

# Supplementary Material for UCDSC: Open Set UnCertainty aware Deep Simplex Classifier for Medical Image Datasets

WACV 2026 Submission 558

## Used Class Splits:

To evaluate the performance of our proposed method in an open-set recognition (OSR) setting, we utilize class splits for all datasets.

For each dataset, the classes are divided into two distinct sets for each trial: *Known Classes* and *Unknown Classes*.

Detailed splits for the BloodMNIST, OCTMNIST, DermaMNIST, TissueMNIST [2] and ASC datasets are presented in Tables 1, 2, 3, 4 and 5 respectively. The following splits were generated randomly, but kept consistent during ablation and hyperparameter tuning.

Table 1: Five trials on the BloodMNIST dataset ( $K = 5$ ).

Trial	Known Classes	Unknown Classes
1	[6, 1, 3, 7]	[0, 2, 4, 5]
2	[7, 2, 1, 6]	[4, 3, 0, 5]
3	[2, 3, 6, 1]	[0, 7, 4, 5]
4	[3, 1, 2, 7]	[5, 4, 6, 0]
5	[0, 7, 4, 6]	[1, 3, 5, 2]

Table 2: Three trials on the OCTMNIST dataset ( $K = 3$ ).

Trial	Known Classes	Unknown Classes
1	[3, 2]	[0, 1]
2	[3, 0]	[1, 2]
3	[3, 1]	[0, 2]

In OCTMNIST, the healthy class (class 3) is always kept in the known set, which is consistent with real circumstances.

Table 3: Four trials on the DermaMNIST dataset ( $K = 4$ ).

Trial	Known Classes	Unknown Classes
1	[0, 2, 5, 6]	[1, 3, 4]
2	[1, 5, 2, 4]	[3, 6, 0]
3	[5, 1, 6, 4]	[0, 3, 2]
4	[0, 3, 5, 6]	[2, 4, 1]

“eof

Table 4: Five trials on the TissueMNIST dataset ( $K = 5$ )

Trial	Known Classes	Unknown Classes
1	[0, 6, 3, 7]	[1, 2, 4, 5]
2	[0, 6, 7, 4]	[1, 3, 2, 5]
3	[4, 3, 0, 7]	[6, 1, 2, 5]
4	[1, 3, 0, 6]	[2, 4, 7, 5]
5	[6, 7, 3, 4]	[0, 1, 2, 5]

Table 5: Four trials on the Augmented Skin Conditions dataset ( $K = 4$ )

Trial	Known Classes	Unknown Classes
1	[0, 1, 2]	[3, 4, 5]
2	[0, 3, 4]	[1, 2, 5]
3	[1, 2, 5]	[0, 3, 4]
4	[0, 2, 4]	[1, 3, 5]

Table 6: Class-wise distribution of samples for the training sets

Dataset	Class Distribution (Number of Samples per Class)
BloodMNIST	Class 0: 852, Class 1: 2181, Class 2: 1085, Class 3: 2026, Class 4: 849, Class 5: 993, Class 6: 2330, Class 7: 1643
OCMNIST	Class 0: 33484, Class 1: 10213, Class 2: 7754, Class 3: 46026
DermaMNIST	Class 0: 228, Class 1: 359, Class 2: 769, Class 3: 80, Class 4: 779, Class 5: 4693, Class 6: 99
TissueMNIST	Class 0: 53075, Class 1: 7814, Class 2: 5866, Class 3: 15406, Class 4: 11789, Class 5: 7705, Class 6: 39203, Class 7: 24608
Augmented Skin Conditions	Class 0: 320, Class 1: 319, Class 2: 319, Class 3: 319, Class 4: 319, Class 5: 319

