SkinSmart Diagnostics

StackFuel Portfolio Course
Aug. 2024

Friday, 23 August 2024

Motivation

What are we trying to solve?

Problem

Manual visual examination of skin malformations is subjective and error-prone.

Frequency

Skin cancer is the most common cancer in Germany, with over 270,000 new cases annually, including more than 30,000 cases of malignant melanoma, the deadliest form.

Importance

~3,000 people die from skin cancer each year in Germany, primarily due to malignant melanoma.

Early detection of skin cancer is critical for better patient outcomes. This tool aids in identifying potentially cancerous lesions early, thus improving the accuracy of initial diagnoses at the GP level.

Project Vision

Diagnosis of **Skin Cancer** using **Image Recognition Neural Networks**

Data

10,000 images with metadata from a Harvard study. Source

Goals

Primary: Classify skin lesions into **7** possible diagnosis, some cancerous.

Secondary: Enable selection and training of different AI models within the application.

Features

- Graphic User Interface (**GUI**):
  - model selection, training, and diagnosis.
- Store and manage skin image data and metadata in a **Diagnosis**.
- Support for running the application on Linux (Debian).

Product Users

General Practitioners / Dermatologists:

- **Select a case involving a skin malformation:** Users can upload images of skin lesions and access patient metadata to assess the case.
- **Receive a diagnosis with associated probabilities:** The system provides a ranked list of possible diagnoses, each with a probability score, helping clinicians determine the likelihood of each condition.
- **Decision Support:** Based on the diagnosis, the tool offers guidance on whether to refer the patient to a specialist, potentially reducing unnecessary referrals.

Data Scientists / Specialized Dermatologists:

- **Serve as admin users to train new AI models:** Admins can manage and preprocess new datasets, fine-tune existing models, and experiment with different machine learning algorithms to improve accuracy.
- **Oversee the updating and improving of diagnostic models:** Admins evaluate model performance using validation metrics and update the models as new data becomes available, ensuring the system evolves with medical advancements.
- **Model Validation and Testing:** Specialized users validate model predictions against known outcomes and ensure compliance with medical standards before deploying updates.

