Topology-Aware Deep Models for Skin Lesion Classification

Sayoni Chakraborty*, Philmore Koung*, and Baris Coskunuzer[0000−0001−7462−8819]

Dept. Math. Sciences, University of Texas at Dallas, Richardson, TX 75080 USA
{sayoni.chakraborty, philmore.koung, coskunuz}@utdallas.edu

Abstract. Skin cancer is among the most prevalent and potentially deadly cancers worldwide, with early detection essential for effective treatment, particularly for aggressive types such as melanoma. Deep learning (DL) models have shown strong performance in skin lesion classification tasks, yet they often struggle to capture the complex geometric and topological structures present in dermoscopic images. In this study, we propose a hybrid classification framework that combines topological descriptors with CNNs and Vision Transformers to improve diagnostic performance across multiple categories of skin lesions. By extracting topological signatures, we quantify shape and connectivity patterns that are often overlooked by standard convolutional neural networks. Our experiments in multiple publicly available dermatology datasets demonstrate that topological models perform competitively on their own, and their integration with DL models consistently improves classification metrics. These results establish topological features as a valuable complement to deep learning in the diagnosis of skin cancer.

Keywords: Topological data analysis · Cubical Persistence · Skin cancer detection · Dermoscopic image analysis · CNNs · Vision Transformers

1 Introduction

Skin cancer is one of the most common and potentially lethal malignancies worldwide, making early and accurate detection crucial for reducing morbidity and mortality. Clinical decision support systems (CDSS) powered by machine learning (ML) have shown great promise in automating the analysis of dermoscopic images, with deep learning (DL) models achieving dermatologist-level performance in various lesion classification tasks [9,36]. However, these conventional DL approaches are based primarily on pixel-level features and often do not fully capture the complex shape and boundary patterns that distinguish malignant from benign lesions [16].

Topological data analysis (TDA) offers a complementary perspective by extracting robust scale-invariant descriptors of the geometric and connectivity structure of an image [5]. Persistent homology, in particular, encodes how features such as connected components and holes evolve across intensity thresholds, producing high-dimensional topological signatures that are insensitive to changes in noise and illumination. Recent work has demonstrated the utility of TDA in medical imaging contexts such as tumor margin delineation and histopathology analysis [17].

* Equal contribution.

In this paper, we bring together the strengths of TDA and DL in a unified CDSS for skin cancer detection. We extract persistence-based features from dermoscopic images of multiple types of lesion, including melanoma, basal cell carcinoma, squamous cell carcinoma, and benign nevi, and evaluate both standalone topological classifiers and their integration with state-of-the-art convolutional and transformer-based architectures. By encoding the persistence diagrams as sequential inputs to a transformer, we leverage its capacity to model long-range dependencies among topological events.

Our contributions can be summarized as follows:

- We present the first comprehensive evaluation of topological data analysis for automated skin lesion classification across multiple cancerous and benign categories.
- We propose a novel fusion strategy that encodes persistence outputs as sequences and integrates them with transformers to capture global topological interactions.
- We show that topological models rival conventional DL methods and that combining them with CNNs and ViTs consistently improves classification performance.
- Our results highlight the potential of topological features to enhance the accuracy and robustness of deep learning–based CDSS for skin cancer detection.

2 Related Work

Machine Learning Methods in Skin Cancer Detection. Machine learning and deep learning have revolutionized automated skin lesion analysis. Esteva et al. trained a CNN on over 100 000 clinical and dermoscopic images to match expert performance across more than a dozen lesion categories [14]. Subsequent work fine-tuned ImageNet-pretrained ResNets for marked accuracy gains [28], confirmed cross-cohort robustness [3], and leveraged large benchmarks like HAM10000 alongside ensemble methods to boost generalization [37,9]. Despite these advances, challenges persist in data heterogeneity, interpretability, and clinical integration [29]. Recent strategies include fusing dermoscopy with patient metadata for better risk stratification [21] and adopting transformer-based models to capture global context [39]. Yet most approaches remain pixel-centric, and the use of topological data analysis to encode shape and connectivity is still largely unexplored in skin cancer diagnostics.

