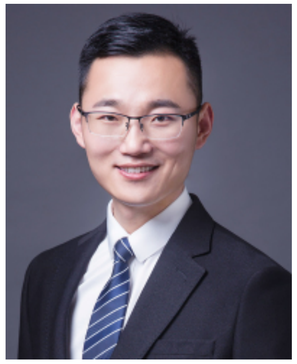
Meta Ordinal Weighting Net For Improving Lung Nodule Classification

Yiming Lei
Shanghai Key Laboratory of Intelligent Information Processing
School of Computer Science, Fudan University

Yiming Lei



Hongming Shan*

Junping Zhang

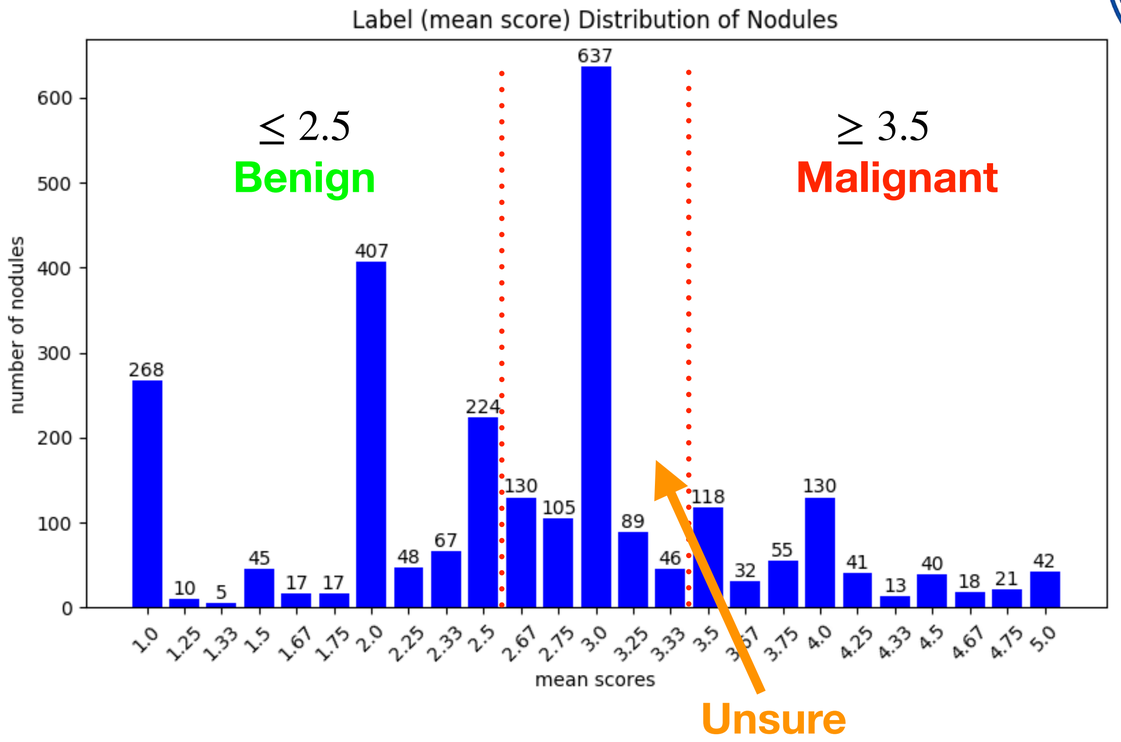
Lung nodule classification

Cross-entropy loss:

$$L_{CE} = -\sum_{c=1}^{C} y_c log\hat{p}_c$$

Ordinal regression:

$$h(x) = r_1 + \eta \sum_{k=1}^{K-1} 1[g_k(x) > 0.5]$$



Malignancy progression:

Benign → Unsure → Malignant

	Benign	Unsure	Malignant
Classification label:	**0**	**1**	**2**
Ordinal label:	**[0 0]**	**[0 1]**	**[1 1]**

K is the number of classes

$$\mathcal{Y} = \{r_1, r_2, \ldots, r_K\}$$

$$\downarrow$$

$$\boldsymbol{o} = (o^1, o^2, \ldots, o^{K-1})^{\mathrm{T}} \in \mathcal{O}$$

Proposed method: Meta ordinal set (MOS)

- Ordinal relationship is resided in the data itself: align each training sample with a MOS (**meta knowledge**)

