Controlling Class Layout for Deep Ordinal Classification via Constrained Proxies Learning

Cong Wang, Zhiwei Jiang, Yafeng Yin, Zifeng Cheng, Shiping Ge, Qing Gu

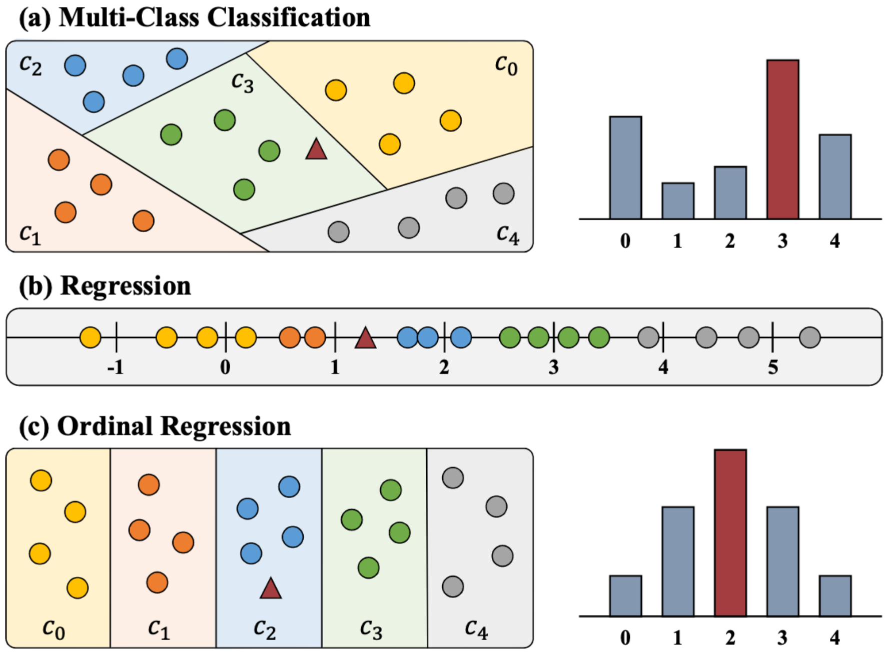
State Key Laboratory for Novel Software Technology, Nanjing University, China

Outline

1. Introduction
2. Our Method
3. Experiments
4. Conclusion

Ordinal Classification Task

- It aims to predict the label of samples on the ordinal scale.
- It is a learning paradigm lying between classification and regression.
- Compared with classification, the classes are naturally ordered.
- Compared with regression, the number of classes is finite, the distance between adjacent classes is undefined.



Existing Methods

- Existing methods seek to learn the feature space specific to ordinal classification, which fall into two fashions: classification & regression.

➢ For the case of classification:

- Both feature space and output label distribution don't show ordinal property.
- Researchers proposed to make implicit ordinal constraints on feature space by recoding the labels.
- The feature space is constrained in a **SOFT** way by constraining the output label distribution.

➢ For the case of regression:

- The samples are mapped into a one-dimensional space, which is ordered in nature.
- The samples are regressed into the continuous real numbers, which need to be discretized into classes by the learned boundaries.
- The feature space is constrained in a **HARD** way by utilizing the ordinal nature of the one-dimensional space.

Motivation

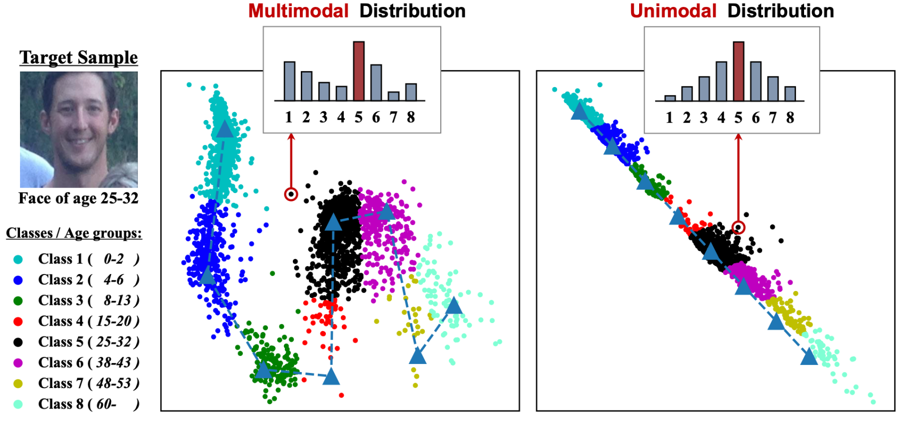
- We consider whether we can explicitly constrain the global layout of samples in the feature space to make it reflect the ordinal nature of classes.

