Fuzzlang: Generating Compilation Errors to Teach ML Code Fixes

Baodi Shan[1,2], Johannes Doerfert[2]

[1]Stony Brook University [2]Lawrence Livermore National Laboratory

Contact Email

Abstract

Fuzzlang, a Clang Python wrapper, generates diverse compilation errors by modifying commands or source code, collecting error messages, code and AST. Its dataset helps ML models improve performance in code correction tasks.
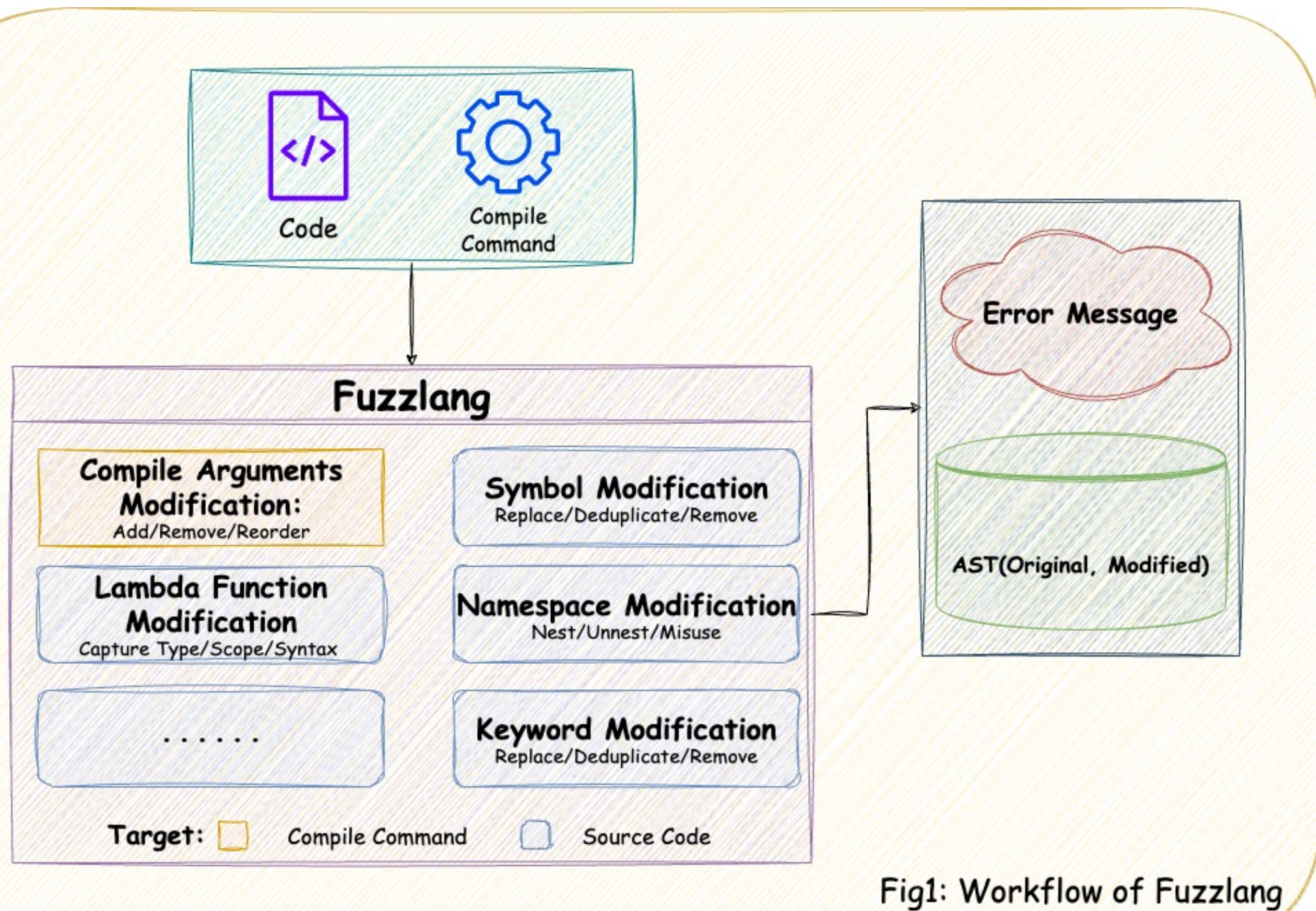
Overview of Fuzzlang

Fig1: Workflow of Fuzzlang

Fuzzlang integrates into C/C++ projects by replacing the standard compiler and injecting fuzzing errors into commands or source code. It collects error data and AST information during compilation and compares AST changes for successful builds.

