Topological Deep Learning Challenge 2026: Bridging the Gap

Guillermo Bernárdez*, Lev Telyatnikov*, Mathilde Papillon*, Marco Montagna, Louisa Cornelis, Louis Van Langendonck, Antonio Purificato, Bruno Neri, David Leko, Donatella Genovese, Eleonora Grassucci, Jan-Willem Van Looy, Johan Mathe, Lucia Testa, Luka Benić, Melanie Weber, Olga Zaghen, Pavlo Vasylenko, Raffael Theiler, Scott Mahan, Gaurav Khanal, Valentina Sánchez, Dario Loi, Gabriele Onorato, Prasanna Devadiga, Taejin Paik, Anier Velasco Sotomayor, Jonas Müller, Aarav Gupta, Riyan Patel, Grold Otieno Mboya, Pushparaj Devadiga, Adam Gohain, Dominik Filipiak, Michael Banf, Alireza Mohammadian, Amirreza Akbari, Gabriel Chaviano Díaz, Marcos Chávez Mesa, Ankit Grover, Rémi Bourgerie, Alessio Comparini, Ankit Kumar, Mohammad Yusuf, Devesh Kaushal, Francesco Orsi, Ghazal Mostofi, Kolya Lettl, Salwa Khattami, Gabriel Franco, Yelizaveta Falkouskaya, Jorge Luiz Franco, Kajal Singla, Fernando David Espinosa Iñiguez, Raj Patil, Eric Dolores-Cuenca, Luigi Petti, Chenguang Xu, Stefano Blando, Christine Zeh, Lennart Bastian, Kaden Zheng, Vihan Tiwari, Sud Athreya, Hisataka Kajita, Hugo Micheron, Rémi Devaux, Antoine Jardin, Tegan Emerson, Olga Fink, Nina Miolane.

* Main organizers, equal contribution. Corresponding email:
`topological.intelligence@gmail.com`

Abstract

This paper describes the 2026 edition of the *Topological Deep Learning Challenge: Bridging the Gap*, hosted at the second Topology, Algebra, and Geometry in Data Science (TAG-DS) Conference. Where previous editions grew the modelling and data ecosystem of Topological Deep Learning (TDL), this edition set out to place Graph Neural Networks (GNNs) and Topological Neural Networks (TNNs) side by side under a single evaluation pipeline. Participants implemented recent state-of-the-art architectures in `TopoBench`, with Track 1 covering GNNs and Track 2 covering TNNs. They evaluated these implementations on shared synthetic benchmarks generated with `GraphUniverse`, which independently controls homophily, average degree and degree-distribution heaviness. The challenge yielded 68 qualifying submissions from 34 teams, spanning the graph, simplicial, cellular, hypergraph and combinatorial domains. For the first time, submissions were also tracked on a public leaderboard, which we summarize and analyze here. Our analysis finds that the differences between the submitted tracks are strongly task-dependent: the mean accuracy gap is modest on community detection, while on triangle counting the Track 1 mean error is 1.85 times the Track 2 mean. Several structurally informed GNNs also outperform the Track 2 mean on triangle counting.

© 2026 G. Bernárdez, L. Telyatnikov, M. Papillon et al.

1. Introduction

Many physical systems, from molecules to galaxies, are driven by interactions among their components: chemical interactions between atoms, functional connectivity between brain regions, or gravitational forces between stars. Conventionally, these relations are modeled with graphs, allowing Graph Neural Networks (GNNs) (Scarselli et al., 2008) to extract key insights (Kipf and Welling, 2017). However, these systems also involve richer interactions that extend beyond the pairwise relations graphs encode. Topological Neural Networks (TNNs) have emerged to address this limitation by accounting for higher-order interactions (Hajij et al., 2022; Bodnar, 2022; Papillon et al., 2023b), extending message passing to simplicial, cellular, hypergraph and combinatorial domains using concepts from algebraic topology (Bick et al., 2023; Barbarossa and Sardellitti, 2020).

Despite this shared lineage, the GNN and Topological Deep Learning (TDL) communities have largely operated in parallel. The two fields are often presented as distinct paradigms, yet they are deeply intertwined, and a growing body of work studies the interplay between standard message passing and higher-order representations. Comparisons between GNNs and TNNs remain difficult because published results typically use different codebases, datasets, tasks, and tuning budgets. This leaves a central empirical question: *when does higher-order structure improve performance?*

Building on the three previous editions of the challenge (Papillon et al., 2023a; Bernárdez et al., 2024; Bernárdez et al., 2026), the *Topological Deep Learning Challenge 2026* set the stage for closing this gap. Participants were asked to implement recent state-of-the-art architectures from either literature into `TopoBench` (Telyatnikov et al., 2025), the open-source platform for developing and benchmarking TDL models, and to evaluate them through a single shared pipeline on synthetic graphs generated with `GraphUniverse` (Van Langendonck et al., 2026). Because `GraphUniverse` generates graphs with prescribed structural properties, the resulting benchmark allows us to examine how each architecture responds to changes in homophily, average degree and degree-distribution heaviness under a controlled generative design.

The challenge concluded with 68 valid submissions from 34 teams, submitted by 45 registered participants. Thirty-seven submissions implemented graph architectures in Track 1; thirty-one implemented higher-order architectures in Track 2, distributed across the simplicial, cellular, hypergraph and combinatorial domains. For the first time in the history of the challenge, submissions were additionally evaluated on a common benchmark suite and tracked on a public leaderboard. This paper details the challenge setup and submission requirements, lists all submissions and winners, and presents an analysis of the leaderboard results.

2. Setup of the Challenge

The challenge[1] was hosted at the second Topology, Algebra, and Geometry in Data Science (TAG-DS) Conference and sponsored by Arlequin AI and New Theory. Participants were

1. Challenge repository: https://github.com/geometric-intelligence/TopoBench. Challenge website: https://geometric-intelligence.github.io/topobench/tdl-challenge-2026/index.html.

asked to implement a recent state-of-the-art message-passing model into `TopoBench` in one of two tracks:

- **Track 1 — Graph Neural Networks (GNNs).** Architectures operating on graph structures, including message-passing networks and graph transformers with structural encodings.
- **Track 2 — Topological Neural Networks (TNNs).** Higher-order message-passing models leveraging rich topological domains, namely simplicial, cellular, hypergraph and combinatorial complexes.

While the tracks were judged separately, the cross-track analysis uses a common evaluation protocol and controlled dataset family.

2.1. Guidelines

Participation was free and open to everyone. To enroll, participants had to open a Pull Request (PR) on `TopoBench`, fill out the Registration Google Form with PR and team information, and tag the PR with the appropriate track label (`track-1-gnn` or `track-2-tnn`). Teams were accepted with a maximum of two members, though larger teams were possible upon approval. A single team was permitted to submit multiple PRs, both within and across tracks. The final submission deadline was August 1st, 2026 (AoE).

Submissions were required to implement models already proposed in the literature and to cite the associated publication or pre-print in the PR description. Each PR could contain at most one core model architecture; models with several significant variants were asked to supply a separate configuration file per variant. Participants were encouraged to submit early, both to leave time to resolve integration issues and because, where multiple high-quality submissions covered the same architecture, earlier submissions were given priority.

Consistent with the aims of an open environment for sharing participation, this activity was completely voluntary. However, due to legal restrictions, individuals affiliated with institutions on the Restricted Foreign Research Institutions list were not eligible for reward outcomes, including co-authorship on this white paper.

2.2. Submission Requirements

A submission consisted of a Pull Request to `TopoBench` repository, titled following the format `Track: [Track1|Track2]; Team Name: <team name>; Model: <model name>`. To be valid, the PR had to pass all integration tests, including linting and formatting. Three components were required of every submission, across tracks:

1. **Backbone.** The model implemented as a `torch.nn.Module` and placed in `topobench/nn/backbones/{domain}/{model_name}.py`, where {`domain`} is one of `graph`, `simplicial`, `cell`, `hypergraph`, `combinatorial` or `non_relational`.
2. **Hydra configuration.** A YAML file at `configs/model/{domain}/{model_name}.yaml` specifying the full `TBModel` composition: feature encoder, backbone, backbone wrapper and readout.

3. **Results file.** The `results.json` produced by running the provided evaluation notebook, `TopoBench/2026_tdl_challenge/run_evaluation.ipynb`, on the submitted model. This file records the model's performance on the benchmark tasks and available resource measurements, and is what feeds the public leaderboard.

Beyond the backbone, participants were encouraged to exploit `TopoBench`'s modular architecture by contributing custom feature encoders, readouts, backbone wrappers or losses where their chosen model required them, each inheriting from the corresponding abstract base class and discovered automatically by the framework. This was a deliberate design choice: it enriches the ecosystem with reusable components rather than with monolithic model files.

Testing. All contributed files had to pass the pre-existing unit tests, and any method or class not already covered had to be accompanied by new tests mirroring the source directory structure. Coverage, measured with Codecov, had to match or exceed 93%. Each submission also had to fill in `test/pipeline/test_pipeline.py` to demonstrate that the full training pipeline runs end-to-end, and the evaluation notebook had to run successfully with the submitted model.

