

Deep learning-based classification of acanthosis nigricans as a clinical sign of high insulin resistance in type 2 diabetes mellitus

Annisa Darmawahyuni^{1,2}, Siti Nurmaini^{1,2}, Sutarno Sutarno^{1,2}, Anggun Islami^{1,2},
Muhammad Raffly Riyansyah Putera^{1,2}, Muhammad Mukhlashyn Az Zahrawi³

¹Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya, Palembang, Indonesia

²Artificial Intelligence-Medical Center of Excellence, Universitas Sriwijaya, Palembang, Indonesia

³Department of Informatics Engineering, Faculty of Computer Science, Sriwijaya University, Palembang, Indonesia

Article Info

Article history:

Received Mar 2, 2026

Revised Aug 27, 2026

Accepted Sep 3, 2026

Keywords:

Acanthosis nigricans

Artificial intelligence

Deep learning

Health

Type 2 diabetes mellitus

ABSTRACT

Acanthosis nigricans (AN) is a cutaneous manifestation strongly linked to insulin resistance and often serves as an early indicator of type 2 diabetes mellitus (T2DM). Early detection of AN is therefore essential for preventive healthcare. Conventional diagnosis typically relies on visual inspection and grading, which may be subjective and inconsistent. To address this limitation, this study proposes a computer-aided diagnostic framework based on deep learning for automatic AN detection from skin fold images. A comprehensive evaluation was conducted using nine widely adopted deep learning architectures, namely VGG16, VGG19, DenseNet121, DenseNet201, ResNet50, ResNet101, MobileNetV2, Xception, and the vision transformer (ViT). To ensure a fair comparison, all models were trained under identical experimental settings. Model performance was assessed using standard classification metrics. The comparative analysis demonstrates that ResNet101 outperformed the other architectures, achieving an accuracy of 94.12% for AN detection, with a precision and recall of 96.55% and an F1-score of 93.33%. Overall, the results highlight the robustness and effectiveness of convolutional neural network (CNN) based models, particularly ResNet101, as non-invasive and objective tools for automated AN screening. This approach has the potential to support early detection of individuals at risk of T2DM and to enhance preventive healthcare delivery.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Annisa Darmawahyuni

Artificial Intelligence-Medical Center of Excellence, Universitas Sriwijaya

Palembang, 30139, Indonesia

Email: riset.annisadarmawahyuni@gmail.com

1. INTRODUCTION

Acanthosis nigricans (AN) is one of the crucial signs of high insulin resistance that is involved in the pathogenesis of type 2 diabetes mellitus (T2DM) [1], [2]. Since increased insulin resistance is usually a harbinger of T2DM, the detection of insulin resistance at an early age allows us to intervene and institute appropriate interventions [3], [4]. Clinically, the neck, axilla, and knuckles are the most common sites linked to insulin resistance; however, reports of unexpected areas, including the face, have also been made [5]–[7]. Early detection of insulin resistance may be possible with the AN observation and presence as a screening method.

AN was first described in association with obesity in 1889 by Gerson Unna and Pollitzer [8], [9]. The condition is characterized by hyperpigmented lesions occurring predominantly in skin folds [10].

The most prominent clinical manifestation is cutaneous hyperpigmentation, which has often been misinterpreted as a sign of poor personal hygiene. Assessment of AN is generally performed noninvasively through visual inspection and grading, although skin biopsy may occasionally be employed for diagnostic confirmation [11]. To quantify disease severity, Burke *et al.* [12] proposed a grading scale that considers the distribution, texture, and extent of hyperpigmentation. However, consistent and reproducible results have been achieved primarily by grading the severity of AN at the neck using a scale ranging from 0 to 4 [13]. In contrast to earlier methods, which focused solely on determining the presence or absence of AN, this approach provides a semiquantitative classification system [14].

Despite its clinical relevance, conventional AN assessment remains inherently subjective and heavily dependent on clinician experience, lighting conditions, and patient skin tone [15]. These factors can introduce substantial inter- and intra-observer variability, particularly in early or mild cases where visual differences are subtle [6]. Such variability poses challenges for consistent diagnosis, longitudinal monitoring, and large-scale epidemiological screening, especially in primary care and community-based settings. From a public health perspective, the global increase in obesity and metabolic syndrome has led to a parallel rise in insulin resistance and T2DM, including among children and adolescents [16]. Early identification of individuals at risk is therefore essential to reduce long-term complications and healthcare burden. In this context, AN represents a visible and non-invasive biomarker that can be leveraged for early screening; however, its potential remains underutilized due to the lack of objective and scalable assessment tools [17].

