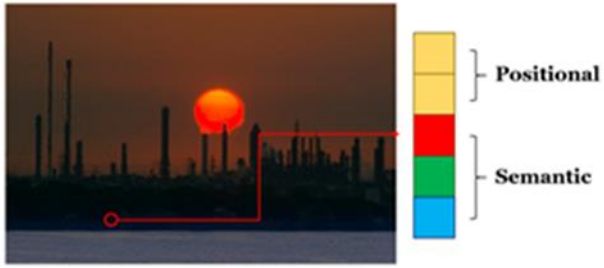
Multi-Instance Multi-Label Position-Aware Doubly Graph Convolutional Networks

Zhi LI, Teng ZHANG ,Yilin WANG, Caiwu JIANG, Xuanhua SHI, Hai JIN

Frontiers of Computer Science, DOI: 10.1007/s11704-025-41159-w

Problems & Ideas

- Problems in Conventional MIML Learning:
  - Positional Ignorance: Traditional methods neglect spatial relationships (e.g., pixel positions in images).
  - Flat Label Modeling: Labels are treated as independent, ignoring hierarchical dependencies (e.g., "sunrise" vs. "noon" depends on sub-concepts like "horizon").
- Proposed Solution: padGCN:
  - Position-Aware Graphs: Organize instances into graphs using positional (e.g., coordinates) and semantic features.
  - Hyperbolic Hierarchy: Model labels as a tripartite graph (sub-sub-concepts → sub-concepts → labels) in hyperbolic space.



padGCN aggregates instances based on positional proximity (e.g., sun and horizon).

Main Contributions

- Key Results:
  - Performance Boost: Average 5% improvement over SOTA on 6 datasets (MS-COCO, FLICKR25K, etc.).
  - Robustness: Outperforms baselines in 30/42 experiments (pairwise t-test, $p < 0.05$).

- Core Conclusion:
  - Positional Awareness and Hyperbolic Label Hierarchy are critical for real-world MIML tasks (e.g., autonomous driving, medical imaging).

Table 2 This table presents the performance of **padGCN** compared to nine baseline methods on Scene data set.

Metrics	HL ↓	OE ↓	CE ↓	RL ↓	AP ↑	AR ↑	AF_1 ↑	w/t/l
padGCN	**0.111±0.023**	**0.156±0.053**	**0.536±0.119**	**0.073±0.029**	**0.903±0.034**	**0.766±0.058**	**0.828±0.048**	
MIMLBoost	0.193±0.007•	0.347±0.019•	0.984±0.049•	0.178±0.011•	0.779±0.012•	0.433±0.027•	0.556±0.023•	6/0/0
MIMLSVM	0.189±0.009•	0.354±0.022•	1.087±0.047•	0.201±0.011•	0.765±0.013•	0.556±0.020•	0.644±0.018•	6/0/0
MIMLfast	0.188±0.009•	0.351±0.023•	0.207±0.012	0.189±0.014•	0.770±0.015•	0.477±0.273•	0.607±0.142•	5/1/0
DeepMIML	0.137±0.038•	0.287±0.120•	0.969±0.290•	0.181±0.074•	0.807±0.079•	0.536±0.181•	0.633±0.164•	6/0/0
M3GN	0.140±0.011•	0.215±0.025•	0.669±0.064•	0.102±0.014•	0.865±0.015•	0.670±0.036•	0.754±0.028•	6/0/0
MIML-GAN	0.253±0.075•	0.417±0.079•	1.195±0.249•	0.228±0.061•	0.556±0.069•	0.663±0.074•	0.587±0.049•	6/0/0
MIML-LLMC	0.250±0.005•	0.717±0.024•	0.730±0.010•	0.477±0.014•	0.503±0.016•	0.497±0.018•	0.503±0.016•	6/0/0
npGCN	0.135±0.013•	0.234±0.038•	0.825±0.113•	0.137±0.026•	0.864±0.026•	0.608±0.045•	0.734±0.036•	6/0/0
esGCN	0.147±0.028•	0.265±0.079•	1.038±0.127•	0.133±0.035•	0.889±0.032•	0.669±0.166•	0.796±0.052•	6/0/0

padGCN (blue) achieves lower Hamming Loss and higher F_1-score than SOTA methods (gray).