Graph Your Own Prompt

Xi Ding, Lei Wang, Piotr Koniusz, Yongsheng Gao

NEURAL INFORMATION PROCESSING SYSTEMS

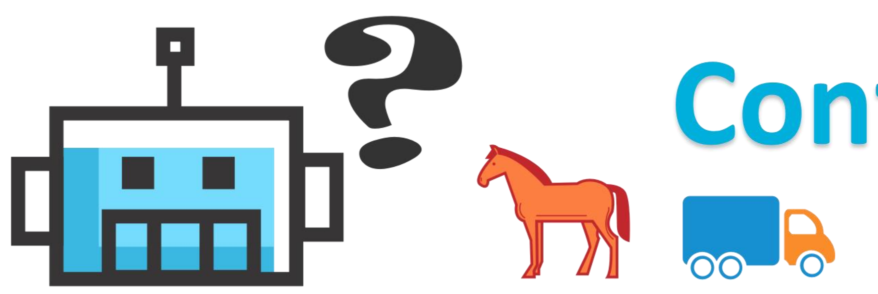
Griffith UNIVERSITY Queensland, Australia · CSIRO · DATA 61 · Australian National University · UNSW SYDNEY

Code

Paper

Motivation

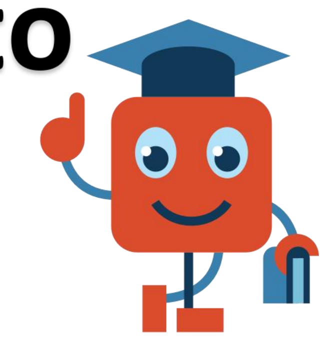
- ✓ Deep networks learn rich features, but these features often do not match semantic class structure.
- ✓ Samples predicted as the same class may still appear far apart in feature space, hurting generalization.



Confused in abstract space

Four legs? Hmm... A car? Or a horse?

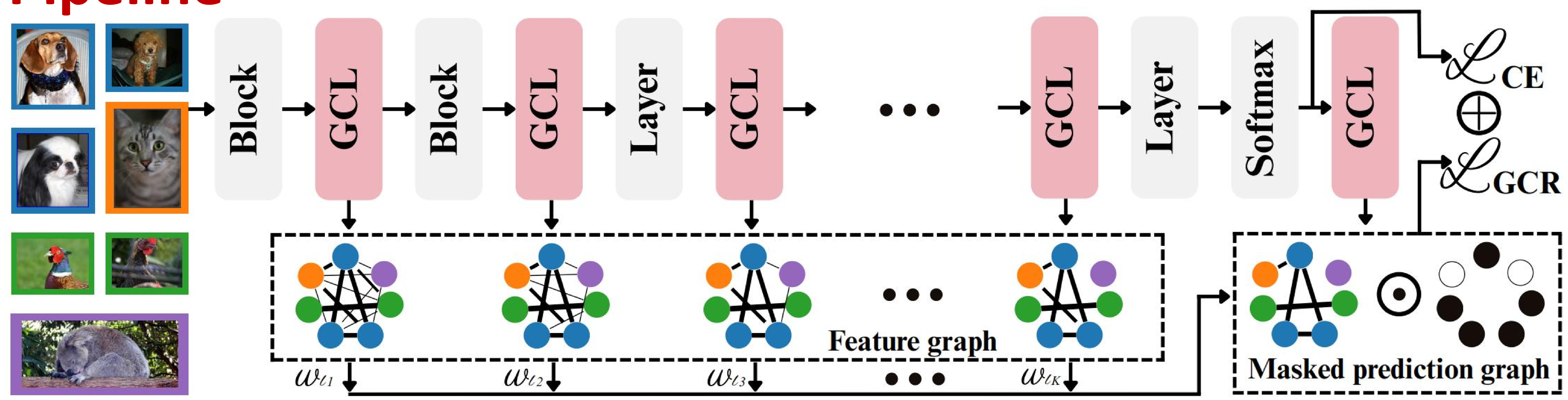
Why not use your own predictions to refine and clean feature structure?



