Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.1, February 2026, pp. 378~387
e-ISSN: 2548-6861

Classification of Facial Acne Types Based on Self-Supervised Learning using DINOv2

Gantari Chardaputeri[1]*, Angelina Pramana Thenata[2]*
* Informatics, Universitas Bunda Mulia
0212gantari@gmail.com [1], angelina.pramana@outlook.com [2]

Article Info

Article history:

Received 2025-11-25
Revised 2026-01-05
Accepted 2026-01-08

Keyword:

DINOv2,
Self-Supervised Learning,
LoRA,
Acne Type Classification,
Computer Vision.

ABSTRACT

Acne is a common inflammatory skin condition that can affect an individual's psychological well-being and overall quality of life. The inability to independently recognize specific types of acne often leads to the use of inappropriate skincare products. This situation highlights the need for an image-based classification system that can provide accurate visual identification. The self-supervised learning method Distillation with NO Labels, version 2 (DINOv2), is employed as a feature extractor to classify four types of acne—Acne fulminans, Acne nodules, Papules, and Pustules—using the "skin-90" dataset. The fine-tuning process is conducted through a Parameter-Efficient Fine-Tuning (PEFT) approach using Low-Rank Adaptation (LoRA) to adjust the model's visual representations to the acne domain without updating all parameters in full, followed by integration with a classification head. The results show that the model achieves an accuracy of 90.70%, with precision, recall, and F1-score values of 90.64%, 90.68%, and 90.57%, respectively. The findings suggest that the proposed architectural design and training configuration are suitable for capturing relevant visual patterns of acne, while further validation is required to assess robustness across more diverse data distributions.

This is an open access article under the CC–BY-SA license.

I. INTRODUCTION

Acne is an inflammatory skin disorder with a high prevalence among adolescents and is still commonly observed in some adults [1], [2]. This condition affects not only skin health but also influences an individual's confidence, psychological well-being, and overall quality of life [3], [4]. A major challenge in acne treatment lies in the limited ability of individuals to identify the specific types of acne accurately. Misclassification often leads to the use of inappropriate skincare products, which may worsen the skin condition and result in unnecessary time and financial costs. This situation highlights the need for an image-based classification system that can consistently and accurately identify various types of acne.

Advancements in computer vision and artificial intelligence have enabled the development of automated image-based classification systems capable of identifying distinctive visual patterns through feature extraction [5], [6]. In dermatological image analysis, however, the availability of large-scale annotated datasets remains limited. Facial acne images often exhibit substantial intra-class variation and inter-class similarity, which poses challenges for supervised learning approaches that rely heavily on labelled data.

To address these limitations, self-supervised learning has emerged as a promising paradigm for learning robust visual representations without requiring manual annotations [7], [8]. Distillation with NO Labels version 2 (DINOv2) is a Vision Transformer (ViT)-based foundation model pretrained on large-scale unlabeled data to produce stable and transferable representations [9]. Self-supervised learning approaches such as SimCLR, MoCo, and BYOL primarily rely on instance-level contrastive objectives or projection heads to learn visual representations [10], [11], [12]. Although these methods have demonstrated strong performance in general image recognition tasks, their effectiveness is influenced by data augmentation strategies and the availability of sufficient training samples. In medical image analysis, where subtle inter-class differences and high intra-class variability are common, such reliance may limit the learning of dense and

semantically rich representations under constrained data conditions.

In response to these limitations, previous studies have reported that DINOv2 exhibits strong generalization across diverse visual domains, including medical imaging and dermatological analysis, as well as visual pattern retrieval tasks such as Content-Based Image Retrieval (CBIR) [13], [14]. Unlike contrastive self-supervised methods, DINOv2 is designed to learn dense visual representations without explicit negative pairs, which supports more stable feature transfer to fine-grained classification tasks. In this study, the pretrained DINOv2 model is adapted using a Parameter-Efficient Fine-Tuning (PEFT) strategy based on Low-Rank Adaptation (LoRA), which updates only a limited subset of parameters to specialize the representations for the acne classification task [15].

Prior research has developed automatic acne classification systems using supervised CNN-based transfer learning, such as MobileNetV2 combined with K-Fold cross-validation on the "skin-90" dataset [16]. Although these approaches demonstrate promising results, the learned representations remain dependent on labelled data. They may be limited in capturing complex visual variations in facial acne images, particularly under constrained data settings.

In contrast to supervised CNN-based methods, this study explores the use of a self-supervised vision foundation model, namely DINOv2, adapted through PEFT for facial acne classification. To the best of our knowledge, this study represents the first investigation of adapting a self-supervised vision foundation model with LoRA for acne classification on the "skin-90" dataset. This work, therefore, evaluates the applicability of self-supervised vision foundation models in a domain that has not been explicitly examined in prior "skin-90"-based studies.

This study aims to evaluate the performance of DINOv2, adapted with LoRA, in classifying four types of acne in the "skin-90" dataset: Acne fulminans, Acne nodules, Papules, and Pustules [16]. The evaluation is conducted using a confusion matrix and the metrics of accuracy, precision, recall, and F1-score to provide a comprehensive assessment of the model's predictive quality [17].

The findings of this study are expected to contribute to the development of a more accurate and efficient image-based acne classification system. The results also have the potential to serve as an educational tool for understanding skin conditions and supporting the selection of relevant and safe skincare ingredients based on the classified acne type. Thus, this research provides not only academic value but also practical benefits in enhancing technology-based skincare literacy.

