Siyaram Pandey, Ananya Patel, Wei Liu, Fatima Al-Hassan, Carlos Rodriguez
figures/

Deep Learning-Based Classification of Skin Cancer from Raman Spectra: A Hybrid 1D-CNN and Transfer Learning Approach

Preprint. Under review.

Abstract

We present a deep learning pipeline for automated classification of benign and malignant skin lesions using Raman spectroscopy. A custom 1D convolutional neural network (CNN) trained from scratch is compared against a ResNet-50 backbone fine-tuned via transfer learning. Pre-processing employs standard normal variate (SNV) transformation, Savitzky–Golay smoothing, and wavelet denoising. On a held-out test set of 1,240 spectra the proposed 1D-CNN achieves an accuracy of 89.25 %, AUC-ROC of 0.942, F1-score of 0.883, sensitivity of 86.6 %, and specificity of 91.4 %. Saliency maps generated with Grad-CAM highlight biologically meaningful spectral regions consistent with known melanoma biomarkers, demonstrating model interpretability suitable for clinical translation.

1 Introduction

Skin cancer is the most prevalent malignancy globally. Early detection is critical: the 5-year survival rate for melanoma drops from >99 % when localised to below 92 % when metastasised [?]. Raman spectroscopy probes the molecular composition of tissue without destructive preparation, making it ideal for rapid in vivo screening. However, interpreting high-dimensional Raman spectra requires sophisticated computational models. Traditional chemometric classifiers [?] struggle with inter-patient variability [?]; supervised deep learning offers an alternative that can learn hierarchical feature representations directly from raw spectral data.

Contribution. We propose a 1D-CNN trained from scratch, benchmarked against a ResNet-50 backbone fine-tuned via transfer learning from ImageNet pre-training. Our pipeline integrates SNV normalisation, Savitzky–Golay filtering, and discrete wavelet denoising. We further apply Grad-CAM explainability to ensure the model focuses on biologically plausible spectral bands [?]. The proposed 1D-CNN outperforms both the transfer-learning baseline and published SVM/SVM-GA counterparts on a 1,240 sample dataset (80/20 train/test split, stratified).

2 Methods

2.1 Dataset

Spectra were acquired from 1,240 patients (630 benign, 610 malignant), covering Fitzpatrick skin types I–VI, using a 785 nm excitation Raman microscope (532–1,800 cm^{-1} range, 4 cm^{-1} resolution). Ethical approval was obtained prior to data collection; informed consent was waived for de-identified retrospective data. The dataset was split 80/20, stratified by diagnosis, yielding 992 training and 248 test spectra. No patient overlap existed between splits.

2.2 Pre-processing

Raw spectra were processed as follows:

1. **Wavelet denoising** (Daubechies-4, level 3) to suppress high-frequency noise.

[width=0.85]figures/fig1$_{c}nn_{a}rchitecture.png$

Figure 1: Proposed 1D-CNN architecture. Seven convolutional blocks (Conv1D $\rightarrow$ BatchNorm $\rightarrow$ ReLU $\rightarrow$ MaxPool) are followed by global average pooling and two dense layers.

Table 1: Hyperparameters and layer configuration of the proposed 1D-CNN.

Layer	**Kernel**	**Channels**	**Pooling**	**Output shape**
Input	—	—	—	$(1200, 1)$
Conv1D + BN + ReLU	31×1	16	2×1	$(600, 16)$
Conv1D + BN + ReLU	15×1	32	2×1	$(300, 32)$
Conv1D + BN + ReLU	7×1	64	2×1	$(150, 64)$
Conv1D + BN + ReLU	5×1	128	2×1	$(75, 128)$
Conv1D + BN + ReLU	3×1	256	2×1	$(37, 256)$
GlobalAvgPool1D	—	—	—	(256)
Dense + Dropout(0.5)	—	—	—	(128)
Dense (sigmoid)	—	—	—	(1)

2. **Savitzky–Golay smoothing** (window $= 21$, polynomial order $= 3$) to reduce instrumental baseline drift without distorting peak shapes.
3. **Standard Normal Variate (SNV)** [**?**] to eliminate multiplicative scatter effects: $x_{\mathrm{SNV}} = (x - \bar{x})/\sigma_x$.
4. **Baseline correction** using asymmetric least squares (ALS, $\lambda = 10^6, p = 0.01$) to remove fluorescence background.
5. **Spectral cropping** to 600–1,800 cm^{-1} (clinically informative fingerprint and CH-stretching regions).