➢ Unconstrained Feature Space

- The layout of samples can hardly guarantee the ordinal nature of classes.
- The samples of some faraway classes may be closely distributed, which results in multimodal probability distributions.

➢ Ordinal-Constrained Feature Space

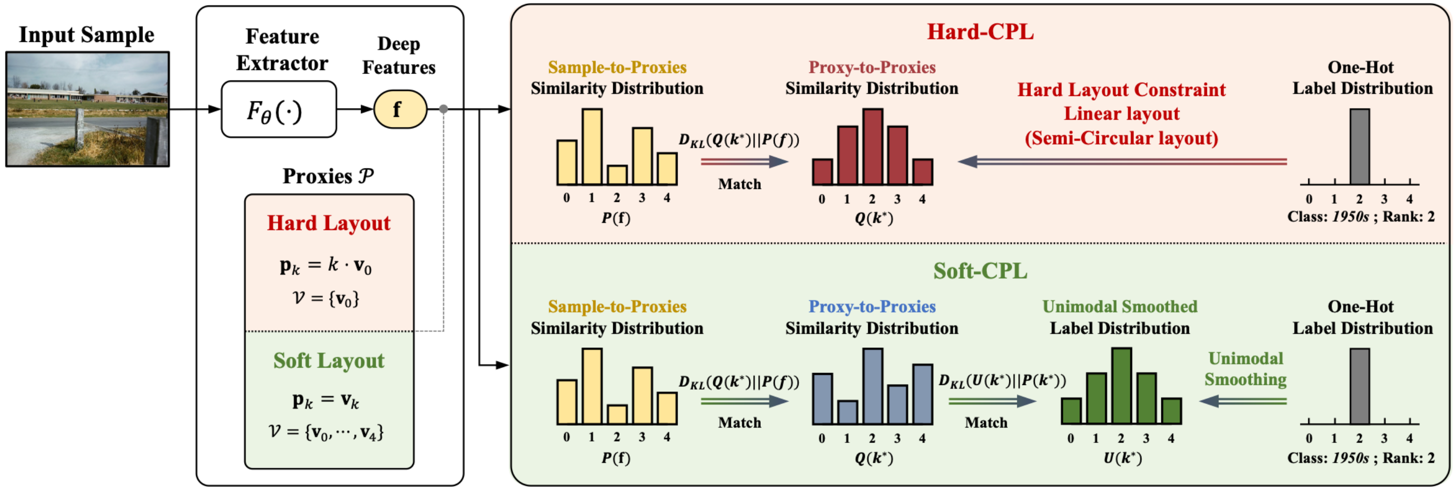
- The sample clusters are arranged in class order in the feature space.
- The samples can always get the unimodal probability distribution (the ideal probability distribution).
- With such ordinal constrained layout, ordinal nature of classes can be guaranteed.



(a) Unconstrained Feature Space
(b) Ordinal Constrained Feature Space *(Our CPL)*

Constrained Proxies Learning (CPL)

- CPL learns a proxy for each class in feature space so as to make samples belonging to the same class can be closely clustered together around the corresponding proxy.
- The basic objective is to encourage the sample feature to be close to the target proxy and to be far away from other proxies according to their relative ordinal distance with the target proxy in the feature space.
- The basic loss function of our CPL is to encourage the sample-to-proxies similarity distribution to match the proxy-to-proxies similarity distribution.



