

DermDepth: Toward Monocular Metric Scale 3D Reconstruction Models for Dermatology

Héctor Carrión^{1*}, Narges Norouzi²

University of California, Santa Cruz¹, University of California, Berkeley²,
hcarri@ucsc.edu*

Abstract. Dermatological practice routinely involves measuring and tracking lesion size, morphology and texture, as critical components of wound or skin cancer screening, monitoring and diagnosis. To accomplish this task, practitioners often image the skin surface with commonly available off-the-shelf camera sensors. This has led to an overwhelming research focus on 2D methods while these objectives naturally benefit from 3D information. In this paper, we demonstrate that dense monocular 3D reconstructions, metric scale measurements and rich surface normal texture estimates are achievable for both dermoscopic and macroscopic cases without the need for additional hardware or multiple captures. We present **DermDepth**, the first single-view metric scale 3D model for the dermatological domain and **D-Synth**, the first synthetic dermoscopic dataset with pixel-perfect 3D information. Our experiments show training DermDepth on D-Synth corrects metric scale error from over $16\times$ to under $1.1\times$ for real dermoscopic data, while preserving geometric quality and increasing texture richness. Fine-tuning on a small amount of real clinical samples generalizes our method across three real-world benchmarks spanning the few mm to hundred cm range, diverse skin-tones, chronic wound cases and produces measurements broadly consistent with disease size reported in medical literature. All code, data and models are available at <https://github.com/hectorcarri@dermdepth>.

Keywords: Metric-Scale 3D Reconstruction · Synthetic Data · Fairness

1 Introduction

Skin lesions and chronic wound conditions impact the lives of millions of patients all over the world with an estimated 3 billion people lacking access to adequate dermatological care, especially in poor communities [10]. Non-melanoma skin cancer is the most common malignant tumor with its incidence increasing year over year and melanoma cancers are highly accretive, meaning morphology undergoes progressive growth [28]. Early detection significantly increases survival rates [6], emphasizing the scale and importance of active Asymmetry, Border, Color, Diameter and Evolution (ABCDE measurements) in dermatology [8]; features that Artificial Intelligence (AI) algorithms could facilitate tracking.

As such, dermatology is often a measurement problem: clinicians screen, triage, and monitor lesions or wounds by tracking changes in size, border irregularity, surface morphology, and texture over time [4]. These characteristics

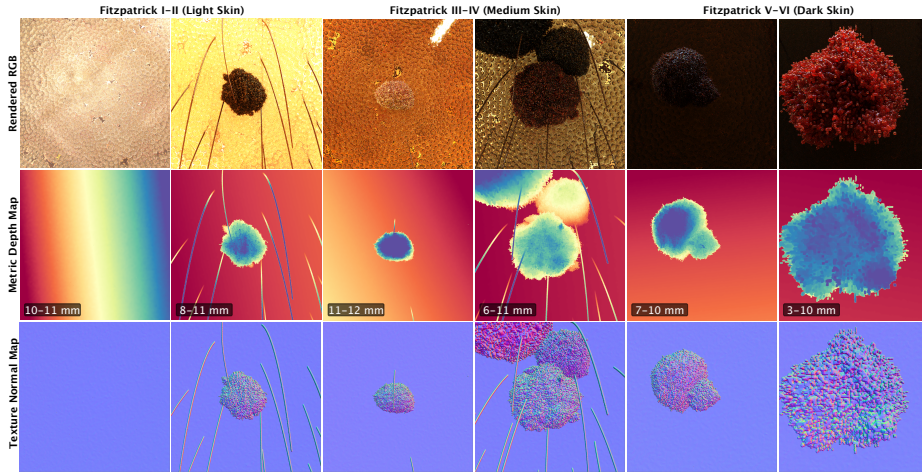


Fig. 1. D-Synth Dataset. Rendered skin-lesion synthetics from our D-Synth pipeline with dense pixel-perfect 3D information and texture normal maps.

are inherently 3D, yet the standard information capture at point of care remains 2D imaging, likely due to cost and inaccessibility of 3D systems. Critically, these measurements require scale consistency, ideally in metric space, thus 2D images are sometimes paired with rulers. However, clinically valuable information such as elevation, indentation or surface roughness are missed by simple width estimates.