2.3. Evaluation Method

To decide the winners in each track, a panel of `TopoBench` maintainers and collaborators voted on their preferred submissions, and the Condorcet method was used to produce a ranking. Reviewers were asked to base their rankings on the following criteria, which were visible to participants throughout:

- **Correctness:** Does the submission correctly implement the model as described in the literature? Modifications required to respect `TopoBench`'s computational constraints were permitted.
- **Code quality:** How readable and clean is the implementation? How well does the submission respect the requirements, including unit tests and memory usage?
- **Benchmark on GraphUniverse datasets:** Is the model correctly benchmarked on the provided datasets? Does it produce reasonable results?
- **Documentation & tests:** Do docstrings clearly describe the code, and do they make explicit reference to the original equations in the associated paper and/or the original implementation? Are unit tests robust?

The award criteria assess the quality and reusability of model implementations, not their final benchmark performance. No minimum training performance was required, and the highest leaderboard score did not determine the winner. The leaderboard supports the comparative analysis in Section 5.

For the first time, reviewers were permitted to use AI tools to help manage the volume of submissions. In Track 1, an AI reviewer agent was given the same instructions as the human reviewers. It graded each submission against the rubric and flagged the bottom tier of submissions. Track 1 reviewers could use these flags at their discretion. Reviewers were asked to disclose their use of AI in their responses.

2.4. Prizes

With thanks to our sponsors, this edition offered cash prizes of $1,000 USD for first place and $400 USD for second place in each track, with Track 1 sponsored by New Theory and Track 2 by Arlequin AI. This year, honorable mentions also carried cash prizes, with $700 USD split across outstanding submissions. In addition, selected teams were invited for research internships at the Geometric Intelligence Lab (University of California, Santa Barbara) and the Intelligent Maintenance and Operations Systems (IMOS) Lab (École Polytechnique Fédérale de Lausanne, Switzerland).

2.5. Software Practices

All submitted code had to comply with the challenge's GitHub Action workflow, passing all tests, linting and formatting checks. To ensure consistency, participants working in the simplicial, cellular or combinatorial domains were asked to use TopoNetX's (Hajij et al., 2024) classes to manage the corresponding complexes.

3. Benchmarking Protocol

The defining novelty of this edition is the use of a shared evaluation pipeline and dataset family for the cross-track comparison. This section describes that protocol.

3.1. Controlled Synthetic Data

Evaluation datasets were generated with `GraphUniverse` (Van Langendonck et al., 2026), a framework for sampling synthetic graphs with prescribed structural properties. Three structural axes were varied independently:

- **Homophily**, at three levels: low (0–0.1), mid (0.4–0.6) and high (0.9–1.0).
- **Average degree**, at two levels: low (1–2.5) and high (4–5).
- **Power-law exponent** γ of the degree distribution, at two levels: low (1.5–2, heavy-tailed) and high (4–5, homogeneous).

The full grid therefore contains $3 \times 2 \times 2 = 12$ configurations. The prescribed evaluation uses three training seeds per configuration, totaling 36 runs per task and 72 per model variant. This controlled design allows us to examine how model performance changes across prescribed structural regimes.

3.2. Tasks and Metrics

We consider two tasks, probing community structure and local substructure:

- **Community detection**, a node-classification task scored by accuracy (higher is better). This task is well aligned with the inductive bias of standard message passing: it rewards models that aggregate and smooth information over neighbourhoods.

- **Triangle counting**, a regression task scored by $\mathrm{MSE}_{\mathrm{test}} / \sum_{G \in \mathcal{D}_{\mathrm{test}}} \tau(G)$ (lower is better), where $\tau(G)$ counts triangles in each test graph's undirected edge structure after preprocessing. Counting triangles is provably beyond the reach of 1-WL-bounded message-passing networks.

3.3. Leaderboard

Results were published on a public leaderboard[2]. For each structural setting, scores are reported as the mean ± standard deviation of the in-distribution test metric across three training seeds; summaries across settings include up to 36 runs per task. The analysis uses the submitted result reports, with one variant per PR selected as described in Section 5.

Two caveats underscore these results. First, the scores describe the implementations and configurations submitted by participants rather than the best achievable performance of each method. Second, implementations were contributed by different participants with differing familiarity with the underlying papers. While the review process checked for correctness as part of the rubric, residual implementation variance potentially impacts leaderboard performance.

4. Submissions and Winners

The challenge received 68 valid submissions from 34 teams, submitted by 45 registered participants. Participation was close to evenly split between the two tracks: 37 submissions in Track 1 and 31 in Track 2. Within Track 2, submissions covered all four higher-order domains supported by `TopoBench`: ten simplicial, six cellular, seven hypergraph and eight combinatorial. Table 1 lists all submitted models, grouped by track and domain, with references to the underlying methods and links to their implementations. Additionally, Table 4 in Appendix A lists the co-authors belonging to each team, as well as the awards each team received.

2. https://geometric-intelligence.github.io/topobench/leaderboard/index.html

Track 1 — GNNs

Graph	37 submissions
Model	*PR*
AdvDIFFormer (Wu et al., 2025)	399
BuNN (Bamberger et al., 2025)	319
CoGNN (Finkelshtein et al., 2024)	369
Conn-NSD (Barbero et al., 2022a)	327
DNSD (Bourgerie et al., 2026)	375
DSNN (Fiorini et al., 2026)	412
ESC-GNN (Yan et al., 2024)	419
FAGCN (Bo et al., 2021)	382
FAGCN (Bo et al., 2021)	388
FiGURe (Ekbote et al., 2023)	360
GAMMA (Ghazizadeh et al., 2025)	411
GATE (Mustafa and Burkholz, 2024)	355
GAUGE (Sun et al., 2026)	396
GFT (Wang et al., 2024)	334
GPR-GNN (Chien et al., 2021)	352
GREAD (Choi et al., 2023)	415
GREAD (Choi et al., 2023)	420
GRIT (Ma et al., 2023)	370
GSN (Bouritsas et al., 2023)	407
HeroFilter (Zhang et al., 2025a)	384
HiD-Net (Li et al., 2024)	335
HOD-GNN (Eitan et al., 2026)	402
HyMN (Southern et al., 2025)	421
ISN (Hernandez Caralt et al., 2026)	374
JacobiConv (Wang and Zhang, 2022)	383
LINKX (Lim et al., 2021)	361
MTGCN (Pei et al., 2024)	400
NSD (Bodnar et al., 2022)	332
NSP (Suk et al., 2022)	326
NSP + Holonomy (Suk et al., 2022)	373
OGFormer (Zhang et al., 2025b)	378
PolyFilter GNN (Favard) (Guo and Wei, 2023)	354
Polynormer (Deng et al., 2024)	345
r-lMPNN (Paolino et al., 2024)	404
SGFormer (Wu et al., 2023)	359
SheafAN (Barbero et al., 2022b)	322
TopoPolynormer (Deng et al., 2024)	353

Track 2 — TNNs

Simplicial	10 submissions
Model	*PR*
Bi-SCNN (Yan and Kuruoglu, 2025)	387
COSIMO (Einizade et al., 2025)	392
CTNN (Hajij et al., 2025)	405
Dir-SNN (Lecha et al., 2025)	328
DirSNN (memory-safe, harmonic-aware) (Lecha et al., 2025)	350
GSAN (Battiloro et al., 2024)	365
HiGCN (Huang et al., 2024)	364
MPSN (Bodnar et al., 2021b)	376
MPSN (Bodnar et al., 2021b)	393
PolynomialFilterTNN (SNN) (Ebli et al., 2020)	414
Cellular	6 submissions
AirTNN (Fiorellino et al., 2024)	329
AirTNN (Fiorellino et al., 2024)	394
Cellular Transformer (Ballester et al., 2024)	371
CIN (Bodnar et al., 2021a)	363
DiffLift-SMCN (Franco et al., 2026; Eitan et al., 2025)	406
SheafTSP (Tandon et al., 2026)	410
Hypergraph	7 submissions
DPHGNN (Saxena et al., 2024)	398
DSHN (Mule et al., 2026)	408
EHNN (Kim et al., 2022)	366
HyperGCN (Yadati et al., 2019)	413
PhenomNN (Wang et al., 2023)	367
SheafHyperGNN (Duta et al., 2023)	321
WHNN (Duta and Liò, 2025)	417
Combinatorial	8 submissions
CTNN (copresheaf, gated) (Hajij et al., 2025)	418
ETNN (Battiloro et al., 2025)	320
ETNN-CoordinatePolicy (structural-LapPE) (Battiloro et al., 2025)	391
ETNN-LapPE (Battiloro et al., 2025)	362
HMC (Hajij et al., 2022)	416
SMCN (Eitan et al., 2025)	397
TopNets (Verma et al., 2024)	333
TopoU-Net (Gaurav et al., 2026)	372

Table 1: All 68 submissions, grouped by track and domain, with models ordered alphabetically within each domain. Citations identify the underlying methods, and PR numbers link to each implementation. Repeated models denote separate submissions. Appendix A lists teams; Appendix B reports scores for the 65 entries in the comparative analysis.