Recent advances in artificial intelligence (AI) and computer vision offer promising opportunities to overcome these limitations. Deep learning methods, particularly convolutional neural networks (CNNs), have demonstrated superior performance in medical image analysis by automatically learning discriminative features from raw images [18]. Unlike traditional image processing approaches that rely on handcrafted features, CNNs can capture complex spatial patterns, texture variations, and color distributions that are critical for dermatological assessment [18]–[25]. Several studies have successfully applied deep learning to skin-related conditions, including melanoma detection, assessment of psoriasis severity, and analysis of hyperpigmentation, achieving performance comparable to or exceeding that of human experts. These findings highlight the suitability of CNN-based models for analyzing subtle dermatological cues, such as those present in AN, where variations in pigmentation and skin texture may be difficult to quantify visually [26].

Nevertheless, automated AN detection presents unique challenges. Variability in skin tone, illumination, camera quality, and anatomical location can significantly affect image appearance. Furthermore, the scarcity of publicly available, well-annotated AN dataset limits the generalizability and reproducibility of existing approaches. Addressing these challenges requires a systematic evaluation of robust model architectures that can generalize well across diverse imaging conditions. In recent years, transformer-based models, such as the vision transformer (ViT), have emerged as powerful alternatives to CNNs in computer vision tasks [27]–[30]. By modeling long-range dependencies through self-attention mechanisms, transformers offer enhanced capability in capturing global contextual information, which may complement or outperform traditional CNN architectures in certain medical imaging applications. However, their effectiveness for dermatological image classification, particularly in detecting AN, remains underexplored. Therefore, this study proposes a computer-aided diagnostic framework for detecting AN from human skin fold images using deep learning. The proposed approach evaluates nine widely adopted architectures, including VGG16, VGG19, DenseNet121, DenseNet201, ResNet50, ResNet101, MobileNetV2, Xception, and ViT. By comparing CNN-based and transformer-based models under the same experimental setting, this study provides a comprehensive analysis of their relative strengths and limitations for AN classification. The main contributions of this study are twofold:

- Propose a comparative analysis of state-of-the-art CNN and transformer-based architectures for AN classification, and
- Develop a robust deep learning framework that demonstrates the feasibility of automated, non-invasive screening of AN.

2. MATERIAL AND METHOD

The research methodology can be presented in Figure 1, which consists of: 1) the data preparation of AN images sourced from public data at Roboflow [31], 2) the image preprocessing that includes the image filtering for denoising, resizing, augmentation, and image normalization, 3) the proposed deep learning architecture for AN classification, and 4) the evaluation of the proposed model.

2.1. Data preparation

The processes in developing AN AI model for early detection of DMT2 through visual identification of AN images as shown in Figure 2. The process starts with the input of a raw image obtained from the posterior area of the patient's neck, which is a common location for visual indicators of AN. The researcher

has downloaded the AN image used in this study, sourced from public data at Roboflow [32]. AN lesion can appear in various areas of the body, especially in skin folds such as the back of the neck, armpits, groin, elbows, and under the breasts. Overall, the initial dataset consists of 800 images divided into three subsets with proportions of 560 images for training data, 160 images for validation data, and 80 images for testing data.

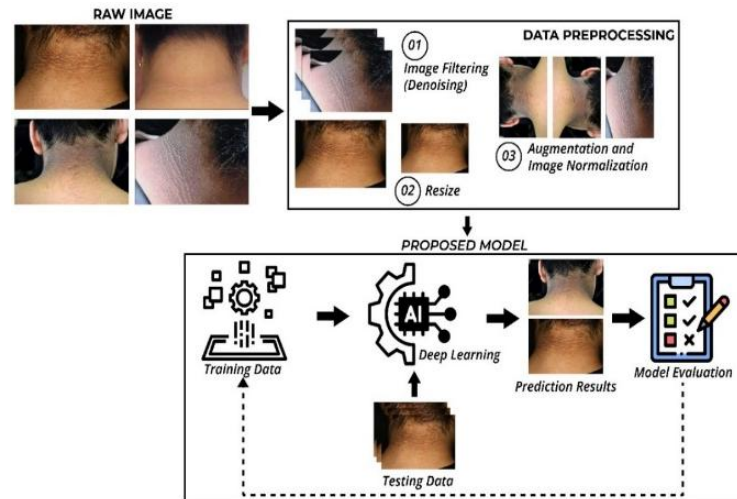


Figure 1. The methodology research of AN and normal condition classification



Figure 2. AN condition of the posterior area of the adult neck [32]