II. METHODS

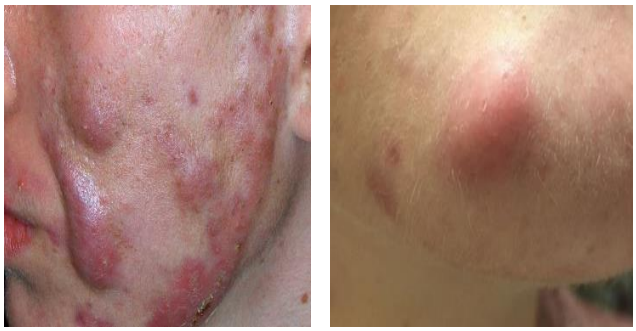
This study adopts a quantitative experimental methodology to evaluate the effectiveness of a self-supervised learning model for acne type classification. The computational experiments involve training, validation, and testing phases conducted on the "skin-90" dataset, which consists of facial images containing visible acne. The dataset is obtained from Kaggle, and its class distribution is summarized in Table 1. As shown in the table, the number of samples is uneven across acne categories, with certain classes represented by fewer images. This imbalance reflects the natural variation in acne occurrence and constitutes an important characteristic of the dataset considered in this study.

TABLE 1.
TOTAL OF IMAGES FOR EACH ACNE TYPE

Acne Types	Data
Acne fulminans	70
Acne nodules	71
Papules	70
Pustules	70

The images exhibit variability in lighting conditions, facial pose, and image quality, which introduces visual diversity but also potential sources of bias. Furthermore, the dataset does not provide metadata related to subject demographics, skin tone, or acquisition settings, limiting the assessment of demographic representativeness and potential sampling bias. As a secondary dataset collected from public sources, these characteristics should be considered when interpreting the experimental results.

Only images that display human faces with visible acne types were included to ensure that the analysis focuses on the visual characteristics of acne severity, without considering factors such as age, gender, or ethnicity. DINOv2 is used for feature extraction, while LoRA is applied to fine-tune the model to the acne domain. The main objective is to assess the model's performance across four acne classes: Acne fulminans, Acne nodules, Papules, and Pustules. To illustrate the visual diversity among these categories, examples are presented in Figure 1.



(a) (b)

(c) (d)

Figure 1. Visualization of Acne Types: (a) Acne fulminans, (b) Acne nodules, (c) Papule, and (d) Pustules

Figure 1(a) presents the severe acne form Acne fulminans, characterized by intense inflammation and coalescent nodulocystic lesions [18], [19]. Figure 1(b) shows Acne nodules, which are inflammatory lesions consisting of large, firm, and painful lumps that develop beneath the skin's surface and may lead to permanent scarring [18], [20]. The papule shown in Figure 1(c) is a type of acne that appears as a small red or pink bump resulting from inflammation in a clogged hair follicle. Despite its relatively small size, a papule can be painful when touched [18]. Pustules, illustrated in Figure 1(d), are pus-filled lesions with a red and tender base [18], [21].

The selection of these four acne classes in this study is based on the consideration that acne variations often confuse individuals without dermatological knowledge when identified. Although each category has distinct characteristics, these differences frequently appear similar during everyday visual observation. This situation indicates that lay assessments are prone to misclassification.

The flowchart is used to visually and systematically represent the stages of the research, facilitating a clearer understanding of the overall process. Each stage is arranged sequentially, from pre-processing to model evaluation, thereby providing a complete overview of the workflow in developing the acne classification system. The flowchart is presented in Figure 2. The process begins with dataset collection.

A. Pre-processing

The pre-processing stage begins with a systematic division of the dataset into training, validation, and testing subsets using a 70:15:15 split ratio. This strategy is designed to preserve the relative class proportions, ensuring that each acne category is represented consistently across all subsets. The resulting sample distribution for each split is summarized in Table 2.

TABLE 2.
DISTRIBUTION OF IMAGES FOR EACH ACNE TYPE IN THE TRAINING, VALIDATION, AND TESTING SUBSETS

Acne Types	Data		
	Training	Validation	Testing
Acne fulminans	49	11	10
Acne nodules	49	11	11
Papules	49	10	11
Pustules	49	10	11

To improve model generalization and robustness against real-world variations, a comprehensive data augmentation pipeline is applied during training. The pipeline incorporates geometric transformations, including random rotations, affine transformations, horizontal and vertical flipping, and random resized cropping, to enhance spatial invariance. Photometric augmentations such as color jittering, Gaussian blur, Contrast-Limited Adaptive Histogram Equalization (CLAHE), adaptive sharpening, and the injection of light noise are employed to account for variations in illumination, contrast, and image quality.

Further robustness is introduced through perspective distortion and random erasing, which simulate partial occlusions and missing visual information. In addition to instance-level augmentations, MixUp is applied as a sample-level regularization technique by linearly combining pairs of images and their corresponding labels, encouraging smoother decision boundaries. Policy-based augmentation is incorporated using TrivialAugment, allowing stochastic selection of transformation operations with varying magnitudes.

All images are resized to a fixed resolution prior to normalization to ensure architectural compatibility and consistent spatial representation. Finally, input tensors are normalized using the mean and standard deviation derived from ImageNet normalization statistics, which are adopted to maintain compatibility with pretrained backbone models.

