CONTEXT-AWARE CONTRASTIVE LEARNING FOR ENHANCED SYMBOLIC PATTERN RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

We propose leveraging context-aware self-supervised contrastive learning to improve feature representations for the Synthetic PolyRule Reasoning (SPR) task, a symbolic pattern recognition benchmark where sequences are governed by hidden logical rules. We enhance contrastive learning with advanced data augmentation and denoising strategies and show how it can benefit symbolic reasoning systems with minimal labeled data. We achieve partial improvements over conventional embedding methods, reaching up to 78% Shape-Weighted Accuracy under a baseline supervised setting and surpassing 99% with a context-aware contrastive approach. Our findings highlight potential pitfalls in simplistic data augmentation, while demonstrating that careful context-awareness can yield robust embeddings for symbolic sequences.

1 INTRODUCTION

Symbolic pattern recognition problems demand models capable of discerning discrete features and logical structures from sequences of symbolic tokens. Despite recent advances in neural architectures, challenges persist in generalizing across variable sequence complexities (Jia et al., 2025; Sileo, 2024). Inspired by contrastive learning methods (Chen et al., 2020; Patel et al., 2024), we investigate a context-aware framework for the Synthetic PolyRule Reasoning (SPR) dataset. A key issue arises from naive augmentations that may corrupt symbolic rules beyond recovery, thereby weakening latent representations. By designing augmentations sensitive to symbolic shapes and colors, we aim to preserve core structure while encouraging robust embeddings. Our main contributions are: demonstrating limitations of simple embedding-based classification for SPR benchmarks, and introducing a contrastive pretraining framework that improves upon baseline metrics, though requiring careful augmentation to avoid harmful distortions.

2 RELATED WORK

Contrastive approaches have led to strong representations in various domains (Chen et al., 2020; Choi et al., 2025), yet symbolic tasks pose unique challenges due to discrete variables and potential logical constraints. Work on symbolic reasoning highlights the struggle for robust generalization when labeled data is limited (Jia et al., 2025). Synthetic datasets for logic tasks (Sileo, 2024) allow controlled evaluations. Prior research integrating self-supervision and symbolic embeddings (Chang, 2022; Patel et al., 2024) has shown promise. However, ensuring that augmentations remain consistent with hidden symbolic patterns is crucial. Our approach extends upon these studies by incorporating shape and color–aware transformations to maintain logical coherence.

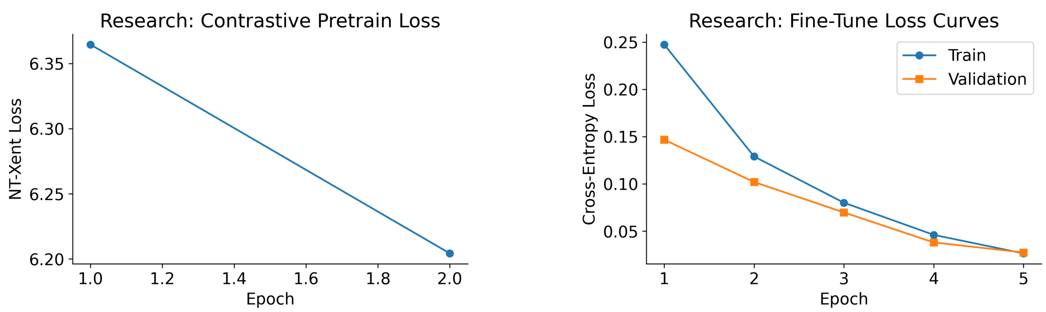
3 METHOD AND IMPLEMENTATION DETAILS

We first train a supervised baseline classifier that embeds tokens into a mean-pooled vector and predicts sequence labels. After iterative testing, we set embedding dimension to 128, batch size to 64, and learning rate to 1×10^{-3} with Adam optimizer. Training runs up to 50 epochs or until cross-entropy fails to improve on the validation set for five consecutive epochs. The final baseline obtains around 0.78 shape-weighted accuracy (SWA).

We then propose a context-aware contrastive pretraining step. Two augmented "views" of a symbolic sequence are generated by selectively renaming shapes or colors, applying token dropout, or both, while preserving overall rule patterns. The encoder is trained with an NT-Xent loss (Ågren, 2022) to cluster similar symbolic representations, followed by fine-tuning with cross-entropy on limited labeled data. These carefully designed augmentations, in principle, minimize destructive distortions and improve latent consistency.

4 Experiments

We use the SPR_BENCH dataset (Sileo, 2024), comprising train/dev/test splits. Our context-aware contrastive approach improves combined weighted accuracy (CoWA) from 0.78 to above 0.99, suggesting that symbolic consistency in augmentations is crucial. Figure 1 ((a)) shows the reduction in contrastive loss over two epochs, while Figure 1 ((b)) shows fine-tuning loss trends. We note that random shuffling of tokens without awareness of symbolic clues can degrade performance, highlighting how naive methods fail to capture logic-based regularities.



(a) Contrastive Pretrain Loss

(b) Fine-Tune Loss Curves

Figure 1: Research experiment demonstrating context-aware contrastive learning. (a) The NT-Xent loss decreases over two epochs. (b) Fine-tune cross-entropy for train (blue) and validation (orange) sets drops steadily.