Dataset & Evaluation

Using Fuzzlang to compile LLVM, we generated a Fuzzlang-LLVM dataset with over 77,000 errors from 3,000 files.
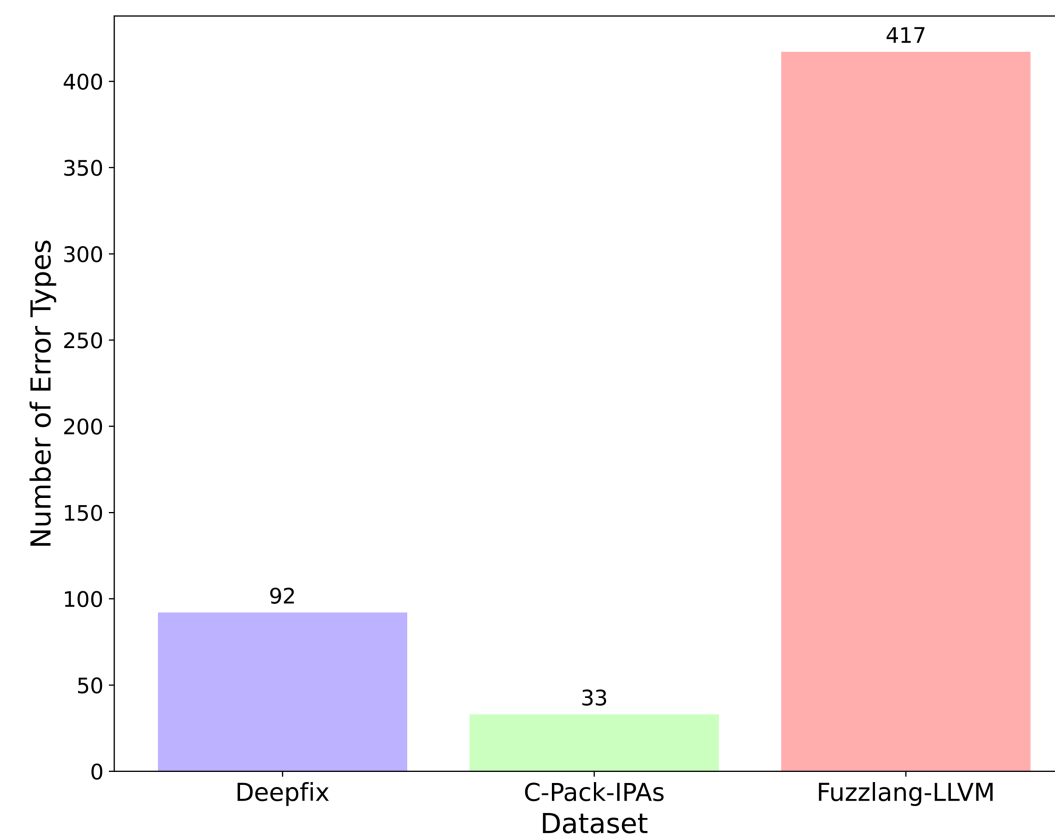

