

Your Paper Title: A Descriptive Subtitle That Conveys the Core Contribution

First Author*, Second Author[†], Third Author*

*Department of Computer Science, University Name, City, Country

Email: {first.author, third.author}@university.edu

[†]Department of Electrical Engineering, Another University, City, Country

Email: second.author@another.edu

Abstract—State the problem you are addressing in 1–2 sentences. Describe the limitation of existing approaches. Present your proposed method or contribution concisely. Summarize the key experimental findings with specific quantitative results (e.g., “achieves 94.3% accuracy, a 3.7% improvement over the state of the art”). Conclude with the broader significance or implication of your work.

Index Terms—topological data analysis, feature extraction, persistent homology, machine learning, image classification

I. INTRODUCTION

Provide the broad context for your research area. Explain why this field is important and what real-world applications motivate the work. Use citations to establish that this is an active area of research [1].

Narrow the focus to the specific problem you address. Describe what has been done and what remains unsolved. Identify the gap in the literature that your work fills.

Discuss the limitations of current approaches. Be specific and cite relevant work [2], [3]. Explain why these limitations matter in practice.

Introduce your proposed approach. Clearly state your contributions. It is common to use a bulleted list:

The main contributions of this paper are as follows:

- We propose [method name], a novel approach to [problem] that addresses [limitation].
- We introduce [specific technical contribution], which enables [benefit].
- We conduct extensive experiments on [datasets/benchmarks], demonstrating [quantitative improvement] over state-of-the-art methods.

The remainder of this paper is organized as follows. Section II reviews related work. Section III presents the proposed methodology. Section IV describes the experimental setup. Section V discusses the results. Section VI concludes the paper and outlines future work.

II. RELATED WORK

A. First Category of Related Work

Discuss the first category of related approaches. Cite and briefly summarize key papers [4], [5]. Note their strengths and limitations.

[Architecture Diagram Placeholder]
Replace with your figure

Fig. 1: Overview of the proposed architecture. The input data flows through [component 1], [component 2], and produces [output]. Best viewed in color.

B. Second Category of Related Work

Discuss the second category. Compare and contrast approaches within this category [6].

C. Summary and Positioning

Briefly summarize the landscape and position your work relative to existing approaches. Make it clear how your contribution differs from and improves upon prior work.

III. PROPOSED METHODOLOGY

A. Problem Formulation

Let $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ denote the input dataset where each $x_i \in \mathbb{R}^d$. We aim to learn a mapping $f: \mathbb{R}^d \rightarrow \mathbb{R}^k$ such that:

$$\min_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \mathcal{L}(f(x_i), y_i) + \lambda \Omega(f), \quad (1)$$

where $\mathcal{L}$ is the loss function, y_i is the ground truth label, and $\Omega(f)$ is a regularization term with hyperparameter $\lambda > 0$.

B. Proposed Approach

Describe your method in detail. Use subsections to break it into logical components.

As illustrated in Fig. 1, our approach consists of three main components: (1) feature extraction, (2) representation learning, and (3) classification.

Algorithm 1: Proposed Method Name

Input: Dataset $\mathcal{X}$, learning rate η , epochs T **Output:** Trained model parameters θ^*

```

1 Initialize parameters  $\theta_0$  randomly;
2 for  $t = 1$  to  $T$  do
3   Sample mini-batch  $\mathcal{B} \subset \mathcal{X}$ ;
4   Compute loss:  $\ell \leftarrow \frac{1}{|\mathcal{B}|} \sum_{(x,y) \in \mathcal{B}} \mathcal{L}(f_{\theta}(x), y)$ ;
5   Compute gradients:  $g \leftarrow \nabla_{\theta} \ell$ ;
6   Update parameters:  $\theta_t \leftarrow \theta_{t-1} - \eta \cdot g$ ;
7 return  $\theta^* = \theta_T$ ;

```

TABLE I: Summary of Datasets Used in Experiments

Dataset	Samples	Features	Classes
Dataset A	10,000	784	10
Dataset B	50,000	2048	100
Dataset C	5,000	512	5

C. Algorithm

The complete procedure is summarized in Algorithm 1. The time complexity is $O(ndk)$ per epoch, where n is the number of samples, d is the feature dimension, and k is the number of classes.

