

Confusion Matrix Evaluation and Autoregressive Rollout Plan

Zhang Yiran

November 25, 2025

Overview

This report summarizes two new analyses extending last week's results on local-neighborhood prediction in Conway's Game of Life:

- a detailed evaluation of confusion matrices for the 3×3 and 5×5 MLP models, assessing whether high accuracy arises from genuine pattern recognition rather than majority-class bias;
- a planned autoregressive rollout experiment to test whether the learned spatial predictors implicitly approximate the underlying Game-of-Life transition rule.

All model architectures, training protocols, and dataset-generation details remain unchanged from the previous report and are omitted here.

1 Confusion Matrix Analysis

To determine whether the MLP predictions truly reflect the structure of local microenvironments, we computed confusion matrices for all (regime, density) configurations, aggregated over three seeds. Two summary figures are provided below: one for the 3×3 model and one for the 5×5 model.

Across all settings, both models maintain a healthy balance between true positives (TP) and true negatives (TN), and false positives (FP) and false negatives (FN) remain comparatively small. Crucially, no configuration collapses into trivial behavior such as predicting all-dead or all-alive. This indicates that the high accuracies reported last week are not the result of majority-class shortcuts.

The larger 5×5 neighborhood exhibits consistently lower FN counts than 3×3 , especially in the early and mid regimes. This suggests improved sensitivity to living-cell patterns and matches the accuracy gains previously observed.

Confusion Matrices — MLP with 3x3 Neighborhood

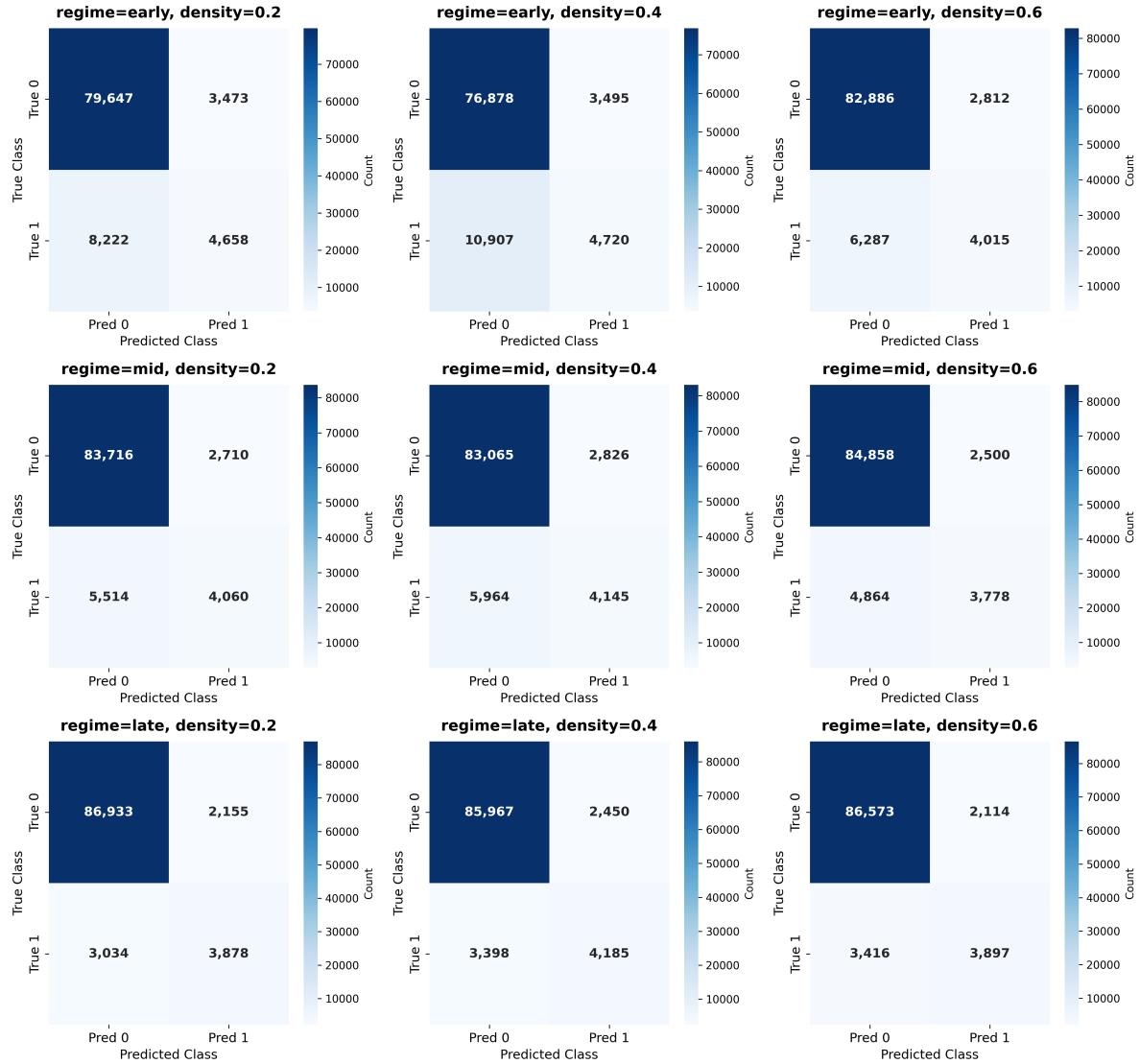


Figure 1: Confusion matrices for the 3×3 MLP model across all (regime, density) combinations, aggregated over three seeds.

