PFIA 2024 - Journée Santé & IA

Interpretable AI for Dermoscopy Images of Pigmented Skin Lesions

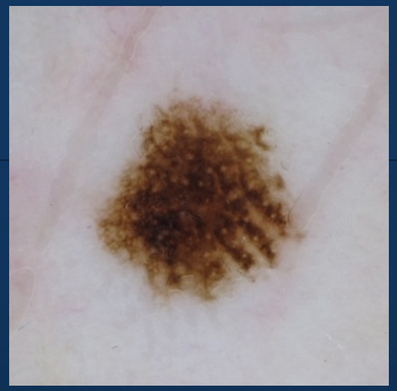
Marianne Defresne, Elise Coutier, **Paul Fricker,** Folkert Blok, Hang Nguyen

Context

Skin cancer prediction

Goal = **early** detection of skin cancer (melanoma)

Horizon Europe iToBoS (Intelligent Total Body Scanner for Early Detection of Melanoma)



- CNNs outperform dermatologists (Haenssle et al. 2018)

Skin cancer prediction

Goal = **early** detection of skin cancer (melanoma)

Horizon Europe iToBoS (Intelligent Total Body Scanner for Early Detection of Melanoma)

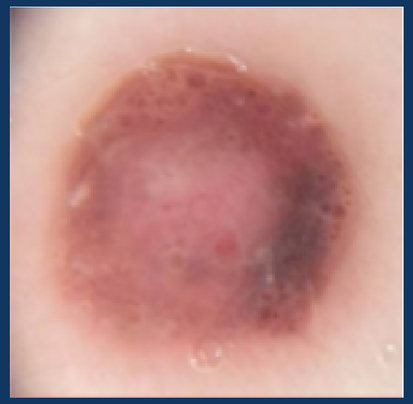
- Our classification model[1]:

- CNNs outperform dermatologists (Haenssle et al. 2018)

[1]belleProTM app

The model exploits spurious correlations

How **biased** is our dataset? Method inspired from (Bissoto et al. 2019)



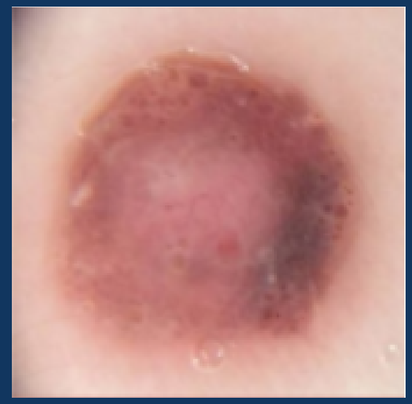
Baseline

Box model

Context

The model exploits spurious correlations

How **biased** is our dataset? Method inspired from (Bissoto et al. 2019)



Baseline

Box model

-> The model uses irrelevant information

- Bias, lack of robustness

Goal: control **what information** is used

	Baseline	Box model	Random	Doctors
Accuracy	73%	35%	17.8%	-
AUROC	0.92	0.68	-	0.67 [4]

Goal

Aligning machine with human decision

Dermatologists use comprehensive rules (ABCD, 7-point checklist)

