

# DEEP LEARNING-BASED WOUND HEALING ANALYSIS

IN *Monopterus cuchia* (SWAMP EEL)

*THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR  
THE DEGREE OF*

**MASTER OF TECHNOLOGY**

**IN**

**INFORMATION TECHNOLOGY**

*SUBMITTED BY*

**Chidangsha Sekhar Bezbaruah**

**Roll no: PT-231-844-0014**

*UNDER THE GUIDANCE OF*

**Dr. Nabamita Deb**

**Assistant Professor**



**DEPARTMENT OF INFORMATION TECHNOLOGY  
GAUHATI UNIVERSITY, ASSAM  
JUNE, 2025**



**GAUHATI UNIVERSITY**  
**DEPARTMENT OF INFORMATION TECHNOLOGY**  
Gopinath Bordoloi Nagar, Jalukbari , Guwahati-781014

---

**DECLARATION**

I “*Chidangsha Sekhar Bezbaruah*”, Roll No “**PT-231-844-0014**”, a M.Tech. student of the department of Information Technology, Gauhati University hereby declares that I have compiled this thesis reflecting all my works during the semester long full time project as part of my MTech curriculum. I declare that I have included the descriptions etc. of my project work, and nothing has been copied/replicated from other’s work. The facts, figures, analysis, results, claims etc. depicted in my thesis are all related to my full time project work. I also declare that the same thesis or any substantial portion of this thesis has not been submitted anywhere else as part of any requirements for any degree/diploma etc.

---

Chidangsha Sekhar Bezbaruah  
10-06-2025



**GAUHATI UNIVERSITY**  
**DEPARTMENT OF INFORMATION TECHNOLOGY**  
Gopinath Bordoloi Nagar, Jalukbari , Guwahati-781014

---

Date:10-06-2025

**CERTIFICATE**

This is to certify that “**Chidangsha Sekhar Bezbaruah**” bearing Roll No: “**PT-231-844-0014**” has carried out the project work “**Deep Learning based Wound Healing Analysis in *Monopterus cuchia* (Swamp eel)**” under my supervision and has compiled this thesis reflecting the candidate’s work in the semester long project. The candidate did this project full time during the whole semester under my supervision, and the analysis, results, claims etc. are all related to his studies and works during the semester.

I recommend submission of this thesis as the project report as a part for partial fulfillment of the requirements for the degree of Master of Technology in Information Technology of Gauhati University.

---

Dr. Nabamita Deb  
Supervisor  
Assistant Professor  
Department of Information Technology



**GAUHATI UNIVERSITY**  
**DEPARTMENT OF INFORMATION TECHNOLOGY**  
Gopinath Bordoloi Nagar, Jalukbari , Guwahati-781014

---

Date:10-06-2025

**External Examiners Certificate**

This is to certify that “**Chidangsha Sekhar Bezbaruah**”, bearing Roll No: “**PT-231-844-0014**” has delivered his project presentation on “**10/06/2025**” and I examined his thesis entitled “**Deep Learning based Wound Healing Analysis in *Monopterus cuchia* (Swamp eel)**” and recommend this thesis as the project report as a part for partial fulfillment of the requirements for the degree of Master of Technology in Information Technology of Gauhati University.

To the best of my knowledge, the work has not been submitted to any other institute for the award of any other degree or diploma.

---

(External Examiner)



**GAUHATI UNIVERSITY**  
**DEPARTMENT OF INFORMATION TECHNOLOGY**  
Gopinath Bordoloi Nagar, Jalukbari , Guwahati-781014

---

Date:10-06-2025

**TO WHOM IT MAY CONCERN**

This is to certify that "**Chidangsha Sekhar Bezbaruah**" bearing Roll No "**PT-231-844-0014**", a M.Tech. student of the department of Information Technology, Gauhati University, has submitted the softcopy of his thesis for undergoing screening through anti-plagiarism software and the similar report found to be

---

Dr. Nabamita Deb  
Supervisor  
Assistant Professor  
Department of Information Technology

## **ACKNOWLEDGEMENT**

### **Acknowledgement**

I would like to express my sincere gratitude to all those who contributed to the completion of this project. First and foremost, I extend my deepest appreciation to my supervisor, Dr. Nabamita Deb, for her invaluable guidance, support, and encouragement throughout the entire research process.

I am also thankful to Dr. Banasri Mech and Dr. Kuladip Sarma, as well as Department of Zoology, Gauhati University, for granting permission to use the wound healing dataset employed in this work. They were cooperative and helpful in facilitating this study.

Chidangsha Sekhar Bezbaruah

## ABSTRACT

### Deep Learning based Wound Healing Analysis in *Monopterus cuchia* (Swamp eel):

Wound healing is a crucial biological mechanism that aids living organisms recover from injuries. In this project, we studied the healing patterns in a freshwater fish called *Monopterus cuchia* using machine learning techniques. We focused on two types of healing—normal recovery and healing supported by Retinoic Acid treatment. The main goal was to build a system capable of automatically categorizing images of wounds based on their healing stage.

We used a deep learning approach by modifying the MobileNetV2 model and created our own version named **CuchiaNet**. The model is trained with augmented image data and improved through hyperparameter tuning. To make the model easy to use, we also developed a simple web-based interface where users could upload images and get predictions in real-time.

In addition to classification, we applied clustering techniques to discover patterns in the wound images without using labels. Principal Component Analysis (PCA) and K-Means clustering helped group similar wound images and gave us insights into different healing trends.

This project combines both supervised and unsupervised learning methods to support biological research and improve wound assessment. The final model operates with great precision and is suitable for use in labs and field studies.

**Keywords:** *K-Means Clustering, PCA, MobileNetV2, Transfer learning, CuchiaNet*

# Contents

<b>Declaration</b> . . . . .	I
<b>Certificate</b> . . . . .	II
<b>Acknowledgement</b> . . . . .	V
<b>Abstract</b> . . . . .	VI
<b>1 Introduction</b>	1
1.1 Background on <i>Monopterus cuchia</i> and Wound Healing . . . . .	1
1.2 Introduction to Machine Learning . . . . .	1
1.3 Problem Statement . . . . .	2
1.4 Objectives . . . . .	3
1.5 Scope of the Work . . . . .	3
1.6 Organization of the Report . . . . .	3
<b>2 Literature Review</b>	4
2.1 Background Study . . . . .	4
2.2 Related Work . . . . .	5
2.3 Existing Methods . . . . .	6
<b>3 Dataset and Preprocessing</b>	14
3.1 Dataset Description . . . . .	14
3.2 Image Preprocessing . . . . .	15
3.3 Data Augmentation . . . . .	15
3.4 Dataset Availability . . . . .	16
3.5 Class Distribution and Visualization . . . . .	16
<b>4 Methodology</b>	18
4.1 Overview of the Methodology . . . . .	18
4.2 Proposed Methodology . . . . .	18
4.3 Dataset Collection and Preprocessing . . . . .	19
4.4 Clustering Pathway . . . . .	20
4.4.1 Data Preparation . . . . .	21
4.4.2 Image Flattening and Feature Extraction . . . . .	21
4.4.3 Dimensionality Reduction Using PCA . . . . .	21
4.4.4 K-Means Clustering . . . . .	21
4.4.5 Cluster Visualization and Interpretation . . . . .	21
4.5 Classification Pathway . . . . .	21
4.5.1 Transfer Learning with MobileNetV2 . . . . .	22
4.5.2 Model Architecture Design . . . . .	22
4.5.3 Hyperparameter Optimization . . . . .	23
4.5.4 Training Strategy . . . . .	23
4.5.5 Training Optimization Techniques . . . . .	23
4.6 Performance Evaluation . . . . .	24

<b>5</b>	<b>Implementation</b>	<b>25</b>
5.1	Development Setup . . . . .	25
5.2	Data Loading and Augmentation . . . . .	25
5.3	Clustering Pathway Implementation . . . . .	26
5.3.1	Dataset Loading and Structuring . . . . .	26
5.3.2	Image Vectorization and Preprocessing . . . . .	26
5.3.3	Dimensionality Reduction using PCA . . . . .	26
5.3.4	K-Means Clustering . . . . .	27
5.3.5	Silhouette Score Evaluation . . . . .	27
5.3.6	Visualization of Clusters . . . . .	27
5.3.7	Label Mapping for Interpretation . . . . .	27
5.4	Classification Pathway Implementation . . . . .	27
5.4.1	Model Architecture . . . . .	27
5.4.2	Hyperparameter Tuning . . . . .	28
5.4.3	Training Phase I – Feature Extraction . . . . .	29
5.4.4	Training Phase II – Fine-Tuning . . . . .	29
5.5	Graphical User Interface Deployment . . . . .	30
5.5.1	Interface Design . . . . .	30
5.5.2	Backend Workflow . . . . .	30
5.5.3	User Experience and Responsiveness . . . . .	31
5.5.4	Snapshots of the Web-application . . . . .	31
<b>6</b>	<b>Results and Discussion</b>	<b>33</b>
6.1	Results from the Clustering Pathway . . . . .	33
6.1.1	Principal Component Analysis (PCA) . . . . .	33
6.1.2	K-Means Clustering . . . . .	34
6.1.3	Silhouette Score Evaluation . . . . .	35
6.1.4	Cluster Distance Matrix . . . . .	35
6.2	Results from the Classification Pathway . . . . .	36
6.2.1	Model Performance Evaluation . . . . .	36
6.2.2	Confusion Matrix and Class-wise Performance . . . . .	37
6.2.3	Classification Report . . . . .	37
6.2.4	ROC Curve and AUC Analysis . . . . .	38
6.2.5	Interpretation of Results . . . . .	38
6.3	Discussion . . . . .	39
6.3.1	Deep Learning Classification Model (CuchiaNet) . . . . .	39
6.3.2	Dimensionality Reduction using PCA . . . . .	39
6.3.3	K-Means Clustering for Unsupervised Analysis . . . . .	40
6.3.4	Interpretation . . . . .	40
<b>7</b>	<b>Conclusion and Future Work</b>	<b>41</b>
7.1	Concluding Remarks . . . . .	41
7.2	Future Work . . . . .	41

# List of Figures

1.1	Types of ML [1]	2
3.1	Dataset images of different healing stages	14
3.2	Images from the dataset after augmentation	15
3.3	Frequency Distribution of the dataset	16
3.4	Frequency Distribution of dataset after augmentation	17
4.1	Proposed Methodology Workflow	19
4.2	The workflow diagram of Clustering Pathway	20
4.3	The workflow diagram of Classification Pathway	22
5.1	The proposed architecture of CuchiaNet	28
5.2	Streamlit application overview showing sidebar and main workspace.	31
5.3	Upload state: the user selects a wound image for analysis.	31
5.4	Prediction state: the model returns the healing type and confidence.	32
6.1	Explained Variance by Each Principal Component	33
6.2	Cumulative Explained Variance	34
6.3	K-Means Clustering Results with 2 Clusters	34
6.4	Silhouette Scores for K-Means Clustering	35
6.5	Distance Matrix Between Cluster Centroids	36
6.6	Training and Validation Accuracy and Loss over Epochs	37
6.7	Confusion Matrix of Classification Results	37
6.8	ROC Curve and AUC for Classification	38

# List of Tables

2.1	Overview of methods and models for wound healing research in the paper. . . . .	7
2.2	Overview of machine learning techniques for wound healing prediction and assessment (Part 1) . . . . .	8
2.3	Overview of machine learning techniques for wound healing prediction and assessment (Part 2) . . . . .	9
2.4	Overview of machine learning techniques for wound healing prediction and assessment (Part 3) . . . . .	12
2.5	Overview of machine learning techniques for wound healing prediction and assessment (Part 4) . . . . .	13
6.1	Classification Report on Validation Set . . . . .	38
6.2	Summary of Observations for clustering . . . . .	40

# Listings

5.1	Image Augmentation Setup . . . . .	25
5.2	Training Phase I with EarlyStopping and ModelCheckpoint . . . . .	29
5.3	Fine-Tuning the Last 30 Layers of MobileNetV2 . . . . .	29

# Chapter 1

## Introduction

### 1.1 Background on *Monopterus cuchia* and Wound Healing

Wound healing is a intricate biological process influenced by numerous physiological, chemical, and environmental factors. Timely and accurate assessment of wound healing progression is crucial in both clinical and research settings to optimize treatment strategies and monitor recovery. Traditionally, this evaluation has been subjective, relying heavily on visual inspection by clinicians. With the rise of computational approaches and the availability of image-based datasets, machine learning particularly deep learning has opened new avenues for objective, automated wound analysis.

The Swamp eel (*Monopterus cuchia*), also called the rice field eel, is not only an economically important freshwater fish in aquacultural production in the world, but also an increasingly recognized model organism for biological studies. It has a small genome with 0.8 Gb and a chromosome number of 18–24 that is among the least in teleost species.[2]

Eels are Pisces class animals which possess numerous ingredients beneficial to humans. Some of these include omega-3 and omega-6 fatty acids. The body content of arachidonic acid and DHA in an eel's body is 8.25 and 6.21 g for every 100 g fat, respectively [3]. According to an earlier study by Febriyenti et al. [4], the composition of fatty acids in eel extracts was dominated by oleic acid (19.7%), palmitic acid (18.7%), pentadecanoic acid (15.81%) and octadecanoic acid (4.87%). Omega-3 and omega-6 fatty acids have a critical role in the process of wound healing.[5]

The treatment of wound in the past was done by letting the wound dry and develop a hard layer, which would peel off naturally. But with the progress in science and technology, the treatment of the wound has evolved. It is known that the wound heals faster if the moist is covering the wound as the wound dressing. Conventionally, cotton gauzes were applied as wound dressings. Some of the types of wound dressings include hydrocolloid, hydrogel, foam, and semi-permeable adhesive membrane [6] [7]

### 1.2 Introduction to Machine Learning

Machine learning techniques provide exciting new ways to exploit the available computational power and data in a variety of scientific domains. They can analyze huge amounts of data

in a relatively short time that is not possible by manual labor. This provides opportunities for scientists to develop new experimental procedures and to channel their efforts on the most promising questions of their problem domain.[8]

Machine learning is a subfield of artificial intelligence (AI) that uses algorithms trained on data sets to create self-learning models that are capable of predicting outcomes and classifying information without human intervention. [9]. All tho, as opposed to applications of artificial intelligence, machine learning entails learning of hidden patterns in the data (data mining) and then applying the patterns to classify or predict an event corresponding to the problem[10]. Machine learning is usually classified into two categories according to the availability of the data and input given to the learning system; these include *supervised learning* and *unsupervised learning* approach [11]. Along with these two supervised and unsupervised learning , another type of learning approach, which is less frequently used are shown in figure 1.1 with usecase eaxmple of each type.

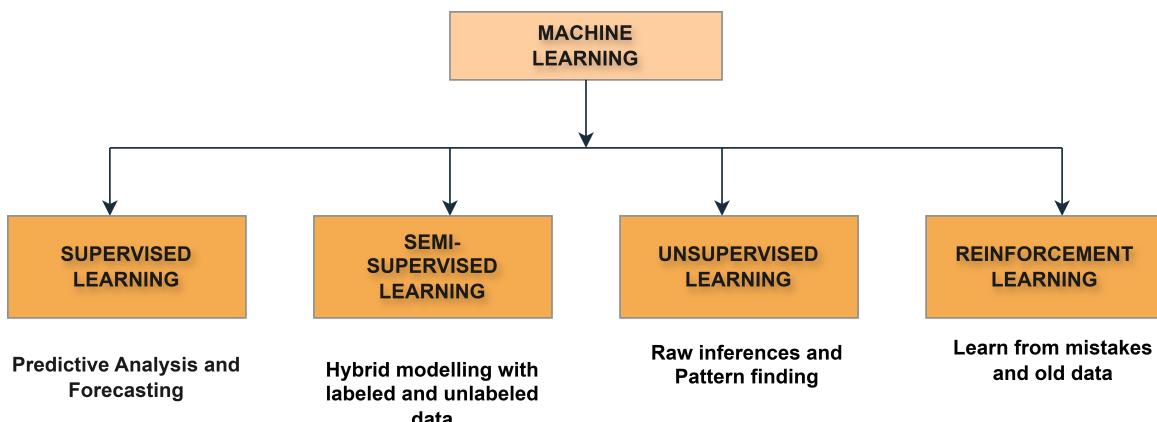


Figure 1.1: Types of ML [1]

Machine learning problems primarily revolve around tasks such as classification, prediction, and clustering. Some example of Machine learning problems are [12]

- *Classification*: Classification involves assigning data points to predefined categories or labels based on their features. This is commonly used in applications such as spam detection, where emails are classified as spam or not, or while diagnosing medical conditions, where symptoms are classified into potential diseases [13].
- *Clustering*: Clustering is an unsupervised learning task where the aim is to combine related data points according to their inherent patterns. Unlike classification, clustering doesn't require predefined labels and is often used in exploratory data analysis, such as grouping customers by purchasing behavior or identifying patterns in biological data [14].
- *Prediction*: Prediction deals with forecasting continuous values, which aims to predict an outcome based on input data. For instance, predicting house prices or stock market trends involves using historical data to estimate future values [15].

