Optimized Few-Shot Learning with Adversarial Feature Hallucination Networks for Multiclass Skin Disease Prediction Modeling

Billy Nicholas Panggiri
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
billy.panggiri@binus.ac.id

Steven Yap
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
steven092@binus.ac.id

Nur Afny Catur Andryani
Doctor of Computer Science Department
BINUS Graduate Program, BINUS University
Jakarta, Indonesia
nur.andryani@binus.edu

Retno Danarti
Indonesia Pediatric Dermatology Research Group
Faculty of Medicine
Gadjah Mada University
Yogyakarta, Indonesia
danarti@ugm.ac.id

2025 International Conference on Information and Communication Technology (ICoICT) | 979-8-3315-0323-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICoICT66265.2025.11192898

***Abstract*—Deep Learning (DL) is a popular learning method known for its effectiveness in classifying data and its ability to handle large input datasets. However, in real-world applications, the amount of data recorded and understood by computers may not always be abundant or sufficient especially on medical data. To address this issue, the Few-Shot Learning (FSL) technique can serve as a solution, as it enables classification by learning from only a few labeled samples. Dealing with the case of prediction modeling for skin's diseases identification, FSL has potential on addressing its limitation of valid annotated data. This paper presents the study of FSL potential on skin's diseases prediction modeling by adopting Adversarial Feature Hallucination Networks (AFHN) to optimize the FSL proposed framework performance. This study conducts a comparative analysis of several setups on the ISIC 2019 dataset to demonstrate the performance differences and advantages of the proposed model. The final results indicate an improvement in performance subject to AFHN optimization. The results showed that the model with feature synthesis achieved an accuracy of 76.99%, and this result was further supported by the evaluation metrics.**

***Keywords*—FSL, AFHN, cWGAN, hallucination based, skin disease**

I. Introduction

The skin is the outermost part of our body, proving that it is an inseparable part of our daily lives. Ironically, because it is the outermost layer, this also makes it the most vulnerable part to damage. The skin serves as the first line of defense against external chemical, physical, and microbial threats. Damage to the skin or the emergence of skin diseases can be caused by several factors, such as exposure to ultraviolet radiation, tanning, family history, environmental factors, alcohol, and more. Each of these factors can impact skin health, which in turn can affect a person's quality of life, mental health, and social interactions [1].

The World Health Organization [2] states that "Overall, skin conditions are estimated to affect 1.8 billion people at any point in time." This indicates that skin diseases are a common issue faced by every country, whether developed or developing. An experienced and skilled doctor is required to accurately identify a skin disease. However, not everyone has access to such expertise due to various factors. Therefore, an alternative system is needed to address this issue, allowing skin diseases to be identified as early as possible [3].

The era of machine learning and artificial intelligence is a rapidly advancing technology, including in the medical field. This technology can serve as a solution for classifying skin diseases. Such a system can learn from samples or images of a disease, classify them, and identify the condition. However, there are still critical requirements for developing this technology, one of which is the availability of samples or datasets as the foundation for learning. In the medical field, the process of collecting data or samples to create datasets presents its challenges [4].

Several factors present challenges in the process of collecting such datasets:

- **Diversity and variation in skin conditions:** This is influenced by factors such as skin color, environmental conditions, and age. Each of these elements contributes to variations that can pose challenges in classifying a disease [5].
- **Challenges in the annotation process:** In practice, the samples required by a machine may vary, and accurate, precise images are needed. Achieving this requires collaboration between medical experts and technology developers to reach the desired goal. This collection process can be both time-consuming and costly [6], [7].
- **Rare diseases:** Due to differences in environmental

979-8-3315-0323-9/25/$31.00 ©2025 IEEE

conditions and daily lifestyles, it is not uncommon for variations of rare diseases to occur. This can result in very few samples for certain classes [8].

These factors often result in imbalanced datasets, where certain classes have many samples while others have few, hindering model performance. To overcome this, methods such as Few Shot Learning have emerged, enabling disease recognition from limited data.

