INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Evidence-Driven Differential Diagnosis of Malignant Melanoma

Naren Akash R J, Anirudh Kaushik and Jayanthi Sivaswamy
Center for Visual Information Technology, IIIT Hyderabad, India

Evidence-Based Diagnosis

EBM involves using the medical literature more effectively in guiding medical practice.

Clinicians leverage **richer set of data and information** in the diagnostic process.

History

Physical Exam

Medical Labs

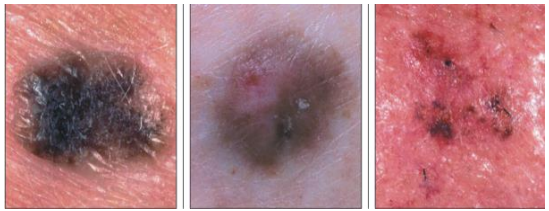
Imaging

Pathology

Genetics

Differential diagnosis is a two-word summary of how doctors think.
First consider the patient. Next, the patient's symptoms. Finally, the general environment.

Complexity of Melanoma Recognition



Melanoma – or mimic?

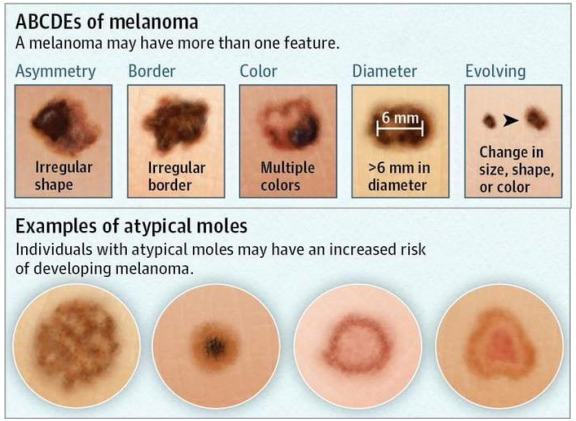
Melanoma can masquerade as benign lesions.
Benign pigmented lesions can resemble melanoma.

Benign lesions that are considered as melanoma mimickers

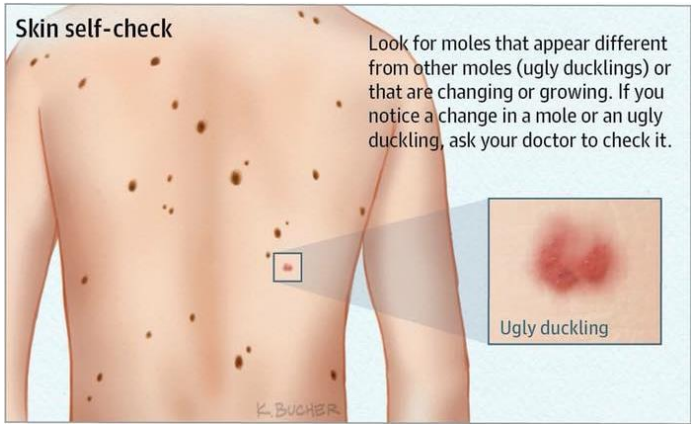
Nevi, atypical melanocytic proliferation, caf´e-au-lait macule, lentigo NOS, lentigo simplex, solar lentigo, lichenoid keratosis, and seborrheic keratosis

Differential Recognition of Malignant Melanoma

ABCDE and ugly duckling rules are complementary to one another.



The ABCDE mnemonic helps to outline the **physical characteristics of skin lesions**, which helps in determining whether it has features of early melanoma.



Most nevi in a patient tend to be similar and can be grouped into a few *perceived similarity clusters* (PSC) based on morphological similarity.

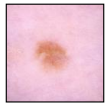
Any nevus that deviates from a consistent pattern within an individual is an outlier or an ugly duckling which is taken to be a suspicious lesion.

Higher Order Thinking of MM Differential Recognition

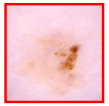
01 Lesion Focused Analysis

02 Patient Context Integration

03 Population Level Reasoning

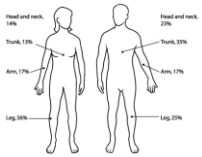


Assymetry?
Border?
Color?
Diameter?



Assymetry?
Border?
Color?
Diameter?

ugly duckling



Below the age of 50, women have a higher risk of developing melanoma than men.

After the age of 50, men are more likely to develop melanoma than women.

Previous Works in Melanoma Recognition

Existing deep learning methods are largely lesion focused approaches.
Includes seven-point checklist, hierarchical structures, lesion segmentation, and ABCD-based medical representations.

Most methods have not fully leveraged the clinician's comprehensive diagnostic process and strategy.
Although CI-Net utilizes some strategy, its focuses only on individual lesion characteristics.

Approaches to model patient context assumes fixed number of lesions.
UDTR is designed for a fixed number of lesions and uses repeated sampling and truncation.