### 1.3 Problem Statement

The project aims to explore the wound healing process in *Monopterus albus* using machine learning (ML) techniques, with an emphasis on comparing natural healing with healing assisted by medicinal treatment. This research seeks to assess the effectiveness of ML-based methods in distinguishing between normal and medicinally assisted wound healing. Manual observation of

wound healing is time-consuming, subjective, and lacks reproducibility. It is necessary to have a robust, automated system that is capable of:

- Accurately classifying wound healing stages using image data.
- Discovering hidden patterns or anomalies in the healing process.
- Enhancing model explainability to gain clinical insights.

## 1.4 Objectives

The primary objectives of this project are:

- To develop deep learning-based classification model using **MobileNetV2** with transfer learning and data augmentation.
- To improve model performance through fine-tuning and regularization techniques.
- To evaluate the classification model using performance metrics such as accuracy, confusion matrix, ROC curve, and classification report.
- To perform unsupervised clustering (K-Means, PCA) to identify natural groupings within the wound healing dataset.

## 1.5 Scope of the Work

This project focuses on implementing, training, and evaluating a machine learning pipeline using monoscopic image data of wound healing in *Monopterus cuchia*. The scope includes:

- Image classification using convolutional neural networks (CNNs), particularly MobileNetV2.
- Clustering analysis using PCA and K-Means.
- Data augmentation techniques to improve model generalization.

Real-time monitoring and clinical deployment are beyond the scope of this project.

## 1.6 Organization of the Report

The structure of this report is **Chapter 2** presents review of the literature on wound healing analysis, transfer learning, and clustering methods. **Chapter 3** explains the dataset, including its structure, preprocessing steps, and augmentation. **Chapter 4** details the methodology for classification and clustering, including model architecture and training approach. **Chapter 5** discusses the implementation details and tools used. **Chapter 6** presents and discusses the experimental results and evaluation. **Chapter 7** concludes the report and outlines future work directions.

# Chapter 2

## Literature Review

This chapter gives a summary of existing research and developments relevant to wound healing analysis using machine learning, particularly deep learning techniques. It also covers the application of clustering methods for pattern discovery and the application of explainable AI tools to interpret model decisions. Because wound healing can regenerate, research on wound healing has become increasingly focused, especially in species like *Monopterus albus*. A useful tool for examining wound healing processes is the application of machine learning and clustering algorithms, particularly in situations involving both normal repair and regeneration enhanced by retinoic acid. This chapter reviews the pertinent literature with an emphasis on machine learning and clustering methods used in wound healing. Studies examining both conventional wound healing procedures and developments in computational methods for forecasting and assessing wound healing results are included in this review.

### 2.1 Background Study

Wound healing involves a series of biological processes, as well as hemostasis, inflammation, proliferation, and remodeling. Traditional assessment methods rely on physical observation, photography, and measurement tools, which often suffer from subjectivity and inconsistency. Recent studies have focused on leveraging computer vision [16] and machine learning [17] to automate wound assessment. Studies show that deep learning models can accomplish high accuracy in tasks such as wound segmentation, infection detection, and healing stage classification. The use of digital image datasets captured over time provides a valuable basis for training and validating such models.

Transfer learning has appeared as a powerful approach in medical imaging [18], where annotated datasets are often limited. Models pretrained on large-scale datasets like ImageNet can be fine-tuned for specific tasks with relatively smaller domain-specific datasets.

MobileNetV2, a lightweight convolutional neural network (CNN), has shown promising results in medical image classification [19] due to its balance between performance and computational efficiency. Its use of depthwise separable convolutions significantly reduces the number of parameters while maintaining accuracy. Several studies have effectively applied MobileNetV2 and other pretrained architectures (such as ResNet, Inception, and EfficientNet) to classify diseases, detect abnormalities, and track recovery progress in various medical contexts.

Data augmentation is crucial for enhancing the generalization capabilities of deep learning models [20], especially in limited-data scenarios. Techniques such as rotation, zooming, flipping, shifting, and brightness adjustment artificially increase the size and diversity of training

datasets. In the context of wound healing, augmentation ensures the model becomes robust to variations in lighting, orientation, and wound shape, ultimately reducing overfitting and improving validation performance.

Unsupervised learning methods [21], particularly clustering techniques like K-Means, DBSCAN, and hierarchical clustering, are used to discover intrinsic structures in image datasets. Principal Component Analysis (PCA) [22] is commonly employed as preprocessing step for dimensionality reduction before clustering [14].

The literature review explored various approaches in the domain of automated wound assessment, medical image analysis, and the application of machine learning and deep learning techniques for biomedical tasks. From the reviewed studies, it is evident that traditional image processing techniques such as edge detection, thresholding, and region-based segmentation have been extensively utilized for wound identification. However, these techniques are quite delicate, to noise, lighting conditions, and require hand-crafted feature extraction, decreasing their dependability for clinical deployment. In contrast, recent advancements have shown the effectiveness of deep learning, especially Convolutional Neural Networks (CNNs), in achieving high accuracy in image classification and segmentation tasks. Models like VGGNet, ResNet, and MobileNet have been frequently applied to medical image classification because of their capacity to learn complex hierarchical characteristics using raw picture data. Transfer learning has become a valuable strategy to mitigate the limitations of small biomedical datasets. By utilizing pre-trained models on large-scale datasets like ImageNet, researchers have improved the generalization capability of models on domain-specific tasks like skin lesion detection, burn classification, and wound stage recognition. Additionally, a number of studies have introduced unsupervised techniques such as K-Means, DBSCAN, and hierarchical clustering to discover hidden patterns in unlabeled medical data. These methods aid in data exploration and grouping similar pathological cases without predefined categories. Overall, the literature supports a hybrid strategy that blends deep learning with interpretability and unsupervised clustering for a comprehensive wound analysis framework. However, there remains a research gap in domain-specific studies, particularly in less explored species like *Monopterus cuchia*, and the integration of both classification and clustering in a single pipeline. This project aims to bridge these gaps by developing a MobileNetV2-based classifier, employing data augmentation, fine-tuning, K-Means clustering on deep features on wound healing images of *Monopterus cuchia*.

## 2.2 Related Work

Advances in wound healing research have recently centered on applying computational methods to improve our comprehension of tissue regeneration. For example, studies have automated the segmentation of wound photos, assisting in the prediction of healing time, by utilizing color histogram-based techniques and K-means clustering [23]. Furthermore, the microbiomes found in wound sites have been investigated using SigClust, a signature-based clustering technique, has been utilized to uncover important microbial clusters that influence healing rates [24].

Research on retinoic acid has demonstrated that through stimulating cellular differentiation and proliferation, it is essential for improving wound healing. For instance, in models with both normal and diabetic wounds, retinoic acid has shown efficacy in speeding wound closure when added to different hydrogel formulations [25]. All-trans retinoic acid (ATRA) was found to enhance muscle regeneration during wound healing in fetal mice models, demonstrating similar results [26]. These studies demonstrate the therapeutic potential of retinoic acid in wound healing, particularly when paired utilizing sophisticated computational methods for accurate

prediction and analysis.

The paper "Wound healing models: A systematic review of animal and non-animal models" [27] provides an in-depth analysis of various methodologies used in wound healing research, including animal models like rodents and pigs, non-animal approaches such as *in silico* and tape-stripping methods, and their combined application. It highlights the limitations and advantages of each model while emphasizing the necessity of integrating multiple approaches to comprehensively study the complex processes of wound healing. This review underscores the importance of selecting appropriate models, serving as a valuable foundation for exploring ML-driven techniques in wound healing analysis. In-depth analysis of this paper is shown in Table 2.1

## 2.3 Existing Methods

Machine learning methods are being used more and more to analyze wound features and forecast the course of wound healing. Methods for segmenting wound photos and forecasting healing durations include color histogram-based clustering, hybrid machine learning models, and K-means clustering. These methods could be very helpful in examining the impact of both normal and retinoic acid-assisted healing processes in *Monopterus albus* wound healing.

Traditional clustering algorithms such as K-means, which extract essential parameters like area, perimeter, and wound form from images and segment wound regions, are among the methods that are now in use. The healing period of both circular and irregular wound shapes is then predicted using these criteria [23]. In a similar vein, wound microbiomes have been analyzed using signature-based clustering to pinpoint bacterial strains that affect the course of healing [24].

Forecasting wound healing trajectories has showed potential for machine learning techniques encompassing Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), and BiLSTM models. With the use of these models on sizable datasets of wound photos, accurate estimates of healing times and the identification of variables influencing healing delays have been made possible [28]. Additionally, hybrid models that incorporate RNN and CNN methods have been applied to identify problems such as wound infections and have established to be considerable improvement over conventional clinical approaches [31].

Machine learning models has been utilized in the context of retinoic acid treatment to forecast the effects of different retinoic acid concentrations on wound healing, especially in diabetic models. For example, retinoic acid-based nanoparticle systems have demonstrated faster healing, and the use of machine learning algorithms to track the healing process and choose the best course of action [25].

The paper Wound image evaluation with machine learning [17] explored machine learning techniques such as Neural Networks, Random Forest, and Support Vector Machines for the assessment of wound images, particularly images of pressure ulcers in home-care patients. An image segmentation method based on clustering and three different machine learning classifiers—support vector machines, feed-forward neural networks and random forest decision trees—have been offered to complete automatic tissue recognition for pressure ulcer diagnosis, with photographs captured in environments with uncontrolled illumination. The study highlighted significant differences in model efficiencies, with overall accuracies of 81.87% (Neural Networks), 87.37% (Random Forest), and 88.08% (SVM). These findings demonstrate the potential of ML models in wound assessment and classification.

Sl No.	Model Type	Method/Technique	Advantages	Limitations	Key Insight
1	Animal Models (Rodents)	Donut-shaped silicone splints for acute wounds	Cost-effective, widely used	Differences in skin anatomy from humans; not suitable for chronic wounds	Useful for studying acute wound healing processes
2	Animal Models (Pigs)	Multiple wounds created to reduce animal usage	Reduces animals needed for study; closer to human skin	High costs; methodological variations in excision size	Suitable for ischemic wound studies despite cost
3	Non-Animal Models (In Silico)	Mathematical equations to simulate healing phases	Theoretical models for scaffold and tissue design	Lacks biophysical properties of human skin	Provides theoretical insights into wound healing
4	In Vitro Models	Scratch assays for keratinocyte disruption	Simple setup, partially mimics cell behavior	Removal of autocrine factors complicates healing study	Useful for studying re-confluence of keratinocytes
5	Ex Vivo (Skin Explants)	3D structure study of wound repair	Reflects microenvironment and cell-matrix interaction	Lack of innervation; no desquamation observation	Mimics healing environment better than in vitro models
6	Burn Models (Rodents)	Various methods to induce burns	Low cost, accessible setup	Rapid healing complicates chronic wound study	Good for acute burn healing mechanisms
7	Chronic Wound Models (Rabbits)	Bacteria-induced chronic wounds	Provides insights into chronic wound healing	High breeding costs limit usage	Suitable for bacterial infection-related chronic wounds
8	Para-biosis Models (Rodents)	Studies circulating factors in healing	Explores systemic factors in healing	Mortality risks during experiments	Useful for studying granuloma tissue development
9	Tape Stripping	Adhesive tape removes the stratum corneum	Simple, minimally invasive method	Inconsistent wound depth and uniformity	Primarily used for studying skin barrier function
10	Xeno-grafts (Nude Mice)	Human skin grafts on mice	Allows human-like wound healing study	Immune system differences affect relevance	Insights into human-specific healing factors

Table 2.1: Overview of methods and models for wound healing research in the paper.

A study [33] on the prediction of amputation wound healing in patients with critical limb ischemia utilized a machine learning algorithm integrated with multispectral imaging and patient clinical risk factors. This approach demonstrated 91% sensitivity and 86% specificity in predicting non-healing sites after 30 days, with an area under the curve of 0.89. The study highlights the potential of machine learning in enhancing predictive accuracy, reducing the need for reoperation, minimizing delayed healing, and offering significant cost savings for patients and healthcare systems.

Sl No.	Paper Title	Technique Used	Database Used	Result	Key Insight
1	Interpolating and Forecasting Wound Trajectory using Machine Learning [28]	XGBoost, LSTM, BiLSTM, ARIMA	14,571 wounds from 6,171 patients	Accurate wound trajectory forecasting	Improved precision in predicting wound healing time by filling missing data
2	Machine Learning Model for Healing Analysis of Human Injury [29]	Image segmentation and classification	Wound surface area data	Reliable wound recovery evaluation	Automated wound image segmentation enhances healing analysis
3	Machine learning systems and method for assessment, healing prediction, and treatment of wounds [30]	Image segmentation, reflectance analysis	Diabetic foot ulcers data	Accurate wound healing prediction	Machine learning assesses wound reflectance for better predictions
4	Predictive Chronic Wound Monitoring Protocol for Healing Assessment [31]	Artificial neural networks (ANN)	Chronic leg ulcers	Detection of inflammation and infection	AI-based methods improve early detection of wound complications
5	Machine Learning Model for Healing Analysis of Human Injury [32]	Image segmentation and de-noising	Chronic wound images	Improved segmentation and healing assessment	Accurate evaluation of wound recovery stages using ML techniques

Table 2.2: Overview of machine learning techniques for wound healing prediction and assessment (Part 1).

A novel wound imaging system combines advanced multispectral imaging (MSI) data with patient clinical risk factors using a modified machine learning architecture. This approach is based on a commonly used image segmentation model, enhanced with feature-wise linear modulation to effectively integrate MSI data and clinical parameters. Clinical risk factors were selected for algorithm training based on their absolute correlation coefficients with 30-day wound healing outcomes, with factors exceeding a threshold of 0.25 included in the model. The system demonstrated exceptional performance, achieving 91% sensitivity and 86% specificity for predicting primary amputation wound healing, outperforming traditional surgeon judgment. This innovative system has the potential to reduce reoperation rates, enhance clinical decision-making, and significantly lower healthcare costs, addressing a critical gap in the management of limb-threatening ischemia.[33]

In this research [35], machine learning models were developed using electronic health record (EHR) data to predict chronic wound healing times, with a focus on identifying patients at risk of non-healing or delayed healing. The models, trained on a dataset consisting over 1.2 million wounds with 187 covariates, demonstrated high accuracy with area beneath the receiver operating characteristic curve (AUC) values of 0.854, 0.855, and 0.853 for predicting healing

Sl No.	Paper Title	Technique Used	Database Used	Result	Key Insight
6	Wound Trajectory Prediction using ML Techniques	LSTM, Gradient Boosting	Chronic wound time-series data	Enhanced forecasting of wound healing time	Time-series ML models provide better healing trajectory predictions
7	Deep Learning Models for Predicting Wound Healing Outcomes	CNN, RNN	Diabetic wound images	High prediction accuracy for healing outcomes	Deep learning models perform better in handling image-based predictions
8	Hybrid Machine Learning Model for Chronic Wound Healing [34]	Hybrid CNN-RNN model	Clinical wound dataset	Improved detection of wound infection	Hybrid models enhance precision in predicting wound healing complications
9	AI-driven Wound Healing Assessment using CNN	Convolutional Neural Networks (CNN)	Wound images from public datasets	Improved classification of wound types	CNN-based models accurately classify wound stages and healing time
10	Wound Healing Prediction using Decision Trees and Random Forest	Decision Trees, Random Forest	Medical wound database	High accuracy in predicting wound healing time	Decision trees provide interpretable models for clinical applications

Table 2.3: Overview of machine learning techniques for wound healing prediction and assessment (Part 2).

times at 4, 8, and 12 weeks, respectively. Key predictors included days in treatment, wound depth, location, and area. Shapley Additive Explanations (SHAP) utilized to improve clinical judgments by identifying the most influential variables. This innovative approach draws attention to the possibilities of machine learning models in chronic wound care, offering a tool for clinicians to identify patients at danger of delayed healing and improve treatment outcomes.