→ **Dermatologist** →

Asymmetry
Border
Color
Differential structures

→ **Melanoma**

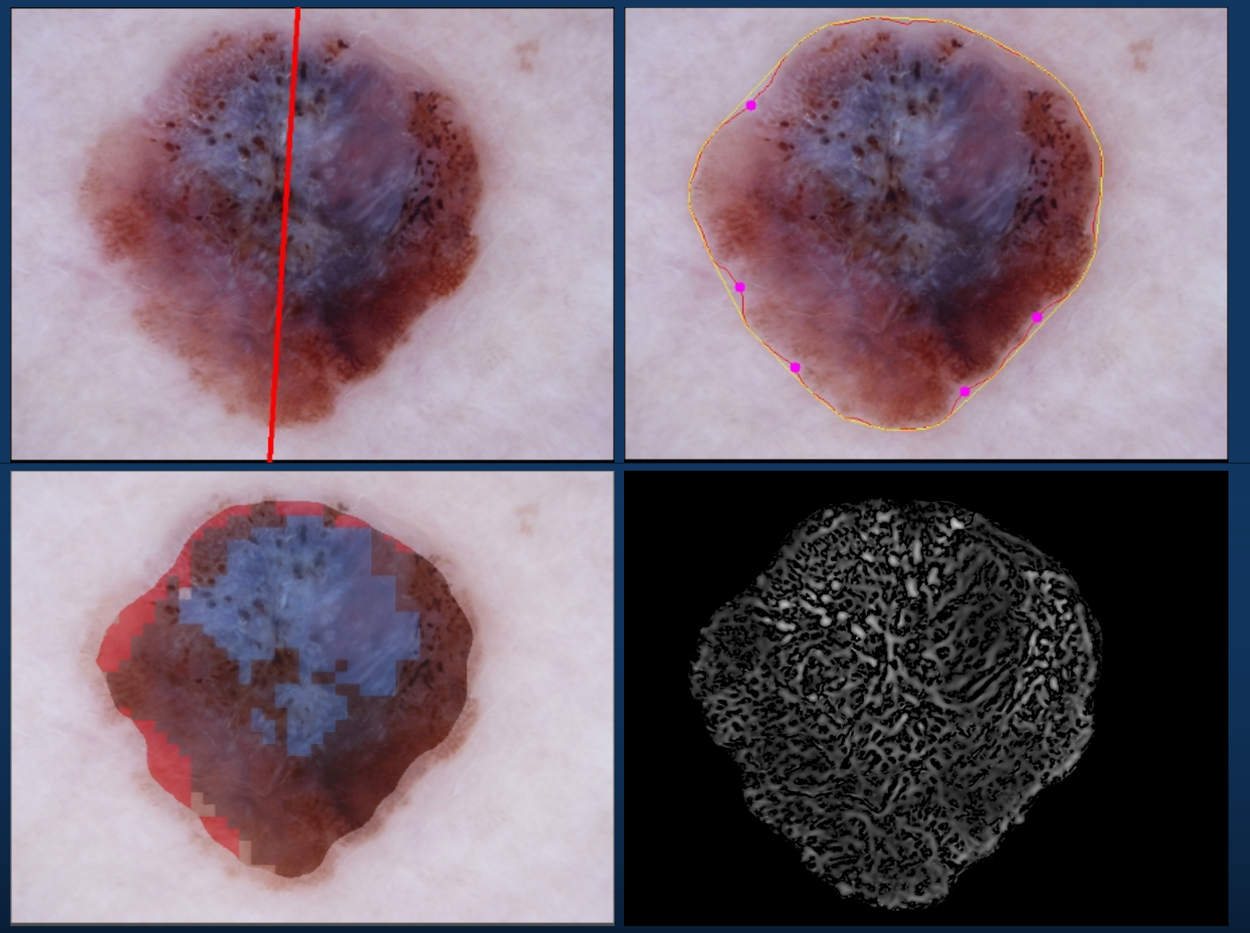
Our contributions:

1) **Tool for practitioner**: visualizing ABCD criteria
2) Build an **interpretable** classifier using concepts

Visualizing ABCD criteria

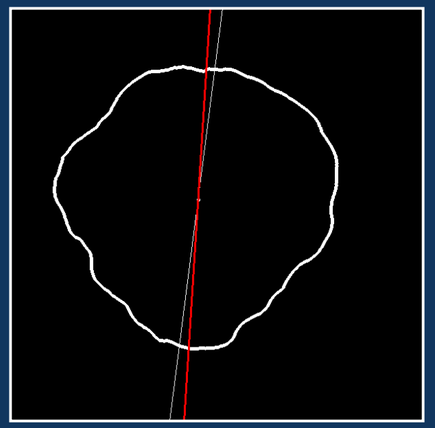
Based on the 4 criteria of the **ABCD** rule from (Stolz et al. 1991):