2.2. Data preprocessing

Before model training, several preprocessing steps are carried out to prepare the data to be clean, uniform, and in accordance with the model input format. First, a image filtering process to exclude irrelevant images, such as blurry, dark, or unclear images, mislabeled images, image sizes that are too small, and images without annotations or with inaccurate annotations. Second, the entire remaining image is then resized to a size of 395 x 395 pixels, as it is close enough to your average image height of 395.69 pixels. Third, image augmentation to improve the robustness of the model. The augmentation process included random rotations of up to 20, horizontal and vertical translations of up to 10%, shear transformations with a shear intensity of 0.2, and zooming in or out by up to 20%. In addition, horizontal flipping was performed to enhance variability. For regions outside the original image that appeared as a result of these transformations, the nearest pixel values were used to fill the empty areas (fill mode = 'nearest'). Fourth, image normalization from the range [0, 255] to [0, 1] using ImageDataGenerator(rescale=1./255). The sample of before and after data preprocessing can be presented in Figures 3 and 4.

Figures 3(a) and 4(a) present the original AN and normal condition images, respectively, prior to preprocessing. These raw AN and normal condition images show variations in lighting, resolution, and

background, which may hinder consistent analysis. The AN images capture the affected areas around the neck, highlighting the darkened and thickened skin characteristic of AN. Figures 3(b) and 4(b) displays the same images after undergoing preprocessing techniques, including image filtering, resizing, augmentation, and normalization. As a result, the processed images appear clearer, standardized in size, and with more uniform brightness and contrast. These enhancements improve the visual quality and ensure greater consistency, making the dataset more suitable for further computational analysis.

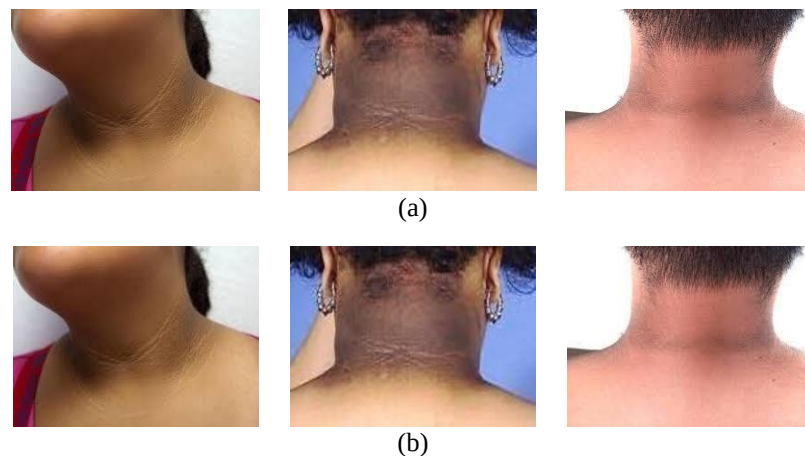


Figure 3. AN condition before and after data preprocessing (image filtering, resize, augmentation and normalization) (a) before data preprocessing and (b) after data preprocessing

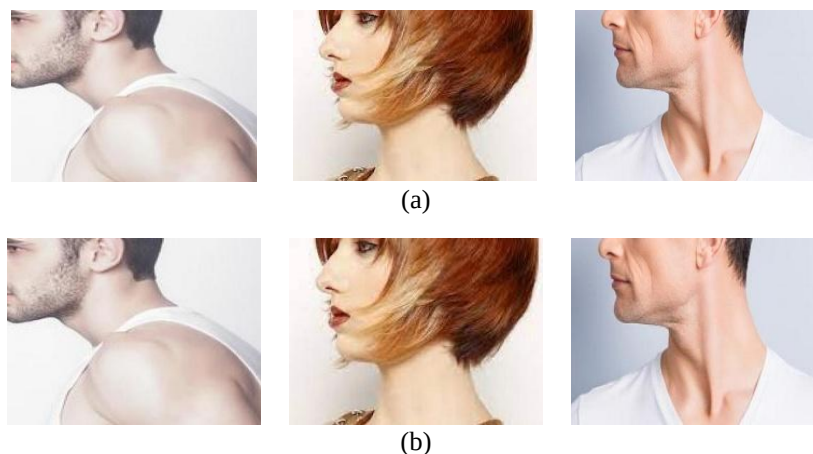


Figure 4. Normal condition before and after data preprocessing (image filtering, resize, augmentation, and normalization) (a) before data preprocessing and (b) after data preprocessing

2.3. Deep learning architectures

This study experimented with nine deep learning architectures, namely VGG16, VGG19, DenseNet121, DenseNet201, ResNet50, ResNet101, MobileNetV2, Xception, and the ViT. These models represent diverse architectural paradigms, including classical CNN with stacked convolutional layers (VGG16 and VGG19). These residual learning frameworks enable stable training of profound networks (ResNet50 and ResNet101), densely connected architectures that encourage feature reuse and parameter efficiency (DenseNet121 and DenseNet201), lightweight models optimized for computational efficiency (MobileNetV2), and advanced convolutional designs based on depthwise separable convolutions (Xception). In contrast to convolution-based approaches, the ViT employs a self-attention mechanism to capture global contextual relationships across image patches. All architectures were trained from scratch without pretrained weights, ensuring that the learned representations are fully adapted to the characteristics of the local dataset and enabling a fair comparison of their intrinsic learning capabilities under identical experimental conditions. The descriptions of these architectures are listed in Table 1.