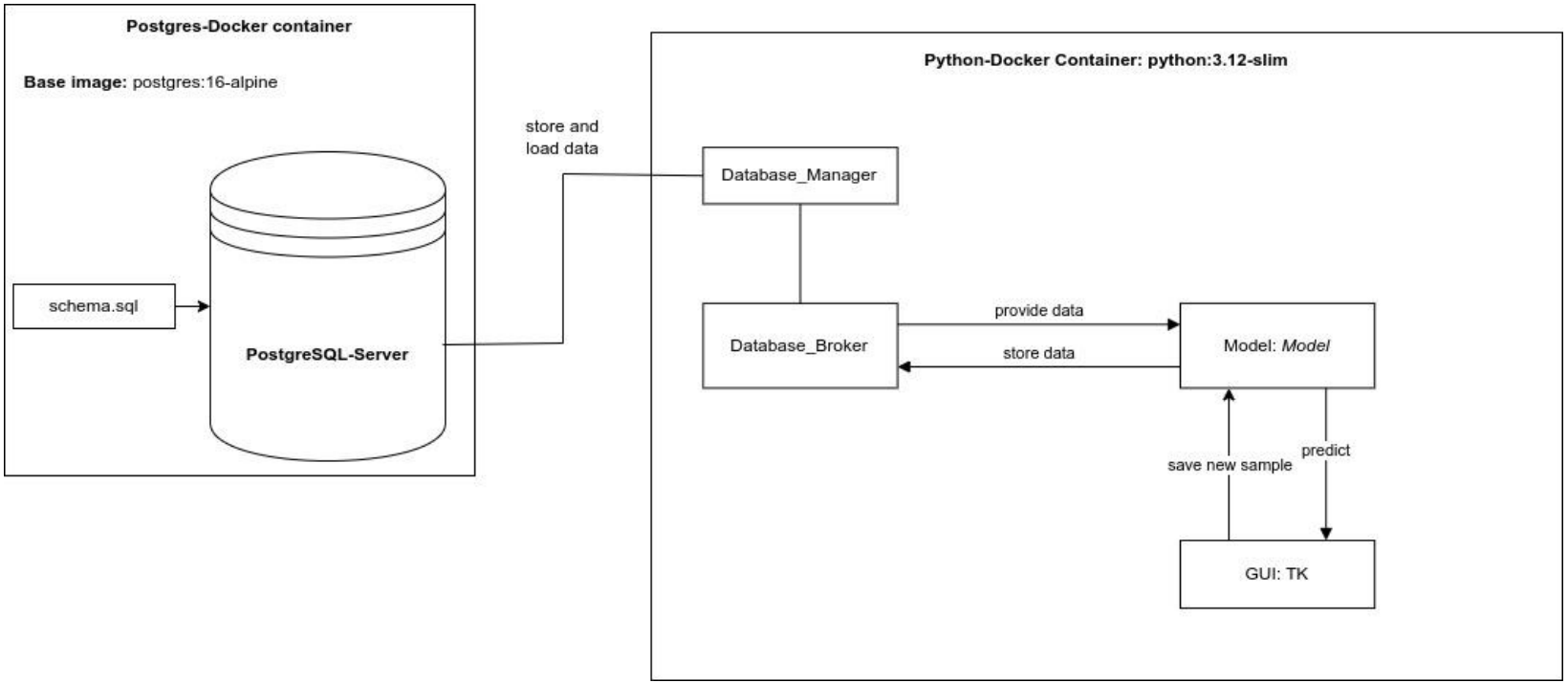
Our Team

- **Marcel Sonnenschein (MSc. Physics)**
  - **Roles:** Scrum Master, Infrastructure Manager
  - **Tasks:** Developing Project Infrastructure, Developing Database

- **Mahtab Lashgari (MSc. Neurobiology)**
  - **Roles:** Data Scientist, Frontend developer
  - **Tasks:** EDA, Machine Learning on Metadata, GUI development

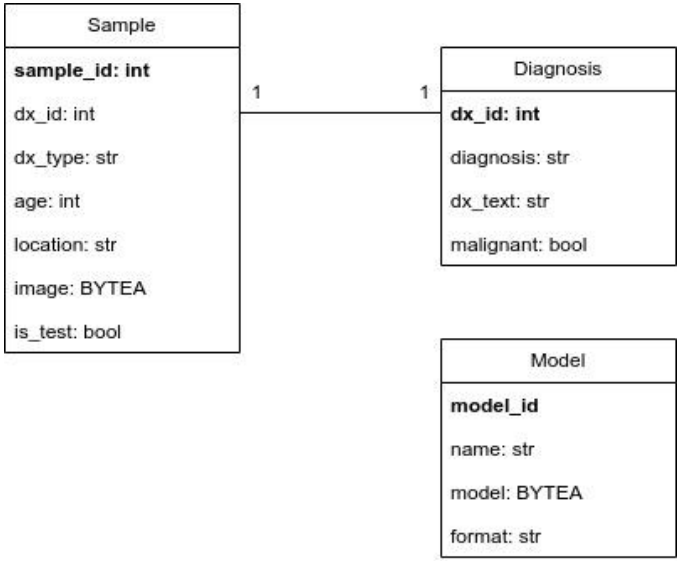
- **Francisco J. Arriaza G. (PhD. Biochemistry)**
  - **Roles:** Quality Control, Deep Learning Expert
  - **Tasks:** Develop Image Classification Models, Final Ensemble

Software structure



Database schema

- Simple schema with 3 Tables
- No normalization adjusted for performance reasons
- Sanity check outsourced to handler class (Database_Broker)

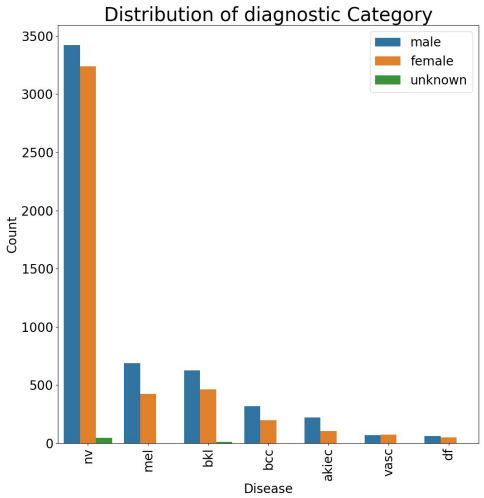


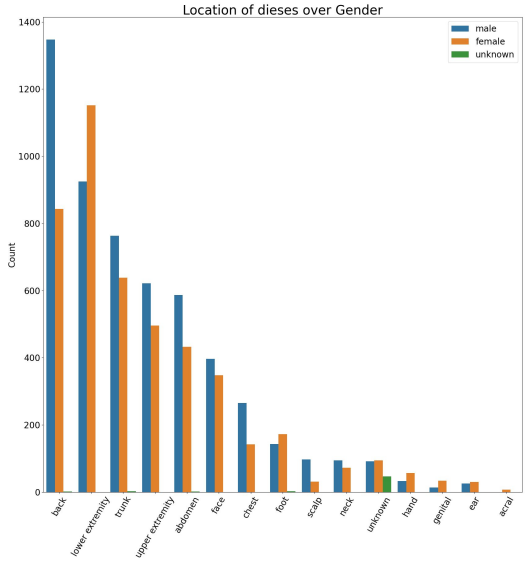
EDA of Our Dataset

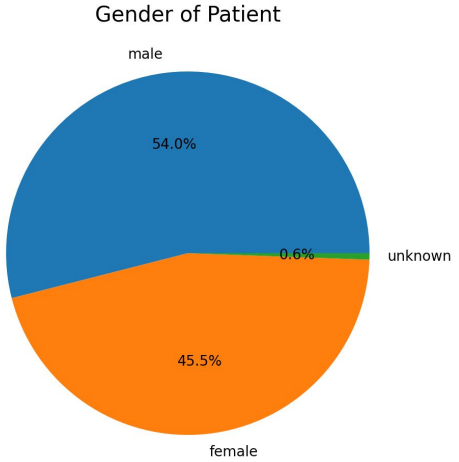
Metadata: contains 10015 dermatoscopic images, csv file with demographic information of each lesion