Topological Machine Learning in Medical Image Analysis. Persistent homology (PH) offers robust, shape-based descriptors that are resilient to noise and photometric variation, making it a powerful tool in medical imaging. PH-based features have been applied in diverse biomedical contexts, including modeling cell development [25], delineating tumor margins [30], analyzing brain connectivity [31], and extracting multiscale genomic signatures [23]. For a broader overview, see Skaf et al. [34]. Recently, *topological deep learning* has emerged as a paradigm for integrating persistence features into neural networks. Architectures by [17,1] embed persistence summaries to improve segmentation [20,33] and classification [6,19]. We extend this framework to skin cancer detection, extracting PH from dermoscopic images and evaluating both standalone and hybrid models combining PH with CNNs and ViTs. Although prior studies have applied TDA to melanoma detection [24,8], this work presents the first comprehensive evaluation of topological descriptors for skin lesion classification, demonstrating both competitive standalone performance and consistent enhancement of deep learning–based diagnostics.

3 Methodology

Our methodology involves two key steps. First, we extract topological feature vectors from dermascopic images. Then, we evaluate their standalone performance in ML models and integrate them with the latest DL models to assess their impact on improving SOTA performance.

3.1 Topological Feature Vectors for Dermascopic Images

PH is a powerful mathematical tool in TDA for analyzing complex data structures. It identifies hidden patterns at multiple resolutions and effectively extracts features from various data formats, including point clouds and networks [10]. This paper focuses on its application in image analysis, specifically *cubical persistence*, a variant of PH. While we provide an accessible overview, deeper insights can be found in [12]. PH follows a three-step procedure:

- **Filtration**: Inducing a sequence of nested topological spaces from the data.
- **Persistence Diagrams**: Recording the topological changes within this sequence.
- **Vectorization**: Converting these diagrams into vectors to be utilized in ML models.

Step 1 - Constructing Filtrations. Since PH essentially functions as a mechanism for monitoring the progression of topological characteristics within a sequence of simplicial complexes, constructing this sequence stands out as a crucial step. In image analysis, the common approach is to generate a nested sequence of binary images, also known as *cubical complexes*. To achieve this from a given color (or grayscale) image $\mathcal{X}$ (with dimensions $r \times s$), one needs to select a specific color channel (e.g., red, blue, green, or grayscale). The color values γ_{ij} of individual pixels $\Delta_{ij} \subset \mathcal{X}$ are then utilized. Specifically, for a sequence of color values $(0 = t_1 < t_2 < \cdots < t_N = 255)$, a nested sequence of binary images $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \cdots \subset \mathcal{X}_N$ is obtained, where $\mathcal{X}_n = \{\Delta_{ij} \subset \mathcal{X} \mid \gamma_{ij} \leq t_n\}$ (See Figure 1). In particular, this involves starting with a blank $r \times s$ image and progressively activating (coloring black) pixels as their grayscale values reach the specified threshold t_n. This process, known as *sublevel filtration*, is conducted on $\mathcal{X}$ relative to a designated function (in this instance, grayscale).

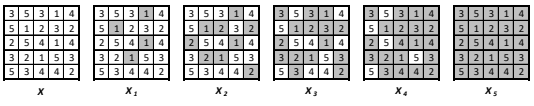

Fig. 1: For the 5×5 image $\mathcal{X}$ with the given pixel values, **the sublevel filtration** is the sequence of binary images $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \cdots \subset \mathcal{X}_5$.

Step 2 - Persistence Diagrams. PH traces the development of topological characteristics across the filtration sequence $\{\mathcal{X}_n\}$ and presents it through a *persistence diagram* (PD). Specifically, if a topological feature σ emerges in $\mathcal{X}_m$ and disappears in $\mathcal{X}_n$ with $1 \leq m < n \leq N$, the thresholds t_m and t_n are denoted as the *birth time* b_σ and *death time* d_σ of σ, respectively ($b_\sigma = t_m$ and $d_\sigma = t_n$). Therefore, PD contains all such 2-tuples $\mathrm{PD}_k(\mathcal{X}) = \{(b_\sigma, d_\sigma)\}$ where k represents the dimension of the topological features. The interval $d_\sigma - b_\sigma$ is termed as the *lifespan* of σ. Formally, the k^{th} persistence diagram can be defined as $\mathrm{PD}_k(\mathcal{X}) = \{(b_\sigma, d_\sigma) \mid \sigma \in H_k(\mathcal{X}_n)$ for $b_\sigma \leq t_n < d_\sigma\}$, where $H_k(\mathcal{X}_n)$ denotes the k^{th} homology group of the cubical complex $\mathcal{X}_n$. Thus, $\mathrm{PD}_k(\mathcal{X})$ contains 2-tuples indicating the birth and death times of k-dimensional voids