Table 1. The description of deep learning architectures

Model	Architecture type	Key idea	Main characteristics
VGG16	CNN	Stacked 3×3 convolutions	Simple and stable, but computationally heavy
VGG19	CNN	Deeper VGG network	Richer features with higher memory cost
ResNet50	Residual CNN	Skip (residual) connections	Stable deep training, widely used
ResNet101	Residual CNN	Deeper residual learning	Strong representation for complex tasks
DenseNet121	Dense CNN	Dense feature concatenation	Parameter-efficient with feature reuse
DenseNet201	Dense CNN	Deeper dense connections	Improved representation, higher cost
MobileNetV2	Lightweight CNN	Depthwise separable convolutions	Fast and efficient for limited resources
Xception	Advanced CNN	Extreme depthwise separable conv	High accuracy–efficiency trade-off
Vision Transformer (ViT)	Transformer	Self-attention over image patches	Strong global context modeling

2.4. Platform

The experimental setup consisted of an Intel Core i9-14900K CPU (32 cores, 3.2 GHz), 32 GB of RAM, and an NVIDIA A10 GPU with 24 GB of memory. All experiments were conducted on Windows 11 Pro 64-bit using Python 3.9.18, with model implementation in PyTorch and supporting libraries including NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, and Roboflow

3. RESULTS AND DISCUSSION

The total dataset for AN and normal conditions consists of 800 images, with 70% allocated to the training set, 20% to the validation set, and the remaining 10% to the testing/unseen set. The study explored nine common CNN's architectures for AN classification, i.e., VGG16, VGG19, DenseNet121, DenseNet201, ResNet50, ResNet101, MobileNetV2, Xception, and ViT. The comparison of each architecture is evaluated to obtain the best model. The architectures used the same hyperparameters, such as 100 epochs, the adaptive moment estimation (Adam) optimizer, and binary cross-entropy loss function. The performance results of AN and normal condition (healthy, H) classification using a testing/unseen set can be listed in Table 2. Based on the comparison of several deep learning models in Table 2, it can be seen that ResNet101 demonstrates the best overall performance. For the AN class, this model achieves a very high 94.12% accuracy with balanced both precision and recall of 96.55%, resulting in an excellent F1-score of 93.33%. Similarly, for the H class, ResNet101 also performs consistently well, reaching 90.91% accuracy and a strong F1-score of 93.02%. These results indicate that ResNet101 is highly reliable and effective in distinguishing between abnormal and healthy cases.

Table 2. The performance results of an and normal condition (Healthy, H) classification using a Testing/Unseen

Model	Class	Performance (%)			
		Accuracy	Precision	Recall	F1-Score
VGG16	AN	90.20	96.30	86.67	91.23
	H		83.33	95.24	88.89
VGG19	AN	86.27	89.66	86.67	88.14
	H		81.82	85.71	83.72
DenseNet121	AN	84.31	92.31	80.00	85.71
	H		76.00	90.48	82.61
DenseNet201	AN	88.24	92.86	86.67	89.66
	H		82.61	90.48	86.36
ResNet50	AN	88.24	90.00	90.00	90.00
	H		85.71	85.71	85.71
ResNet101	AN	94.12	96.55	96.55	93.33
	H		90.91	95.24	93.02
MobileNet V2	AN	88.24	96.15	83.33	89.29
	H		80.00	95.24	86.96
Xception	AN	86.27	100.00	76.67	86.79
	H		75.00	100.00	85.71
ViT	AN	86.27	87.10	90.00	88.52
	H		85.00	80.95	82.93

The confusion matrix (CM) heatmap of nine common CNN's architecture for AN classification can be displayed in Figure 5. Figure 5(a)-(i) displayed among the tested models, ResNet101 achieved the highest overall performance, with an accuracy of 94.12% for AN and 90.91% for H. The CM shows that ResNet101 correctly classified 28 abnormal and 20 healthy samples, with only three misclassifications in total.