4.1. Winners and Honorable Mentions

The winners were announced publicly at the second TAG-DS Conference, on social media, and on the official challenge website. We emphasize that regardless of final rankings, the organizers and reviewers were impressed by the quality of the submissions and their contribution to the `TopoBench` ecosystem. We warmly congratulate all participants.

Track 1 — Graph Neural Networks.

- **First place ($1,000):** *DSNN*, Team *howyadoin* (Dario Loi and Gabriele Onorato; PR #412).
- **Joint second place ($200 each):** *GRIT*, Team *TripleA* (Amirreza Akbari and Alireza Mohammadian; PR #370), and *HOD-GNN*, Team *SweetLesson* (Aarav Gupta and Riyan Patel; PR #402).

Track 2 — Topological Neural Networks. The Condorcet evaluation resulted in a tie, and the Track 2 prize was divided equally between the two co-winners:

- **Joint first place ($700 each):** *ETNN-CoordinatePolicy*, Team *E(n)igma* (Gaurav Khanal; PR #391), and *TopoU-Net*, Team *TripleA* (Amirreza Akbari and Alireza Mohammadian; PR #372).

Honorable mentions.

- **Excellent portfolio ($150 each):** Teams *howyadoin* (Dario Loi and Gabriele Onorato) and *TripleA* (Amirreza Akbari and Alireza Mohammadian), recognised for multiple high-quality contributions across both tracks.
- **Novel infrastructure with differentiable liftings ($200):** *DiffLift-SMCN*, Team *One Ring to Lift Them All* (Jorge Luiz Franco; PR #406).
- **Outstanding supplementary analysis ($200):** *Neural Sheaf Propagation + Holonomy Measurement*, Team *Sheafu* (Ankit Grover and Rémi Bourgerie; PR #373 with a companion paper Grover and Bourgerie (2026)).

5. Leaderboard Results and Analysis

This section presents the first cross-track comparison in the history of the challenge. We include 65 entries[3]: 34 in Track 1 and 31 in Track 2. For each PR, we consider variants of the submitted model with complete benchmark results, select the variant with the highest mean in-distribution community-detection accuracy over the full evaluation grid, and use that same variant for both tasks[4]. For each task, we first average the three seed scores within each setting, then average the 12 setting means to obtain one score per PR. Track and domain means give each PR equal weight. Table 2 reports the per-track aggregates, and Table 3 breaks them down by topological domain. Complete leaderboards for both tasks are given in Appendix B.

3. The main analysis excludes PRs #319 and #327 because their reported test-graph triangle totals differ from those of the reference cohort, and PR #382 because it covers all 12 structural settings with only one training seed.
4. For PR #391, this is the structural-LapPE variant.

Track	n	All	Homophily			Avg. degree	
			Low	Mid	High	Low	High
Community detection — accuracy (↑)							
Track 1 (GNN)	34	0.4241	0.3088	0.3869	0.5765	0.3991	0.4491
Track 2 (TNN)	31	0.4343	0.3061	0.4000	0.5970	0.4099	0.4588
Triangle counting — MSE per triangle (↓)							
Track 1 (GNN)	34	0.7097	0.0887	0.6099	1.4304	0.3579	1.0614
Track 2 (TNN)	31	0.3828	0.0428	0.3251	0.7805	0.1858	0.5798

Table 2: Mean in-distribution test score per track, averaged over submissions in the 65-entry analysis cohort (Section 5). Columns marginalise the GraphUniverse configuration grid over homophily and average degree. PRs #319, #327 and #382 are excluded.

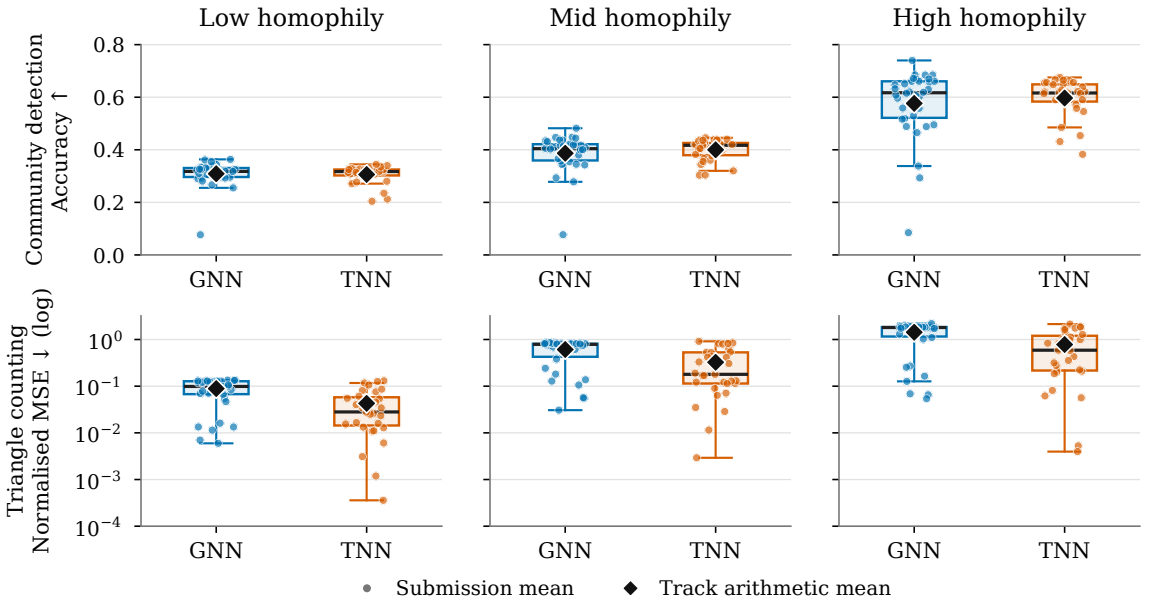


Figure 1: Submission-score distributions by homophily, complementing Table 2. Each dot is one submission's mean over four degree–power-law settings and three seeds per setting (34 GNNs; 31 TNNs). Boxes span the interquartile range, with a line marking the median. Whiskers extend to the most extreme observations within 1.5 interquartile ranges below the first quartile or above the third quartile. Diamonds mark arithmetic track means. Axes are shared within each task; triangle-counting error is shown on a logarithmic scale.

		Community detection (↑)				Triangle counting (↓)			
Domain	n	**All**	h_{lo}	h_{mid}	h_{hi}	**All**	h_{lo}	h_{mid}	h_{hi}
Graph (T1)	34	0.4241	0.3088	0.3869	0.5765	0.7097	0.0887	0.6099	1.4304
Simplicial	10	0.4411	0.3141	0.4030	0.6062	0.2586	0.0119	0.2184	0.5455
Cell	6	0.4356	0.2846	0.3994	0.6227	0.3144	0.0467	0.2652	0.6314
Hypergraph	7	0.4451	0.3196	0.4067	0.6090	0.6505	0.0867	0.5253	1.3393
Combinatorial	8	0.4155	0.3004	0.3907	0.5555	0.3551	0.0403	0.3281	0.6970

Table 3: Mean in-distribution test score per topological domain, broken down by homophily regime. Track 1 submissions all operate on the graph domain. The 65-entry analysis cohort is defined in Section 5.

5.1. Performance Differences Across Tasks

The clearest takeaway from these results is that the mean performance advantage of TNNs over GNNs is task-dependent. For example, it is modest on community detection and more pronounced on triangle counting.

On **community detection**, Track 2 attains a mean accuracy of 0.4343 against 0.4241 for Track 1, a modest descriptive difference of 1.02 percentage points. The best individual result comes from Track 1: GRIT (PR #370) reaches 0.5207, ahead of the best Track 2 model, HiGCN (PR #364), at 0.4869.

On **triangle counting**, the picture changes sharply. Track 2 attains a mean error of 0.3828 against 0.7097 for Track 1; the Track 1 mean is 1.85 times the Track 2 mean. The best submission overall is a combinatorial model, CTNN (PR #418), at 0.0024—more than two orders of magnitude below the Track 1 mean. This is consistent with the motivation behind encoding higher-order structure. We note that the track comparison does not isolate the contribution of that structure from other architectural and implementation choices.

Table 3 gives a more detailed view. On triangle counting, simplicial submissions have the lowest mean error (0.2586), followed by cellular (0.3144), combinatorial (0.3551), hypergraph (0.6505) and graph submissions (0.7097).

5.2. Structural Information Beyond Track Labels

Going beyond track aggregations, we find that several of the strongest triangle-counting submissions are Track 1 models. For example, TopoPolynormer (PR #353, 0.0451), ESC-GNN (PR #419, 0.0552), GSN (PR #407, 0.0665), GRIT (PR #370, 0.0693) and HOD-GNN (PR #402, 0.0802) all outperform the Track 2 mean, and several outperform the majority of Track 2 submissions. A common property of these architectures is that while they operate on graph-based domains, they inject substructure or structural-encoding information, such as subgraph counts, random-walk encodings, or explicit higher-order descriptors.

