International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 10 | Oct - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Deep Learning-Based Psoriasis Severity Staging: A Comprehensive Survey

Akshada Jagtap, Shravani Jarhad, Sharavari Gade, Anurag More, Rageshri Bakare
Department of Electronics and Communication Engineering MIT School of Engineering and Science

Abstract — Psoriasis is a chronic, immune-mediated skin disorder affecting approximately 2 to 4 percent of the global population. Its clinical severity is quantified using the Psoriasis Area and Severity Index (PASI), a procedure that is labour-intensive, subjective, and prone to significant inter-observer disagreement. Over the past decade, deep learning techniques have emerged as promising tools for automating the analysis of dermoscopic and clinical skin photographs, offering objective and reproducible severity estimates. This paper presents a structured survey of twenty representative studies on psoriasis image analysis published between 2017 and 2023, covering convolutional transfer learning architectures, lesion segmentation frameworks, hybrid sequential models, and attention-based interpretability methods. Each study is examined for its dataset characteristics, architectural design, reported performance, and documented limitations. The survey identifies four persistent structural gaps — the dominance of binary detection over multi-tier severity staging, insufficient attention to the mild-to-moderate boundary, limited annotated dataset availability, and the absence of standardised evaluation benchmarks — and maps these to a proposed research direction targeting end-to-end three-class severity staging using an EfficientNetB3 backbone with multi-step preprocessing and lesion-aware augmentation.

Keywords — *psoriasis severity classification; EfficientNetB3; convolutional neural networks; transfer learning; PASI; CLAHE preprocessing; skin disease detection; medical image analysis.*

I. INTRODUCTION

Skin disorders collectively represent one of the most prevalent categories of chronic disease worldwide, placing a sustained burden on patients and healthcare systems alike. Among these, psoriasis occupies a position of particular clinical significance owing to its complex immunopathology, its wide phenotypic variability, and the profound impact it exerts on quality of life. The condition is driven by dysregulated T-cell-mediated immune responses that accelerate epidermal turnover, producing characteristic erythematous, hyperkeratotic plaques that may range from small localised patches to confluent involvement of large body-surface areas.

These presentations span a continuous clinical spectrum formally subdivided into mild, moderate, and severe categories using the Psoriasis Area and Severity Index (PASI), which aggregates measurements of redness, plaque thickness, surface scaling, and the proportion of affected body area into a composite numeric score. Therapeutic decision-making is directly stratified by severity tier: mild disease is typically managed with topical corticosteroids or vitamin D analogues, while moderate-to-severe presentations may require phototherapy, systemic immunomodulators, or high-cost biologic agents. Erroneous severity staging therefore carries direct clinical and economic consequences.

Despite its clinical centrality, manual PASI scoring is time-intensive and subjective. Reported inter-rater correlation coefficients as low as 0.70 demonstrate that experienced dermatologists may arrive at meaningfully different scores for the same patient, undermining consistency across sites and complicating trial endpoint adjudication. Convolutional neural networks offer a compelling alternative: they can learn hierarchical visual representations capturing erythema gradients, scaling texture, and plaque morphology directly from pixel data, closely paralleling the perceptual features exploited by clinicians. The landmark demonstration by Esteva et al. that a deeply trained network could match board-certified dermatologist accuracy in skin cancer classification catalysed a wave of research targeting psoriasis and other inflammatory dermatoses.

This paper presents a structured survey of twenty representative studies on deep learning-based psoriasis image analysis published between 2017 and 2023. The remainder of the paper is organised as follows: Section II provides a detailed literature review; Section III states the problem; Section IV outlines the proposed research direction; and Section V concludes.

