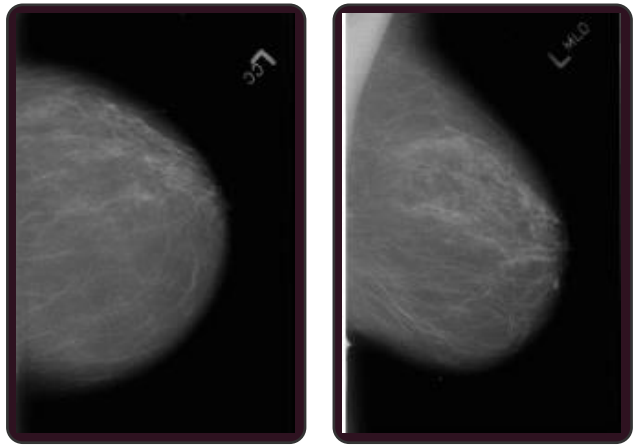
ICPR 2026

Complexity-Efficient Deep Learning for Breast Cancer Detection Using BI-RADS Descriptors

Structured expert knowledge beats architectural complexity.



MLO CC

Gil Ben-Artzi

Department of Computer Science, Ariel University, Israel
gilba@g.ariel.ac.il

The GAP: two ways to read a mammogram

The radiologist

- Reads with a standardised vocabulary
- Lesion shape, margin, density, calcification morphology and distribution
- These words are the reasoning — they drive the report and the management decision

The deep network

- Reads raw pixels, end to end
- Progress = more capacity, higher resolution, more views (CC + MLO)
- No access to the explicit clinical concepts — it must rediscover them from data

BACKGROUND

BI-RADS in one slide

Breast Imaging Reporting and Data System — the American College of Radiology's standardised lexicon for describing what a lesion looks like.

Descriptors → our input

- Mass shape: round / oval / irregular
- Margin: circumscribed / obscured / spiculated
- Calcification type: amorphous / pleomorphic …
- Distribution, breast density

Assessment categories → excluded

- BI-RADS 0–6 encode overall suspicion
- They are the radiologist's conclusion, not an observation
- Using them would leak the label — we deliberately exclude them

The Research Question

The discrepancy between data-driven learning and expert-guided interpretation raises a fundamental question: to what extent can explicit radiological domain knowledge contribute to deep visual representations?

Methodology: Evaluate systematically the interaction between BI-RADS descriptors as inputs and Deep-learning models

HOW PRIOR WORK USED BI-RADS

As output

Predicted as an auxiliary task for post-hoc explainability.
BI-RADS-Net, DeepMiner

As input · ultrasound

Tried before — no consistent gain over image-only models.
Carrilero-Mardones et al.

As input · mammography

Gains reported, but for one fixed configuration only.
Ben-Artzi et al., 2024

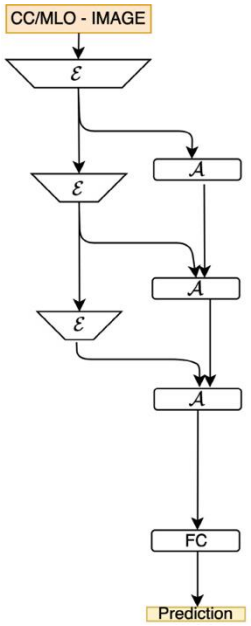
Why this is not trivial: descriptor assignment is subjective — inter-reader agreement is only slight to fair, even among experienced radiologists (Lee et al., 2017).