More than 50% of lesion are confirmed through histopathology

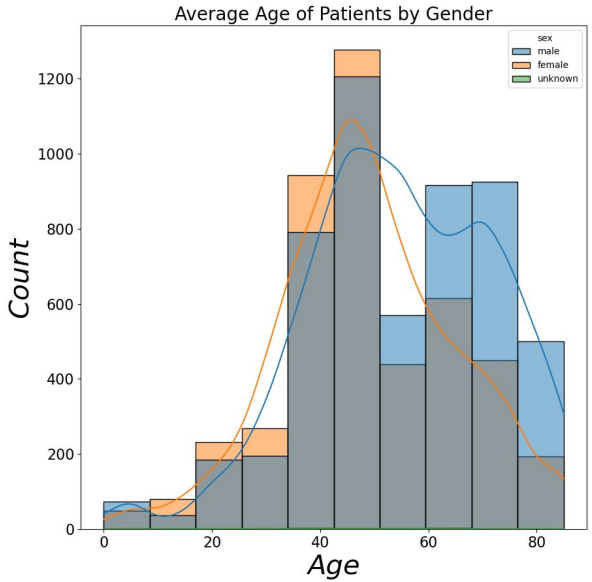
Visualize of Frequency distribution of different classes (diseases) and visualizes the gender distribution of patients

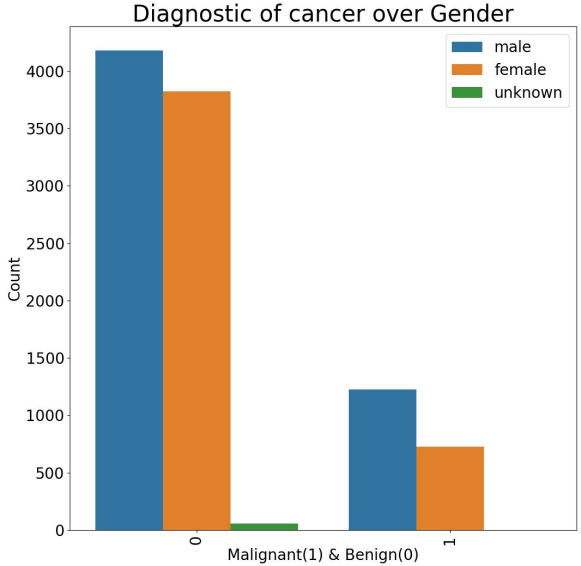




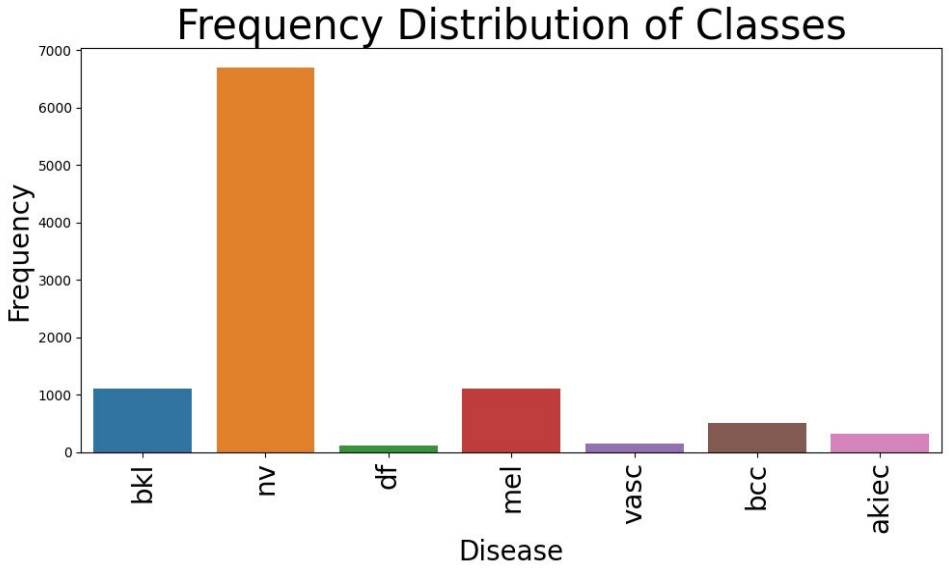


Our Data


Average Age of Patients by Gender
sex
male
female
unknown
Count
Age
0
200
400
600
800
1000
1200
0
20
40
60
80


Diagnostic of cancer over Gender
male
female
unknown
Count
Malignant(1) & Benign(0)
0
500
1000
1500
2000
2500
3000
3500
4000
0
1

