West Visayas State University
(Formerly Iloilo Normal School)
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Luna St., La Paz, Iloilo City 5000
Iloilo, Philippines
* Trunkline: (063) (033) 320-0870 to 78 loc. 1403 * Telefax No.: (033) 320-0879
* Website: www.wvsu.edu.ph * Email Address: cict@wvsu.edu.ph

BAGONG PILIPINAS

DermaLense:
An AI Powered Screening for Canine Skin Diseases

A Final Project
Presented to the Faculty of the
College of Information and Communications Technology
West Visayas State University
La Paz, Iloilo City

In Partial Fulfillment
of the Requirements for the Subject
CCS 248: Artificial Neural Networks

BREY, JUSTIN JONES P.
JUANICO, MICHELLE JOY C.
LEDESMA, ALLYN RALF T.

DECEMBER 2025

Problem Overview and Justification

Canine dermatological conditions are highly prevalent, affecting an estimated 20% to 75% of dogs, making it one of the most common reasons for veterinary consultations, yet they present a significant diagnostic challenge due to their visual ambiguity. Conditions such as Demodicosis, Fungal Infections, Hypersensitivity, Ringworm, and Dermatitis often have identical clinical signs, including erythema (redness), alopecia (hair loss), and scaling, making it difficult for pet owners and even general practitioners to distinguish them without invasive diagnostic tests like skin scrapings.

In remote areas, access to veterinary care is scarce or expensive. Consequently, many dogs suffer from prolonged conditions that are treatable due to the lack of immediate and accurate diagnostic support. The owners often attempt to treat minor skin issues at home using over-the-counter products. However, mistaking a fungal infection for an allergic reaction and treating it with steroids can worsen the infection, leading to severe secondary complications.

There is an urgent need for an accessible and automated screening tool that can classify these visually similar skin conditions to guide owners toward appropriate care and prevent the worsening of these manageable diseases.

This project aims to build an automated image classification system using Residual Neural Network (ResNet-50) to identify dog skin disease. Resnet is specifically designed to handle complex image recognition tasks. Unlike simpler models, ResNet uses skip connections that allow it to learn very deep and detailed features. This is necessary because the visual difference between a mite infestation (Demodicosis) and a bacterial infection (Dermatitis) can be very subtle and hard for standard cameras to pick up.

Data Selection and Validation

1. **Data Aggregation**

To achieve the six-class classification capability, we will aggregate data from three distinct open-source repositories. Since the primary dataset source does not contain enough images of canine dermatological conditions, we will create a Unified Master Dataset by merging the following sources:

- Mohmmed, Y. (2025). *Dog's skin diseases (Image Dataset)* [Data set]. Kaggle. https://www.kaggle.com/datasets/youssefmohmmed/dogs-skin-diseases-image-dataset
- Motiani, Y. (2023). *Dogs Skin disease dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/yashmotiani/dogs-skin-disease-dataset

- Dog Skin Disease Dermatosis. (2025). *Dog Skin Disease Dataset* [Data set]. Roboflow Universe. https://universe.roboflow.com/dog-skin-disease-dermatosis/dog-skin-disease-dataset

2. Dataset Composition and Organization

The project utilizes the DermaLens dataset, a curated collection of canine dermatological imagery organized into three distinct partitions: Training, Validation, and Testing. The dataset comprises a total of 5,530 validated images, categorized into six specific classes to support the project's multi-class classification objective.

- Total Volume: 5,530 Images
- Split Strategy:
  - Training Set: 3,911 images (approx. 70%) – Used for model feature learning.
  - Validation Set: 1,023 images (approx. 18.5%) – Used for hyperparameter tuning and preventing overfitting.
  - Test Set:596 images (approx. 10.8%) – Reserved for final unbiased performance evaluation.

3. Distribution Analysis

The dataset covers five distinct pathological conditions and one healthy control group. An analysis of the training set reveals a class imbalance, with Hypersensitivity being the minority class.

Skin Disease Class	Image Count (Train)	Distribution Status
Ringworm	923	Dominant Class
Healthy	864	High Representation

Dermatitis	728	Balanced
Demodicosis	632	Balanced
Fungal Infections	471	Moderate Representation
Hypersensitivity	293	**Minority Class**

4. Data Integrity and Quality Assurance

To ensure the reliability of the input data, a rigorous validation protocol was implemented prior to model training.

- Integrity Checking (check_dataset_integrity.py): A custom validation script utilizing the Python Imaging Library (PIL) was executed to verify file readability.
    - Format Verification: Validated strict adherence to .jpg, .jpeg, and .png formats.
    - Corruption Scanning: Simulated "open and read" operations on every file to detect header corruption or truncated data.
- Validation Results: The dataset achieved a 100% integrity pass rate according to the dataset_integrity_report.json, with zero corrupted files detected across all 5,530 images.
- Runtime Safeguards (data_loader.py): The data loading pipeline includes exception handling (try-catch blocks) to gracefully manage any unforeseen loading errors during runtime, ensuring that a single faulty file does not crash the training process.

5. Preprocessing and Augmentation Pipeline

To prepare the raw imagery for the ResNet architecture, the following standardization and augmentation steps were applied:

A. Standardization (Preprocessing)

- Resolution: All input images were resized to 224 × 224 pixels to match the input layer requirements of the ResNet50 architecture.
- Normalization: Pixel values were normalized using the standard ImageNet statistics (Mean: [0.485, 0.456, 0.406], Std: [0.229, 0.224, 0.225]) to accelerate gradient convergence.
- Color Space: All images were converted to RGB to ensure consistent 3-channel input.

