

Image Based Detection of Nail Melanoma Using Deep Learning Techniques

Ayca Öztürk - University of Würzburg, Germany

Introduction

- Nail Apparatus Melanoma (NAM):** a rare but serious type of melanoma (skin cancer)
- Key Sign:** pigmented nail band
- Challenge:** delayed diagnosis, due to unclear symptoms (e.g., subtle signs, similar to other diseases)
- Outcome:** poorer prognosis (e.g., surgery, amputation)
- AI:** can support early diagnosis

Motivation & Aim

Why Deep Learning?

- Nail diseases are image-based
- Clinical images are easy to collect
- Deep Learning achieves expert-level accuracy

Aim of the project:

- Multi-stage deep-learning pipeline for automatic NAM detection

1. Dataset

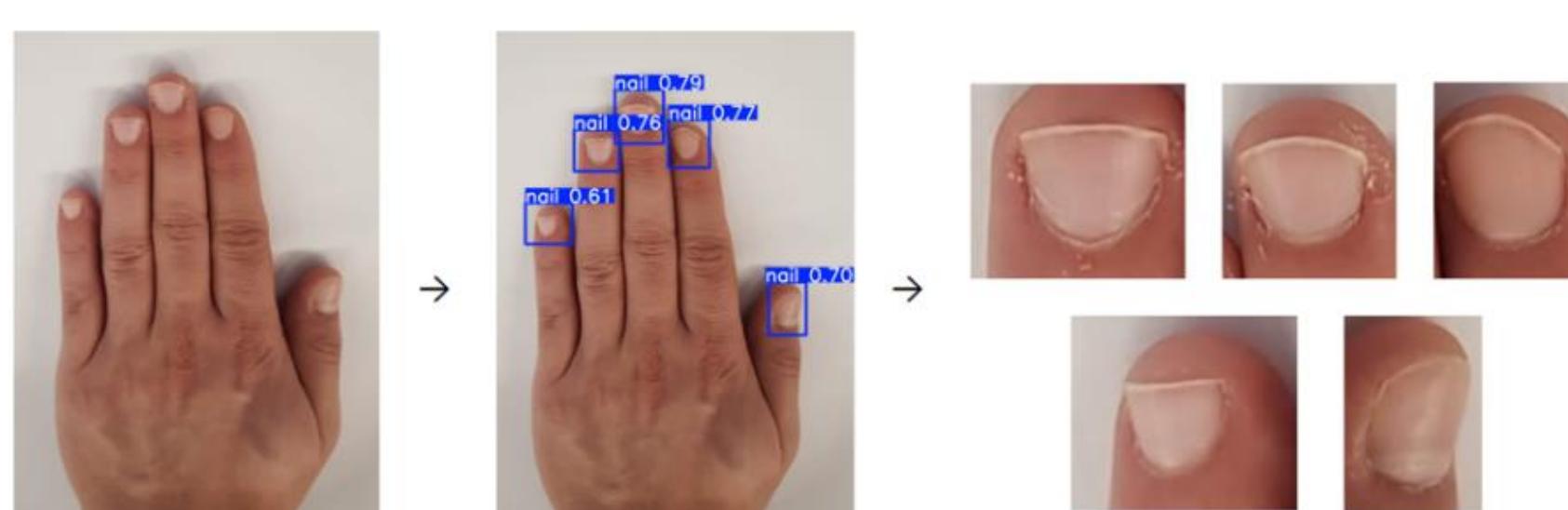
Sources: Universitätsklinikum Würzburg, Kaggle, Google Images

Class	Number of Images
Nail Melanoma	473
Healthy	491
Onychomycosis	574
Nail Psoriasis	684
Onycholysis	110
Beau's Lines	102
Koilonychia	38
Leukonychia	31
in total	2503

Methodology

2. Nail Region Detection using YOLOv8

- YOLOv8 detects bounding boxes for nails
- Nails get cropped & resized for standardization



3. Training Deep Learning Classification Models

Classes: healthy vs melanoma

- | | |
|--|--|
| CNNs | ViTs |
| <ul style="list-style-type: none"> ResNet-18 VGG-19 EfficientNet-B0 Custom CNN | <ul style="list-style-type: none"> ViT-B/16 |

Hybrid CNN

- YOLOv8-cls

4. Evaluating the Models in Binary and Multi-Class Settings

Binary Classification

Classes:
melanoma vs
non-melanoma
(healthy + other
diseases)