These results show that a graph-domain model equipped with structural information can perform competitively on triangle counting. Higher-order domains offer another way to supply structural information, but performance also varies substantially among TNN submissions. Establishing the separate contributions of the lifting, structural encoding and message-passing scheme via controlled ablations is a promising avenue for future work.

5.3. Structural Sensitivity

The controlled grid allows us to ask how each family responds to the structural axes. Figure 1 reveals substantial overlap between GNNs and TNNs on community detection. At high homophily, their median accuracies are close (0.6170 and 0.6161, respectively), despite the larger difference between their means. On triangle counting, TNNs achieve lower median error across all three homophily regimes, showing that their advantage is also reflected in the medians. Nevertheless, performance varies considerably within each family, and several GNN submissions achieve errors below the TNN median in every homophily regime.

Homophily. On community detection, both tracks lose accuracy as homophily drops: from 0.5765 to 0.3088 for Track 1 and from 0.5970 to 0.3061 for Track 2. The Track 2 mean exceeds the Track 1 mean by 2.04 percentage points at high homophily and 1.31 at mid homophily, while Track 1 leads by 0.28 percentage points at low homophily. Thus, the observed Track 2 advantage does not extend to the low-homophily regime on this task. On triangle counting, the direction is reversed: mean error decreases as homophily falls, from 1.4304 to 0.0887 for Track 1 and from 0.7805 to 0.0428 for Track 2.

Average degree and degree distribution. Both tracks perform better on community detection at high average degree than at low (0.4491 versus 0.3991 for Track 1; 0.4588 versus 0.4099 for Track 2), consistent with denser neighborhoods carrying more signal. On triangle counting the pattern inverts, with normalized error rising at high degree and at high homophily in both tracks. Heavy-tailed degree distributions (low γ) also yield higher mean normalized error than homogeneous ones in both tracks.

5.4. Leaderboard Position and Award Outcome

Because both a leaderboard and a set of awards appear in this paper for the first time, we make their relationship explicit. As described in Section 2.3, awards were decided by Condorcet vote on implementation quality, not on benchmark score, and the two rankings differ considerably: the Track 1 winner, DSNN (PR #412), ranks 62nd of 65 on community detection, and the Track 2 co-winner TopoU-Net (PR #372) ranks 59th. The awards recognize implementation quality and reusability; Tables 5 and 6 describe performance on the two benchmark tasks. Award outcomes and leaderboard positions should therefore be interpreted separately.

6. Discussion & Concluding Remarks

This edition set out to answer two questions: how specific graph properties affect classical GNNs versus higher-order TNNs, and under what data regimes TNNs provide capabilities that standard GNNs do not. The results above provide a first, qualified answer to both.

On the first question, the community-detection gap narrows as homophily falls, with a small reversal in favour of GNNs at low homophily; a higher-order domain alone therefore does not establish robustness in that regime. On the second question, the submitted Track 2 cohort has lower mean error on triangle counting, while its mean accuracy advantage on community detection is small. The strong triangle-counting results of several Track 1 models suggest that both higher-order architectures and structurally informed graph models merit

consideration for this task. Extending this observation to other motifs or tasks requires further evaluation.

We highlight the finding that substructure-aware GNNs close much of the gap on triangle counting as it points to a possible reframing of the comparison. Rather than a strict dichotomy between two model families, the results motivate studying the structural information supplied to models in both tracks. Both communities therefore face the same practical question: how to give networks access to structure beyond edges.

Several limitations bound these conclusions. The benchmark comprises two tasks on one family of synthetic graphs; performance on real-world data with correlated structural properties may differ. The results concern the submitted configurations, and their relative performance may change under different tuning budgets. Selecting variants by reported test accuracy can favour PRs with more variants and make community-detection performance estimates optimistic. Implementation quality, while reviewed, inevitably varies. Moreover, the aggregate statistics we report summarize a heterogeneous population of architectures, so track-level means should be interpreted as descriptions of the submissions rather than defined properties of the underlying model classes. Extending the benchmark suite with additional tasks and re-running it with per-model tuning are the natural next steps. Both are straightforward within `TopoBench`.

The challenge yielded 68 valid submissions from 34 teams, enriching `TopoBench` with implementations from both the GNN and TDL literatures. It also produced the first cross-track comparison in this challenge series under a shared benchmarking protocol, supported by a public leaderboard. We hope this infrastructure hosted in `TopoBench` will support research within the communities, facilitate comparisons between models, and encourage the exchange of methods and ideas between them.

Acknowledgments

We thank Arlequin AI and New Theory for sponsoring this edition of the challenge, the Geometric Intelligence Lab (UC Santa Barbara) and the IMOS Lab (EPFL) for supporting the internship opportunities, and the TAG-DS organizers for hosting the challenge. We are grateful to all `TopoBench` maintainers and collaborators who served on the review panel, and above all to the participants, whose contributions constitute the substance of this paper. N. M., G. B., and M. P. acknowledge partial support from NSF grant 2602079 and the Noyce foundation.

References

Rubén Ballester, Pablo Hernández-García, Mathilde Papillon, Claudio Battiloro, Nina Miolane, Tolga Birdal, Carles Casacuberta, Sergio Escalera, and Mustafa Hajij. Attending to Topological Spaces: The Cellular Transformer. arXiv preprint arXiv:2405.14094, 2024. URL https://arxiv.org/abs/2405.14094.

Jacob Bamberger, Federico Barbero, Xiaowen Dong, and Michael M. Bronstein. Bundle Neural Network for message diffusion on graphs. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=scI9307PLG.

Sergio Barbarossa and Stefania Sardellitti. Topological signal processing over simplicial complexes. *IEEE Transactions on Signal Processing*, 68:2992–3007, 2020.

Federico Barbero, Cristian Bodnar, Haitz Sáez de Ocáriz Borde, Michael Bronstein, Petar Veličković, and Pietro Liò. Sheaf Neural Networks with Connection Laplacians. In *Proceedings of Topological, Algebraic, and Geometric Learning Workshops 2022*, volume 196 of *Proceedings of Machine Learning Research*, pages 28–36. PMLR, 2022a. URL https://proceedings.mlr.press/v196/barbero22a.html.

Federico Barbero, Cristian Bodnar, Haitz Sáez de Ocáriz Borde, and Pietro Liò. Sheaf Attention Networks. NeurIPS 2022 Workshop on Symmetry and Geometry in Neural Representations, 2022b. URL https://openreview.net/forum?id=LIDvgVjpkZr.

Claudio Battiloro, Lucia Testa, Lorenzo Giusti, Stefania Sardellitti, Paolo Di Lorenzo, and Sergio Barbarossa. Generalized Simplicial Attention Neural Networks. *IEEE Transactions on Signal and Information Processing over Networks*, 10:833–850, 2024. doi: 10.1109/TSIPN.2024.3485473. URL https://ieeexplore.ieee.org/document/10735150/.

Claudio Battiloro, Ege Karaismailoğlu, Mauricio Tec, George Dasoulas, Michelle Audirac, and Francesca Dominici. E(n) Equivariant Topological Neural Networks. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=Ax3uliEBVR.

Guillermo Bernárdez, Lev Telyatnikov, Marco Montagna, Federica Baccini, Mathilde Papillon, Miquel Ferriol-Galmés, Mustafa Hajij, Theodore Papamarkou, Maria Sofia Bucarelli, Olga Zaghen, et al. Icml topological deep learning challenge 2024: Beyond the graph domain. In *Geometry-grounded Representation Learning and Generative Modeling Workshop (GRaM) at ICML 2024*, pages 420–428. PMLR, 2024.

Guillermo Bernárdez, Lev Telyatnikov, Mathilde Papillon, Marco Montagna, Raffael Theiler, Louisa Cornelis, Johan Mathe, Miquel Ferriol, Pavlo Vasylenko, Jan-Willem Van Looy, Lucia Testa, Bruno Neri, Donatella Genovese, Melanie Weber, Amaury Wei, Alessio Devoto, Alexander Weers, Robert Jankowski, Loris Cino, David Leko, Michael Banf, Jonas Müller, Thomas Grapentin, Taejin Paik, Abhijeet Dutta, Hugo Walter, Thomas Vaitses Fontanari, Ali Ghasemi, Dario Loi, Haitz Sáez de Ocáriz Borde, Gabriela Aguilar-Argüello, Giovanni B. da Rosa, Théo Saulus, Eric Rubiel Dolores-Cuenca, Leonardo Di Nino, Pierrick Leroy, Mario Edoardo Pandolfo, Andrea Cavallo, Yu Qin, Pavel Snopov, Amirreza Akbari,

Ixchel Meza-Chávez, Louis Van Langendonck, Jared Able, Maria Yuffa Meshcheryakova, Henry Tsay, Luka Benić, Dominik Filipiak, Patrick Liu, Huidong Liang, Alexsandro Santos da Rosa Jr., Tiziana Cattai, Henrique M. Borges, Enrico Grimaldi, Manuel Lecha, Claudio Battiloro, Xuan-Chen Liu, Raj Deshpande, Graham Johnson, Igor Morgunov, Hugo Micheron, Rémi Devaux, Antoine Jardin, Tegan Emerson, Olga Fink, and Nina Miolane. Topological deep learning challenge 2025: Expanding the data landscape. In Guillermo Bernardez Gil, Mitchell Black, Alexander Cloninger, Timothy Doster, Tegan Emerson, Ines García-Rodondo, Chester Holtz, Mit Kotak, Henry Kvinge, Gal Mishne, Mathilde Papillon, Alison Pouplin, Katie Rainey, Bastian Rieck, Lev Telyatnikov, Eric Yeats, Qingsong Wang, Yusu Wang, and Jeremy Wayland, editors, *Proceedings of the 1st Conference on Topology, Algebra, and Geometry in Data Science(TAG-DS 2025)*, volume 321 of *Proceedings of Machine Learning Research*, pages 4–14. PMLR, 01–02 Dec 2026. URL https://proceedings.mlr.press/v321/bernardez26a.html.