The primary bottleneck toward training and evaluating 3D AI methods in dermatology is the difficulty of collecting suitable datasets [30]. Purpose-built devices have enabled limited 3D dermatological capture: SKINL2 [12] provides dermoscopic depth via plenoptic camera and WoundsDB [18] captures clinical wounds via time-of-flight, demonstrating significant downstream benefits: increased classification accuracy ($\uparrow 10\%$), sensitivity ($\uparrow 29\%$) [26], and $3\times$ more accurate area estimation [18], but require special hardware. Existing monocular approaches are limited to coarse elevation classification [2], relative-scale defocusing [25], or multi-view private-data methods [32]. Synthetic dermatological data [27,20] has been explored for 2D tasks but not for metric 3D depth.

Meanwhile, natural-image monocular depth estimation has advanced rapidly by leveraging large vision transformers [24]. Recent methods claim metric depth output: MoGe-2 [29] and MapAnything [19] predict scale via decoupled heads, Depth Anything v3 [21] uses canonical camera transformations, and Pixel-Perfect Depth [31] refines depth via diffusion generation. Whether these models generalize to dermatological images remains underexplored. In order to address these challenges, we present the following contributions:

1. **D-Synth**, a novel rendered synthetic dataset leveraging knowledge informed lesion growth models [20], contributing pixel-perfect 3D depth in metric scale, surface normals and camera intrinsics, as shown in Figure 1.

Table 1. DermDepth Datasets. SKINL2 is stratified by version, WoundsDB by patient, and DDI by Fitzpatrick skin tone.

Dataset	Modality	Total	Train	Test	3D GT source
D-Synth	Synth. dermoscopy	4,000	3,000	1,000	Rendered depth
SKINL2	Dermoscopy	375	263	112	Plenoptic camera
WoundsDB	Clinical wound	77	50	27	ToF sensor
DDI	Clinical + ruler	47	33	14	Ruler annotation

2. **DermDepth**, a model and fine-tuning approach for affine-invariant 3D prediction methods, correcting metric scale error across benchmarks by up to a factor of $14\times$, while improving surface normal detail and performing equitably across Fitzpatrick skin tones (disparity: $10.9 \rightarrow 1.02$).
3. **Systematic evaluation** of metric scale accuracy on dermatological images across 4 foundation models, 3 evaluation datasets, and 2 imaging modalities.
4. **Downstream validation** of metric accuracy for 3D lesion measurement (area, width, volume) from single photographs, validated against ruler ground truth and literature review.

2 Method

2.1 Training and Evaluation Datasets

We extend [20], which renders dermoscopic-style samples following biologically motivated lesion growth patterns. Originally, the method was designed to export 2D images and binary masks; we additionally contribute functionality to export per-pixel depth in metric scale, camera intrinsics and normal maps, while adding camera angles and multiple lesions. We synthesize 4,000 samples across 131 melanin levels and 240 lesion morphologies at 12–20 mm distance with 75° field of view. We call this new dataset D-Synth (Figure 1). We retain realism features including layered melanosome, blood and lipid models in epidermis, dermis and hypodermis, alongside physics-based light scatter across wavelengths and skin tones, however, we note a photorealism gap which real data can alleviate.

To bridge this sim-to-real gap and evaluate across clinical distances, we process SKINL2 and WoundsDB for fine-tuning and evaluation. SKINL2 plenoptic depth maps exhibit local planar noise [22] and WoundsDB time-of-flight maps are sparse, low-resolution with spatial offset from the RGB sensor. These issues may contribute noise to geometric evaluation but do not impact metric scale.

Diverse Dermatology Images (DDI) [11] is a fairness-focused 2D dataset akin to smartphone capture during first-pass clinical examination, sitting between SKINL2 and WoundsDB in terms of sampling distance. This dataset is collected without 3D information, however, we find 47 samples include a fully visible medical ruler with known dimensions [1] and matching segmentation masks [9]. This allows us to estimate the metric scale of the scene and evaluate fairness across skin-tones. Table 1 summarises all datasets and their train/test splits.