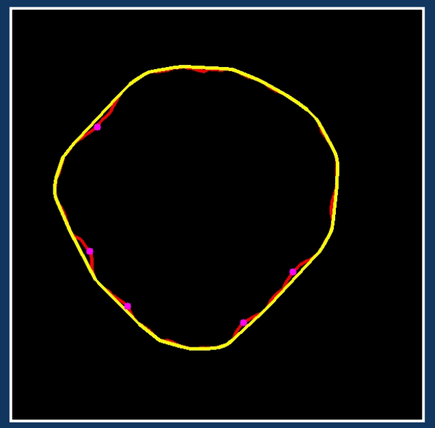
- ➔ **A**symmetry
- ➔ **B**order irregularity
- ➔ **C**olors
- ➔ **D**ifferential structures

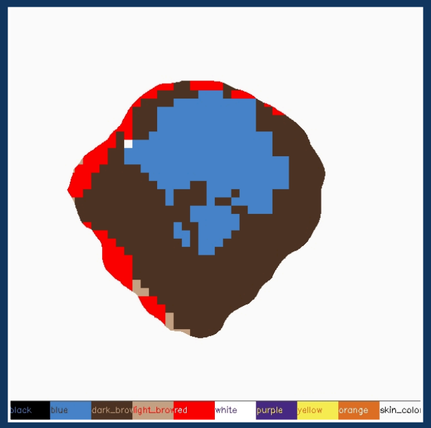


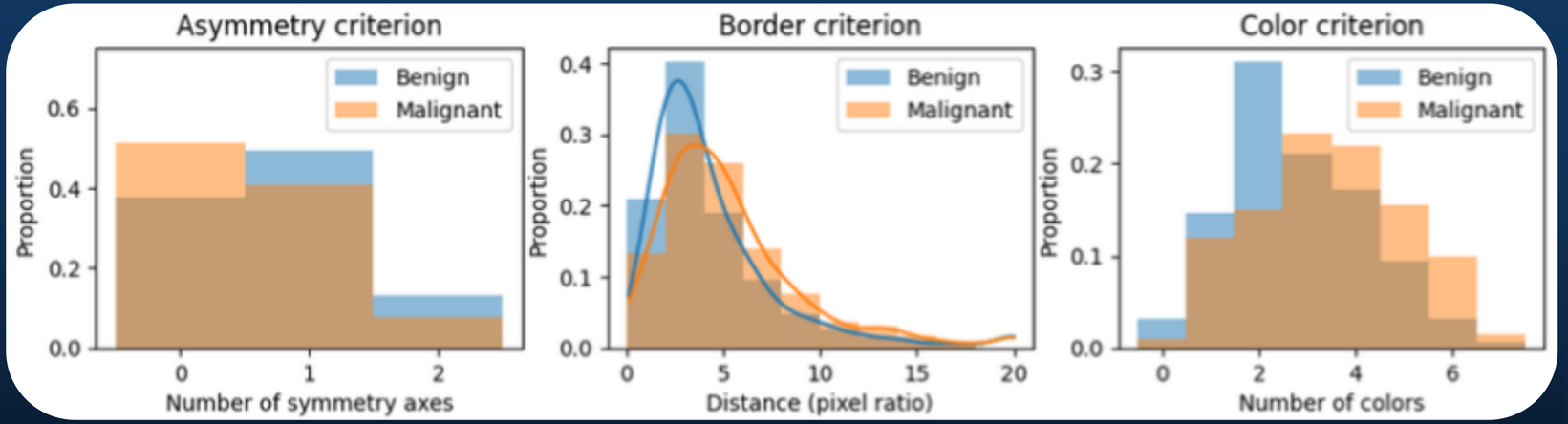
Assessing ABC criteria

- ➔ **A**: Fewer symmetry axes in malignant lesions
- ➔ **B**: Borders are smoother in benign lesions
- ➔ **C**: More color variations within malignant lesions





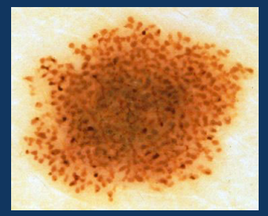




Has the model learned medical concepts?

