



ANCIENT TAMIL DECODING



Submitted by

VARSITHA A	(714023104157)
VISWANATH N	(714023104174)
YAZHINI R	(714023104176)
YESWANTH S	(714023104178)

**In partial fulfilment for the
award of the degree**

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

SRI SHAKTHI

INSTITUTE OF ENGINEERING AND TECHNOLOGY,

An Autonomous Institution,

Accredited by NAAC with “A” Grade

NOVEMBER 2025

BONAFIDE CERTIFICATE

BONAFIDE CERTIFICATE

Certified that this report titled “**ANCIENT TAMIL DECODING**” is the bonafide work of **VARSITHA A(714023104157),VISWANATH N (714023104174), YAZHINI R (714023104176),YESWANTH S** who carried out the work under my supervision.

SIGNATURE

Dr.K.E.Kannammal

SUPERVISOR

Professor and Head,
Department of CSE,
Sri Shakthi Institute of
Engineering and Technology,
Coimbatore – 641062.

SIGNATURE

Dr.K.E.Kannammal

HEAD OF THE DEPARTMENT

Professor and Head,
Department of CSE,
Sri Shakthi Institute of
Engineering and Technology,
Coimbatore – 641062.

Submitted for the project work viva-voce Examination held on.....

INTERNAL EXAMINER

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

ACKNOWLEDGEMENT

First and foremost, I would like to thank God Almighty for giving me the strength. Without his blessings this achievement would not have been possible.

We express our deepest gratitude to our Chairman **Dr.S.Thangavelu** for his continuous encouragement and support throughout our course of study.

We are thankful to our Secretary **Mr.T.Dheepan** for his unwavering support during the entire course of this project work.

We are also thankful to our Joint Secretary **Mr.T.Sheelan** for his support during the entire course of this project work.

We are highly indebted to Principal , **Dr. N. K. Sakthivel, M.Tech., Ph.D., MCSI, MISTE**, for his support during the tenure of the project.

We are deeply indebted to our Head of the Department, Computer Science and Engineering, **Dr.K.E.Kannammal**, for providing us with the necessary facilities.

It's a great pleasure to thank our Project Guide **Dr.K.E.Kannammal** for her valuable technical suggestions and continuous guidance throughout this project work.

We are also thankful to our Project Coordinator **Mrs.M.Manimegala** Assistant Professor for providing us with necessary facilities and encouragement.

We solemnly extend our thanks to all the teachers and non-teaching staff of our department, family and friends for their valuable support.

VARSITHA A

VISWANATH N

YAZHINI R

YESWANTH S

ABSTRACT

ABSTRACT

The preservation and interpretation of ancient scripts remain one of the most significant challenges in the fields of digital humanities and computational linguistics. Among these, the ancient Tamil script, found on inscriptions, palm-leaf manuscripts, and copper plates, holds immense cultural, historical, and linguistic value. However, deciphering these scripts manually is time-consuming and demands specialized expertise. The *Ancient Tamil Decoding* project addresses this problem by leveraging deep learning techniques to automate the recognition and translation of ancient Tamil characters into modern Tamil text.

This study proposes a novel hybrid model combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to accurately recognize characters from degraded and stylized ancient scripts. The project follows a modular methodology that includes the collection of a labeled dataset from high-resolution ancient manuscript images, preprocessing (resizing, normalization, augmentation), label encoding, and splitting of data into training and validation sets. The CNN layers extract spatial features from the character images, while the LSTM layers capture sequential dependencies for improved contextual accuracy during prediction.

Once trained using TensorFlow/Keras, the model is capable of classifying individual characters from unseen images with high accuracy. These characters are then mapped to their modern Tamil equivalents and assembled to form grammatically correct sentences. The final output is enhanced using post-processing techniques such as spacing correction and Tamil font rendering. Tools such as OpenCV, Pandas, and Google Colab were employed throughout the workflow.

TABLE OF CONTENT

TABLE OF CONTENT

CHAPTER	TITLE	PAGE NO
	ABSTRACT	
1	INTRODUCTION	1
2	LITERATURE SURVEY	5
	2.1 INTRODUCTION TO SCRIPT RECOGNITION AND OCR	
	2.2 RESEARCH ON TAMIL SCRIPT RECOGNITION	
	2.3 DEEP LEARNING FOR SCRIPT RECOGNITION	
	2.4 CHALLENGES IN ANCIENT SCRIPT RECOGNITION	
	2.5 EXISTING SYSTEMS AND TOOLS	
	2.6 RELATED WORKS ON INDIC SCRIPTS	
	2.7 GAPS IN EXISTING RESEARCH	

3 SYSTEM ANALYSIS 12

4 PROPOSED SYSTEM 17

4.1 EXISTING SYSTEM

4.2 PROPOSED SYSTEM

4.3 DATASET PREPARATION AND PREPROCESSING

4.4 YOLOV8 MODEL ARCHITECTURE

4.5 INTEGRATION WITH ROBOFLOW AND POST-PROCESSING

4.6 ADVANTAGES OF THE PROPOSED SYSTEM

5 IMPLEMENTATION 23

5.1 SYSTEM ARCHITECTURE AND SCALABILITY

5.2 RESEARCHER DASHBOARD AND FEATURES

5.3 CUSTOMIZATION AND NOTIFICATIONS

5.4 PRIVACY AND SECURITY

5.5 DEPLOYMENT, MAINTENANCE, AND SUPPORT

6	RESULT ANALYSIS	31
	6.1 DATASET OVERVIEW	
	6.2 EXPERIMENTAL SETUP	
	6.3 RESULT INTERPRETATION	
	6.4 USER FEEDBACK AND USABILITY EVALUATION	
	6.5 ERROR ANALYSIS	
7	CONCLUSION	36
8	FUTURE WORK	39
	REFERENCES	41

INTRODUCTION