$\{\sigma\}$ (such as connected components, holes, and cavities) in the filtration sequence $\{\mathcal{X}_n\}$. For instance, for $\mathcal{X}$ in Figure 1, $\mathrm{PD}_0(\mathcal{X}) = \{(1,\infty),(1,3),(1,3),(2,3)\}$ represents the connected components, while $\mathrm{PD}_1(\mathcal{X}) = \{(2,5),(2,3),(4,5)\}$ illustrates the holes in the corresponding binary images in Figure 1.

Step 3 - Vectorization. Persistence Diagrams (PDs), consisting of collections of 2-tuples, are impractical for ML tools. A common alternative is *vectorization* [2], which converts PD information into vectors or functions. One widely used method is the *Betti vector*, which tracks the number of *alive* topological features at each threshold. It is a step function where $\beta_0(t_n)$ counts connected components in the binary image $\mathcal{X}_n$, and $\beta_1(t_n)$ counts holes (loops). In ML, Betti functions are typically represented as vectors $\overrightarrow{\beta_k} = [\beta_k(t_1)\dots\beta_k(t_N)]$. For example, in Figure 1, $\overrightarrow{\beta_0}(\mathcal{X}) = [3\ 2\ 1\ 1\ 1]$ shows the connected components, while $\overrightarrow{\beta_1}(\mathcal{X}) = [0\ 2\ 1\ 2\ 0]$ represents the holes.

Other PD vectorization methods include persistence images [1], landscapes [4], silhouettes [7], and kernel methods [2]. However, in this work, we primarily use Betti vectors due to their computational efficiency, interpretability, and flexibility to be represented as sequences rather than just vectors, making them well-suited for SOTA ML approaches such as transformers.

3.2 ML and DL models

We employ topological vectors in two distinct approaches to thoroughly evaluate their effectiveness in skin image analysis.

Basic ML Model. In our basic model, we evaluate the standalone performance of topological vectors by directly feeding them into ML classifiers. Using the procedure outlined in Section 3.1, we extract topological feature vectors from each image via sublevel filtration applied to each color channel. Since our vectorization approach employs Betti vectors, these representations can be treated both as static feature vectors and sequential data. To leverage this dual nature, we integrate them into state-of-the-art ML models: a multi-layer perceptron for topological embeddings (PH+MLP) and a Transformer-based sequential classifier for structured sequence modeling (PH+TR). This setup enables us to assess the optimal utilization of topological features across different ML architectures.

Hybrid Deep Learning Models. In our hybrid models, we evaluate the improvements topological vectors bring to SOTA DL models. In order to test this direction, we used pre-trained CNN models, and Vision Transformers.

Topo-ViT Model. We propose a dual-branch Transformer architecture, **Topo-ViT**, that integrates TDA with vision-based representation learning for skin lesion classification. The model consists of two parallel encoders: one for dermoscopic images and a second for PH features.

The image branch uses a Vision Transformer backbone (e.g., DaViT-Tiny) to convert a 224×224 RGB image into a sequence of patch tokens, each embedded into a vector of dimension d, capturing spatial and textural patterns.