No attempt has been made by any approach so far to take into account a richer set of information that clinicians rely on for melanoma diagnosis.
These include lesion counts in a patient, which can be variable, lesion location in the body and patient demographic information.

Our aim is to understand how the addition of specific information influences the decision-making process.
An understanding of the sensitivity-specificity trade-off when considering different information can make a method more transparent.

Our Contributions

01

A modular, **multi-level framework** to holistically integrate evidence at multiple levels (lesion, patient and population).

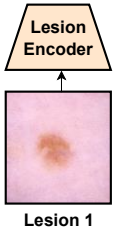
02

A solution based on a *masked transformer* to utilize **variable-count context lesions** from a patient along with their **anatomic location** and **metadata** (age and sex).

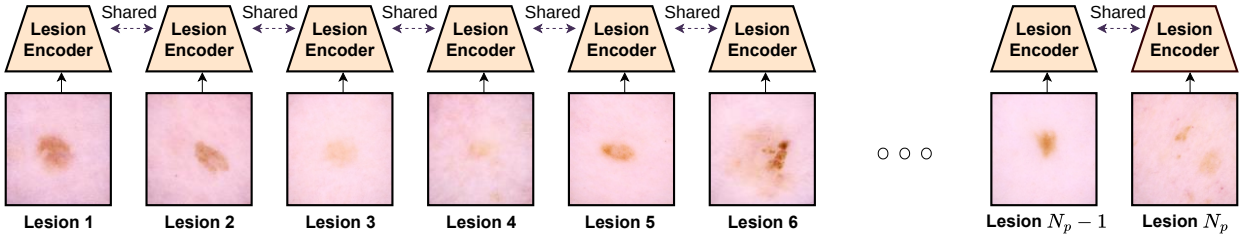
03

Insights on the role of various information in melanoma recognition, based on validation results on the SIIM-ISIC 2020 dataset.

