



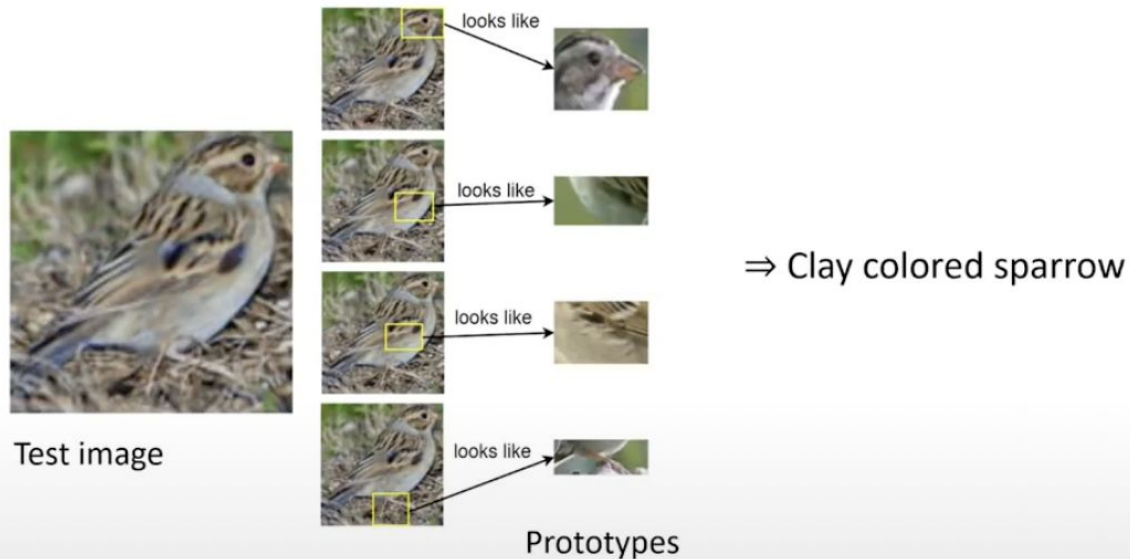
Klinikum rechts der Isar
Technische Universität München



Inherently interpretable ProtoPNets for medical imaging

Presenters: Hanad Abdullahi, Tuğcan Hoşer, Carlotta Hölzle
Supervisor: Sarah Lockfisch

ProtoPNet - General idea of prototype learning



Interpretability



A necessary equivalence for ProtoPNet is the following:

Two images patches look similar to a ProtoPNet \leftrightarrow Two image patches look similar to a human.

Structure



Until 29.07.24

Quantitative Analysis

Outline of qualitative Analysis

Research Ideas



Performance
Comparison



Impact of Prototype
Quantity



Usefulness of
Prototypes

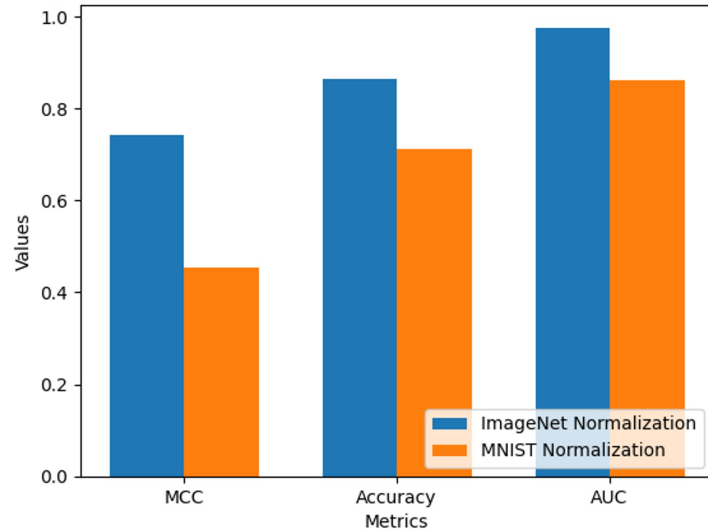
How performant are prototypical architectures in medical image classification, and does the integration of prototypical heatmaps have utility among users in Blood / Derma medical task?