2.2 Model Architecture and Training Strategy

MoGe-2 and Pixel-Perfect Depth [29,31] predict affine-invariant depth maps which can then be scaled to metric space. This metric scaling factor is predicted via dedicated output heads by both MoGe-2 and MapAnything [19] with the former also able to predict normal maps via dedicated convolutional head. Given that real dermatology datasources may have accurate scale but noisy depth (and thus normals), we elect MoGe-2 as our base architecture due to its design flexibility.

When training on D-Synth we freeze the backbone encoder and fine-tune the scale and normal output heads. The scale head is composed of a 3-layer Multi-Layer Perceptron (MLP) that takes the CLS token from the base encoder as input and outputs a single scalar in log-space. The scale factor is then multiplied into the predicted affine-invariant point cloud to convert from relative to metric units, the loss is thus defined as:

$$\mathcal{L}_{\text{scale}} = (\log \hat{s} - \log s^*)^2 \quad (1)$$

Where $\hat{s}$ is the predicted scale factor and s^* is the ground-truth value. This formulation makes the loss symmetric with respect to over- or under-estimation, which is useful since monocular scale can be ambiguous. The normal output head is composed of a ConvStack [29] decoder similar to the 3D head, inputting feature maps and outputting a 3-channel per-pixel normal vector that is L2-normalized to unit length. The normal loss formula is:

$$\mathcal{L}_{\text{normal}} = \frac{1}{HW} \sum_i \angle(\hat{\mathbf{n}}_i, \mathbf{n}_i^*)^2 \quad (2)$$

Where H, W are image dimensions, $\hat{\mathbf{n}}_i$ and $\mathbf{n}_i^*$ are the predicted and ground-truth unit surface normals at pixel i , and $\angle(\cdot, \cdot)$ denotes the angle in radians between two vectors. When training on SKINL2, WoundsDB or DDI, given their noisy depth and normal maps, we choose to freeze the normal output head as well.

We split datasets 70% for training and 30% for testing stratified across versions for SKINL2, patient cases for WoundsDB (so that no patient appears in both splits) and skin-tones for DDI. We train for 1,000 steps with learning rate 10^{-4} , AdamW optimiser, and batch size 4 on a single A100 GPU. Exponential moving average (EMA, decay 0.999) is applied to all parameters; we find EMA essential for stable convergence, as raw checkpoints exhibit oscillation and over-correction. The same protocol is used for real-data training stages, with learning rate reduced to 5×10^{-5} and 2,000 steps per stage.

2.3 DDI Fairness and Lesion Measurements

We estimate lesion width, area and volume for held-out DDI samples with visible rulers for validation. Given predicted metric depth and estimated camera intrinsics, we back-project to a 3D point cloud and compute: surface area via cross-product triangle integration, lesion width as the major PCA axis, and volume as the integral between the lesion surface and a fitted boundary plane.

Table 2. Scale Results. Foundation models systematically miss scale on dermatological data despite strong geometry ($\text{SI-}\delta_1$). DermDepth fine-tunes only 2.1 M parameters (0.6%), with significant scale correction on all benchmarks. Scale ratio target is $1.0\times$; DDI ratio target is $1.0\times$. DermDepth_S: synthetic data only. Eval on held-out test sets.