Basic Loss

$$Q_k(k^*) = \frac{\exp(\mathrm{sim}(\mathbf{p}_{k^*}, \mathbf{p}_k))}{\sum_{k'=0}^{K-1} \exp(\mathrm{sim}(\mathbf{p}_{k^*}, \mathbf{p}_{k'}))}$$

$$P_k(\mathbf{f}) = \frac{\exp(\mathrm{sim}(\mathbf{f}, \mathbf{p}_k))}{\sum_{k'=0}^{K-1} \exp(\mathrm{sim}(\mathbf{f}, \mathbf{p}_{k'}))}$$

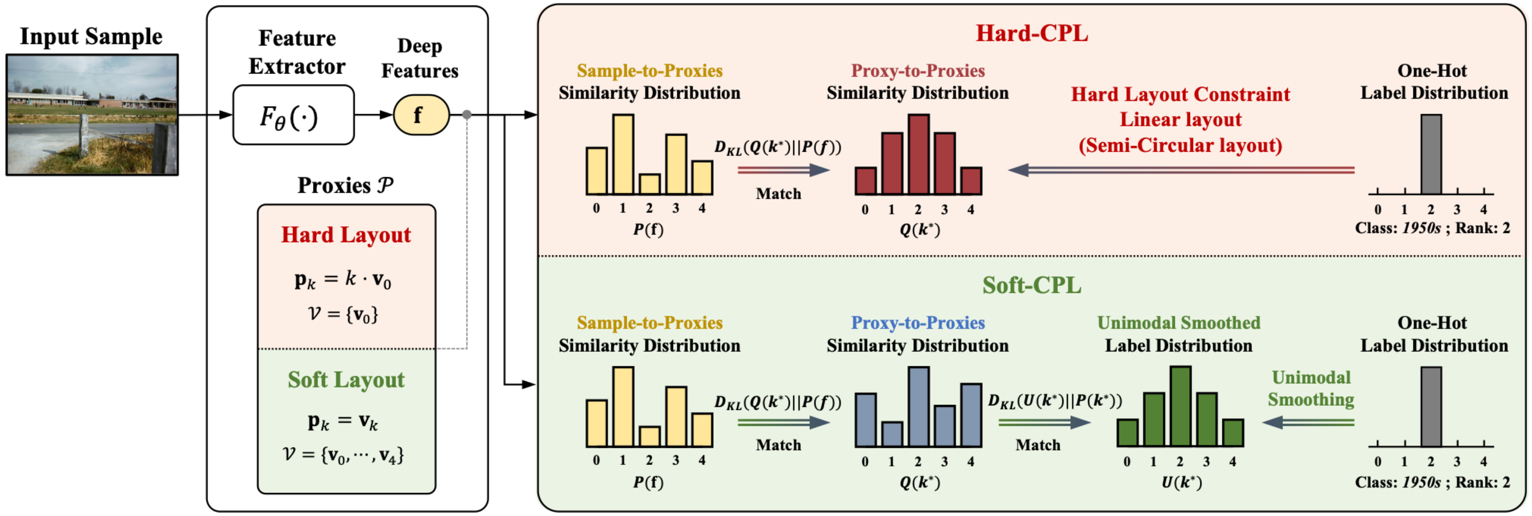
$$\mathcal{L}_{\text{basic}}(\mathbf{f}, k^*) = D_{\text{KL}}[Q(k^*) \| P(\mathbf{f})]$$

Constrained Proxies Learning (CPL)

- CPL aims to constrain the global layout of proxies in feature space to make it more suitable for ordinal classification.
- Two strategies of layout constraint are considered: hard layout constraint (Hard-CPL) and soft layout constraint (Soft-CPL).

➢ **Hard-CPL**: proxies are constrained to be generated in a specific way so that they can be placed in a predefined ordinal layout.

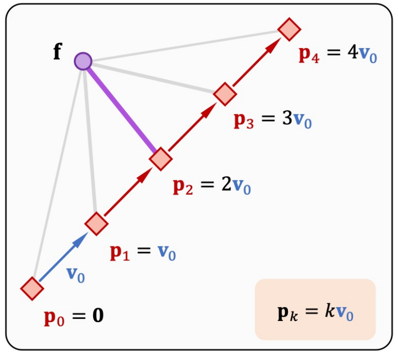
➢ **Soft-CPL**: proxies are constrained to be placed in an ordinal layout corresponding to a specific unimodal distribution.