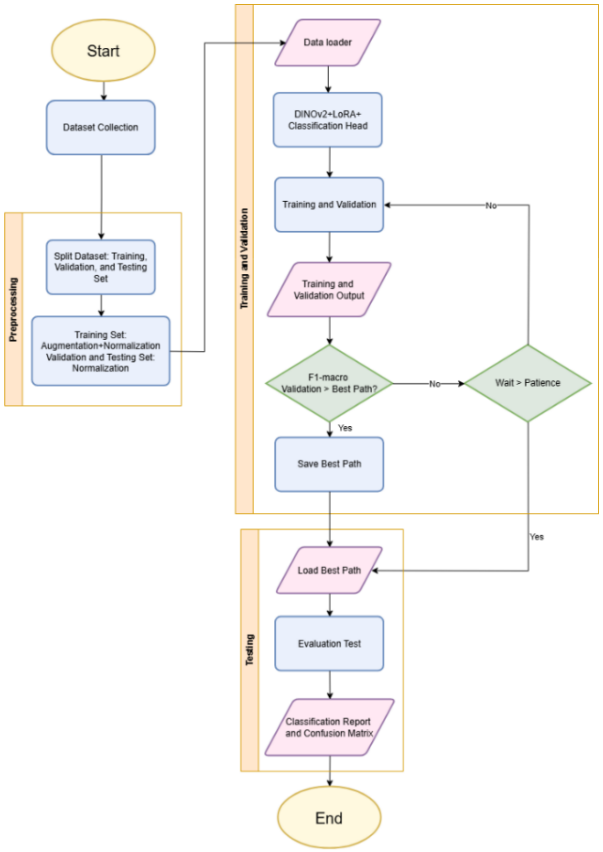


Figure 2. System Workflow Flowchart

The validation and testing subsets undergo only normalization to ensure consistent scaling without altering the intrinsic characteristics of each acne type. Every image is transformed using the same statistical parameters applied to the training subset. Table 3 presents the detailed distribution of acne images across the training, validation, and testing subsets after augmentations and normalization.

TABLE 3.
DISTRIBUTION OF IMAGES FOR EACH ACNE TYPE IN THE TRAINING, VALIDATION, AND TESTING SUBSETS AFTER AUGMENTATIONS AND NORMALIZATION

Acne Types	Data		
	Training	Validation	Testing
Acne fulminans	294	10	11
Acne nodules	294	11	11
Papules	294	11	10
Pustules	294	10	11

The data splitting is carried out using a stratified approach to maintain proportional representation across all acne categories, thereby preserving class balance and ensuring consistent diversity throughout the experimental workflow.

B. Training and Validation

The training and validation processes were conducted for 20 epochs, during which the system generated indicators such as loss, accuracy, precision, recall, and F1-score. An early stopping mechanism with a patience threshold of five epochs was applied to reduce overfitting. If the macro F1-score on the validation set exceeded the previous best value, the model parameters were saved as the optimal checkpoint. Conversely, if no improvement occurred for five consecutive epochs, the training and validation processes were automatically terminated. This optimization strategy ensures that the final model reflects a configuration that is well-generalized, stable, and maintains a balanced trade-off between sensitivity and specificity across all types of acne.

C. Testing

After determining the model with optimal performance, the process continued with the testing stage. The evaluation was conducted on the testing subset, which consists of data unseen during training and validation, thereby ensuring an objective assessment of the model's real-world generalization capability. The outputs include a classification report encompassing accuracy, precision, recall, and F1-score, as well as a confusion matrix that visually illustrates the correspondence between predicted labels and actual labels.

D. Website Implementation

The system implementation was carried out through the development of a website that serves as the main interface for users to classify acne types and obtain recommended skincare ingredients. The website was built using a combination of HTML, CSS, JavaScript, and Bootstrap to produce a responsive, consistent, and user-friendly interface. HTML is used to construct the page structure, while CSS manages the layout and visual aspects through color settings, typography, and interface elements. The use of the Bootstrap framework accelerates the development of user interface components through standardized, ready-to-use elements, ensuring visual consistency across all pages.

JavaScript handles the interactive logic, including image input processing, classification function calls, prediction display, and the management of dynamic elements on the website. Integration with Supabase enables the system to authenticate users and manage the skincare ingredient recommendation tables. Supabase was selected because it provides a serverless backend service with fast APIs, supported by a PostgreSQL database, and equipped with Row-Level Security (RLS) mechanisms. The integration is carried out through the JavaScript SDK provided by Supabase, allowing authentication, data retrieval, and table updates to be performed directly from the client side without adding infrastructure complexity.

The Maximum Softmax Percentage (MSP) is applied to determine whether an image can be classified into one of the four acne categories or falls outside the model's scope [22]. In this system, a threshold of 0.70 is used, obtained by calculating the minimum MSP value from correct predictions and the maximum MSP value from incorrect predictions, then adjusting it according to the application's requirements. This adjustment is intended to reduce false positives for data that do not belong to the target classes.

All website functionalities were tested using the black-box testing method to ensure alignment between the system outputs and the predefined functional requirements [23]. The testing was carried out by designing various scenarios, which are presented in Table 4.

TABLE 4.
BLACKBOX TESTING

No.	Tested Features	Scenario	Input	Expected Output
1.	Account Registration (*Sign Up*)	The user registers with valid data.	A name, an unregistered email, and a valid password.	The account is successfully created, a registration success notification appears, and a verification link is sent to the user's email.
2.		The user registers using an email that is already registered.	A name, an email that has already been used, and a valid password.	The system rejects the registration and displays an error message indicating that the email is already registered.
3.	System Login (*Sign In*)	The user signs in with the correct credentials.	A valid email and the corresponding password.	The user is directed to the main page according to their role (user/admin).