Quantitative Evaluation

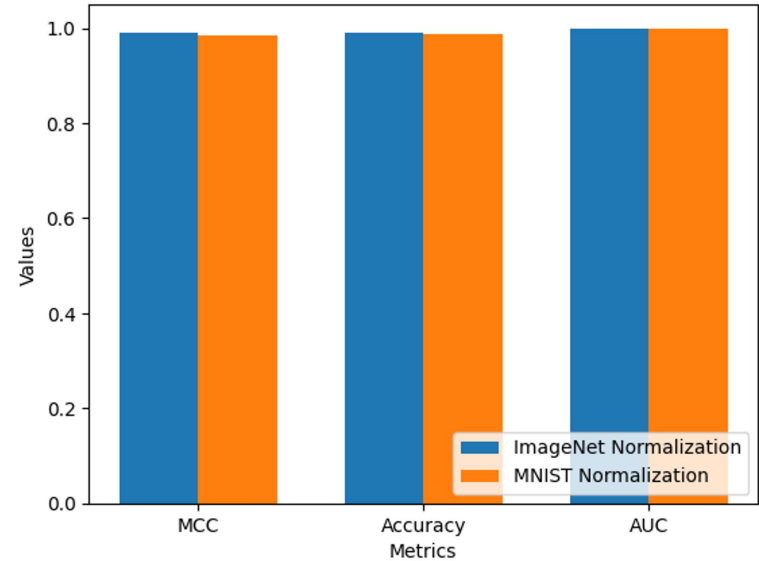
Quantitative Evaluation - normalisation



DermaMNIST



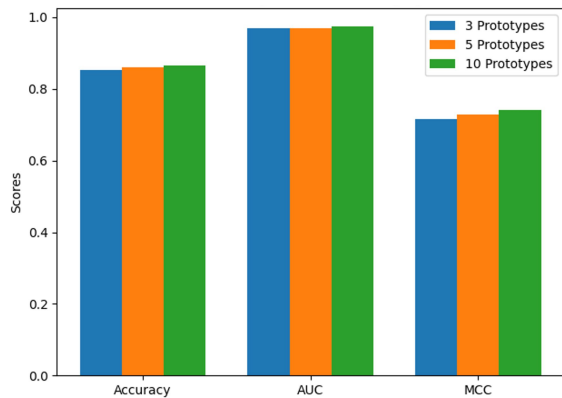
BloodMNIST



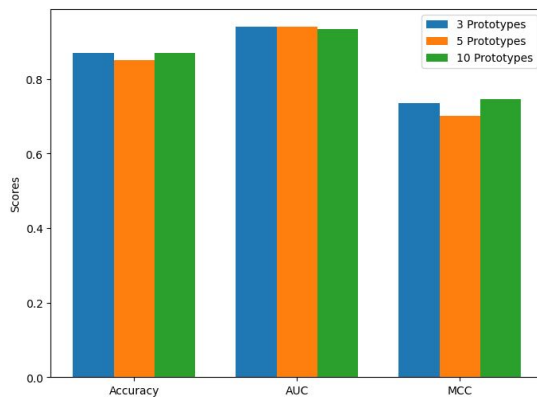
Quantitative Evaluation - number of prototypes - overall trend



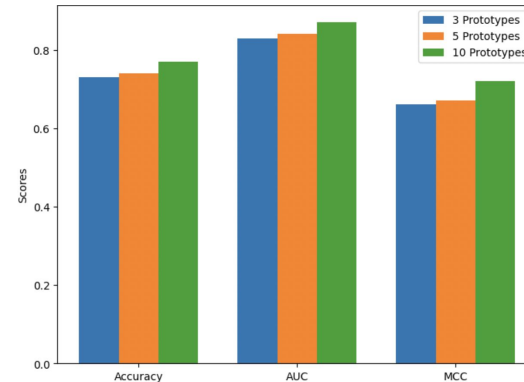
DermaMNIST



ProtoPFormer



ProtoASNet



PiPNet

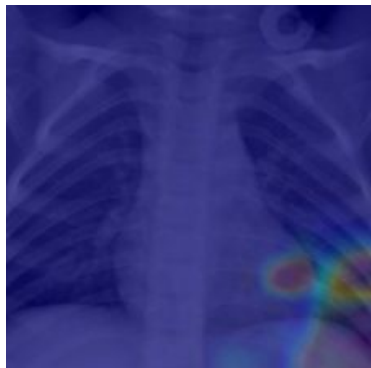
Quantitative Evaluation - number of prototypes - conclusion