Method	Params	SKINL2 (dermoscopy)		WoundsDB (clinical)		DDI
		Scale r	$\text{SI-}\delta_1$ (%)	Scale r	$\text{SI-}\delta_1$ (%)	Ratio ρ
DA3 [21]	1.4 B	4.16	99.6	0.67	89.1	53.6
MapAnything [19]	1.1 B	10.88	99.3	<u>0.75</u>	80.7	156.1
PPD [31]	831 M	16.21	92.0	0.66	<u>91.3</u>	90.4
MoGe-2 [29]	331 M	16.10	<u>100.0</u>	0.62	91.1	81.0
DermDepth _S	2.1 M ft	1.11	<u>100.0</u>	0.28	91.1	<u>9.2</u>
DermDepth	2.1 M ft	<u>0.87</u>	<u>100.0</u>	0.91	92.6	1.95

3 Results

3.1 Baseline Methods and Metrics

We benchmark four recent models that claim metric depth output: **Depth Anything 3** [21] (DA3-METRIC-LARGE; 1.4 B params), **MapAnything** [19] (1.1 B), **Pixel-Perfect Depth** [31] (PPD; 831 M, diffusion refinement with MoGe-2 scaling as done on the official codebase), and **MoGe-2** [29] (331 M).

We report three complementary metrics: **Scale ratio** $r = \text{median}(\hat{d}_i/d_i^*)$, where $r=1.0$ is perfect; **SI- δ_1** = percentage of pixels where the scale-aligned depth ratio falls within $[1/1.25, 1.25]$, measuring geometry quality independent of global scale; and **DDI ratio** $\rho = A_{\text{pred}}/A_{\text{GT}}$, measuring predicted ruler area relative to the known ruler area of 6.6 cm^2 .

3.2 Experimental Results

Table 2 demonstrates that every evaluated foundation model produces systematically incorrect metric scale on dermatological images. All four baselines overestimate dermoscopic depth by $4\text{--}16\times$ (SKINL2) and underestimate clinical wound depth at $0.6\text{--}0.75\times$ (WoundsDB), with DDI ruler-area ratios ranging from $54\times$ to $156\times$. Crucially, geometry quality is high across all models ($\text{SI-}\delta_1 \geq 92\%$), confirming that the 3D structure is captured correctly, supporting our decision to freeze the majority of parameters, focusing on scale and normal map training.

PPD internally uses MoGe-2 for metric grounding and consequently inherits near-identical scale bias ($16.2\times \approx 16.1\times$ on SKINL2), while its diffusion refinement slightly degrades geometry ($\text{SI-}\delta_1$: 92.0% vs. 100.0%). DA3, despite having the most parameters, overestimates by $4.2\times$.

DermDepth corrects scale across all three benchmarks. With synthetic data alone, SKINL2 scale drops from $16.1\times$ to $1.11\times$ while high-quality geometry is preserved ($\text{SI-}\delta_1=100\%$ and $\text{SI-AbsRel}=0.017$). With progressive real-data training, DermDepth achieves $0.87\times$ on SKINL2, $0.91\times$ on WoundsDB, and $1.95\times$ on

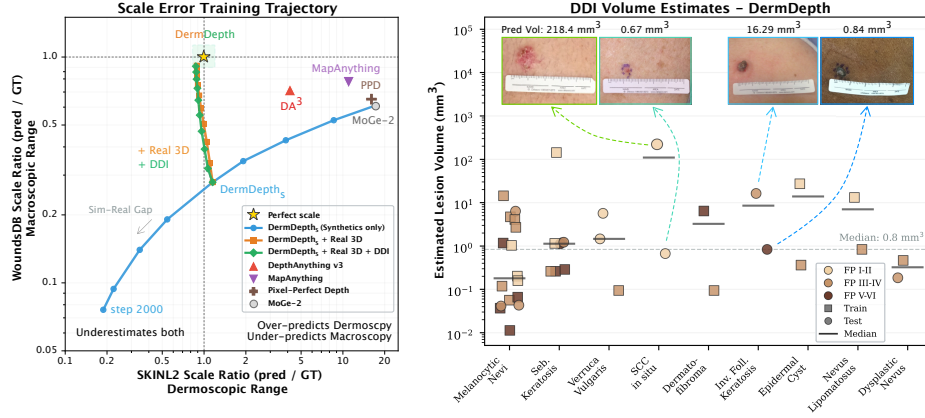


Fig. 2. Scale Correction Training (Left), Predicted Volume (Right). We measure dermoscopic range (cm-level) over-estimation and macroscopic range (m-level) underestimation for all foundation models tested. Training on D-Synth corrects dermoscopic error, adding real data corrects macroscopic error (left). We demonstrate DermDepth volume measurements on DDI held-out test data (right).

