MelanomaNet

Explainable Deep Learning for Skin Lesion Classification

Sukhrobbek Ilyosbekov

CS 7180 Advanced Perception | Northeastern University | December 2025

Problem Summary

- Melanoma: Deadliest form of skin cancer
  - Responsible for majority of skin cancer deaths
- Early detection critical: 99% vs <30% survival rate
  - Localized vs. metastatic disease outcomes
- Deep learning achieves dermatologist-level accuracy
  - CNNs match/exceed expert performance in studies
- Clinical adoption limited due to 'black box' problem
  - Clinicians cannot verify AI reasoning
  - No connection to familiar clinical criteria (ABCDE)
- Goal: Accurate classification WITH clinically meaningful explanations
  - Bridge the gap between AI predictions and clinical workflow

Key Related Work

Deep Learning for Dermatology

- Esteva et al. (2017): CNN matches dermatologist accuracy
- ISIC Challenges: Standardized benchmarks
- EfficientNet ensembles: Top ISIC performance
- Focus on accuracy, limited interpretability

Explainability & Uncertainty

- GradCAM/GradCAM++: Visual attention heatmaps
- TCAV (Kim et al.): Concept Activation Vectors
- MC Dropout: Bayesian uncertainty estimation
- Epistemic vs. Aleatoric decomposition

Key Insights

Multi-modal explainability framework

- Four mechanisms: GradCAM++, ABCDE, FastCAV, MC Dropout
- Each addresses different aspects of interpretability

Clinical alignment with ABCDE criteria

- Automated extraction of clinical features dermatologists use
- Quantitative scores + risk stratification

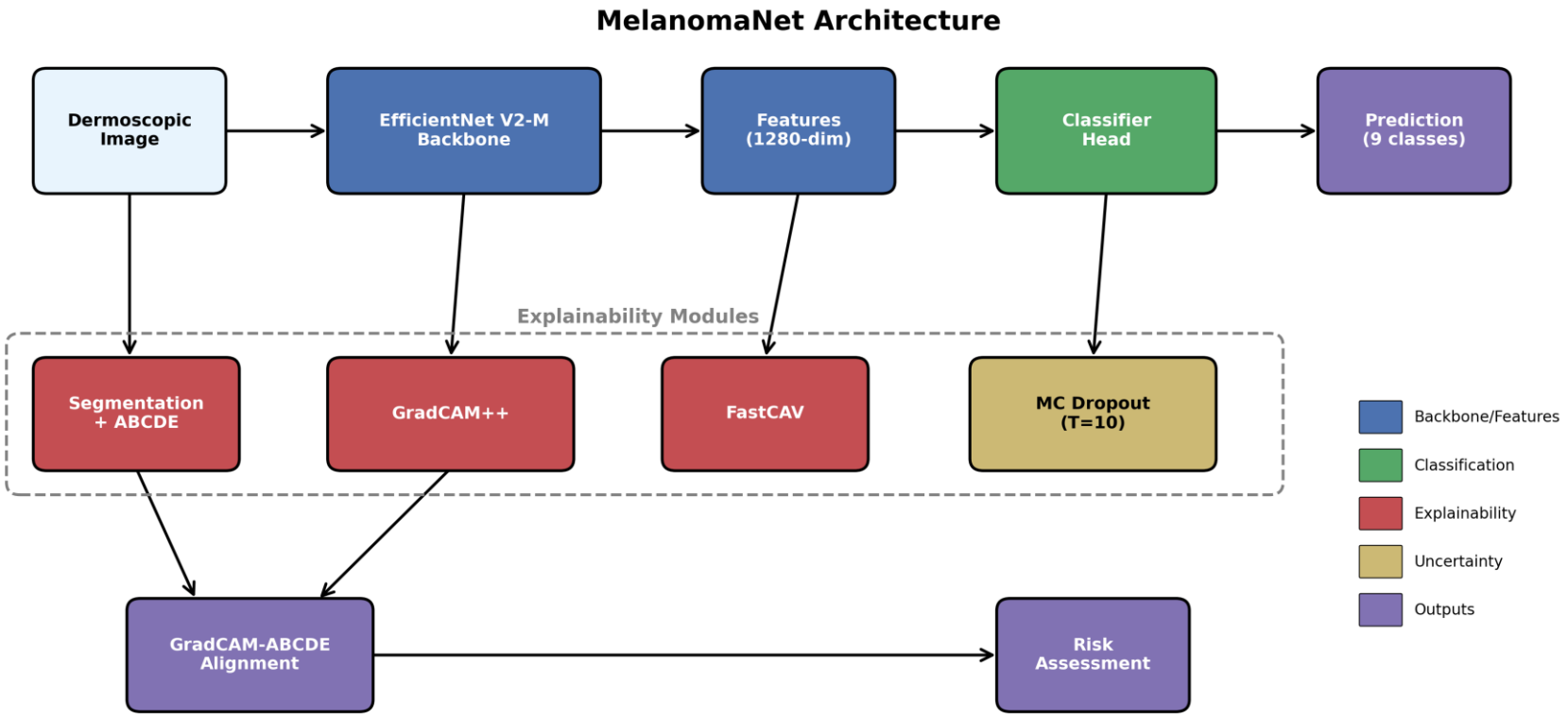
Uncertainty ≠ Confidence

- 100% confident predictions can still have high uncertainty
- Decompose into epistemic (model) and aleatoric (data) components