The **7-point checklist** (7PCL)

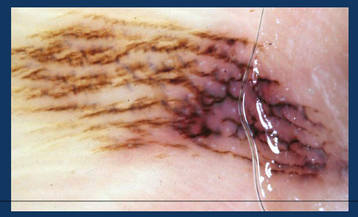
- Annotated in the Atlas dataset (Argenziano et al. 2022)
- Represented as concepts
  - CAV (Concept Activation Vector)



Dots and globules

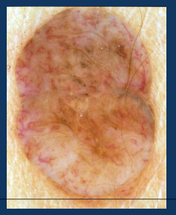
Pigmented network

Streaks

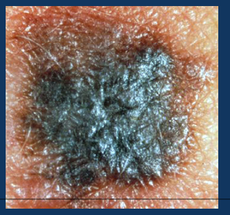


Regression structures

Pigmentation



Vascular structure

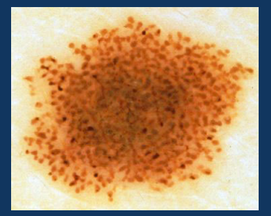


Blue-whitish veil

Has the model learned medical concepts?

The **7-point checklist** (7PCL)

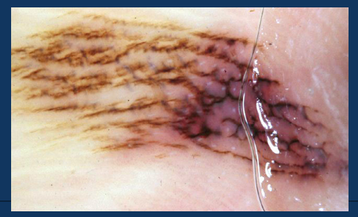
- Annotated in the Atlas dataset (Argenziano et al. 2022)
- Represented as concepts
  - CAV (Concept Activation Vector)



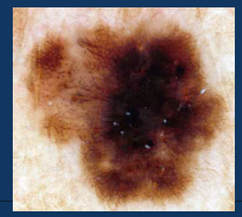
Dots and globules

Pigmented network

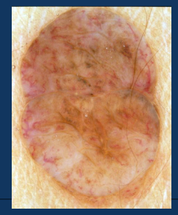
Streaks



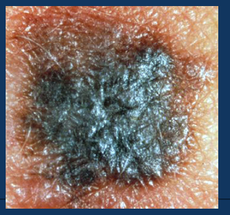
Regression structures



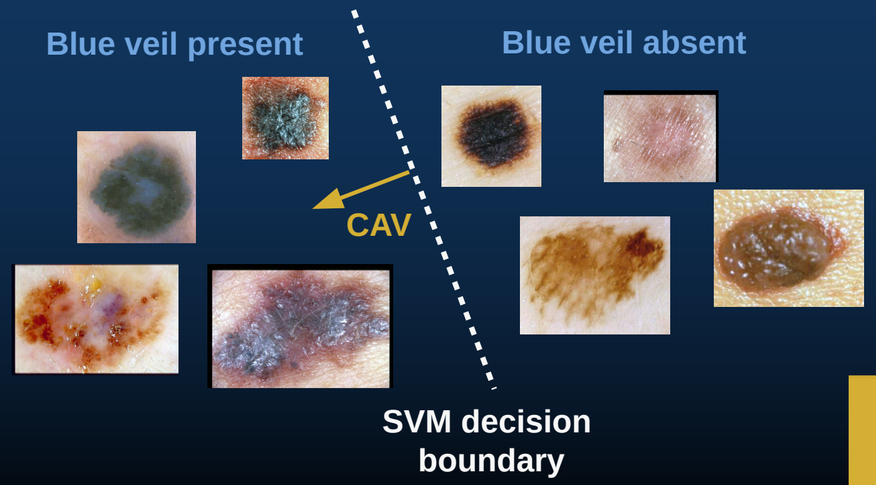
Pigmentation



Vascular structure



Blue-whitish veil



Has the model learned medical concepts?