Few Shot Learning (FSL), is a machine learning system focused on creating technology capable of classifying objects using only a small amount of training data, yet producing outputs that classify objects in a generalizable manner [9].

In order to construct prediction modeling for skin's diseases identification with limited annotated data, this research object to adopt FSL framework. Motivated by the limited use of hallucination-based methods in few-shot medical image classification [10], we propose a feature-level hallucination approach using a GAN to synthesize support features. An optimization approach by implementing Adversarial Feature Hallucination Networks (AFHN) [11] is conducted. The experimentation presents the ability of AFHN to improve the FSL performance on identifying multi-classes skin's diseases.

II. Related Work

Technological advancements have led to the widespread use of Machine Learning, including mobile-based skin disease detection [12]. In dermatology, ML has shown performance comparable to general medical professionals [13], [14]. Despite its potential, implementing machine learning is challenging due to various requirements and costs [15], particularly the need for sufficient and relevant training data. In the medical field, data availability remains limited due to the rarity of certain diseases, emerging variants, and other factors [16].

Without meeting the requirements mentioned above, the output from machine learning could produce models that are not adaptive and are prone to making errors [17]. Due to the need for datasets and the scarcity of data for certain groups of skin diseases, researchers have been exploring other machine learning techniques that enable learning with fewer sample data as the basis for training [18]. One potential solution technique that can be applied is Few-Shot Learning.

Few-Shot Learning (FSL) enables classification on new tasks using minimal training data [19]. While FSL has been applied in the medical field, its use has largely focused on severe conditions like skin cancer [20]. However, research on FSL in dermatology remains limited, particularly for general skin diseases.

This research is important to conduct, considering the potential of transfer learning for each skin disease [16] and the challenges in diagnosing diseases with similar characteristics but different types [21]. Given these considerations, our goal is to conduct research focused on utilizing the FSL technique to classify a broader range of skin diseases.

III. Few Shot Learning

Few-Shot Learning (FSL) is a technique in machine learning aimed at training a model to recognize or classify new data using only a few training samples [22]. FSL is generally applied in cases where data collection is challenging, such as due to the rarity of an object or, as in this study, the rarity of a disease.

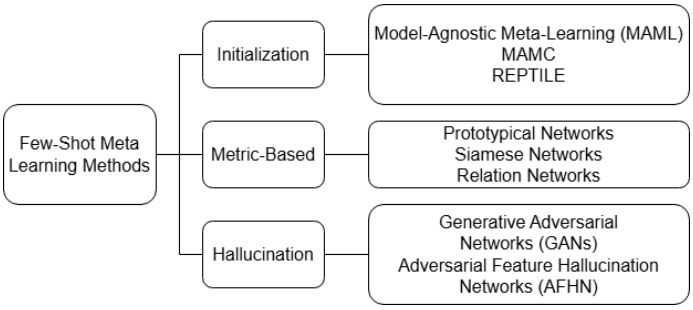

Fig. 1: Few Shot Meta Learning Methods

Based on the learning approach, FSL techniques can be divided into meta-learning and non-meta-learning frameworks. Meta-learning is a technique that focuses on training models or algorithms that can efficiently adapt to new tasks using a small number of examples, some distribution methods in this technique can be seen in Fig. 1. In contrast, non-meta-learning refers to techniques aimed at creating models that focus on specific tasks and typically utilize task-specific datasets [23]. Therefore, in this research, we apply hallucination-based learning, which falls under the meta-learning category. Several studies [10] related to this approach are as follows:

A. Initialization Based Methods

The initialization-based method focuses on creating a model, often referred to as a *meta-model*, which is trained using a parameter initialization system with a support set for fine-tuning the model and a query set to test the model's performance after fine-tuning. The ultimate goal of this approach is to realize the idea of "learning to learn" [24]. The workflow consists of the following stages:

1) Training Process: This stage involves the formation of the model, commonly referred to as *meta-training*. The training process consists of two types of iterations: the inner loop and the outer loop.

a) Inner Loop:: The model receives input from the support set, which serves as a task for the model. Each task triggers fine-tuning that generates gradients to adjust the parameters, enabling better generalization to the task. To evaluate performance after fine-tuning, the query set is used.

b) Outer Loop:: This stage utilizes multiple support sets and query sets with the aim of updating the meta-initialization used at the start of the inner loop.