The topology branch processes a 400-dimensional PH vector computed from sublevel filtrations across color channels. This vector is reshaped into a sequence of smaller

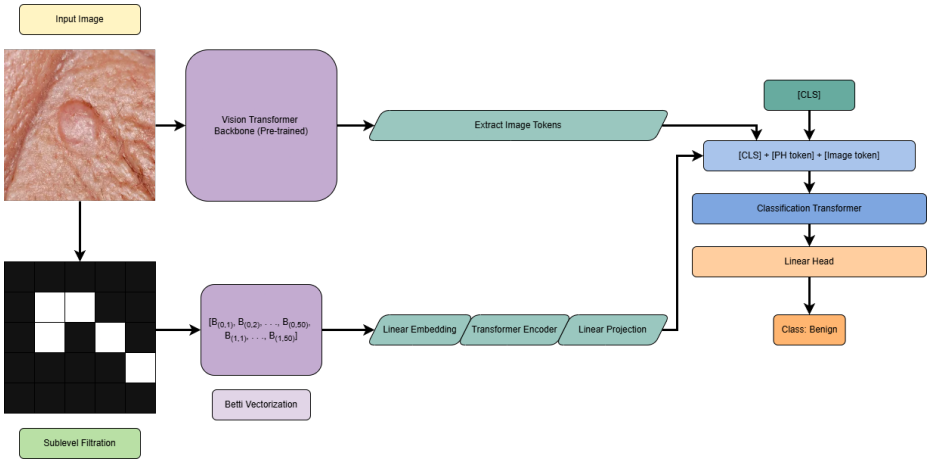

Fig. 2: **Topo-ViT Model.** Hybrid DL framework integrating topological features with Vision Transformers. PH extracts topological feature vectors from dermascopic images, which are concatenated with deep feature embeddings extracted via hierarchical self-attention mechanisms in vision Transformer blocks.

tokens each linearly projected and passed through a shallow transformer encoder. These PH tokens are then up-projected to match the embedding dimension of the image tokens.

A learnable `[CLS]` token is prepended to the combined sequence:

$$[\texttt{CLS}] + [\text{image tokens}] + [\text{PH tokens}]$$

This fused token sequence is passed through a shared classifier transformer composed of multiple layers of self-attention. By allowing all tokens to attend to each other, the `[CLS]` token aggregates joint information from both visual and topological modalities.

The final output of the `[CLS]` token is passed through a linear classification head to generate logits for skin disease prediction. By combining complementary cues—texture from images and structure from topology—**Topo-ViT** improves generalization, robustness to noise, and interpretability in melanoma and basal cell carcinoma detection tasks.

4 Experiments

4.1 Experimental Setup

Datasets. We evaluated our models using four publicly available skin image datasets that vary in size, class granularity (Table 1). Our primary dataset, *ISIC2018*, used in the ISIC 2018 Challenge [9], originally derived from the HAM10000 dataset [37] and a benchmark task in the MedMNIST v2 collection [38]. It comprises 10K dermoscopic RGB images across seven diagnostic classes. We used the dataset for both 7-class multiclass classification and binary classification, where all malignant lesions were grouped and contrasted against benign cases.

To evaluate performance of our models on smaller datasets, we used the *PH*^{2} dataset [27], which includes 200 dermoscopic images labeled as *Common Nevus*, *Atypical Nevus*, or *Melanoma*. We also used the *MedNode* dataset [15], containing 170 macroscopic images categorized as *Benign Nevus* or *Melanoma*, and the *Diverse Dermatology Images (DDI)* dataset [11], which includes 656 clinical images across two classes: *Benign* and *Malignant*. All images were resized to 224×224 pixels for consistency across models.

Table 1: Summary of dermatoscopic datasets used in this study.

Dataset	Class	# Images
ISIC2018	Melanocytic Nevus	6,705
	Melanoma	1,113
	Basal Cell Carcinoma	514
	Actinic Keratoses	327
	Benign Keratosis	1,099
	Dermatofibroma	115
	Vascular Lesions	142
PH^2	Common Nevus	80
	Atypical Nevus	80
	Melanoma	40
MedNode	Benign Nevus	100
	Melanoma	70
DDI	Benign	485
	Malignant	171

Topological Vectors. As described in Section 3.1, we apply sublevel filtration to each color channel (R, G, B, and an additional G) using 50 threshold levels per channel. This process generates 50-dimensional $\overrightarrow{\beta_0}$ and $\overrightarrow{\beta_1}$ vectors for each channel. By concatenating these vectors across all channels, we construct a 400-dimensional topological vector $\widehat{\beta}(\mathcal{X})$ for each image $\mathcal{X}$. Notably, since $\widehat{\beta}(\mathcal{X})$ is directly induced by the filtration process, it can also be interpreted as a sequential representation.