Christian Bick, Elizabeth Gross, Heather A Harrington, and Michael T Schaub. What are higher-order networks? *SIAM Review*, 65(3):686–731, 2023.

Deyu Bo, Xiao Wang, Chuan Shi, and Huawei Shen. Beyond Low-frequency Information in Graph Convolutional Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(5):3950–3957, 2021. doi: 10.1609/aaai.v35i5.16514. URL https://ojs.aaai.org/index.php/AAAI/article/view/16514.

Cristian Bodnar. *Topological Deep Learning: Graphs, Complexes, Sheaves*. PhD thesis, Apollo - University of Cambridge Repository, 2022. URL https://www.repository.cam.ac.uk/handle/1810/350982.

Cristian Bodnar, Fabrizio Frasca, Nina Otter, Yuguang Wang, Pietro Liò, Guido F. Montúfar, and Michael Bronstein. Weisfeiler and Lehman Go Cellular: CW Networks. In *Advances in Neural Information Processing Systems*, volume 34, pages 2625–2640, 2021a. URL https://proceedings.neurips.cc/paper/2021/hash/157792e4abb490f99dbd738483e0d2d4-Abstract.html.

Cristian Bodnar, Fabrizio Frasca, Yuguang Wang, Nina Otter, Guido F. Montúfar, Pietro Liò, and Michael Bronstein. Weisfeiler and Lehman Go Topological: Message Passing Simplicial Networks. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 1026–1037. PMLR, 2021b. URL https://proceedings.mlr.press/v139/bodnar21a.html.

Cristian Bodnar, Francesco Di Giovanni, Benjamin Chamberlain, Pietro Liò, and Michael Bronstein. Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs. In *Advances in Neural Information Processing Systems*, volume 35, pages 18527–18541, 2022. doi: 10.52202/068431-1346. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/75c45fca2aa416ada062b26cc4fb7641-Abstract-Conference.html.

Rémi Bourgerie, Šarūnas Girdzijauskas, and Viktoria Fodor. Deep Neural Sheaf Diffusion. arXiv preprint arXiv:2605.19021, 2026. URL https://arxiv.org/abs/2605.19021.

Giorgos Bouritsas, Fabrizio Frasca, Stefanos Zafeiriou, and Michael M. Bronstein. Improving Graph Neural Network Expressivity via Subgraph Isomorphism Counting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):657–668, 2023. doi: 10.1109/TPAMI.2022.3154319. URL https://ieeexplore.ieee.org/document/9721082/.

Eli Chien, Jianhao Peng, Pan Li, and Olgica Milenkovic. Adaptive Universal Generalized PageRank Graph Neural Network. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=n6jl7fLxrP.

Jeongwhan Choi, Seoyoung Hong, Noseong Park, and Sung-Bae Cho. GREAD: Graph Neural Reaction-Diffusion Networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 5722–5747. PMLR, 2023. URL https://proceedings.mlr.press/v202/choi23a.html.

Chenhui Deng, Zichao Yue, and Zhiru Zhang. Polynormer: Polynomial-Expressive Graph Transformer in Linear Time. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=hmv1LpNfXa.

Iulia Duta and Pietro Liò. Wasserstein Hypergraph Neural Network. arXiv preprint arXiv:2506.09682, 2025. URL https://arxiv.org/abs/2506.09682.

Iulia Duta, Giulia Cassarà, Fabrizio Silvestri, and Pietro Liò. Sheaf Hypergraph Networks. In *Advances in Neural Information Processing Systems*, volume 36, pages 12087–12099, 2023. doi: 10.52202/075280-0529. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/27f243af2887d7f248f518d9b967a882-Abstract-Conference.html.

Stefania Ebli, Michaël Defferrard, and Gard Spreemann. Simplicial Neural Networks. arXiv preprint arXiv:2010.03633, 2020. URL https://arxiv.org/abs/2010.03633.

Aref Einizade, Dorina Thanou, Fragkiskos D. Malliaros, and Jhony H. Giraldo. Continuous Simplicial Neural Networks. In *Advances in Neural Information Processing Systems*, volume 38, pages 59467–59495, 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/55c944c9d908af245fa658b0e02b9003-Abstract-Conference.html.

Yam Eitan, Yoav Gelberg, Guy Bar-Shalom, Fabrizio Frasca, Michael M. Bronstein, and Haggai Maron. Topological Blindspots: Understanding and Extending Topological Deep Learning Through the Lens of Expressivity. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=EzjsoomYEb.

Yam Eitan, Moshe Eliasof, Yoav Gelberg, Fabrizio Frasca, Guy Bar-Shalom, and Haggai Maron. On The Expressive Power of GNN Derivatives. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=Tvat33IDmK.

Chanakya Ekbote, Ajinkya Deshpande, Arun Iyer, Sundararajan Sellamanickam, and Ramakrishna Bairi. FiGURe: Simple and Efficient Unsupervised Node Representations with Filter Augmentations. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/6f479ea488e0908ac8b1b37b27fd134c-Abstract-Conference.html.

Ben Finkelshtein, Xingyue Huang, Michael M. Bronstein, and Ismail Ilkan Ceylan. Cooperative Graph Neural Networks. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 13633–13659. PMLR, 2024. URL https://proceedings.mlr.press/v235/finkelshtein24a.html.

Simone Fiorellino, Claudio Battiloro, and Paolo Di Lorenzo. Topological Neural Networks over the Air. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 12986–12990. IEEE, 2024. doi: 10.1109/ICASSP48485.2024.10446693. URL https://ieeexplore.ieee.org/document/10446693.

Stefano Fiorini, Hakan Aktas, Iulia Duta, Pietro Morerio, Alessio Del Bue, Pietro Liò, and Stefano Coniglio. Sheaves Reloaded: A Direction Awakening. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=iDiiETH7Qv.

Jorge Luiz Franco, Gabriel Duarte, Alexander V Nikitin, Moacir A Ponti, Diego Mesquita, and Amauri H Souza. Differentiable Lifting for Topological Neural Networks. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=eC89CbINIw.

Gaurav Gaurav, Ibrahem ALJabea, Yaroslav Zakomornyy, Eric Frank, Mohamed Elhamdadi, Theodore Papamarkou, and Mustafa Hajij. TopoU-Net: A U-Net Architecture for Topological Domains. arXiv preprint arXiv:2605.10091, 2026. URL https://arxiv.org/abs/2605.10091.

Amir Ghazizadeh, Rickard Ewetz, and Hao Zheng. GAMMA: Gated Multi-hop Message Passing for Homophily-Agnostic Node Representation in GNNs. In *Advances in Neural Information Processing Systems*, volume 38, pages 11726–11777, 2025. doi: 10.52202/085713-0393. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/1129f729097a28bfb5b836ec4bf94478-Abstract-Conference.html.

Ankit Grover and Rémi Bourgerie. Do sheaf neural networks use holonomy? a measure–intervene–control study, 2026. URL https://arxiv.org/abs/2607.19514.

Yuhe Guo and Zhewei Wei. Graph Neural Networks with Learnable and Optimal Polynomial Bases. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 12077–12097. PMLR, 2023. URL https://proceedings.mlr.press/v202/guo23i.html.

Mustafa Hajij, Ghada Zamzmi, Theodore Papamarkou, Nina Miolane, Aldo Guzmán-Sáenz, Karthikeyan Natesan Ramamurthy, Tolga Birdal, Tamal K. Dey, Soham Mukherjee, Shreyas N. Samaga, Neal Livesay, Robin Walters, Paul Rosen, and Michael T. Schaub. Topological Deep Learning: Going Beyond Graph Data. arXiv preprint arXiv:2206.00606, 2022. URL https://arxiv.org/abs/2206.00606.

Mustafa Hajij, Mathilde Papillon, Florian Frantzen, Jens Agerberg, Ibrahem AlJabea, Rubén Ballester, Claudio Battiloro, Guillermo Bernárdez, Tolga Birdal, Aiden Brent, et al.

