Alma Mater Studiorum · Universitá di Bologna

Department of Computer Science and Engineering
Master Degree in Artificial Intelligence

SkinScan - Recognition of pigmented skin lesions with Vision Transformers and Bayesian Networks

Applicant
Marco Acerbis

Supervisor
Prof. Samuele Salti

Introduction

- In recent years, computer vision techniques based on Deep Learning have found extensive application in medical diagnosis;
- The ongoing challenge in modern medicine remains the effective detection of cancer;
- SkinScan aims to provide reliable, efficient, and cost-effective tools to assist physicians.

Introduction

SkinScan harnesses the capabilities of two modules:

- Image Classifier;
- Bayesian Network.

Image Classifier

Models

While CNNs proved to be an excellent tool to help physician with image classification, in recent years Vision Transformers have matched the performance of classical convolution approaches in almost every task.
Given the aims of this project, we evaluated three different neural network architectures:

- Swin-Transformer;
- ConvNeXt;
- Swiftformer.

Image Classifier

Swin-Transformer

Swin-Transformer introduced a series of innovations:

- **Hierarchical Feature Maps** to handle objects at different scales;

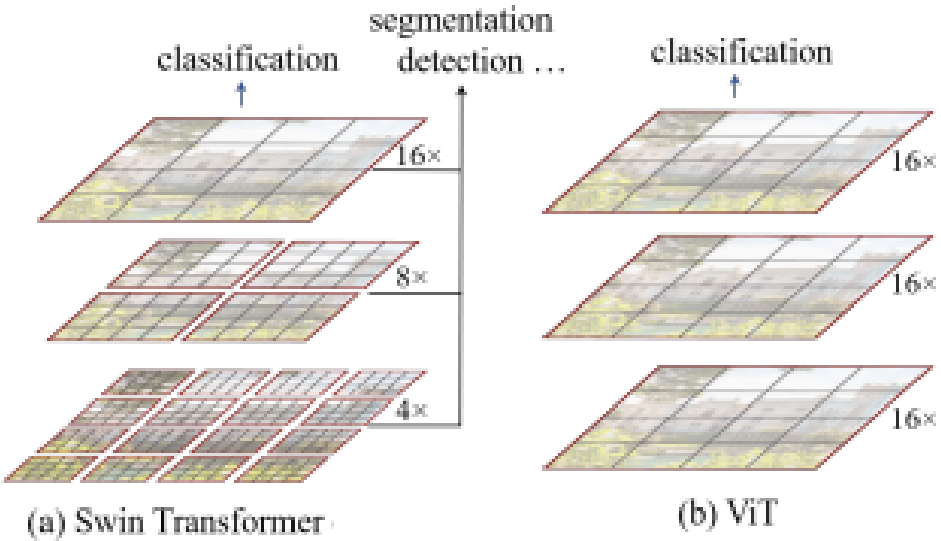


Image Classifier

Swin-Transformer

Swin-Transformer introduced a series of innovations:

- **Hierarchical Feature Maps** to handle objects at different scales;
- **Shifted Window** approach for local computation of self-attention on non-overlapping windows.

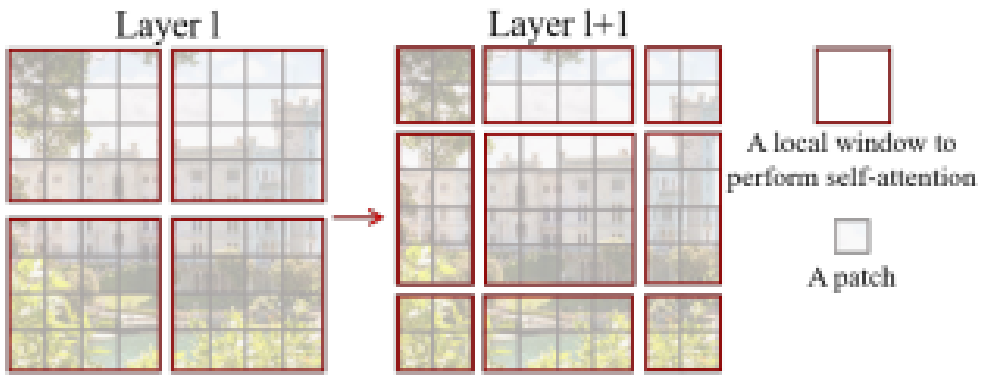


Image Classifier

Swin-Transformer

Swin-Transformer introduced a series of innovations:

- **Hierarchical Feature Maps** to handle objects at different scales;
- **Shifted Window** approach for local computation of self-attention on non-overlapping windows.

$$\Omega(\text{MSA}) = 4hwC^2 + 2(hw)^2C, \tag{1}$$

$$\Omega(\text{W-MSA}) = 4whC^2 + 2M^2hwC. \tag{2}$$

Image Classifier

ConvNeXt

ConvNeXt wants to refresh the ResNet architecture with changes **inspired** by the dominant Transformers:

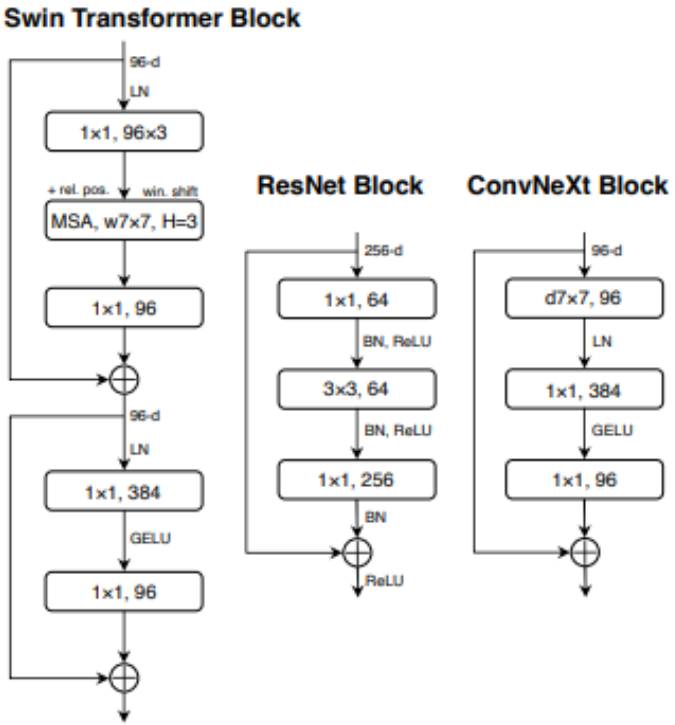


Image Classifier

Swiftformer

Since transformers are limited by the quadratic complexity of self-attention, SwiftFormer aims to introduce an **efficient additive attention mechanism**:

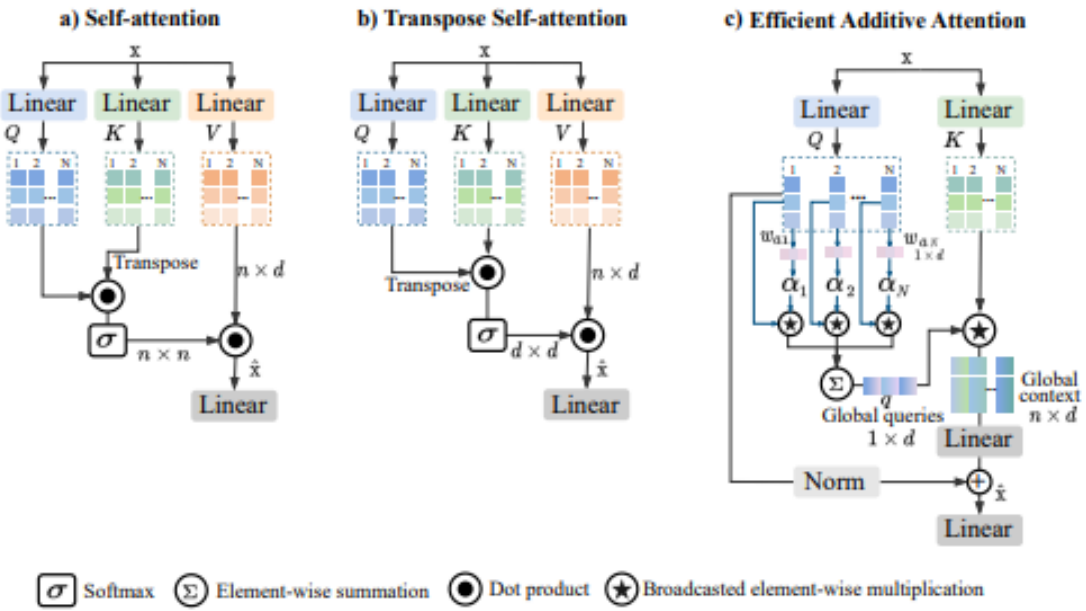


Image Classifier

Explainable AI

- A big obstacle that AI applications have to face, especially in the medical field, is that they are usually seen as a black-box;
- Many physicians remain sceptical and prefer to not follow or even do the opposite of what an algorithm suggests;
- At the same time, some of these biases find some reflection in reality;
- We decided to introduce two tools dedicated to the explanation of the provided classification.