The final result is the evaluation of the parameters from each inner loop, where the model has parameters that can be generalized to different tasks.

2) Testing Process: The testing process is the stage where the trained model is evaluated to ensure its performance on unseen tasks.

B. Metric Based Methods

This classification technique is carried out by comparing images based on proximity using a distance metric [25]. The workflow includes:

1) Feature Extraction: This stage involves extracting features from the data in the support set and the query set.

2) Distance Measurement: The extracted features of the support set are compared with those of the query set based on their proximity using a distance metric.

3) Evaluation Process: This stage classifies the query set based on comparisons with the support set, using distance metrics to assign each query image to the appropriate class.

C. Hallucination Based Methods

This technique is based on the concept of augmenting image classes with few samples by generating synthetic data. The synthetic data helps address the lack of samples in certain classes [26]. This method can reduce the risk of overfitting and ensure a more balanced sample distribution across all classes.

The common stages in this method include:

1) Feature Extraction: This process involves extracting features from the dataset to represent the data.

2) Generating Synthetic Data: Extracted features help generate synthetic data similar to the original.

3) Data Expansion: The synthetic data is combined with the original data to represent minority classes better, with the goal of balancing the data distribution.

4) Training Classifier: This stage trains the model using the expanded dataset for class classification tasks.

5) Testing: This involves evaluating the model's performance on new or unseen data.

Adversarial Feature Hallucination Few Shot Learning (AFHN), is a hallucination-based method that focuses on generating synthetic data for limited samples, thereby expanding the data distribution and increasing intraclass variation within those classes [11].

Unlike previous hallucination-based methods that aim to synthesize new images through augmentation systems, AFHN adopts the concept of augmenting and synthesizing the features of an image. This approach helps reduce computational load and emphasizes the critical information within an image [11]. In the medical field, this technique is useful for synthesizing variations in image information, such as lighting, angles, or other important details .

Ultimately, the primary goal of the above steps is to create a model with a more general learning foundation due to the presence of intraclass variation in each class within the support set. This enables the classification model to accurately classify unseen or new classes when new data is introduced.

IV. Methodology

This section describes the research methodology used to realize the proposed idea. Fig. 2 presents the general research flow conducted in this project, starting from the data preparation stage, data synthesis, and data classification using a prototypical network, up to the evaluation phase.

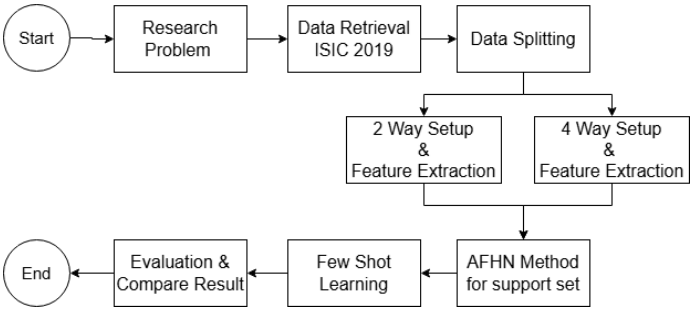

Fig. 2: Research's Flowchart

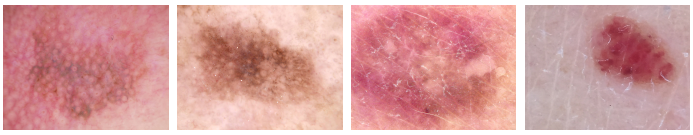

Fig. 3: ISIC 2019 Dataset

A. About Dataset

During the testing process of the model we developed, we evaluated it using several datasets related to skin diseases, including the International Skin Imaging Collaboration (ISIC) 2019 [27] [28] [29]. The dataset representations are shown in Fig. 3 and Table I.

