Controlling Class Layout for Deep Ordinal Classification via Constrained Proxies Learning

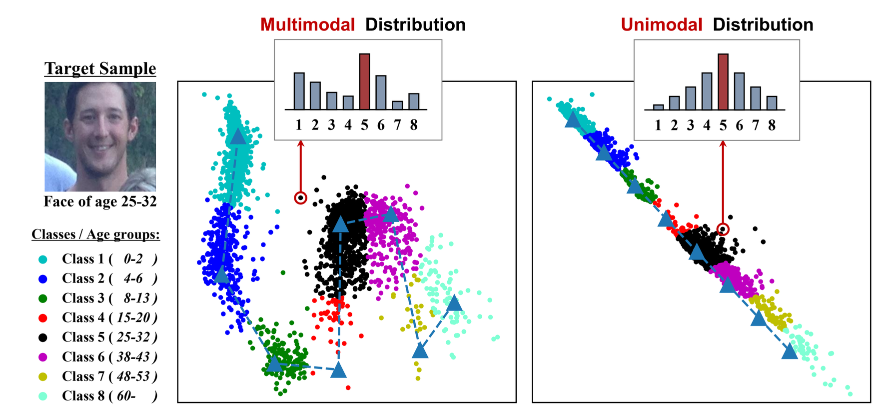
Cong Wang, Zhiwei Jiang, Yafeng Yin, Zifeng Cheng, Shiping Ge, Qing Gu
State Key Laboratory for Novel Software Technology, Nanjing University, China

1. Introduction

- Ordinal classification aims to predict the label of samples on the ordinal scale.
- Existing methods seek to learn the specific feature space, which fall into two fashions: classification & regression.
- In this work, we propose constrained proxies learning (CPL) to explicitly control the global layout of classes, making it more suitable for ordinal classification.

2. Motivation

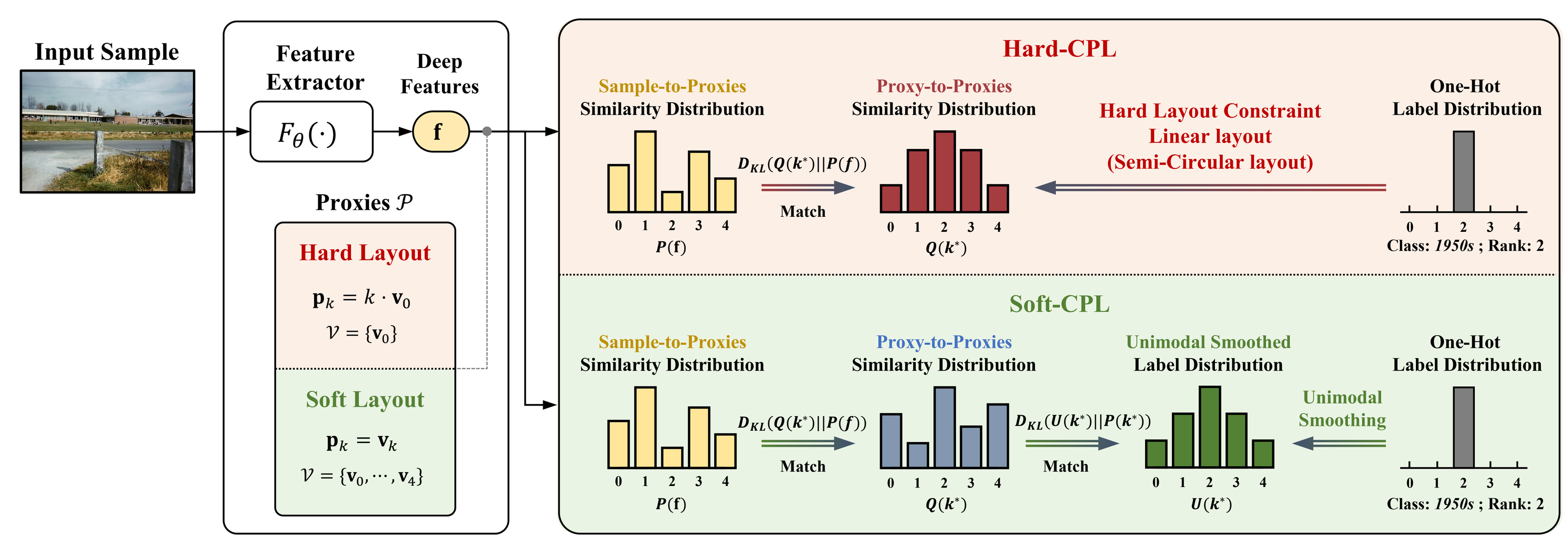
- The global layout of samples in the feature space is explicitly constrained to make it reflect the ordinal nature of classes.



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

3. 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.
- CPL aims to constrain the global layout of proxies in feature space to make it more suitable for ordinal classification.
- 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.