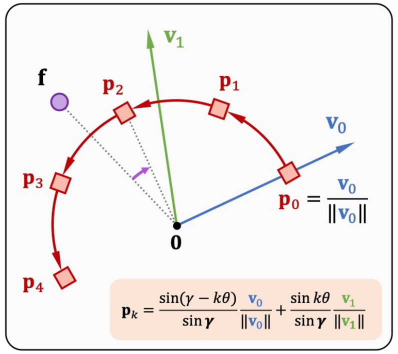
Hard-CPL

- Proxies are constrained to be generated in a specific way so that they can be placed in a predefined ordinal layout.
- Considering that the ordinal layout is different under different metrics, two instantiations are provided:
  - A linear layout specific to the Euclidean distance metric (**H-L**);
  - A semicircular layout specific to the cosine similarity metric (**H-S**).
- Only the basic loss is used for model training of Hard-CPL. $\mathcal{L}_{\mathrm{H}} = \mathcal{L}_{\mathrm{basic}}$

The generated proxies of (a) H-L and (b) H-S.



(a) Linear Layout



(b) Semicircular Layout

Soft-CPL

- For Soft-CPL, we relax the hard layout constraint, allowing proxies not to be placed in strict layout.
- We allow the proxies to be learned freely and only constrain that the proxy layout should always produce unimodal proxy-to-proxies similarity distribution for each proxy.
- To constrain the proxy-to-proxies similarity distribution $Q(k^*)$ to be unimodal, we define a unimodal smoothed label distribution $U_k(k^*)$ by a unimodal smoothing function $E(\cdot;\cdot)$ and the softmax function:

$$U_k(k^*) = \frac{\exp\left(E(k;k^*)\right)}{\sum_{k'=0}^{K-1} \exp\left(E(k';k^*)\right)}$$

- The loss function uses an extra unimodal loss function:

$$\mathcal{L}_{\text{unimodal}}(k^*) = D_{\text{KL}}[U(k^*) \| Q(k^*)]; \quad \mathcal{L}_{\text{S}} = \mathcal{L}_{\text{basic}} + \alpha \mathcal{L}_{\text{unimodal}}$$

- For the unimodal smoothing function $E(\cdot;\cdot)$, we consider two classic unimodal distributions as examples:
  - Soft-CPL based on Poisson distribution: $E(k;k^*) = \frac{1}{\tau_p} \cdot \text{LPMF}\left(k; k^* + \frac{1}{2}\right)$;
  - Soft-CPL based on Binomial distribution: $E(k;k^*) = \frac{1}{\tau_b} \cdot \text{LPMF}\left(k; K-1, \frac{2k+1}{2K}\right)$.

Datasets & Evaluation Metrics

➢ Datasets

- **Historical Color** is a small and balanced ordinal classification dataset which contains images captured on five decades, from *1930s* to *1970s*, each of which has 265 images.
- **Adience Face** contains 26,580 face photos from 2,284 subjects. The dataset is divided into 8 age groups, which are *0-2*, *4-6*, *8-13*, *15-20*, *25-32*, *38-43*, *48-53*, and *elder than 60 years old*, respectively.
- **Image Aesthetics** provides 13,774 Flickr image URLs. The dataset contains four categories of images, namely **nature**, **animals**, **people**, and **urban**. The quality of each image is scored by at least five graders on the five scales, i.e., *unacceptable*, *flawed*, *ordinary*, *professional*, and *exceptional*.

➢ Evaluation Metrics

- **Accuracy**
- **Mean Absolute Error (MAE)**

Performance Comparison

- Compared with all baseline methods, the proposed CPL achieves overall better performance.
- Hard-CPL generally achieves better MAE than Soft-CPL, while Soft-CPL generally achieves better accuracy than Hard-CPL.
- For Soft-CPL, the strategy using Euclidean distance achieves better results than that using cosine similarity.

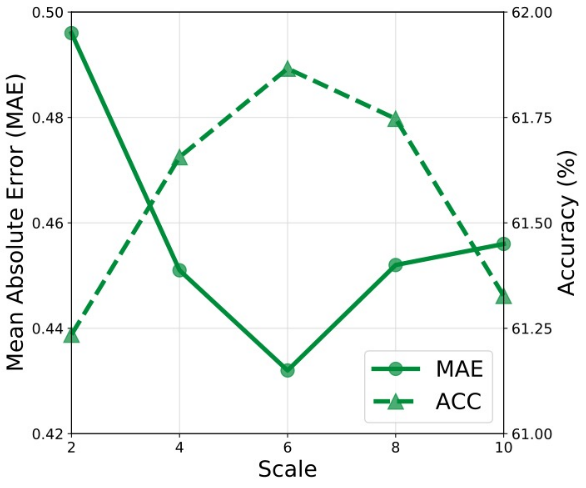
Table 1: The performance (accuracy and MAE) of all comparison methods on **Historical Color** dataset and **Adience Face** dataset. The feature extractors are all VGG-16. The best measures are **in bold**, and the second best measures are underlined.

