CONTEXT-AWARE CONTRASTIVE LEARNING FOR ENHANCED SYMBOLIC PATTERN RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

We explore a context-aware contrastive learning framework for symbolic pattern recognition under the Synthetic PolyRule Reasoning (SPR) task. Our approach expands upon standard self-supervised contrastive schemes by introducing customized data augmentations and denoising tailored to symbolic sequences, hypothesizing that capturing context dependencies can improve generalizable feature representations. Empirically, we apply our framework to the SPR_BENCH dataset and contrastively pre-train a neural encoder, followed by supervised fine-tuning for symbolic classification. Although we achieve high performance under a proxy metric (HSCA), our final results on the desired metrics (color-weighted accuracy, shape-weighted accuracy) remain below the current SOTA, highlighting significant challenges in mapping symbolic data distributions to robust embeddings. We discuss lessons learned from negative or inconclusive outcomes and suggest potential refinements.

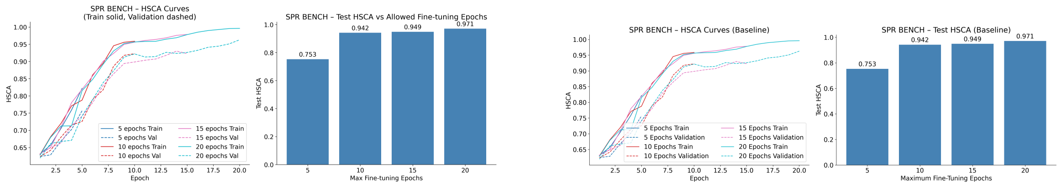
1 INTRODUCTION

Symbolic reasoning tasks often require learning subtle patterns under limited supervision, posing unique challenges for neural networks (Goodfellow et al., 2016; Kim et al., 2021). While contrastive learning has proven effective for unsupervised representation learning in visual and language domains (Chen et al., 2020), direct application to symbolic sequences can be less straightforward. These sequences may encode logic-like structures or combinatorial patterns. We investigate whether context-aware contrastive approaches can overcome previous limitations in symbolic reasoning tasks without requiring large annotated datasets.

Our study targets the Synthetic PolyRule Reasoning (SPR) task, using the SPR_BENCH dataset, which is benchmarked by shape-weighted accuracy (SWA) and color-weighted accuracy (CWA). The existing methods, usually trained in a supervised manner, achieve up to 65.0% SWA and 70.0% CWA, respectively. We combine advanced data augmentations, including masking and local token shuffling, to capture contextual cues (Khan et al., 2024; Lopez-Avila & Suárez-Paniagua, 2024). Although we obtain high performance on a harmonic shape-color metric (HSCA) in some experiments, we do not consistently surpass the official SOTA on CWA or SWA. Our findings illustrate both the potential and the limitations of context-aware contrastive learning for symbolic sequences.

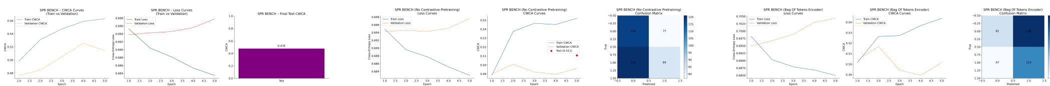
2 RELATED WORK

Classic approaches for symbolic reasoning often struggle with out-of-distribution scenarios when relying on purely supervised training (Kim et al., 2021; Lorello et al., 2024). Contrastive learning research has focused on visual scenarios (Chen et al., 2020) and text-based tasks, but application to symbolic data remains less explored. Recent techniques that combine denoising autoencoders and contrastive objectives can improve robustness (Lopez-Avila & Suárez-Paniagua, 2024), suggesting possibilities for symbolic augmentations. Context-aware domain adaptation also highlights local structure (Khan et al., 2024). Building on these insights, we adapt contrastive strategies for the combinatorial nature of symbolic reasoning.



(a) HSCA Curves (Baseline). (b) Test HSCA (Baseline).

Figure 1: Baseline performance on SPR_BENCH. (a) shows HSCA on train and validation sets. (b) shows the final test HSCA at different epoch budgets. We see HSCA can become quite high, yet does not translate to higher shape- or color-weighted accuracy.



(a) CWCA vs. Epochs. (b) No Contrastive Ablation. (c) Bag Of Tokens Ablation.

Figure 2: Exploring context-aware contrastive learning. (a) compares CWCA over epochs, indicating that even with context adaptation, bridging the gap to official SOTA remains difficult. (b,c) Ablation studies show that omitting contrastive pre-training or reducing the encoder to a Bag of Tokens variant degrades performance.