No.	Tested Features	Scenario	Input	Expected Output
4.		The user signs in with an incorrect password.	A valid email and an incorrect password.	The system displays an error message indicating that the sign-in attempt has failed.
5.	Remember Me Feature	The user enables the "remember me" option during sign-in.	A valid email and password, with the "remember me" option checked.	The system stores the authentication token, allowing the user to remain signed in for the next session.
6.	Forgot Password	The user submits a password reset request.	Registered email.	The system sends a reset link to the email and allows the password to be reset.
7.	Upload Facial Image	The user uploads an acne facial image in a supported format.	Image file in JPG/PNG format.	The system accepts the file, displays a preview, and is ready to perform the classification
8.	Acne Classification Process	The user clicks the "Classify" button after uploading a valid image.	Facial image file.	The system displays the acne type, the confidence score, and the recommended skincare ingredients.
9.	Skincare Ingredient Recommendations	The system displays recommendations based on the classification results.	Predicted acne type.	The recommended skincare ingredients appear according to the acne type.
10.	Admin Model Evaluation Page	The admin opens the model evaluation page.	Valid admin credentials.	The system displays the model metadata, class distribution, dataset image previews, model metrics (accuracy, precision, recall, F1-score), the classification report, and the confusion matrix.
11.	Management of Skincare Ingredient Recommendations	The admin updates the list of skincare ingredient recommendations.	New recommendation data.	The system saves the changes and consistently displays the updated list.

This approach is relevant because it does not require examining the source code; instead, it evaluates the system's behaviour based on its responses to user inputs.

III. RESULT AND DISCUSSION

A. *Method Success Evaluation*

The evaluation of the method's performance was conducted to assess the model's generalization ability. Figures 3 to 5, along with Tables 5 and 6, are used to present the dynamics of the metrics during training, the patterns of misclassification across each subset, and the consistency of the model's confidence levels in the final predictions.

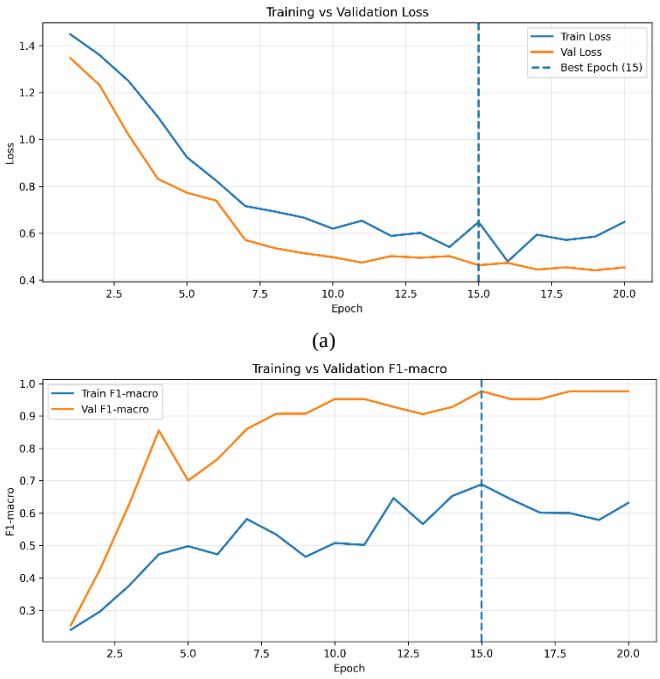


Figure 3. Training and Validation Result Graphs: (a) Loss and (b) F1-macro

Figure 3(a) illustrates the dynamics of the loss values for both the training and validation data over 20 epochs. During the initial training phase, both curves exhibit a pronounced decrease, indicating rapid adaptation of model parameters to the underlying data patterns. In subsequent epochs, the loss reduction becomes more gradual, reflecting a stabilization of the learning process.

Throughout training, the training loss remains slightly higher than the validation loss, while both curves follow similar trajectories with a relatively stable gap. This pattern suggests aligned optimization behaviour between the training and validation phases rather than evidence of performance divergence. Overall, the loss trends indicate a controlled training process with no apparent signs of unstable learning, while not implying uniform performance across all acne classes.

Figure 3(b) presents the progression of the model's F1-macro over 20 epochs. Both curves show an upward trend as training progresses, indicating an overall improvement in the model's classification performance as measured by the macro

F1 metric. During the early epochs, the validation F1-macro increases more noticeably, suggesting effective initial adaptation to the data structure.

In later epochs, the training F1-macro improves more gradually with minor fluctuations, while the validation F1-macro reaches its highest value at epoch 15, which is therefore selected as the best-performing checkpoint. The similar trajectories of the training and validation curves reflect stable evaluation behavior at the metric level rather than uniform performance across individual classes. Overall, the trend indicates a controlled learning process that enhances predictive quality up to an optimal point as defined by the validation F1-macro.

TABLE 5.
EVALUATION RESULTS PER EPOCH IN THE VALIDATION PROCESS

Epochs	Loss	Accuracy (%)	Precision (%)	Recall (%)	F1-Score Macro (%)
1	1.3469	28.6	26.8	27.7	25.3
2	1.2322	47.6	45.4	46.6	42.7
3	1.0186	64.3	71.0	63.9	62.9
4	0.8322	85.7	86.6	85.5	85.5
5	0.7727	71.4	75.4	71.4	70.0
6	0.7395	76.2	83.2	75.9	76.7
7	0.5715	85.7	87.9	86.1	86.0
8	0.5369	90.5	91.5	90.7	90.7
9	0.5151	90.5	92.9	90.9	90.7
10	0.4983	95.2	95.4	95.2	95.2
11	0.4752	95.2	95.6	95.2	95.2
12	0.5027	92.9	93.8	93.0	92.8
13	0.4957	90.5	91.9	90.2	90.5
14	0.5027	92.9	94.2	93.2	92.8
15	**0.4640**	**97.6**	**97.7**	**97.7**	**97.6**
16	0.4740	95.2	95.4	95.2	95.2
17	0.4454	95.2	95.4	95.2	95.2
18	0.4548	97.6	97.7	97.7	97.6
19	0.4424	97.6	97.7	97.7	97.6
20	0.4547	97.6	97.7	97.7	97.6

The per-epoch validation results presented in Table 3 demonstrate a clear improvement in model performance as reflected by the progression of evaluation metrics across training epochs. At the first epoch, the validation loss remains relatively high at 1.3469, accompanied by an accuracy of 28.6% and a macro-averaged F1 score of 25.3%, indicating limited initial predictive capability.