Methods			Historical Color		Adience Face	
			Accuracy (%) ↑	MAE ↓	Accuracy (%) ↑	MAE ↓
Classification (Liu, Kong, and Goh 2018)			48.94 ± 2.54	0.89 ± 0.06	54.0 ± 6.3	0.61 ± 0.08
Regression (Niu et al. 2016)			42.24 ± 2.91	0.79 ± 0.03	56.3 ± 4.9	0.56 ± 0.07
Ranking (Li et al. 2021)			44.67 ± 4.24	0.81 ± 0.06	56.7 ± 6.0	0.54 ± 0.08
CNNPOR (Liu, Kong, and Goh 2018)			50.12 ± 2.65	0.82 ± 0.05	57.4 ± 5.8	0.55 ± 0.08
GP-DNNOR (Liu, Wang, and Kong 2019)			46.60 ± 2.98	0.76 ± 0.05	57.4 ± 5.5	0.54 ± 0.07
SORD (Diaz and Marathe 2019)			–	–	59.6 ± 3.6	0.49 ± 0.05
POEs (Li et al. 2021)			54.68 ± 3.21	0.66 ± 0.05	60.5 ± 4.4	0.47 ± 0.06
UPL		**Euclidean Distance**	52.20 ± 3.84	0.71 ± 0.07	58.1 ± 3.2	0.48 ± 0.05
		Cosine Similarity	51.32 ± 2.99	0.74 ± 0.05	56.8 ± 4.5	0.51 ± 0.07
CPL	**Hard-Linear**	**Euclidean Distance**	55.71 ± 3.20	**0.63 ± 0.06**	61.6 ± 2.6	**0.43 ± 0.04**
	Hard-Semicircular	**Cosine Similarity**	55.41 ± 3.21	0.64 ± 0.06	61.8 ± 3.1	**0.43 ± 0.04**
	Soft-Poisson	**Euclidean Distance**	57.28 ± 3.41	0.65 ± 0.07	61.3 ± 3.7	0.45 ± 0.05
		Cosine Similarity	56.99 ± 2.44	0.65 ± 0.05	61.1 ± 4.0	0.46 ± 0.05
	Soft-Binomial	**Euclidean Distance**	**57.96 ± 3.14**	0.66 ± 0.08	**62.1 ± 3.6**	0.44 ± 0.04
		Cosine Similarity	57.66 ± 3.11	0.65 ± 0.06	61.9 ± 4.5	0.44 ± 0.05

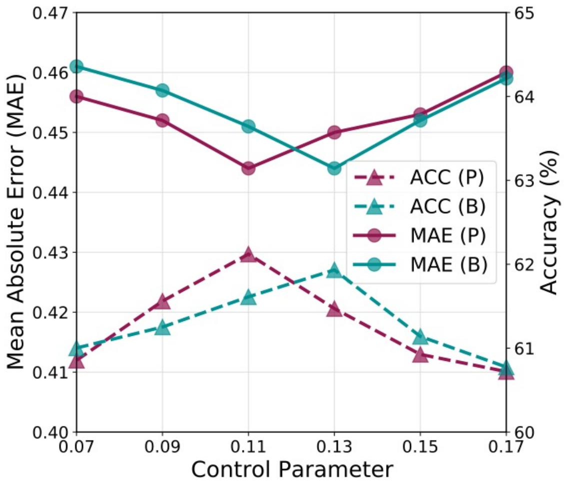
Table 2: The performance (accuracy and MAE) of all comparison methods on **Image Aesthetics** dataset. The feature extractors are all VGG-16. The best measures are **in bold**, and the second best measures are underlined.