Multi-Class Classification

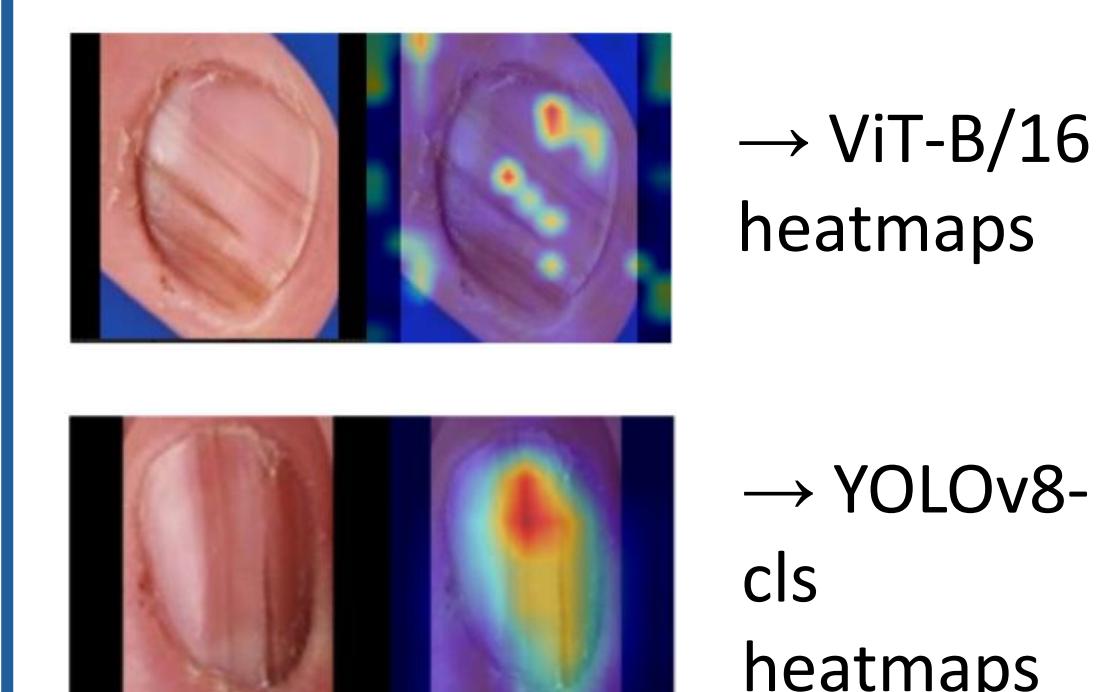
Classes:
healthy,
melanoma,
other

5. Fine-tuning & Metrics

- Best Models, which are fine-tuned:** ViT-B/16, YOLOv8-cls
- Tuned Hyperparameters :** epochs, batch size, learning rate
- Evaluation Metrics:** confusion matrix, precision, recall, F1-score (balance between precision and recall), accuracy