B. Implementation Steps

In this section, we describe the implementation steps required to develop our model. The process involves data splitting, feature extraction, feature synthesis, classification, and evaluation. The overall workflow is depicted in Fig. 4.

TABLE I. Class Distribution In The ISIC 2019 dataset

Class	Sample
Melanocytic nevus	12,875
Melanoma	4,523
Basal cell carcinoma	3,323
Benign keratosis	2,624
Actinic keratosis	868
Squamous cell carcinoma	628
Vascular lesion	253
Dermatofibroma	239

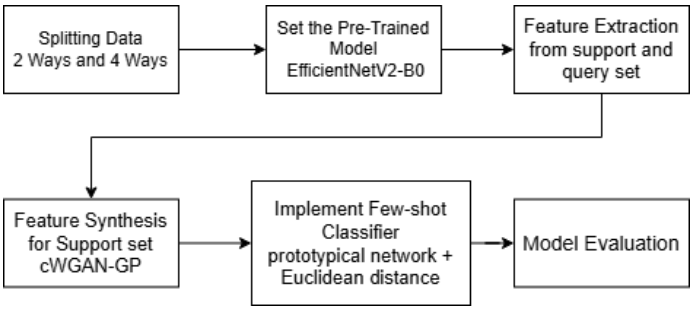

Fig. 4: AFHN's Workflow

1) Preparing Data: This stage involves preparing the data to align with the objectives of the code. Several adjustments are made, including:

- Incorporating the UNK class from the test data into the training data.
- Splitting the data into support set and query set for few-shot learning with 4-way and 2-way settings.

2) Training the Pre-trained Model: The initial model is built using EfficientNetV2-B0 as a feature extractor, chosen for its effectiveness in dermatology datasets [30]. Available in Keras via `tensorflow.keras.applications`, this model is pre-trained on ImageNet, providing a strong foundation for fine-tuning.

To adapt it for skin lesion classification, the following modifications are applied:

- The model is initialized with pre-trained ImageNet weights, removing its original classification layers.
- A global average pooling layer and a fully connected dense layer with softmax activation are added for multi-class classification.
- Adam optimizer with a learning rate of 0.0001 is used for training stability.
- The model is fine-tuned on the ISIC 2019 dataset (80-20 train validation split) for 10 epochs.
- Input images are resized to 224×224 and preprocessed using EfficientNetV2's built-in function.

3) Feature Loading: The extracted features and labels are stored as a NumPy dictionary. The features are represented as a multidimensional array, while the labels are stored as a separate array for class assignments.

4) Feature Synthesis for Support Set: Feature synthesis is performed using a Conditional Wasserstein Generative Adversarial Network with Gradient Penalty (cWGAN-GP). Unlike traditional GANs, WGAN-GP stabilizes training by enforcing a Lipschitz constraint using a gradient penalty term, while the conditional mechanism allows the generator to produce class-specific features.

The generator (G) takes a random noise vector ($z \in \mathbb{R}^{100}$) and a class label as input, generating synthetic feature vectors of dimension 1280, which match the output of EfficientNetV2-B0. The discriminator (D) receives both real/synthetic features and their corresponding class labels, learning to distinguish between them.

Key training components:

- Loss Function: The discriminator loss is computed as:

$$L_D = \mathbb{E}[D(x,y)] - \mathbb{E}[D(G(z,y),y)] + \lambda_{GP} \cdot GP \quad (1)$$

  where x is a real feature, $G(z,y)$ is a generated feature, and GP is the gradient penalty term. The generator loss is defined as:

$$L_G = -\mathbb{E}[D(G(z,y),y)] \quad (2)$$

- Gradient Penalty: The discriminator is regularized by enforcing a Lipschitz constraint via gradient penalty:

$$GP = \mathbb{E}\left[(\|\nabla_{\hat{x}} D(\hat{x},y)\|_2 - 1)^2\right] \quad (3)$$