Topox: a suite of python packages for machine learning on topological domains. *Journal of Machine Learning Research*, 25(374):1–8, 2024.

Mustafa Hajij, Lennart Bastian, Sarah Osentoski, Hardik Kabaria, John L. Davenport, Sheik Dawood, Balaji Cherukuri, Joseph G. Kocheemoolayil, Nastaran Shahmansouri, Adrian Lew, Theodore Papamarkou, and Tolga Birdal. Copresheaf Topological Neural Networks: A Generalized Deep Learning Framework. In *Advances in Neural Information Processing Systems*, volume 38, pages 149759–149803, 2025. doi: 10.52202/085713-5009. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/dc62cd4ec77f3bbec8e6245d0bd91d08-Abstract-Conference.html.

Ferran Hernandez Caralt, Mar Gonzàlez i Català, Adrián Bazaga, and Pietro Liò. On the Necessity of Learnable Sheaf Laplacians. arXiv preprint arXiv:2603.05395, 2026. URL https://arxiv.org/abs/2603.05395.

Yiming Huang, Yujie Zeng, Qiang Wu, and Linyuan Lü. Higher-Order Graph Convolutional Network with Flower-Petals Laplacians on Simplicial Complexes. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(11):12653–12661, 2024. doi: 10.1609/aaai.v38i11.29160. URL https://ojs.aaai.org/index.php/AAAI/article/view/29160.

Jinwoo Kim, Saeyoon Oh, Sungjun Cho, and Seunghoon Hong. Equivariant Hypergraph Neural Networks. In *Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXI*, volume 13681 of *Lecture Notes in Computer Science*, pages 86–103. Springer, Cham, 2022. doi: 10.1007/978-3-031-19803-8_6. URL https://link.springer.com/chapter/10.1007/978-3-031-19803-8_6.

Thomas N. Kipf and Max Welling. Semi-Supervised Classification with Graph Convolutional Networks. In *ICLR*, 2017.

Manuel Lecha, Andrea Cavallo, Francesca Dominici, Elvin Isufi, and Claudio Battiloro. Higher-Order Topological Directionality and Directed Simplicial Neural Networks. In *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2025. doi: 10.1109/ICASSP49660.2025.10890264. URL https://doi.org/10.1109/ICASSP49660.2025.10890264.

Yibo Li, Xiao Wang, Hongrui Liu, and Chuan Shi. A Generalized Neural Diffusion Framework on Graphs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(8):8707–8715, 2024. doi: 10.1609/aaai.v38i8.28716. URL https://ojs.aaai.org/index.php/AAAI/article/view/28716.

Derek Lim, Felix Hohne, Xiuyu Li, Sijia Linda Huang, Vaishnavi Gupta, Omkar Bhalerao, and Ser Nam Lim. Large Scale Learning on Non-Homophilous Graphs: New Benchmarks and Strong Simple Methods. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/ae816a80e4c1c56caa2eb4e1819cbb2f-Abstract.html.

Liheng Ma, Chen Lin, Derek Lim, Adriana Romero-Soriano, Puneet K. Dokania, Mark Coates, Philip Torr, and Ser-Nam Lim. Graph Inductive Biases in Transformers without

Message Passing. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 23321–23337. PMLR, 2023. URL https://proceedings.mlr.press/v202/ma23c.html.

Emanuele Mule, Stefano Fiorini, Antonio Purificato, Federico Siciliano, Stefano Coniglio, and Fabrizio Silvestri. Directional Sheaf Hypergraph Networks: Unifying Learning on Directed and Undirected Hypergraphs. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=d1hQkcDy7c.

Nimrah Mustafa and Rebekka Burkholz. GATE: How to Keep Out Intrusive Neighbors. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 36990–37015. PMLR, 2024. URL https://proceedings.mlr.press/v235/mustafa24a.html.

Raffaele Paolino, Sohir Maskey, Pascal Welke, and Gitta Kutyniok. Weisfeiler and Leman Go Loopy: A New Hierarchy for Graph Representational Learning. In *Advances in Neural Information Processing Systems*, volume 37, pages 120780–120831, 2024. doi: 10.52202/079017-3838. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/dad28e90cd2c8caedf362d49c4d99e70-Abstract-Conference.html.

Mathilde Papillon, Mustafa Hajij, Audun Myers, Helen Jenne, Johan Mathe, Theodore Papamarkou, Aldo Guzmán-Sáenz, Neal Livesay, Tamal Dey, Abraham Rabinowitz, Aiden Brent, Alessandro Salatiello, Alexander Nikitin, Ali Zia, Claudio Battiloro, Dmitrii Gavrilev, German Magai, Gleb Bazhenov, Guillermo Bernardez, Indro Spinelli, Jens Agerberg, Kalyan Nadimpalli, Lev Telyatninkov, Luca Scofano, Lucia Testa, Manuel Lecha, Maosheng Yang, Mohammed Hassanin, Odin Hoff Gardaa, Olga Zaghen, Paul Hausner, Paul Snopoff, Rubén Ballester, Sadrodin Barikbin, Sergio Escalera, Simone Fiorellino, Henry Kvinge, Karthikeyan Natesan Ramamurthy, Paul Rosen, Robin Walters, Shreyas N. Samaga, Soham Mukherjee, Sophia Sanborn, Tegan Emerson, Timothy Doster, Tolga Birdal, Abdelwahed Khamis, Simone Scardapane, Suraj Singh, Tatiana Malygina, Yixiao Yue, and Nina Miolane. Icml 2023 topological deep learning challenge: Design and results. In Timothy Doster, Tegan Emerson, Henry Kvinge, Nina Miolane, Mathilde Papillon, Bastian Rieck, and Sophia Sanborn, editors, *Proceedings of 2nd Annual Workshop on Topology, Algebra, and Geometry in Machine Learning (TAG-ML)*, volume 221 of *Proceedings of Machine Learning Research*, pages 3–8. PMLR, 28 Jul 2023a. URL https://proceedings.mlr.press/v221/papillon23a.html.

Mathilde Papillon, Sophia Sanborn, Mustafa Hajij, and Nina Miolane. Architectures of topological deep learning: A survey of message-passing topological neural networks. *arXiv preprint arXiv:2304.10031*, 2023b.

Hongbin Pei, Yu Li, Huiqi Deng, Jingxin Hai, Pinghui Wang, Jie Ma, Jing Tao, Yuheng Xiong, and Xiaohong Guan. Multi-Track Message Passing: Tackling Oversmoothing and Oversquashing in Graph Learning via Preventing Heterophily Mixing. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 40078–40091. PMLR, 2024. URL https://proceedings.mlr.press/v235/pei24a.html.

Siddhant Saxena, Shounak Ghatak, Raghu Kolla, Debashis Mukherjee, and Tanmoy Chakraborty. DPHGNN: A Dual Perspective Hypergraph Neural Networks. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 2548–2559. ACM, 2024. doi: 10.1145/3637528.3672047. URL https://dl.acm.org/doi/10.1145/3637528.3672047.

Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE transactions on neural networks*, 20(1): 61–80, 2008.

Joshua Southern, Yam Eitan, Guy Bar-Shalom, Michael M. Bronstein, Haggai Maron, and Fabrizio Frasca. Balancing Efficiency and Expressiveness: Subgraph GNNs with Walk-Based Centrality. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 56615–56649. PMLR, 2025. URL https://proceedings.mlr.press/v267/southern25a.html.

Julian Suk, Lorenzo Giusti, Tamir Hemo, Miguel Lopez, Konstantinos Barmpas, and Cristian Bodnar. Surfing on the Neural Sheaf. NeurIPS 2022 Workshop on Symmetry and Geometry in Neural Representations, 2022. URL https://openreview.net/forum?id=xOXFkyRzTlu.

Li Sun, Zhenhao Huang, Yiding Wang, Qin Chen, Pietro Liò, and Philip S. Yu. Are Common Substructures Transferable? Riemannian Graph Foundation Model with Neural Vector Bundles. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, pages 116668–116694. PMLR, 2026. URL https://proceedings.mlr.press/v306/sun26o.html.

Kartik Tandon, Julian Gould, Tanishq Bhatia, Francesca Dominici, Alejandro Ribeiro, and Claudio Battiloro. Consistent Geometric Deep Learning via Hilbert Bundles and Cellular Sheaves. arXiv preprint arXiv:2605.06395, 2026. URL https://arxiv.org/abs/2605.06395.

Lev Telyatnikov, Guillermo Bernardez, Marco Montagna, Mustafa Hajij, Martin Carrasco, Pavlo Vasylenko, Mathilde Papillon, Ghada Zamzmi, Michael T Schaub, Jonas Verhellen, Pavel Snopov, Bertran Miquel-Oliver, Manel Gil-Sorribes, Alexis Molina, Victor Guallar, Theodore Long, Julian Suk, Patryk Rygiel, Alexander V Nikitin, Giordan Escalona, Michael Banf, Dominik Filipiak, Liliya Imasheva, Max Schattauer, Alvaro L. Martinez, Halley Fritze, Marissa Masden, Valentina Sánchez, Manuel Lecha, Andrea Cavallo, Claudio Battiloro, Matthew Piekenbrock, Mauricio Tec, George Dasoulas, Nina Miolane, Simone Scardapane, and Theodore Papamarkou. Topobench: A framework for benchmarking topological deep learning. *Journal of Data-centric Machine Learning Research*, 2025. URL https://openreview.net/forum?id=07sTzyEVtY.