A substantial performance gain becomes evident at epoch 4, where the loss decreases to 0.8322, and the macro F1 score rises to 85.5%, alongside an accuracy of 85.7%. Between epochs 8 and 11, the validation loss further stabilizes within the range of 0.5369–0.4752, while accuracy improves to 90.5%–95.2%. During this interval, the macro F1 score consistently exceeds 90%, reflecting strong overall classification performance at the metric level.

The highest validation performance is achieved at epoch 15, with a loss value of 0.4640, accuracy of 97.6%, precision of 97.7%, recall of 97.7%, and a peak macro F1 score of 97.6%. This epoch is therefore selected as the optimal checkpoint based on validation results. In subsequent epochs, the macro F1 score remains high but does not surpass the performance observed at epoch 15, suggesting that the model has reached its optimal validation performance under the current training configuration.

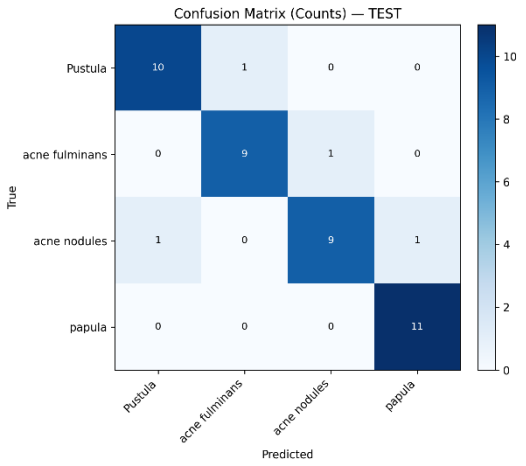


Figure 4. Confusion Matrix Results in the Testing Process

The confusion matrix on the testing set, presented in Figure 4, indicates that not all samples are correctly classified according to their respective categories. Several misclassifications are observed, including instances where Pustules are predicted as Acne fulminans, Acne fulminans as Acne nodules, and Acne nodules as either Pustules or Papules.

These error patterns suggest the presence of overlapping visual characteristics among certain acne types, particularly in cases where morphological differences are subtle. Despite these misclassifications, the number of correctly predicted samples remains dominant across the testing set, indicating that the model captures general discriminative patterns within the data.

Nevertheless, the observed confusion across specific classes highlights that the learned representations do not uniformly separate all acne categories. Therefore, the confusion matrix should be interpreted as evidence of strong overall predictive capability rather than consistent class-wise performance.

TABLE 6.
CLASSIFICATION REPORT RESULTS IN THE TESTING PROCESS

Acne Types	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Acne fulminans	90.70	90.0	90.0	90.0
Acne nodules	90.70	90.0	81.82	85.71
Papules	90.70	91.67	100.0	95.65
Pustules	90.70	90.91	90.91	90.91

The classification report presented in Table 6 complements the confusion matrix analysis by providing class-wise

performance metrics on the testing set. The overall accuracy remains constant at 90.70% across classes, while precision, recall, and F1-score exhibit noticeable variation among acne categories.

The Papules class achieves the highest recall value of 1.00, indicating that all Papules samples are correctly identified under the current evaluation setting. In contrast, the Acne nodules class records a lower recall of 81.82% and an F1-score of 85.71%, suggesting greater classification difficulty for this category. This discrepancy may reflect visual overlap with other acne types or a more limited number of representative samples.

These results indicate that, although the model demonstrates strong overall predictive performance on the testing data, class-wise behavior is not uniform. Therefore, the reported metrics should be interpreted as evidence of effective aggregate performance rather than perfectly consistent classification across all acne categories or guaranteed robustness under different data distributions.

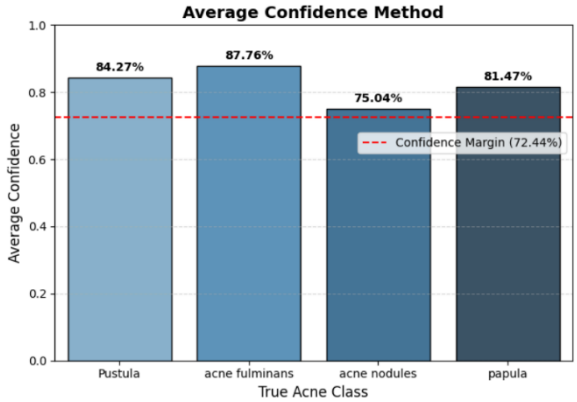


Figure 5. Result of Average Confidence Method

The average confidence analysis is illustrated in Figure 5 to examine the model's confidence level for predictions across acne classes. This measure is computed as the mean predicted probability assigned to correctly classified samples in the testing set, based on their ground-truth labels. A confidence margin of 72.44% is adopted as a reference threshold representing a minimally acceptable confidence level. All classes exhibit average confidence values above this margin, with scores of 84.27% for Pustules, 87.76% for Acne fulminans, 75.04% for Acne nodules, and 81.47% for Papules.