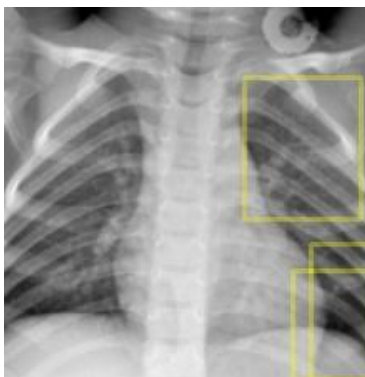
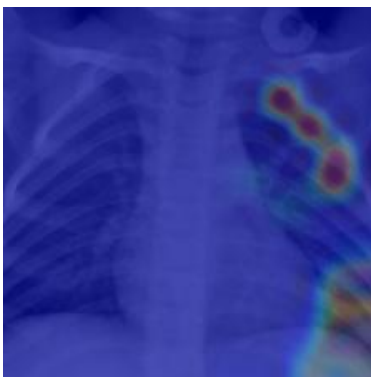
	Multi- vs. Binary	RGB vs. BW
ProtoPFormer	Binary: 3 Multi: 5 or 10	BW: 3 RGB: 5 or 10
ProtoASNet	Binary: 3 or 10 Multi: 3 or 5	BW: 3 or 5 or 10 RGB: 3 or 5
PiPNet	Binary: 3 Multi: 5 or 10	BW: 3 or 5 or 10 RGB: 5 or 10