Database balance



- Nevus (nv) is the most frequent class (~65% of datapoints)
- Unbalanced dataset that corresponds to clinical observations
- One effective way : SMOTE (Synthetic Minority Over-sampling Technique)
- that helps create more diverse and realistic examples for the minority class. Significantly improve the class balance in our dataset

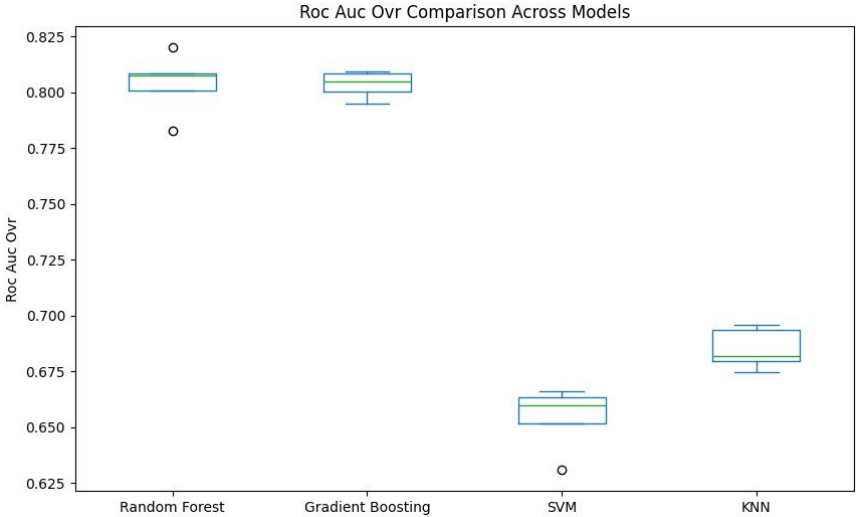
ML workflow

- Import Packages
- Data-cleaning
- Exploratory Data Analysis
- Data Visualization
- Class Imbalance Handling (Oversampling with Methods like SMOTE)

- Train-Test-Split (test-size=20%)
- Model Selection and Import
- Metric Definition
- Performance Analysis
- Selection of the Best Model

- Validation
- Performance metrics
- Visualization
- Comparison between Predictions and Actual Outcomes

Performance Analysis and Comparison



The Random Forest and Gradient Boosting models stand out as the top performers based on this analysis.
Random Forest shows a balanced performance with high ROC-AUC score, highest accuracy and F1-score

Modell	Genauigkeit	F1-Score	ROC-AUC-Score
RF	0.705941	0.641709	0.828809
Gradient Boosting	0.703944	0.636592	0.833333
SVM	0.667998	0.535038	0.652275
KNN	0.677983	0.629808	0.688642

Image classification

a. Actinic keratosis
b. Basal cell carcinoma
c. Benign keratosis-like
d. Dermatofibroma
e. Malignant melanoma
f. Melanocytic nevi
g. Vascular lesions

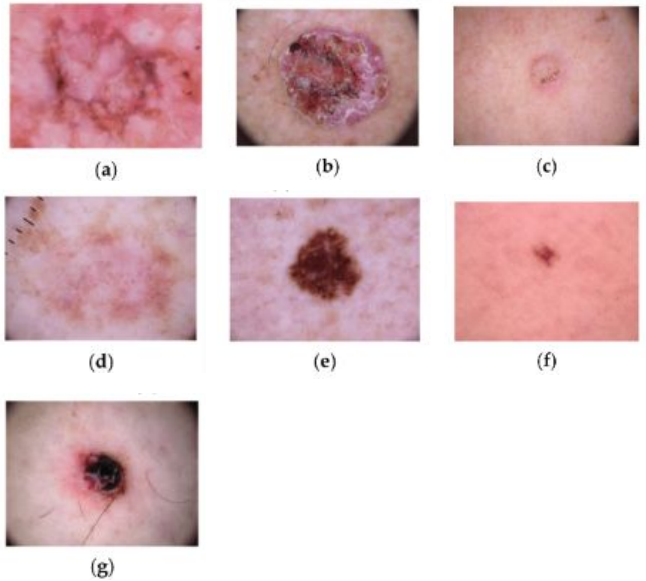
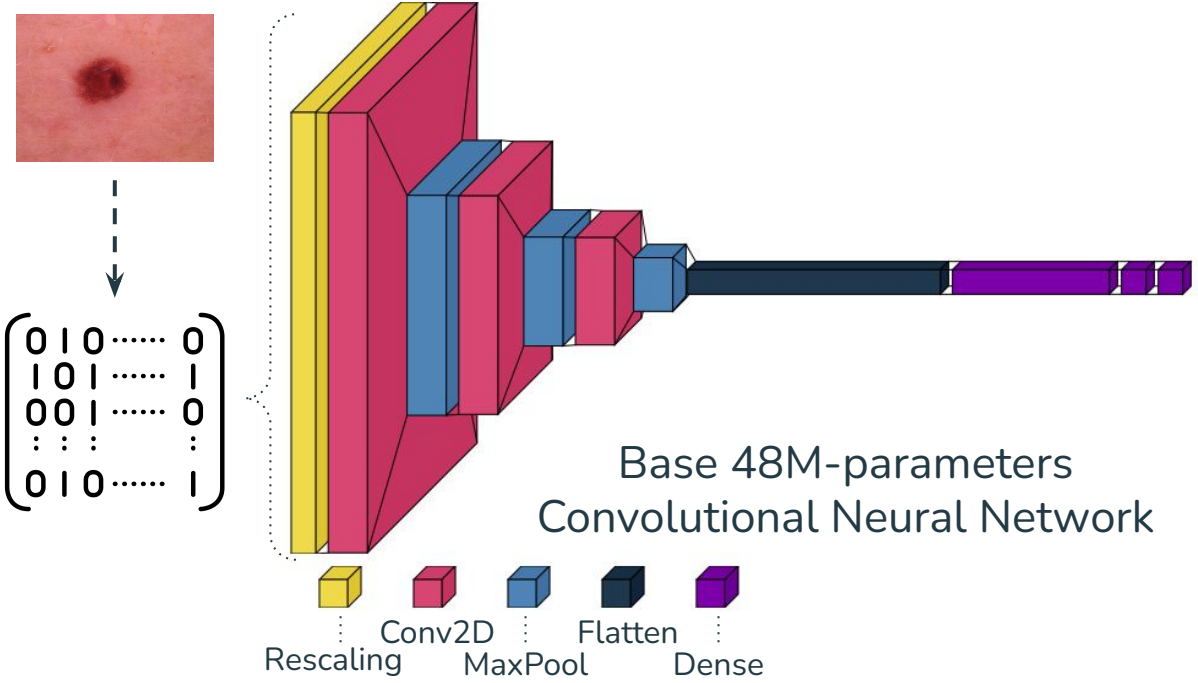