## Background Dataset

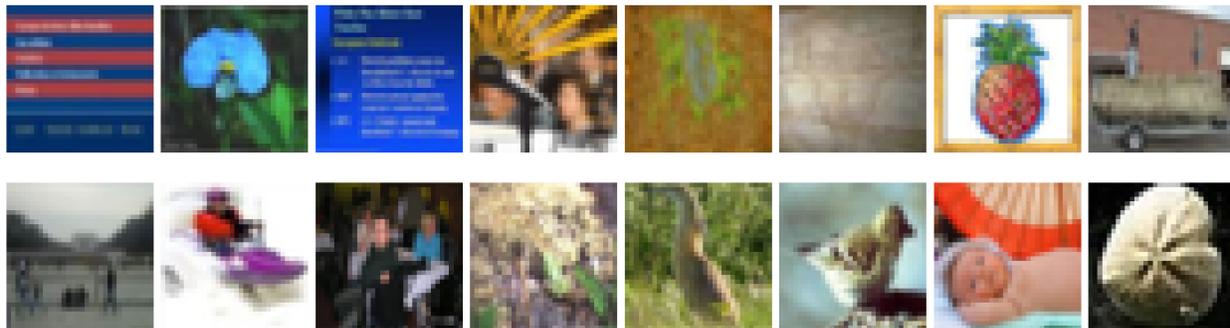


Figure 1: Samples from 300k Random images dataset [1], which is used as the background set for Outlier Exposure

## TSNE Plots

We also observed the t-SNE plots in figure 2 corresponding to the output of the second last layer (last layer of embedding) of our model on BloodMNIST, OCTMNIST, DermaMNIST and TissueMNIST datasets. These plots are consistent with the OSCR scores using UCDSC.

## Hyperparameter Tuning

We show the hyperparameter plots with respect to the AUROC metrics in figures 3 respectively.

## Augmented Skin Conditions Image Dataset

This dataset is available at: <https://www.kaggle.com/datasets/syedalinaqvi/augmented-skin-conditions-image-dataset/data>

For our experiments, the dataset was partitioned into training and testing sets using an 80:20 ratio

## Instructions for running code

### Basic Training

Run the main training script with default parameters:

```
python NirvanaOSR.py --dataset dataset-name --dataroot ./data --outf ./results
```

### Hyperparameter Configuration

The available command-line arguments for configuring hyperparameters are described in Table 7.

## References

- [1] Dan Hendrycks, Mantas Mazeika, and Thomas Dietterich. Deep anomaly detection with outlier exposure. In *Proceedings of the 7th International Conference on Learning Representations (ICLR)*, 2019.