DDI, bringing all three benchmarks close to metric scale, as shown in Figure 2 (left).

3.3 Training Strategy Ablations

In the small-data-large-model setting it is common to freeze most model weights and train only output heads; we ablate this choice. Under synthetic-only training, unfreezing more parameters (2.1–326 M) progressively harms scale and geometry on real data: 3,000 synthetic samples cannot re-learn the full decoder, and the sim-real gap causes overfitting (Table 3, rows 1–4). D-Synth covers only dermoscopic distances (~ 12 –20 mm); adding real SKINL2 ($\sim$ few-hundred mm) and WoundsDB (~ 1 –1.5 m) depth bridges the macroscopic gap but worsens DDI, which sits between the two in sampling distance. Including 33 DDI pseudo-GT samples anchored to ruler measurements generalizes scale prediction, without compromising other benchmarks (Table 3, rows 2, 5–6), justifying our output-head-only and progressive real data training design.

3.4 Skin Tone Fairness

Equitable performance across skin tones is essential for clinical deployment of dermatology methods [11]. Table 4 evaluates DDI ruler-area ratio stratified by Fitzpatrick group. MapAnything shows the strongest bias ($211.88\times$ on FP V–VI vs. $88.53\times$ on FP III–IV), while the most uniform baseline, MoGe-2, has a spread of 10.90. Diverse skin tones in our synthetic training data could be valuable to reduce bias as DermDepth_s lowers disparity to 1.70, while training on real data leads to 1.02, with per-group ratios between $1.39\times$ and $2.41\times$.

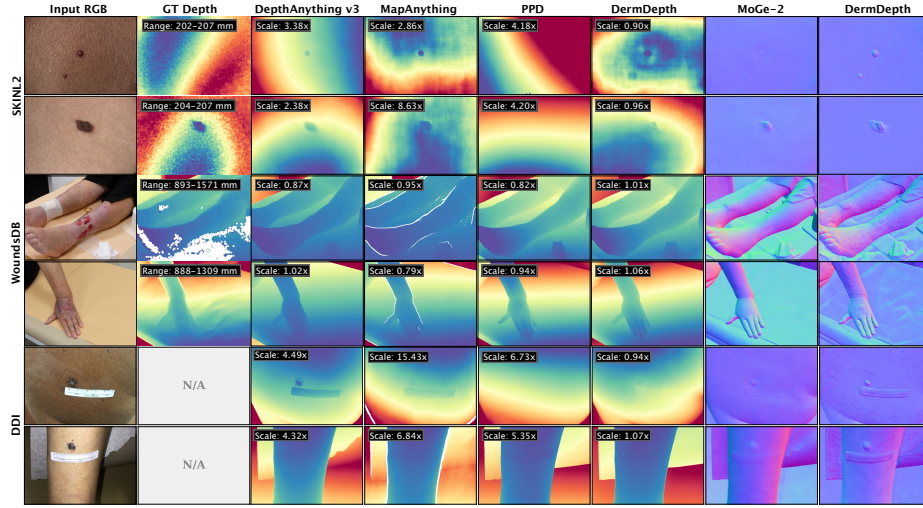


Fig. 3. Predicted Depth, Scale and Normals. We visualize performance across three datasets at different metric distances. We find DermDepth produces more accurate and consistent scale and normal maps with finer lesion detail. We note DDI was not collected alongside 3D information, thus we estimate scale error from known ruler area. Ground-truth may show dataset-specific sensor noise, as discussed in Sect. 2.1.

3.5 Normal Map Qualitative Results

Surface normal maps encode local 3D orientation at each pixel, capturing microgeometric features invisible in RGB or depth alone. As shown in Figure 3, DermDepth produces normal maps with noticeably finer lesion detail compared to the baseline predictions. The improved normals better delineate lesion boundaries, surface roughness and elevation changes, features directly relevant to dermatological assessment. In particular, the Asymmetry and Border components of ABCDE screening [8] and texture features such as ulceration, scaling and papillomatous morphology manifest as distinctive normal map patterns. Quantitative normal descriptors derived from these maps could supplement existing 2D analysis by capturing surface morphology currently assessable only through physical examination or specialized hardware.