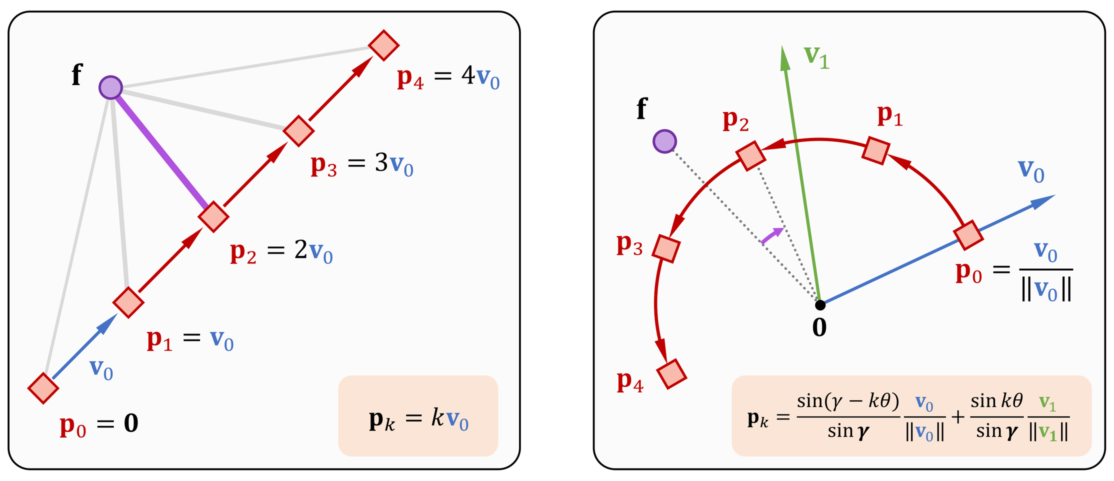


Basic Loss

$$Q_k(k^*) = \frac{\exp(\mathrm{sim}(\mathbf{p}_{k^*}\cdot\mathbf{p}_k))}{\sum_{k'=0}^{K-1}\exp(\mathrm{sim}(\mathbf{p}_{k^*}\cdot\mathbf{p}_{k'}))};\quad P_k(\mathbf{f}) = \frac{\exp(\mathrm{sim}(\mathbf{f}\cdot\mathbf{p}_k))}{\sum_{k'=0}^{K-1}\exp(\mathrm{sim}(\mathbf{f}\cdot\mathbf{p}_{k'}))};\quad \mathcal{L}_{\mathrm{basic}}(\mathbf{f},k^*) = D_{\mathrm{KL}}[Q(k^*)\|P(\mathbf{f})]$$

4. Hard-CPL

- Proxies are constrained to be generated in a specific way so that they can be placed in a predefined ordinal layout.
- Two instantiations: the linear layout specific to the Euclidean distance metric (**H-L**); and the semicircular layout specific to the cosine similarity metric (**H-S**).



(a) Linear Layout (b) Semicircular Layout

$$\mathcal{L}_{\mathrm{H}} = \mathcal{L}_{\mathrm{basic}}$$

5. Soft-CPL

- Proxies can be learned freely. The proxy layout is constrained to produce unimodal proxy-to-proxies similarity distribution for each proxy.
- To constrain the proxy-to-proxies similarity distribution to be unimodal, we define a unimodal smoothed label distribution $U_k(k^*)$ by a unimodal smoothing function $E(\cdot;\cdot)$.
- For the unimodal smoothing function, two classic unimodal distributions are considered as examples: the Poisson distribution and the Binomial distribution.

$$U_k(k^*) = \frac{\exp(E(k;k^*))}{\sum_{k'=0}^{K-1}\exp(E(k';k^*))};\quad \mathcal{L}_{\mathrm{unimodal}}(k^*) = D_{\mathrm{KL}}[U(k^*)\|Q(k^*)];\quad \mathcal{L}_{\mathrm{S}} = \mathcal{L}_{\mathrm{basic}} + \alpha\mathcal{L}_{\mathrm{unimodal}}$$

6. Experiments

- **Performance Comparison**

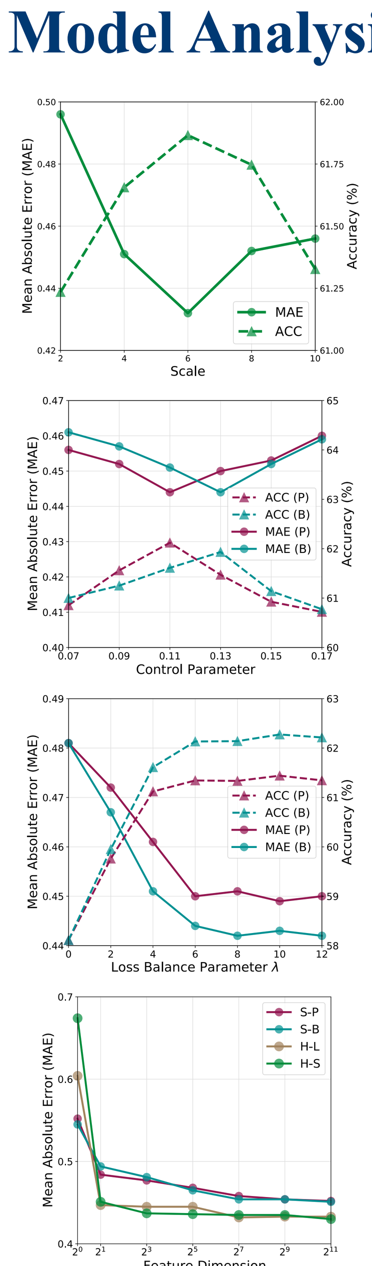
Table 1: The performance (accuracy and MAE) of all comparison methods on **Historical Color** dataset and **Adience Face** dataset. The feature extractors are all VGG-16. The best measures are **in bold**, and the second best measures are underlined.