  where $\hat{x}$ is an interpolation between real and generated features.
- Hyperparameters: A 100-dimensional latent space to provide the generator with sufficient randomness for producing realistic features, a common choice in successful GAN models, particularly in medical imaging tasks such as DSCLPGAN [31]. The generator outputs 1280-dimensional feature vectors to match the output of EfficientNetV2-B0, our selected feature extractor, ensuring seamless integration into the classification pipeline, as demonstrated by Singh, Guleria, and Sharma [32]. The model is trained with a batch size of 128 for 5000 epochs, using separate learning rates of 10^{-4} for the generator and 5×10^{-5} for the discriminator. We apply a gradient penalty coefficient (λ_{GP}) of 10 and update the discriminator three times for every generator update to stabilize training.

During training, real features are perturbed by Gaussian noise to improve discriminator robustness. The final trained generator synthesizes class-conditional features, enriching the query set for improved few-shot learning.

5) Few-Shot Classifier: The classification is performed using a prototypical network , which computes a class prototype as the mean feature vector of all samples within a class. The query sample is then classified based on the **Euclidean distance** between its representation of features and the prototypes of the class. The model is evaluated under three settings:

- **1-shot learning:** The classifier is trained with only one labeled example per class.
- **5-shot learning:** The classifier is trained with five labeled examples per class.
- **10-shot learning:** The classifier is trained with ten labeled examples per class.

Formally, the prototype for class k is computed as:

$$c_k = \frac{1}{|S_k|} \sum_{x_i \in S_k} f(x_i) \quad (4)$$

where S_k represents the support set samples of class k, and $f(x_i)$ is the feature embedding of sample x_i.

The classification of a query sample x_q is determined by the closest prototype using Euclidean distance:

$$d(x_q, c_k) = \|f(x_q) - c_k\|_2^2 \quad (5)$$

where the predicted class is given by:

$$\hat{y} = \arg\min_k d(x_q, c_k) \quad (6)$$

As shown in Equations (5) and (6), the classification process depends on the feature representations of query samples. To enhance these representations, we employ a Conditional WGAN-GP, which synthesizes query features using adversarial learning. The training procedure follows an iterative optimization of the generator and discriminator networks, as described in Algorithm 1.

After training the generator and discriminator, the next step involves evaluating the model's performance under various few-shot learning scenarios.

Algorithm 1 Conditional WGAN-GP Training

Input: Feature dataset $D_t = \{X_t, Y_t\}$, hyperparameters λ, α, β, number of epochs E, batch size B, latent dimension L, feature dimension F.
Output: Trained Generator G and Discriminator D.

1. Train G and D with dataset D_t.
while training is active **do**
  // Train D **while keeping** G **fixed**
  2. Draw batch (X, Y) from dataset.
  for each batch (X, Y) **do**
    3. Sample latent noise $Z \sim \mathcal{N}(0, 1)$ of size (B, L).
    4. Generate fake features $\tilde{X} = G(Z, Y)$.
    5. Compute D scores $D(X, Y)$ and $D(\tilde{X}, Y)$.
    6. Compute GP using real and fake features.
    7. Update D by maximizing the objective:

$$\mathcal{L}_D = \mathbb{E}[D(X, Y)] - \mathbb{E}[D(\tilde{X}, Y)] + \lambda GP.$$

  end for
  // Train G **periodically while keeping** D **fixed**
  if $e \bmod 3 = 0$ **then**
    8. Sample latent noise $Z \sim \mathcal{N}(0, 1)$.
    9. Generate fake features $\tilde{X} = G(Z, Y)$.
    10. Compute generator loss $\mathcal{L}_G = -\mathbb{E}[D(\tilde{X}, Y)]$.
    11. Update G by minimizing $\mathcal{L}_G$.
  end if
end while
12. Save trained models G and D **to disk**.

C. Simulation Setup

In this section, we outline the experimental setup used to evaluate the model. We utilize the ISIC 2019 dataset and perform simulations using different few-shot learning configurations, including 2-way and 4-way tasks with varying numbers of shots. Table II summarizes the key aspects of our simulation setup.