TABLE I. PASI-Based Severity Classification

Stage	PASI Range	Key Visual Characteristics
Mild	0 – 10	Localised plaques, minimal scaling, low erythema
Moderate	10 – 20	Broader lesions, moderate erythema, partial scaling
Severe	> 20	Extensive confluent plaques, dense hyperkeratosis

II. LITERATURE REVIEW

A. Overview: The review covers convolutional transfer learning classifiers, segmentation-assisted pipelines, attention-based interpretability frameworks, hybrid spatial-sequential architectures, unsupervised clustering methods, and ensemble approaches published over a six-year window. Studies were selected to represent the principal methodological directions in psoriasis image analysis, with emphasis on those that explicitly address severity staging or contribute techniques directly applicable to it.

TABLE II. Common Augmentation Techniques in Reviewed Works

Technique	Typical Parameter	Clinical Motivation
Horiz./Vert. Flip	p = 0.5	Mirror plaque orientations
Random Rotation	+/- 25 deg	Camera angle variability
Zoom / Scale	0.8x – 1.2x	Lesion size variability
Brightness Adjust.	+/- 20%	Illumination differences
Gaussian Noise	sigma = 0.01	Sensor noise robustness
Colour Jitter	Hue +/- 0.1	Skin tone variation
CLAHE Enhancement	Clip = 2.0	Erythema contrast boost

B. Deep Learning Techniques for Psoriasis Severity Analysis: The following papers are reviewed in reference order, each examined for architectural design, dataset characteristics, reported performance, and documented limitations.

[1] Maulana et al. (2023) — *Comparative Deep Learning for Three-Class Severity Classification* Maulana and colleagues conducted one of the more disciplined comparative studies in the psoriasis severity classification literature, evaluating ResNet50, VGG19, MobileNetV3, and EfficientNetB0 under a controlled design that held preprocessing, augmentation, dataset split, and training budget constant across all models. By isolating architectural choice as the primary variable, the study generated directly comparable performance estimates — a methodological rigour notably absent in much of the broader literature. EfficientNetB0 and ResNet50 consistently outperformed the lighter MobileNetV3, which the authors attributed to their deeper representational capacity and more efficient gradient flow. EfficientNetB0 achieved the highest accuracy overall, confirming the compound-scaling principle for medical image classification. However, the study also surfaced an important practical constraint: achieving reported accuracy levels demanded substantial labelled data and GPU compute resources frequently unavailable in lower-resource clinical environments. Detailed per-class breakdowns for the clinically critical moderate tier were not reported, limiting assessment of where failures concentrate.

[2] Xing et al. (2022) — *Deep Learning for Automated PASI Score Regression* Rather than treating severity assessment as a discrete classification problem, Xing and colleagues framed it as continuous regression, directly predicting numerical PASI scores from multi-region clinical trial images. Their pipeline separately estimated erythema intensity, plaque thickness, scaling density, and affected body-surface area across anatomical regions using a segmentation-assisted architecture that first delineated lesion boundaries before extracting region-specific feature vectors for regression. Predicted composite scores demonstrated strong correlation with dermatologist-assigned values, providing compelling evidence that deep learning can serve as a viable computational surrogate for manual PASI computation. The regression formulation preserves the continuous nature of the severity spectrum rather than collapsing it into discrete bins, offering greater clinical granularity. Key challenges included performance degradation under variable illumination and the scarcity of large-scale, standardised multi-region psoriasis datasets with paired component annotations — a demand substantially higher than that of image-level classification approaches.

[3] Aijaz et al. (2022) — *CNN-LSTM Hybrid for Psoriasis Subtype Classification* Aijaz and colleagues addressed psoriasis subtype classification — covering plaque, guttate, inverse, pustular, and erythrodermic forms — by combining convolutional feature extraction with long short-term memory sequence modelling. The architecture treated spatial feature sequences from convolutional layers as temporal inputs to the LSTM, enabling the network to capture contextual relationships between spatially separated regions that global pooling discards. The hybrid design yielded measurable accuracy improvements over conventional machine learning baselines. However, the system showed notable sensitivity to image quality variations, and annotated data for rarer subtypes — particularly erythrodermic and pustular presentations — remained insufficient for robust learning. This scarcity problem is especially acute for rare clinical variants where even multi-institutional collaboration may fail to accumulate statistically adequate samples within a reasonable collection window.