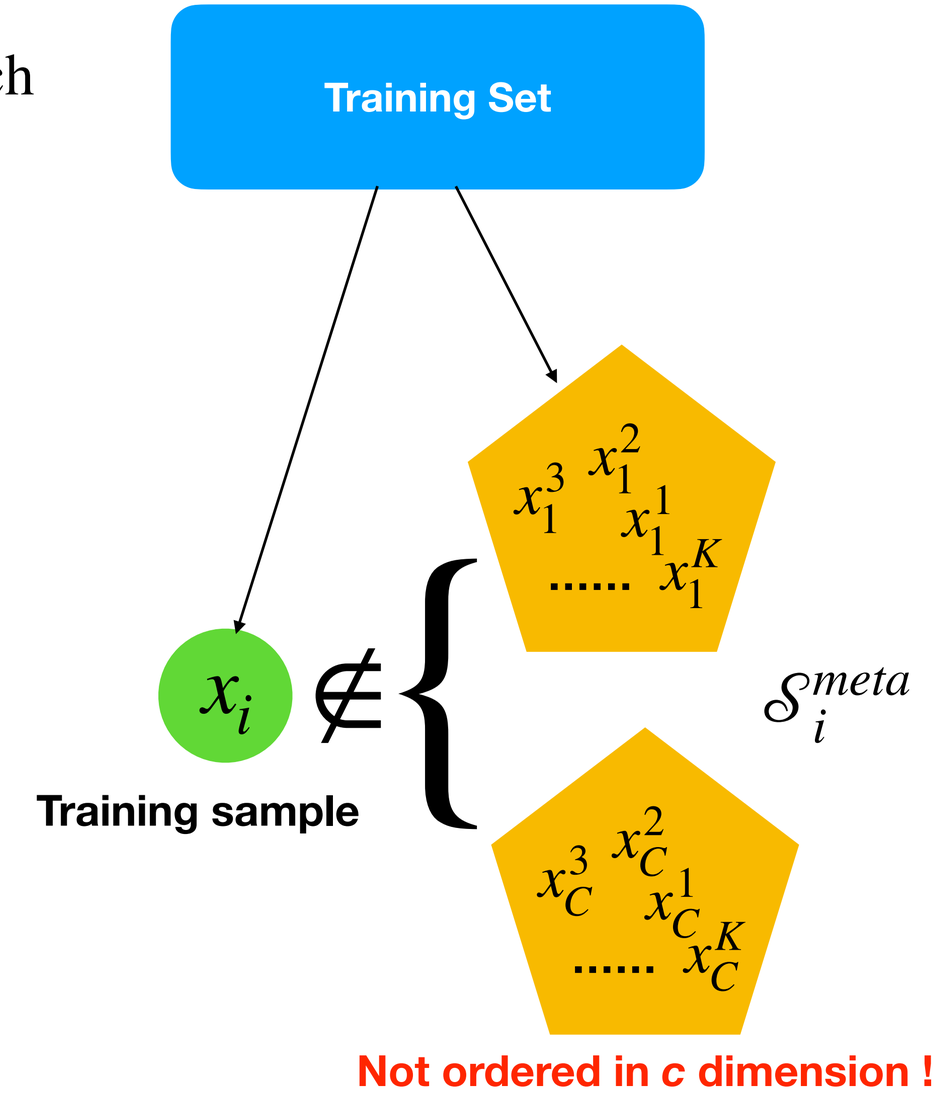
- Definition of MOS for the ***i***-th training sample:

$$\mathcal{S}_i^{meta} = \{x_c^k\}_{c=1,...,C,k=1,...,K}$$

C is the number of classes.

K denotes the number of meta samples for class c.

For a training sample x_i , $x_c^k (c \in 1,2,...,C)$ is randomly sampled from the training set, and $x_i \notin \mathcal{S}_i^{meta}$



Proposed method: Meta cross-entropy (MCE)