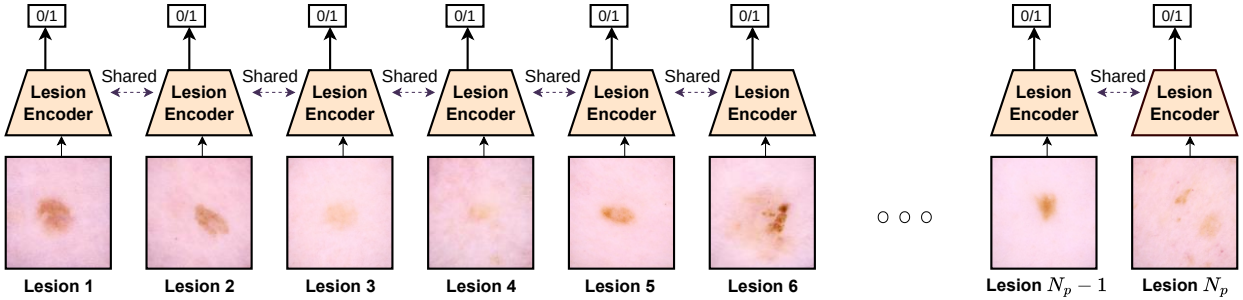
Our Proposed MelDD Framework



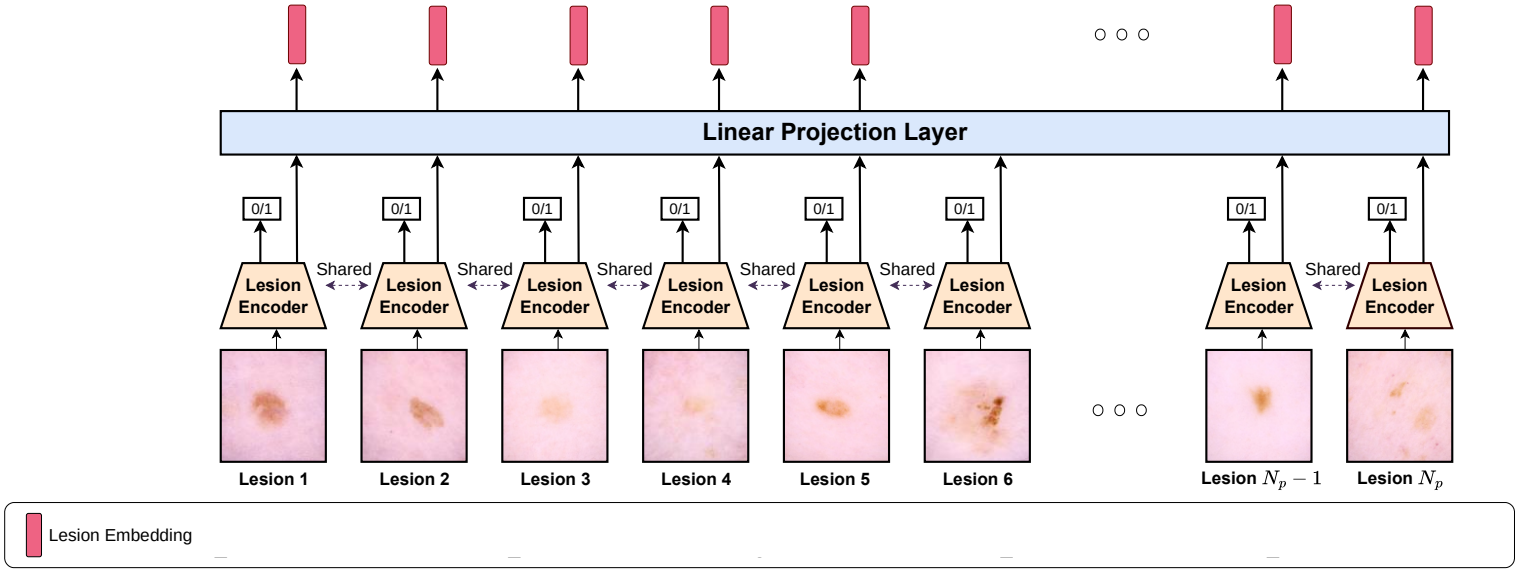
Our Proposed MelDD Framework



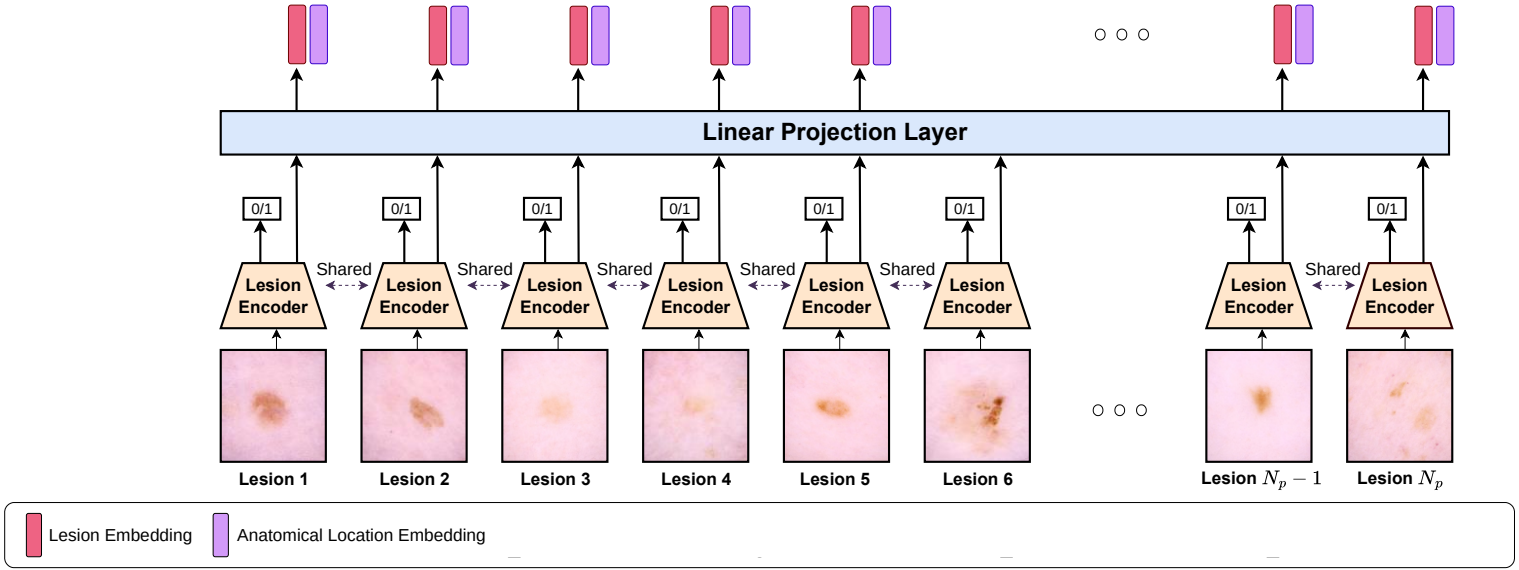
Our Proposed MelDD Framework



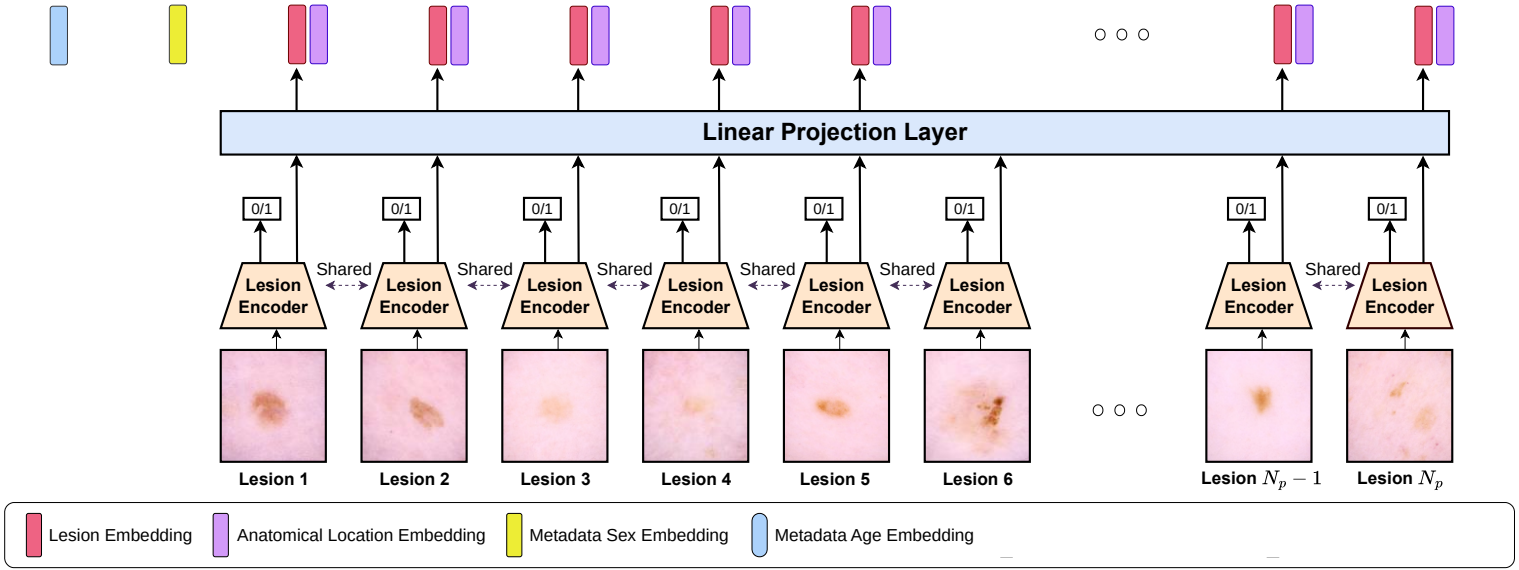
Our Proposed MelDD Framework



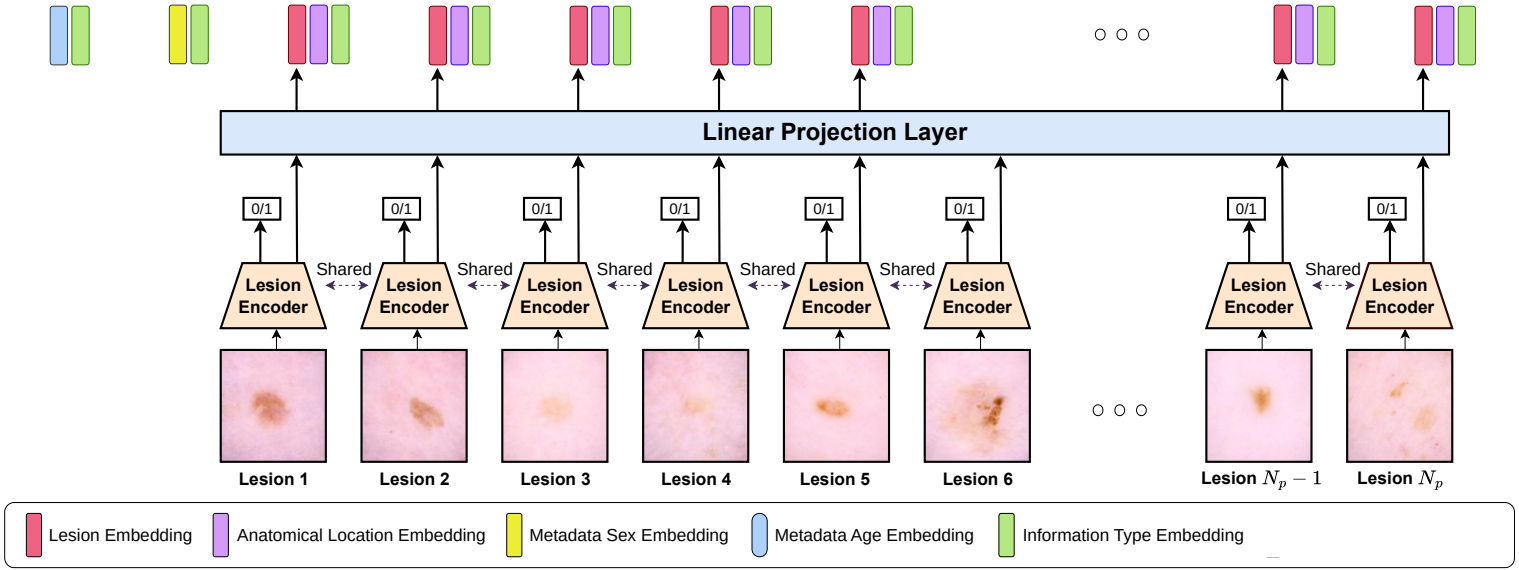
Our Proposed MelDD Framework



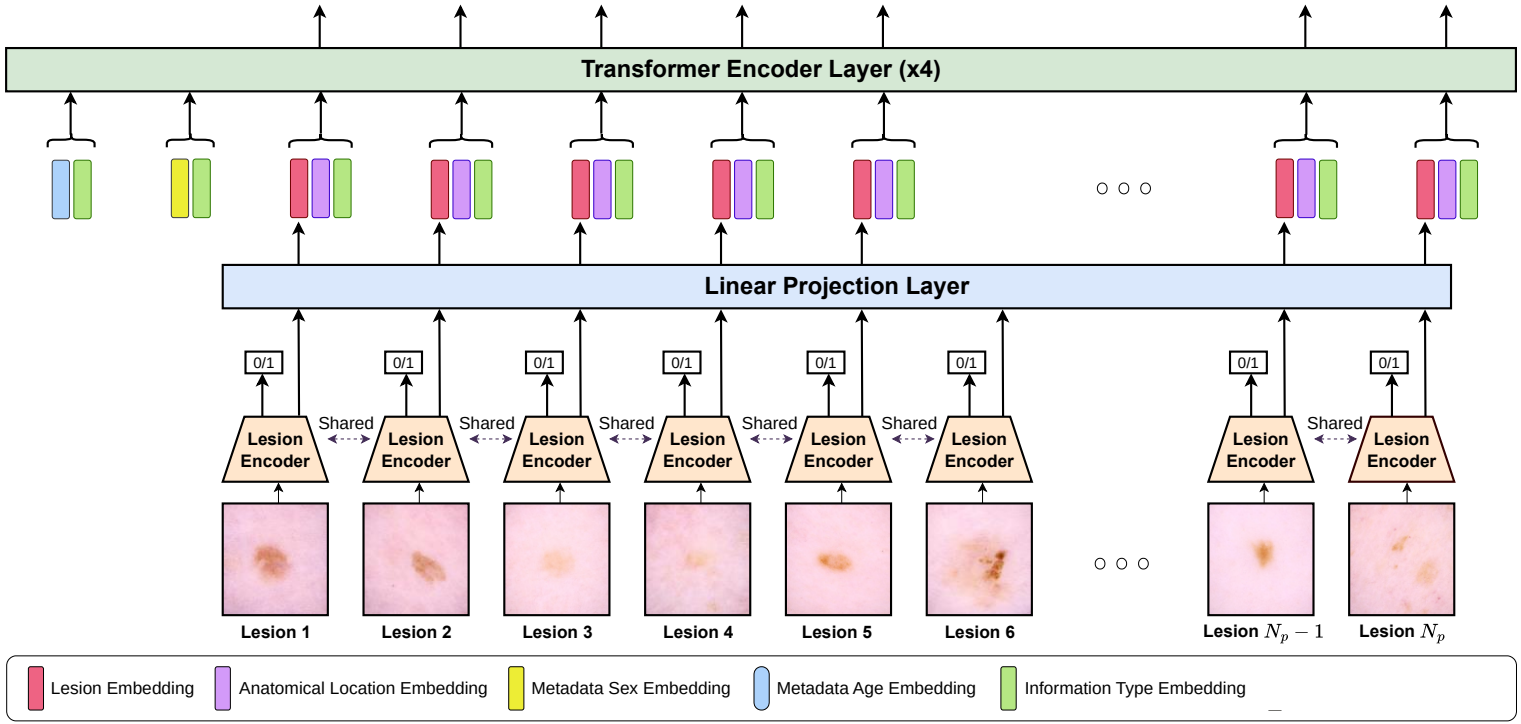