[4] Goswami & Sharma (2022) — *YOLO-Powered Two-Stage Detection and Classification* Goswami and Sharma proposed a two-stage architecture that explicitly separated lesion localisation from severity classification, addressing a fundamental limitation of whole-image classifiers: susceptibility to confounding background content in clinical photographs where large areas of unaffected skin occupy significant input portions. A YOLO-based detector first localised psoriatic plaques and generated standardised bounding-box crops, which were then passed to a downstream severity classifier trained exclusively on lesion

content. Focusing the classifier's receptive field on plaque regions improved robustness in heterogeneous background environments and reduced the influence of irrelevant contextual variation. The design also enabled faster preliminary screening, practically important for telehealth applications where images may be captured by non-specialist users under uncontrolled conditions. Limitations included the additional bounding-box annotation overhead required for the detection stage and the compounding of errors: detector mislocalisation propagates as input noise to the classifier with no downstream correction mechanism.

[5] Bolia & Joshi (2023) — *Hybrid EfficientNet-LSTM for Dermoscopic Classification* Bolia and Joshi extended hybrid spatial-sequential modelling to a dermoscopic setting by combining EfficientNet's compound-scaled feature extraction with an LSTM component that modelled contextual relationships across extracted feature channels. While EfficientNet's spatial convolutions efficiently capture local texture, erythema gradients, and plaque morphology, they do not inherently model relational structure between feature channels — the gap the LSTM component was designed to fill. The hybrid architecture yielded accuracy improvements over standalone EfficientNet baselines on their dermoscopic dataset. The authors concluded that spatial-sequential hybridisation meaningfully benefits tasks where spatial context alone does not fully characterise severity. The principal limitation was dataset size: dermoscopic collections specific to psoriasis remain substantially smaller than those available for skin cancer tasks, restricting the hybrid architecture's ability to demonstrate its advantages over a well-regularised single-stage classifier.

[6] Lin et al. (2022) — *Transfer Learning for Plaque Boundary Segmentation* Lin and colleagues focused on the lesion segmentation sub-problem, applying transfer learning to initialise a segmentation network that delineated plaque boundaries in clinical photographs. Pretrained backbone weights substantially reduced training time and improved boundary delineation accuracy relative to random initialisation — particularly for moderate-severity cases where plaque margins grade continuously into perilesional skin. The improved segmentation quality was shown to support downstream severity monitoring by enabling reliable extraction of lesion area and morphological features that serve as key inputs to both automated classification and manual PASI estimation. The study made an important theoretical contribution by demonstrating transfer learning's effectiveness for dense prediction tasks, not only classification. Limitations included residual difficulty with low-contrast perilesional boundaries and the absence of downstream validation demonstrating that improved segmentation accuracy translated into measurable gains in an integrated severity staging pipeline.

[7] Semwal et al. (2022) — *PsLSNet: Modified U-Net for Psoriasis Segmentation* Semwal and colleagues introduced PsLSNet, a purpose-built segmentation architecture derived from U-Net with modified skip-connection topology designed specifically for psoriasis plaque delineation. The restructured encoder-decoder pathways improved pixel-level labelling accuracy and boundary precision across a range of plaque morphologies. The study made a compelling case that high-quality lesion segmentation is a foundational prerequisite for reliable severity quantification: the inherent boundary ambiguity of moderate presentations is exacerbated rather than resolved by imprecise segmentation, and segmentation errors compound downstream into classification inaccuracies. PsLSNet outperformed the baseline U-Net on Dice coefficient and boundary-focused metrics. However, improvements were most pronounced for clearly demarcated morphologies; for the diffuse, irregular presentations characteristic of moderate severity, performance gains were smaller, suggesting the moderate boundary problem may require solutions beyond architectural improvements to segmentation alone.