Qualitative Evaluation

Qualitative Evaluation - Heatmaps vs. Bounding Boxes

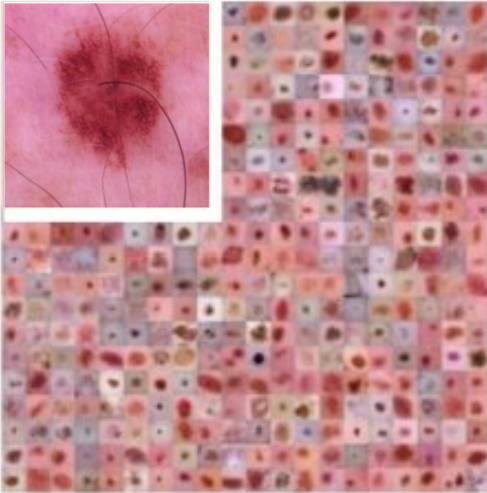


category1



category0

DermaMNIST

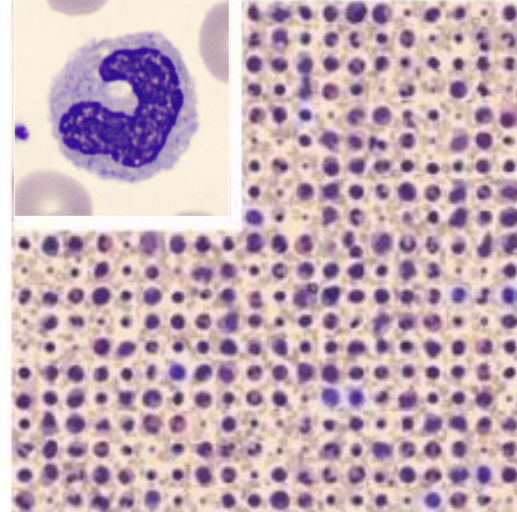


multi-class

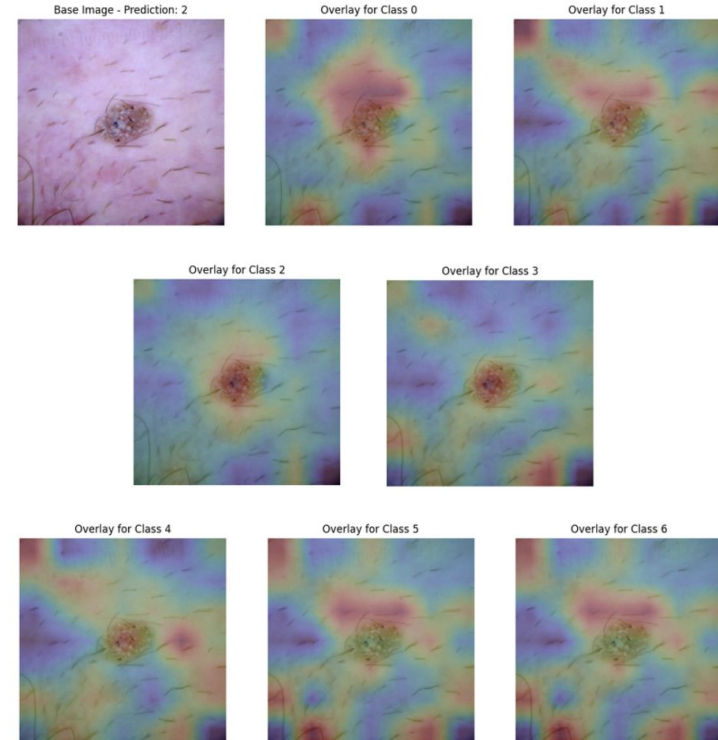
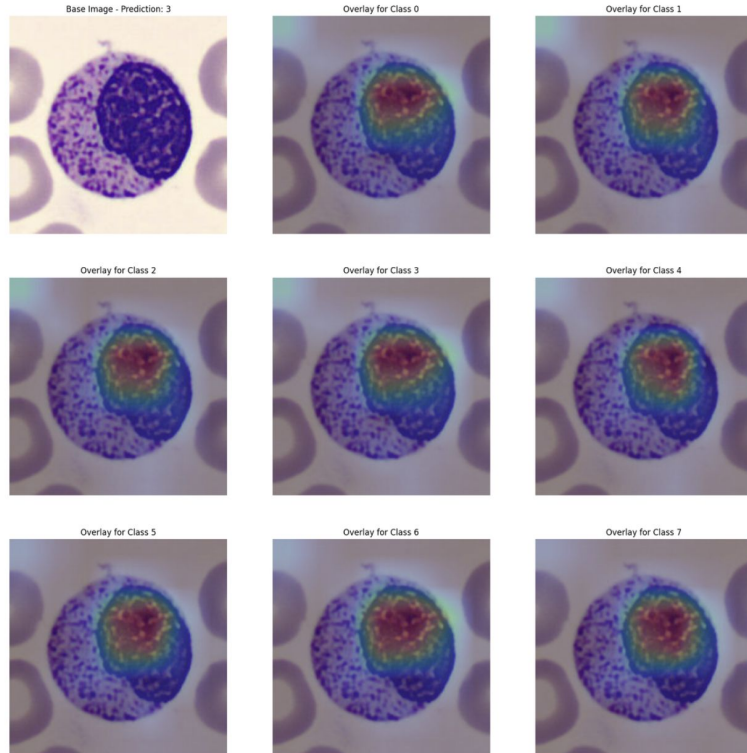
classification less
dependent on high
resolution