Image classification model

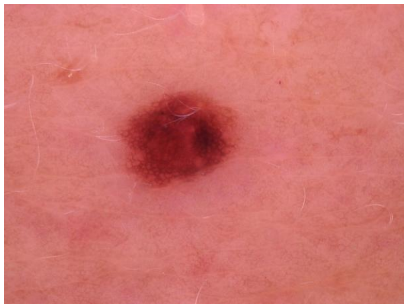


- Actinic Keratosis / Intraepithelial Carcinoma
- Basal Cell Carcinoma
- Benign Keratosis-like Lesions
- Dermatofibroma
- Melanoma
- Melanocytic Nevi (Moles)
- Vascular Lesions

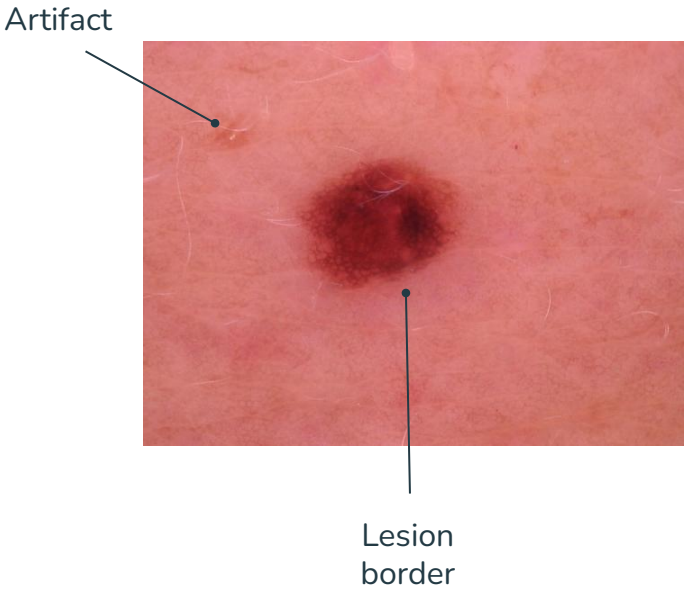
ROC AUC ~ 0.6

Feature extraction

- Images are not aligned or centered
- How to remove artifacts?
- How to enhance features?