6. Explainable AI (XAI)

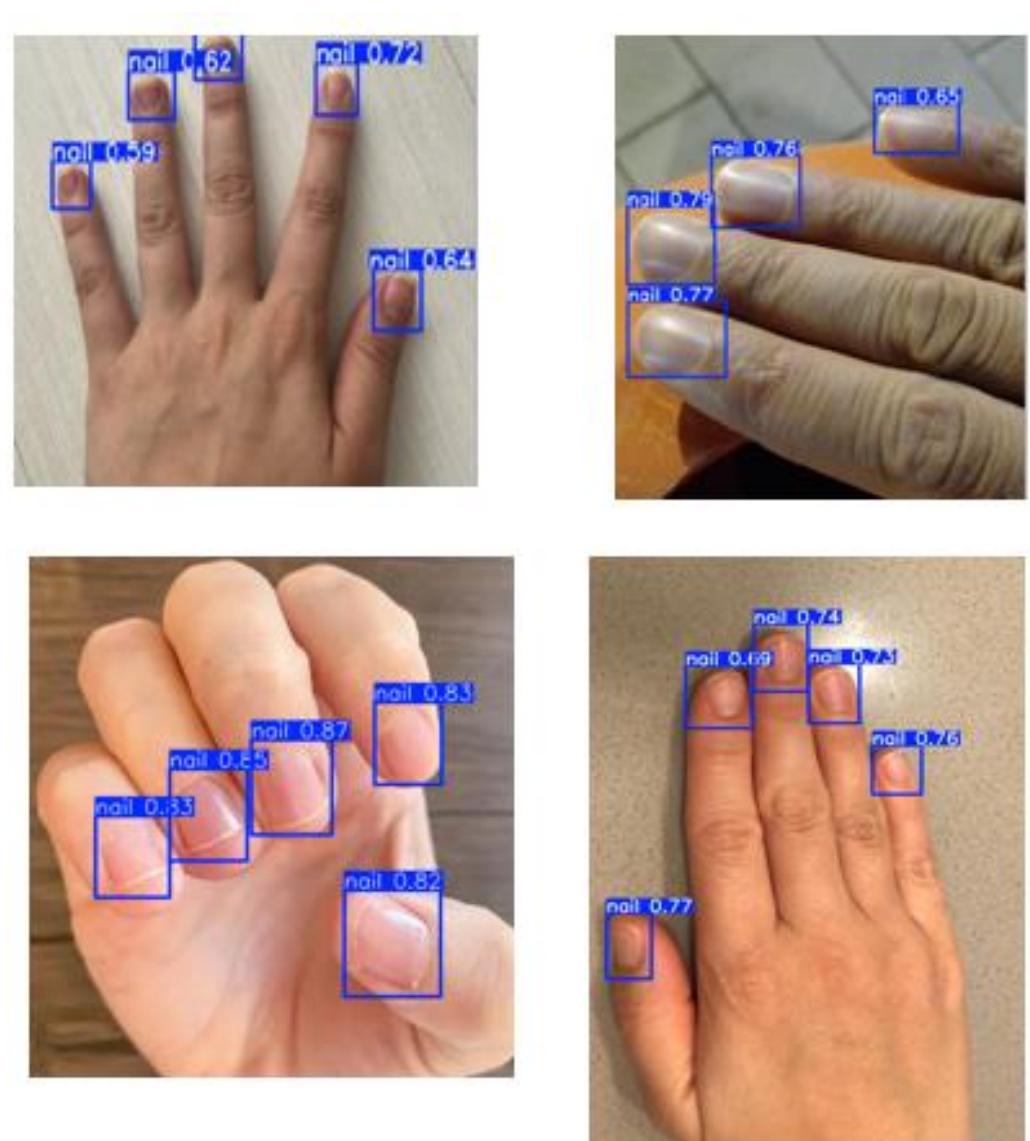
- Generation of heatmaps to visualize model focus regions