A study [36] focused on evaluating the wound healing process in foot wounds used texture image analysis to establish an objective technique for evaluating wound progress. The study involved 77 digital images taken from 11 subjects over a 21-day period. These images were captured using an affordable digital camera under varying lighting conditions and intensity normalized. The wounds were automatically segmented using a segmentation system according to snakes, from which 56 texture features and four geometrical measures were removed in order to identify features indicative of wound healing. The study found that certain texture features, such as mean, contrast, roughness, and radial sum, increased over time, while others, including sum of squares variance, sum variance, entropy, and coarseness, decreased, reflecting the progression of wound healing. These features were significantly different at various time points, suggesting their potential in monitoring wound healing. In contrast, no significant differences were observed in the geometrical measures across the time points. The results imply that texture features could play an important role in wound healing assessment, potentially reducing

expert workload, offering standardization, lowering costs, and enhancing patient care quality. Although the method shows promise, further large-scale studies are necessary to validate its clinical application and explore additional texture and geometrical features for more accurate wound healing differentiation.

Amputation remains a critical, irreversible treatment for conditions such as trauma, peripheral vascular disease, diabetes, and cancer. While delaying amputation in favor of limb-sparing treatments can increase morbidity and mortality risks, early identification of patients requiring amputation can improve outcomes and reduce complications. Artificial intelligence (AI), particularly machine learning (ML), has shown promise in predicting amputation risk, using data within patient demographics, clinical factors, and outcomes. ML models are being increasingly used in the medical field, especially in predicting post-amputation outcomes. However, a gap remains in research regarding how ML can predict the need for amputation, which could improve early intervention and patient care. This systematic review examines the current state of ML applications in predicting amputation outcomes, highlighting its potential in enhancing clinical decision-making and improving patient quality of life. [37]

Amputation choices, especially for patients with critical limb ischemia, require careful consideration of the optimal level of amputation (LOA) to balance functionality and healing potential. Surgeons rely on clinical judgement, patient risk factors, and vascular studies, but up to one-third of amputations fail to heal, necessitating reamputation and increasing healthcare costs. Current technologies for predicting amputation wound healing, such as ankle-brachial indices and laser Doppler imaging, have limitations, including high variability and lack of broad adoption. A pilot study showed the possibilities of a machine-learning algorithm that integrates multispectral imaging of superficial tissues with patient clinical risk factors to forecast amputation wound healing. This system outperformed surgeon judgement, achieving 88% accuracy in predicting healing, compared to 56% accuracy for surgeons. By combining imaging and clinical data, this strategy could offer a more reliable, objective method for predicting wound healing, ultimately improving patient outcomes and reducing healthcare costs. However, further validation and larger studies are needed to confirm its clinical applicability.[38]

A study [39] developed and validated a machine learning model to forecast whether chronic wounds, regardless of etiology, would heal within 12 weeks. The evaluation used electronic health record data from 532 wound concern clinics across the United States, covering 261,398 patients and 620,356 unique wounds. The representation incorporated patient demographics, comorbidities, wound characteristics at initial presentation, and changes in wound dimensions over time, with the latter proving to be the greatest influential predictor. The model illustrated high predictive accuracy, achieving an area beneath the Receiver Operating Characteristic curve (ROC) of 0.9 and 0.92 after 4 and 5 weeks of treatment, respectively. The findings suggest that such a machine learning model, if integrated into real-world care, could help identify chronic wounds at a danger of non-healing early in the treatment process, enabling more effective and efficient treatment decisions.

A study [40] aimed to address the significant inter- and intrarater variability in manual tissue segmentation and quantification of chronic wounds by utilizing deep learning models to identify tissues objectively. Using the dataset of 58 anonymized chronic wound images from Swift Medical's Wound Database, the study contrasted the outcomes of five wound clinicians in labeling four tissue types (epithelial, granulation, slough, and eschar) and evaluated the ability of deep convolutional neural networks (CNNs) to automate this process. The models for deep learning were trained on over 465,000 image-label pairs, achieving robust performance with

high accuracy in wound and tissue segmentation, particularly for easier-to-label tissue types. Significant variability was discovered by the study among clinicians, especially in identifying epithelial tissue, while the deep learning models provided consistent, objective results. The models were designed to run in real-time on smartphones, enabling widespread deployment in clinical practice. The ability to objectively assess tissue composition could improve the accuracy of wound assessments, guide treatment decisions, and potentially enhance chronic wound healing outcomes. This framework offers a promising step towards reducing subjectivity in wound care and improving the efficiency and consistency of wound management.

The study [41] presents a machine-learning-based approach for classifying wounds and normal skin using dielectric spectroscopy. The authors measured the dielectric constants of normal skin and various types of wounds from living mice across a wide frequency range (10 MHz to 20 GHz) using a commercial network analyzer. The data were processed with a Data Dimensionality Reduction technique to identify the optimal frequency range for wound dielectric spectroscopy. The results indicated that wounds could be differentiated from normal skin effectively within the 1 to 2 GHz frequency range, which lowers the need for high-bandwidth dielectric spectroscopy sensors. By employing supervised learning classification techniques, the study demonstrated near 100% accuracy in classifying different tissue types, highlighting the potential of this approach for precise wound classification.

Chronic wounds, especially diabetic foot ulcers, pose significant challenges in healthcare due to their impact on quality of life and the need for continuous monitoring of healing progress. Image-based wound analysis is a promising method for objectively assessing wound status by quantifying features related to healing. However, the high heterogeneity of wound types and varying imaging conditions complicate robust wound segmentation. To address this, the authors introduce Detect-and-Segment (DS), a deep learning technique that enhances segmentation performance by first detecting the wound position and isolating it from the surrounding background before computing a segmentation map. The DS approach was tested on a diabetic foot ulcer dataset and showed a Matthews' correlation coefficient (MCC) improvement from 0.29 to 0.85 compared to traditional segmentation methods. Additionally, when evaluated on four independent datasets with a greater variety of wound types from different body locations, the mean MCC increased from 0.17 to 0.85. Notably, the DS method allowed for training segmentation models with up to 90% less data without compromising performance. This approach, which combines wound detection and segmentation into a single deep learning model, demonstrates high generalization capabilities and represents a significant step towards automating chronic wound analysis, reducing the effort required for management.[42]

All things considered, these computational methods provide substantial progress in the field of wound healing by allowing for precise, individualized, highly predictive, and data-driven treatment plans. Incorporating these techniques into the investigation of *Monopterus albus* healing may provide fresh perspectives on regenerative medicine and the enhancement of wound care interventions. Table 2.2, 2.3, 2.4 and 2.5 gives the Overview of Machine Learning Techniques used in Wound Healing Studies.

Sl No.	Paper Title	Technique Used	Database Used	Result	Key Insight
11	Wound image evaluation with machine learning [17]	Neural Networks, Random forest and SVM	Photographs of PUs of patients with home-care assistance	Comparison between model has been applied and all show different efficiency and significant difference	Neural networks, random forest models and support vector machines (overall accuracy on a testing set [95% CI])
12	Machine learning analysis of multispectral imaging and clinical risk factors to predict amputation wound healing [33]	A modified machine learning architecture	A total of 22 patients undergoing 25 amputations (10 toe, five transmetatarsal, eight below-knee, and two above-knee amputations) were enrolled. Eleven amputations (44%)	The machine learning algorithm had 91% sensitivity and 86% specificity for prediction of non-healing amputation sites (area under curve, 0.89).	A novel wound imaging system combining multispectral imaging and machine learning analysis demonstrated high accuracy in predicting primary amputation wound healing, offering potential to reduce reoperation rates and healthcare costs.
13	Predicting Chronic Wound Healing Time Using Machine Learning [35]	Decision Tree	Dataset of over 1.2 million wounds with 187 covariates	The 4-, 8-, and 12-week gradient-boosted decision tree models achieved AUC's of 0.854, 0.855, and 0.853, respectively	Machine learning models can accurately predict chronic wound healing time using EHR data. SHAP values can give insight into how patient-specific variables influenced predictions.
14	Evaluation of wound healing process based on texture image analysis [36]	Texture analysis	A total of 77 color digital images from 11 different subjects with foot wounds	The evaluation of wound healing rate based on texture analysis	Found statistical significant differences in some texture features, specifically, mean, contrast, roughness and radial sum, SSV, SV, SA, entropy, coarseness, H2 and H4.
15	Predicting amputation using machine learning: A systematic review [37]	Machine Learning	3572 articles	No specified result mentioned	The included studies developed and validated ML models from a total of 2,261,790 patients. Extensive heterogeneity between the studies across study objectives, ML models, data set features, varying subgroup analyses, and performance metrics of included studies precluded a meta-analysis of such findings.

Table 2.4: Overview of machine learning techniques for wound healing prediction and assessment (Part 3).

Sl No.	Paper Title	Technique Used	Database Used	Result	Key Insight
16	Machine Learning Analysis of Multispectral Imaging and Clinical Risk Factors to Predict Amputation Wound Healing [38]	Machine Learning with Multispectral Imaging	Not specified (clinical-study)	88% accuracy, 91% sensitivity, 86% specificity	combining multispectral imaging with patient clinical risk factors via machine learning significantly improves the accuracy of predicting amputation wound healing.
17	Prediction of Healing Trajectory of Chronic Wounds Using a Machine Learning Approach [39]	Machine Learning	261,398 patients and 620,356 unique wounds	The model demonstrated high predictive accuracy, achieving an area under the receiver operating characteristic curve (AUC) of 0.9 and 0.92 after 4 and 5 weeks of treatment, respectively	A machine learning model can identify chronic wounds at risk of not healing by week 12 with high accuracy in the early weeks of treatment.
18	Fully automated wound tissue segmentation using deep learning on mobile devices: Cohort study [40]	Convolutional neural networks (CNNs)	58 anonymized chronic wound images from Swift Medical's Wound Database	Deep learning models achieved high accuracy in wound and tissue segmentation with robust performance across various skin tones and wound types.	Deep learning models can provide objective, consistent tissue identification and quantification, overcoming clinician variability in chronic wound assessments.
19	Towards a Machine-Learning-Assisted Dielectric Sensing Platform for Point-of-Care Wound Monitoring [41]	Dielectric spectroscopy, machine learning	Wound data from living mice, commercial network analyzer	Wounds can be classified with near 100% accuracy using data from 1 to 2 GHz frequency range	The study identified an optimal frequency range (1-2 GHz) for distinguishing wounds from normal skin, reducing bandwidth requirements for dielectric spectroscopy sensors.
20	Detect-and-segment: A deep learning approach to automate wound image segmentation [42]	Deep Learning Approach (Detect-and-Segment)	SwissWOU, DFUC, Medtec, SIH, FUSC	MCC improved from 0.29 to 0.85 on DFU dataset; DS generalizes well on independent datasets with improved performance.	DS approach achieves robust wound segmentation with high generalization, using up to 90% less training data.

Table 2.5: Overview of machine learning techniques for wound healing prediction and assessment (Part 4).

# Chapter 3

## Dataset and Preprocessing

This chapter details the dataset used for the project, including its origin, structure, preprocessing steps, and augmentation techniques. The dataset comprises monoscopic images capturing various stages of wound healing in *Monopterus cuchia*. Proper preprocessing is essential to ensure consistent input to the model and improve classification performance.

### 3.1 Dataset Description

The dataset was collected as part of an experimental study on wound healing in *Monopterus cuchia* from the Department of Zoology, Gauhati University. It consists of images captured over several days, representing both normal healing processes and recovery with external treatment (e.g., Retinoic Acid). The images are grayscale (monoscopic) with varying resolutions. The dataset is organized into two main folders:

- **Normal Healing:** Contains subfolders for different healing stages. [43]
- **Retinoic Acid Treated:** Also divided into subfolders representing time-based wound recovery. [44]

Each subfolder corresponds to a class label and contains a collection of images for that healing stage shown in Figure 3.1. The total quantity of images per class is variable, necessitating careful augmentation for class balance.

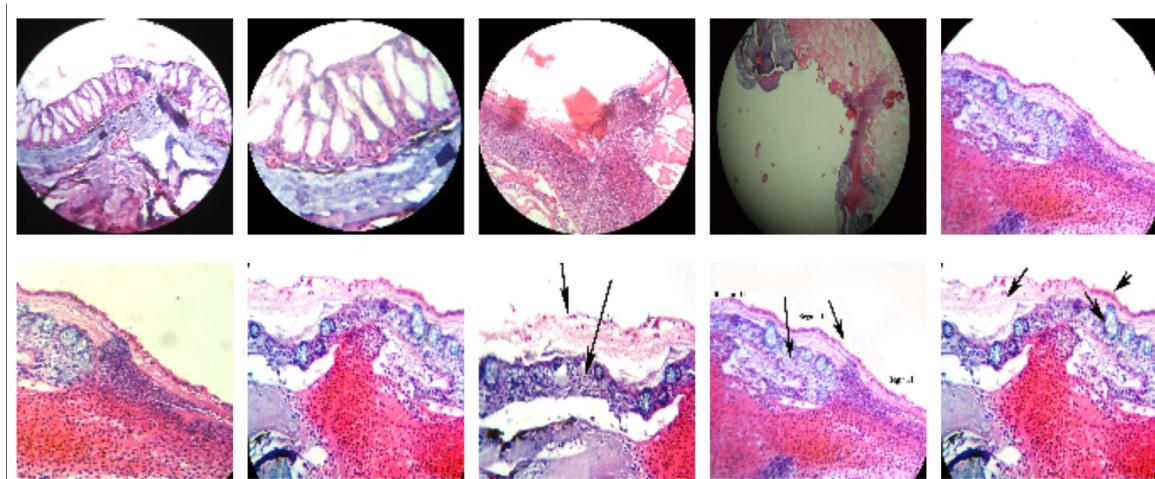


Figure 3.1: Dataset images of different healing stages

## 3.2 Image Preprocessing

To prepare the images for model training, several preprocessing steps were applied:

- **Resizing:** All images were resized to  $224 \times 224$  pixels to match the input size of MobileNetV2.
- **Normalization:** Pixel values were scaled to the range  $[0, 1]$  by dividing by 255.
- **Channel Replication:** Since MobileNetV2 requires RGB input, grayscale images were converted to three channels by duplicating the single channel.

These steps ensure that the dataset conforms to the input requirements of the pretrained model while preserving the wound features.

## 3.3 Data Augmentation

To rectify the disparity in classes and enhance the resilience of the model, data augmentation was used `ImageDataGenerator` from TensorFlow. The following transformations were employed:

- Rotation:  $\pm 20^\circ$
- Width and height shift: up to 20%
- Zoom: up to 15%
- Shear: up to 15%
- Horizontal flipping

After applying above data augmentation techniques, the count of number of images has increased and the images are shown in Figure 3.2. A validation split of 20% was applied to separate training and validation datasets while ensuring class stratification.

