ST5-base Model Performance Across Different Evaluation Scenarios
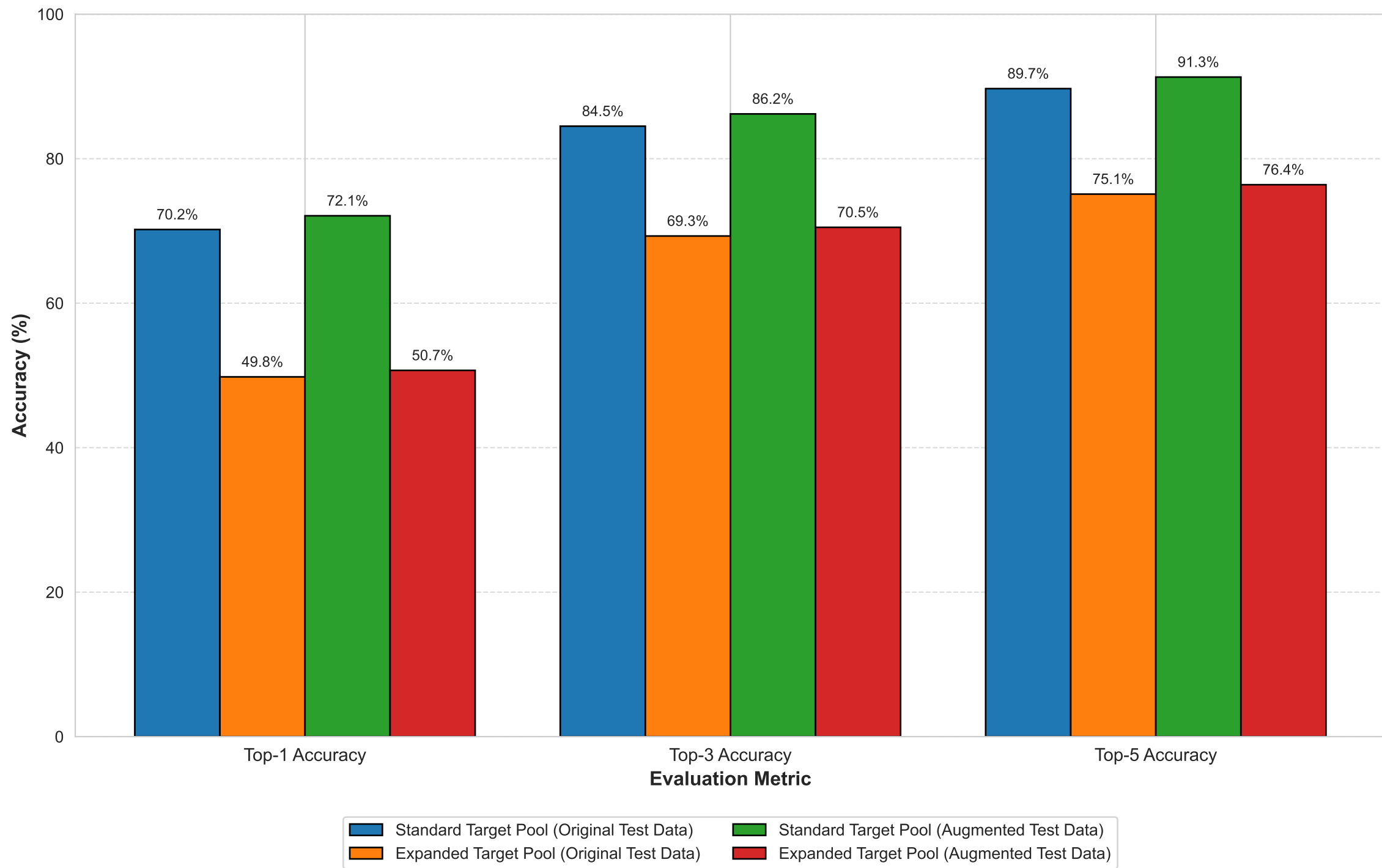
Data from reproduced ST5-base model (5-fold cross-validation mean)