Methods			Accuracy (%) ↑					MAE ↓				
			Nature	Animals	Urban	People	Overall	Nature	Animals	Urban	People	Overall
Classification (Liu, Kong, and Goh 2018)			70.97	68.02	68.19	71.63	69.45	0.305	0.342	0.374	0.412	0.376
Regression (Li et al. 2021)			71.52	70.72	71.22	69.72	70.80	0.378	0.397	0.387	0.400	0.390
Ranking (Niu et al. 2016)			69.81	69.10	66.49	66.49	68.96	0.313	0.331	0.349	0.312	0.326
CNNPOR (Liu, Kong, and Goh 2018)			71.86	69.32	69.09	69.94	70.05	0.294	0.322	0.325	0.321	0.316
SORD (Diaz and Marathe 2019)			73.59	70.29	73.25	70.59	72.03	0.271	0.308	**0.276**	0.309	0.290
POEs (Li et al. 2021)			73.62	71.14	72.78	72.22	72.44	0.273	0.299	0.281	0.293	0.287
UPL		**Euclidean Distance**	71.82	68.21	69.24	68.98	69.56	0.283	0.343	0.313	0.341	0.320
		Cosine Similarity	72.88	68.68	69.88	69.81	70.31	0.284	0.325	0.311	0.352	0.318
CPL	**Hard-Linear**	**Euclidean Distance**	74.43	72.11	72.99	72.53	73.02	**0.260**	**0.289**	0.283	0.287	**0.280**
	Hard-Semicircular	**Cosine Similarity**	74.35	71.50	72.91	72.33	72.77	0.262	0.297	0.288	0.290	0.284
	Soft-Poisson	**Euclidean Distance**	74.46	71.73	72.94	72.45	72.90	0.267	0.302	0.281	0.297	0.287
		Cosine Similarity	74.53	71.39	72.97	72.38	72.82	0.270	0.299	0.287	**0.286**	0.286
	Soft-Binomial	**Euclidean Distance**	**74.97**	**72.61**	**73.28**	72.61	**73.37**	0.262	0.297	0.285	0.299	0.286
		Cosine Similarity	74.62	72.28	73.20	**72.74**	73.21	0.265	0.301	0.286	0.294	0.287

