



UNDERGRADUATE THESIS

CLASSIFICATION OF CAT BREED TYPE IMAGES USING GLCM AND DECISION TREE METHODS

DIMAS ADITYA PUTRA

NPM 19081010190

THESIS ADVISORS

Fetty Tri Anggraeny, S.Kom, M.Kom

Chrystia Aji Putra, S.Kom, M.Kom

**MINISTRY OF EDUCATION, CULTURE, RESEARCH, AND TECHNOLOGY
NATIONAL DEVELOPMENT UNIVERSITY OF VETERANS EAST JAVA
FACULTY OF COMPUTER SCIENCE
INFORMATICS STUDY PROGRAM
SURABAYA
2026**

APPROVAL SHEET

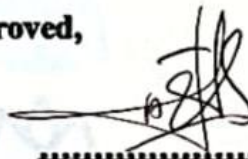
CLASSIFICATION OF CAT BREED TYPE IMAGES USING GLCM AND DECISION TREE METHODS

By:
DIMAS ADITYA PUTRA
NPM. 19081010190

Has been defended before, and accepted by, the Board of Assessors of the Thesis Examination of the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jawa Timur, on June 2, 2026:

Approved,

Fetty Tri Anggraeny, S.Kom., M.Kom.
NIP. 19820211 202121 2 005



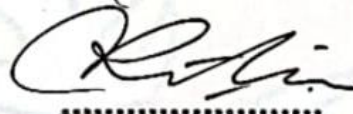
(Advisor I)

Chrystia Aji Putra S.Kom., M.Kom.
NIP. 19861008 202121 1 001



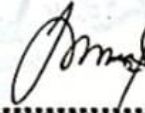
(Advisor II)

Retno Mumpuni, S.Kom., M.Sc.
NIP. 19870716 202521 2 045



(Head Assessor)

Achmad Junaidi, S.Kom., M.Kom
NIP. 1978111 0202521 1 048



(Assessor I)

Acknowledge by,

Dean of the Faculty of Computer Science




Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

APPROVAL SHEET

CLASSIFICATION OF CAT BREED TYPE IMAGES USING GLCM AND DECISION TREE METHODS

By:
DIMAS ADITYA PUTRA
NPM. 19081010190

Approved to proceed to the Thesis Examination

Approved by,

Coordinator of Informatics Study Program
Faculty of Computer Science



Dr. Intan Yuniar Purbasari, S.Kom. M.Sc.
NIP. 19800602 202521 2 029

STATEMENT OF ORIGINALITY

The undersigned:

Student Name : DIMAS ADITYA PUTRA

Study Programs : Informatics

Supervisor : Fetty Tri Anggraeny, S.Kom, M.Kom
Chrystia Aji Putra, S.Kom, M.Kom

It is hereby stated that the contents of part or all of the dissertation with the title:

CLASSIFICATION OF CAT BREED TYPE IMAGES USING GLCM AND DECISION TREE METHODS

is truly the work of an independent intellectual, completed without the use of unauthorized materials and is not the work of another party that I acknowledge as my own work. All references cited or referenced have been written in full in the bibliography. If it turns out that this statement is not true, I am willing to accept sanctions in accordance with the applicable regulations.



Surabaya, 17 Juli 2026
Who made the statement,


DIMAS ADITYA PUTRA
NPM. 19081010190

ABSTRACT

Student Name /NPM : Dimas Aditya Putra / 19081010190
Thesis Title : Image Classification of Cat Breed Types Using
GLCM and Decision Tree Methods
Supervisor : 1. Fetty Tri Anggraeny, S.Kom, M.Kom
2. Chrystia Aji Putra, S.Kom, M.Kom

This study aims to develop a classification system for cat breed images using the Gray Level Co-occurrence Matrix (GLCM) method as a texture feature extraction technique and the Decision Tree algorithm as a classification method. The dataset used consists of a number of cat imagery from several breeds that have gone through the preprocessing stage, such as resizing and image conversion to grayscale.

The texture features produced from GLCM include contrast, correlation, energy, and homogeneity, which are then used as inputs in the classification process using Decision Tree. The built model is evaluated using testing methods such as accuracy, precision, and recall to measure system performance.

The results of the study show that the combination of the GLCM and Decision Tree methods is able to provide a fairly good classification performance in distinguishing cat breeds based on imagery. The system is expected to help in the process of automatic identification of cat breeds and can be further developed with more complex methods to improve accuracy.

Keywords : Image Classification, Cat Breed, GLCM, Decision Tree, Digital Image Processing

FOREWORD

We thank Allah SWT, God Almighty, because thanks to His blessings, grace, and pleasure we were able to complete this thesis report smoothly. Only because of His permission and kindness that we can feel every day, the author was able to complete the thesis with the title:

"Image Classification of Cat Breed Types Using GLCM and Decision Tree Methods"

There is a lot of support or assistance obtained during this thesis research process. With respect and gratitude, the author also wishes to all related parties who helped in the process of completing this research.

The author realizes that there are still many shortcomings in this study considering the limitations of the author's abilities and knowledge. Therefore, the author accepts all forms of constructive criticism and suggestions from all parties in the improvement of this thesis.