Evaluation. For the ISIC2018 dataset, which includes predefined training (7,007), validation (1,003), and test (2,005) splits [38], we used these official partitions without modification to ensure consistency and reproducibility. For all smaller datasets (PH^2, MedNode, and DDI), we employed an 80:20 stratified split to preserve class distribution between training and testing sets. Final metrics included AUC, accuracy, recall, specificity, precision, and F1 (Table 4).

Hyperparameters. For ISIC2018, Topo-CNNs were trained on 224×224 images with batch size 64 for up to 50 epochs, using early stopping (patience 5) and LR scheduling. The ImageNet-pretrained CNN backbone was frozen; its global-average-pooled features were concatenated with 400-D Betti vectors and fed through two FC layers (256, 128 units) with batch norm and dropout (0.3, 0.2). We optimized with Adam (LR=1e-3), using sparse categorical cross-entropy for the 7-class task and binary cross-entropy (threshold 0.35) for the binary task.

Our default Topo-ViT reshapes each 400-D PH vector into 25 tokens of 16 D, projects them to 128 D, and encodes with a 4-head transformer. The resulting topological embeddings are linearly mapped to the ViT's feature width (768 D, or 512 D for MobileViT) and concatenated with image patch tokens plus a learnable class token. A 6-layer transformer decoder (width 768, 8 heads) fuses these tokens, with uniform dropout (0.3) across all embedding layers, transformer blocks, and the classification head. Our code is available at the following link: [1].

[1] Code link: https://anonymous.4open.science/r/TopoDerma-28EB/

Table 2: AUC results across all datasets. T-CNN and T-ViT denote the corresponding CNN and ViT architectures augmented with topological vectors via attention. The final column shows the mean AUC improvement of each topological model over its vanilla counterpart.

Model	ISIC (7)	ISIC (2)	DDI (2)	PH^2 (3)	MedNode (2)	Av. Imp
PH+MLP	87.90	83.06	66.03	81.21	74.29	
PH+TR	87.42	81.60	66.63	83.42	87.50	2.82
Inception v3 [35]	93.17	85.30	60.02	71.18	74.28	
T-Inception v3	89.03	84.64	73.58	75.69	80.00	3.80
DenseNet121 [18]	92.92	85.52	61.10	66.84	81.43	
T-DenseNet121	91.47	89.28	71.78	82.81	74.28	4.36
MobileNet v2 [32]	61.28	79.01	50.78	66.84	77.14	
T-MobileNet v2	89.28	87.89	73.10	80.38	57.14	10.55
MobileViT [26]	94.09	93.78	69.06	82.81	90.71	
T-MobileViT	95.25	92.06	70.80	89.41	92.14	1.84
Swin v2 [22]	97.95	94.95	78.24	87.57	88.93	
T-Swin v2	**98.01**	95.48	79.62	87.76	94.29	1.50
DaViT [13]	94.93	95.57	76.32	83.46	91.43	
T-DaViT	95.55	**95.73**	**83.33**	**89.50**	**94.64**	3.41

4.2 Results and Discussion

Tables 2 and 4 report the performances for baseline and topologically-enhanced models across five skin-lesion datasets. Across the large ISIC2018 benchmark, topological augmentation yields modest gains (+1–4 pts), while CNNs and ViTs augmented with persistence features consistently outperform their vanilla counterparts. For example, T-DenseNet121 and T-Inception v3 improve by +4.36 and +3.80 AUC points on average, and even the lightweight MobileNetV2 backbone jumps by +10.55 pts when fused with Betti vectors. Standalone topological classifiers remain competitive.

Table 3: MedMNIST baselines for ISIC 2018 (7) [38].

Model	AUC	ACC
ResNet-18 (28)	91.7	73.5
ResNet-18 (224)	92.0	75.4
ResNet-50 (28)	91.3	73.5
ResNet-50 (224)	91.2	73.1
auto-sklearn	90.2	71.9
AutoKeras	91.5	74.9
Google AutoML Vision	91.4	76.8
T-Swin	98.0	88.2

Notably, the benefits of topological vectors are more pronounced on smaller datasets (DDI, PH^2, and MedNode), where deep models often struggle to generalize. In these low-data regimes, persistence-based features provide robust, scale-invariant signals that help guide learning. As a result, AUC gains often exceed 10 points—for example, +10.68 for T-DenseNet121 on DDI and +15.97 on PH^2—while improvements on the larger ISIC dataset remain modest. T-DaViT achieves the highest AUCs across all three small datasets (83.33, 89.50, and 94.64), highlighting the value of topological augmentation in data-scarce settings. Compared to MedMNIST baselines on DermaMNIST (ISIC2018 (7)) [38], our Topo-ViT models also achieve superior performance.