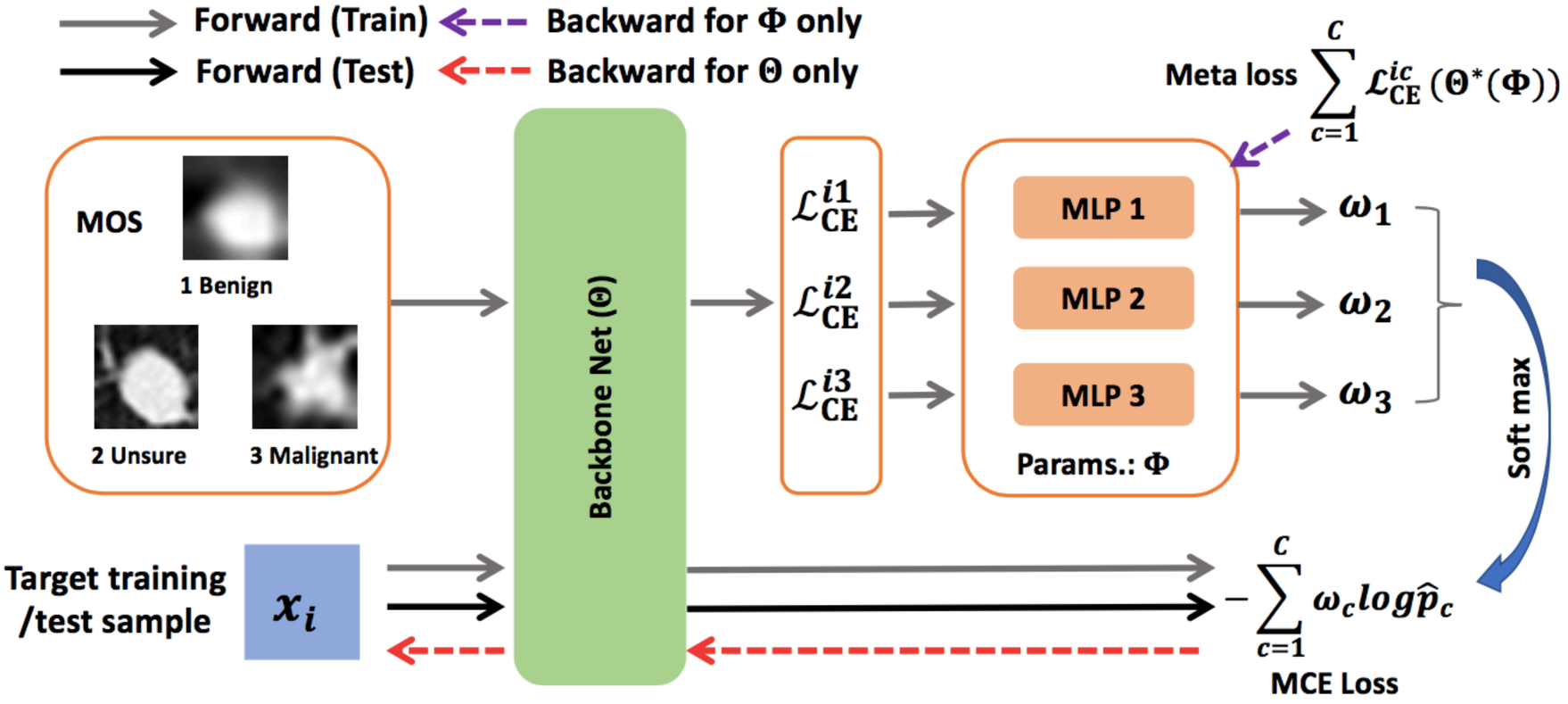
- In order to enable the designed model to absorb the meta knowledge maintained by the MOS, we propose to let the knowledge of each class to affect the loss and the back-propagation.

- The proposed meta cross-entropy loss:

$$L_{MCE} = -\sum_{c=1}^{C} \omega_c log\hat{p}_c \quad \Longleftrightarrow \quad L_{CE} = -\sum_{c=1}^{C} y_c log\hat{p}_c$$

- Compared with the conventional CE loss, the MCE loss enables the training samples to be supervised by the corresponding meta data (class-specific knowledge).

Proposed method: MOW-Net



Optimization objectives:

$$\Theta^*(\Phi) = \arg\min_{\Theta} \mathcal{L}_{\mathrm{MCE}}(\Theta;\Phi) = \frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_{\mathrm{MCE}}^{i} \qquad \Phi^* = \arg\min_{\Phi} \frac{1}{M}\sum_{j=1}^{M}\sum_{c=1}^{C}\mathcal{L}_{\mathrm{CE}}^{j,c}(\Theta^*(\Phi))$$

$$= -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\underbrace{V_c(\mathcal{L}_{\mathrm{CE}}^{i,c}(\Theta);\Phi)}_{\omega_c}\cdot\log\hat{p}_{i,c}(\Theta)$$

Experiments: Classification

- We use the LIDC-IDRI dataset which is partitioned into 3 categories: Benign, Unsure, and Malignant.

- We compared our MOW-Net with the CE loss based method and the state-of-the-art ordinal regression methods.

Table 1. Results of classification on LIDC-IDRI dataset. Following [9], the values with underlines indicate the best results while less important in the clinical diagnosis.