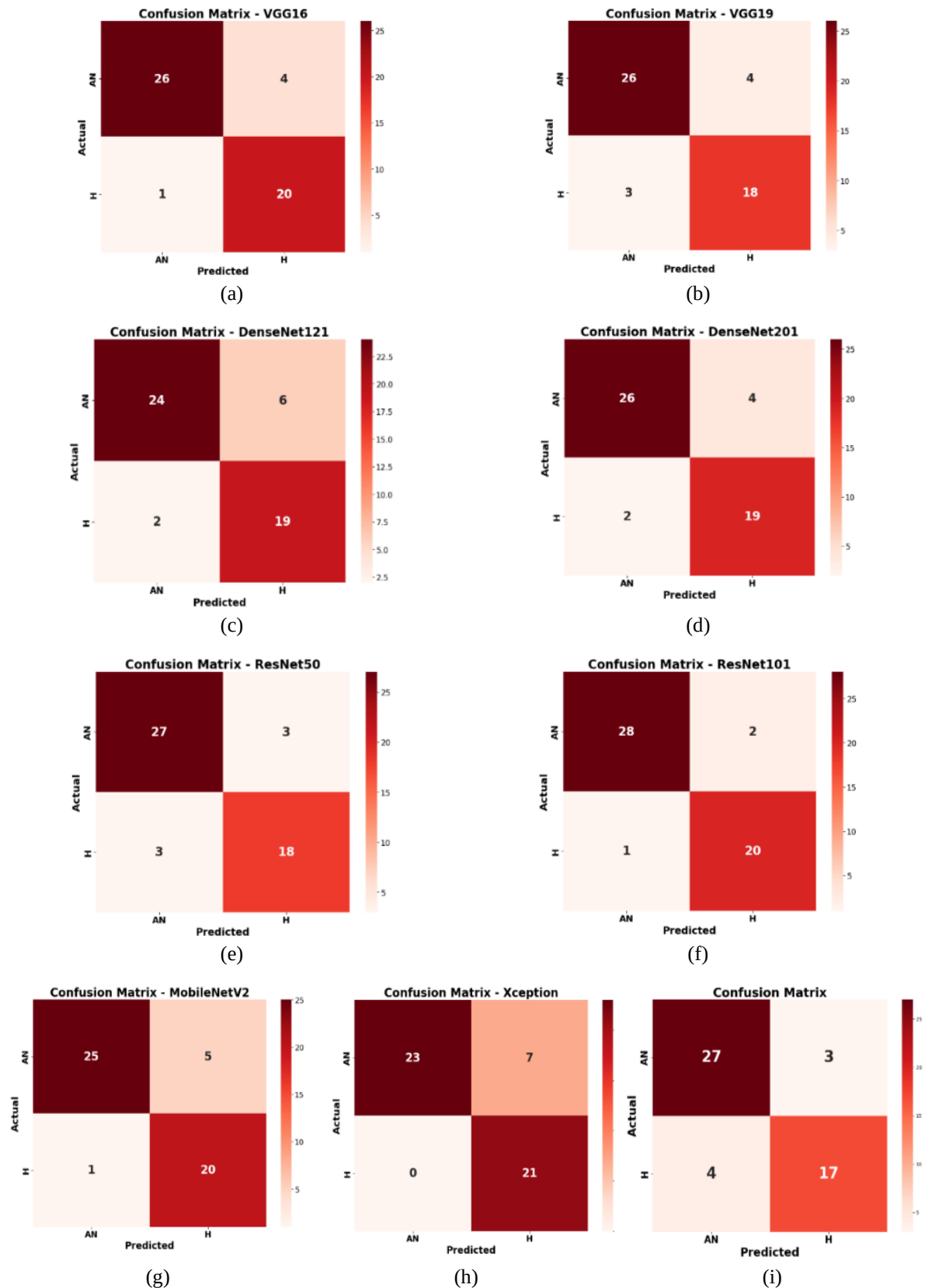


Figure 5. The CM heatmap of nine common CNN's architecture for AN classification (a) VGG16, (b) VGG19, (c) DenseNet121, (d) DenseNet201, (e) ResNet50, (f) ResNet101, (g) MobileNetV2, (h) Xception, and (i) ViT]

These results are consistent with its superior precision, recall, and F1-scores, confirming ResNet101 as the most robust and reliable model across both classes. VGG16 also produced competitive results, correctly identifying 26 AN and 20 H class. The model demonstrated excellent precision, with only one false positive for the H class. However, its recall was slightly lower due to four abnormal samples misclassified as healthy. Despite this limitation, VGG16 remains a strong alternative, especially when computational simplicity is considered. ResNet50 exhibited similarly balanced outcomes, correctly classifying 27 AN and 18 H cases. Although slightly less accurate than ResNet101, it still provides consistent and stable classification performance.