Table 4: Performance comparison of models across all datasets. T-CNN and T-ViT models denote the corresponding CNN and ViT architectures augmented with topological vectors.

	ISIC2018 (7-class)					**ISIC2018 (Binary)**				
Model	AUC	Acc	Sens	Spec	F-1	AUC	Acc	Sens	Spec	F-1
PH+MLP	87.90	71.67	71.67	91.27	68.56	83.06	80.30	64.80	84.07	56.26
PH+TR	87.42	70.97	36.52	91.93	39.48	81.60	80.00	65.17	65.17	44.38
Inception v3	93.17	75.81	57.44	93.01	50.90	85.30	82.29	17.60	98.02	27.99
T-Inception v3	89.03	72.91	33.23	65.38	36.19	84.64	81.64	60.20	86.85	56.19
DenseNet121	92.92	69.58	27.83	87.11	31.83	85.52	82.49	52.81	79.79	45.92
T-DenseNet121	91.47	75.52	45.48	85.71	48.52	89.28	83.04	68.62	86.54	61.27
MobileNetV2	61.28	64.39	17.61	85.75	33.38	79.01	58.2	91.33	50.15	46.07
T-MobileNetV2	89.28	74.36	36.83	34.21	39.80	87.89	83.59	67.61	87.47	61.69
MobileViT	94.09	87.23	70.87	96.60	72.32	93.78	88.98	82.63	82.63	82.53
T-MobileViT	95.25	82.54	56.15	95.13	60.93	92.06	86.13	80.67	80.67	66.90
DaViT	94.93	89.38	76.13	97.21	78.79	95.57	91.67	85.46	85.46	86.40
T-DaViT	95.55	85.69	70.88	96.44	71.00	95.73	91.32	85.24	85.24	77.23
Swin v2	97.95	87.38	75.95	96.61	77.81	94.95	89.68	86.63	86.63	84.51
T-Swin v2	98.01	88.18	74.60	96.76	78.09	95.48	90.92	84.51	84.51	76.12

	DDI (Binary)					**PH^2 (3-class)**					**MedNode (Binary)**				
Model	AUC	Acc	Sens.	Spec.	F1	AUC	Acc	Sens.	Spec.	F1	AUC	Acc	Sens.	Spec.	F1
PH+MLP	66.03	74.24	47.06	83.67	48.48	81.21	70.00	70.00	84.38	69.94	74.29	67.65	78.57	60.00	66.67
PH+TR	66.63	75.76	35.29	89.80	42.86	83.42	72.50	70.83	85.42	71.47	87.50	79.41	78.57	80.00	75.86
DenseNet121	61.10	77.27	11.76	95.91	19.04	66.84	45.00	45.83	70.14	47.19	81.43	76.47	100.00	42.85	83.33
T-DenseNet121	71.78	66.67	58.82	69.38	47.62	82.81	55.00	58.33	77.08	49.62	74.28	70.58	85.71	60.00	70.58
InceptionV3	60.02	66.67	41.17	75.51	38.89	71.18	55.00	54.16	75.00	56.50	74.28	58.82	80.00	58.57	69.56
T-InceptionV3	73.58	74.24	76.47	73.47	60.46	75.69	50.05	54.16	75.00	48.90	80.00	76.47	85.71	70.00	75.00
MobileNetV2	50.78	75.75	17.64	95.91	27.27	66.84	55.00	58.33	75.69	56.11	77.14	70.58	60.00	85.71	70.58
T-MobileNetV2	73.10	69.69	64.71	71.43	52.38	80.38	65.00	62.50	80.56	63.59	57.14	47.06	85.71	50.00	66.67
MobileViT	69.06	72.73	61.46	61.46	41.94	82.81	67.50	70.83	82.99	68.54	90.71	82.35	82.86	82.86	80.00
T-MobileViT	70.80	75.76	75.76	65.43	48.39	89.41	75.00	79.17	87.15	76.17	92.14	79.41	77.14	77.14	72.00
DaViT	76.32	73.48	62.94	62.94	44.44	83.46	70.00	72.92	83.68	72.92	91.43	79.41	81.43	81.43	78.79
T-DaViT	83.33	72.50	75.00	85.42	73.92	89.50	77.50	79.17	87.50	80.32	94.64	94.12	92.86	92.86	92.31
Swin v2	78.24	80.30	68.49	68.49	53.57	87.57	70.00	72.92	84.03	71.74	88.93	73.53	74.29	74.29	70.97
T-Swin v2	79.62	81.06	70.92	70.92	57.63	87.76	72.50	72.92	85.07	73.89	94.29	91.18	89.29	89.29	88.00