[8] Shrivastava et al. (2021) — *Classical Machine Learning with Handcrafted Features* Shrivastava and colleagues pursued a classical machine learning approach, constructing feature vectors from texture descriptors, colour histograms, and morphological measurements extracted from clinical images, then feeding these representations to a support vector machine classifier. The system achieved competitive accuracy on a constrained dataset, confirming that handcrafted feature engineering remains viable when labelled data is insufficient for effective deep model training — a practically relevant finding for smaller clinical institutions with limited collection resources. Limitations were equally clear: performance degraded substantially on lower-quality images where texture and colour features were distorted by illumination artefacts, and generalisation beyond the training distribution was constrained by the fixed feature representation's inability to adapt to visual patterns absent from the training set.

[9] Jung & Wu (2021) — *Unsupervised Lesion Identification via K-Means Clustering* Jung and Wu developed an unsupervised framework requiring no labelled training data, combining k-means pixel clustering with morphological image processing to segment psoriatic lesion regions. The clustering step grouped pixels by colour and texture similarity, separating erythematous plaque regions from surrounding skin based on spectral signatures, while morphological operators refined boundaries and removed spurious small regions. The approach demonstrated improved lesion extraction over simple thresholding and required no annotation effort during training. The fundamental limitation, however, was that the pipeline could not reliably discriminate severity levels from segmented regions alone: having identified plaque areas, the system lacked the supervised signal needed to associate visual features with severity categories. The study therefore functions as a detection tool rather than a staging tool, with

an additional supervised classification stage required for clinical severity quantification.

[10] Amruthalingam et al. (2021) — *Specialised CNN for Pustular Psoriasis Quantification* Amruthalingam and colleagues addressed pustular psoriasis — a clinically distinct subtype underserved by general-purpose severity metrics. Unlike plaque psoriasis, where PASI criteria apply naturally, pustular presentations are characterised by discrete fluid-filled lesions whose severity is better captured by pustule count, density, and spatial distribution. Their convolutional framework estimated these subtype-specific features, producing severity estimates in strong agreement with expert clinical assessments. By reducing the variability inherent in manual count-based assessment, the study demonstrated that lesion-type-specific architectures can outperform general-purpose classifiers on narrow subtypes with distinctive visual characteristics. The principal limitation was scope: the specialised architecture and feature extraction approach are not transferable to plaque-type psoriasis, which dominates clinical caseloads.

[11] Kamran et al. (2022) — *PSO-Net: Attention-Gated Interpretable Severity Assessment* Kamran and colleagues addressed the opacity of model decision-making — a recognised barrier to clinical adoption — through PSO-Net, an attention-gated architecture designed to simultaneously improve classification accuracy and produce spatial attention maps highlighting image regions most influential to each severity prediction. By making the model's reasoning spatially explicit, PSO-Net enables clinicians to verify whether the network attends to clinically meaningful plaque regions rather than incidental background features, providing a form of mechanistic transparency that standard classifiers cannot offer. Attention-weighted predictions showed reduced inter-prediction variability compared with standard classification outputs, suggesting that focusing gradient signals on diagnostically relevant regions improves consistency as well as interpretability. Limitations included computational overhead of the attention mechanism and the challenge of validating attention map quality against clinical gold standards in the absence of established ground-truth region annotations.

[12] Milani et al. (2023) — *Comparative Evaluation: ResNet50, InceptionV3, VGG19* Milani and colleagues conducted a controlled comparative evaluation of three standard architectures on a dermoscopic psoriasis dataset. InceptionV3 achieved the highest F1-score, a result attributed to its multi-scale convolution architecture: by processing features at multiple receptive field sizes simultaneously, Inception modules capture both fine-grained textural details relevant to scaling density and coarser morphological patterns relevant to plaque extent within a single forward pass. This multi-scale property is particularly advantageous for psoriasis severity tasks, where discriminative features operate simultaneously at the cellular texture and macroscopic lesion morphology scales. The study's primary contribution was identifying dataset scale as the principal performance bottleneck: all three architectures showed saturation under available data, and multi-institutional image collections were recommended for future benchmarks.

