



**សាកលវិទ្យាល័យភូមិន្ទភ្នំពេញ**

**ROYAL UNIVERSITY OF PHNOM PENH**

**ប្រធានបទនិក្ខេបបទ**

**A novel approach to Khmer Handwritten Text  
Recognition using TrOCR Architecture**

**Mr. Vitou Soy**

**A Thesis**

**In Partial Fulfilment of the Requirement for the Degree of  
Bachelor of Engineering in Information-Technology-Engineering  
Academic Year 2022 - 2025**

**Examination committee: Mr. Sokchea Kor (Advisor)  
Mr. Chanpiseth Chap (committee)  
Mrs. Daly Chea (committee)  
Dr. ....**

## មូលន័យសង្ខេប

ការស្រាវជ្រាវនេះផ្ដោតលើវិធីសាស្ត្រថ្មីសម្រាប់ការទទួលស្គាល់អត្ថបទសរសេរដោយ  
ដៃគាតាខ្មែរ។ វាប្រើបច្ចេកវិទ្យា TrOCR ដែលបានបណ្តុះបណ្តាលជាមុនលើទិន្នន័យ  
ពហុភាសាថ្នាក់ធំ ហើយបន្តដំណើរការបណ្តុះបណ្តាលបន្ថែមជាផ្នែកបន្ថែម។  
ទិន្នន័យសម្រាប់ការបណ្តុះបណ្តាលត្រូវបានរៀបចំយ៉ាងប្រុងប្រយ័ត្ន  
ដើម្បីអាចផ្គូផ្គងនឹងលក្ខណៈពិសេសនៃអក្សរខ្មែរ។ លទ្ធផលនៃការសាកល្បងបង្ហាញថា  
TrOCR អាចផ្តល់នូវលទ្ធភាពស្គាល់អក្សរដែលមានតម្លៃខ្ពស់ជាងវិធីបុរាណ។

# Abstract

A novel approach to Khmer handwritten text recognition is proposed using a fine-tuned TrOCR transformer model. The system leverages large-scale multilingual pretraining and supervised fine-tuning to handle complex character structures in Khmer script. This thesis presents the dataset preparation, model architecture, training methodology, and performance evaluation using standard benchmarks. The results show significant improvements in recognition accuracy compared to traditional OCR techniques.

# Contents

1	Introduction	4
---	--------------	---

# Chapter 1

## Introduction

This is the introduction chapter.