Our Proposed MelDD Framework



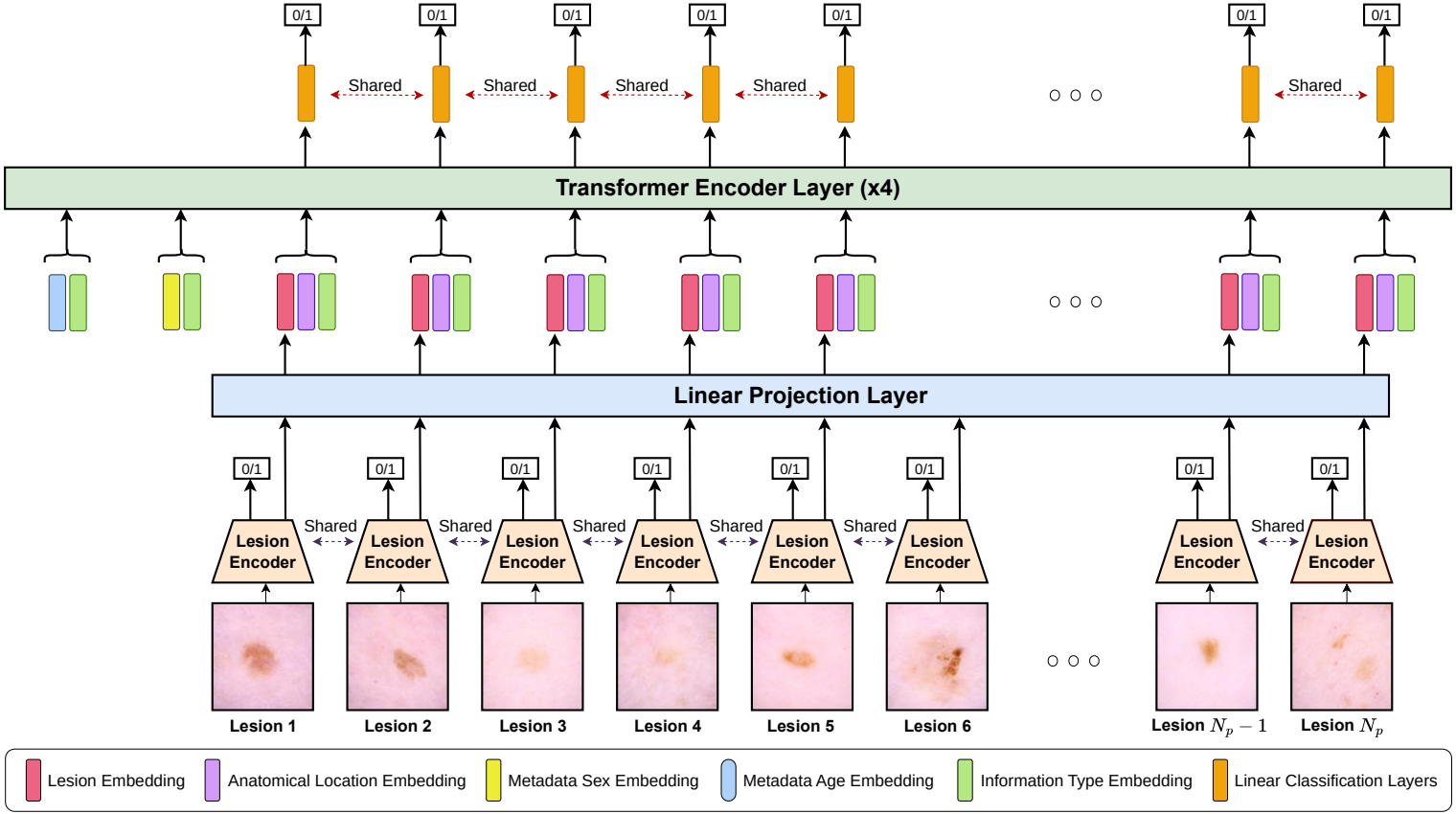
Our Proposed MelDD Framework



Our Proposed MelDD Framework



Our Proposed MelDD Framework



Dataset and Metrics

SIIM-ISIC 2020 dataset includes 2,056 patients with 428 individuals exhibit at least one melanoma.
Comprises of 33,126 dermoscopic images, including 584 histopathologically confirmed melanomas.
Is **severely imbalanced**, with melanomas accounting for only 1.8% of the samples.

In addition to image data, metadata pertaining to age and sex and general anatomic site of lesion present.

Many SOTA models for ISIC 2020 classification focus on **optimizing the area under the ROC curve (AUC).**

- May be inappropriate since AUC is not clinically interpretable.
- Identical AUC values yet perform differently at clinically significant thresholds.

Reports a high AUC score but exhibits poor sensitivity, making it unsuitable for clinical use in MM recognition.

	BACC	SN	SP	AUC
CI-Net	0.6200	0.3220	0.9180	0.9230

We opt to optimize the **balanced accuracy (BACC) at the Youden's J index**.
May be more clinically meaningful for a small and imbalanced dataset with low melanoma prevalence.

