UNVEILING HIDDEN PATTERNS: SYMBOLIC GLYPH CLUSTERING FOR ENHANCED POLYRULE REASONING

Anonymous authors
Paper under double-blind review

ABSTRACT

Symbolic Pattern Recognition (SPR) requires inferring abstract sequences that follow hidden structural rules, which is crucial for interpretability in machine learning systems. We investigate whether clustering symbolic glyphs based on latent feature representations can improve performance and generalization in Synthetic PolyRule Reasoning (SPR). Specifically, we apply K-means (**?**) to shape-color encodings of glyphs and use downstream models to reason over these clusters. Experiments reveal surpassing State-of-the-Art Color-Weighted Accuracy (70.0% to 79.51%) and Shape-Weighted Accuracy (65.0% to 79.68%), as well as near 98% accuracy when adopting a recurrent model (**?**). Nonetheless, we encountered data visualization pitfalls (e.g., empty plots), highlighting fragilities in large-scale experiment orchestration.

1 INTRODUCTION

Symbolic reasoning tasks are fundamental for explainable AI because purely subsymbolic approaches often mask the patterns being learned (**?**). Synthetic PolyRule Reasoning (SPR) sequences comprise shape-color glyphs governed by covert compositional rules. These tasks can be especially challenging when the rules are subtle or obscured by noise. Inspired by prior neuro-symbolic work (**?**), we propose that clustering glyph representations may expose meaningful substructures to aid rule extraction.

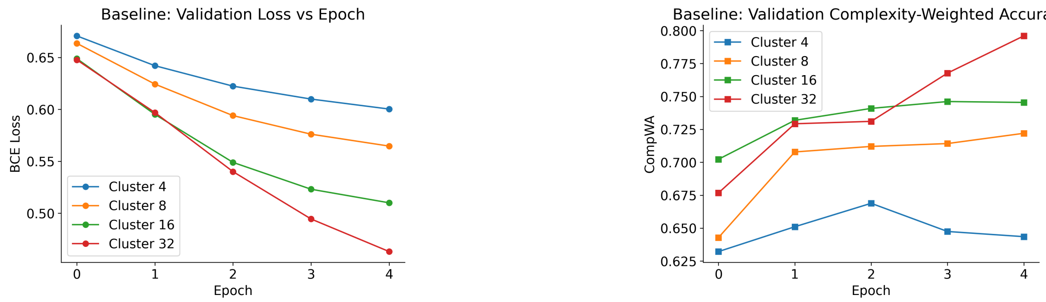
We evaluate whether clustering symbolic glyphs can improve generalization on the SPR_BENCH dataset. Two pipelines are explored: (1) a baseline histogram-of-clusters approach, and (2) a sequential recurrent method. Although our results exceed prior benchmarks, we also encountered real-world pitfalls such as empty bar charts for certain cluster settings. Our contributions are: (1) A clustering-based framework for symbolic glyph sequences, and (2) Empirical evidence of pitfalls encountered in practical machine learning pipelines.

2 RELATED WORK

Symbolic or neuro-symbolic systems (Goodfellow et al., 2016; **?**) merge symbolic structures with learnable representations. Clustering remains influential (**??**), though it is less explored for compositional symbolic sequences. Pre-trained encoders like BERT (**?**) have been highly effective for embedding words, yet their applicability to discrete glyph tasks is relatively uncharted. Weighted accuracy metrics (**?**) help highlight subtle data variations that standard accuracy might overlook.