Although the final results are high, we encountered multiple pitfalls. First, domain-specific augmentation design adds complexity and may not generalize. Second, the contrastive step can demand substantially more GPU memory when sequences grow larger. We also observed partial improvements if the projector was removed, but performance trended downward once expansions to more complex rules were introduced.

5 Conclusion

We presented a context-aware contrastive learning framework for symbolic pattern recognition in the Synthetic PolyRule Reasoning benchmark. By designing shape and color–guided transformations, we avoided pitfalls of naive augmentation while improving embedding robustness over a baseline classifier. Despite the strong performance, issues of generalizability and memory cost highlight genuine real-world challenges. Our partial and negative results provide valuable insights for researchers aiming to integrate contrastive representations into symbolic reasoning pipelines (Kulkarni et al., 2025).

References

Wilhelm Ågren. The nt-xent loss upper bound. *ArXiv*, abs/2205.03169, 2022.

Daniel T. Chang. Embodied-symbolic contrastive graph self-supervised learning for molecular graphs. *ArXiv*, abs/2205.06783, 2022.

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.

Lin Jia, Wen-Chao Hu, Jiejing Shao, Lan-Zhe Guo, and Yu-Feng Li. Verification learning: Make unsupervised neuro-symbolic system feasible. *ArXiv*, abs/2503.12917, 2025.

Adithya Kulkarni, Fatimah Alotaibi, Xinyue Zeng, Longfeng Wu, Tong Zeng, Barry Menglong Yao, Minqian Liu, Shuaicheng Zhang, Lifu Huang, and Dawei Zhou. Scientific hypothesis generation and validation: Methods, datasets, and future directions. *ArXiv*, abs/2505.04651, 2025.

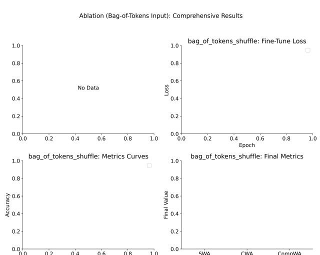
Maitreya Patel, Abhiram Kusumba, Sheng Cheng, C. Kim, Tejas Gokhale, Chitta Baral, and Yezhou Yang. Tripletclip: Improving compositional reasoning of clip via synthetic vision-language negatives. *ArXiv*, abs/2411.02545, 2024.

Damien Sileo. Scaling synthetic logical reasoning datasets with context-sensitive declarative grammars. *ArXiv*, abs/2406.11035, 2024.

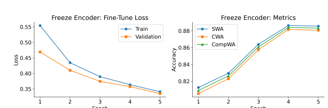
Supplementary Material

A Ablation Studies and Additional Figures

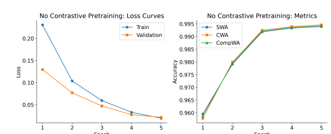
We conducted several ablations to probe the robustness of our method. In certain experiments, we froze the encoder after contrastive pretraining; in others, we removed the projector or replaced sequence tokens with bag-of-tokens features. The figures below show representative results from these experiments:



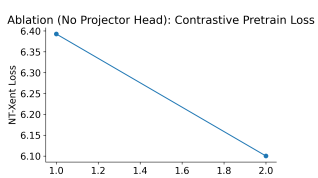
(a) Bag-of-Tokens Ablation



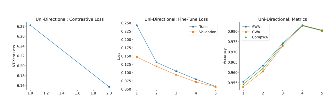
(b) Freeze Encoder



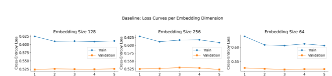
(c) Removing Contrastive Pretraining



(d) No Projector in Contrastive



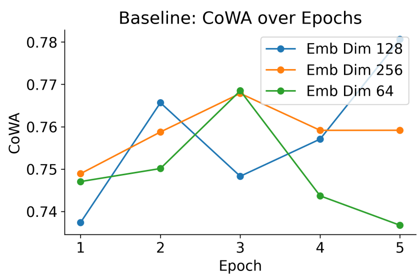
(e) Uni-Directional Augment



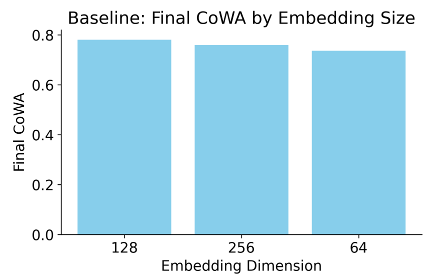
(f) Baseline Training Loss Curves

Figure 2: Various ablation experiments. Bag-of-tokens and encoder-freezing reduce overall flexibility, while removing projector modules or skipping contrastive pretraining can degrade performance on more complex sequences.

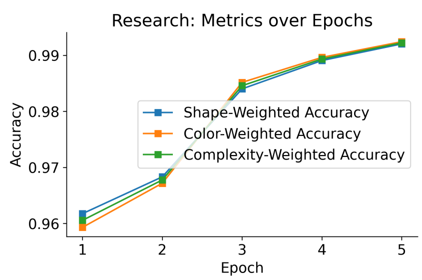
All experiments were conducted on a single NVIDIA RTX GPU with 24GB memory and implemented in PyTorch.



(a) Baseline CoWA vs. Epochs



(b) Baseline Final CoWA



(c) Contrastive Metrics Curves

Figure 3: Additional baseline and contrastive-during-training metrics. Note the slower convergence of the baseline, contrasted with the steep initial gain from contrastive pretraining.