Louis Emiel T Van Langendonck, Guillermo Bernárdez, Nina Miolane, and Pere Barlet-Ros. Graphuniverse: Synthetic graph generation for evaluating inductive generalization. In *International Conference on Learning Representations*, volume 2026, pages 138611–138642, 2026.

Yogesh Verma, Amauri H Souza, and Vikas Garg. Topological Neural Networks go Persistent, Equivariant, and Continuous. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 49388–49407. PMLR, 2024. URL https://proceedings.mlr.press/v235/verma24a.html.

Xiyuan Wang and Muhan Zhang. How Powerful are Spectral Graph Neural Networks. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 23341–23362. PMLR, 2022. URL https://proceedings.mlr.press/v162/wang22am.html.

Yuxin Wang, Quan Gan, Xipeng Qiu, Xuanjing Huang, and David Wipf. From Hypergraph Energy Functions to Hypergraph Neural Networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 35605–35623. PMLR, 2023. URL https://proceedings.mlr.press/v202/wang23d.html.

Zehong Wang, Zheyuan Zhang, Nitesh V Chawla, Chuxu Zhang, and Yanfang Ye. GFT: Graph Foundation Model with Transferable Tree Vocabulary. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-3412. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/c23ccf9eedf87e4380e92b75b24955bb-Abstract-Conference.html.

Qitian Wu, Wentao Zhao, Chenxiao Yang, Hengrui Zhang, Fan Nie, Haitian Jiang, Yatao Bian, and Junchi Yan. SGFormer: Simplifying and Empowering Transformers for Large-Graph Representations. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-2826. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/cc57fac10eacadb3b72a907ac48f9a98-Abstract.html.

Qitian Wu, Chenxiao Yang, Kaipeng Zeng, and Michael M. Bronstein. Supercharging Graph Transformers with Advective Diffusion. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 67959–67985. PMLR, 2025. URL https://proceedings.mlr.press/v267/wu25an.html.

Naganand Yadati, Madhav Nimishakavi, Prateek Yadav, Vikram Nitin, Anand Louis, and Partha Talukdar. HyperGCN: A New Method For Training Graph Convolutional Networks on Hypergraphs. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/1efa39bcaec6f3900149160693694536-Abstract.html.

Yi Yan and Ercan Engin Kuruoglu. Binarized Simplicial Convolutional Neural Networks. *Neural Networks*, 183:106928, 2025. doi: 10.1016/j.neunet.2024.106928. URL https://www.sciencedirect.com/science/article/abs/pii/S0893608024008578.

Zuoyu Yan, Junru Zhou, Liangcai Gao, Zhi Tang, and Muhan Zhang. An Efficient Subgraph GNN with Provable Substructure Counting Power. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3702–3713. ACM, 2024. doi: 10.1145/3637528.3671731. URL https://dl.acm.org/doi/10.1145/3637528.3671731.

Shuaicheng Zhang, Haohui Wang, Junhong Lin, Xiaojie Guo, Yada Zhu, Si Zhang, Dongqi Fu, and Dawei Zhou. HeroFilter: Adaptive Spectral Graph Filter for Varying Heterophilic Relations. In *Advances in Neural Information Processing Systems*, volume 38, 2025a. URL https://openreview.net/forum?id=77Kw7b7EYb.

Yu Zhang, Xin Li, Yaoqun Xu, Xitong Xu, and Zhe Wang. A graph transformer with optimized attention scores for node classification. *Scientific Reports*, 15:30015, 2025b. doi: 10.1038/s41598-025-15551-2. URL https://www.nature.com/articles/s41598-025-15551-2.

Appendix A. Teams and Co-authors

Table 4 associates each participating team with the individual co-authors of this paper who belong to it, and lists the pull requests and awards of each team. Participants who submitted under more than one team appear once per team.

Team	Co-authors	Track	PR numbers	Award
AdamGohain	Adam Gohain	T1	359	
CoPreSheafers	Christine Zeh, Lennart Bastian	T2	418	
CtrlAltMe	Salwa Khattami	T2	397	
DeveshKaushal	Devesh Kaushal	T2	392	
E(n)igma	Gaurav Khanal	T2	320, 362, 391	T2 joint 1st (ETNN-CoordinatePolicy)
Fernando Espinosa	Fernando David Espinosa Iñiguez	T2	410	
Francesco Orsi	Francesco Orsi	T2	393, 394	
GiggiolIGriggio	Alessio Comparini	T2	376	
Grold	Grold Otieno Mboya	T2	350	
howyadoin	Dario Loi, Gabriele Onorato	T1+T2	322, 326, 405, 408, 412	T1 1st (DSNN); HM portfolio
Kajolya	Kolya Lettl, Kajal Singla	T1	407	
kxdenz	Kaden Zheng	T1	419	
LangDiff	Jonas Müller	T1	335, 399, 400, 420	
Lettlini	Kolya Lettl, Ghazal Mostofi	T1	396	
LoopCounter	Sud Athreya	T2	364, 365, 366, 367	
MISN_AY	Mohammad Yusuf, Ankit Kumar	T1+T2	382, 383, 387, 388	
One Ring to Lift Them All	Jorge Luiz Franco	T2	406	HM differentiable liftings
Oversmooth operators	Gabriel Franco, Yelizaveta Falkouskaya	T1+T2	398, 404	
Perelyn Research	Michael Banf, Dominik Filipiak	T1	361	
Prasanna28Devadiga	Prasanna Devadiga	T1	327	
PushparajD	Pushparaj Devadiga	T1	352, 355	
r2	Anier Velasco Sotomayor	T1+T2	329, 354, 360, 414	
Raj Patil	Raj Patil	T1	411	
rayquaza	Luigi Petti, Eric Dolores-Cuenca	T2	413	
Remy	Stefano Blando	T2	417	
s/pairwise/ho	Valentina Sánchez	T2	321	
Sheafu	Ankit Grover, Rémi Bourgerie	T1	373, 374, 375	HM supplementary analysis
SweetLesson	Aarav Gupta, Riyan Patel	T1	345, 353, 402	T1 joint 2nd (HOD-GNN)
TJPaik	Taejin Paik	T1+T2	328, 332, 333, 334	
TopoCoop	Marcos Chávez Mesa, Gabriel Chaviano Díaz	T1+T2	369, 371	
topolab	Chenguang Xu	T1+T2	415, 416	
TripleA	Amirreza Akbari, Alireza Mohammadian	T1+T2	363, 370, 372, 378, 384	T1 joint 2nd (GRIT); T2 joint 1st (TopoU-Net); HM portfolio
weeyev	Vihan Tiwari	T1	421	
yeahkitory	Hisataka Kajita	T1	319	

Table 4: Teams and their co-authors, in alphabetical order of team name. T1 and T2 denote Track 1 (GNN) and Track 2 (TNN); HM denotes an honorable mention (Section 4). A dash marks a team for which no author information was provided.

Appendix B. Full Leaderboards

Tables 5 and 6 give the rankings of the 65 entries included in the comparative analysis, with variant selection and exclusions as described in Section 5. Scores are means over the 12 GraphUniverse settings and three training seeds per setting; rankings use unrounded scores.