easy task

BloodMNIST



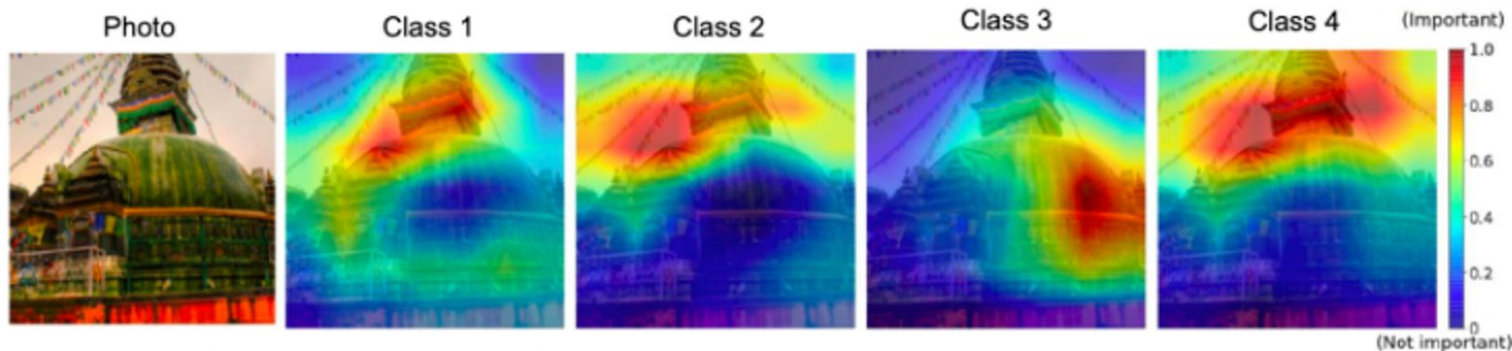
Qualitative Evaluation - dataset



UI for GradCAM distinction study

Task: Select the class you think is correct.

For each photo, we show explanations for the model's 4 predictions.



Q. Which class do you think is correct?

☐ 1 ☐ 2 ☐ 3 ☐ 4

Struggles



Visualization Prototypes

Normalization

Human Studies

heatmaps

synthesis

no boundary
boxes

no “learned”
prototypes

ProtoASNet

ProtPFormer

MedMNIST Normalisation **NOT**
available

1. newly computed
2. rerun networks



ImageNET default
normalization better
performance

1. Not many publications
2. Which datasets to use
 - a. **Derma or Blood**



Until 29.07.24

Until End

Qualitative Evaluation

Report + Code Pipeline

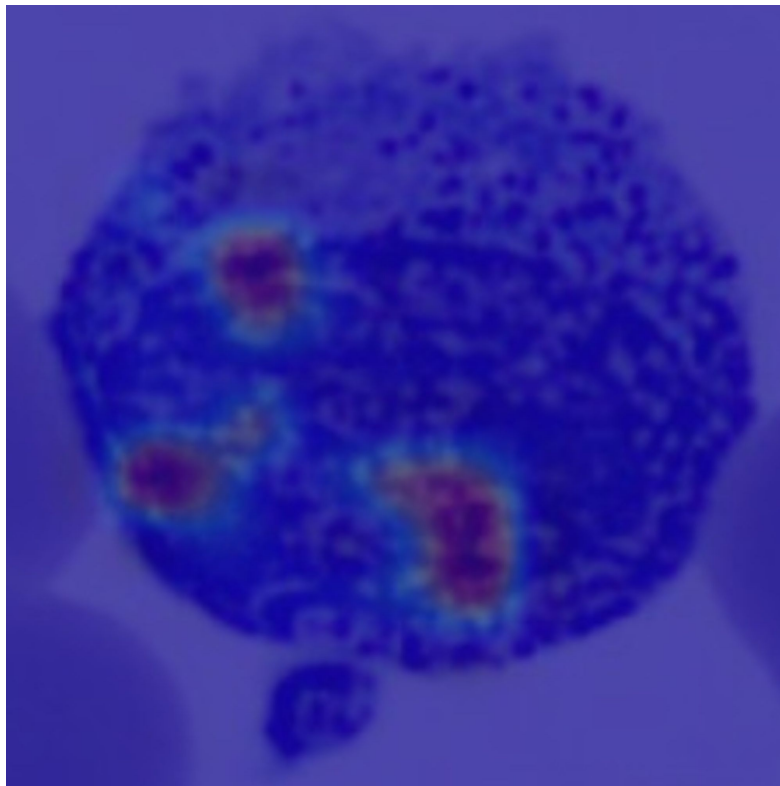
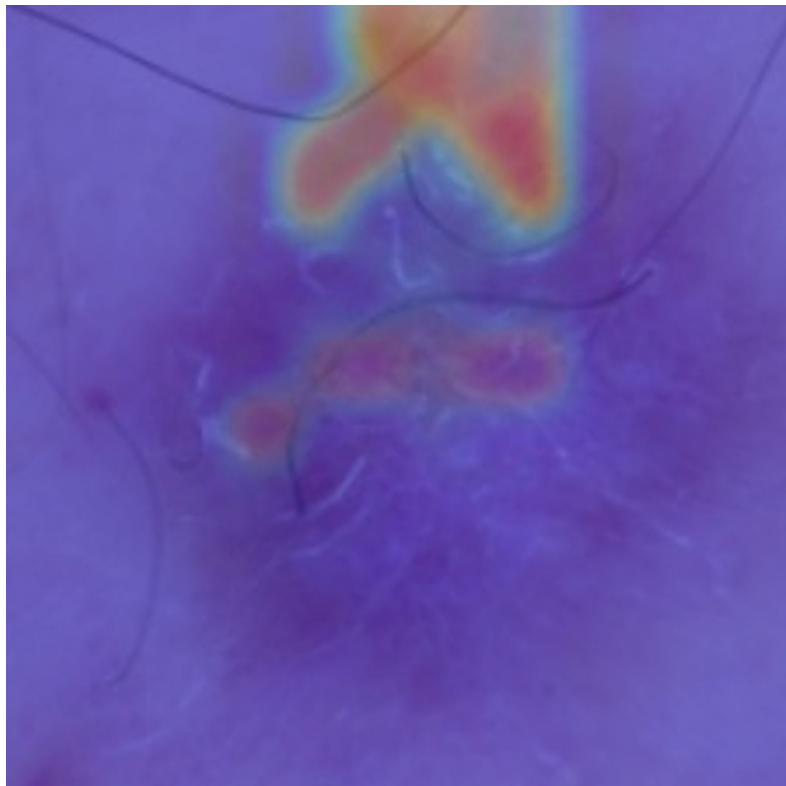