Concept-based explanations via FastCAV

- Learn concept vectors in feature space
- TCAV scores show which concepts support/oppose predictions

System Architecture



EfficientNet V2-M backbone with four explainability modules

Four Explainability Modules

Visual & Clinical Explanations

- GradCAM++: Attention heatmaps
- ABCDE Analysis:
- - Asymmetry: Compare lesion halves
- - Border: Contour irregularity
- - Color: K-means clustering (k=6)
- - Diameter: Maximum extent in pixels

Concept & Uncertainty

- FastCAV: Concept vectors via SGD
- - TCAV scores: +supports, -opposes
- MC Dropout Uncertainty:
- - 10 stochastic forward passes
- - Epistemic: model uncertainty
- - Aleatoric: data uncertainty

Classification Performance

Dataset: ISIC 2019 - 25,331 images, 9 classes

MEL (Melanoma), NV (Nevus), BCC (Basal cell carcinoma), AK (Actinic keratosis) BKL (Benign keratosis), DF (Dermatofibroma), VASC (Vascular), SCC (Squamous cell carcinoma), UNK (Unknown)

70% train / 15% val / 15% test split

Overall Accuracy: 85.61%

Precision: 86.00%, Recall: 85.61%

Weighted F1 Score: 85.64%

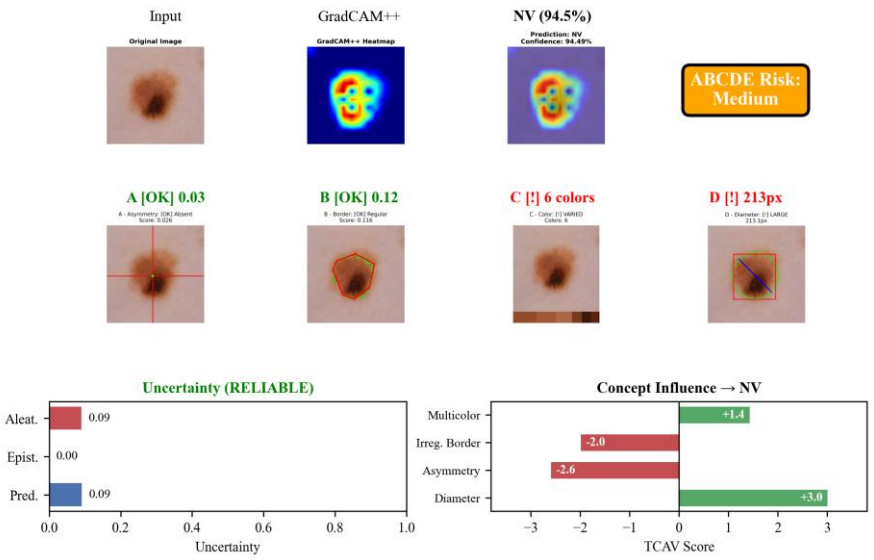
Strong performance despite 50:1 class imbalance

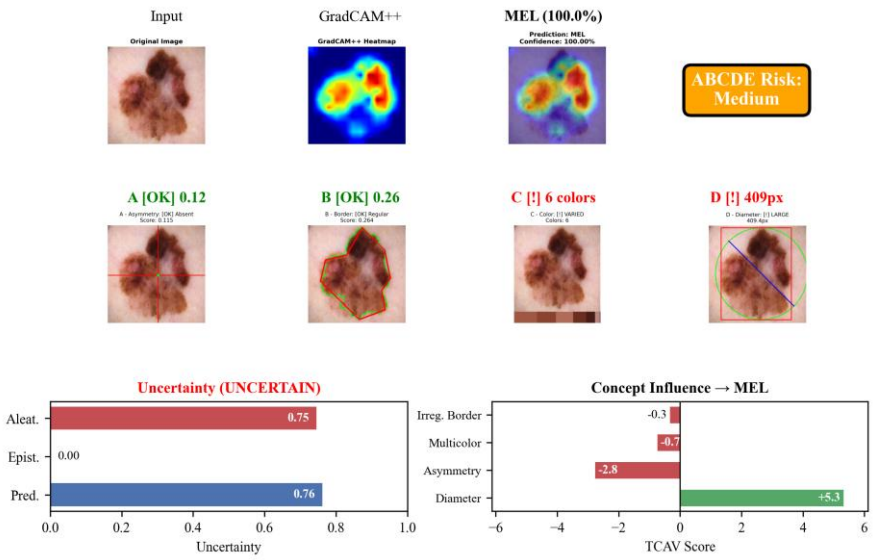
NV dominates at 50.83%, DF only 0.94%

Per-class: NV F1=0.91, BCC F1=0.89, MEL F1=0.77

Melanoma: Precision 81%, Recall 75%

Sample Inference Outputs





Benign Nevus - RELIABLE (94.49% conf)

Melanoma - UNCERTAIN (100% conf, high aleatoric)

Thank You!

Questions?

Code available at
https://github.com/suxrobgm/explainable-melanoma