3.6 Clinical Use Case: 3D Lesion Measurement

We validate clinical utility by measuring lesion geometry in DDI images with known-sized rulers. We highlight test-set cases of Squamous Cell Carcinoma and Inverted Follicular Keratosis to visually validate reasonable volumetric differences (Figure 2, right). To verify if DermDepth width estimates are clinically plausible, we collect dermatological literature that has recorded observations of test-set DDI lesions. Across 8 diagnostic categories (14 lesions), 10 of 14 fall within published typical ranges, while the remaining 4 could be clinically viable

Table 3. Training ablations, trainable parameters and real-data value.

Configuration	Scale ratio r ($\rightarrow 1.0$)			SI- δ_1 (%)	
	SKINL2	WoundsDB	DDI	SKINL2	WoundsDB
Baseline (MoGe-2)	16.10	0.62	81.0	100.0	91.1
Scale head, synth	1.11	0.28	9.2	100.0	91.1
+ Decoder, synth	0.14	0.06	0.76	99.9	85.8
+ Full model, synth	0.09	0.04	0.13	99.8	77.5
Scale head, + real depth	0.89	0.92	18.3	100.0	91.1
Scale head, + DDI pseudo	<u>0.87</u>	<u>0.91</u>	1.95	100.0	92.6

Table 4. DDI ruler-area ratio by Fitzpatrick skin tone. DermDepth achieves equitable performance across skin tones, reducing the disparity (max – min per-group ratio). DermDepth_S: synthetic data only; DermDepth: synth + SKINL2, DDI, WDB.

Method	Overall	FPI I–II	FP III–IV	FP V–VI	Disparity
DA3 [21]	53.60×	73.00×	58.60×	38.10×	34.90
MapAnything [19]	156.11×	156.11×	88.53×	211.88×	123.35
PPD [31]	90.36×	100.44×	100.10×	80.26×	20.18
MoGe-2 [29]	81.00×	88.50×	84.50×	77.60×	10.90
DermDepth _S	9.20×	9.20×	10.00×	8.30×	1.70
DermDepth	1.95×	1.92×	2.41×	1.39×	1.02

boundary cases: an early-stage keratoacanthoma (7.1 mm vs. classic mature 10–20 mm), an early-detection Bowen’s lesion (4.5 mm vs. median 10 mm), a dysplastic nevus at the ≥ 5 mm diagnostic threshold (5.1 mm), and a large melanocytic nevus in the top 2% by size (9.0 mm). These ranges along with citations are shown in Table 5.

4 Discussion and Conclusion

Limitations. D-Synth currently renders dermoscopic-distance scenes (12–20 mm), so macroscopic clinical distances must be covered by real training data, a gap we bridge with WoundsDB and SKINL2, but which limits the synthetic component to close-range correction. The real-data ground-truth itself carries noise. These issues affect geometric evaluation (surface normals in particular) but do not compromise metric scale. Finally, DDI pseudo-GT is derived from ruler segmentation and known ruler dimensions; segmentation errors or partial ruler occlusion introduce uncertainty, and the DDI test set ($n=14$) remains small, limiting statistical significance for DDI results, an issue common under limited fairness data. We strongly believe the community should collect more data with clean 3D information.

Table 5. Literature validation of DermDepth lesion width predictions. Predicted widths for DDI test-set lesions compared to published clinical size ranges.