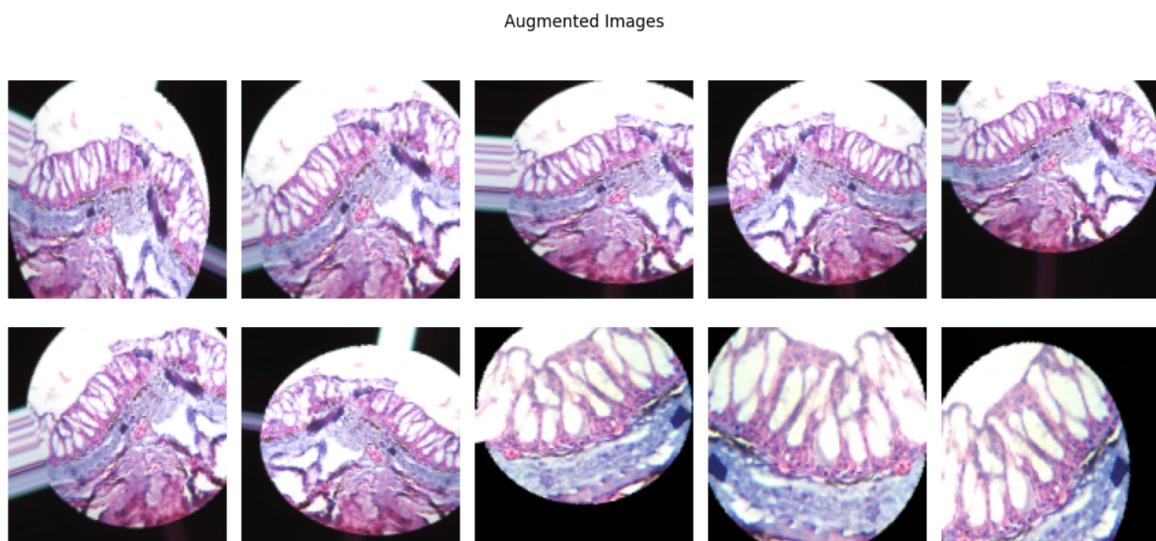


Figure 3.2: Images from the dataset after augmentation

## 3.4 Dataset Availability

The datasets created and examined for this work are publicly accessible at Zenodo. This includes the data for normal wound healing in *Monopterus cuchia* [43] and the data pertaining to retinoic acid-based wound healing in the same species [44]. The public availability of this data ensures transparency and allows for further research and validation of the findings presented in this report.

## 3.5 Class Distribution and Visualization

The original dataset had varying image counts across classes shown in figure 3.3. The distribution after augmentation was visualized using bar plots 3.4 to understand class imbalance.

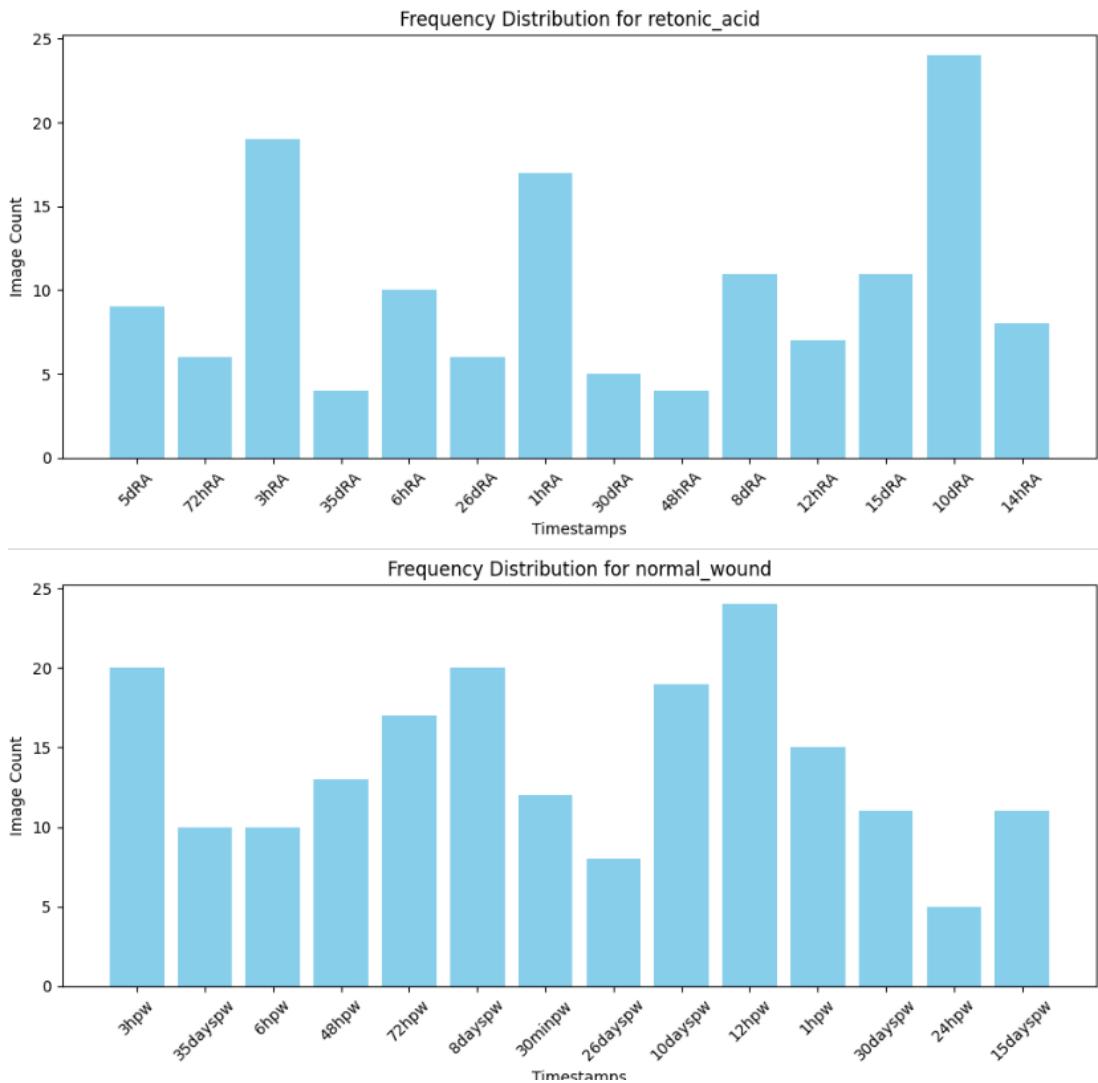


Figure 3.3: Frequency Distribution of the dataset

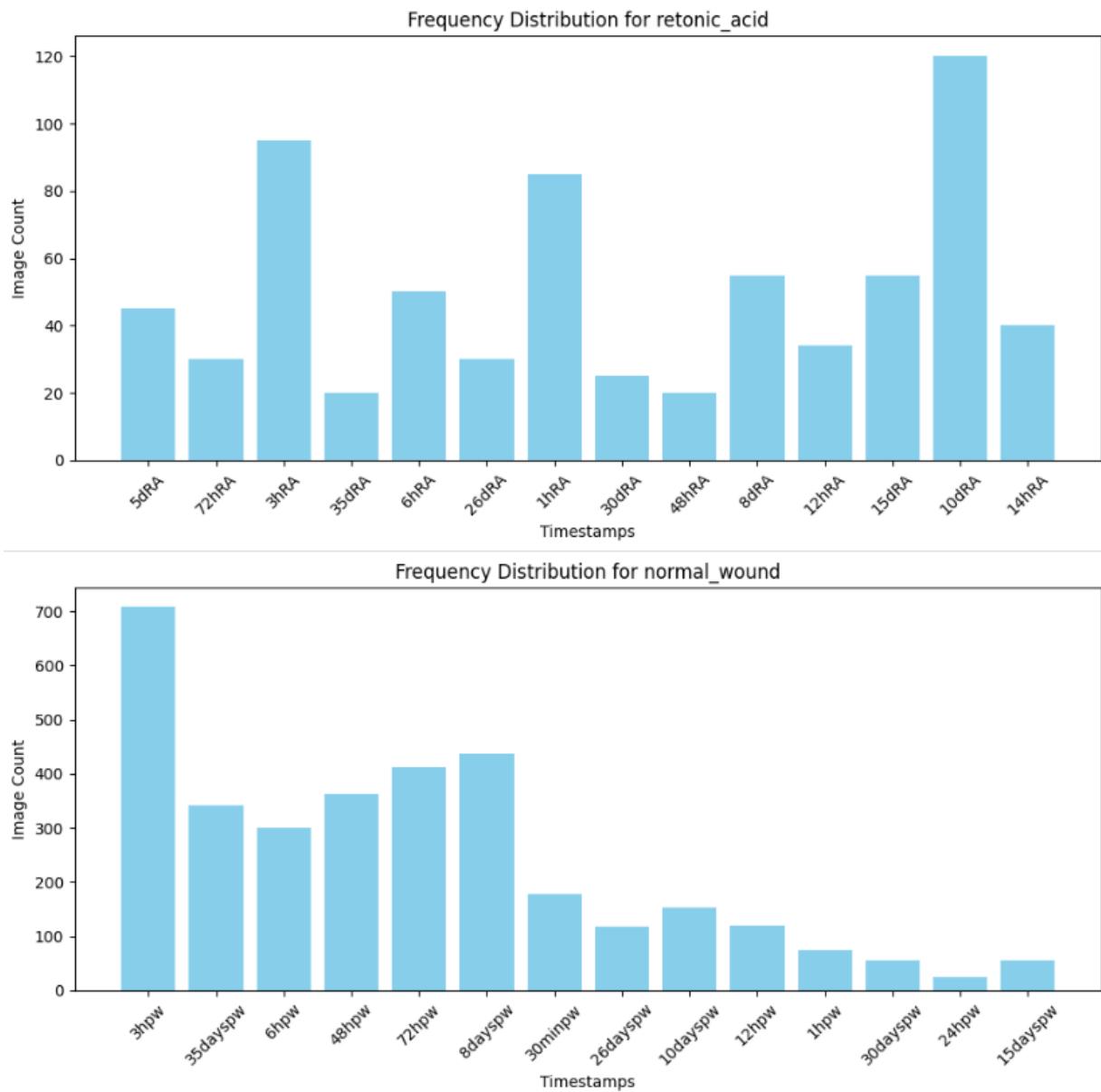


Figure 3.4: Frequency Distribution of dataset after augmentation

# Chapter 4

## Methodology

The methodology of this project focuses on developing an automated system for classifying wound healing stages in *Monopterus cuchia*, using both supervised learning (deep learning classification) and unsupervised learning (clustering). This multi-step approach enables not only accurate classification but also exploratory pattern discovery. The methodology involves dataset preparation, preprocessing, model design using transfer learning, fine-tuning and clustering.

### 4.1 Overview of the Methodology

The core components of the methodology are as follows:

- Dataset Collection
- Dataset Preparation and Augmentation
- Clustering
  - Feature Extraction
  - Dimensionality reduction
  - K-Means clustering and Analysis
- Deep Learning Model Design for Classification
  - Hyperparameter Tuning
  - Transfer Learning and Fine-Tuning
  - Model Training
- Performance Evaluation

### 4.2 Proposed Methodology

The methodology adopted in this project is a dual-path pipeline combining supervised deep learning and unsupervised learning to analyze wound healing stages in *Monopterus cuchia*. The entire workflow is depicted in Figure 4.1, comprising data preparation, classification using a fine-tuned MobileNetV2 model (CuchiaNet), and clustering using K-Means on extracted features.

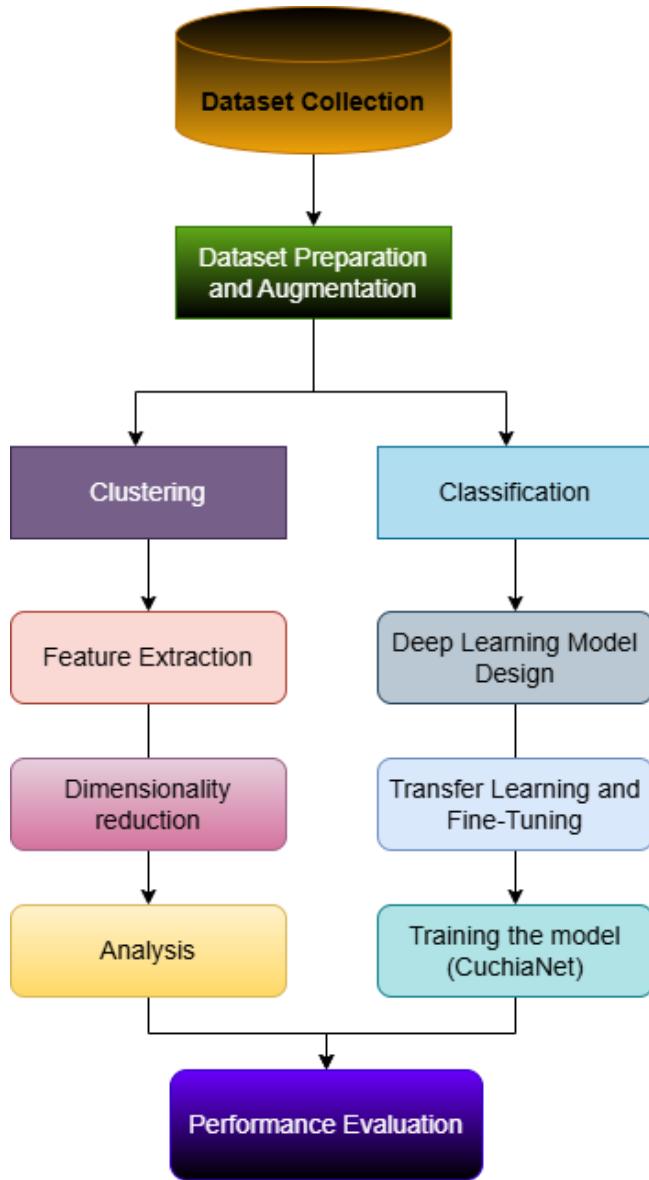


Figure 4.1: Proposed Methodology Workflow

### 4.3 Dataset Collection and Preprocessing

The dataset was collected as part of an experimental study on wound healing in *Monopterus cuchia* from the Department of Zoology, Gauhati University. It consists of images captured over several days, representing both normal healing processes and recovery with external treatment (e.g., Retinoic Acid). The images are grayscale (monoscopic) with varying resolutions a significant challenges on training deep learning models.

To defeat this limitation and improve model robustness, an extensive data augmentation pipeline was applied using the `ImageDataGenerator` class from TensorFlow's Keras API. Augmentation techniques included:

- **Random Rotations:** Images were rotated within a range of degrees to account for orientation variability during image capture.
- **Horizontal and Vertical Flipping:** Simulated mirrored wound appearances to diversify spatial perspectives.

- **Width and Height Shifts:** Applied random translations to mimic misaligned captures and increase positional tolerance.
- **Zooming:** Random zoom levels were introduced to ensure scale invariance.
- **Shearing:** Slight shearing distortions were added to simulate natural shape deformations.

These augmentations not only expanded the effective size of the dataset but also improved the model's ability to generalize across variations that might be encountered in real-world scenarios. All images were resized to a standard input shape (e.g., 224x224 pixels) and normalized to scale pixel values between 0 and 1.

This preprocessing phase played a crucial role in preparing the dataset for both supervised classification and unsupervised clustering tasks, ensuring that the input to the neural network was both varied and representative of the biological variability observed during wound recovery.

## 4.4 Clustering Pathway

The unsupervised clustering pathway in the study is created to identify hidden patterns in wound healing progression without relying on predefined class labels. The use of clustering complements the supervised classification approach by offering an exploratory view of the dataset, potentially revealing sub-stages or variations in healing behavior. The workflow of the clustering pathway is shown in Figure 4.2

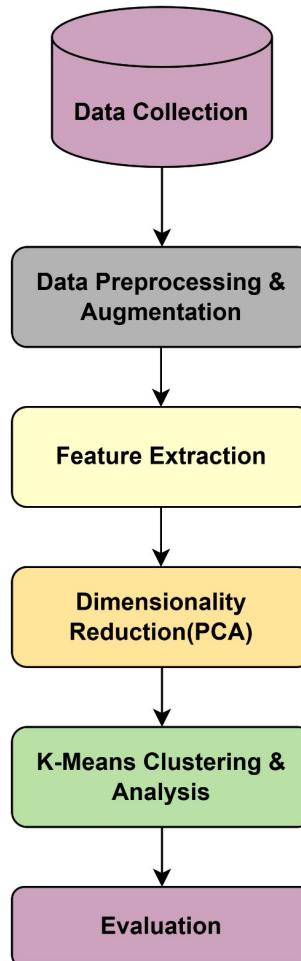


Figure 4.2: The workflow diagram of Clustering Pathway

#### 4.4.1 Data Preparation

Images from two distinct healing conditions normal wound healing and Retinoic Acid-assisted healing were collected and structured under labeled folders. However, unlike the classification pathway, during clustering, these labels weren't utilized. The dataset was programmatically loaded by traversing directory paths, where each image was resized to a fixed dimension of  $128 \times 128$  pixels to ensure uniformity in feature dimensions.