Image Classifier

Grad-CAM

Gradient-weighted Class Activation Mapping (Grad-CAM) is used to create localization map that underlines the most relevant pixels used by a CNN to classify an image.

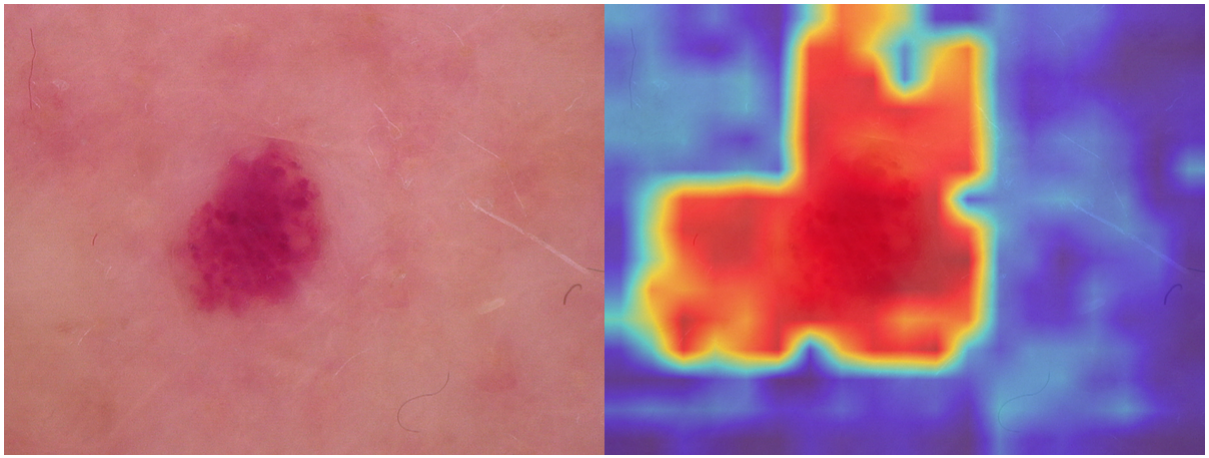


Image Classifier

Grad-CAM

Gradient-weighted Class Activation Mapping (Grad-CAM) is used to create localization map that underlines the most relevant pixels used by a CNN to classify an image.

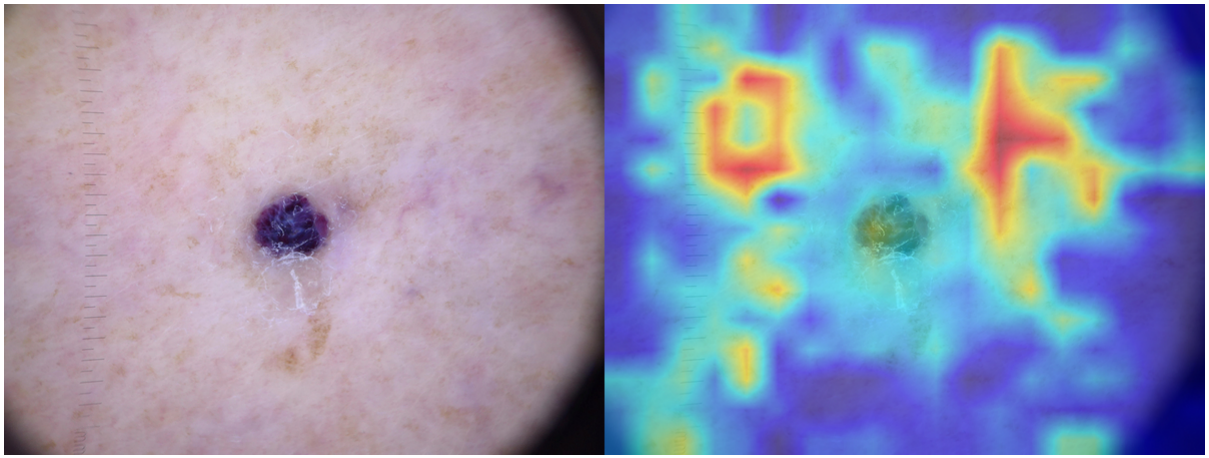
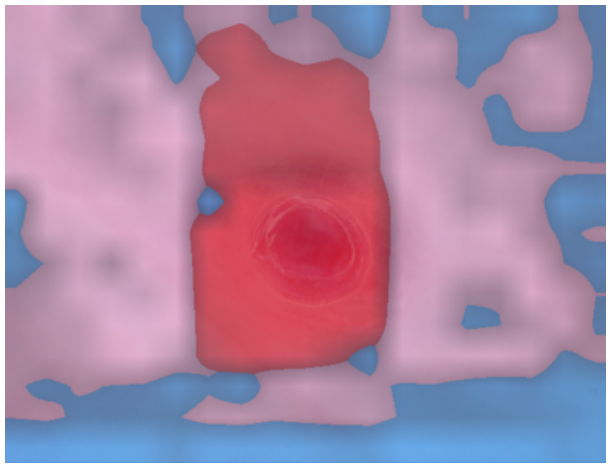