Strength

Lightweight **Model-agnostic** **Parameter-free** **Easy to plug in**

Pipeline



Self-prompting: The model learns from its own outputs, reinforcing semantic structure

Method

We use cosine similarity with non-negative values:

$$F_{ij}^{(l)} = \mathrm{ReLU}(\cos(x_i^{(l)}, x_j^{(l)})), \quad i, j = 1, \ldots, n. \tag{1}$$

From the prediction logits $Z = [z_1^\top, \ldots, z_n^\top]^\top$ of the same batch:

- apply softmax to obtain class probability vectors $p_i = \mathrm{softmax}(z_i)$,
- compute pairwise cosine similarity between prediction vectors:

$$S_{ij} = \mathrm{ReLU}(\cos(p_i, p_j)). \tag{2}$$

To focus on reliable semantic relations, we build a binary mask $M \in \{0,1\}^{n \times n}$:

$$M_{ij} = \begin{cases} 1, & \text{if } y_i = y_j, \\ 0, & \text{otherwise}. \end{cases} \tag{3}$$

The masked prediction graph $P \in \mathbb{R}^{n \times n}$ is then

$$P_{ij} = M_{ij} \odot S_{ij}, \tag{4}$$

where $\odot$ denotes elementwise multiplication.

The layer-wise **graph consistency loss** is

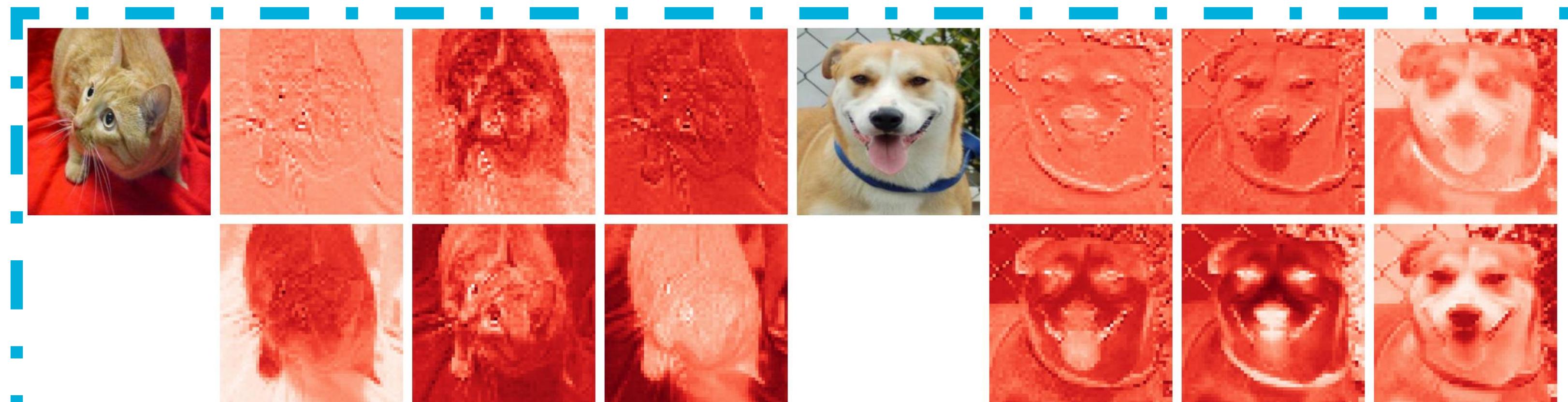
$$\mathcal{L}_{\mathrm{GCR}}^{(l)} = \|\mathrm{triu}(F^{(l)}) - \mathrm{triu}(P)\|_F^2. \tag{5}$$