Comparison of Melanoma Recognition Performance

	P	V	L	M	BACC	SN	SP	AUC
Variant 0	✗	-	✗	✗	0.7649	**0.8867**	0.6431	0.8371
Variant 1	✓	✓	✗	✗	0.7841	0.8679	0.7003	0.8558
Variant 2	✓	✓	✓	✗	**0.7904**	0.8274	**0.7534**	**0.8612**
Variant 3	✓	✓	✗	✓	0.7867	0.8843	0.6890	0.8544
Variant 4	✓	✓	✓	✓	0.7793	0.8761	0.6825	0.8504

P = patient context, V = varying lesion count, L = anatomic location, M = metadata
BACC = balanced accuracy, SN = sensitivity, SP = specificity (at Youden's J statistic cut-off)

Performance Improvement with Additional Information

	P	V	L	M	BACC	SN	SP	AUC
Variant 0	✗	-	✗	✗	0.7649	0.8867	0.6431	0.8371
Variant 1	✓	✓	✗	✗	+2.51%	−2.12%	+8.89%	+2.23%
Variant 2	✓	✓	✓	✗	**+3.33%**	−6.69%	**+17.15%**	**+2.88%**
Variant 3	✓	✓	✗	✓	+2.85%	−0.27%	+7.14%	+2.07%
Variant 4	✓	✓	✓	✓	+1.88%	−1.20%	+6.13%	+1.59%

P = patient context, V = varying lesion count, L = anatomic location, M = metadata
BACC = balanced accuracy, SN = sensitivity, SP = specificity (at Youden's J statistic cut-off)

Which information for a specific use case?

	P	V	L	M	BACC	SN	SP	AUC
Variant 0	✗	-	✗	✗	0.7649	0.8867	0.6431	0.8371
Variant 1	✓	✓	✗	✗	+2.51%	−2.12%	+8.89%	+2.23%
Variant 2	✓	✓	✓	✗	**+3.33%**	−6.69%	**+17.15%**	**+2.88%**
Variant 3	✓	✓	✗	✓	+2.85%	−0.27%	+7.14%	+2.07%
Variant 4	✓	✓	✓	✓	+1.88%	−1.20%	+6.13%	+1.59%

Relative importance of SN and SP varies based on priorities.

01

A high SP value will be required to avoid overdiagnosis and needless biopsies.
MelDD-V2 is a good choice. Patient context and anatomic site of lesions do play a crucial role.

02

A higher SN is preferable if the application scenario is screening.
MelDD-V3 is a good choice. Using metadata instead of lesion location may be preferable.
This suggests patient sex and age do play a key role in improving SN.

Intuitively, combining all information should be beneficial to performance which is not seen.
Skew in the melanoma cases in the dataset. Stratification in the data split needs to be explored.

Comparison with State-of-the-Art Solutions

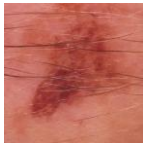
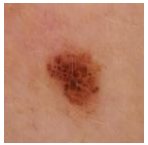
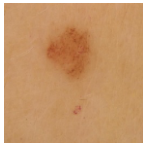
	P	V	L	M	BACC	SN	SP	AUC
Variant 0	✗	-	✗	✗	0.7649	**0.8867**	0.6431	0.8371
Variant 1	✓	✓	✗	✗	0.7841	0.8679	0.7003	0.8558
Variant 2	✓	✓	✓	✗	**0.7904**	0.8274	**0.7534**	**0.8612**
Variant 3	✓	✓	✗	✓	0.7867	0.8843	0.6890	0.8544
Variant 4	✓	✓	✓	✓	0.7793	0.8761	0.6825	0.8504
CI-Net	✗	-	✗	✗	0.6200	0.3220	0.9180	0.9230
UDTR-L	✓	✗	✗	✗	0.7564	0.7522	0.7605	0.8493
UDTR-Ad	✓	✗	✗	✗	0.7094	0.7922	0.6266	0.7634
UDTR-F	✓	✗	✗	✗	0.8183	0.8164	0.8202	0.8964

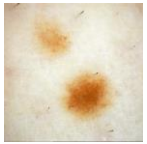
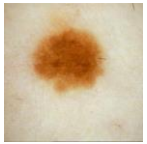
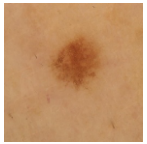
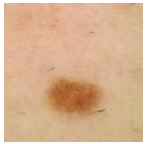
P = patient context, V = varying lesion count, L = anatomic location, M = metadata
BACC = balanced accuracy, SN = sensitivity, SP = specificity (at Youden's J statistic cut-off)

Prediction Changes with Patient Context

Multiple atypical lesions reduce suspicion of malignancy in an additional atypical lesion.
A morphologically typical lesion distinct in the nevus landscape is considered suspicious.