The highest confidence score is observed for the Acne fulminans class. This pattern may be attributed to its distinctive visual characteristics, such as pronounced inflammation, which can facilitate clearer separation from other acne categories. However, average confidence values should be interpreted cautiously, as high confidence does not necessarily imply uniform reliability across all samples or complete robustness to unseen data variations.

Overall, the confidence analysis suggests that the LoRA-based adaptation enables the model to maintain relatively well-separated feature representations, particularly for classes with more homogeneous visual patterns, while acknowledging residual uncertainty for categories with greater visual overlap.

The relatively high performance metrics observed in this study warrant careful interpretation, particularly given the limited size of the dataset. Nevertheless, several methodological factors may help mitigate the risk of overfitting. No substantial divergence was observed between training and validation performance during model optimization, indicating stable learning behavior rather than excessive memorization. In addition, the use of Parameter-Efficient Fine-Tuning through LoRA constrains the number of trainable parameters, thereby reducing model complexity and functioning as an implicit regularization mechanism. This strategy enables domain adaptation while preserving the pretrained representations learned by the DINOv2 foundation model, which itself benefits from large-scale self-supervised pretraining on diverse unlabelled data.

Despite these considerations, characteristics of the "skin-90" dataset may still introduce potential biases. The class distribution is imbalanced, with certain acne types represented by fewer samples, which may limit the model's ability to learn distinctive features for underrepresented categories. Moreover, visual similarities among acne classes pose additional challenges. Papules and Pustules often share comparable morphological traits, such as size and color, while differing mainly in subtle surface characteristics. Severe acne conditions may also exhibit overlapping visual patterns under varying lighting and acquisition conditions, increasing the likelihood of misclassification, particularly for minority classes. The absence of demographic metadata, including skin tone and acquisition context, further constrains the assessment of representativeness and generalizability to broader populations. Consequently, aggregate metrics such as accuracy and F1-score may not fully reflect class-specific performance.

In addition, this study does not include direct experimental comparisons with supervised convolutional or transformer-based baseline models trained on the same dataset. As a result, the observed performance cannot be interpreted as a comparative advantage attributable exclusively to the use of DINOv2 and LoRA. Instead, the findings demonstrate the feasibility of adapting a self-supervised vision foundation model for acne classification under limited data conditions. Incorporating baseline models in future work would enable a more comprehensive and comparative evaluation of the proposed approach.

B. Website Implementation

The results of the implementation and testing on the website are presented in Figure 6 and Table 7.

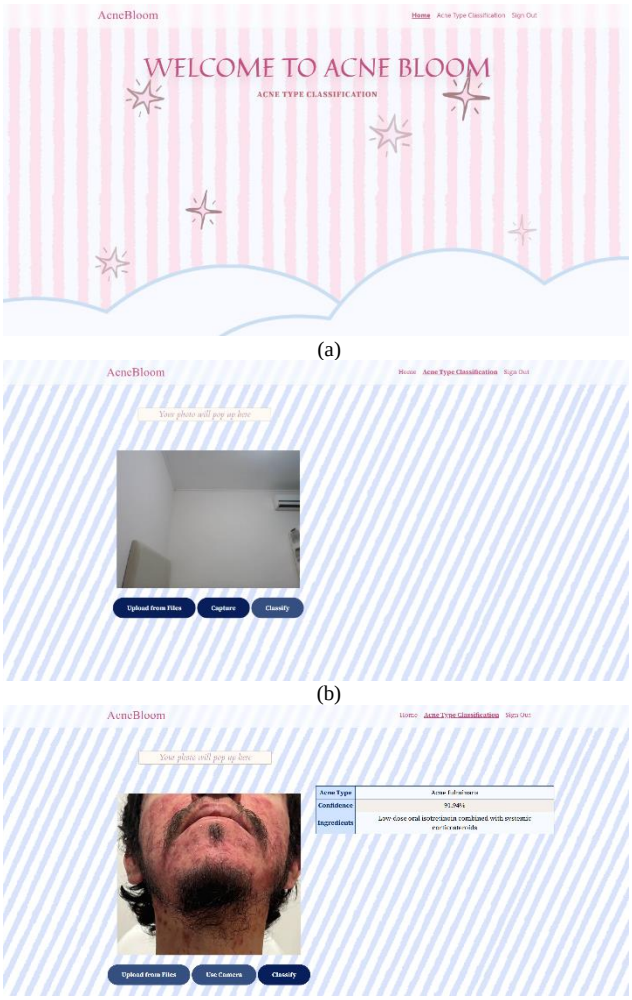


Figure 6. Result of Website Implementation: (a) Home Page, (b) Classification Page using Camera, and (c) Classification Page with the Result

Figure 6(a) shows the "Home" page, which serves as the main display after users successfully log into the application. The text "WELCOME TO ACNE BLOOM" is placed at the center as the primary welcoming element, accompanied by the phrase "ACNE TYPE CLASSIFICATION," emphasizing the application's role as an acne-type classification system. At the top of the page, a navigation bar contains the menus "Home," "Acne Type Classification," and "Sign Out," allowing users to easily access the main features. On the "Home" page for administrators, the navigation bar includes "Home," "Model Evaluation," "Skincare Ingredients Recommendation," and "Sign Out."

Although the system interface includes a feature labeled "Skincare Ingredients Recommendation," this component is intended solely for educational and informational purposes. The recommendations are derived from a literature-based mapping of acne categories and skincare ingredients, as reported in peer-reviewed dermatological studies, rather than from validated clinical guidelines or direct expert consultation [19], [20], [24]. Therefore, this feature should not be interpreted as providing medical or therapeutic advice. The proposed system is designed as a supportive learning tool to enhance general understanding of acne types, not as a diagnostic or clinical decision-making system. Users are advised to consult qualified dermatology professionals for medical evaluation and treatment decisions.

