological image classification model: Contrastive Learning with Adaptive Stain Separation and MixUp (CLASS-M), that does not require extensively labeled datasets.
- We apply our novel contrastive learning on Hematoxylin images and Eosin images after performing slide-level stain separation. We also use pseudo-labeling with MixUp to further improve classification accuracy.
- We provide the new Utah clear cell renal cell carcinoma (ccRCC) dataset which has 49 WSIs with patch-level labels, and patch-level labels for 150 WSIs from the TCGA ccRCC dataset.
- We test various state-of-the-art semi-supervised learning and self-supervised learning methods on both ccRCC datasets and demonstrate that CLASS-M outperforms other state-of-the-art models.

## Model design

Contrastive loss:

$$\mathcal{L}_{c.t.}(x_i) = \max\left(\| f_H(x_i) - f_E(x_i) \|_2 - \| f_H(x_i) - f_E(x_k) \|_2 + m, 0\right)$$

Pseudo-labeling:

$$\bar{y}_i = \frac{1}{K} \sum_{1 \le k \le K} P_{model}(x_{i,k}),\ x_i \in U$$

$$y_i = sharpen(\bar{y}_i, T), \text{where } sharpen(y, T)_{(c)} = \frac{{y_{(c)}}^{\frac{1}{T}}}{\sum_{j=1}^{C} {y_{(j)}}^{\frac{1}{T}}}$$

MixUp samples:

$$x' = \lambda' x_i + (1 - \lambda') x_j$$

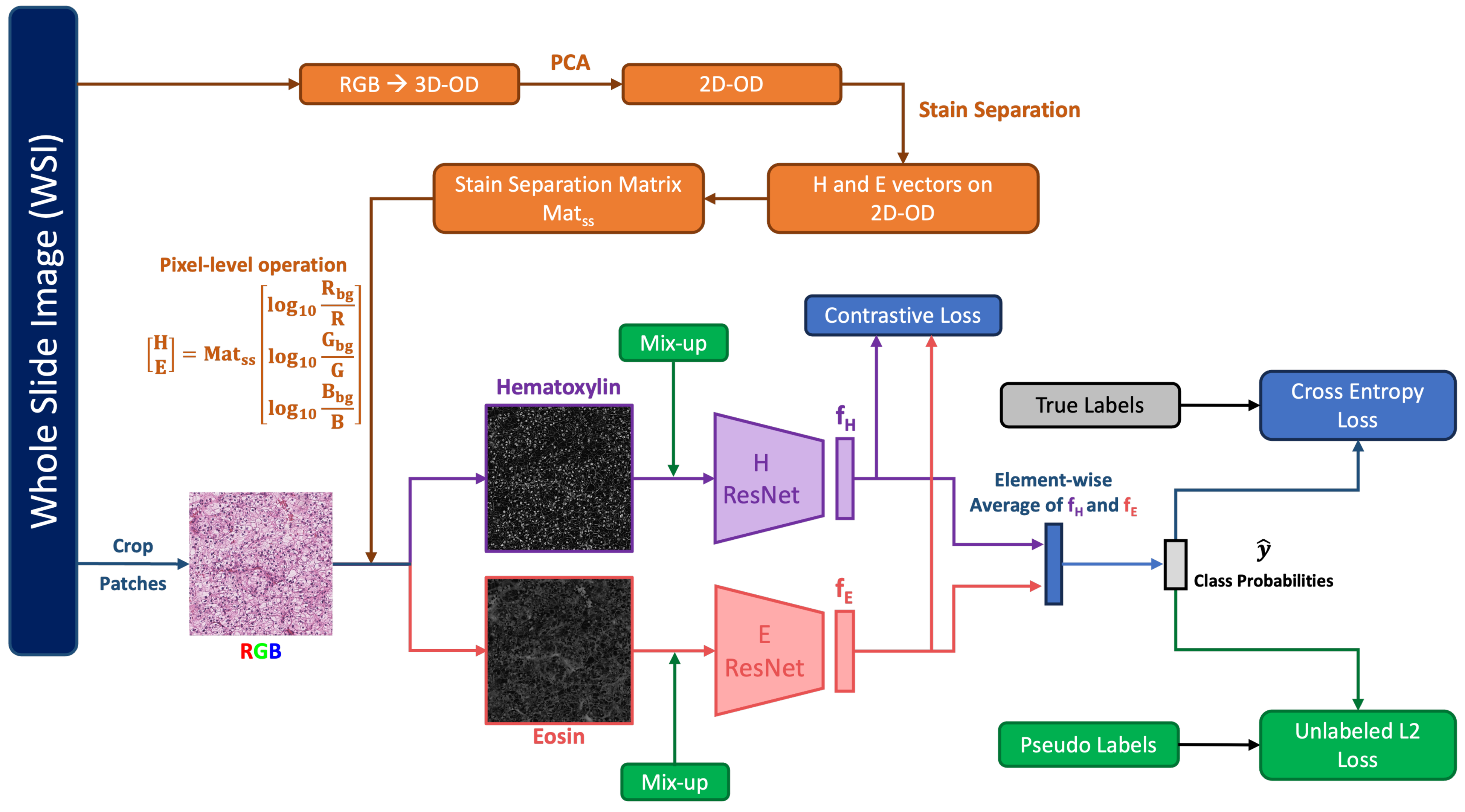
$$y' = \lambda' y_i + (1 - \lambda') y_j$$