IV. EXPERIMENTS

A. Datasets

We evaluate our method on three benchmark datasets, summarized in Table I.

B. Baseline Methods

We compare against the following state-of-the-art methods:

- **Method A** [7]: Brief description.
- **Method B** [8]: Brief description.
- **Method C** [9]: Brief description.

C. Implementation Details

All experiments are conducted using Python 3.10 with PyTorch 2.0 on a single NVIDIA RTX 3090 GPU. We use Adam optimizer with learning rate $\eta = 10^{-3}$ and batch size 64. Training runs for 100 epochs with early stopping (patience of 10 epochs). All results are averaged over 5 independent runs with different random seeds.

D. Evaluation Metrics

We report accuracy, precision, recall, and F1-score (macro-averaged). For computational efficiency, we also report training time and inference latency.

V. RESULTS AND DISCUSSION

A. Main Results

Table II presents the main experimental results.

Our method achieves the best performance across all metrics on Dataset A, with a 1.2% improvement in accuracy over the strongest baseline (Method C).

TABLE II: Comparison with State-of-the-Art Methods. Best results in **bold**, second best underlined.

Method	Accuracy	Precision	Recall	F1
Method A	91.2	90.8	91.0	90.9
Method B	92.5	<u>92.1</u>	92.3	92.2
Method C	<u>93.1</u>	91.9	<u>93.0</u>	<u>92.4</u>
Ours	94.3	94.0	94.1	94.0

B. Ablation Study

To understand the contribution of each component, we conduct ablation experiments by systematically removing one component at a time.

C. Qualitative Analysis

Provide visual examples, attention maps, failure cases, or other qualitative evidence that helps readers understand when and why your method works.

D. Computational Cost

Discuss training time, inference speed, memory usage, and model size. Compare with baselines.

VI. CONCLUSION

In this paper, we proposed [method name], a novel approach for [problem]. Our method addresses the limitation of [existing gap] by introducing [key innovation]. Extensive experiments on [number] benchmark datasets demonstrate that our approach outperforms state-of-the-art methods by up to [X]% in [metric].

Limitations. We acknowledge the following limitations: (1) our method assumes [assumption]; (2) scalability to [larger scale] has not been tested.

Future Work. We plan to extend this work by: (1) [direction 1]; (2) [direction 2]; (3) [direction 3].

ACKNOWLEDGMENTS

The authors would like to thank [funding agency] for financial support under grant [number]. We also thank [names] for helpful discussions.

REFERENCES

- [1] A. Author and B. Coauthor, "Title of foundational paper," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, pp. 123–135, Mar. 2020.
- [2] C. Researcher, "Title of paper showing limitation 1," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2021, pp. 456–463.
- [3] D. Scholar and E. Collaborator, "Title of paper showing limitation 2," *Pattern Recognition*, vol. 118, p. 108024, 2021.
- [4] F. Scientist, "Title of related work 1," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2022, pp. 789–798.
- [5] G. Expert and H. Partner, "Title of related work 2," *Neural Networks*, vol. 150, pp. 201–215, 2022.
- [6] I. Pioneer, "Title of related work 3," *arXiv preprint arXiv:2301.00001*, 2023.
- [7] J. Baseline and K. Author, "Baseline method A," in *Proc. Advances Neural Inf. Process. Syst. (NeurIPS)*, 2022, pp. 1234–1245.
- [8] L. Comparison, "Baseline method B," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 5, pp. 2345–2358, 2023.
- [9] M. Reference, "Baseline method C," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2022, pp. 567–578.