DenseNet201 classified 26 AN and 19 H class correctly, while DenseNet121 and VGG19 showed more frequent misclassifications, particularly for the abnormal class. This reduction in sensitivity indicates a weaker ability to detect abnormalities compared to the ResNet and VGG architectures. In contrast, MobileNetV2 and Xception displayed a distinctive trend. MobileNetV2 achieved very high precision for the H class, correctly classifying 20 out of 21 H class, but it misclassified five AN class as healthy. Xception demonstrated an even stronger precision-oriented behavior, achieving perfect classification of all 21 H class with zero false negatives. However, it misclassified seven AN case, leading to reduced recall. These findings suggest that MobileNetV2 and Xception may be advantageous in applications where minimizing false positives is prioritized, but their lower sensitivity to abnormal cases poses a risk in clinical screening scenarios where recall is critical. ViT provided balanced yet less consistent performance compared to ResNet101. It correctly classified 27 AN and 17 H class, with higher misclassification rates for the H class. Although ViT achieved reasonable F1-scores, it did not outperform the convolutional architectures, particularly the deeper ResNet models. Overall, the results indicate that ResNet101 offers the most balanced and accurate classification, making it the most suitable model for medical screening tasks where both high recall and precision are crucial.

Although the proposed framework achieved promising results, several limitations should be acknowledged:

- a) The dataset size may be limited, which could affect the generalizability of the model to broader populations with diverse demographics, skin tones, and imaging conditions,
- b) The evaluation was conducted under controlled experimental settings, and external validation using independent clinical datasets was not performed. This restricts the ability to confirm robustness in real-world environments, and
- c) The study focused solely on image-based classification without incorporating clinical or biochemical data, which may provide additional context for accurate diagnosis of insulin resistance.

4. CONCLUSION

This study demonstrated the feasibility of applying deep learning methods for the automated detection of AN from skin fold images. By conducting a comparative evaluation of nine state-of-the-art CNN and transformer-based architectures, the findings highlight the superiority of ResNet101, which achieved the highest accuracy, precision, recall, and F1-scores for both AN and normal conditions. These results underscore the potential of CNN-based approaches, particularly ResNet101, as reliable and non-invasive tools for early screening of AN.

The outcomes further suggest that deep learning frameworks can complement traditional clinical examinations by offering objective, reproducible, and scalable diagnostic support. Importantly, the integration of such models into clinical or telemedicine platforms could enhance early detection of insulin resistance, thereby contributing to preventive measures against T2DM. In addition, further validation in real-world clinical settings is essential to confirm the robustness and usability of the proposed framework. Future work may also explore the integration of multimodal data, such as clinical history and laboratory values, to strengthen diagnostic accuracy.

ACKNOWLEDGMENTS

We thank the Intelligent System Research Group (ISysRG) and Artificial Intelligence-Medical Center of Excellence (AIMed CoE), Universitas Sriwijaya, Indonesia, for its full support of our research infrastructure.

FUNDING INFORMATION

The research/publication of this article was funded by the 2025 Competitive Grant of Universitas Sriwijaya. In accordance with the Rector's Decree Number: 0028/UN9/LPPM.PT/2025, On September 17, 2025.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Annisa Darmawahyuni	✓	✓			✓				✓					✓
Siti Nurmaini					✓					✓				
Sutarno Sutarno					✓		✓	✓						
Anggun Islami					✓		✓	✓						
Muhammad Raffly			✓					✓			✓			
Riyansyah Putera														
Muhammad			✓					✓			✓			
Mukhlashyn Az														
Zahrawi														

C : **C**onceptualization

M : **M**ethodology

So : **S**oftware

Va : **V**alidation

Fo : **F**ormal analysis

I : **I**nvestigation

R : **R**esources

D : **D**ata Curation

O : Writing - **O**riginal Draft

E : Writing - Review & **E**diting

Vi : **V**isualization

Su : **S**upervision

P : **P**roject administration

Fu : **F**unding acquisition

CONFLICT OF INTEREST STATEMENT

No other potential conflict of interest relevant to this article was reported.

INFORMED CONSENT

There were no human subjects.

DATA AVAILABILITY

The datasets generated and/or analysed during the current study are available in the Roboflow (<https://roboflow.com/>).

