

# 先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	ソフトウェア設計学 (飯田 元 (教授))		
学籍番号	2111030	提出日	令和 5年 1月 19日
学生氏名	井原 輝人		
論文題目	データの依存関係を考慮した深層学習モデルに基づく自動コードレビュー手法		
要旨			
<p>コードレビューは、開発者がソースコードを精査して、バグや可読性の問題を抽出する品質管理プロセスである。現行のコードレビューは人手による精査作業であり、膨大な労力を要すると指摘されている。コードレビューに要する労力を低減するため、開発者が作成したソースコードに対して、レビューコメントや修正案を自動で生成する自動コードレビュー技術が注目を集めており、既存研究では様々な深層学習モデルが応用されている。</p> <p>しかしながら、既存研究のほとんどが変数名や関数名の文字列情報に基づいた深層学習モデルを用いており、データの処理手順を考慮していない。コードレビューではデータの処理手順の正しさが重要な指摘対象であるが、変数名や関数名の文字列情報のみでは正しい指摘を予測できない可能性がある。</p> <p>本研究では、データの依存関係を捉えることに長けたGraphCodeBERTを応用し、データの処理手順を考慮したコードレビュー自動化の実現を目指す。評価実験では、オープンソースソフトウェアから収集されたレビューデータセットを用いて、従来手法で用いられたTransformerおよびCodeT5との比較評価を行った。評価実験の結果、モデル出力と正解ラベルの完全一致率(EMスコア)が従来手法と比較し、5.58%向上することを確認した。</p>			