Confusion Matrices — MLP with 5×5 Neighborhood

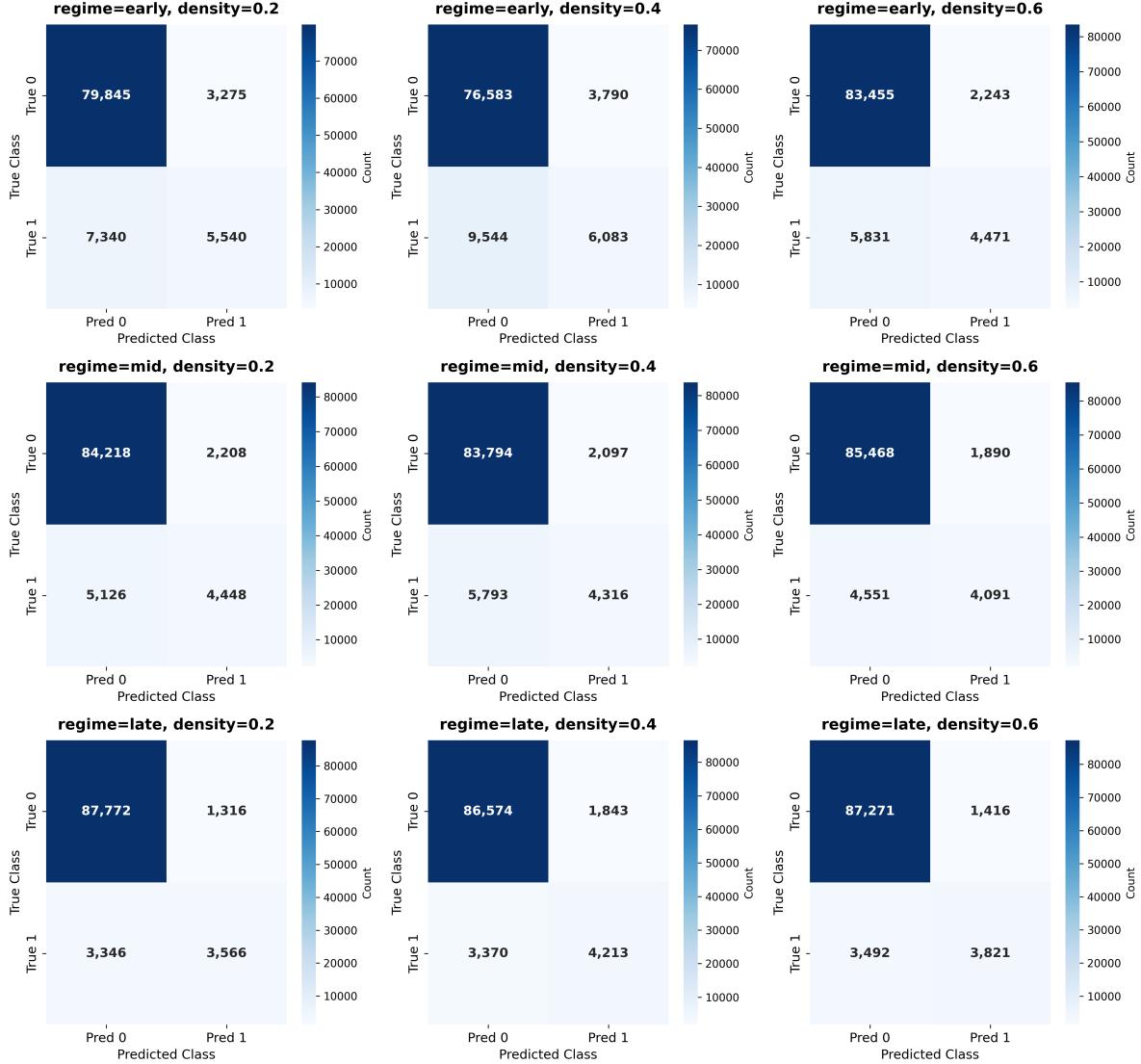


Figure 2: Confusion matrices for the 5×5 MLP model across all (regime, density) combinations. Lower FN rates compared to Figure 1 highlight the benefit of enlarged local context.

Overall, the confusion matrix analysis confirms that:

- model performance is not dominated by the majority class;
- 5×5 neighborhoods provide more informative spatial context than 3×3 ;
- improvements in accuracy reflect genuine gains in sensitivity to alive cells.

2 Autoregressive Rollout (In Progress)

Although the MLP models were trained without any temporal information and only to predict the *current* state of the center cell, they may implicitly approximate the local transition rule of the Game of Life. To evaluate this possibility, we will perform an autoregressive rollout:

1. Starting from a full board state S_t , predict the next board \hat{S}_{t+1} using the trained MLP applied at every cell.
2. Iteratively feed predictions back into the model to generate $\hat{S}_{t+2}, \hat{S}_{t+3}, \dots$
3. Compare the predicted trajectories to the true Game-of-Life evolution, examining stability of static structures, oscillators, and moving patterns.

Results and rollout visualizations will be included in next week's report.

Summary

This week's analysis confirms that the MLP models genuinely learn spatial microenvironment features rather than exploiting class imbalance. The planned autoregressive rollout will further test whether such local predictors capture aspects of the underlying dynamics, and will serve as a baseline for upcoming GeoGNN experiments.