REFERENCES

- [1] M. Bhagyanathan, D. Dhayanithy, V. A. Parambath, and R. Bijayraj, "Acanthosis nigricans: A screening test for insulin resistance--An important risk factor for diabetes mellitus type-2," *Journal of Business Research*, vol. 6, no. 1, pp. 43–46, 2017, doi: 10.4103/2249-4863.214961.
- [2] N. A. Álvarez-Villalobos *et al.*, "Acanthosis nigricans in middle-age adults: A highly prevalent and specific clinical sign of insulin resistance," *International Journal of Innovation Studies*, vol. 74, no. 3, p. e13453, 2020, doi: 10.1111/ijcp.13453.
- [3] W. Ogawa *et al.*, "New classification and diagnostic criteria for insulin resistance syndrome," *Endocrine Journal*, vol. 69, no. 2, pp. 107–113, 2022, doi: 10.1507/endocrj.EJ21-0725.
- [4] A. M. Angelidi, A. Filippaios, C. S. Mantzoros, and others, "Severe insulin resistance syndromes," *Journal of Clinical Investigation*, vol. 131, no. 4, 2021, doi: 10.1172/JCI142245.
- [5] A.-M. Radu, M. Carsote, M. C. Dumitrascu, and F. Sandru, "Acanthosis nigricans: pointer of endocrine entities," *Diagnostics*, vol. 12, no. 10, p. 2519, 2022, doi: 10.3390/diagnostics12102519.
- [6] A. Das *et al.*, "Acanthosis nigricans: a review," *Journal of Cosmetic Dermatology*, vol. 19, no. 8, pp. 1857–1865, 2020, doi: 10.1111/jocd.13544.
- [7] A. K. C. Leung, J. M. Lam, B. Barankin, K. F. Leong, and K. L. Hon, "Acanthosis nigricans: an updated review," *Current Pediatric Reviews*, vol. 19, no. 1, pp. 68–82, 2023, doi: 10.2174/1573396318666220429085231.
- [8] P. G. Unna, *The Histopathology of the Diseases of the Skin*. WF Clay, 1896.
- [9] S. Pollitzer, "Paul Gerson UNna, MD 1850-1929," *Archives of Dermatology and Syphilology*, vol. 19, no. 4, p. NP–678, 1929, doi: 10.1001/archderm.1929.02380220143009.
- [10] S. Zhang *et al.*, "Global hotspots and trends in Acanthosis nigricans research: A bibliometric and visualized analysis," *Skin Research and Technology*, vol. 30, no. 9, p. e70054, 2024, doi: 10.1111/srt.70054.
- [11] M.-L. Popa, A. C. Popa, C. Tanase, and A.-A. Gheorghisan-Galateanu, "Acanthosis nigricans: To be or not to be afraid," *Oncology Letters*, vol. 17, no. 5, pp. 4133–4138, 2019, doi: 10.3892/ol.2018.9736.
- [12] J. P. Burke, D. E. Hale, H. P. Hazuda, and M. P. Stern, "A quantitative scale of acanthosis nigricans," *Diabetes Care*, vol. 22, no. 10, pp. 1655–1659, 1999, doi: 10.2337/diacare.22.10.1655.
- [13] B. A. Rosner and others, *Fundamentals of biostatistics*, vol. 6. Thomson-Brooks/Cole Belmont, CA, 2006.
- [14] J. A. Hud, J. B. Cohen, J. M. Wagner, and P. D. Cruz, "Prevalence and significance of acanthosis nigricans in an adult obese population," *Archives of Dermatology*, vol. 128, no. 7, pp. 941–944, 1992, doi: 10.1001/archderm.1992.01680170073009.