Method	Accuracy	Benign			Malignant			Unsure		
		P	R	F1	P	R	F1	P	R	F1
CE Loss	0.517	0.538	0.668	0.596	0.562	0.495	0.526	0.456	0.360	0.402
Poisson [11]	0.542	0.548	<u>0.794</u>	0.648	0.568	0.624	0.594	0.489	0.220	0.303
NSB [10]	0.553	0.565	0.641	0.601	0.566	0.594	0.580	0.527	0.435	0.476
UDM [9]	0.548	0.541	0.767	0.635	<u>0.712</u>	0.515	0.598	0.474	0.320	0.382
CORF [12]	0.559	0.590	0.627	0.608	0.704	0.495	0.581	0.476	0.515	0.495
MOW-Net ($k = 1$)	0.629	0.752	0.489	0.592	0.558	**0.851**	0.675	0.600	0.675	0.635
MOW-Net ($k = 5$)	0.672	0.764	0.596	0.670	0.600	0.802	0.686	<u>0.642</u>	0.690	0.665
MOW-Net ($k = 10$)	**0.687**	**0.768**	0.623	**0.688**	0.668	0.705	**0.686**	0.606	**0.792**	**0.687**

Experiments: Analysis on Learned Weights

- At epoch 45, the malignant samples are clustered again and the unsure samples are more centralized than that of the previous epochs.

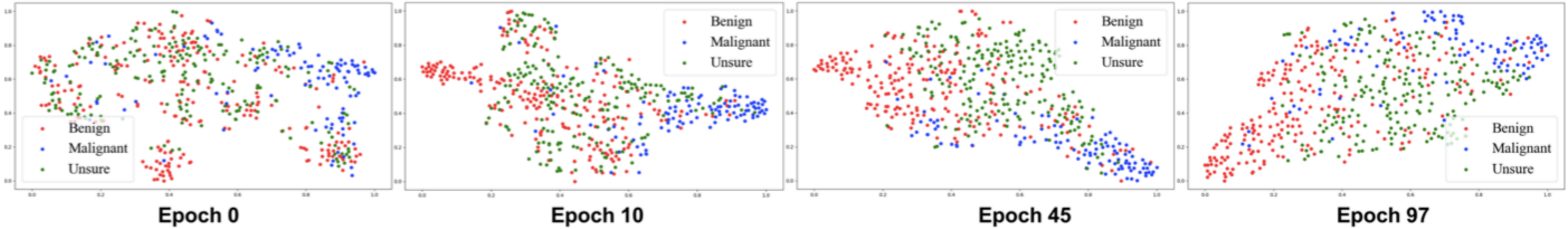


Fig. 2. The visualization results on the testing set using t-SNE.

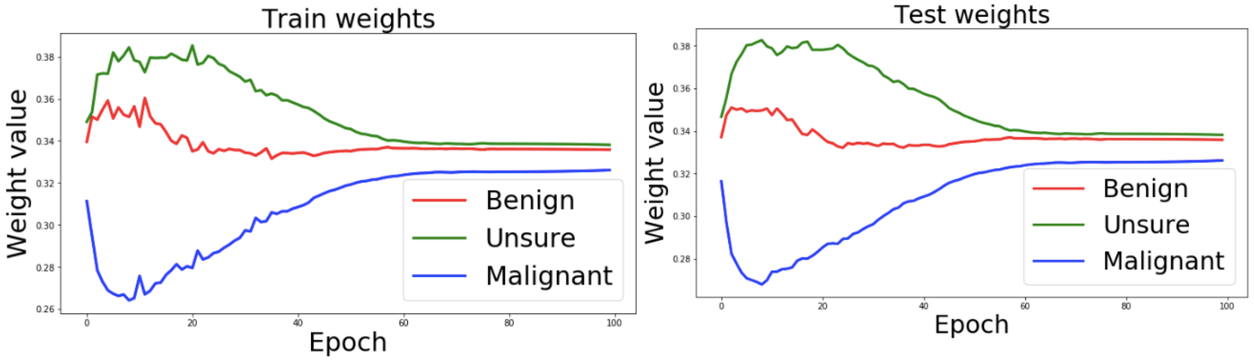


Fig. 3. The variations of the learned weights for all classes.

- The weights of different classes varies across the training, and finally they tends to have similar values and become stable.

Conclusion

- We propose to align each training sample with a meta ordinal set (MOS), which acts as the meta knowledge.
- The proposed meta cross-entropy (MCE) loss enables the MOS to affect the back-propagation during the meta-learning scheme.
- Experimental results demonstrate that the MOW-Net framework surpasses other methods on lung nodule classification, and the weighting scheme evaluated on the weights learned from MOS samples reflects the model bias during training.

Thanks for listening!

Any Questions?

E-mail : ymlei17@fudan.edu.cn