Methods			Historical Color Accuracy (%) ↑	Historical Color MAE ↓	Adience Face Accuracy (%) ↑	Adience Face 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

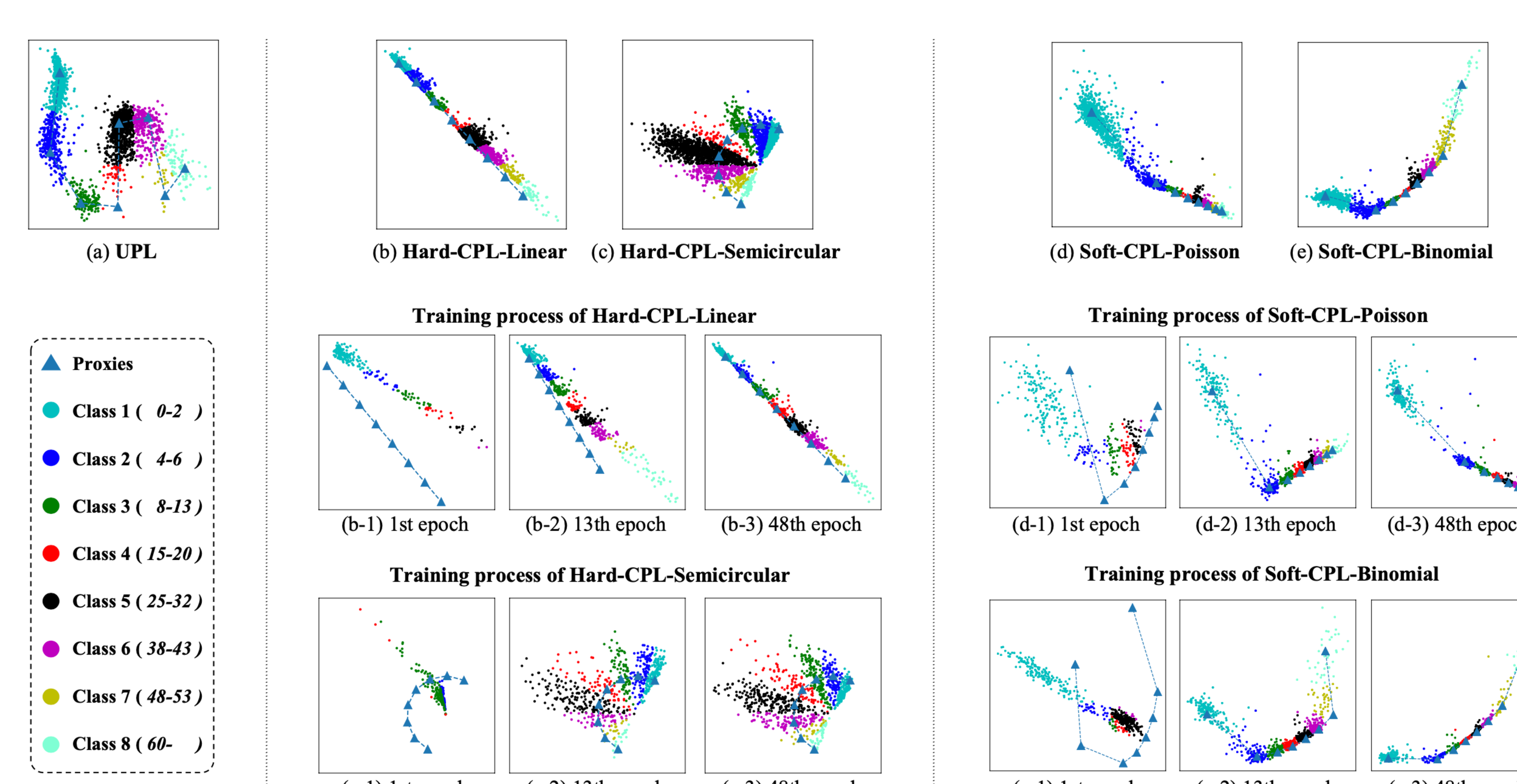
Table 2: The performance (accuracy and MAE) of all comparison methods on **Image Aesthetics** dataset. The feature extractors are all VGG-16. The best measures are **in bold**, and the second best measures are underlined.

Methods			Accuracy (%) ↑ Nature	Animals	Urban	People	Overall	MAE ↓ 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**



- **Visualization**



- For the visualization of Hard-CPL and Soft-CPL, proxies and feature clusters are both arranged in expected ordinal layouts.