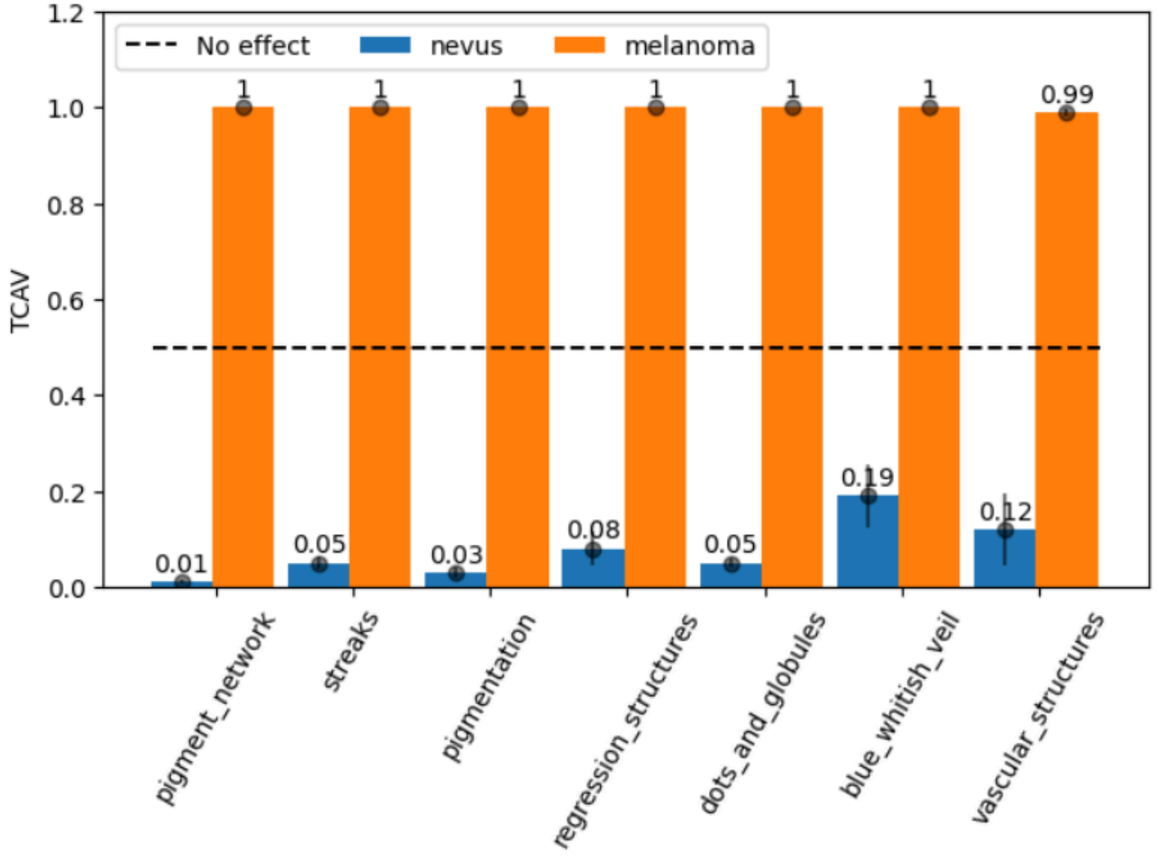
The **7-point checklist** (7PCL)

- Annotated in the Atlas dataset (Argenziano et al. 2022)
- Represented as concepts
  - CAV (Concept Activation Vector)

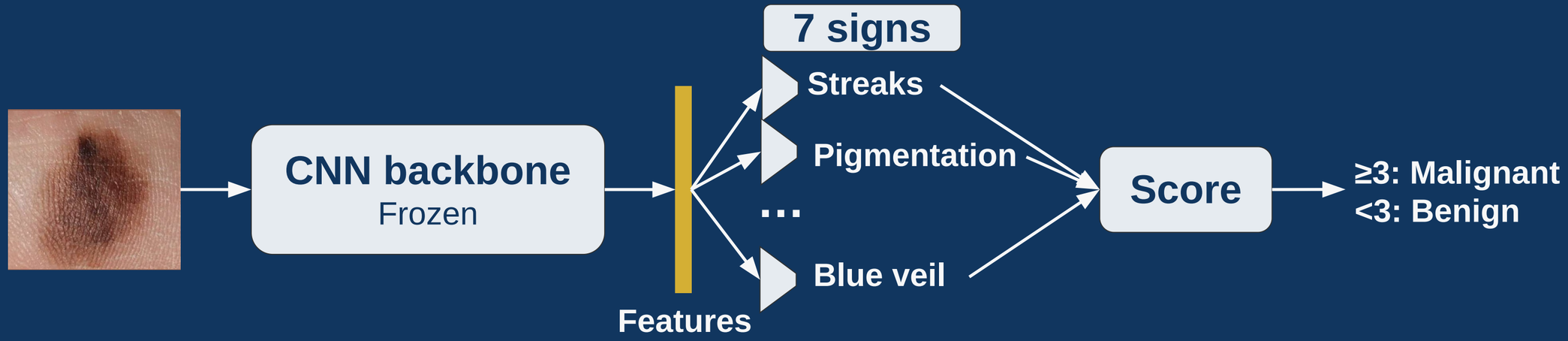
How important is each concept for the class melanoma?