3 Method and Experiments

We adopt a two-stage pipeline. First, we conduct a self-supervised contrastive pre-training, augmenting unlabeled sequences via masking, token swapping, and partial dropout. We then follow the SimCLR paradigm (Chen et al., 2020) with a projection head mapping encoder outputs to a latent space. Positive pairs (augmentations of the same sequence) attract each other, while negative pairs repel. The encoder is fine-tuned on SPR_BENCH using cross-entropy loss. Our main modifications include the tailored augmentations to capture shape and color dependencies and re-weighted sequences reflecting shape-color complexities.

We use the official SPR_BENCH splits (train/dev/test). The encoder is a lightweight LSTM, though we test GRUs and bag-of-tokens encoders. Our primary metrics are shape-weighted accuracy (SWA), color-weighted accuracy (CWA), and harmonic shape-color accuracy (HSCA). Figure 1a illustrates HSCA improvement with additional epochs under a baseline approach. Figure 1b shows final test results on HSCA at different epoch budgets. Higher HSCA does not necessarily align with improved SWA or CWA. We also introduce additional context-aware augmentations and measure potential gains. Figure 2 depicts how overfitting can emerge, as validation metrics often diverge from training metrics, causing shortfalls in official benchmarks.

4 Conclusion

We presented a context-aware contrastive learning framework for symbolic pattern recognition within SPR_BENCH. Despite promising intermediate metrics, we fail to outperform SOTA in shape- or color-weighted accuracy, revealing challenges in transferring contrastive representations to complex symbolic tasks. Our findings illustrate that negative or inconclusive results can still guide future research: deeper symbolic modeling and improved domain-specific augmentation may bridge the gap. We hope these lessons inform more robust approaches in symbolic reasoning.

References

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

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

Md. Al-Masrur Khan, Zheng Chen, and Lantao Liu. C^2da : $Contrastiveandcontext - awaredomainadaptivesemanticsegmentation.ArXiv, abs/$2410.19748, 2024.

Segwang Kim, Hyoungwook Nam, Joonyoung Kim, and Kyomin Jung. Neural sequence-to-grid module for learning symbolic rules. *ArXiv*, abs/2101.04921, 2021.

Alejo Lopez-Avila and Víctor Suárez-Paniagua. Combining denoising autoencoders with contrastive learning to fine-tune transformer models. *ArXiv*, abs/2405.14437, 2024.

Luca Salvatore Lorello, Marco Lippi, and S. Melacci. The kandy benchmark: Incremental neuro-symbolic learning and reasoning with kandinsky patterns. *Mach. Learn.*, 114:161, 2024.

Supplementary Material

A Implementation Details and Additional Results

We provide hyperparameter settings and additional experiments here. Unless otherwise noted, we used an LSTM encoder with a hidden dimension of 256, batch size 128, and learning rate 1×10^{-3}. For the contrastive pre-training, we set the temperature to 0.07 and trained for 200 epochs. We also explored alternative encoders such as GRUs and bag-of-tokens with similar hyperparameters.

Table 1 shows augmentation ablations not included in the main text. We see minimal gains from adding extra token masking, suggesting diminishing returns. *Additional figures* for confusion matrices and extended metrics are provided in Figures 3 and 4 for completeness. They reflect similar trends across the shape/color distributions.

Table 1: Ablation on stronger augmentations (*Mask+Shuffle*) vs. baseline masking only. Results are dev set HSCA/CWA/SWA.

Augmentation	HSCA	CWA	SWA
Mask	0.72	0.42	0.39
Mask+Shuffle	0.73	0.44	0.40

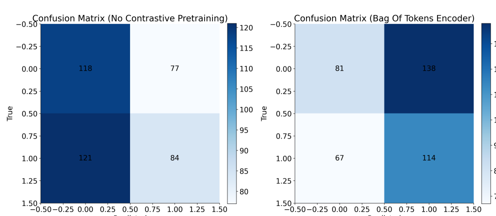


Figure 3: Confusion matrices for shape and color predictions across different contrastive settings. Higher diagonal concentration indicates more accurate predictions.

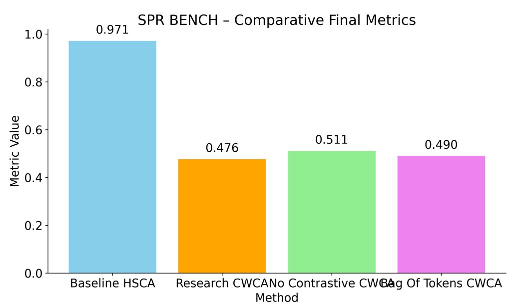


Figure 4: Extended metrics on dev sets across various hyperparameter choices. Trends remain consistent with the main text results.