B. Training Augmentation to improve the model's generalization and robustness against variations in lighting and orientation, the following transformations were applied dynamically during training:

- Geometric Transformations: Random horizontal/vertical flips and random rotations (±15°) to simulate different camera angles.
- Crop & Scale: Random resized cropping (scale 0.8–1.0) to force the model to focus on lesion details rather than the image background.
- Photometric Distortions: Color jittering (adjusting brightness, contrast, saturation, and hue) and Gaussian blur to simulate varying photo qualities typical of user-uploaded images.

Model Architecture Selection and Design

The neural network architecture for DermaLens was centered around the use of ResNet-50, an architecture known for its strong feature extraction capabilities and stability in image classification tasks. ResNet-50 was chosen because its deep residual structure allows the model to capture both fine-grained and large-scale visual patterns that are essential for distinguishing dermatological conditions, which often vary in texture, color distribution, and lesion shape. Its depth and residual connections support efficient gradient flow, enabling the network to learn complex features without performance degradation.

The model design begins with standardized preprocessing, where images are resized to the required input dimensions and normalized to ensure stable and consistent training. ResNet-50's convolutional layers then extract hierarchical features, progressing from low-level edges to high-level lesion representations. These features are passed to a classifier head that applies global average pooling to reduce spatial dimensions, followed by fully connected layers with dropout to minimize overfitting. Categorical cross-entropy is used as the primary loss function, while regularization methods such as dropout and weight decay strengthen the model's generalization capability.

Model Training and Hyperparameter Tuning

The DermaLens model was trained from scratch to ensure that its parameters were fully tailored to the dermatological classification task without relying on features learned from unrelated datasets. This approach required more careful optimization but allowed the network to develop representations specific to skin lesions characteristics, improving its ability to distinguish subtle visual differences across classes.

AdamW optimizer was selected because it provides adaptive learning rates while applying decoupled weight decay, resulting in more stable updates and stronger regularization than standard Adam. This made it well suited for training a deep model from scratch, where controlling overfitting and maintaining smooth convergence are essential.

Hyperparameter tuning was carried out through systematic experimentation, focusing primarily on learning rate, batch size, and regularization strength. Learning rates between 1e-5 and 1e-2 were tested, with 0.01 (1e-2) producing the most stable and consistent convergence across the entire network. Batch sizes of 8, 16, and 32 were evaluated, with a batch size of 16 offering the best balance of memory efficiency and gradient reliability on the available GPU.

Regularization was refined by experimenting with multiple dropout rates, where a value of 0.5 in the classifier head most effectively minimized overfitting. Weight decay was set to 0.0001 (1e-4) to further enhance generalization. To prevent stagnation and adapt the learning rate based on validation performance, a ReduceLROnPlateau scheduler was applied, reducing the learning rate by a factor of 0.1 when validation loss plateaued for 5 consecutive epochs, with a minimum learning rate threshold of 1e-6.

The final configuration used AdamW with a uniform learning rate of 0.01, a batch size of 16, dropout of 0.5, weight decay of 0.0001, and ReduceLROnPlateau scheduling. Training was conducted for 143 epochs without early stopping, with the best model saved at epoch 122 based on validation accuracy. This setup ensured comprehensive training and robust generalization suited to dermatological image classification.

Results, Evaluation, and Analysis

The performance of the DermaLens model was evaluated across training, validation, and test datasets to determine its overall effectiveness and generalization capability. Throughout 143 epochs of training, the model demonstrated smooth and stable convergence. Training loss steadily decreased from 1.791 at epoch 1 to 0.380 at epoch 143, while validation loss decreased from 1.673 to 0.462, indicating that the model maintained a balanced fit. The best

validation performance was achieved at epoch 122 with a validation accuracy of 87.59% and validation F1-score of 87.38%. At this checkpoint, the training accuracy was 85.58% with a training loss of 0.419.

Evaluation on the held-out test set further confirmed the model's generalization performance. The model achieved a test accuracy of 86.61% with a test F1-score of 86.31%. These results reflect well-balanced performance, with validation and test metrics showing close alignment (87.59% validation vs 86.61% test accuracy), indicating effective generalization without significant overfitting. The confusion matrix revealed consistent accuracy across most categories, though it also highlighted slight misclassifications in underrepresented classes, emphasizing the ongoing impact of class imbalance in the dataset..

Overall, the model exhibited clear strengths through its strong F1-score (87.38% validation, 86.31% test) and consistent alignment between validation and test results (87.59% vs 86.61% accuracy), demonstrating robust performance and effective generalization. However, limitations were also observed, particularly in classes with fewer examples, where performance variations suggested that additional class balancing strategies may be beneficial. Addressing these gaps through techniques such as oversampling, focal loss, or domain-specific augmentation could further stabilize predictions. Moving forward, incorporating stratified k-fold cross-validation and exploring self-supervised pretraining may provide more comprehensive performance estimates and improved feature extraction for dermatological images.

Conclusion

The DermaLens project demonstrates how deep learning can be effectively applied to dermatological image classification through a carefully designed model architecture, a fully-from-scratch training approach, and systematic hyperparameter tuning. The resulting performance, reflected in strong accuracy (86.61% test, 87.59% validation) and F1-scores (86.31% test, 87.38% validation), confirms the model's robustness and strong generalization capability, indicating its potential reliability in clinical decision support

Although the model performs well, areas for improvement remain, particularly in addressing class imbalance and enhancing robustness through domain-specific data augmentations. Exploring more advanced training strategies, such as self-supervised pretraining, may also strengthen feature extraction in future iterations. These refinements could further increase the model’s stability and adaptability in real-world clinical settings

Overall, DermaLens provides a solid foundation for dermatological AI systems and demonstrates clear potential to assist clinicians in diagnosing and classifying skin conditions with greater accuracy and consistency.