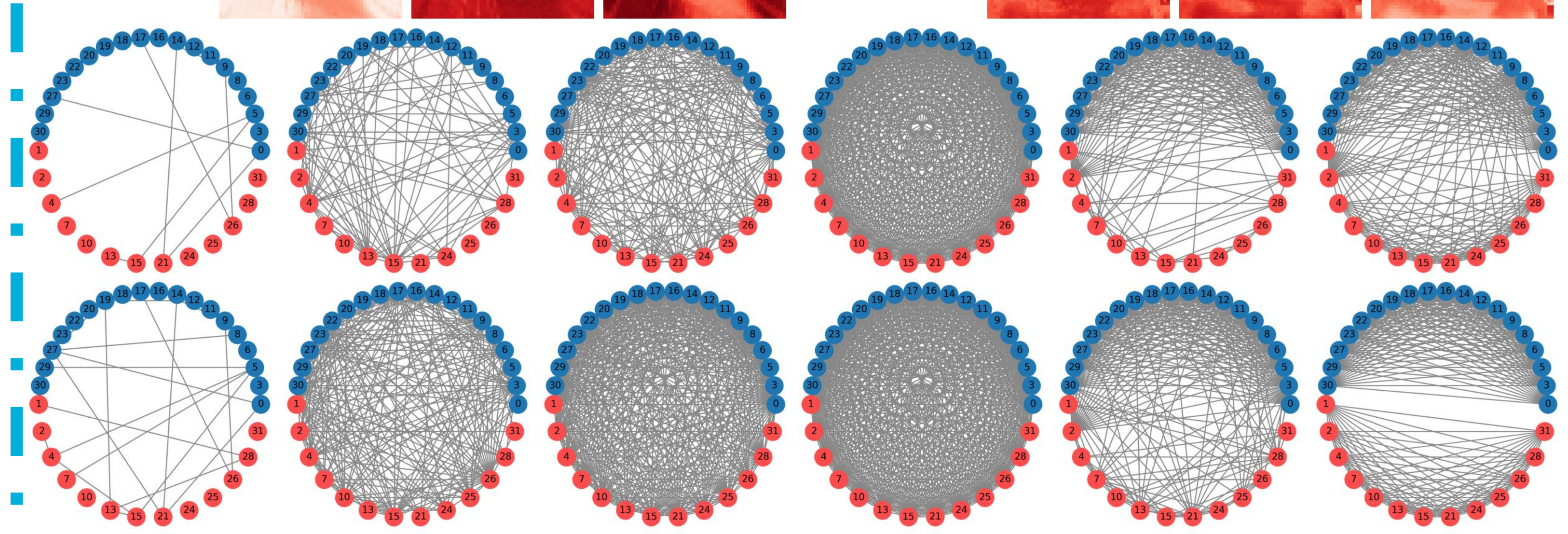
For a set of layers $\{1, \ldots, K\}$, compute a graph consistency loss at each layer and combine them:

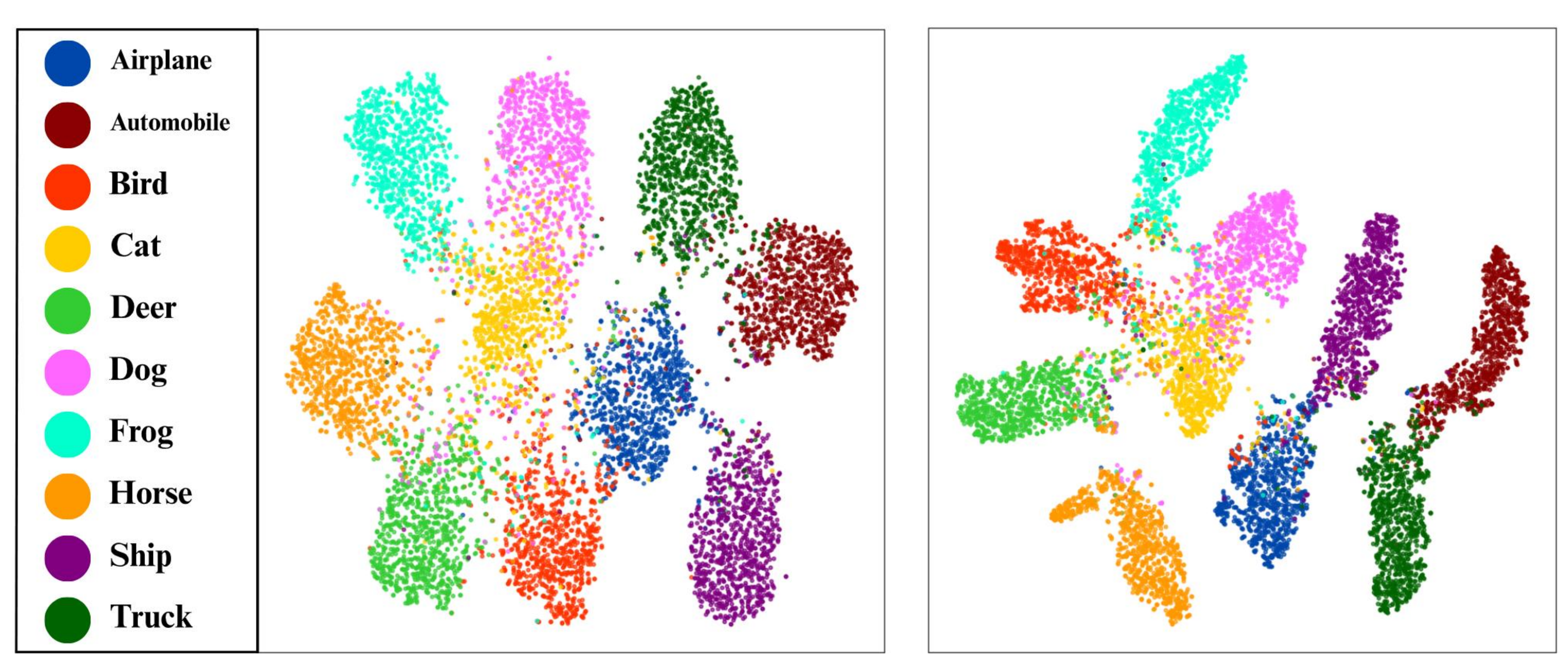
$$\mathcal{L}_{\mathrm{GCR}} = \sum_{l=1}^{K} w_l \|\mathrm{triu}(F^{(l)}) - \mathrm{triu}(P)\|_F^2, \tag{6}$$

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{CE}} + \lambda\, \mathcal{L}_{\mathrm{GCR}}$$