#	Model	Track	Team	Score ↑
1	Graph Inductive Bias Transformer (GRIT)	T1 · GNN	TripleA	0.5207
2	Filter Bank GNN (FiGURe)	T1 · GNN	r2	0.4986
3	PolyFilter GNN (Favard)	T1 · GNN	r2	0.4977
4	HiGCN	T2 · TNN	LoopCounter	0.4869
5	GREAD	T1 · GNN	topolab	0.4816
6	TopoPolynormer	T1 · GNN	SweetLesson	0.4798
7	Deep Neural Sheaf Diffusion (DNSD)	T1 · GNN	Sheafu	0.4743
8	ESC-GNN	T1 · GNN	kxdenz	0.4743
9	MPSN	T2 · TNN	GiggioIlGriggio	0.4742
10	SheafTSP	T2 · TNN	Fernando Espinosa	0.4735
11	SheafHyperGNN	T2 · TNN	s/pairwise/ho	0.4728
12	DiffLift-SMCN	T2 · TNN	One Ring to Lift Them All	0.4719
13	Generalized PageRank GNN	T1 · GNN	PushparajD	0.4697
14	MTGCN	T1 · GNN	LangDiff	0.4697
15	CTNN (copresheaf, gated)	T2 · TNN	CoPreSheafers	0.4690
16	HiD-Net	T1 · GNN	LangDiff	0.4685
17	Cell Isomorphism Network (CIN)	T2 · TNN	TripleA	0.4681
18	JacobiConv	T1 · GNN	MISN_AY	0.4649
19	r-ℓMPNN	T1 · GNN	Oversmooth operators	0.4648
20	DPHGNN	T2 · TNN	Oversmooth operators	0.4641
21	DSHN	T2 · TNN	howyadoin	0.4633
22	HMC	T2 · TNN	topolab	0.4596
23	CoGNN	T1 · GNN	TopoCoop	0.4585
24	HeroFilter	T1 · GNN	TripleA	0.4574
25	OGFormer	T1 · GNN	TripleA	0.4574
26	EHNN	T2 · TNN	LoopCounter	0.4566
27	CTNN	T2 · TNN	howyadoin	0.4561
28	GSAN	T2 · TNN	LoopCounter	0.4556
29	Polynormer	T1 · GNN	SweetLesson	0.4550
30	ETNN-LapPE	T2 · TNN	E(n)igma	0.4545
31	ETNN-CoordinatePolicy (structural-LapPE)	T2 · TNN	E(n)igma	0.4543
32	ETNN	T2 · TNN	E(n)igma	0.4534
33	GATE	T1 · GNN	PushparajD	0.4457
34	GFT	T1 · GNN	TJPaik	0.4451
35	WHNN	T2 · TNN	Remy	0.4435
36	DirSNN (Memory-Safe, Harmonic-Aware)	T2 · TNN	Grold	0.4414
37	HyMN	T1 · GNN	weeyev	0.4408
38	Cellular Transformer	T2 · TNN	TopoCoop	0.4400
39	HOD-GNN	T1 · GNN	SweetLesson	0.4395
40	MPSN	T2 · TNN	Francesco Orsi	0.4387
41	SGFormer	T1 · GNN	AdamGohain	0.4375
42	PolynomialFilterTNN (SNN)	T2 · TNN	r2	0.4368
43	COSIMO	T2 · TNN	DeveshKaushal	0.4333

#	Model	Track	Team	Score ↑
44	PhenomNN	T2 · TNN	LoopCounter	0.4308
45	GAMMA	T1 · GNN	Raj Patil	0.4257
46	GAUGE (Riemannian Graph Foundation Model)	T1 · GNN	Lettlini	0.4219
47	Neural Sheaf Diffusion (NSD)	T1 · GNN	TJPaik	0.4113
48	Bi-SCNN	T2 · TNN	MISN_AY	0.4057
49	AdvDIFFormer	T1 · GNN	LangDiff	0.4023
50	AirTNN	T2 · TNN	Francesco Orsi	0.4007
51	GREAD	T1 · GNN	LangDiff	0.3937
52	SheafAN	T1 · GNN	howyadoin	0.3887
53	Neural Sheaf Propagation	T1 · GNN	howyadoin	0.3852
54	HyperGCN	T2 · TNN	rayquaza	0.3847
55	Directed Simplicial Neural Networks (Dir-SNN)	T2 · TNN	TJPaik	0.3821
56	Neural Sheaf Propagation + Holonomy Measurement	T1 · GNN	Sheafu	0.3817
57	FAGCN	T1 · GNN	MISN_AY	0.3813
58	SMCN	T2 · TNN	CtrlAltMe	0.3697
59	TopoU-Net	T2 · TNN	TripleA	0.3666
60	LINKX	T1 · GNN	Perelyn Research	0.3628
61	AirTNN	T2 · TNN	r2	0.3594
62	DSNN	T1 · GNN	howyadoin	0.3081
63	TopNets	T2 · TNN	TJPaik	0.2970
64	Identity Sheaf Network (ISN)	T1 · GNN	Sheafu	0.2755
65	GSN	T1 · GNN	Kajolya	0.0795

Table 5: Community-detection leaderboard for the 65-entry analysis cohort (accuracy, higher is better; Section 5). Ranks use full-precision means.

#	Model	Track	Team	Score ↓
1	CTNN (copresheaf, gated)	T2 · TNN	CoPreSheafers	0.0024
2	SheafTSP	T2 · TNN	Fernando Espinosa	0.0108
3	MPSN	T2 · TNN	GiggioIlGriggio	0.0287
4	MPSN	T2 · TNN	Francesco Orsi	0.0333
5	TopoPolynormer	T1 · GNN	SweetLesson	0.0451
6	COSIMO	T2 · TNN	DeveshKaushal	0.0504
7	ESC-GNN	T1 · GNN	kxdenz	0.0552
8	GSN	T1 · GNN	Kajolya	0.0665
9	Graph Inductive Bias Transformer (GRIT)	T1 · GNN	TripleA	0.0693
10	HOD-GNN	T1 · GNN	SweetLesson	0.0802
11	ETNN-LapPE	T2 · TNN	E(n)igma	0.1147
12	TopNets	T2 · TNN	TJPaik	0.1192
13	DiffLift-SMCN	T2 · TNN	One Ring to Lift Them All	0.1201
14	ETNN-CoordinatePolicy (structural-LapPE)	T2 · TNN	E(n)igma	0.1211
15	ETNN	T2 · TNN	E(n)igma	0.1213
16	r-ℓMPNN	T1 · GNN	Oversmooth operators	0.1382
17	HyMN	T1 · GNN	weeyev	0.1497
18	Bi-SCNN	T2 · TNN	MISN_AY	0.1498
19	DirSNN (Memory-Safe, Harmonic-Aware)	T2 · TNN	Grold	0.1543
20	EHNN	T2 · TNN	LoopCounter	0.1672
21	Directed Simplicial Neural Networks (Dir-SNN)	T2 · TNN	TJPaik	0.1716
22	GSAN	T2 · TNN	LoopCounter	0.2570
23	Cell Isomorphism Network (CIN)	T2 · TNN	TripleA	0.2638
24	DPHGNN	T2 · TNN	Oversmooth operators	0.4021
25	PolynomialFilterTNN (SNN)	T2 · TNN	r2	0.4197
26	AirTNN	T2 · TNN	Francesco Orsi	0.4538
27	HMC	T2 · TNN	topolab	0.4843
28	GFT	T1 · GNN	TJPaik	0.4911
29	Cellular Transformer	T2 · TNN	TopoCoop	0.4982
30	CTNN	T2 · TNN	howyadoin	0.5174
31	Neural Sheaf Diffusion (NSD)	T1 · GNN	TJPaik	0.5378
32	AirTNN	T2 · TNN	r2	0.5401
33	WHNN	T2 · TNN	Remy	0.5492
34	Identity Sheaf Network (ISN)	T1 · GNN	Sheafu	0.6044
35	SheafHyperGNN	T2 · TNN	s/pairwise/ho	0.6742
36	HiGCN	T2 · TNN	LoopCounter	0.8038
37	SMCN	T2 · TNN	CtrlAltMe	0.8219
38	Neural Sheaf Propagation	T1 · GNN	howyadoin	0.8292
39	Filter Bank GNN (FiGURe)	T1 · GNN	r2	0.8396
40	PolyFilter GNN (Favard)	T1 · GNN	r2	0.8425
41	Deep Neural Sheaf Diffusion (DNSD)	T1 · GNN	Sheafu	0.8658
42	Polynormer	T1 · GNN	SweetLesson	0.8721
43	HiD-Net	T1 · GNN	LangDiff	0.8753
44	GAUGE (Riemannian Graph Foundation Model)	T1 · GNN	Lettlini	0.8758
45	GREAD	T1 · GNN	topolab	0.8773
46	DSHN	T2 · TNN	howyadoin	0.9073
47	DSNN	T1 · GNN	howyadoin	0.9099
48	LINKX	T1 · GNN	Perelyn Research	0.9132

#	Model	Track	Team	Score ↓
49	Neural Sheaf Propagation + Holonomy Measurement	T1 · GNN	Sheafu	0.9158
50	MTGCN	T1 · GNN	LangDiff	0.9160
51	GATE	T1 · GNN	PushparajD	0.9223
52	PhenomNN	T2 · TNN	LoopCounter	0.9243
53	CoGNN	T1 · GNN	TopoCoop	0.9286
54	HyperGCN	T2 · TNN	rayquaza	0.9289
55	GAMMA	T1 · GNN	Raj Patil	0.9301
56	GREAD	T1 · GNN	LangDiff	0.9315
57	Generalized PageRank GNN	T1 · GNN	PushparajD	0.9325
58	AdvDIFFormer	T1 · GNN	LangDiff	0.9342
59	SheafAN	T1 · GNN	howyadoin	0.9345
60	JacobiConv	T1 · GNN	MISN_AY	0.9378
61	HeroFilter	T1 · GNN	TripleA	0.9470
62	OGFormer	T1 · GNN	TripleA	0.9509
63	SGFormer	T1 · GNN	AdamGohain	0.9658
64	FAGCN	T1 · GNN	MISN_AY	1.0438
65	TopoU-Net	T2 · TNN	TripleA	1.0560

Table 6: Triangle-counting leaderboard for the 65-entry analysis cohort (MSE per triangle, lower is better; Section 5). Ranks use full-precision means.