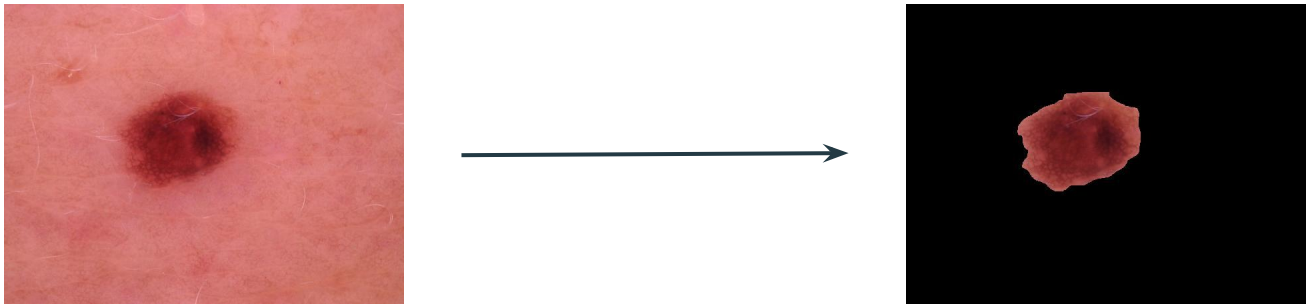
Feature extraction



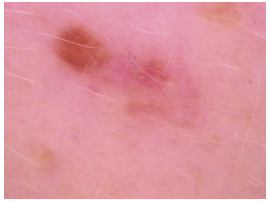
Feature extraction

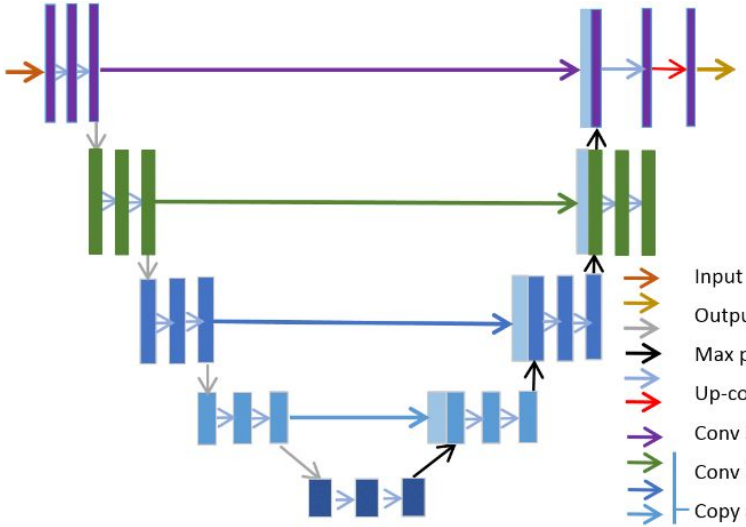
- Images are not aligned or centered
- How to remove artifacts?
- How to enhance features?

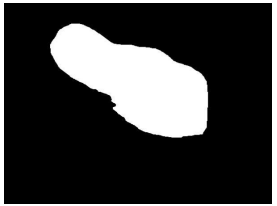
- Remove everything that is not a feature of interest