Total loss:

$$\mathcal{L} = \underbrace{\sum_{x_i \in L'} \frac{y_i \log \hat{y}_i}{|L'|}}_{\text{Label prediction loss (labeled set)}} + \lambda_{U'} \underbrace{\sum_{x_i \in U'} \frac{\| y_i - \hat{y}_i \|_2{}^2}{C|U'|}}_{\text{Label prediction loss (unlabeled set)}} + \lambda_C \underbrace{\sum_{x_i \in L' \cup U'} \mathcal{L}_{c.t.}(x_i)}_{\text{Contrastive loss (labeled + unlabeled set)}}$$

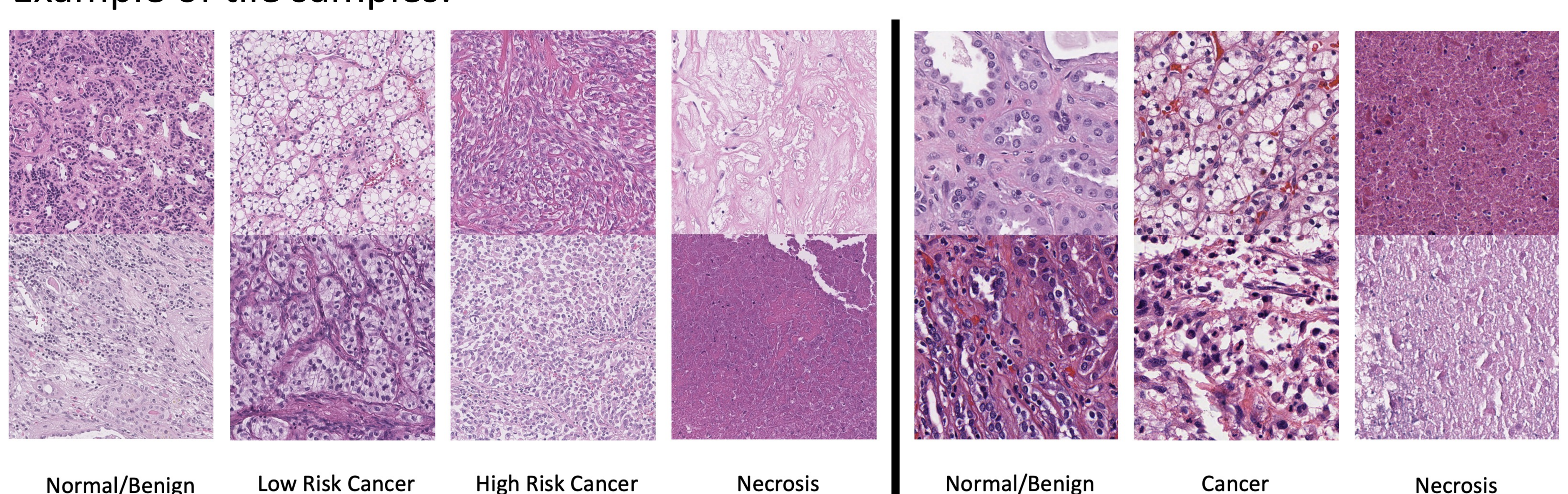
- Paper link:
  arxiv.org/abs/2312.06978
- Code link:
  github.com/BzhangURU/Paper_CLASS-M/tree/main