About the patient	
Age	56 years old
Sex	Male

	head/neck	head/neck	torso	torso	torso	torso	torso	oral/genital	palm/soles	lower extremity
										
Lesion Only	malignant	malignant	malignant	benign	malignant	malignant	malignant	malignant	benign	malignant
+ Patient Context	malignant	malignant	benign	benign	benign	malignant	benign	malignant	benign	benign

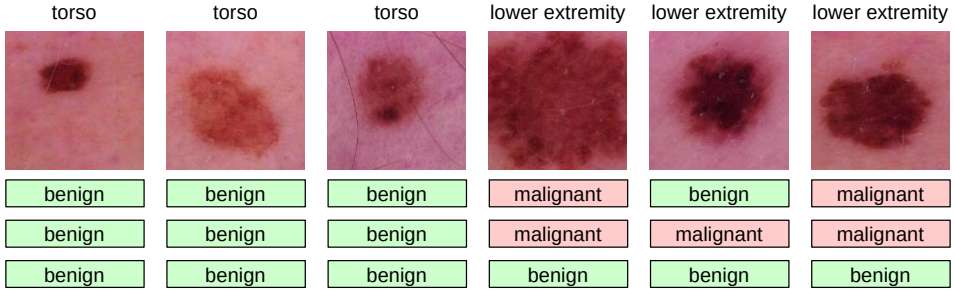
	lower extremity	lower extremity	lower extremity	lower extremity	lower extremity	lower extremity	lower extremity	lower extremity	upper extremity	upper extremity
										
Lesion Only	benign	benign	malignant	malignant	malignant	malignant	benign	benign	malignant	malignant
+ Patient Context	benign	malignant	benign	benign	benign	malignant	benign	benign	malignant	malignant

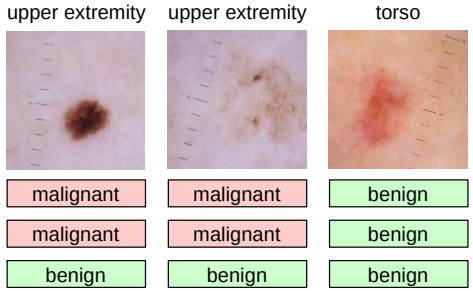
Prediction Changes with Patient Context and Lesion Location

Incorporating location could prevent the misclassification of benign lesions by considering the specific anatomical characteristics that differentiate suspicious lesions in different locations.

About the patient	
Age	78 years old
Sex	Male

About the patient	
Age	65 years old
Sex	Female

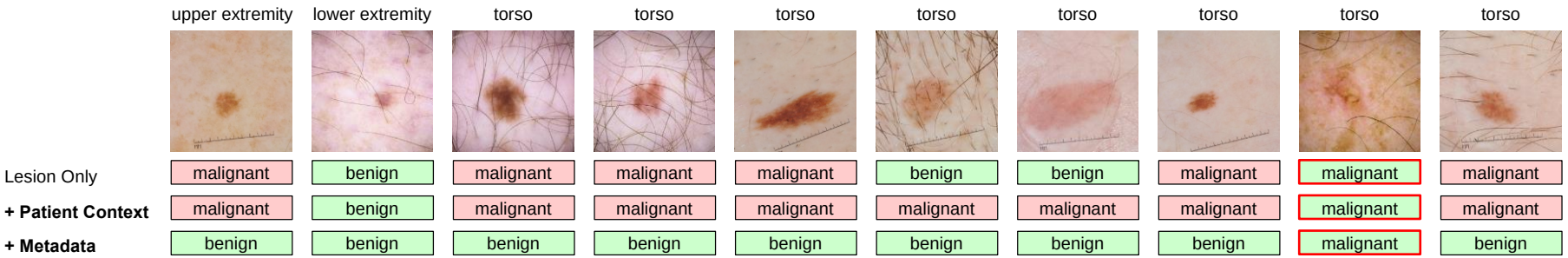




Prediction Changes with Patient Context and Metadata

The patient demographics could help the model **correlate lesion characteristics with susceptibility to risk factors**, avoiding misdiagnosis of benign lesions as malignant based on a **better understanding of patient-specific factors**.

About the patient	
Age	46 years old
Sex	Male



Conclusion

01 A **modular, multi-level framework** for melanoma diagnosis, inspired by clinical reasoning and utilizing multiple sources of information, **integrating lesion, patient, and population levels**.

02 Since the number of lesions in unknown, MelDD employs a *masked transformer* to seamlessly incorporate **variable lesion counts**, enabling flexible integration of patient context information.

03 Results show the **differential roles** played by additional information.
- **context** and **location** leads to a **significant improvement in SP** with a marginal dip in SN.
- **metadata** serves to restore SN to that of the baseline with a **modest increase in SP**.

04 Optimizing BACC at Youden's J index aids in gaining **better control over SP and SN variations**, avoiding the big SP-SN tradeoff seen with conventional AUC optimization.

05 Our solution offers a **transparent decision support system** for melanoma recognition, supporting clinicians in evidence-based decision-making.

Thanks

Do you have any questions?

Evidence-Driven Differential Diagnosis of Malignant Melanoma

Naren Akash R J, Anirudh Kaushik and Jayanthi Sivaswamy
Center for Visual Information Technology, IIIT Hyderabad, India

Project Page

Code on GitHub