5 Conclusion

We presented a topological deep learning framework for skin cancer detection, combining persistent homology–based features with convolutional and transformer architectures. Our experiments demonstrate that topological signatures alone are competitive, and their integration with deep models consistently improves classification accuracy across diverse skin lesion types. By capturing global shape and connectivity information, topological features enhance both robustness and interpretability, key limitations of conventional deep learning. In future work, we plan to extend our framework to multimodal data by incorporating patient metadata and clinical images, explore end-to-end differentiable topological layers, and validate our approach through prospective clinical studies. This study highlights the potential of TDA to improve diagnostic performance and interpretability in dermatological AI systems.

Acknowledgements. This work was partially supported by National Science Foundation under grants DMS-2220613, and DMS-2229417.

References

1. Adams, H., et al.: Persistence images: A stable vector representation of persistent homology. Journal of Machine Learning Research **18**(8), 1–35 (2017)
2. Ali, D., et al.: A survey of vectorization methods in topological data analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence (2023)
3. Brinker, T.J., et al.: CNN outperforms dermatologists in melanoma classification: A prospective, multicenter, noninferiority study. JAMA Dermatology **155**(11), 922–926 (2019)
4. Bubenik, P., Dłotko, P.: A persistence landscapes toolbox for topological statistics. Journal of Symbolic Computation **78**, 91–114 (2017)
5. Carlsson, G.: Topology and data. Bulletin of the American Mathematical Society **46**(2), 255–308 (2009)
6. Chachólski, W., Hiraoka, Y., Obayashi, I.: Topological data analysis for complex medical imaging. IEEE Transactions on Medical Imaging **38**(6), 1389–1398 (2019)
7. Chazal, F., Fasy, B.T., Lecci, F., Rinaldo, A., Wasserman, L.: Stochastic convergence of persistence landscapes and silhouettes. In: SoCG. pp. 474–483 (2014)
8. Chung, Y.M., Hu, C.S., Lawson, A., Smyth, C.: Topological approaches to skin disease image analysis. In: 2018 IEEE International Conference on Big Data (Big Data). pp. 100–105. IEEE (2018)
9. Codella, N.C.F., et al.: Skin lesion analysis toward melanoma detection: Challenge at ISBI 2018. In: ISBI. pp. 168–172 (2018)
10. Coskunuzer, B., Akçora, C.G.: Topological methods in machine learning: A tutorial for practitioners. arXiv preprint arXiv:2409.02901 (2024)
11. Daneshjou, R., et al.: Disparities in dermatology ai performance on a diverse, curated clinical image set. Science Advances **8**(6), eabj7790 (2022)
12. Dey, T.K., Wang, Y.: Computational topology for data analysis. Cambridge University Press (2022)
13. Ding, M., et al.: Davit: Dual attention vision transformers. In: ECCV. pp. 74–92. Springer (2022)
14. Esteva, A., et al.: Dermatologist-level classification of skin cancer with deep neural networks. Nature **542**(7639), 115–118 (2017)
15. Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M.F., Petkov, N.: Med-node: A computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert Systems with Applications **42**(19), 6578–6585 (2015)
16. Gutman, D.A., et al.: Skin lesion analysis toward melanoma detection: A challenge at the 2016 international symposium on biomedical imaging (isbi). In: ISBI. pp. 168–172 (2016)
17. Hofer, C., Kwitt, R., Niethammer, M., Uhl, A.: Deep learning with topological signatures. In: Advances in Neural Information Processing Systems. vol. 30 (2017)
18. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: CVPR. pp. 4700–4708 (2017)
19. Johnson, K., Söderkvist, J., Hansson, P.: Application of persistent homology in histopathological grading. Computers in Biology and Medicine **145**, 105379 (2022)
20. Kahle, R., Rieck, B., Moor, M., Horn, M.: Topological regularization for deep image segmentation. In: Proceedings of the IEEE International Symposium on Biomedical Imaging. pp. 235–239 (2021)