[13] Ahmmed et al. (2022) — *VGG16 Transfer Learning for Psoriasis and Actinic Keratosis* Ahmmed and colleagues applied VGG16 with ImageNet-pretrained weights to a joint classification task covering psoriasis and actinic keratosis, investigating the extent to which pretrained representations reduce effective training data requirements. Results confirmed that transfer learning significantly improved accuracy and convergence speed relative to random initialisation — consistent with the broader transfer learning literature. A persistent challenge was class imbalance between psoriasis and actinic keratosis samples, which degraded recall on the minority class as the loss gradient was dominated by the majority class. The authors recommended cost-sensitive loss functions or synthetic minority oversampling as remedies, noting that this imbalance problem is representative of the distributional challenges encountered in real-world clinical datasets where disease prevalence varies by orders of magnitude.

[14] Garakani et al. (2022) — *Similarity-Clustering for Standardised Severity Scoring* Garakani and colleagues proposed a similarity-clustering methodology that grouped patient images by visual feature similarity — derived from compact feature embeddings — and assigned standardised severity scores based on cluster membership. By reframing the problem as one of consistency rather than absolute accuracy, the approach measurably reduced both intra-observer and inter-observer PASI scoring variation without requiring end-to-end image classification, offering a lower-cost standardisation path for institutions lacking deep learning infrastructure. Limitations included the sensitivity of cluster quality to the choice of embedding and similarity metric, and the absence of a mechanism for handling out-of-distribution images that fall between established clusters and cannot be confidently assigned to any severity group without clinician review.

[15] Esteva et al. (2017) — *Dermatologist-Level Skin Cancer Classification* The work of Esteva and colleagues, published in Nature, occupies a foundational position in the deep learning dermatology literature. Training a deeply convolutional network on over 129,000 clinical skin disease images, they demonstrated classification of malignant melanoma and non-melanoma skin cancers at accuracy matching a panel of 21 board-certified dermatologists on a standardised test set. Beyond its specific technical findings, the study established deep learning as a legitimate clinical decision-support modality at specialist-level performance, and formalised a methodological template — large curated datasets, ImageNet pretraining, and cross-validation against clinical experts — that has since become the standard framework for disease-specific architecture development. Its

influence catalysed the wave of psoriasis-focused transfer learning research that all subsequent reviewed studies either build upon or respond to.

[16] Mahbod et al. (2019) — *Hybrid Multi-Architecture Feature Fusion* Mahbod and colleagues investigated hybrid configurations combining feature representations extracted from multiple architecturally distinct pretrained networks, motivated by the hypothesis that diverse inductive biases capture complementary aspects of lesion visual structure. Concatenating or averaging feature vectors from architecturally different backbones before downstream classification achieved improved robustness and reduced per-class error rates relative to any single constituent model. The work foreshadowed the ensemble methodology that became standard in medical imaging benchmarks where single-model performance improvements plateaued. The key insight — that architectural diversity in feature extraction translates into classification robustness via complementary error patterns — has direct implications for psoriasis staging, where moderate-tier ambiguity may benefit more from ensemble consistency than from incremental single-model gains.

[17] Brinker et al. (2019) — *Deep Learning versus 136 Dermatologists in Melanoma Classification* Brinker and colleagues extended the Esteva benchmark to a considerably larger clinical comparison cohort, evaluating a deep learning classifier against 136 board-certified dermatologists on standardised dermoscopic melanoma classification. The automated system matched or exceeded dermatologist-level accuracy across the full cohort, providing substantially stronger statistical evidence than previous smaller comparisons. The authors positioned their findings carefully: rather than advocating autonomous algorithmic diagnosis, they recommended deploying deep learning as a supplementary decision-support tool that augments clinician judgment, particularly for ambiguous cases where patient history, physical examination context, and clinical intuition remain indispensable. This framing — deep learning as decision support rather than autonomous replacement — has become the dominant paradigm for medical AI deployment and directly informs the clinical utility framing of psoriasis severity staging systems.