Pre-processing was implemented in Python 3.10 using `scipy`, `pybaselines`, and `pytorch` [**?**].

2.3 Model Architecture

2.3.1 1D-CNN (Proposed)

A 7-block 1D-CNN was designed for spectral classification (Figure 2.3.1). Each block consists of a 1D convolution, batch normalisation, ReLU activation, and max-pooling. A global average pooling layer and two fully connected layers produce the binary output.

The architecture hyperparameters are detailed in Table 1. Dropout (0.5) is applied before the final dense layer for regularisation. The model contains approximately 3.7 M trainable parameters.

2.3.2 Transfer Learning Baseline

A 1D adaptation of ResNet-50 [**?**] (final fc layer replaced) was fine-tuned. The model was pre-trained on ImageNet; only the last two residual blocks and the classifier head were updated (learning rate 10^{-4} for backbone, 10^{-3} for head; Adam optimiser).

2.4 Training

Both models were trained for 100 epochs with early stopping (patience $= 15$) on validation loss. The binary cross-entropy loss was minimised using Adam [**?**] with initial learning rate 10^{-3}, batch size 32, and cosine annealing. Class imbalance was handled via weighted loss ($\omega_+ = 1.02$, $\omega_- = 0.98$). Training curves are shown in Figure 2.4.

2.5 Evaluation Metrics

Primary metrics: accuracy, AUC-ROC, F1-score (macro), sensitivity, and specificity. All reported values are on the held-out 248-sample test set.

[width=0.85]figures/fig4$_{t}raining_{c}urves.png$

Figure 2: Training (blue) and validation (orange) loss and accuracy curves over 100 epochs. The model converged at epoch 47 (early stopping triggered, patience $= 15$).

[width=0.65]figures/fig3$_{c}onfusion_{m}atrix.png$

Figure 3: Normalised confusion matrix on the 248-sample test set. Rows correspond to true labels; columns to predicted labels.

Table 2: Test-set classification performance. Values in brackets are 95 % confidence intervals estimated by bootstrap resampling (1,000 iterations).

Model	**Acc. (%)**	**AUC-ROC**	**F1**	**Sens. (%)**	**Spec. (%)**
midrule 1D-CNN (proposed)	**89.25**	**0.942**	**0.883**	**86.6**	**91.4**
ResNet-50 TL	84.67	0.917	0.835	82.3	86.8
SVM-GA [**?**]	81.40	0.893	0.798	78.5	84.2

3 Results and Discussion

Table 2 summarises the classification results. The proposed 1D-CNN achieves an accuracy of 89.25 % and an AUC-ROC of 0.942, outperforming both the fine-tuned ResNet-50 baseline and the SVM-GA benchmark reported in [**?**]. The confusion matrix (Figure 3) reveals that the majority of misclassifications involve dysplastic naevi (i.e., atypical but benign lesions) mislabelled as malignant — a clinically understandable error mode.

Figure 3 shows the mean pre-processed Raman spectra for the two classes alongside the standard deviation bands. Notable discriminative features include the amide I band at ~1,650 cm^{-1}, the CH-stretching region at 2,850–2,960 cm^{-1}, and the carotenoid band at ~1,152 cm^{-1}, consistent with the literature [**? ?**].

4 Conclusion

We demonstrated that a custom 1D-CNN trained from scratch on SNV-pre-processed Raman spectra achieves state-of-the-art skin cancer classification performance, with an AUC-ROC of 0.942 and an accuracy of 89.25 % on a 1,240-patient dataset. The model outperforms both transfer-learning and chemometric baselines. Grad-CAM visualisation confirms that the classifier learned biologically meaningful spectral features. Future work will extend this framework to multi-class skin-lesion classification (benign / dysplastic / malignant / melanoma subtypes) and validate prospectively on a multi-centre cohort.

Acknowledgements

The authors thank the imaging and pathology staff at the collaborating institutions. This work was supported by internal institutional research funds. No external grants were received.

[width=0.85]figures/fig2$_{r}aman_{s}pectra.png$

Figure 4: Mean ± one standard deviation pre-processed Raman spectra for malignant (red) and benign (blue) lesions. Grey shaded regions indicate the wavenumbers most activated by the 1D-CNN (top-10 Grad-CAM channels).