V. Result & Discussion

Formal classification performance metric are measured to evaluate the proposed FSL with optimization framework. The simulation's records are presented on Table III and Table IV. Those two tables comparing the FSL performance with and without optimization procedure on skin's diseases identification.

Additionally, we applied two different settings for running FSL, namely using the FSL process with a query set in 2-way and 4-way configurations, to compare the effectiveness of the model with different input setups.

TABLE II. Simulation Setup

Simulation Component	Details
Dataset	ISIC 2019
Ways & Shots	2-Way: 1, 5, 10 shots
	4-Way: 1, 5, 10 shots
Hardware	NVIDIA GeForce RTX 3060 Laptop GPU
CUDA Version	12.4
Evaluation Metrics	Precision, Recall, F1-score, Accuracy
Number of Evaluations	100

TABLE III. Prototypical Network FSL

Dataset	Ways & Shots	Accuracy	Precision	Recall	F1-score
ISIC (2019)	2 way 1 shot	56.50%	56.53%	56.50%	56.45%
	2 way 5 shot	71.70%	71.76%	71.70%	71.68%
	2 way 10 shot	74.60%	74.60%	74.60%	74.60%
	4 way 1 shot	36.00%	31.16%	31.00%	30.99%
	4 way 5 shot	44.75%	44.22%	44.20%	44.18%
	4 way 10 shot	49.00%	48.99%	48.92%	48.95%

TABLE IV. Prototypical Network FSL With AFHN Optimization

Dataset	Ways & Shots	Accuracy	Precision	Recall	F1-score
ISIC (2019)	2 way 1 shot	61.31%	61.99%	61.18%	60.61%
	2 way 5 shot	70.84%	71.44%	70.75%	71.12%
	2 way 10 shot	76.99%	76.22%	75.76%	75.71%
	4 way 1 shot	37.11%	37.21%	37.14%	34.82%
	4 way 5 shot	45.68%	45.74%	46.57%	44.80%
	4 way 10 shot	48.94%	48.25%	49.35%	47.81%

The first setup was conducted by selecting the seven classes with the largest sample sizes as the support set, while the remaining classes were placed into the query set. The results showed that the model could compete with the performance of models without synthetic data augmentation. The model with feature synthesis achieved an accuracy of 76.99%, and this result was further supported by other convincing evaluation metrics in table IV.

Meanwhile, in the 4-Way setup, four classes with the smallest sample sizes were assigned to the query set, while the remaining classes formed the support set. The results showed a performance decline compared to the model without feature synthesis this due the overlap happen.

the simulation results show the quite significance improvement of AFHN optimization is presented on 1-shot setting. The improvement around 5% can be achieved to this extremely low number of annotated data. However, the AFHN optimization is not significantly improve the performance of FSL on this skin's diseases identification setting with higher number of annotated dataset.

VI. Conclusion

FSL is a commonly used approach to addressing the challenge of data scarcity, allowing a model to perform similarly to humans by recognizing objects with only a few reference samples. In the medical field, where certain samples are rare and difficult to obtain, FSL is promising to explore. This research focuses on the study of FSL improvement by adopting the AFHN optimization procedure to handle multiclass skin disease prediction modeling with a limited data set. The series of simulation setup had been executed. It can be seen that AFHN consistently improves FSL performance in almost all settings in the ISIC skin disease dataset. However, we should further consider the more computational cost for 2 percent improvement on average.

ACKNOWLEDGMENT

This research is part of collaborative research project among BINUS University, Pusat Riset Komputasi BRIN (National Research and Innovation Agency) and Kelompok Studi Dermatologi Anak Per doski(Indonesian Society of Dermatology and Venereology) entitled "AI based pediatric Teledermatology development in Indonesia"

REFERENCES

[1] F. A. Mphande and F. A. Mphande, "Types of skin disease and their causes," *Skin Disorders in Vulnerable Populations: Causes, Impacts and Challenges*, pp. 23–64, 2020.