[18] Akrout et al. (2018) — *Automatic Skin Lesion Classification under Transfer Learning* Akrout and colleagues examined several pretrained backbone architectures for skin lesion classification under a transfer learning paradigm, systematically documenting accuracy gains over from-scratch training across multiple model families. The study confirmed consistent improvement from pretrained feature reuse across all evaluated architectures, reinforcing the near-universal applicability of transfer learning in medical imaging contexts where labelled dataset sizes are constrained. A practically significant finding was the identification of high GPU memory requirements as a barrier to deployment in resource-limited settings: larger, more accurate architectures that benefit most from transfer learning impose the greatest hardware demands, creating a tension between classification performance and point-of-care accessibility directly relevant to psoriasis staging system design.

[19] Harangi (2018) — *Ensemble Deep Learning for Medical Skin Classification* Harangi assembled an ensemble of architecturally diverse deep learning models for medical skin image classification, demonstrating that prediction averaging across constituent models substantially reduced classification error compared with any individual model. Robustness improvements were particularly pronounced for minority-class and boundary-case samples — precisely the categories that single models disproportionately misclassify — because diverse models tend to make uncorrelated errors on ambiguous inputs, allowing majority voting to recover correct predictions where individual models fail. For psoriasis staging, where moderate-severity boundary cases represent the primary error source, ensemble methods offer a principled mechanism for targeted error reduction without requiring additional labelled data. The practical cost is increased inference complexity and latency.

[20] Li et al. (2022) — *Systematic Review of Psoriasis Image Analysis* Li and colleagues synthesised the broader body of psoriasis image analysis research in a comprehensive systematic review spanning lesion segmentation, disease recognition, continuous severity scoring, and multi-class classification across a large number of published studies. The review provided the field with a structured map of methodological approaches, reported performance ranges, and dataset characteristics. The persistent obstacles identified — limited and poorly standardised labelled datasets, high computational demands, imaging heterogeneity across acquisition devices, and the absence of universally adopted evaluation protocols — are properties of the overall research ecosystem rather than limitations of any particular architecture. The review reinforced that no single methodological innovation has resolved these structural challenges, and emphasised that progress requires coordinated effort across dataset curation, evaluation standardisation, and efficient architecture design.

C. Synthesis and Gap Analysis: Across the twenty studies reviewed, convolutional transfer learning, lesion segmentation, hybrid spatial-sequential architectures, attention mechanisms, and ensemble averaging emerge as the dominant methodological building blocks. Four persistent structural gaps run through the collective literature and have not been fully resolved by any individual work.

The most consistently observed limitation is the framing of psoriasis image analysis as binary disease detection rather than multi-tier severity staging. Only a subset of the reviewed studies explicitly addressed three-class severity classification, and fewer still reported per-class performance

breakdowns adequate for assessing clinical utility at the moderate-severity boundary — the tier most critical to therapeutic decision-making and where classification errors concentrate. Moderate psoriasis, occupying the phenotypic overlap zone between mild and severe presentations, represents the hardest classification target and the one with the greatest clinical consequence when misclassified.

Dataset scale and standardisation represent the second persistent gap. Even the most recent studies operated on datasets of a few hundred to a few thousand images, with ground-truth annotations assigned under varying PASI interpretation criteria. The absence of multi-institutional benchmark datasets with standardised protocols makes it impossible to determine whether performance differences between architectures reflect genuine modelling advances or favourable dataset characteristics. A third gap concerns preprocessing: many studies adopted standard natural-image pipelines without adapting them to the specific challenges of clinical dermatological photographs — hair and specular artefacts, illumination variability, and the particular colour-texture signatures of erythema and scaling. Finally, explainability mechanisms adequate to support regulatory submissions and build clinical trust remain underexplored across most reviewed architectures.