THE UNIVERSITY OF UTAH




## Experiment settings

Example of tile samples:




(a) Tiles from Utah ccRCC dataset

(b) Tiles from TCGA ccRCC dataset

Details of experiment settings:

Utah ccRCC dataset

| Number of tiles | Normal/Benign | Low Risk Cancer | High Risk Cancer | Necrosis | Unlabeled |
|---|---|---|---|---|---|
| Training | 28,497 | 2,044 | 2,522 | 4,115 | 171,113 |
| Validation | 5,472 | 416 | 334 | 2,495 | |
| Test | 7,263 | 598 | 389 | 924 | |

TCGA ccRCC dataset

| Number of tiles | Normal/Benign | Cancer | Necrosis | Unlabeled |
|---|---|---|---|---|
| Training | 84,578 | 180,471 | 7,932 | 1,373,684 |
| Validation | 19,638 | 79,382 | 1,301 | |
| Test | 15,323 | 62,565 | 6,168 | |

• The work was funded in part by NIH 1R21CA277381. We also acknowledge the support of the Computational Oncology Research Initiative (CORI) at the Huntsman Cancer Institute, ARUP Laboratories, and the Department of Pathology at the University of Utah.

## Experiment results

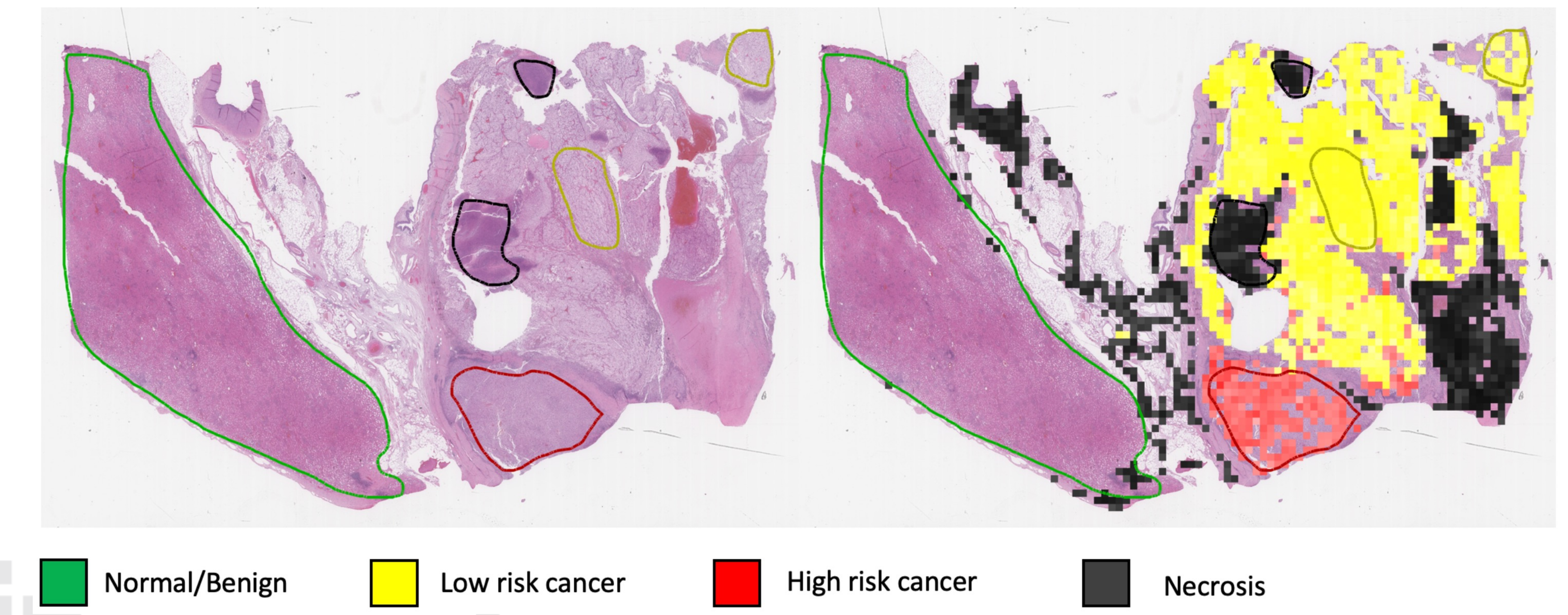
Model comparisons:

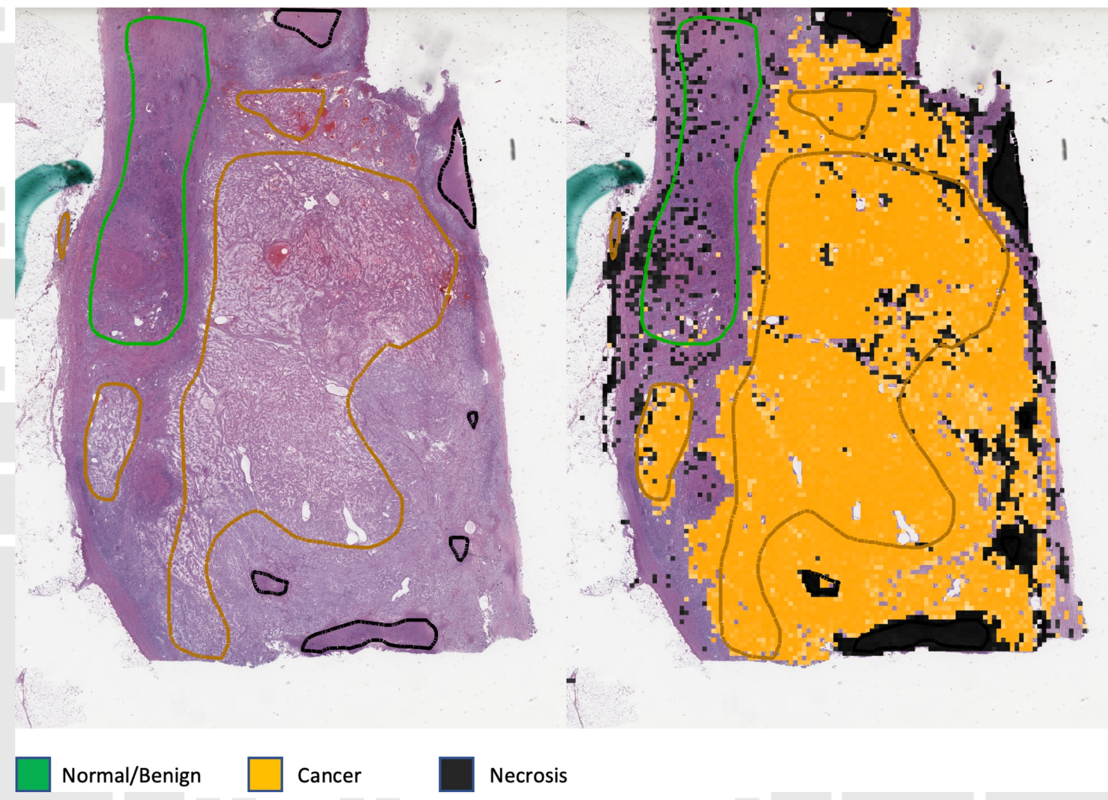
| | Models | Test accuracy (Utah) | Test accuracy (TCGA) |
|---|---|---|---|
| **Supervised** (labeled images only) | ResNet | $88.85 \pm 2.66\%$ | $72.11 \pm 0.41\%$ |
| | ViT | $84.69 \pm 1.33\%$ | $73.50 \pm 0.94\%$ |
| **Self-supervised** (pre-trained on labeled and unlabeled images) | Barlow Twins | $93.42 \pm 0.43\%$ | $77.42 \pm 4.92\%$ |
| | SwAV | $93.87 \pm 0.59\%$ | $82.17 \pm 0.05\%$ |
| | MoCo v3 | $93.91 \pm 0.60\%$ | $78.82 \pm 0.93\%$ |
| | ViT-DINO | $90.53 \pm 1.13\%$ | $79.76 \pm 2.15\%$ |
| **Semi-supervised** (trained on labeled and unlabeled images) | FixMatch | $91.58 \pm 0.65\%$ | $83.34 \pm 2.53\%$ |
| | MixMatch | $92.94 \pm 1.54\%$ | $88.35 \pm 1.39\%$ |
| | CLASS | $94.92 \pm 0.67\%$ | $83.06 \pm 0.47\%$ |
| | CLASS-M | $\mathbf{95.35 \pm 0.46\%}$ | $\mathbf{92.13 \pm 0.89\%}$ |