The classification output is presented in Figure 6(b) on the "Acne Type Classification" page, which displays the system's results after the facial image has been analyzed following the upload process. The results are presented in two main components: an image area on the left, displaying the user's face, and a classification table on the right, which provides information on the acne type, confidence score, and recommended skincare ingredients.

Figure 6(c) illustrates the "Acne Type Classification" page when users directly utilize the camera to capture their facial image as input for the classification process. A display area is provided to show the real-time camera feed.

TABLE 7.
TESTING RESULTS WITH BLACKBOX TESTING

No.	Scenario	Input	Expected Output	Testing Results
1.	The user registers with valid data.	A name, an unregistered email, and a valid password.	The account is successfully created, a registration success notification appears, and a verification link is sent to the user's email.	Successful
2.	The user registers using an email that is already registered.	A name, an email that has already been used, and a valid password.	The system rejects the registration and displays an error message indicating that the email is already registered.	Successful
3.	The user signs in with the correct credentials.	A valid email and the corresponding password.	The user is directed to the main page according to their role (user/admin).	Successful
4.	The user signs in with an incorrect password.	A valid email and an incorrect password.	The system displays an error message indicating that the sign-in attempt has failed.	Successful
5.	The user enables the "remember me" option during sign-in.	A valid email and password, with the "remember me" option checked.	The system stores the authentication token, allowing the user to remain signed in for the next session.	Successful
6.	The user submits a	Registered email.	The system sends a reset link to the	Successful

No.	Scenario	Input	Expected Output	Testing Results
	password reset request.		email and allows the password to be reset.	
7.	The user uploads an acne facial image in a supported format.	Image file in JPG/PNG format.	The system accepts the file, displays a preview, and is ready to perform the classification	Successful
8.	The user clicks the "Classify" button after uploading a valid image.	Facial image file.	The system displays the acne type, the confidence score, and the recommended skincare ingredients.	Successful
9.	The system displays recommendations based on the classification results.	Predicted acne type.	The recommended skincare ingredients appear according to the acne type.	Successful
10.	The admin opens the model evaluation page.	Valid admin credentials.	The system displays the model metadata, class distribution, dataset image previews, model metrics (accuracy, precision, recall, F1-score), the classification report, and the confusion matrix.	Successful
11.	The admin updates the list of skincare ingredient recommendations.	New recommendation data.	The system saves the changes and consistently displays the updated list.	Successful

Table 5 contains three main components in the black-box evaluation: the testing scenarios, the input types, and the expected system outputs. All testing scenarios received a "successful" status. Therefore, the website implementation is deemed to perform very well and is suitable for use.

Despite the promising classification performance, the interpretability of the proposed model remains limited. The DINOv2-based architecture operates as a deep vision transformer, where the learned representations are not directly interpretable in terms of clinically meaningful visual cues. This study does not include explicit visual explanation techniques, such as saliency maps or attention visualizations, which constrains the transparency of the model's decision-making process. Consequently, the predictions should be viewed as algorithmic outputs rather than explanatory clinical evidence.

IV. CONCLUSION

The results of this study indicate that integrating DINOv2 with a Parameter-Efficient Fine-Tuning strategy based on LoRA can achieve strong classification performance on the four evaluated acne categories within the "skin-90" dataset. Freezing the backbone preserves pretrained visual representations, while updating only the adapter layers and classification head provides an effective mechanism for domain adaptation under limited data conditions. The best-performing model was selected based on the stability of validation metrics at their optimal epoch.

Evaluation on the testing set yields an accuracy of 90.70%, with precision, recall, and F1-score values of 90.64%, 90.68%, and 90.57%, respectively. These results reflect reliable predictive performance at the aggregate level within the experimental setting. However, the reported metrics should be interpreted as indicative of effectiveness on the evaluated dataset rather than as evidence of uniform class-wise consistency or broad generalization. The findings suggest that the proposed architectural design and training configuration are suitable for capturing relevant visual patterns of acne, while further validation is required to assess robustness across more diverse data distributions. Future work may include systematic comparisons between LoRA-based parameter-efficient fine-tuning, full fine-tuning, and linear probing to quantitatively assess the trade-offs between computational efficiency and classification performance in acne image analysis.

REFERENCES

[1] G. V. Agustin, M. Ayub, and S. L. Liliawati, "Deteksi dan Klasifikasi Tingkat Keparahan Jerawat: Perbandingan Metode You Only Look Once," *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 3, pp. 2443–2229, 2024, [Online]. Available: http://dx.doi.org/10.28932/jutisi.v10i3.9414

[2] A. Quattrini, C. Boër, T. Leidi, and R. Paydar, "A Deep Learning-Based Facial Acne Classification System," *Clin. Cosmet. Investig. Dermatol.*, vol. 15, no. April, pp. 851–857, 2022, doi: 10.2147/CCID.S360450.

[3] A. Beauty and Y. Erlyana, "Perancangan Buku Berjudul 'My Acne Journey, Kenali, Atasi, Cegah Jerawat' Untuk Remaja Usia 13-18 Tahun Dengan Teknik Ilustrasi Digital," *J. Dimens. DKV Seni Rupa dan Desain*, vol. 7, no. 2, pp. 143–162, 2022, doi: 10.25105/jdd.v7i2.12813.

[4] I. G. Ayu, Y. Wahyuning, N. Made, S. Noviana, and N. T. Aryanata, "Hubungan Antara Citra Tubuh Dengan Kepercayaan Diri Pada Model Di Agency X makan dikarenakan keinginan besar untuk mengakibatkan terjadinya gangguan citra Citra Tubuh dengan Kepercayaan Diri pada Model di Agency X ". Penelitian ini bertujuan untuk mengeta," vol. 4, no. 1, pp. 8–13, 2023.