Klinikum rechts der Isar
Technische Universität München



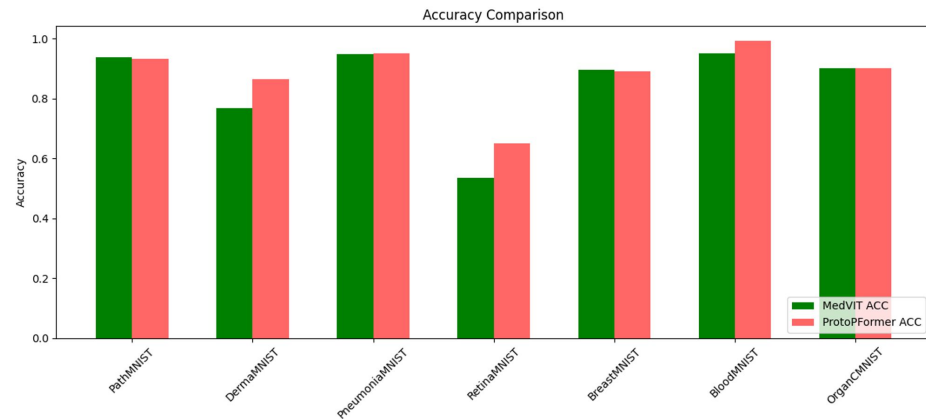
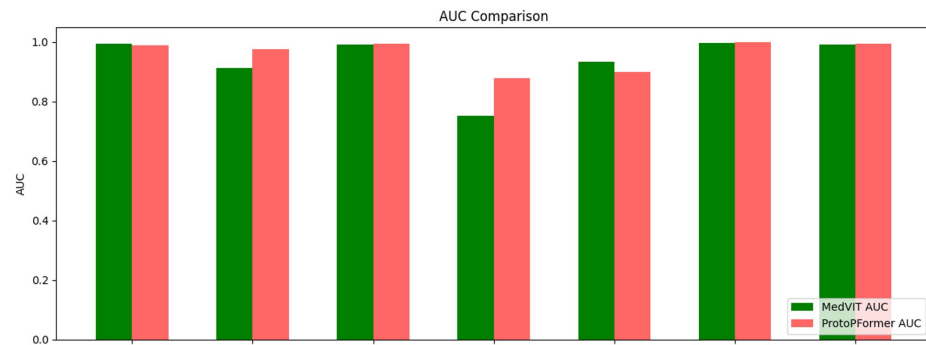
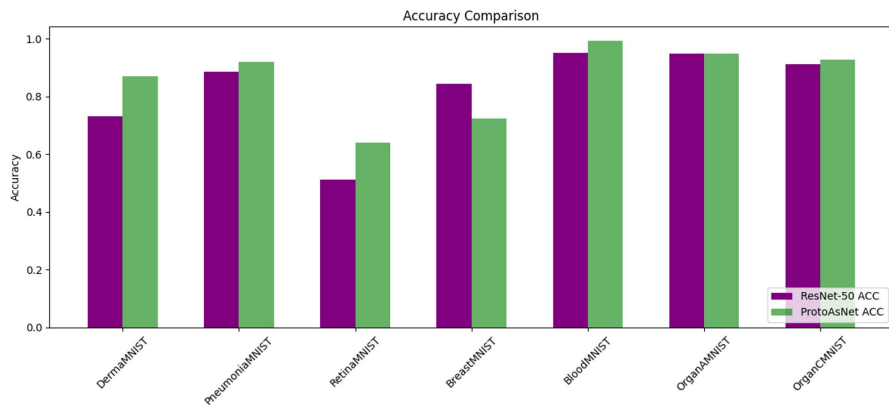
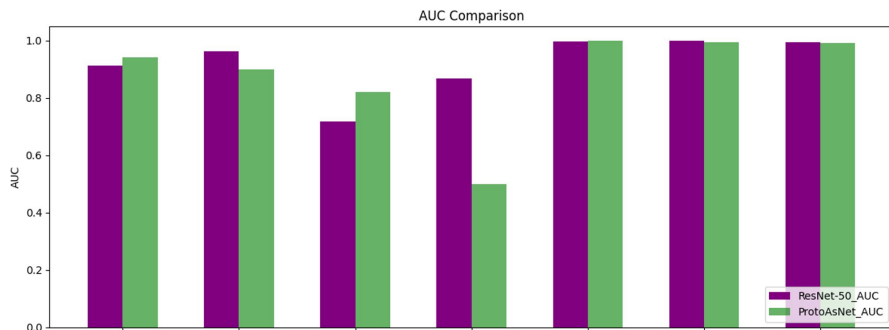
Thank you!