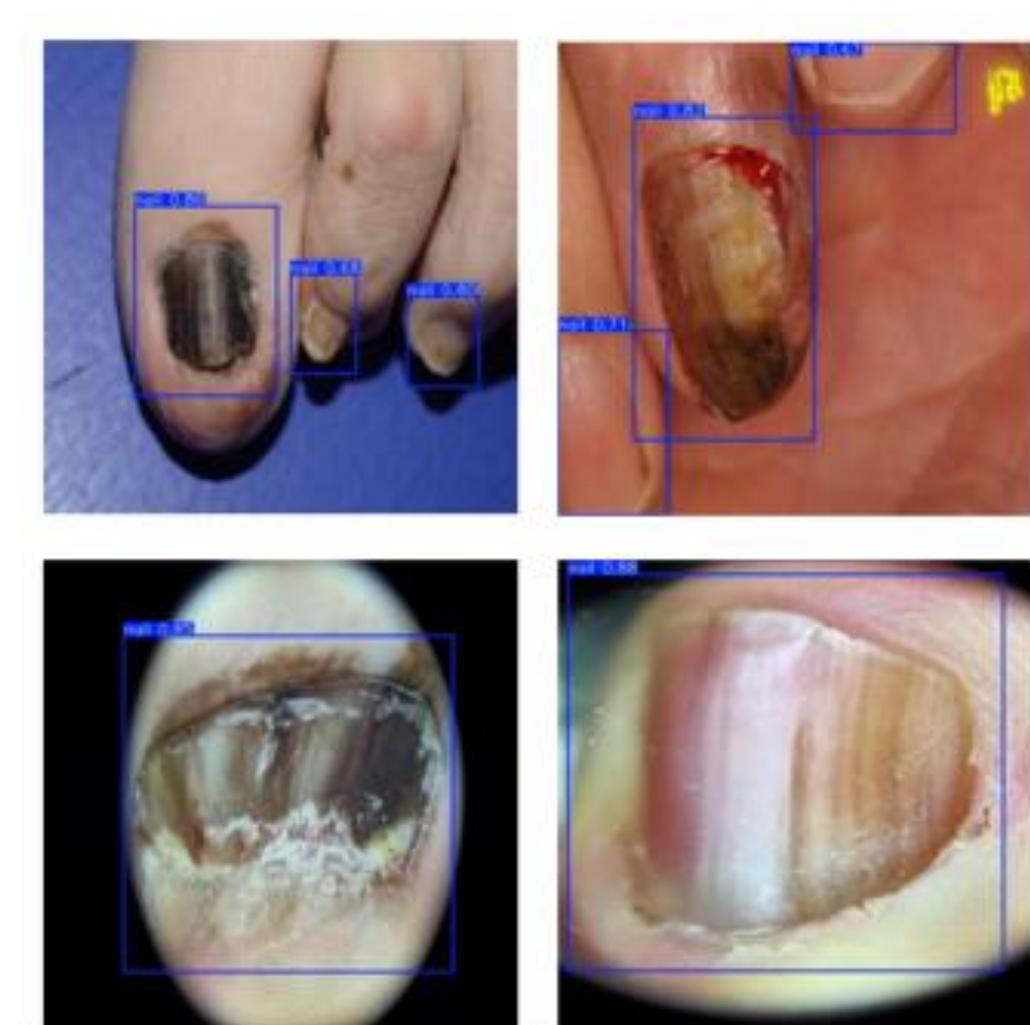
Results

1. Nail Region Detection Results

I. Bounding box detections of healthy nails



II. Bounding box detections of nails with melanoma



2. Classification Performance of the models (healthy vs melanoma nails)

Model	NAM Precision	NAM Recall	NAM F1-Score	Accuracy
ResNet-18	0.940	0.902	0.921	0.960
VGG-19	0	0	0	0.747
EfficientNet-B0	0	0	0	0.747
Custom CNN	1.000	0.313	0.472	0.830
ViT-B/16	1.000	0.990	1.000	0.990
YOLOv8-cls	1.000	0.992	0.996	0.998

→ ResNet-18, ViT-B/16 and YOLOv8-cls are selected for further experiments

3. Classification Performance of the models in binary vs multi-class configurations

I. Multi-Class Classification Results

Model	Class	Precision	Recall	F1-score	Accuracy
ViT-B/16	Healthy	0.97	0.98	0.98	
	Melanoma	0.96	0.94	0.95	0.97
	Other	0.97	0.97	0.97	
YOLOv8-cls	Healthy	0.98	0.99	0.98	
	Melanoma	0.96	0.95	0.96	0.98
	Other	0.99	0.97	0.98	
ResNet-18	Healthy	0.91	0.97	0.94	
	Melanoma	0.83	0.71	0.76	0.90
	Other	0.91	0.90	0.91	

- good performance in both settings
- accuracy remained stable
- ViT-B/16 and YOLOv8-cls outperformed ResNet-18
- melanoma recall** (how many true cases the model detects) decreased in the binary setting
→ **multi-class classification** was selected

II. Binary Classification Results

Model	Class	Precision	Recall	F1-Score	Accuracy
ViT-B/16	melanoma	0.97	0.92	0.95	0.98
	non melanoma	0.98	0.99	0.99	
YOLOv8-cls	melanoma	0.97	0.93	0.95	0.98
	non melanoma	0.98	0.99	0.99	
ResNet-18	melanoma	0.88	0.36	0.52	0.88
	non melanoma	0.88	0.99	0.93	