Image Classifier

Deep Feature Factorization

Deep Feature Factorization (DFF) answers the question *'what does my image classifier see in this picture?'*

Predicted Class	Score
VASC (6)	0.515
NV (5)	0.478
MEL (4)	0.004
BKL (2)	0.001
BCC (1)	9e-4
DF (3)	5.3e-5
AKIEC (0)	1.4e-9



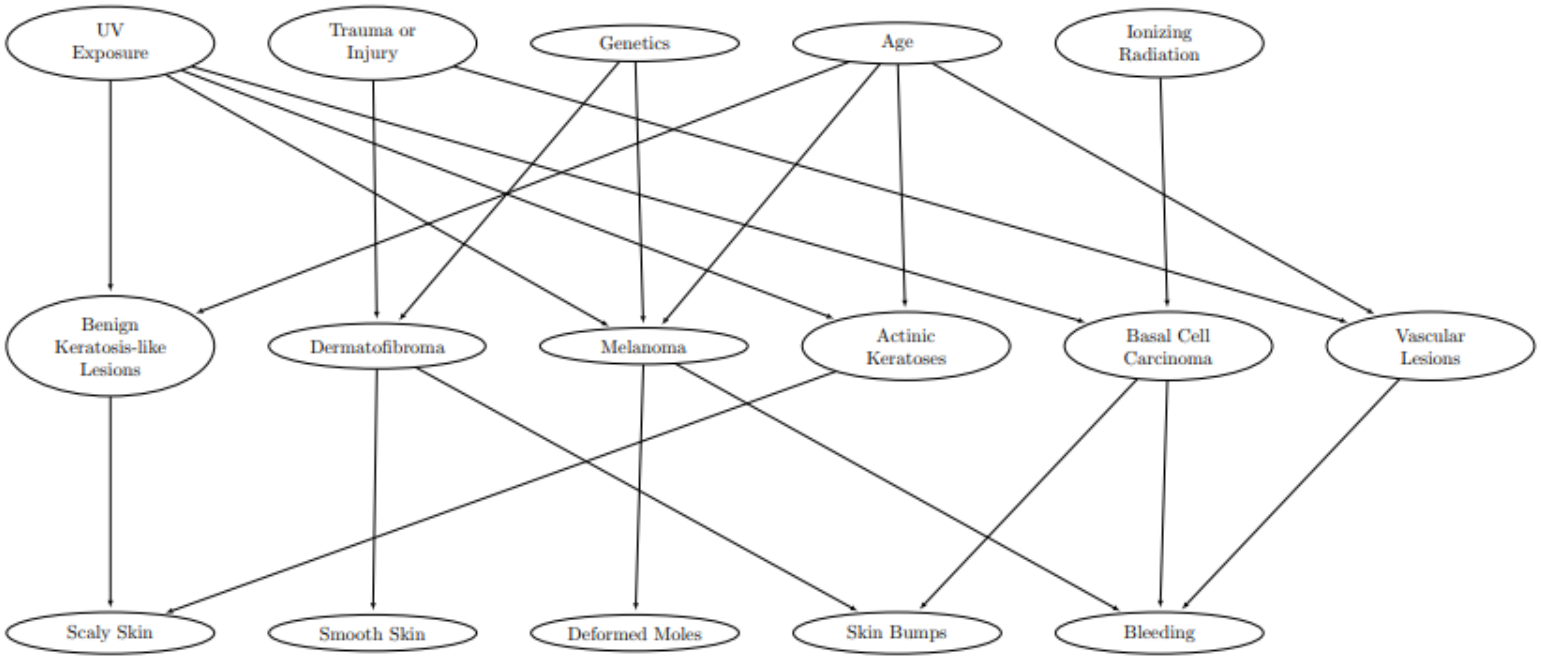
Bayesian Networks

Proposed Model

Bayesian Networks are an example of probabilistic graphical models employed for different tasks in the medical field. In particular, they provide probabilities given the correlation between:

- Risk Factors;
- Diseases;
- Signs.

Bayesian Networks



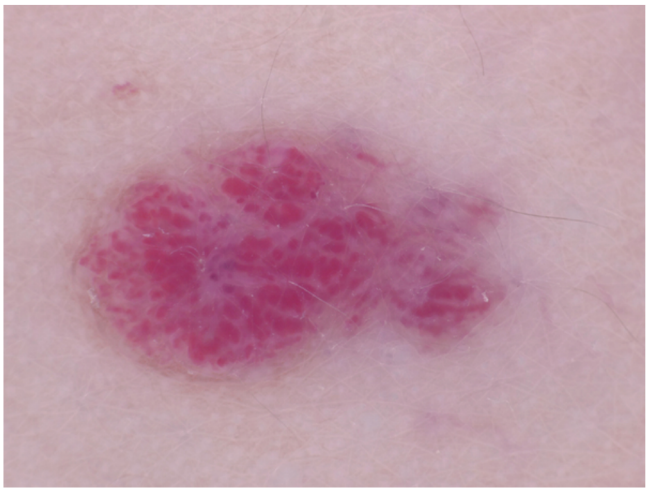
Web App

SkinScan is designed to serve as a readily accessible tool, so the optimal solution is to deploy it through a web app

- **Inference Service**: makes the trained model available on a Kubernets cluster thanks to the Kserve tool;
- **Back-End**: provides the backbone of our system;
- **Fron-End**: the interface that communicates with the final user.

Web App

User Interface



CARICA

Skin Scan

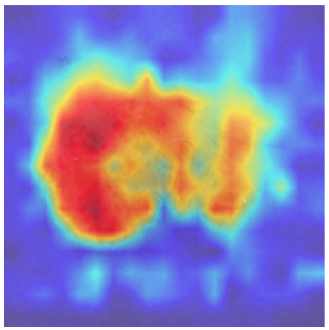
Quanti anni ha il paziente?	UNDER 14	14 - 65	OVER 65
Il paziente si è esposto a raggi solari?	SI	NO	NON LO SO