TCAV: Testing with CAV (Kim et al. 2018)

- % melanoma images with derivative in same direction as CAV



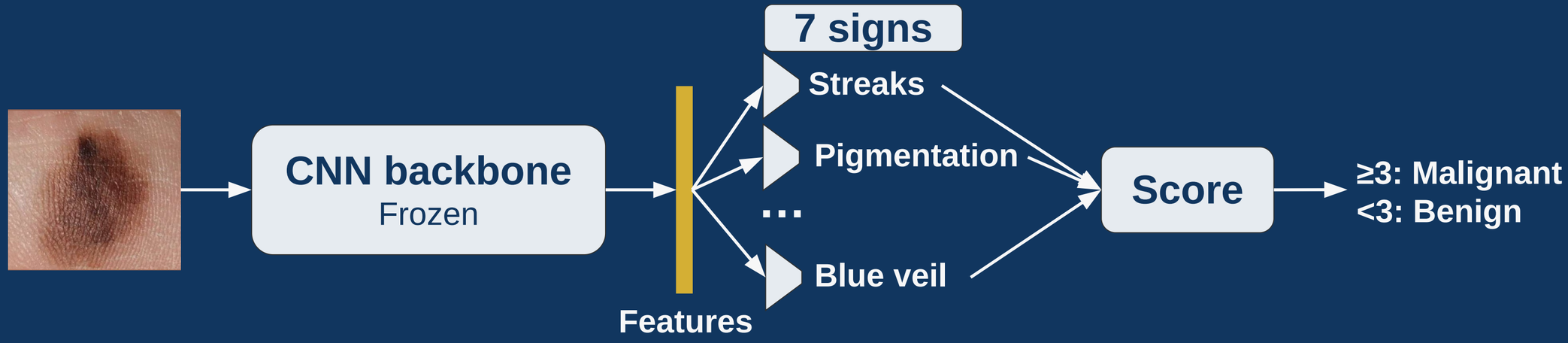
Interpretable concept-based model



Training: sign classification + diagnosis

$$\mathcal{L}(y, \hat{y}) = \sum_{i=1}^{7} CE(y_i, \hat{y}_i) + MSE(\sum_{i=1}^{7} \mathbb{1}(y_i \neq 0), \sum_{i=1}^{7} \mathbb{1}(\hat{y}_i \neq 0))$$

Interpretable concept-based model



Training: sign classification + diagnosis

$$\mathcal{L}(y, \hat{y}) = \sum_{i=1}^{7} CE(y_i, \hat{y}_i) + MSE(\sum_{i=1}^{7} \mathbb{1}(y_i \neq 0), \sum_{i=1}^{7} \mathbb{1}(\hat{y}_i \neq 0))$$

Training	MSE+CE	CE-only
Mean acc.	74.4%	75.9%
MAE	1.41	1.71

7-sign classification

Model	Baseline	Interpretable	CE-only
2-class acc.	85.2%	76.8%	72.8%

Diagnosis accuracy (benign vs malignant)

Conclusion

- Tool for practitioner to visualize ABCD rule
- Analysis of the medical concepts learned
- Interpretable concept-based classification model
  - Trade-off interpretability/accuracy
  - Next: Fit a residual term (Yuksekgonul et al 2023) to recover accuracy

TORUS AI

Thanks for your attention!

Any questions?

marianne.defresne@torus.ai
paul.fricker@torus.ai

Phone
+33 07 49 23 47 15

Website
www.torus.ai

Email Address
contact@torus.ai

Address
12 Av. de l'Europe, 31520 Ramonville-Saint-Agne