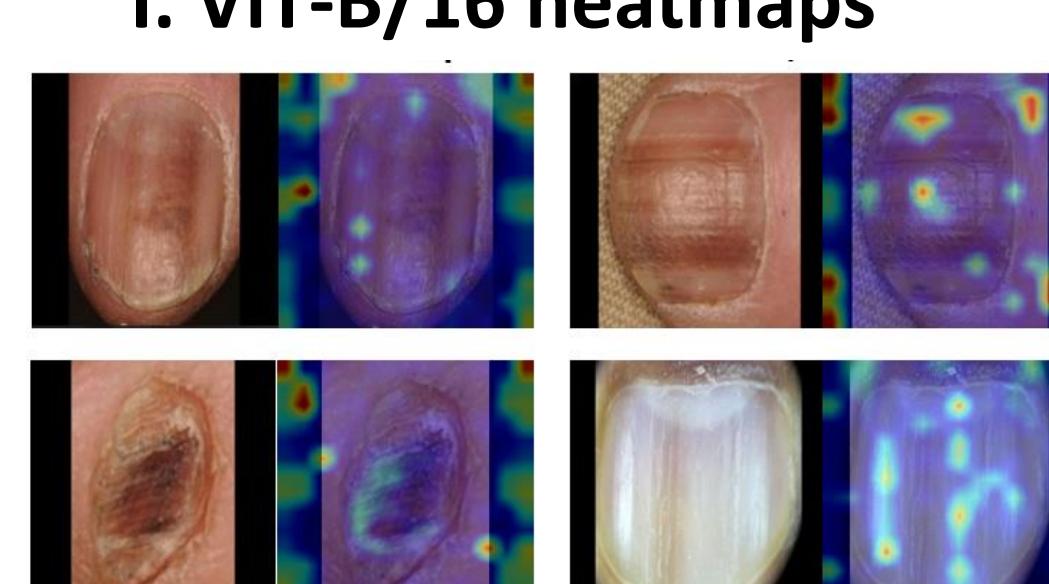
4. Fine-Tuning Results of the selected models

Model	F1-Score (healthy)	F1-Score (melanoma)	F1-Score (other)
ViT-B/16	0.98	0.97	0.98
YOLOv8-cls	0.99	0.96	0.98

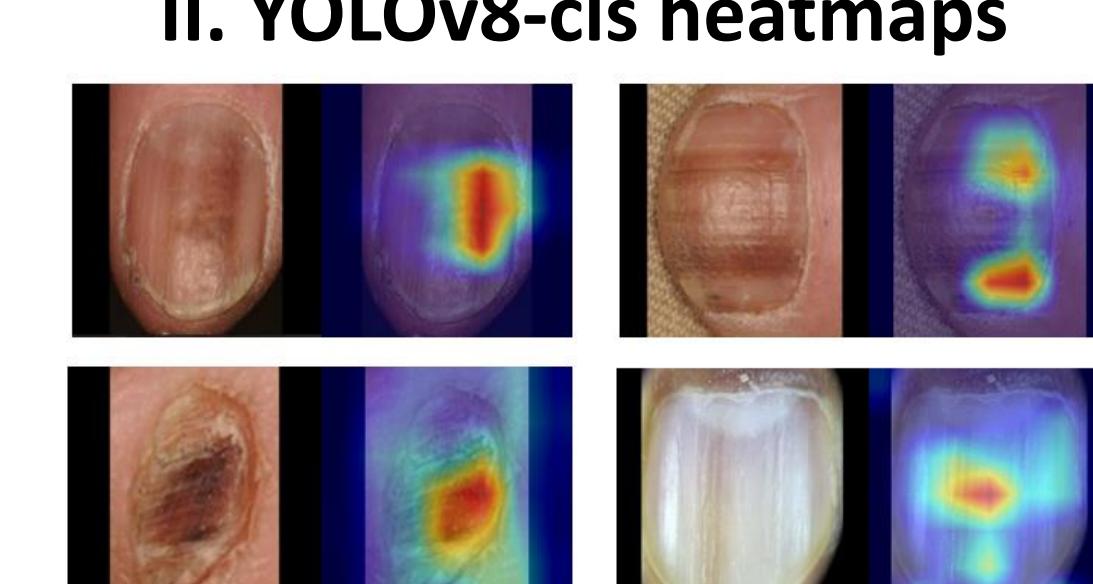
- ViT-B/16's accuracy: from 97% to **98%**
- YOLOv8-cls's accuracy: from 98,2% to **98,6%**

5. Explainable AI: Heatmaps

I. ViT-B/16 heatmaps



II. YOLOv8-cls heatmaps



- diffuse & distributed heatmaps across the nail, highlighting multiple areas
- localized patterns, focusing on salient regions (e.g., pigmentation, lesions, irregular structures)

Conclusion

- The models achieved a high accuracy
- Multi-Class classification provided better results
- Heatmaps (XAI) confirmed focus on clinically relevant regions

→ Deep Learning techniques have strong potential to support earlier NAM detection in clinical practice