21. Li, Y., Shen, L.: Integrating dermoscopic images and patient metadata for enhanced melanoma detection using deep learning. Computers in Biology and Medicine **123**, 103900 (2020). https://doi.org/10.1016/j.compbiomed.2020.103900
22. Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., Ning, J., Cao, Y., Zhang, Z., Dong, L., et al.: Swin transformer v2: Scaling up capacity and resolution. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 12009–12019 (2022)
23. Lum, P., et al.: Extracting insights from the shape of complex data using topology. Scientific Reports **3**, 1236 (2013)
24. Maurya, A., et al.: Hybrid topological data analysis and deep learning for basal cell carcinoma diagnosis. Journal of Imaging Informatics in Medicine **37**(1), 92–106 (2024)
25. McGuirl, M.R., Volkening, A., Sandstede, B.: Topological data analysis of zebrafish patterns. PNAS **117**(10), 5113–5124 (2020)
26. Mehta, S., Rastegari, M.: Mobilevit: Light-weight, general-purpose, and mobile-friendly vision transformer. In: ICLR (2022)
27. Mendonça, T., Ferreira, P.M., Marques, J.S., Marcal, A., Rozeira, J.: Ph^2: A dermoscopic image database for research and benchmarking. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) pp. 5437–5440 (2013)
28. Menegola, A., Tavares, J.M.R.S., Fornaciali, M., Li, L., Avila, S., Valle, E.: Knowledge transfer for melanoma screening with deep learning. arXiv preprint arXiv:1703.02223 (2017)
29. Patel, S., et al.: Artificial intelligence in dermatology: A systematic review of applications and challenges. Journal of the American Medical Informatics Association **28**(8), 1842–1857 (2021)
30. Qaiser, T., et al.: Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features. Medical image analysis **55**, 1–14 (2019)
31. Saggar, M., et al.: Towards a new approach to reveal dynamical organization of the brain using topological data analysis. Nature Communications **9**(1), 1399 (2018)
32. Sandler, M., et al.: MobileNetV2: Inverted residuals and linear bottlenecks. In: CVPR. pp. 4510–4520 (June 2018). https://doi.org/10.1109/CVPR.2018.00474
33. Santhirasekaram, A., Winkler, M., Rockall, A., Glocker, B.: Topology preserving compositionality for robust medical image segmentation. In: Proceedings of the IEEE/CVF Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 999–1008 (2023)
34. Skaf, Y., Laubenbacher, R.C.: Topological data analysis in biomedicine: A review. Journal of Biomedical Informatics **130**, 104082 (2022). https://doi.org/10.1016/j.jbi.2022.104082
35. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)
36. Tang, P., Li, X., He, J., Zheng, Y.: Automated melanoma detection using deep learning based on multiple image modalities. IEEE Journal of Biomedical and Health Informatics **24**(7), 1922–1931 (2020)
37. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multisource dermatoscopic images of common pigmented skin lesions. Scientific Data **5**(1), 180161 (2018)
38. Yang, J., Shi, R., Wei, D., Liu, Z., Zhao, L., Ke, B., Pfister, H., Ni, B.: Medmnist v2-a large-scale lightweight benchmark for 2d and 3d biomedical image classification. Scientific Data **10**(1), 41 (2023), https://medmnist.com
39. Yuan, L., Chen, Y., Wang, T., Yu, W., Shi, Y., Jiang, Z., Tay, F.E.H., Feng, J., Yan, S.: Tokens-to-token vision transformer: Training vision transformers from scratch on imagenet. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 558–567 (October 2021)