[2] World Health Organization, "Report of the first who global meeting on skin-related neglected tropical diseases: Geneva, switzerland, 27–31 march 2023," Geneva, Switzerland, 2023. [Online]. Available: https://www.who.int/publications/i/item/9789240091337

[3] N. A. C. Andryani, S. Achmad, M. K. Ario, S. Sun, L. Daniel, N. Constantine, M. Mazaya, S. P. Gondokaryono, D. Puspitosari *et al.*, "Ai-based paediatric teledermatology analysis and proposed framework," in *2023 IEEE International Biomedical Instrumentation and Technology Conference (IBITeC)*. IEEE, 2023, pp. 165–170.

[4] A. O. A. Deheyab, M. H. Alwan, I. k. A. Rezzaqe, O. A. Mahmood, Y. I. Hammadi, A. N. Kareem, and M. Ibrahim, "An overview of challenges in medical image processing," in *Proceedings of the 6th international conference on future networks & distributed systems*, 2022, pp. 511–516.

[5] D. Tandelon, D. L. Christiono, N. A. C. Andryani, S. P. Gondokaryono, J. I. Rahardja, and R. Danarti, "Ensemble-stacking machine learning model for multi-class skin diseases identification," in *2024 5th International Conference on Biomedical Engineering (IBIOMED)*. IEEE, 2024, pp. 187–192.

[6] O. Jaiyeoba, O. Jaiyeoba, E. Ogbuju, and F. Oladipo, "Ai-based detection techniques for skin diseases: A review of recent methods, datasets, metrics, and challenges," *Journal of Future Artificial Intelligence and Technologies*, vol. 1, no. 3, pp. 318–336, 2024.

[7] A. Kalaivani and S. Karpagavalli, "Detection and classification of skin diseases with ensembles of deep learning networks in medical imaging," *International journal of health sciences*, no. I, p. 431085, 2022.

[8] T.-C. Kuo, P.-H. Wang, Y.-K. Wang, C.-I. Chang, C.-Y. Chang, and Y. J. Tseng, "Rsdb: a rare skin disease database to link drugs with potential drug targets for rare skin diseases," *Scientific Data*, vol. 9, no. 1, p. 521, 2022.

[9] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," *ACM computing surveys (csur)*, vol. 53, no. 3, pp. 1–34, 2020.

[10] E. Pachetti and S. Colantonio, "A systematic review of few-shot learning in medical imaging," *Artificial intelligence in medicine*, p. 102949, 2024.

[11] K. Li, Y. Zhang, K. Li, and Y. Fu, "Adversarial feature hallucination networks for few-shot learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 13 470–13 479.

[12] R. Pangti, J. Mathur, V. Chouhan, S. Kumar, L. Rajput, S. Shah, A. Gupta, A. Dixit, D. Dholakia, S. Gupta *et al.*, "A machine learning-based, decision support, mobile phone application for diagnosis of common dermatological diseases," *Journal of the European Academy of Dermatology and Venereology*, vol. 35, no. 2, pp. 536–545, 2021.

[13] M. P. Salinas, J. Sepúlveda, L. Hidalgo, D. Peirano, M. Morel, P. Uribe, V. Rotemberg, J. Briones, D. Mery, and C. Navarrete-Dechent, "A systematic review and meta-analysis of artificial intelligence versus clinicians for skin cancer diagnosis," *NPJ Digital Medicine*, vol. 7, no. 1, p. 125, 2024.

[14] N. A. C. Andryani, M. F. Hidayatullah, W. Yen, O. J. Terawan, M. Mazaya, D. Puspitosari, and S. P. Gondokaryono, "Analysis of segmentation method for skin's lesions identification," in *2023 International Conference on Information Technology and Computing (ICITCOM)*. IEEE, 2023, pp. 369–374.

[15] H. K. Jeong, C. Park, R. Henao, and M. Kheterpal, "Deep learning in dermatology: a systematic review of current approaches, outcomes, and limitations," *JID Innovations*, vol. 3, no. 1, p. 100150, 2023.