TABLE III. Comparative Summary of Reviewed Literature

Ref	Author(s)	Architecture	Key Contribution	Main Limitation
[1]	Maulana et al.	ResNet50, VGG19, EfficientNetB0	Controlled 4-way arch. comparison	High compute; no per-class detail
[2]	Xing et al.	Seg.-assisted CNN	Continuous PASI regression	Illumination sensitivity
[3]	Aijaz et al.	CNN + LSTM	Subtype classification	Rare subtype data scarcity
[4]	Goswami & Sharma	YOLO + classifier	Two-stage detect-then-classify	Error compounding
[5]	Bolia & Joshi	EfficientNet + LSTM	Channel-context modelling	Small dataset
[6]	Lin et al.	Transfer learning + seg.	Plaque boundary delineation	No end-to-end validation
[7]	Semwal et al.	Modified U-Net	Skip-connection redesign	Marginal gains on diffuse plaques
[8]	Shrivastava et al.	SVM + handcrafted features	Classical pipeline on small data	Poor generalisation
[9]	Jung & Wu	K-means + morphology	Unsupervised lesion detection	Cannot stage severity alone
[10]	Amruthalingam et al.	Pustule detection	Subtype-specific	Not generalisable
		CNN	quantification	to plaque type
[11]	Kamran et al.	Attention-gated CNN	Interpretable spatial attention	Compute overhead
[12]	Milani et al.	ResNet50, InceptionV3, VGG19	Multi-scale InceptionV3 advantage	Dataset size ceiling
[13]	Ahmmed et al.	VGG16 transfer learning	Joint psoriasis-keratosis task	Class imbalance
[14]	Garakani et al.	Similarity clustering	Scoring standardisation	OOD images unhandled
[15]	Esteva et al.	Inception-v3 deep CNN	Dermatologist-level classification	Not psoriasis-specific
[16]	Mahbod et al.	Multi-arch. feature fusion	Hybrid feature extraction	Inference complexity
[17]	Brinker et al.	CNN vs. 136 dermatologists	Large-scale clinical comparison	Support tool framing
[18]	Akrout et al.	Multiple pretrained CNNs	Transfer learning survey	GPU memory barrier
[19]	Harangi	Ensemble deep models	Ensemble reduces boundary errors	Increased latency
[20]	Li et al.	Systematic review	Comprehensive field synthesis	Highlights persistent gaps

III. PROPOSED RESEARCH DIRECTION

Building on the gaps identified in the survey, the proposed research targets an end-to-end deep learning framework for three-class psoriasis severity staging. A multi-step preprocessing sequence — including resolution standardisation, channel-wise normalisation, Contrast-Limited Adaptive Histogram Equalisation in LAB colour space for erythema boundary sharpening, and morphological artefact attenuation — will be applied specifically to enhance the discriminative signals most relevant to severity classification, departing from the standard natural-image preprocessing adopted without modification in several reviewed studies.

A lesion-aware augmentation strategy will expand the effective training distribution by modelling realistic intra-patient variability, with targeted coverage of moderate-severity boundary presentations identified as the primary failure mode across reviewed systems. EfficientNetB3 is selected as the backbone based on its compound-scaling properties and the consistent superiority of EfficientNet variants in reviewed comparative studies. A two-phase fine-tuning protocol — freezing backbone weights during initial

head training, then releasing upper layers for joint optimisation — will balance general feature retention with psoriasis domain adaptation. The framework will be benchmarked against established architectures under identical experimental conditions, with explicit per-class metrics for the moderate tier reported to support clinical utility assessment.

IV. CONCLUSION

This paper has presented a structured survey of twenty deep learning studies on psoriasis image analysis published between 2017 and 2023. The reviewed works collectively demonstrate that transfer learning-based convolutional architectures can achieve clinically meaningful accuracy on severity staging and related tasks, with compound-scaled EfficientNet variants consistently outperforming earlier VGG and MobileNet designs under controlled comparisons.