Qualitative Evaluation - dataset





Quantitative Comparison to state-of-the-art models



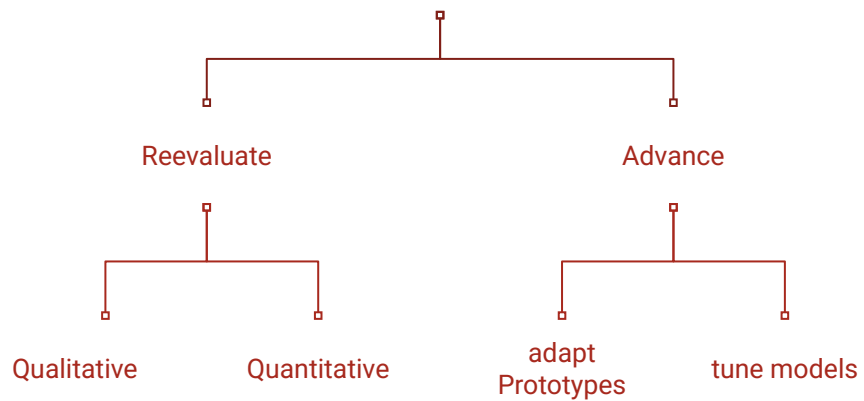
Until 01.07.24

Until 29.07.24

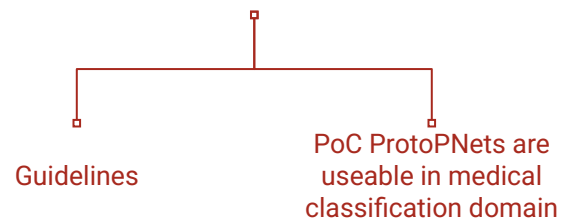
Reevaluate, Advance

Contribute

3rd Sprint



4th Sprint





How effective and interpretable are ProtoPNet architectures in medical image classification, and how can the use of prototypes be optimized to aid clinical decision-making?

1. Can ProtoPNets achieve similar quantitative performance in classification of medical images as state-of-the-art models?
2. What is the use case of prototypes in the application domain?
 - a. How do they need to be presented to be useful for downstream tasks?
 - b. What and when would doctors use prototypes (for)?
3. How does the performance change with change of number of prototypes per class
4. Contribution: given a medical dataset, give a framework with categories what the prototypes could be used for, then give recommendation which network and how many prototypes per class should be selected