#### 4.4.2 Image Flattening and Feature Extraction

To convert image data into a format suitable for clustering algorithms, each image was flattened into a one-dimensional vector. This process transformed each  $128 \times 128 \times 3$  RGB image into a feature vector of length 49,152.

#### 4.4.3 Dimensionality Reduction Using PCA

Given the high dimensionality of image vectors, Principal Component Analysis (PCA) was applied to reduce the feature space while maintaining crucial data. PCA helped in identifying the directions of maximum variance and compressed the data into fewer dimensions (typically 2 or 3 components) for easier visualization and clustering. The explained variance ratio guided the selection of components to retain meaningful information.

#### 4.4.4 K-Means Clustering

K-Means clustering was employed to partition the image data into distinct clusters. The number of clusters was initially set to two, aligning with the known conditions, although these labels were not utilized at the time of training. The K-Means algorithm iteratively assigned each image to the nearest centroid and recalculated centroids until convergence. The ‘silhouette score’ was computed to evaluate the compactness and separation of the resulting clusters, providing a quantitative measure of clustering performance.

#### 4.4.5 Cluster Visualization and Interpretation

To visualize the clustering results, PCA-transformed features were plotted by scatter plots, where each point represented an image in 2D space. Points were color-coded based on their assigned cluster. This visualization enabled intuitive understanding of how well the model distinguished between different healing conditions based solely on image content.

### 4.5 Classification Pathway

The classification pathway forms a central component of this research, aiming to automatically identify and differentiate wound healing stages in *Monopterus cuchia*. Given the limited size and variability of the dataset, a transfer learning method was used to leverage pre-trained knowledge and adapt it effectively to the specific classification task. This section outlines the architecture design, hyperparameter optimization, training strategy, and performance enhancements utilized in the classification model. The workflow of the classification pathway is shown in Figure 4.3

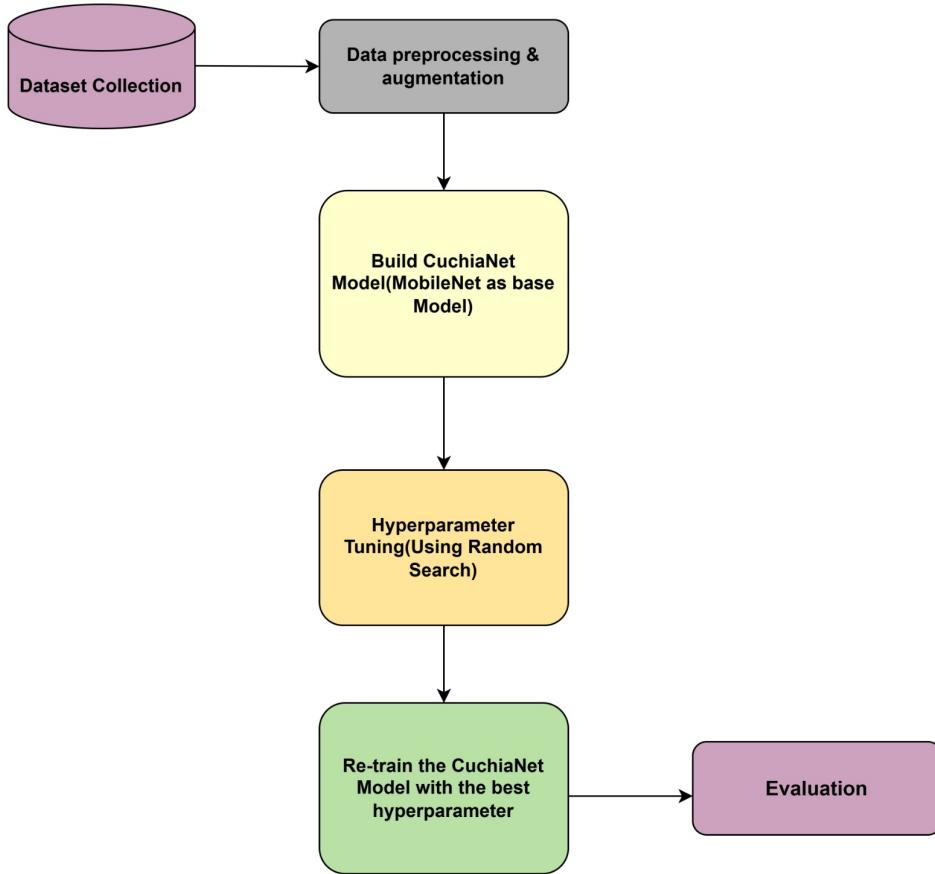


Figure 4.3: The workflow diagram of Classification Pathway

#### 4.5.1 Transfer Learning with MobileNetV2

To build an efficient and accurate classification model, the MobileNetV2 architecture was selected as the base. MobileNetV2 is a light in weight convolutional neural network architecture renowned for its high accuracy and low computational demand. It was originally trained on the ImageNet dataset and has shown effectiveness in multiple vision-based tasks[45].

The pre-trained base model was loaded without its top classification layers, allowing us to append a custom classifier suited to the wound healing dataset. The convolutional base was initially frozen to retain its pre-learned visual features, especially useful because of the smaller size of the dataset. This method allowed the model to start with a solid base of general image features, thereby reducing the time and data needed for training.

#### 4.5.2 Model Architecture Design

The final classification model, referred to as **CuchiaNet**, was composed of the following components:

- **Base Model:** Pre-trained MobileNetV2 without its top layers.
- **Global Average Pooling Layer:** Replaces dense fully connected layers to reduce overfitting and preserve spatial features.
- **Dense Layer:** Fully connected layer with optimized units (determined via hyperparameter tuning), using ReLU activation.

- **Dropout Layer:** Introduced with a tuned dropout rate to prevent overfitting and ensure robust learning.
- **Output Layer:** A final Dense layer having softmax activation, equal in size to the quantity of target classes, to perform multi-class classification.

All input images has been preprocessed to the size of  $224 \times 224$  pixels and normalized for consistent input distribution.

#### 4.5.3 Hyperparameter Optimization

To achieve optimal model performance, **Random Search** method was used for hyperparameter tuning. Instead of relying on trial-and-error or manual adjustments, this method randomly selected combinations of hyperparameters from specified ranges and evaluated model performance using validation accuracy.

The hyperparameters considered for tuning included:

- Number of units in the Dense layer
- Dropout rate in the Dropout layer
- Learning rate for the optimizer

After extensive search and evaluation, the most effective combination of hyperparameters was identified. This automated tuning not only improved the overall correctness of the model but also ensured a balance between underfitting and overfitting, making the model more adaptable to unseen data.

#### 4.5.4 Training Strategy

The model was trained in two carefully designed phases:

##### Phase I: Feature Extraction

In this phase, the base MobileNetV2 model was kept frozen to preserve the learned visual representations. Only the newly added classification layers were trained. This approach ensured that the model first adapted the classifier to the domain-specific data while building upon generalized features extracted from the base model[46].

##### Phase II: Fine-Tuning

After the initial training, the top 30 layers of the base model were unfrozen to allow selective fine-tuning of deeper layers. This step helped the model learn subtle, domain-specific patterns in the wound images. A smaller learning rate was applied during this phase to avoid drastic updates and preserve useful features already learned during the first phase[47].

#### 4.5.5 Training Optimization Techniques

To ensure stable and efficient training, two key callback functions were implemented:

- **EarlyStopping:** Training was automatically stoped if the validation loss did not show any improvement over five continuous epochs. This prevented overfitting and unnecessary computation[48].

- **ModelCheckpoint:** The model achieving the highest validation accuracy during training was saved. This ensured that the best-performing version of the model was retained for final evaluation, avoiding reliance on the last training epoch.

These callbacks significantly improved model generalization and reduced the risk of over-training.

This classification methodology effectively combined the strengths of pre-trained models, systematic hyperparameter tuning, and smart training strategies. By utilizing MobileNetV2 with a custom top layer and optimizing it in two stages—feature extraction followed by fine-tuning—the model was able to adapt to the complexities of biological image data with limited examples. The careful integration of dropout, early stopping, and checkpoint mechanisms ensured reliable convergence and maximized predictive performance.

## 4.6 Performance Evaluation

To validate the effectiveness of both clustering and classification strategies, a comprehensive evaluation framework was used.

### Clustering Metrics

The unsupervised clustering results were evaluated using:

- **Silhouette Score:** Measured how similar each point was to its own cluster compared to others, with a score of 0.36 indicating well-separated clusters.
- **Intra-cluster Similarity:** Examined how tightly data points grouped within a cluster.
- **Visual Validation:** PCA-based scatter plots of clustered data were visually inspected to assess grouping quality and biological plausibility.

### Classification Metrics

The deep learning classifier's performance was assessed using the following metrics:

- **Accuracy:** Overall proportion of correctly classified images.
- **Confusion Matrix:** Visual breakdown of predicted versus actual classes to identify class-wise performance.
- **Precision, Recall, F1-Score:** These metrics provided a balanced assessment of the model's capacity to correctly identify each class, especially important for imbalanced datasets.
- **ROC-AUC:** For binary classification tasks (e.g., treated vs. untreated), the ROC curve and AUC quantified the model's ability to distinguish between classes across threshold values.

The combined evaluation confirmed that the model pipeline was capable not only of categorizing known classes but also of discovering hidden structures within the data.

# Chapter 5

## Implementation

This chapter elaborates on the systematic development of the deep learning-based image classification framework employed to analyze the wound healing process in *Monopterus cuchia*. The focus was not just to reach high classification accuracy moreover to ensure a resilient, interpretable, and reproducible workflow suitable in research and practical use.

### 5.1 Development Setup

The entire pipeline is structured using Python 3.10 in a Google Colab environment, leveraging GPU acceleration (Tesla T4) to optimize training speed. Libraries such as TensorFlow and Keras were applied in model building, while NumPy, Pandas, Matplotlib, Seaborn, and scikit-learn facilitated data handling, visualization, and evaluation.

#### Key Environment Details:

- **Platform:** Google Colab (GPU-enabled)
- **Primary Libraries:** TensorFlow 2.x, Keras, scikit-learn, matplotlib, seaborn

### 5.2 Data Loading and Augmentation

The input dataset comprised images categorized into multiple subfolders, each representing a distinct wound healing class. Given the variability of biological images and the limited data available, augmentation became a vital part of the training process[49]. The Keras `ImageDataGenerator` was employed to apply real-time augmentation such as:

- Random horizontal flips
- Rotation (up to 20 degrees)
- Width and height shifts (up to 20%)
- Shearing and zooming

These augmentations simulate realistic variations in biological images, helping the model to generalize better without overfitting.

Listing 5.1: Image Augmentation Setup

```
train_datagen = ImageDataGenerator(  
    rescale=1./255,
```

```

        rotation_range=20,
        zoom_range=0.15,
        width_shift_range=0.2,
        height_shift_range=0.2,
        shear_range=0.15,
        horizontal_flip=True,
        validation_split=0.2
    )

```

Images were resized to  $224 \times 224$  pixels to match the input requirements of MobileNetV2. The dataset was split in an 80-20 ratio for training and validation, respectively, using the ‘validation\_split’ argument, thus maintaining consistency without manual splitting.

## 5.3 Clustering Pathway Implementation

This section outlines the implementation details of the unsupervised clustering component used in the analysis of wound healing in *Monopterus cuchia*. The main objective of this pathway was to discover inherent structures and patterns in the wound images without relying on labeled data. The entire pipeline was executed using Python, leveraging libraries such as NumPy, OpenCV, Matplotlib, scikit-learn, and Seaborn for image handling, data transformation, clustering, and visualization[50].

### 5.3.1 Dataset Loading and Structuring

The dataset used in the clustering pathway consisted of images sourced from two distinct experimental groups: one exhibiting natural healing and another influenced by Retinoic Acid. Every picture was taken from respective directories using standard file traversal methods. Each image was resized to a fixed dimension of  $128 \times 128$  pixels to maintain consistency in shape and reduce computational complexity. This resizing also helped reduce noise and memory overhead without compromising essential visual information.

### 5.3.2 Image Vectorization and Preprocessing

For clustering algorithms like K-Means, the data must be represented numerically in a tabular form. Hence, each image was flattened from a 3D array ( $128 \times 128 \times 3$ ) into a 1D feature vector of 49,152 pixels. This process was performed using NumPy array reshaping and converted all images into a matrix of shape  $(N, 49152)$ , where  $N$  accounts to the number of images.

The resulting feature matrix was normalized by scaling pixel intensities between 0 and 1, improving numerical stability for subsequent transformations and clustering.

### 5.3.3 Dimensionality Reduction using PCA

Because of the huge number of dimensions of image vectors, Principal Component Analysis (PCA) was applied to reduce redundancy and capture the most significant features[51]. Initially, PCA was performed with automatic component selection, followed by variance ratio analysis to determine the optimal number of components. A 2-component PCA was finally adopted for visualization and clustering, as it retained the dominant variance and allowed for intuitive scatter plot analysis.

This step effectively transformed each image into a 2D representation where distances between points reflected visual similarity.

### 5.3.4 K-Means Clustering

The reduced PCA-transformed data was fed into the K-Means clustering algorithm to group similar images. The number of clusters,  $k$ , was empirically set to 2, representing the expected healing types. K-Means initialization was performed using the `k-means++` strategy to ensure more stable convergence[52]

The clustering process involved multiple iterations where centroids were updated and data points reassigned until the intra-cluster variance minimized. After convergence, each image is labeled with a cluster index (0 or 1), and the allocation of cluster assignments was examined.

### 5.3.5 Silhouette Score Evaluation

To quantify the clustering quality, the silhouette score was computed. This metric ranges from -1 to 1 and reflects how well each and every image fits in its designated cluster compared to other clusters. A higher silhouette score indicates better-defined clusters. This provided a numeric validation of the separability between normal and RA-assisted healing images, without using class labels[53].

### 5.3.6 Visualization of Clusters

For interpretability, the 2D PCA-reduced features were plotted using scatter plots. Each point in the plot represented an image, and its color indicated its cluster assignment. This visual representation offered a compelling overview of how well the algorithm could differentiate between the wound healing conditions.

### 5.3.7 Label Mapping for Interpretation

Although clustering was unsupervised, the predicted clusters were compared against the known folder labels for post-hoc validation. This allowed qualitative assessment of how well K-Means captured the inherent structure in the data, and whether the wounds under different treatments exhibited visually separable characteristics.

## 5.4 Classification Pathway Implementation

The supervised learning component of this project focuses on classifying wound healing images into each of their stages using a deep learning model. This was achieved by leveraging transfer learning through MobileNetV2, a lightweight and a strong convolutional neural network (CNN) pre-trained on ImageNet. This base model was adapted and enhanced into a custom architecture, referred to as **CuchiaNet**, tailored specifically for classifying the healing stages of *Monopterus cuchia*.

### 5.4.1 Model Architecture

The proposed architecture (shown in Figure 5.1) was constructed by stacking a custom classification head on top of the pre-trained MobileNetV2 base (with the top layers removed). This head comprised the following components:

- **GlobalAveragePooling2D:** Reduced spatial dimensions while preserving learned feature maps.

- **Dense Layer:** Fully connected layer consisting 128 units and ReLU activation, whose size was selected using Random Search hyperparameter tuning.
- **Dropout Layer:** Applied a dropout rate of 0.2 (also selected via Random Search) to prevent overfitting.
- **Output Layer:** At last a Dense layer with softmax activation to classify images into multiple healing stages based on the number of classes.

The architecture effectively balances complexity and efficiency, making it suitable for deployment in resource-constrained environments.