Results

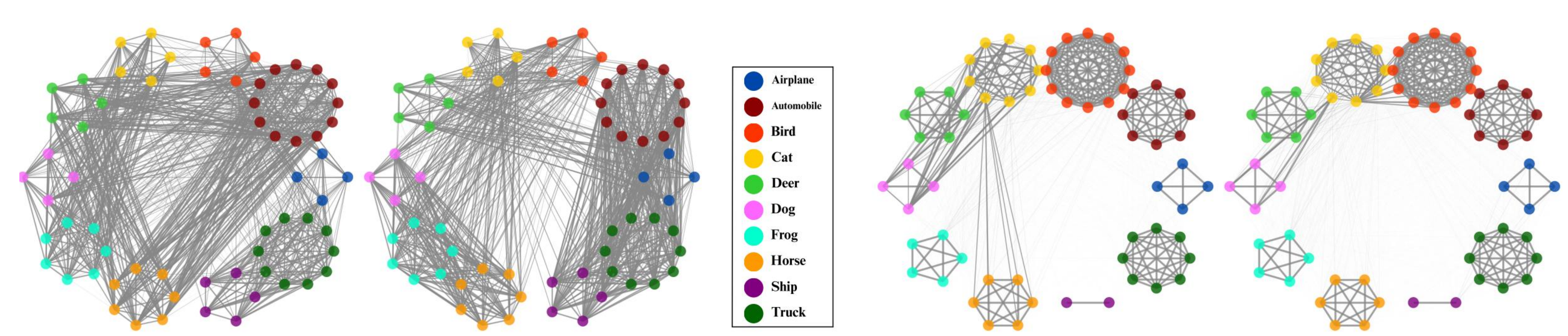






	MAE	MNet	SN	SQNet	GLNet	Rx-50	Rx-101	R34	R50	R101	D121	Mean
Baseline	88.95$_{\pm0.33}$	90.23$_{\pm0.25}$	91.21$_{\pm0.28}$	92.30$_{\pm0.25}$	94.10$_{\pm0.26}$	94.57$_{\pm0.29}$	95.12$_{\pm0.30}$	94.83$_{\pm0.25}$	95.03$_{\pm0.28}$	95.22$_{\pm0.31}$	95.01$_{\pm0.27}$	93.32$_{\pm2.26}$
Early GCL	89.42$_{\pm0.25}$	91.17$_{\pm0.22}$	92.33$_{\pm0.33}$	92.59$_{\pm0.21}$	**94.89**$_{\pm0.23}$	95.48$_{\pm0.22}$	95.63$_{\pm0.25}$	95.55$_{\pm0.18}$	95.57$_{\pm0.23}$	95.39$_{\pm0.26}$	**95.81**$_{\pm0.17}$	93.98$_{\pm2.22}$
Mid GCL	**89.77**$_{\pm0.22}$	91.15$_{\pm0.18}$	92.58$_{\pm0.19}$	92.40$_{\pm0.20}$	94.82$_{\pm0.21}$	95.47$_{\pm0.19}$	95.39$_{\pm0.24}$	**95.69**$_{\pm0.23}$	95.61$_{\pm0.20}$	**95.75**$_{\pm0.17}$	95.51$_{\pm0.22}$	94.01$_{\pm2.15}$
Late GCL	89.70$_{\pm0.29}$	**91.40**$_{\pm0.19}$	92.36$_{\pm0.21}$	92.80$_{\pm0.19}$	94.88$_{\pm0.19}$	95.35$_{\pm0.28}$	**95.71**$_{\pm0.26}$	**95.69**$_{\pm0.19}$	**95.66**$_{\pm0.17}$	95.51$_{\pm0.24}$	95.72$_{\pm0.22}$	**94.07**$_{\pm2.14}$
Early+Mid	89.52$_{\pm0.19}$	90.77$_{\pm0.26}$	92.56$_{\pm0.21}$	92.27$_{\pm0.25}$	94.79$_{\pm0.18}$	95.33$_{\pm0.27}$	95.55$_{\pm0.23}$	95.46$_{\pm0.20}$	95.51$_{\pm0.21}$	95.37$_{\pm0.19}$	95.64$_{\pm0.20}$	93.89$_{\pm2.22}$
Mid+Late	89.59$_{\pm0.28}$	91.23$_{\pm0.20}$	**92.79**$_{\pm0.20}$	**92.86**$_{\pm0.23}$	94.61$_{\pm0.22}$	**95.51**$_{\pm0.19}$	95.38$_{\pm0.27}$	95.45$_{\pm0.18}$	95.33$_{\pm0.26}$	95.52$_{\pm0.14}$	95.70$_{\pm0.19}$	94.00$_{\pm2.09}$
Early+Late	89.64$_{\pm0.25}$	91.03$_{\pm0.24}$	92.30$_{\pm0.28}$	92.70$_{\pm0.23}$	94.69$_{\pm0.20}$	95.40$_{\pm0.20}$	95.35$_{\pm0.23}$	95.66$_{\pm0.21}$	95.31$_{\pm0.25}$	95.49$_{\pm0.16}$	95.53$_{\pm0.22}$	93.92$_{\pm2.14}$
Full GCL	89.55$_{\pm0.23}$	90.99$_{\pm0.18}$	92.48$_{\pm0.19}$	92.65$_{\pm0.20}$	94.57$_{\pm0.21}$	95.50$_{\pm0.19}$	95.34$_{\pm0.20}$	95.48$_{\pm0.17}$	95.62$_{\pm0.18}$	95.38$_{\pm0.21}$	95.51$_{\pm0.20}$	93.92$_{\pm2.15}$