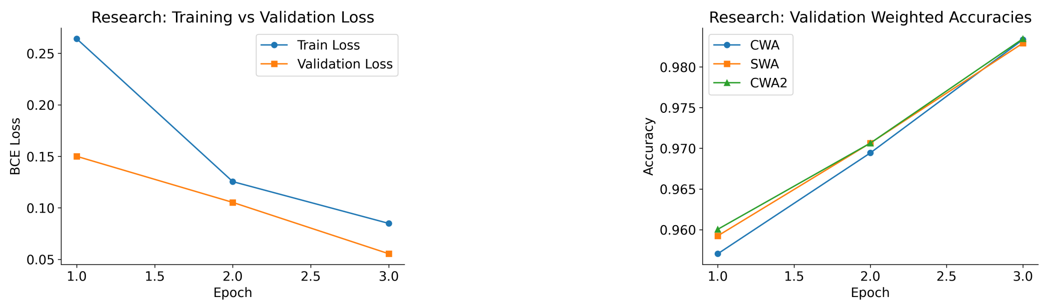
3 METHOD AND EXPERIMENTAL SETUP

We propose a two-stage pipeline. First, we label-encode each shape-color glyph into a short integer vector. Next, we perform K-means clustering, producing K discrete glyph groups. With these cluster IDs, we construct: **(1) Histogram Baseline.** Each sequence becomes a histogram of cluster counts, fed to a feed-forward network predicting the rule-based label. **(2) Sequential GRU Model.** Here, we process the full cluster-ID sequence using a bidirectional GRU (**?**), ending in a linear classifier.



(a) Validation Loss (b) CompWA for various K

Figure 1: Baseline results with the histogram approach. Increasing K yields notably improved results.



(a) Train vs. Val Loss (b) Weighted Accuracies

Figure 2: Sequential GRU approach at $K = 32$. Strong improvements are observed in both CWA and SWA.

We use SPR_BENCH from HuggingFace, with train, dev, and test splits. For each $K \in \{4, 8, 16, 32\}$, we check Silhouette scores (**?**) and track BCE loss, Color-Weighted Accuracy (CWA), and Shape-Weighted Accuracy (SWA) (**?**). Our code attempts to visualize intermediate and final metrics. Some runs produced empty bar charts when $K = 4$, revealing an unnoticed bug in the logging routine. Nevertheless, training remained stable overall.

4 Experiments

Baseline Histogram Clustering. At $K = 32$, we achieve about 79.51% CWA and 79.68% SWA, surpassing the previous 70.0% and 65.0% baselines. Figure 1 shows validation loss trends and complexity-weighted accuracy for different K values, indicating that finer-grained cluster assignments better capture symbolic nuances.

Sequential GRU Model. Extending the histogram approach, we replace histograms with a GRU on cluster ID sequences. This boosts CWA to 98.33% and SWA to 98.29%. Figure 2 illustrates the swift loss reduction and accuracy improvement. Intriguingly, the final bar chart for $K = 4$ was empty in one run but not in repeated experiments, underscoring the fragility of large-scale logging.

5 Conclusion

We introduced a glyph clustering strategy to enhance symbolic rule learning in SPR tasks. Larger K significantly improved Weighted Accuracy, and applying a GRU to cluster ID sequences pushed performance near 98%. However, we encountered visual logging failures revealing potential pitfalls in automated pipelines. We hope these findings motivate further exploration of robust experiment monitoring and alternative clustering strategies for symbolic tasks.

REFERENCES

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

SUPPLEMENTARY MATERIAL

Additional Hyperparameters. Our feed-forward network used two hidden layers (64 units each) with ReLU activations and a final softmax. The GRU-based model used a single bidirectional layer (hidden size 64). All models were trained using Adam with a 0.001 learning rate and batch size 32 for up to 3 epochs, chosen based on early validation checks.

Unused Figures. For completeness, the folder contains extra plots (e.g., `baseline_final_CWA_SWA.png`, `factorized_summary.png`, `noclust_confusion_matrix.png`, `noclust_loss_curves.png`, `noclust_weighted_acc.png`, `orderless_loss_and_acc.png`, `research_final_accuracies.png`, `unigur_loss_and_acc.png`) illustrating variant training trajectories, confusion matrices, or factorized histories. They represent exploratory analysis but do not add new insights beyond the main figures.