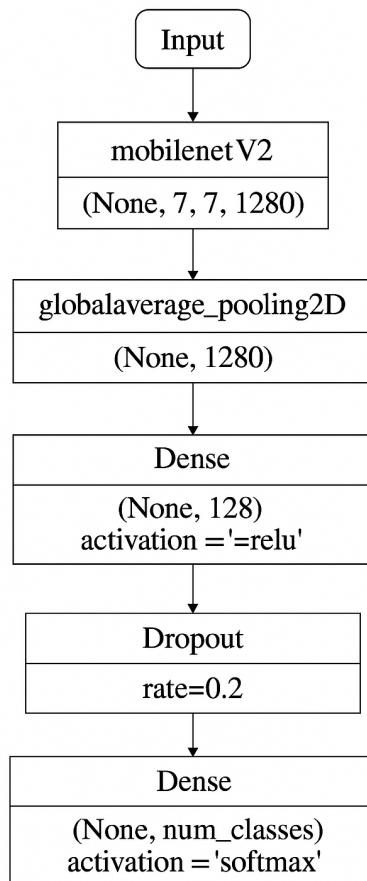


Figure 5.1: The proposed architecture of CuchiaNet

#### 5.4.2 Hyperparameter Tuning

In order to maximize the model's performance, a **Random Search**[54] strategy is used for hyperparameter tuning. This involved randomly sampling combinations of critical parameters and evaluating model performance on a validation set. The following hyperparameters were tuned:

- **Dense Layer Units:** Number of neurons in the Dense layer (best: 128)
- **Dropout Rate:** Prevents overfitting by randomly dropping neurons (best: 0.2)
- **Learning Rate:** Controls how quickly Weights are changed throughout training phase. (best: 0.001)

This methodical tweaking procedure greatly enhanced generalization and convergence speed compared to arbitrary trial-and-error approaches.

### 5.4.3 Training Phase I – Feature Extraction

During the first training stage, every layer of the MobileNetV2 base model were frozen, and only the custom classification head was trained. This phase aimed to utilize pre-learned low-level visual features from ImageNet while customizing the model for wound classification.

To ensure robust training, the following callback mechanisms were implemented:

- **EarlyStopping:** Monitored the validation loss and halted training if no improvement was observed after 5 epochs, reducing overfitting.
- **ModelCheckpoint:** Continuously saved the best model weights based on validation accuracy, ensuring optimal performance retention.

Listing 5.2: Training Phase I with EarlyStopping and ModelCheckpoint

```
early_stop = EarlyStopping(monitor='val_loss',
                           patience=5,
                           restore_best_weights=True)
checkpoint = ModelCheckpoint('best_model.keras',
                             monitor='val_accuracy',
                             save_best_only=True)

history = model.fit(
    train_generator,
    epochs=20,
    validation_data=val_generator,
    callbacks=[early_stop, checkpoint]
)
```

### 5.4.4 Training Phase II – Fine-Tuning

Once the classification head was trained, the model underwent fine-tuning by unfreezing the last 30 layers of the base MobileNetV2 model. This allowed deeper features to adjust to the nuances of wound healing patterns in the dataset, further boosting accuracy.

During fine-tuning, the learning rate was reduced to  $1 \times 10^{-5}$  to avoid drastic updates that could degrade previously learned weights. The same callback strategies—`EarlyStopping` and `ModelCheckpoint`—were reused to maintain training stability and model reliability.

Listing 5.3: Fine-Tuning the Last 30 Layers of MobileNetV2

```
for layer in base_model.layers[:-30]:
    layer.trainable = False
base_model.trainable = True

model.compile(optimizer=Adam(learning_rate=1e-5),
              loss='categorical_crossentropy',
              metrics=['accuracy'])

history_fine = model.fit(
```

```

        train_generator ,
        epochs=20,
        validation_data=val_generator ,
        callbacks=[early_stop , checkpoint]
)

```

## 5.5 Graphical User Interface Deployment

To translate the trained CuchiaNet model into an accessible tool for researchers and field technicians, a lightweight web application is created through **Streamlit**. Streamlit was chosen due to its low boiler-plate code, rapid prototyping capabilities, and seamless integration with Python-based deep-learning workflows. The resulting application enables users to upload microscopic wound images and receive near-instantaneous predictions of the healing stage.

### 5.5.1 Interface Design

Figure 5.2 gives a summary of the application’s layout. The interface is made of two primary regions:

- **Sidebar (left panel).** This panel introduces the project, outlines the two wound-healing categories (*Normal Repaired* vs. *Retinoic Acid Repaired*), and displays model provenance information (MobileNetV2 backbone,  $\geq 90\%$  validation accuracy). It also credits the author and briefly states the academic purpose of the tool.
- **Main panel (central workspace).** Users interact with the core functionality here. A title header and concise instructions guide the user to upload a JPEG or PNG image. Once an image is selected, it is rendered in the browser (maintaining aspect ratio) and a progress spinner indicates that the model is performing inference. Final results are presented as:
  1. A bold, human-readable class label automatically converted into sentence case.
  2. A confidence score expressed as a percentage with two-decimal precision.

### 5.5.2 Backend Workflow

Upon image upload, the following sequence is executed server-side:

1. The image is resized to  $224 \times 224$  pixels to match MobileNetV2’s expected input dimensions.
2. If an alpha channel is present, it is discarded to maintain a three-channel RGB tensor.
3. Pixel values are scaled and centered via `mobilenet`’s built-in `preprocess_input`.
4. The pre-processed tensor is forwarded to the saved `final_cuchianet_model.keras` file, which outputs class-probability logits.
5. The highest-probability class is mapped to a human-friendly label defined in `class_names`, and the corresponding confidence is computed.

### 5.5.3 User Experience and Responsiveness

Inference on a consumer-grade CPU typically completes in under one second, offering a smooth user experience. Streamlit's native caching ensures that repetitive image renders and static assets (e.g., sidebar text) load instantly, while the progress spinner provides immediate feedback during prediction. The minimalistic aesthetic—soft colours, rounded icons, and a single primary action (image upload)—reduces cognitive load and supports rapid decision-making in laboratory contexts.

### 5.5.4 Snapshots of the Web-application

Figures 5.3 and 5.4 illustrate the interface in two states: before image upload and after prediction. These visuals confirm that the tool preserves usability even on modest screen resolutions, an important consideration for field deployment.

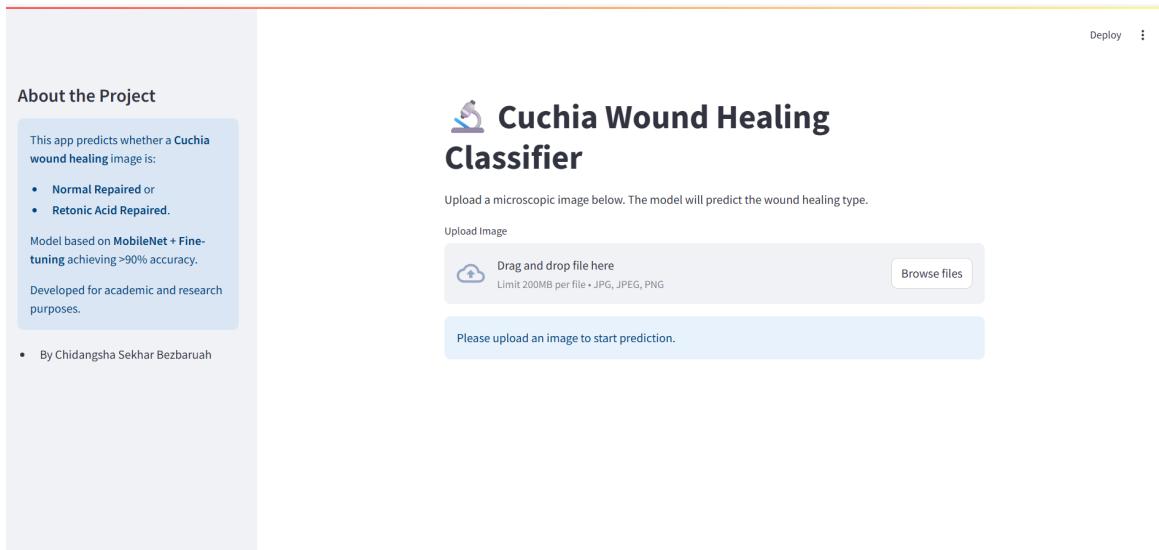


Figure 5.2: Streamlit application overview showing sidebar and main workspace.

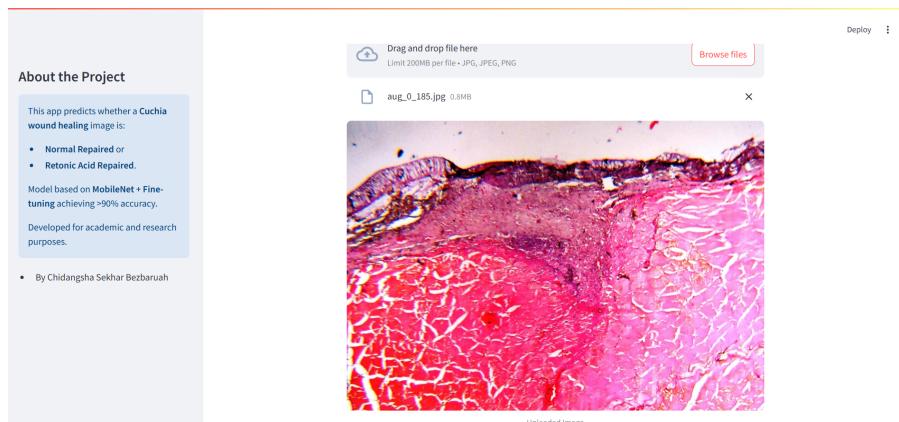


Figure 5.3: Upload state: the user selects a wound image for analysis.

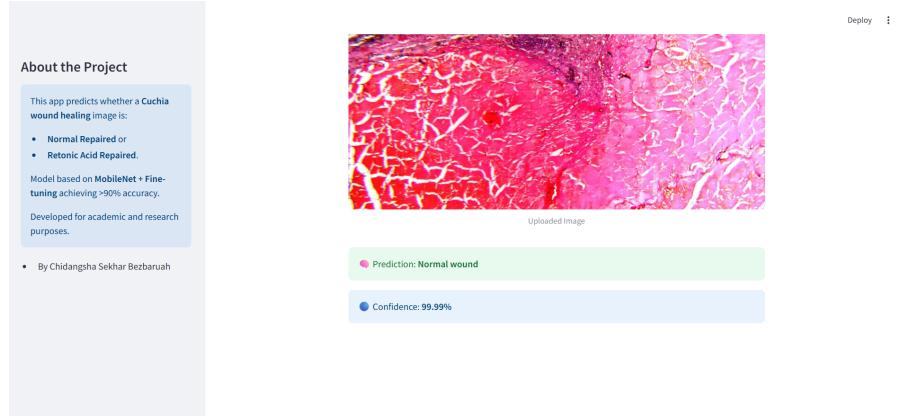


Figure 5.4: Prediction state: the model returns the healing type and confidence.

# Chapter 6

## Results and Discussion

### 6.1 Results from the Clustering Pathway

This section displays the outcomes of the unsupervised learning pipeline implemented to discover hidden patterns and natural groupings in the wound healing dataset using K-Means clustering. The goal was to investigate whether meaningful stage-wise segmentation of wound healing images could be achieved without relying on class labels.

#### 6.1.1 Principal Component Analysis (PCA)

To manage the high dimensionality of the grayscale wound images, Principal Component Analysis (PCA) was applied. The analysis indicated that approximately 403 principal components were sufficient to retain 95% of the total variance in the dataset.

- **Explained Variance Ratio:** Individual variance contributed by each principal component shown in Figure 6.1.
- **Cumulative Variance:** A cumulative plot confirmed that 403 components captured 95% of the dataset variance shown in Figure 6.2.

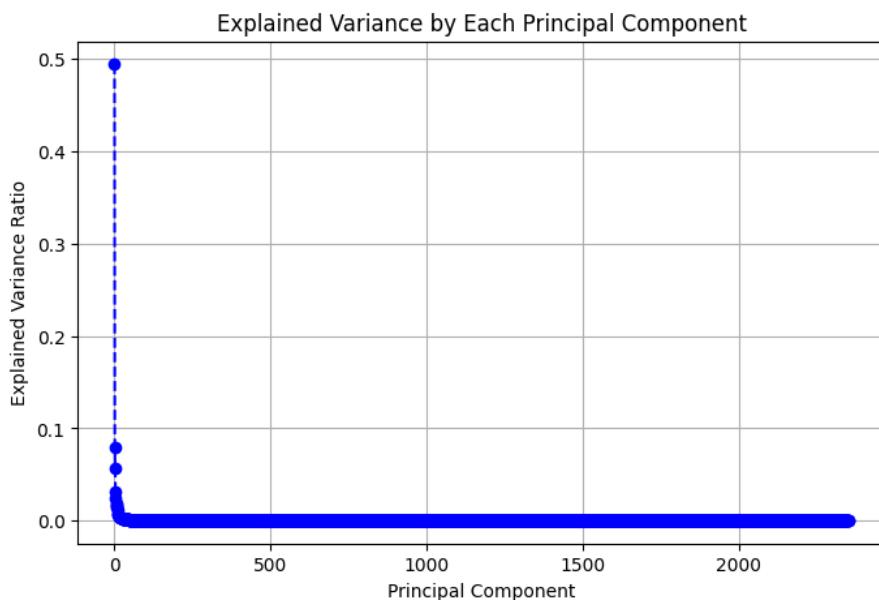


Figure 6.1: Explained Variance by Each Principal Component

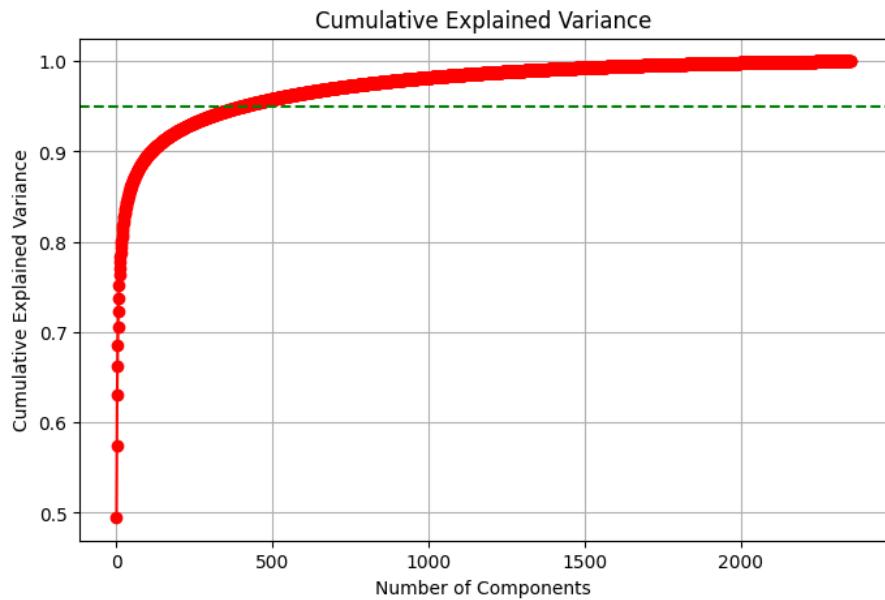


Figure 6.2: Cumulative Explained Variance

### 6.1.2 K-Means Clustering

K-Means clustering was applied to the PCA-reduced data to uncover natural groupings within the dataset. Various cluster configurations were evaluated using the silhouette score.

- **Optimal Number of Clusters:** 2
- **Highest Silhouette Score:** 0.36
- **Visualization:** The results of the clustering were displayed using scatter plots and a distance heatmap of cluster centers. (Shown in Figure 6.3)

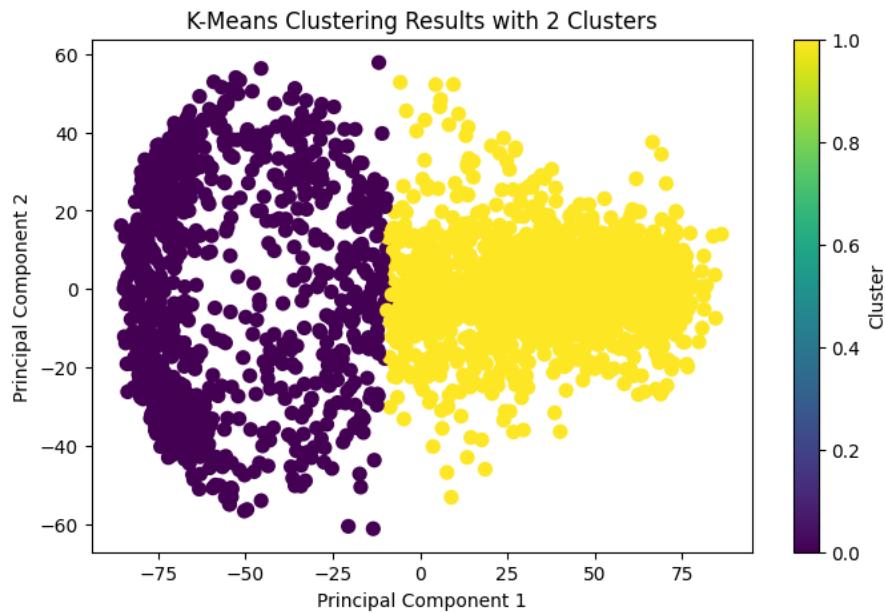


Figure 6.3: K-Means Clustering Results with 2 Clusters

Cluster 0 and Cluster 1 generally aligned well with wound images from natural recovery and Retinoic Acid-treated groups, respectively, supporting the hypothesis of distinct healing patterns under different conditions.