Four persistent structural gaps were identified: the underrepresentation of three-class severity staging relative to binary detection tasks; insufficient attention to the moderate-severity boundary where classification errors concentrate; limited standardised multi-institutional datasets; and the absence of universally adopted evaluation benchmarks enabling reliable cross-study comparison. These gaps collectively motivate a dedicated severity staging pipeline with psoriasis-specific preprocessing, lesion-aware augmentation targeting boundary-case coverage, and rigorous per-class benchmarking. Future extensions include integration of upstream lesion segmentation for spatial attention, mobile deployment, and longitudinal severity tracking within electronic health record systems.

REFERENCES

[1] A. Maulana et al., "Deep learning-based psoriasis severity classification using ResNet50, VGG19, MobileNetV3, and EfficientNetB0," J. Comput. Sci. Inf. Technol., vol. 10, no. 2, pp. 45-60, 2023.

[2] Y. Xing et al., "Deep learning framework for automated PASI score prediction using clinical trial images," IEEE J. Biomed. Health Inform., vol. 26, no. 4, pp. 1523-1532, Apr. 2022.

[3] S. F. Aijaz et al., "Deep learning for psoriasis subtype classification using neural network and LSTM architectures," Comput. Med. Imaging Graph., vol. 98, p. 102066, Jun. 2022.

[4] A. Goswami and N. Sharma, "YOLO-powered deep learning for automated psoriasis classification," in Proc. IEEE BIBM, Las Vegas, NV, USA, 2022, pp. 1432-1438.

[5] C. Bolia and S. Joshi, "Hybrid EfficientNet-LSTM for psoriasis dermoscopic image classification," Expert Syst. Appl., vol. 215, p. 119409, 2023.

[6] G .-S. Lin et al., "Transfer learning for psoriasis image segmentation and analysis," Comput. Biol. Med., vol. 148, p. 105922, Sep. 2022.

[7] V. K. Semwal et al., "PsLSNet: Automated psoriasis lesion segmentation using modified U-Net," Biomed. Signal Process. Control, vol. 78, p. 103979, Aug. 2022.

[8] V. N. Shrivastava et al., "Machine learning-based psoriasis classification using image processing," J. Dermatol. Sci., vol. 89, pp. 112-120, 2021.

[9] J. Jung and C. Wu, "Psoriasis identification using k-means clustering and morphological processing," Comput. Biol. Med., vol. 132, p. 104297, 2021.

[10] L. Amruthalingam et al., "Deep learning for quantification of pustular psoriasis lesions," J. Invest. Dermatol., vol. 141, no. 7, pp. 1812-1820, Jul. 2021.

[11] S. A. Kamran et al., "PSO-Net: Attention-based interpretable deep learning for psoriasis severity," in Proc. IEEE EMBC, Glasgow, UK, 2022, pp. 3512-3518.

[12] A. Milani et al., "Comparative deep learning models for psoriasis classification using dermoscopic images," Skin Res. Technol., vol. 29, no. 3, p. e13259, Mar. 2023.

[13] F. Ahmmed et al., "Transfer learning with VGG16 for psoriasis and actinic keratosis classification," IEEE Access, vol. 10, pp. 54301-54310, 2022.

[14] A. Garakani et al., "Similarity-clustering for psoriasis severity assessment," J. Dermatol. Treat., vol. 33, no. 5, pp. 2240-2248, 2022.

[15] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115-118, Feb. 2017.

[16] A. Mahbod et al., "Skin lesion classification using hybrid deep neural networks," in Proc. ICIAR, 2019, pp. 301-309.

[17] T. J. Brinker et al., "Deep learning outperformed 136 dermatologists in dermoscopic melanoma classification," Eur. J. Cancer, vol. 113, pp. 47-54, May 2019.

[18] B. Akrout, M. Ferjani, and M. Atri, "Automatic skin lesion classification using deep learning," Procedia Comput. Sci., vol. 126, pp. 1917-1926, 2018.

[19] B. Harangi, "Skin lesion classification with ensembles of deep learning architectures," J. Biomed. Inform., vol. 86, pp. 25-32, 2018.

[20] Y. Li et al., "Psoriasis image analysis using machine learning: A systematic review," Comput. Biol. Med., vol. 140, p. 105083, 2022.