Surabaya, _____
Author

TABLE OF CONTENTS

APPROVAL SHEET.....	4
APPROVAL SHEET.....	5
STATEMENT OF ORIGINALITY	6
ABSTRACT	7
FOREWORD	8
TABLE OF CONTENTS.....	9
IMAGE LIST	12
TABLE LIST	13
LIST OF NOTATIONS	14
CHAPTER I.....	1
INTRODUCTION.....	1
1.1 Background	1
1.2 Problem Formulation	2
1.3 Problem Limitations	3
1.4 Research Objectives.....	3
1.5 Research Benefits	4
CHAPTER II	6
LITERATURE REVIEW.....	6
2.1 Previous Research.....	6
2.2 Cat Breeds	7
2.3 Classification.....	9
2.4 Digital Image.....	10

2.4.1 Color Imagery.....	10
2.4.1 Grayscale Image	10
2.4.2 Binary Image	10
2.5 GLCM (Gray Level Co-occurrence Matrix).....	10
2.6 Decision Tree.....	13
2.6.1 Advantages and Disadvantages of Decision Tree.....	13
2.6.2 How Decision Tree Works	14
2.7 C4.5 Algorithm.....	15
2.8 Google Colab.....	15
2.9 Machine Learning.....	16
2.9.1 Supervised Learning	16
2.10 Evaluation of Performers.....	17
2.10.1 Confusion Matrix.....	17
2.10.2 Accuracy.....	18
2.10.3 Accuracy.....	18
2.10.4 Recall.....	18
2.10.5. F-1 Score	19
CHAPTER III.....	20
SYSTEM DESIGN AND IMPLEMENTATION.....	20
3.1 Research Methodology	20
3.1.1 Cat breed imagery dataset.....	20
3.1.2 Grayscale.....	22
3.1.3 Data transformation (Image to Array)	23
3.1.4 GLCM Feature Extraction	24

3.1.5 GLCM values or patterns	27
3.1.6 Training Decision Tree Based on GLCM pattern.....	29
3.1.7 Testing.....	33
3.1.8 Confusion Matrix.....	34
3.2 Trial Scenario	34
CHAPTER IV.....	35
TESTING AND ANALYSIS.....	35
4.1 Data Analysis.....	35
4.2 Data Preparation.....	35
4.3 Preprocessing Data	35
4.4 Model Design	37
4.5 Evaluation.....	42
4.6 Test Scenarios.....	43
4.6.1 Small Dataset Scenarios.	43
4.6.2 Dataset scenarios with 80%:20% data split.	47
4.6.3 Dataset scenario with 75%:25% data split.....	51
4.6.4 Dataset Scenario with 70%:30% data Split.	55
4.7 Perform 10 Runs On Each Scenario	58
CHAPTER V	64
CONCLUSION	64
5.1 Conclusion.....	64
5.2 Suggestions.....	65
BIBLIOGRAPHY	66

IMAGE LIST

Figure 2. 1 GLCM matrix pixel illustration.....	11
Figure 2. 2 GLCM matrix before normalization.....	11
Figure 2. 3 GLCM matrix normalization.....	12
Figure 2. 4 Supervised Learning.....	17
Figure 2. 5 Matrix Confusion Table	18
Figure 3. 1 Research methodology diagram block	20
Figure 3. 2 Flowchart grayscale.....	22
Figure 3. 3 Array transform flowchart.....	23
Figure 3. 4 Flowchart pre-processing and feature extraction	25
Figure 3. 5 Flowchart C4.5	30
Figure 3. 6 Flowchart test C4.5	33
Figure 4. 1 Labeling Results.....	36
Figure 4. 2 Grayscale Image Results	36
Figure 4. 3 GLCM Feature Extraction Results Images	39
Figure 4. 4 Classification Report.....	42
Figure 4. 5 Small Confusion Matrix Datasets.....	42
Figure 4. 6 Small Dataset Classification Results.....	43
Figure 4. 7 GLCM Features.....	45
Figure 4. 8 Confusion Matrix 80:20	46
Figure 4. 9 Split Data Classification Results 80%:20%	47
Figure 4. 10 Features of GLCM 80:20	48
Figure 4. 11 Confusion Matrix 75:25	49
Figure 4. 12 Split Data Classification Results 75%:25%	50
Figure 4. 13 GLCM 75:25 Features.....	52
Figure 4. 14 Confusion Matrix 70:30	53
Figure 4. 15 Split Data Classification Results 70%:30%	54
Figure 4. 16 GLCM 75:25 Features.....	55

TABLE LIST

Table 2. 1 Types of cat breeds	8
Table 3. 1 Example of a cat breed dataset	21
Table 4. 1 Run 10 Small Dataset Scenarios.....	56
Table 4. 2 Run 10 Times Dataset Scenarios with 80% Data Split:20%	58
Table 4. 3 Run 10 Times Dataset Scenarios with 75% Data Split:25%	59
Table 4. 4 Run 10 Times Dataset Scenarios with 70% Data Split:30%	60

LIST OF NOTATIONS

Notation	Remarks
$p(x,y)$	Probability of the appearance of pixel pairs at position (x,y) in the GLCM matrix
x	Pixel row coordinates on a GLCM matrix
y	Pixel column coordinates on the GLCM matrix
d	Distance between pixels (in pixels)
θ	GLCM directional angle orientation (0° , 45° , 90° , 135°)
μ_x	Mean values of rows in GLCM matrix
μ_y	Mean values of columns in the GLCM matrix
σ_x	Standard line deviation in GLCM matrix
σ_y	Standard column deviation in GLCM matrix
S	Datasets in the C4.5 algorithm
A	Attributes used in the Information Gain calculation
n	Number of attribute partitions or number of S partitions
Yes	Number of cases on the i partition
bee	Conditional probability on the calculation of Entropy
TP	True Positive – positive data predicted to be true
TN	True Negative – negative data predicted to be true
FP	False Positive – negative data predicted positive
FN	False Negative – positive data predicted to be negative
RGB	Red Green Blue – color model of color imagery