Ablation studies:

| Contrastive Loss | Augmentations on original RGB | Adaptive stain separation | Pseudo labeling with MixUp | Two views | Test accuracy (Utah) | Test accuracy (TCGA) |
|---|---|---|---|---|---|---|
| ✓ | ✓ | ✓ | ✓ | H/E | **95.35** $\pm$ **0.46** % | **92.13** $\pm$ **0.89** % |
| | ✓ | ✓ | ✓ | H/E | 90.92 $\pm$ 1.18 % | 89.70 $\pm$ 0.77 % |
| ✓ | | ✓ | ✓ | H/E | 94.41 $\pm$ 0.44 % | 91.21 $\pm$ 2.05 % |
| ✓ | ✓ | | ✓ | H/E | 93.97 $\pm$ 0.40 % | 90.97 $\pm$ 1.86 % |
| ✓ | ✓ | ✓ | | H/E | 94.92 $\pm$ 0.67 % | 83.06 $\pm$ 0.47 % |
| ✓ | ✓ | ✓ | | Red/Green | 90.75 $\pm$ 0.13 % | 81.14 $\pm$ 0.34 % |
| ✓ | ✓ | ✓ | | Red/Blue | 89.06 $\pm$ 0.54 % | 80.14 $\pm$ 2.66 % |
| ✓ | ✓ | ✓ | | Green/Blue | 83.43 $\pm$ 3.63 % | 80.25 $\pm$ 1.15 % |
| | ✓ | ✓ | | H/E | 84.57 $\pm$ 2.40 % | 74.46 $\pm$ 2.03 % |
| ✓ | | ✓ | | H/E | 94.99 $\pm$ 0.58 % | 82.78 $\pm$ 1.10 % |
| ✓ | ✓ | | | H/E | 92.43 $\pm$ 0.34 % | 79.89 $\pm$ 0.69 % |
| ✓ | | | | H/E | 91.53 $\pm$ 1.52 % | 75.67 $\pm$ 1.76 % |

Visualization of CLASS-M model prediction on slides:




www.sci.utah.edu