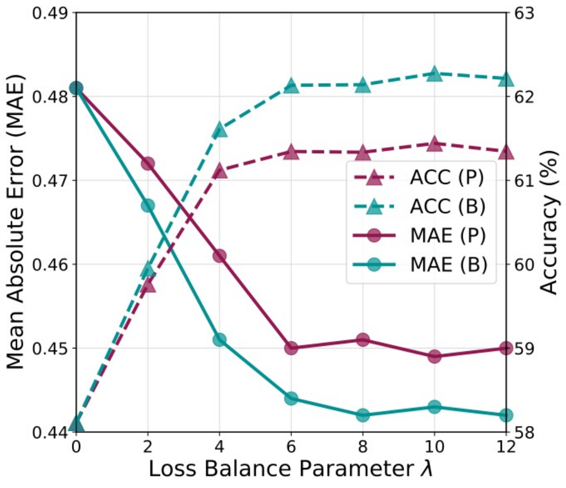
Model Analysis



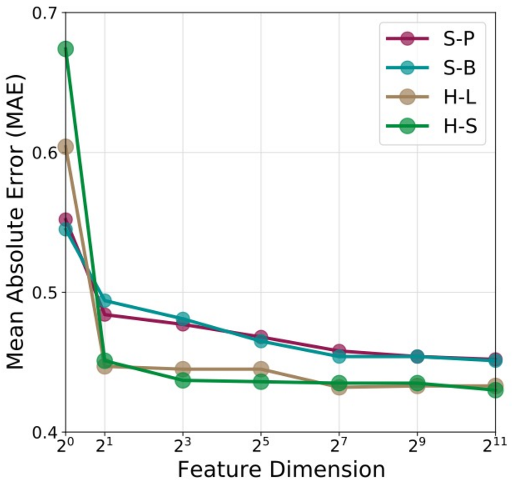
(a) Effect of Scale Parameter



(b) Effect of Control Parameter

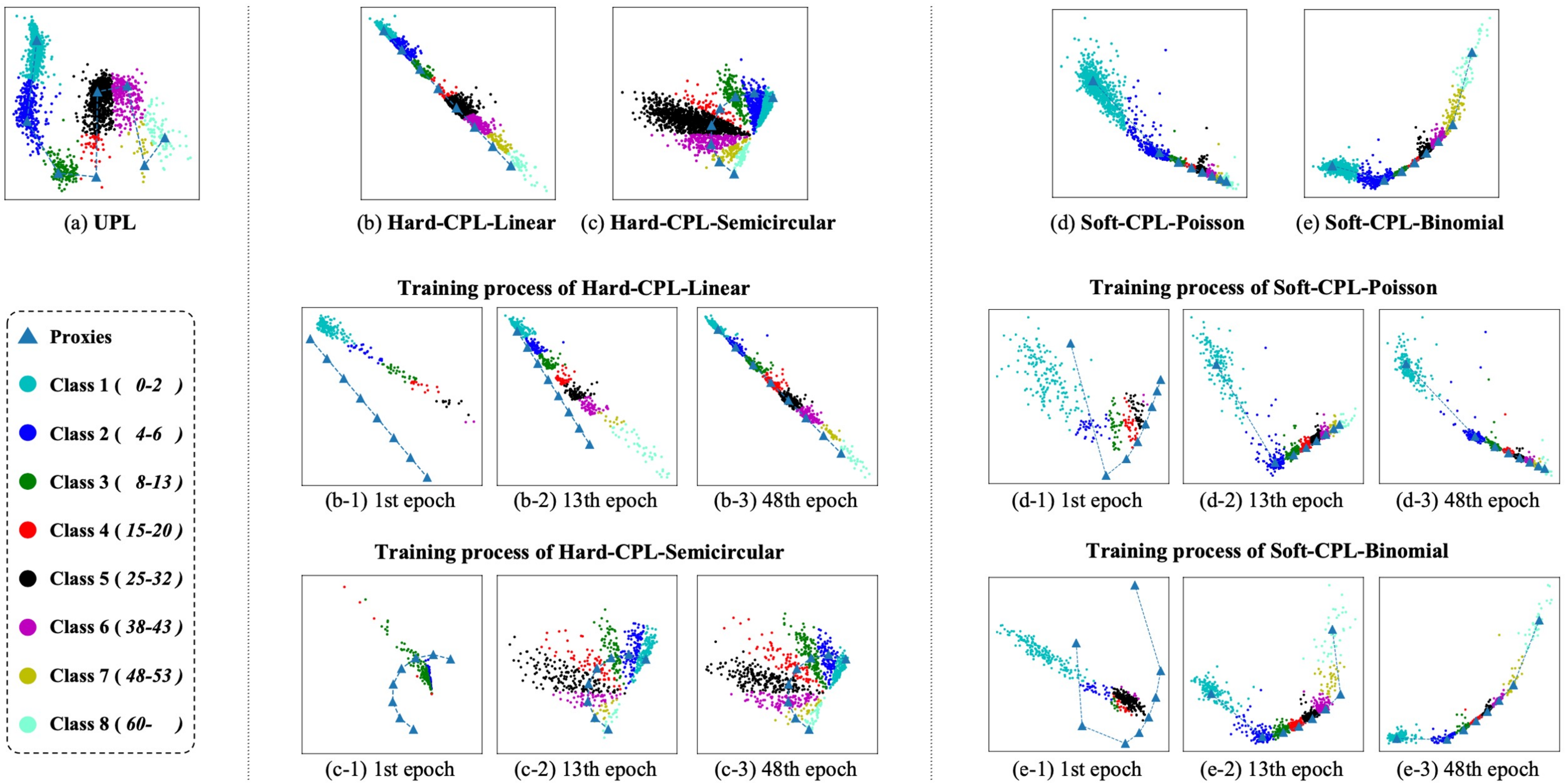


(c) Effect of Tradeoff Parameter



(d) Effect of Feature Dimension

Visualization



Conclusion

➢ Major Contributions

- We propose a constrained proxies learning method to explicitly control the global layout of classes in high-dimensional feature space, making it more suitable for ordinal classification.
- We propose both the hard and soft layout constraints of proxies, and explore some example layouts for both of them (i.e., two strict ordinal layouts for hard constraint and two relaxed ordinal layouts for soft constraint).
- We conduct experiments on three public datasets and show that the proposed CPL achieves better performance than previous ordinal classification methods.

➢ Future Work

- We will try to verify our method on more ordinal classification datasets.
- We will try to improve the ordinal classification performance from the perspective of graph learning.


THE 37TH AAAI CONFERENCE ON
ARTIFICIAL INTELLIGENCE
FEBRUARY 7-14, 2023 · WASHINGTON, DC, USA
WALTER E. WASHINGTON CONVENTION CENTER
AAAI-23

Thanks!