Lesion type	DermDepth (mm)	Published range (mm)	Sources
Seborrheic keratosis	5.9–10.0	1–30; typical 5–10	[15,7]
Melanocytic nevi	1.6–9.0	1–10; 98% ≤ 5	[17,23]
Inv. follicular keratosis	9.2–10.2	3–30; mean 10.1	[16]
Verruca vulgaris	7.2–7.4	2–10	[3]
SCC, keratoacanthoma	7.1	10–20 (mature)	[33]
Bowen’s disease	4.5	1–55; median 10	[14]
Melanoma	9.9	IQR 9–20	[13]
Dysplastic nevus	5.1	≥ 5 ; typical 5–15	[5]

Conclusion. We have shown that foundation metric depth models fail on dermatological images, overestimating dermoscopic depth by $4\text{--}16\times$ while underestimating clinical wound depth. DermDepth corrects this systematic error by fine-tuning a scale prediction head and normal output head (0.6% of parameters) on D-Synth, the first synthetic dermoscopic dataset with pixel-perfect 3D information, followed by progressive real-data refinement with just 346 clinical samples. The result is much improved metric scale across three diverse benchmarks with equitable performance across Fitzpatrick skin tones (disparity: $10.90 \rightarrow 1.02$). DermDepth is memory efficient, and could run on mobile devices. We publish all code, models, and datasets to encourage further research in this direction.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Delasco photo measuring labels. <https://www.delasco.com/products/photo-measuring-labels>, accessed: 2026-02-23
2. Abhishek, K., Hamarneh, G.: Lesion elevation prediction from skin images improves diagnosis. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) ISIC Skin Image Analysis Workshop. vol. 15274, pp. 45–55 (2025)
3. Aboud, A.M.A., Nigam, P.K.: Wart (2023), <https://www.ncbi.nlm.nih.gov/books/NBK431047/>
4. Argenziano, G., Fabbrocini, G., Carli, P., et al.: Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Archives of Dermatology **134**(12), 1563–70 (1998)
5. Baigrie, D., Tanner, L.S.: Dysplastic nevi (2022), <https://www.ncbi.nlm.nih.gov/books/NBK482210/>
6. Balch, C.M., Gershenwald, J.E., Soong, S.J., et al.: Final version of 2009 AJCC Melanoma Staging and Classification. Journal of Clinical Oncology **27**(36), 6199–6206 (2009)
7. Barthelmann, S., Butsch, F., Lang, B.M., et al.: Seborrheic keratosis. JDDG Journal der Deutschen Dermatologischen Gesellschaft **21**(3), 265–277 (2023)