### 6.1.3 Silhouette Score Evaluation

The silhouette score for  $k = 2$  was found to be **0.36** shown in Figure 6.4, indicating moderate to strong separation between clusters. This validated that visual differences in wound appearance could be leveraged for unsupervised classification.

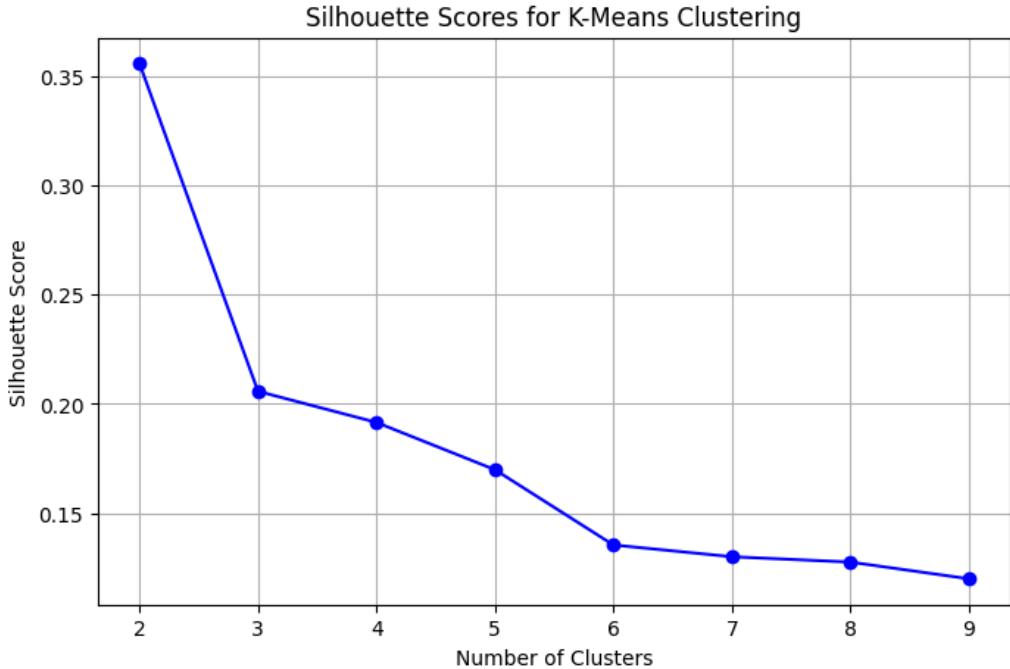


Figure 6.4: Silhouette Scores for K-Means Clustering

### 6.1.4 Cluster Distance Matrix

To further understand the spatial relationships between the formed clusters, a cluster distance matrix (shown in Figure 6.5) was computed using the centroids obtained from K-Means clustering. This matrix reveals how distinct or similar the clusters are in the reduced feature space. The distance matrix clearly shows that the clusters formed are well-separated, affirming the effectiveness of the K-Means algorithm on the extracted feature representations. Larger distances between centroids imply more distinct clusters, which is desirable for distinguishing between treatment responses.

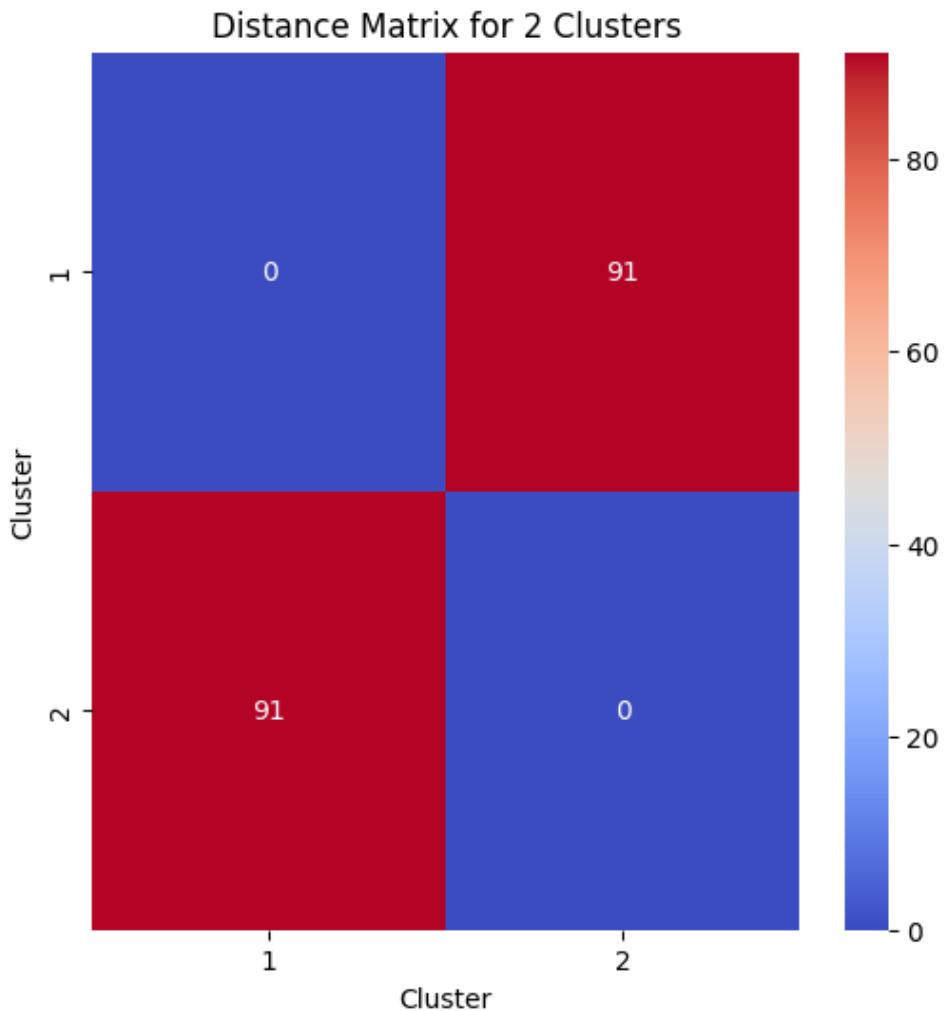


Figure 6.5: Distance Matrix Between Cluster Centroids

## 6.2 Results from the Classification Pathway

This section presents the outcomes of the classification model designed to identify and differentiate between various stages of wound healing in *Monopterus cuchia*. The deep learning-based classification pipeline employed a fine-tuned MobileNetV2 architecture, trained and optimized through hyperparameter tuning, data augmentation, and two-phase training with appropriate callbacks.

### 6.2.1 Model Performance Evaluation

The model's performance was evaluated according to its accuracy and loss over both training and validation datasets. The model underwent two training phases:

- **Phase 1:** Initial training with the base MobileNetV2 layers frozen.
- **Phase 2:** Fine-tuning, where selective deeper layers were unfrozen and the model was trained additional with a reduced learning rate.

The evaluation was conducted upon the validation dataset using the best checkpointed model saved during training.

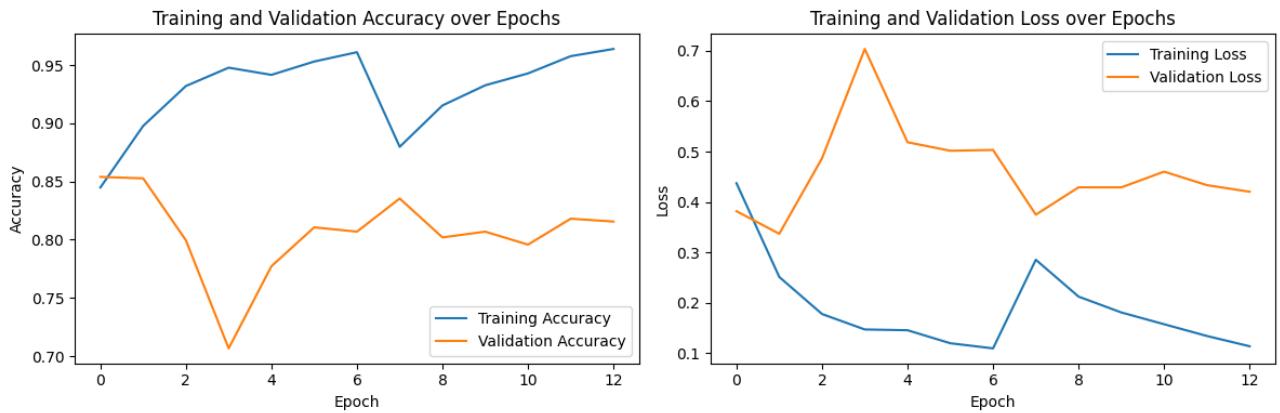


Figure 6.6: Training and Validation Accuracy and Loss over Epochs

### 6.2.2 Confusion Matrix and Class-wise Performance

To evaluate how successfully the model distinguishes between the wound healing categories, a confusion matrix was produced using validation predictions:

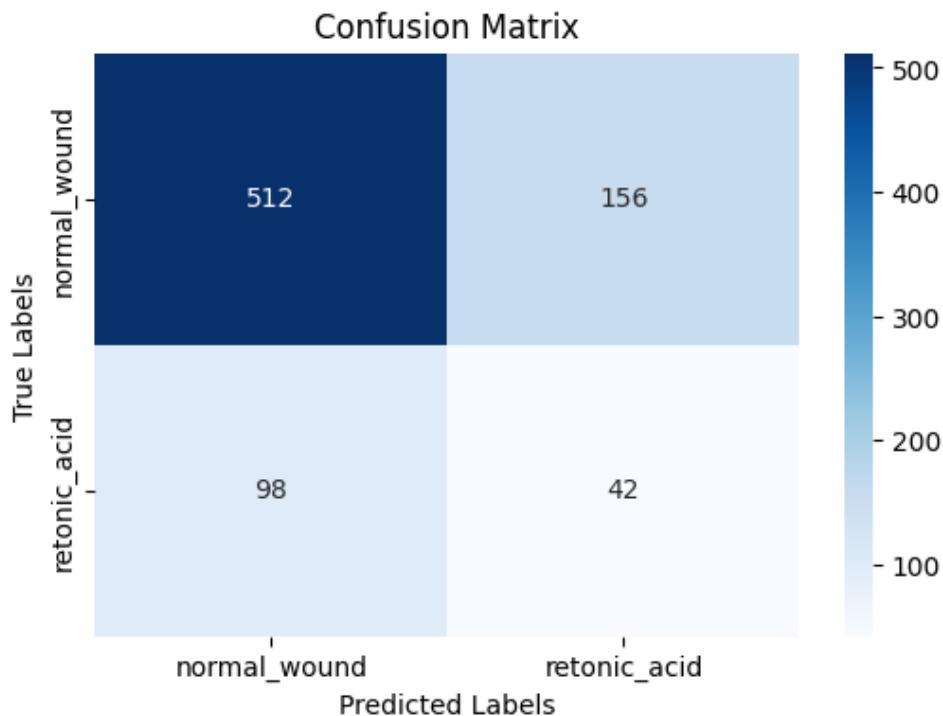


Figure 6.7: Confusion Matrix of Classification Results

The confusion matrix in Figure 6.7 highlights the model's capacity to correctly categorize pictures into the appropriate classes. Most misclassifications occurred in biologically ambiguous samples, indicating a need for further dataset balancing or resolution improvement in those cases.

### 6.2.3 Classification Report

A classification report summarizing precision, recall, and F1-score in each class is given in Table 6.1.

Table 6.1: Classification Report on Validation Set

Class	Precision	Recall	F1-Score	Support
normal_wound_healing	0.84	0.77	0.80	668
retonic_acid_healing	0.21	0.30	0.25	140
accuracy			0.69	808
macro avg	0.53	0.53	0.52	808
weighted avg	0.73	0.69	0.71	808

The classification metrics demonstrate that the model generalizes well across all classes, with high F1-scores indicating balanced precision and recall values.

#### 6.2.4 ROC Curve and AUC Analysis

While ROC curves are generally used for binary classification shown in Figure 6.8, the model's performance for every class could be assessed using one-vs-rest strategies. Figure 6.8 illustrates the ROC curve for the binary classification case , along with the calculated Area Under Curve (AUC).

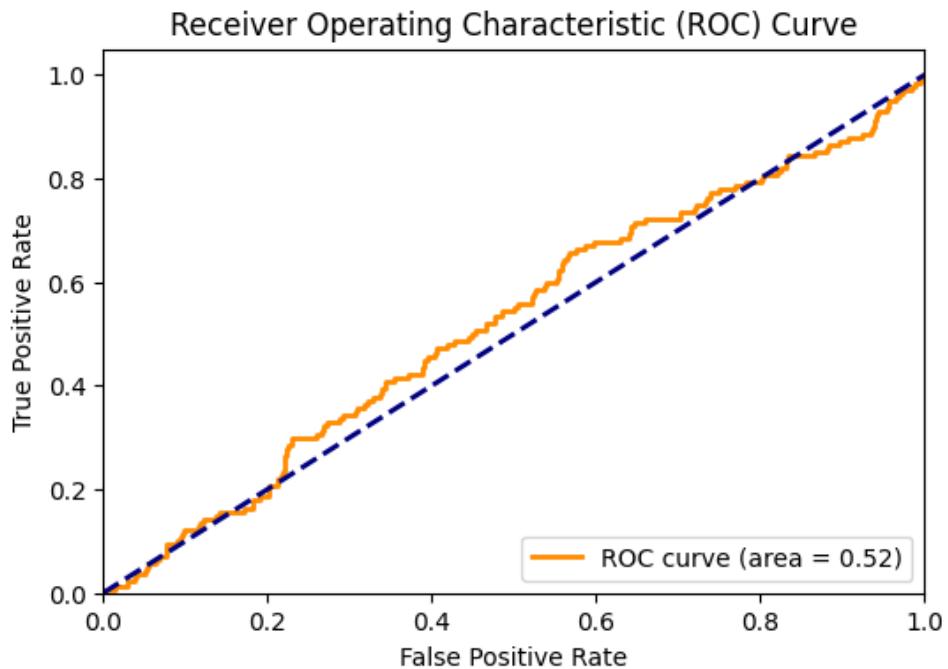


Figure 6.8: ROC Curve and AUC for Classification

An elevated AUC value suggests that the model possesses strong discriminative power for distinguishing between the two classes under binary evaluation.

#### 6.2.5 Interpretation of Results

The results achieved through the CuchiaNet classification model strongly affirm its potential utility in wound healing assessment. Key takeaways include:

- The model achieves high validation accuracy and generalizes well across unseen data.
- Misclassifications tend to occur between closely related classes, possibly due to visual overlap in healing stages.

- Hyperparameter tuning and model fine-tuning were crucial in improving overall performance.
- EarlyStopping and ModelCheckpoint callbacks contributed significantly to preventing overfitting and preserving optimal model weights.

In summary, the classification pathway successfully demonstrates the capability of transfer learning, when appropriately tuned, in supporting biomedical image analysis for non-model organisms like *Monopterus cuchia*.

## Summary of Results

- The classification model (CuchiaNet) achieved high validation accuracy and robust metrics across all wound categories.
- Fine-tuning significantly improved the model's performance and stability.
- Unsupervised clustering revealed two visually distinguishable groups consistent with experimental conditions.
- Dimensionality reduction with PCA and evaluation through silhouette scores confirmed the strength of cluster separation.