Preprocessing: Unet Neural Network



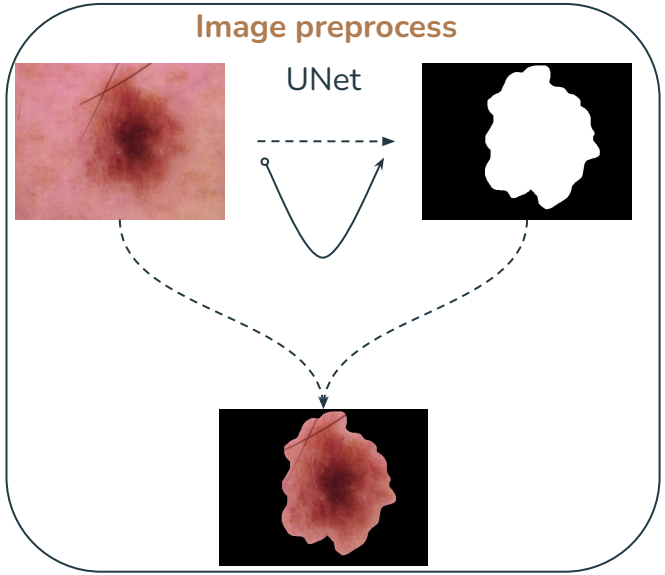




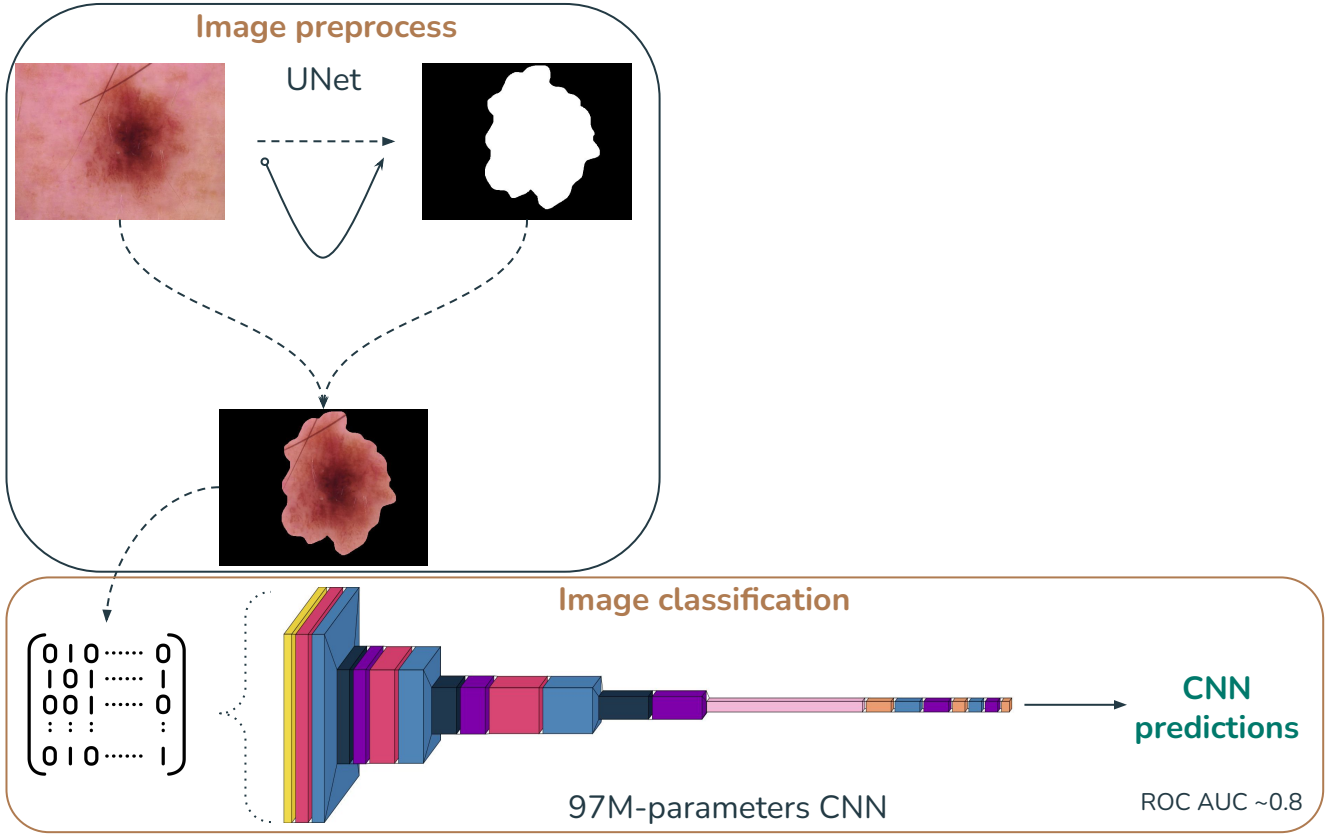
Metric used

$$\text{Dice} = \frac{2\,X\,Area\ of\ overlap}{Total\ area} = \frac{2 \times \text{Prediction} \cap \text{Ground truth}}{\text{Prediction} + \text{Ground truth}} = 99.99\%$$