8. Blum, A., Rassner, G., Garbe, C.: Modified ABC-point list of dermoscopy: A simplified and highly accurate dermoscopic algorithm for the diagnosis of cutaneous melanocytic lesions. *Journal of the American Academy of Dermatology* **48**(5), 672–678 (2003)
9. Carrión, H., Norouzi, N.: FEDD – Fair, Efficient, and diverse diffusion-based lesion segmentation and malignancy Classification. *arXiv (Cornell University)* (2023), <https://arxiv.org/abs/2307.11654>
10. Coustasse, A., Sarkar, R., Abodunde, B., et al.: Use of teledermatology to improve dermatological access in rural areas. *Telemedicine and e-Health* **25**, 1022–1032 (2019)
11. Daneshjou, R., Vodrahalli, K., Novoa, R.A., et al.: Disparities in dermatology ai performance on a diverse, curated clinical image set. *Science Advances* **8** (2022)
12. De Faria, S.M.M., Henrique, M., Filipe, J.N., et al.: Light Field Image Dataset of Skin Lesions. *IEEE* **2019**, 3905–3908 (2019)
13. Dessinioti, C., Plaka, M., Befon, A., et al.: A retrospective study of diameter and breslow thickness in invasive melanomas. *Dermatology* **240**(3), 462–467 (2024)
14. Fougelberg, J., Ek, H., Claeson, M., et al.: Surgery for Bowen Disease: Clinicopathological factors associated with incomplete excision. *Dermatology Practical & Conceptual* **11**(2), e2021046 (2021)
15. Greco, M.J., Bhutta, B.S.: Seborrheic keratosis (2024), <https://www.ncbi.nlm.nih.gov/books/NBK545285/>
16. Hamzelou, S., Aryanian, Z., Maqbool, F., et al.: Assessment of clinical and pathological characteristics of inverted follicular keratosis - five years comprehensive review. *Archives of Dermatological Research* **317**(1), 485 (2025)
17. İyidal, A.Y., Gül, U., Kılıç, A.: Number and size of acquired melanocytic nevi and affecting risk factors in cases admitted to the dermatology clinic. *Advances in Dermatology and Allergology* **5**(5), 375–380 (2016)
18. Juszczak, J.M., Wijata, A., Czajkowska, J., et al.: Wound 3D geometrical feature estimation using poisson reconstruction. *IEEE Access* **9**, 7894–7907 (2020)
19. Keetha, N., Müller, N., Schönberger, J., et al.: MapAnything: Universal feed-forward metric 3D reconstruction. In: *International Conference on 3D Vision (3DV)*. *IEEE* (2026)
20. Kim, A., Saharkhiz, N., Sizikova, E., Lago, M., Sahiner, B., Delfino, J., Badano, A.: S-SYNTH: Knowledge-Based, Synthetic generation of skin images. *arXiv (Cornell University)* (2024), <https://arxiv.org/abs/2408.00191>
21. Lin, H., Chen, S., Liew, J.H., Chen, D.Y., Li, Z., Shi, G., Feng, J., Kang, B.: Depth anything 3: Recovering the visual space from any views. *arXiv preprint arXiv:2511.10647* (2025), <https://arxiv.org/abs/2511.10647>
22. Lourenco, R.M., Tavora, L.M.N., Assuncao, P.A.A., et al.: Enhancement of light field disparity maps by reducing the silhouette effect and plane noise. *Multidimensional Systems and Signal Processing* **33**(2), 1–33 (2022)
23. Muradia, I., Khunger, N., Yadav, A.K.: A clinical, dermoscopic, and histopathological analysis of common acquired melanocytic nevi in skin of color (2022), <https://pmc.ncbi.nlm.nih.gov/articles/PMC9586527/>
24. Oquab, M., Darcet, T., Moutakanni, T., et al.: DINOv2: Learning robust visual features without supervision. *Transactions on Machine Learning Research* (2024)
25. Parekh, P., Oyeleke, R., Vishwanath, T.: The Depth Estimation and Visualization of Dermatological Lesions: Development and Usability Study. *JMIR Dermatology* **7**, e59839 (2024)

26. Pereira, P.M.M., Thomaz, L.A., Tavora, L.M.N., et al.: Melanoma classification using light-Fields with morlet scattering transform and CNN: Surface depth as a valuable tool to increase detection rate. *Medical Image Analysis* **75**, 102254 (2021)
27. Sinha, A., Kawahara, J., Pakzad, A., et al.: Dersynth3d: Synthesis of in-the-wild annotated dermatology images. *Medical Image Analysis* p. 103145 (2024)
28. Wang, M., Gao, X., Zhang, L.: Recent global patterns in skin cancer incidence, mortality, and prevalence. *Chinese Medical Journal* **138**(2), 185–192 (2024)
29. Wang, R., Xu, S., Dong, Y., Deng, Y., Xiang, J., Lv, Z., Sun, G., Tong, X., Yang, J.: Moge-2: Accurate monocular geometry with metric scale and sharp details (2025), <https://arxiv.org/abs/2507.02546>
30. Xie, B.: A Comparative Study of Medical Image Segmentation via ResNet-Driven UNet Architecture. *IEEE* pp. 96–101 (2025)
31. Xu, G., Lin, H., Luo, H., Wang, X., Yao, J., Zhu, L., Pu, Y., Chi, C., Sun, H., Wang, B., et al.: Pixel-perfect depth with semantics-prompted diffusion transformers. *arXiv preprint arXiv:2510.07316* (2025), <https://arxiv.org/abs/2510.07316>
32. Zhang, C., Tan, Z., Yang, X., et al.: From 2D images to 3D model: Segmenting and reconstructing acne for robust skin lesion analysis. *Cognitive Computation* **18**(1) (2025)
33. Zito, P.M., Scharf, R.: Keratoacanthoma (2023), <https://www.ncbi.nlm.nih.gov/books/NBK499931/>