Automatic Code Generation from Graphical User Interface (GUI) Images

Sabthavi. J

219395V

Master of Science in Computer Science

Department of Computer Science & Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

March 2024

Automatic Code Generation from Graphical User Interface (GUI) Images

Sabthavi. J

219395V

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science

Department of Computer Science & Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

March 2024

DECLARATION

I declare that this is my own work, and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the master's thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof. Indika Perera

Signature of the Supervisor: Date:

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to my project supervisor, Prof. Indika Perera, from the Department of Computer Science & Engineering, Faculty of Engineering, University of Moratuwa, for his invaluable support, guidance, and encouragement throughout the project.

I also take this opportunity to convey my special thanks to all my friends, batch mates and colleagues who supported and encouraged me with their best wishes. Finally, I extend my heartfelt gratitude to my parents and family who supported me in numerous ways, contributing significantly to the success of this project.

ABSTRACT

In today's software landscape, Graphical User Interfaces (GUIs) are crucial for improving user experience. As the demand for user-friendly software applications grows, the development of GUI-based software becomes more complex. This rapid growth of GUIs demands efficient methods for translating design representations into functional code. This thesis explores the novel approach of employing an Image Captioning model to automatically generate source code from Graphical User Interface (GUI) images. The openly available Pix2Code dataset, which encompasses screenshots of GUIs and their corresponding DSL code for Android, iOS, and web platforms, is utilized. The proposed model employs a standard encoder-decoder architecture. It integrates a ResNet152 model as the image encoder and an LSTM model as the text decoder to convert GUI images into human-readable DSL code sequences. The selection of the ResNet152 model is a strategic choice driven by its exceptional depth and complexity. The architectural design of ResNet152 enables it to adeptly capture the complex visual characteristics found within GUI images, making it well-suited for the task of code synthesis. GUI images often contain complex design elements such as buttons, text fields, and menus, which require a nuanced understanding to accurately translate into code. The depth of ResNet152 enables it to capture these intricate details effectively, contributing to the precision and accuracy of the code generation process. This thesis also investigates the choice of decoding algorithm, opting for a greedy approach over alternatives like beam search. This decision is informed by the need for simplicity and efficiency in the code generation process. Model's performance is assessed using BLEU scores, demonstrating a noteworthy achievement with a score of 0.78. This thesis provides an in-depth exploration of the technical aspects, including model architecture and training processes, along with insights into encountered challenges. This study significantly contributes to the field of automated code generation by introducing an efficient method for translating GUI images into executable code. The implications of this work extend to streamlining software development processes, reducing manual coding efforts, and enhancing overall productivity. The experimental results showcase the model's proficiency in capturing complex GUI designs and generating accurate source code snippets, paving the way for potential applications in GUI-based software development methodologies. This study contributes to the evolving field of automatic code generation and suggests possibilities for future enhancements in code generation for various platforms.

Keywords: Graphical User Interfaces (GUIs), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Domain Specific Language (DSL), Bilingual Evaluation Understudy (BLEU)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
GUI	Graphical User Interface
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
DSL	Domain Specific Language
BLEU	Bilingual Evaluation Understudy
NLP	Natural Language Processing
PIL	Python Imaging Library
HTML	Hyper Text Markup Language
RQ	Research Question