[16] N. A. C. Andryani, F. H. Juwono, and P. J. Sianturi, "Artificial intelligence approach for automatic multiclass skin diseases identification," *Journal of Engineering Science and Technology*, vol. 18, pp. 1–14, 2023.

[17] E. R. Gordon, M. H. Trager, D. Kontos, C. Weng, L. J. Geskin, L. S. Dugdale, and F. H. Samie, "Ethical considerations for artificial intelligence in dermatology: a scoping review," *British Journal of Dermatology*, p. ljae040, 2024.

[18] L. Alzubaidi, J. Bai, A. Al-Sabaawi, J. Santamaría, A. S. Albahri, B. S. N. Al-dabbagh, M. A. Fadhel, M. Manoufali, J. Zhang, A. H. Al-Timemy *et al.*, "A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications," *Journal of Big Data*, vol. 10, no. 1, p. 46, 2023.

[19] S. Woerner and C. F. Baumgartner, "Navigating data scarcity using foundation models: A benchmark of few-shot and zero-shot learning approaches in medical imaging," in *International Workshop on Foundation Models for General Medical AI*. Springer, 2024, pp. 30–39.

[20] S. Gopikha and M. Balamurugan, "Identification of multi-class skin cancer classification using relation based embedding network with few shot learning," *Webology*, vol. 18, no. 6, pp. 5520–5536, 2021.

[21] C. B. Cobb, C. R. Heath, A. S. Byrd, L. J. McKinley-Grant, V. Callender, A. S. Adamson, S. Brown, S. R. Desai, D. A. Glass, T. Jaleel *et al.*, "The skin of color society's meeting the challenge summit, 2022: diversity in dermatology clinical trials proceedings," *JAMA dermatology*, vol. 159, no. 7, pp. 757–762, 2023.

[22] S. X. Hu, D. Li, J. Stühmer, M. Kim, and T. M. Hospedales, "Pushing the limits of simple pipelines for few-shot learning: External data and fine-tuning make a difference," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 9068–9077.

[23] A. Parnami and M. Lee, "Learning from few examples: A summary of approaches to few-shot learning," *arXiv preprint arXiv:2203.04291*, 2022.

[24] J. Chen, W. Yuan, S. Chen, Z. Hu, and P. Li, "Evo-maml: Meta-learning with evolving gradient," *Electronics*, vol. 12, no. 18, p. 3865, 2023.

[25] M. J. Lee and J. So, "Metric-based learning for nearest-neighbor few-shot image classification," in *2021 International Conference on Information Networking (ICOIN)*. IEEE, 2021, pp. 460–464.

[26] J. Wu, N. Dong, F. Liu, S. Yang, and J. Hu, "Feature hallucination via maximum a posteriori for few-shot learning," *Knowledge-Based Systems*, vol. 225, p. 107129, 2021.

[27] P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, p. 180161, 2018.

[28] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)," *arXiv*, 2017.

[29] C. Hernández-Pérez, M. Combalia, S. Podlipnik, N. C. Codella, V. Rotemberg, A. C. Halpern, O. Reiter, C. Carrera, A. Barreiro, B. Helba, S. Puig, V. Vilaplana, and J. Malvehy, "Bcn20000: Dermoscopic lesions in the wild," *Scientific Data*, vol. 11, no. 1, p. 641, Jun 2024.

[30] A. Winata, N. A. C. Andryani, A. A. Santoso, and F. L. Gaol, "Efficiency analysis of ai model compression for edge teledermatology," in *Proceedings of the 2024 International Conference on Data Engineering and Enterprise System*, 2025, in press.

[31] J. Zafar, V. Koc, and H. Zafar, "Dual-stream contrastive latent learning generative adversarial network for brain image synthesis and tumor classification," *Journal of Imaging*, vol. 11, no. 4, p. 101, 2025.

[32] G. Singh, K. Guleria, and S. Sharma, "A pre-trained efficientnetv2b0 model for the accurate classification of fake and real images," in *2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2024, pp. 1082–1086.