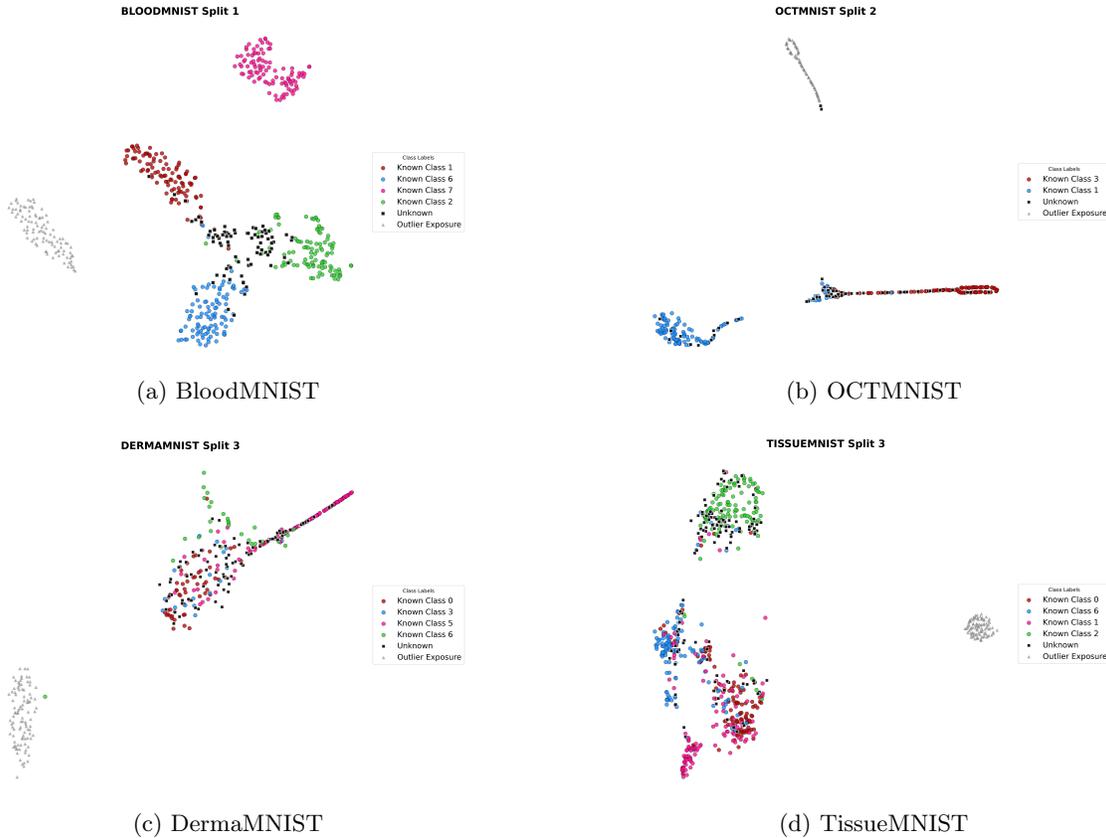


Figure 2: t-SNE visualisations of ResNet model embeddings for all 4 datasets.

Table 7: Descriptions for key hyperparameter command-line arguments.

Parameter	Description
<code>--batch-size</code>	Training batch size
<code>--lr</code>	Learning rate
<code>--max-epoch</code>	Maximum training epochs
<code>--optim</code>	Optimizer to be used
<code>--margin</code>	Margin for loss
<code>--Expand</code>	Expand factor of centers
<code>--uncertainty-weight</code>	Weight for uncertainty loss
<code>--outlier-weight</code>	Weight for outlier triplet loss
<code>--model</code>	Backbone network to be used

Table 8: Hyperparameter values for each dataset.

Dataset	batch-size	margin	expand
BloodMNIST	512	70	100
OCTMNIST	512	70	100
DermaMNIST	512	35	100
TissueMNIST	1024	35	100
ASC	16	70	100

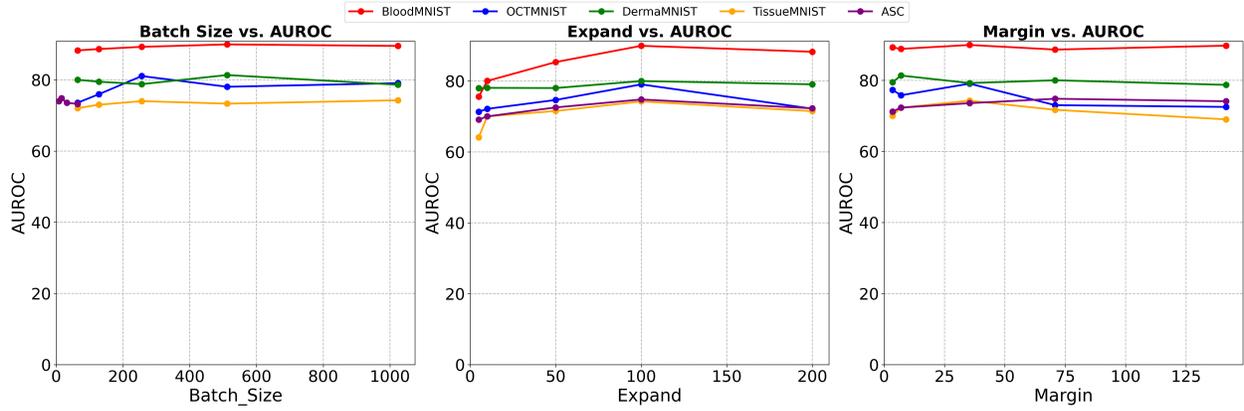


Figure 3: Hyperparameter tuning results for all 5 datasets. These plots illustrate the effect of varying Batch Size, Expand factor and Margin on AUROC metric.

- [2] Jinyu Yang, Yao Shi, Bing Ni, et al. Medmnist v2: A large-scale lightweight benchmark for 2d and 3d biomedical image classification. *Scientific Data*, 8(1):1–14, 2021.