La zona ha subito traumi?	SI	NO	NON LO SO
Il paziente ha una predisposizione genetica?	SI	NO	NON LO SO
Il paziente si è esposto a radiazioni ionizzanti?	SI	NO	NON LO SO
Il paziente ha notato una desquamazione della pelle?	SI	NO	NON LO SO
Il paziente ha notato la pelle lucida/liscia?	SI	NO	NON LO SO
Il paziente ha notato delle protuberanze cutanee?	SI	NO	NON LO SO
Il paziente ha notato sanguinamenti?	SI	NO	NON LO SO

INVIA REIMPOSTA

Web App

User Interface

Immagine spiegata



Rete Neurale

Disease	Likelihood
vascular lesions	0.9092
melanocytic nevi	0.0512
melanoma	0.0261
basal cell carcinoma	0.0103
benign keratosis-like lesions	0.0021
dermatofibroma	0.0010
Bowen's disease	8.3800e-05

Rete Bayesiana

Disease	Likelihood
vascular lesions	0.0011
Bowen's disease	0.1154
basal cell carcinoma	0.0010
benign keratosis-like lesions	0.2219
melanoma	0.0002
dermatofibroma	0.0294

Conclusion

In this thesis, we focused on realizing an efficient and reliable tool to help practitioners and dermatologists in recognizing pigmented skin lesions. As result, we developed the SkinScan system, that implements both Vision Transformers and Bayesian Networks, and deployed it through a Web App.