[5] E. M. Sipayung and E. Christopher R., "Klasifikasi Image Jenis Kayu pada Furnitur dengan Convolutional Neural Network," *J. Telemat.*, vol. 18, no. 2, pp. 82–87, 2024, doi: 10.61769/telematika.v18i2.617.

[6] T. Matius and S. Mulyana, "Penentuan Area Wajah Menggunakan Algoritma Viola-Jones Untuk Aplikasi Manipulasi Warna Citra [Face Area Determination Using Viola-Jones Algorithm For Digital Image Colour Manipulation Application]," *J. Algoritm. Log. dan Komputasi*, no. 01, pp. 743–750, 2025.

[7] L. Jing and Y. Tian, "Self-Supervised Visual Feature Learning with Deep Neural Networks: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 11, pp. 4037–4058, 2021, doi:

10.1109/TPAMI.2020.2992393.
[8] T. Uelwer *et al.*, "A Survey on Self-Supervised Representation Learning," pp. 1–48, 2023, [Online]. Available: http://arxiv.org/abs/2308.11455
[9] M. Baharoon, W. Qureshi, J. Ouyang, Y. Xu, A. Aljouie, and W. Peng, "Towards General Purpose Vision Foundation Models for Medical Image Analysis: An Experimental Study of DINOv2 on Radiology Benchmarks," pp. 1–22, 2023, [Online]. Available: http://arxiv.org/abs/2312.02366
[10] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," *37th Int. Conf. Mach. Learn. ICML 2020*, vol. PartF16814, no. Figure 1, pp. 1575–1585, 2020.
[11] K. He and F. Ai, "Momentum Contrast for Unsupervised Visual Representation Learning".
[12] J. Grill *et al.*, "Bootstrap Your Own Latent A New Approach to Self-Supervised Learning," vol. 200.
[13] J. Mohan, A. Sivasubramanian, S. V., and V. Ravi, "Enhancing skin disease classification leveraging transformer-based deep learning architectures and explainable AI," *Comput. Biol. Med.*, vol. 190, pp. 1–17, 2025, doi: 10.1016/j.compbiomed.2025.110007.
[14] A. R. Rafif Raif, R. E. Putra, A. Prapanca, and A. Qoiriah, "Penerapan DINOv2 pada Content Based Image Retrieval (CBIR) dalam Website Katalog Digital Batik Surabaya," *J. Informatics Comput. Sci.*, vol. 6, no. 03, pp. 678–687, 2024, doi: 10.26740/jinacs.v6n03.p678-687.
[15] B. Cui, M. Islam, L. Bai, and H. Ren, "Surgical-DINO: adapter learning of foundation models for depth estimation in endoscopic surgery," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 19, no. 6, pp. 1013–1020, 2024, doi: 10.1007/s11548-024-03083-5.
[16] I. G. N. L. W. K. Virna Dalira Br Sebayang, "Klasifikasi Jenis Jerawat Berdasarkan Citra Menggunakan Convolutional Neural Network dengan Arsitektur MobileNetV2," *J. FASILKOM*, vol. 14, no. 2, pp. 766–774, 2024, doi: 10.56211/helloworld.v3i2.518.
[17] A. Christopher and T. M. S. Mulyana, "Klasifikasi Tumbuhan Angiospermae Menggunakan Algoritma K-Nearest Neighbor Berdasarkan Pada Bentuk Daun," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 7, no. 4, pp. 1233–1243, 2022, doi: 10.29100/jipi.v7i4.3211.
[18] C. N. Putri, W. D. Qornain, F. Bamahri, G. E. Yuliastuti, and M. Kurniawan, "Klasifikasi Jenis Jerawat pada Data Citra Jerawat Wajah Menggunakan Convolutional Neural Network," *TIN Terap. Inform. Nusant.*, vol. 5, no. 2, pp. 172–181, 2024, doi: 10.47065/tin.v5i2.5231.
[19] J. Woźna, K. Korecka, J. Stępka, A. Bałoniak, R. Żaba, and R. A. Schwartz, "Acne fulminans treatment: case report and literature review," *Front. Med.*, vol. 11, no. July, pp. 1–7, 2024, doi: 10.3389/fmed.2024.1450666.
[20] Y. Li, X. Hu, G. Dong, X. Wang, and T. Liu, "Acne treatment: research progress and new perspectives," *Front. Med.*, vol. 11, no. July, pp. 1–9, 2024, doi: 10.3389/fmed.2024.1425675.
[21] A. Paichitrojjana, "Malassezia Folliculitis: A Review Article," *J. Med. Assoc. Thail.*, vol. 105, no. 2, pp. 160–167, 2022, doi: 10.35755/jmedassocthai.2022.02.13268.
[22] D. This, "Understanding OOD detection methods for image classification networks Eindhoven University of Technology Data Science and Artificial Intelligence," 2024.
[23] S. D. Pratama, L. Lasimin, and M. N. Dadaprawira, "Pengujian Black Box Testing Pada Aplikasi Edu Digital Berbasis Website Menggunakan Metode Equivalence Dan Boundary Value," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 6, no. 2, p. 560, 2023, doi: 10.53513/jsk.v6i2.8166.
[24] R. V Reynolds *et al.*, "FROM THE ACADEMY Executive summary: Guidelines of care for the management of acne vulgaris," *J Am Acad Dermatol*, vol. 90, pp. 1006–1010, 2024, [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/38300170/