Fully optimized model

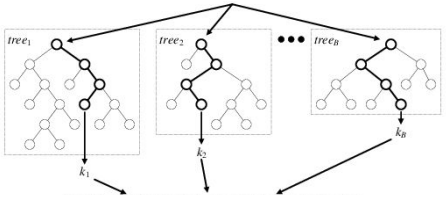


Fully optimized model



Fully optimized model

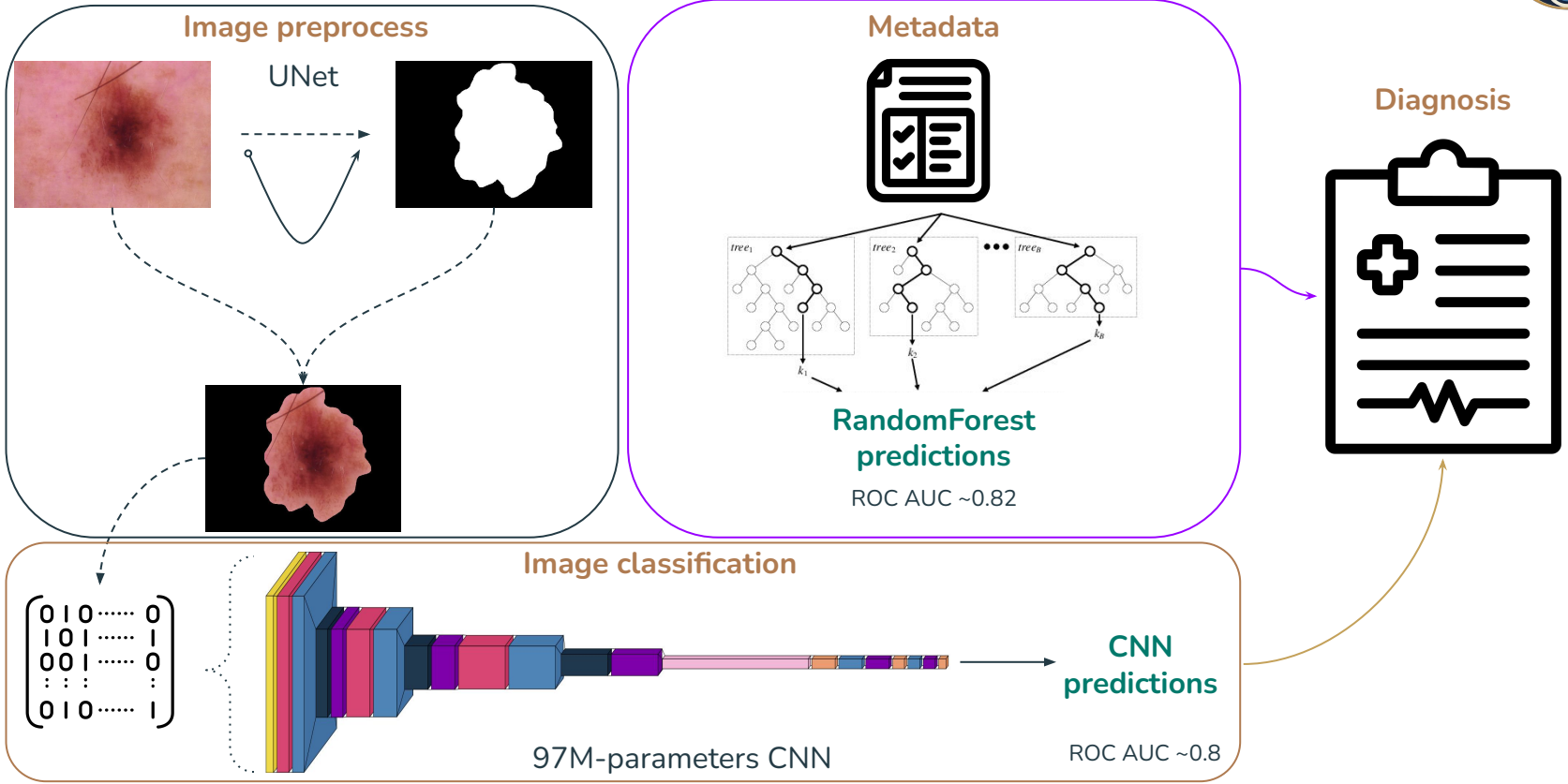
Metadata



RandomForest predictions

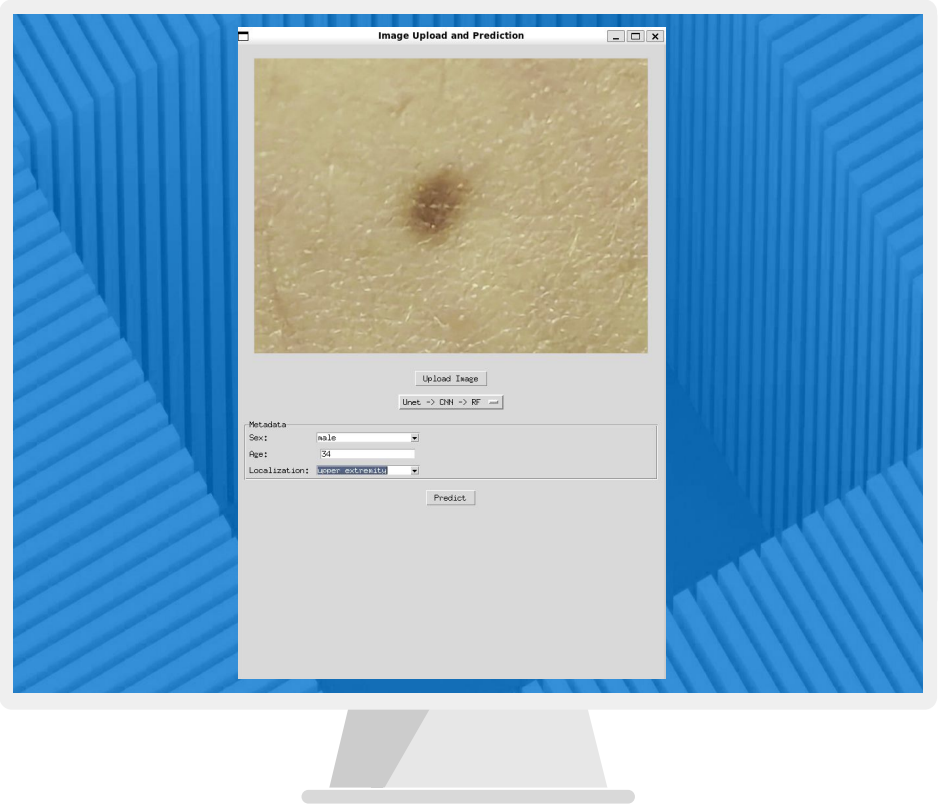
ROC AUC ~0.82

Fully optimized model



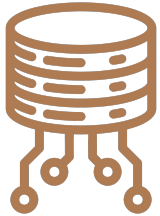
Our Solution

Demo



Outlook

Database

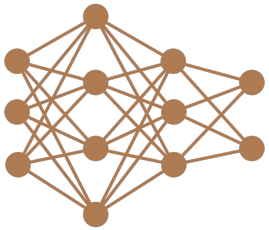


Database integration

Spark instead of psycopg2

Model training inside the app

AI



Optimize Ensemble models

Test different architectures and image augmentation strategies

Thank you!