- [15] A. S. Dhanoo, S. K. Ramroach, F. Hill-Briggs, and B. N. Cockburn, "Grading acanthosis nigricans using a smartphone and color analysis: a novel noninvasive method to screen for impaired glucose tolerance and Type 2 Diabetes," *Diabetes Spectrum*, vol. 37, no. 2, pp. 139–148, 2024, doi: 10.2337/ds23-0042.
- [16] C. Juárez-López, M. Klünder-Klünder, P. Medina-Bravo, A. Madrigal-Azcárate, E. Mass-D'Ávaz, and S. Flores-Huerta, "Insulin resistance and its association with the components of the metabolic syndrome among obese children and adolescents," *BMC Public Health*, vol. 10, no. 1, p. 318, 2010, doi: 10.1186/1471-2458-10-318.
- [17] L. Cveticanin and M. Arsenovic, "Prediction models for diabetes in children and adolescents: a review," *Applied Sciences*, vol. 15, no. 6, p. 2906, 2025, doi: 10.3390/app15062906.
- [18] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, 2018, doi: 10.1007/s13244-018-0639-9.
- [19] J. Du *et al.*, "Super-resolution reconstruction of single anisotropic 3D MR images using residual convolutional neural network," *Neurocomputing*, vol. 392, no. xxxx, pp. 209–220, 2020, doi: 10.1016/j.neucom.2018.10.102.
- [20] S. Li and X. Zhao, "Image-based concrete crack detection using convolutional neural network and exhaustive search technique," *Advances in Civil Engineering*, vol. 2019, 2019, doi: 10.1155/2019/6520620.
- [21] D. R. Nayak, N. Padhy, P. K. Mallick, D. K. Bagal, and S. Kumar, "Brain tumour classification using noble deep learning approach with parametric optimization through metaheuristics approaches," *Computers*, vol. 11, no. 1, 2022, doi: 10.3390/computers11010010.
- [22] R. A. Pratiwi, S. Nurmaini, D. P. Rini, M. N. Rachmatullah, and A. Darmawahyuni, "Deep ensemble learning for skin lesions classification with convolutional neural network," *International Journal of Artificial Intelligence IAES*, vol. 10, no. 3, p. 563, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp563-570.
- [23] J. A. Almaraz-Damian, V. Ponomaryov, S. Sadovnychiy, and H. Castillejos-Fernandez, "Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures," *Entropy*, 2020, doi: 10.3390/E22040484.
- [24] Z. Li, K. C. Koban, T. L. Schenck, R. E. Giunta, Q. Li, and Y. Sun, "Artificial intelligence in dermatology image analysis: current developments and future trends," *Journal of Clinical Medicine*, vol. 11, no. 22, p. 6826, 2022, doi: 10.3390/jcm11226826.
- [25] I. Lihacova *et al.*, "Multi-class CNN for classification of multispectral and autofluorescence skin lesion clinical images," *Journal of Clinical Medicine*, vol. 11, no. 10, p. 2833, 2022, doi: 10.3390/jcm11102833.
- [26] S. S. Noronha, M. A. Mehta, D. Garg, K. Kotecha, and A. Abraham, "Deep learning-based dermatological condition detection: a systematic review with recent methods, datasets, challenges, and future directions," *IEEE Access*, vol. 11, pp. 140348–140381, 2023, doi: 10.1109/ACCESS.2023.3339635.
- [27] A. Khan *et al.*, "A survey of the vision transformers and their CNN-transformer based variants," *Artificial Intelligence Review*, vol. 56, no. Suppl 3, pp. 2917–2970, 2023, doi: 10.1007/s10462-023-10595-0.
- [28] V. Hassija *et al.*, "Transformers for vision: a survey on innovative methods for computer vision," *IEEE Access*, 2025.
- [29] S. Jamil, M. Jalil Piran, and O.-J. Kwon, "A comprehensive survey of transformers for computer vision," *Drones*, vol. 7, no. 5, p. 287, 2023, doi: 10.3390/drones7050287.
- [30] K. Han *et al.*, "A survey on vision transformer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 87–110, 2022, doi: 10.1109/TPAMI.2022.3152247.
- [31] B. Dwyer, J. Nelson, and T. Hansen, "Roboflow (Version 1.0)," 2024.
- [32] Roboflow, "Acanthosis nigricans detection computer vision project," Accessed: Jun. 04, 2025. [Online]. Available: <https://universe.roboflow.com/sliit-gvlap/acanthosis-nigricans-detection-xemph>

BIOGRAPHIES OF AUTHORS



Annisa Darmawahyuni    is currently a lecturer and researcher in the Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya, Indonesia. Her research interest includes biomedical signal/image processing, deep learning, and machine learning. She can be contacted at email: riset.annisadarmawahyuni@gmail.com or annisadarmawahyuni@unsri.ac.id.



Siti Nurmaini    is currently a professor in the Faculty of Computer Science, Universitas Sriwijaya. Her research interests include biomedical engineering, deep learning, machine learning, image processing, control systems, and robotics. She can be contacted at email: sitinurmaini@gmail.com.



Sutarno Sutarno    is currently a lecturer in the Faculty of Computer Science, Universitas Sriwijaya. His research interests include control systems, robotics, deep learning, and machine learning. He can be contacted at email: sutarno@unsri.ac.id.



Anggun Islami    is currently a lecturer and researcher in the Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya, Indonesia. Her research interest includes biomedical image processing, deep learning, and machine learning. She can be contacted at email: anggunislami2@gmail.com.



Muhammad Raffly Riyansyah Putera    is an undergraduate student at the Faculty of Computer Science, Universitas Sriwijaya, Indonesia. His research interest includes machine learning, image processing, and deep learning. He can be contacted at email: rafflyriyansyah10@gmail.com.



Muhammad Mukhlashyn Az Zahrawi    is an undergraduate student at the Faculty of Computer Science, Universitas Sriwijaya, Indonesia. His research interest includes machine learning, image processing, and deep learning. He can be contacted at email: mukhlashyn.azzahrawi@gmail.com.