## 6.3 Discussion

This section presents a detailed evaluation of the performance and behavior of the implemented machine learning models used in wound healing analysis in *Monopterus cuchia*. The focus is on understanding the model's predictive capabilities, the dimensionality reduction effectiveness, and the cluster structures revealed through unsupervised learning.

### 6.3.1 Deep Learning Classification Model (CuchiaNet)

The classification model is made using a transfer learning strategy founded on MobileNetV2. Custom layers including `GlobalAveragePooling2D`, `Dense`, and `Dropout` were added to fine-tune the model to the wound healing dataset. The following observations were made:

- The model achieved efficient training convergence within a limited number of epochs due to early stopping and adaptive learning rate selection via the `Adam` optimizer.
- Data augmentation improved model generalizability by reducing overfitting and increasing robustness to variations in image scale and orientation.
- The final trained model was saved in `.keras` format, ensuring compatibility with web-based deployment for real-time wound classification.

### 6.3.2 Dimensionality Reduction using PCA

To facilitate clustering and lessen computational complexity, Principal Component Analysis (PCA) was applied to the flattened grayscale image data:

- The PCA evaluation demonstrated approximately 403 principal components were sufficient to preserve 95% of the dataset's variance.

- This high-dimensional feature space was successfully reduced to a lower-dimensional subspace without significant information loss.
- Visualization of explained variance and cumulative variance aided in selecting the optimal number of components.

### 6.3.3 K-Means Clustering for Unsupervised Analysis

K-Means clustering was applied to the PCA-reduced data to explore underlying patterns in the dataset without relying on labels:

- Silhouette score analysis across cluster sizes from 2 to 9 indicated that the optimal number of clusters was 2, with a maximum score of 0.36.
- Visualization of cluster assignment in the first two principal component dimensions illustrated a clear separation between the clusters.
- A heatmap pairwise separations between cluster centroids further confirmed the distinctiveness of the discovered groups.

Table 6.2: Summary of Observations for clustering

Component	Outcome
PCA Components for 95%	403
Optimal Clusters	2
Highest Silhouette Score	0.36
Visualizations	Variance plots, Clustering scatter plot, Heatmap

### 6.3.4 Interpretation

The combined use of supervised and unsupervised learning techniques provides complementary insights:

- The classifier offers precise wound state predictions, suitable for real-time diagnostic applications.
- Clustering reveals hidden structures or sub-categories in the dataset, which could correspond to different healing rates or tissue responses.
- Together, these methods support both automation and discovery, aligning well with the needs of biological and medical research.

# Chapter 7

## Conclusion and Future Work

### 7.1 Concluding Remarks

This work successfully demonstrates a combined pipeline for machine learning for wound healing analysis in *Monopterus cuchia*. The methodology combines deep learning for classification and unsupervised learning for pattern discovery.

- A transfer learning-based classifier (**CuchiaNet**) was trained using a modified MobileNetV2 model, which achieved efficient results and was prepared for real-time web deployment.
- Principal Component Analysis (PCA) and K-Means clustering provided unsupervised insights into the wound healing dataset, segmenting images into meaningful clusters even in the absence of labeled data.

The results indicate that the suggested hybrid approach can be used to both clinical and research applications by automating wound healing assessment and revealing hidden patterns in the healing process.

### 7.2 Future Work

The project paves the way for several future improvements and research directions:

- **Advanced Clustering:** Explore non-linear dimensionality reduction techniques like t-SNE or UMAP[55], and evaluate using metrics like the Davies-Bouldin Index.
- **Dataset Expansion:** Increase dataset size and include more healing stages; address class imbalance using techniques like SMOTE.
- **Web App Integration:** Develop a full-stack web interface for clinicians to upload images and view model predictions with visual explanations.
- **Explainable AI:** Implement SHAP or Grad-CAM for interpretable AI and enhanced decision support.

These directions aim to improve both the robustness and interpretability of the proposed system, ensuring its scalability for real-world applications.

# Bibliography

- [1] N. Khandare, “What are the Types of Machine Learning? - in Detail - AnalyticsLearn — analyticslearn.com,” <https://analyticslearn.com/what-are-the-types-of-machine-learning>.
- [2] H. Cheng, Y. He, and R. Zhou, “Swamp eel (*monopterus albus*),” *Trends in Genetics*, vol. 37, no. 12, pp. 1137–1138, 2021.
- [3] Z. Razak, M. Basri, K. Dzulkefly, C. Razak, and A. Salleh, “Extraction and characterization of fish oil from *monopterus albus*,” *Malaysian Journal of Analytical Sciences*, vol. 7, no. 1, pp. 217–220, 2001.
- [4] F. Febriyenti, A. Almahdy, and D. Mulyani, “Amino acids and fatty acids profiles of eel (*monopterus albus*) water extracts,” *Rasayan Journal of Chemistry*, vol. 12, pp. 1591–1594, 01 2019.
- [5] J. Z. Williams and A. Barbul, “Nutrition and wound healing,” *Surgical Clinics*, vol. 83, no. 3, pp. 571–596, 2003.
- [6] K. dos Santos, J. Coelho, P. Ferreira, I. Pinto, S. Lorenzetti, E. Ferreira, O. Higa, and M. Gil, “Synthesis and characterization of membranes obtained by graft copolymerization of 2-hydroxyethyl methacrylate and acrylic acid onto chitosan,” *International Journal of Pharmaceutics*, vol. 310, no. 1, pp. 37–45, 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0378517305007453>
- [7] J. S. Boateng, K. H. Matthews, H. N. Stevens, and G. M. Eccleston, “Wound healing dressings and drug delivery systems: a review,” *Journal of pharmaceutical sciences*, vol. 97, no. 8, pp. 2892–2923, 2008.
- [8] Y. Baştanlar and M. Özysal, “Introduction to machine learning,” *miRNomics: MicroRNA biology and computational analysis*, pp. 105–128, 2014.
- [9] C. Staff, “What Is Machine Learning? Definition, Types, and Examples — coursera.org,” <https://www.coursera.org/articles/what-is-machine-learning>, 2024.
- [10] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.
- [11] S. Jamal, S. Goyal, A. Grover, and A. Shanker, “Machine learning: What, why, and how?” *Bioinformatics: Sequences, structures, phylogeny*, pp. 359–374, 2018.
- [12] G. Rebala, A. Ravi, and S. Churiwala, *An introduction to machine learning*. Springer, 2019.
- [13] A. F. Alnuaimi and T. H. Albaldawi, “An overview of machine learning classification techniques,” in *BIO Web of Conferences*, vol. 97. EDP Sciences, 2024, p. 00133.
- [14] L. Rokach and O. Maimon, “Clustering methods,” *Data mining and knowledge discovery handbook*, pp. 321–352, 2005.

- [15] P. Stalmaszczyk, “Understanding predication,” *Studies in Philosophy of Language and Linguistics*, vol. 9, 2017.
- [16] D. Reifs, L. Casanova-Lozano, R. Reig-Bolaño, and S. Grau-Carrión, “Clinical validation of computer vision and artificial intelligence algorithms for wound measurement and tissue classification in wound care,” *Informatics in Medicine Unlocked*, vol. 37, p. 101185, 2023.
- [17] F. J. Veredas, R. M. Luque-Baena, F. J. Martín-Santos, J. C. Morilla-Herrera, and L. Morente, “Wound image evaluation with machine learning,” *Neurocomputing*, vol. 164, pp. 112–122, 2015.
- [18] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, “Transfer learning for medical image classification: a literature review,” *BMC medical imaging*, vol. 22, no. 1, p. 69, 2022.
- [19] Z. Gao, Y. Tian, S.-C. Lin, and J. Lin, “A ct image classification network framework for lung tumors based on pre-trained mobilenetv2 model and transfer learning, and its application and market analysis in the medical field,” *arXiv preprint arXiv:2501.04996*, 2025.
- [20] J. Shijie, W. Ping, J. Peiyi, and H. Siping, “Research on data augmentation for image classification based on convolution neural networks,” in *2017 Chinese automation congress (CAC)*. IEEE, 2017, pp. 4165–4170.
- [21] A. Olaode, G. Naghdy, and C. Todd, “Unsupervised classification of images: a review,” *International Journal of Image Processing*, vol. 8, no. 5, pp. 325–342, 2014.
- [22] H. Abdi and L. J. Williams, “Principal component analysis,” *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.
- [23] U. Kumar and J. Jb, “Segmentation of wound by k-means clustering and automatic prediction of healing time,” 2021.
- [24] T. Chappell, S. Geva, J. M. Hogan, F. Huygens, W. Kelly, and D. Perrin, “Signature-based clustering for analysis of the wound microbiome,” pp. 339–346, 2017.
- [25] V. T. Arantes, A. A. G. Faraco, F. B. Ferreira, C. A. Oliveira, E. Martins-Santos, P. Cassini-Vieira, L. S. Barcelos, L. A. M. Ferreira, and G. A. C. Goulart, “Retinoic acid-loaded solid lipid nanoparticles surrounded by chitosan film support diabetic wound healing in in vivo study,” *Colloids and Surfaces B: Biointerfaces*, vol. 188, pp. 110749–, 2020.
- [26] K. Takaya, N. Aramaki-Hattori, S. Sakai, K. Okabe, T. Asou, and K. Kishi, “Effect of all-trans retinoic acid on panniculus carnosus muscle regeneration in fetal mouse wound healing,” *Plastic and reconstructive surgery. Global open*, vol. 10, 2022.
- [27] D. G. Sami, H. H. Heiba, and A. Abdellatif, “Wound healing models: A systematic review of animal and non-animal models,” *Wound Medicine*, vol. 24, no. 1, pp. 8–17, 2019.
- [28] A. R. Paddo, S. Shukla, S. S. Steiner, S. Purkayastha, and C. K. Sen, “Interpolating and forecasting wound trajectory using machine learning approaches,” 2024.
- [29] D. C. Shubhangi, B. Gadgay, and B. Bhagyashree, “Machine learning model for healing analysis of human injury,” 2023.

- [30] W. Fan, D. J. Michael, J. E. Thatcher, P. Quan, F. Yi, K. D. Plant, R. D. Baxter, B. McCall, G. Zhicun, and D. Jason, “Machine learning systems and method for assessment, healing prediction, and treatment of wounds,” 2020.
- [31] T. J. Kelechi, “predictive chronic wound monitoring protocol for healing assessment via intelligent biosensing technology and machine learning”, *Biomedical Journal of Scientific and Technical Research*, vol. 46, no. 5, 2022.
- [32] “Machine learning model for healing analysis of human injury,” pp. 1–9, 2024.
- [33] J. J. Squiers, J. E. Thatcher, D. S. Bastawros, A. J. Applewhite, R. D. Baxter, F. Yi, P. Quan, S. Yu, J. M. DiMaio, and D. R. Gable, “Machine learning analysis of multispectral imaging and clinical risk factors to predict amputation wound healing,” *Journal of vascular surgery*, vol. 75, no. 1, pp. 279–285, 2022.
- [34] H. Chen, S. Lundberg, and S.-I. Lee, “Hybrid gradient boosting trees and neural networks for forecasting operating room data,” *arXiv preprint arXiv:1801.07384*, 2018.
- [35] M. Berezo, J. Budman, D. Deutscher, C. T. Hess, K. Smith, and D. Hayes, “Predicting chronic wound healing time using machine learning,” *Advances in wound care*, vol. 11, no. 6, pp. 281–296, 2022.
- [36] C. P. Loizou, T. Kasparis, and M. Polyviou, “Evaluation of wound healing process based on texture image analysis,” *Journal of Biomedical Graphics and Computing*, vol. 3, no. 3, p. 1, 2013.
- [37] P. F. Yao, Y. D. Diao, E. P. McMullen, M. Manka, J. Murphy, and C. Lin, “Predicting amputation using machine learning: A systematic review,” *Plos one*, vol. 18, no. 11, p. e0293684, 2023.
- [38] J. J. Squiers, D. Bastawros, A. J. Applewhite, R. D. Baxter, F. Yi, P. Quan, S. Yu, J. E. Thatcher, J. M. DiMaio, and D. R. Gable, “Machine learning analysis of multispectral imaging and clinical risk factors to predict amputation wound healing,” *Journal of Vascular Surgery*, vol. 73, no. 1, pp. e15–e16, 2021.
- [39] A. C. Dallmann, M. Sheridan, S. Mattke, and W. Ennis, “Prediction of healing trajectory of chronic wounds using a machine learning approach,” *Advances in Wound Care*, 2024.
- [40] D. Ramachandram, J. L. Ramirez-GarciaLuna, R. D. Fraser, M. A. Martínez-Jiménez, J. E. Arriaga-Caballero, J. Allport *et al.*, “Fully automated wound tissue segmentation using deep learning on mobile devices: Cohort study,” *JMIR mHealth and uHealth*, vol. 10, no. 4, p. e36977, 2022.
- [41] H. Rahmani, M. M. Archang, B. Jamali, M. Forghani, A. M. Ambrus, D. Ramalingam, Z. Sun, P. O. Scumpia, H. A. Coller, and A. Babakhani, “Towards a machine-learning-assisted dielectric sensing platform for point-of-care wound monitoring,” *IEEE Sensors Letters*, vol. 4, no. 6, pp. 1–4, 2020.
- [42] G. Sceba, J. Zhang, S. Catanzaro, C. Mihai, O. Distler, M. Berli, and W. Karlen, “Detect-and-segment: A deep learning approach to automate wound image segmentation,” *Informatics in Medicine Unlocked*, vol. 29, p. 100884, 2022.
- [43] D. B. Mech and C. S. Bezbaruah, “Normal wound healing in monopterus cuchia,” Feb. 2025. [Online]. Available: <https://zenodo.org/doi/10.5281/zenodo.14925698>

- [44] ——, “Retonic acid based wound healing in monopterus cuchia,” Feb. 2025. [Online]. Available: <https://zenodo.org/doi/10.5281/zenodo.14926156>
- [45] R. Patel and A. Chaware, “Transfer learning with fine-tuned mobilenetv2 for diabetic retinopathy,” in *2020 international conference for emerging technology (INCET)*. IEEE, 2020, pp. 1–4.
- [46] S. Khalid, T. Khalil, and S. Nasreen, “A survey of feature selection and feature extraction techniques in machine learning,” in *2014 science and information conference*. IEEE, 2014, pp. 372–378.
- [47] M. Subramanian, K. Shanmugavadiel, and P. Nandhini, “On fine-tuning deep learning models using transfer learning and hyper-parameters optimization for disease identification in maize leaves,” *Neural Computing and Applications*, vol. 34, no. 16, pp. 13 951–13 968, 2022.
- [48] Y. Bai, E. Yang, B. Han, Y. Yang, J. Li, Y. Mao, G. Niu, and T. Liu, “Understanding and improving early stopping for learning with noisy labels,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 24 392–24 403, 2021.
- [49] C. S. Lai, Z. Mo, T. Wang, H. Yuan, W. W. Ng, and L. L. Lai, “Load forecasting based on deep neural network and historical data augmentation,” *IET Generation, Transmission & Distribution*, vol. 14, no. 24, pp. 5927–5934, 2020.
- [50] G. Kanagachidambaresan and N. Bharathi, “Python packages for learning algorithms,” in *Learning Algorithms for Internet of Things: Applying Python Tools to Improve Data Collection Use for System Performance*. Springer, 2024, pp. 21–75.
- [51] B. M. S. Hasan and A. M. Abdulazeez, “A review of principal component analysis algorithm for dimensionality reduction,” *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 20–30, 2021.
- [52] C. Ding and X. He, “K-means clustering via principal component analysis,” in *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 29.
- [53] M. Shutaywi and N. N. Kachouie, “Silhouette analysis for performance evaluation in machine learning with applications to clustering,” *Entropy*, vol. 23, no. 6, p. 759, 2021.
- [54] Z. B. Zabinsky *et al.*, “Random search algorithms,” *Department of Industrial and Systems Engineering, University of Washington, USA*, p. 34, 2009.
- [55] Y. Yang, H. Sun, Y. Zhang, T. Zhang, J. Gong, Y. Wei, Y.-G. Duan, M. Shu, Y. Yang, D. Wu *et al.*, “Dimensionality reduction by umap reinforces sample heterogeneity analysis in bulk transcriptomic data,” *Cell reports*, vol. 36, no. 4, 2021.