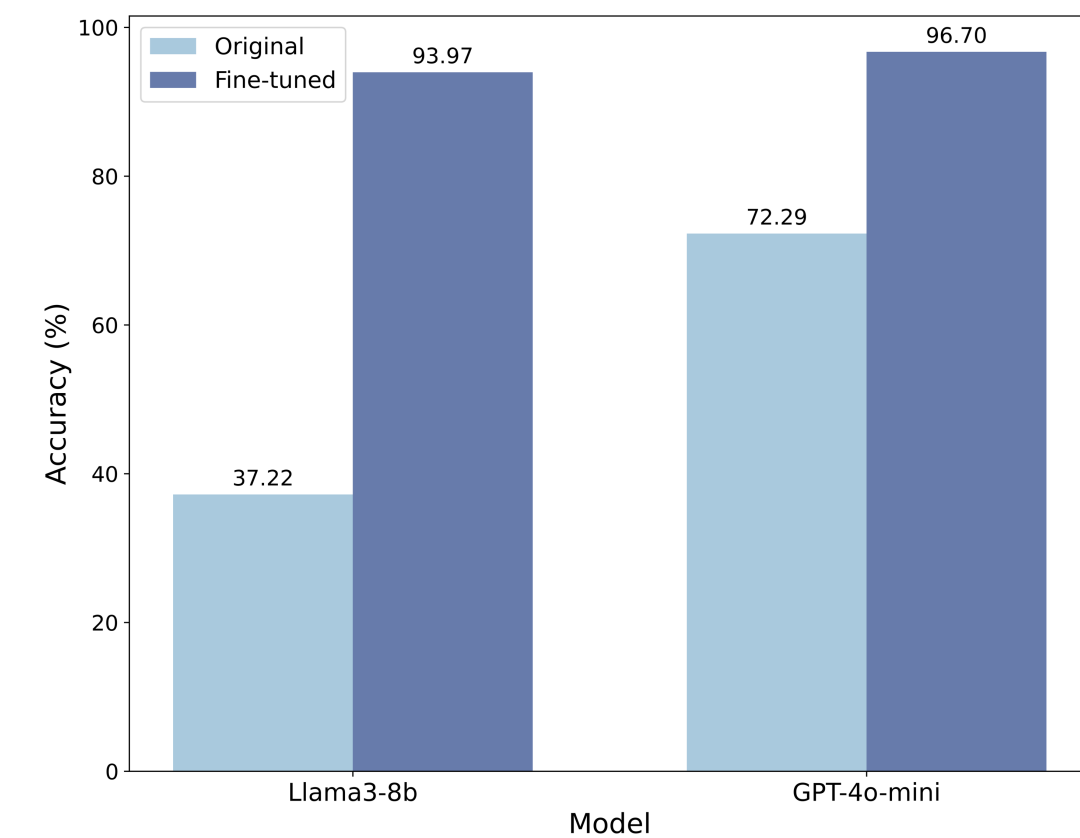Fig 2: The number of error types across datasets

Fig3: Accuracy of LLMs

After fine-tuning on the training dataset, both LLMs demonstrated significant improvements in their code correction capabilities.
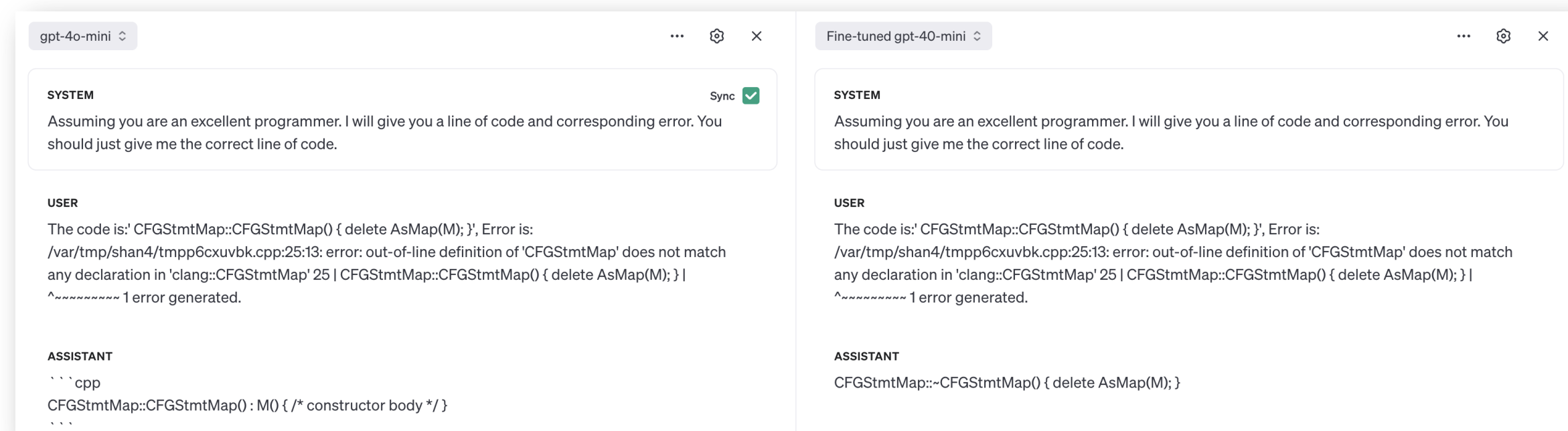

Fig4: Use case comparison between non-fine-tuned and fine-tuned GPT-4o-mini