	MAE	MNet	SN	SQNet	Rx-50	Rx-101	R34	R50	D121	Mean
Baseline	64.29$_{\pm0.34}$	65.95$_{\pm0.25}$	70.11$_{\pm0.30}$	69.43$_{\pm0.27}$	77.75$_{\pm0.29}$	77.83$_{\pm0.30}$	76.82$_{\pm0.28}$	77.31$_{\pm0.29}$	77.09$_{\pm0.27}$	72.95$_{\pm5.50}$
Early GCL	65.05$_{\pm0.29}$	67.45$_{\pm0.21}$	**71.96**$_{\pm0.27}$	70.90$_{\pm0.20}$	79.18$_{\pm0.22}$	79.69$_{\pm0.27}$	77.90$_{\pm0.22}$	79.37$_{\pm0.25}$	79.41$_{\pm0.22}$	74.55$_{\pm5.78}$
Mid GCL	64.99$_{\pm0.30}$	67.88$_{\pm0.21}$	71.89$_{\pm0.24}$	70.21$_{\pm0.25}$	79.07$_{\pm0.19}$	79.28$_{\pm0.26}$	77.83$_{\pm0.20}$	78.90$_{\pm0.24}$	79.26$_{\pm0.21}$	74.37$_{\pm5.66}$
Late GCL	**65.54**$_{\pm0.27}$	68.32$_{\pm0.20}$	71.42$_{\pm0.24}$	70.55$_{\pm0.22}$	**79.54**$_{\pm0.20}$	**79.83**$_{\pm0.21}$	**78.31**$_{\pm0.20}$	**79.42**$_{\pm0.21}$	**79.69**$_{\pm0.23}$	**74.74**$_{\pm5.73}$
Early+Mid	65.23$_{\pm0.31}$	67.62$_{\pm0.24}$	71.50$_{\pm0.28}$	70.47$_{\pm0.19}$	78.90$_{\pm0.18}$	79.25$_{\pm0.20}$	77.41$_{\pm0.19}$	78.58$_{\pm0.24}$	79.22$_{\pm0.20}$	74.28$_{\pm5.56}$
Mid+Late	65.27$_{\pm0.28}$	**68.33**$_{\pm0.19}$	71.63$_{\pm0.28}$	70.30$_{\pm0.22}$	78.91$_{\pm0.17}$	79.57$_{\pm0.21}$	77.30$_{\pm0.20}$	78.85$_{\pm0.22}$	79.54$_{\pm0.24}$	74.41$_{\pm5.55}$
Early+Late	65.22$_{\pm0.21}$	67.25$_{\pm0.21}$	71.55$_{\pm0.27}$	**71.03**$_{\pm0.24}$	79.03$_{\pm0.20}$	79.41$_{\pm0.22}$	78.19$_{\pm0.23}$	78.70$_{\pm0.23}$	79.45$_{\pm0.22}$	74.43$_{\pm5.69}$
Full GCL	65.38$_{\pm0.22}$	68.22$_{\pm0.19}$	71.30$_{\pm0.24}$	70.77$_{\pm0.20}$	79.01$_{\pm0.19}$	79.29$_{\pm0.21}$	77.79$_{\pm0.20}$	78.71$_{\pm0.22}$	79.27$_{\pm0.19}$	74.42$_{\pm5.49}$



(a) DenseNet-121 (b) With our GCLs (c) MobileNet (d) With our GCLs

The relational graphs show that adding GCLs yields cleaner, tighter class clusters with fewer cross-class links, reducing feature noise and aligning features with semantic predictions