METHOD

What we compare

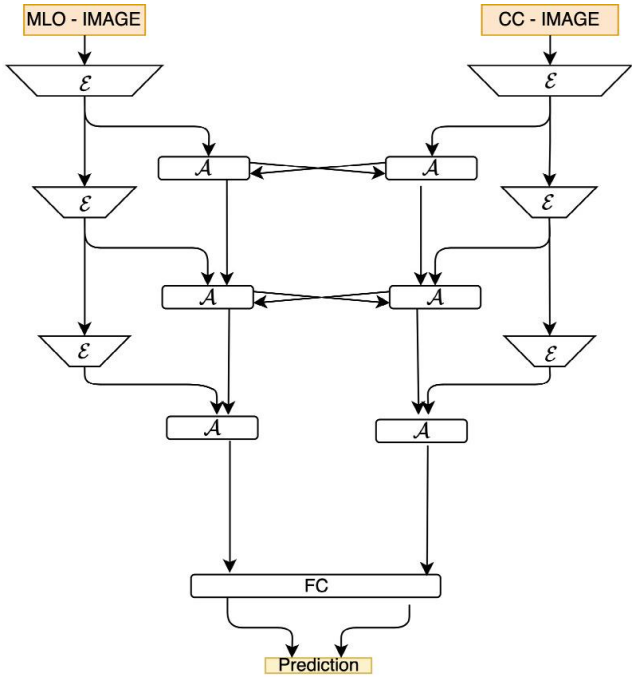
Single view

CC or MLO · self-attention



Multi view

CC + MLO · cross-view attention



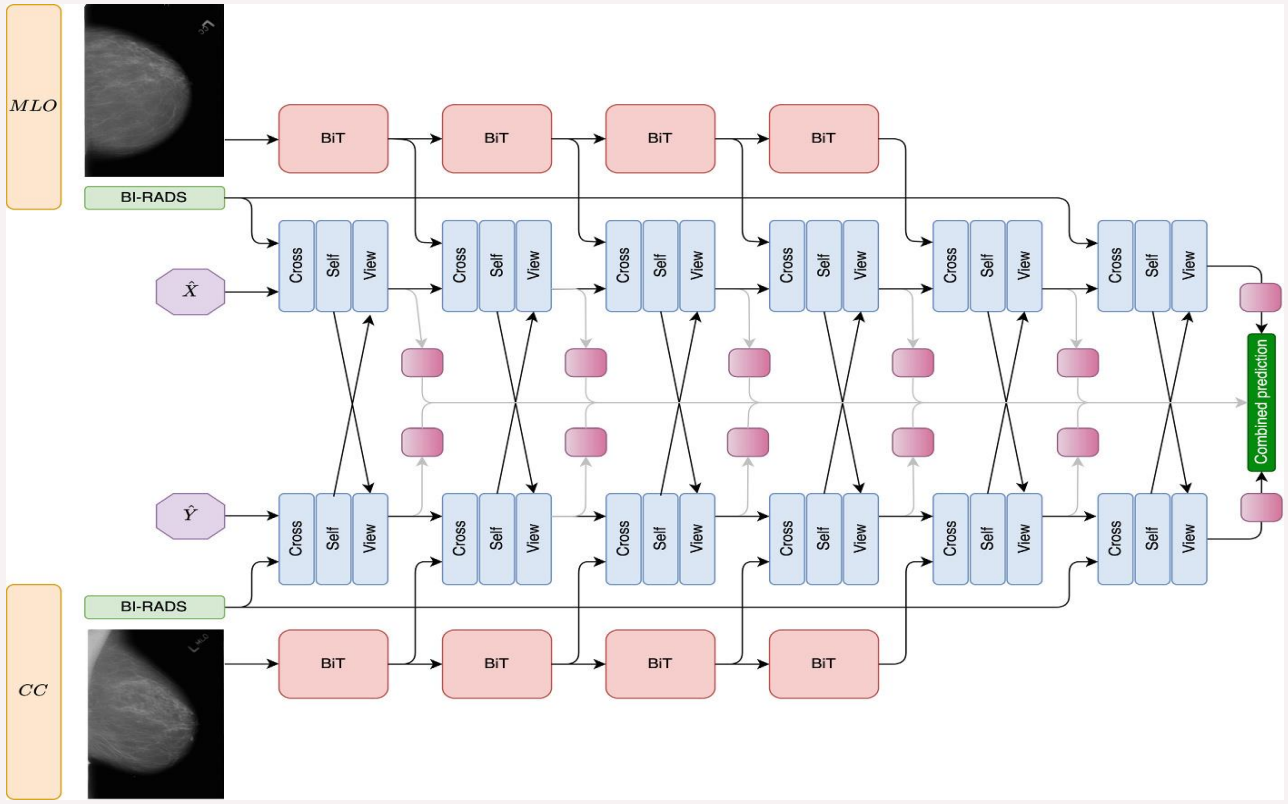
BI-RADS only

Decision tree on the descriptors alone — no pixels at all. Isolates how much of the signal lives in the words.

Each deep model is run with and without descriptors: *5 configurations × 3 lesion tasks*

METHOD

How the descriptors enter the network



Dual-branch model: BiT image features (pink) and BI-RADS vectors meet inside stacked attention layers (blue).

Descriptors → 256-d binary vector

One index per descriptor class. Enters the first attention layer, with a skip connection to the last.

Images → BiT / ResNet-101

Features taken at four resolutions; pre-trained on PatchCamelyon.

6 stacked multi-attention layers

Cross-attention (text ↔ image) · self-attention (within image) · view-attention (CC ↔ MLO).

EXPERIMENTAL SETUP

Data and protocol

1,566
patients

3,568
abnormalities

1,696
masses

1,872
calcifications

CBIS-DDSM

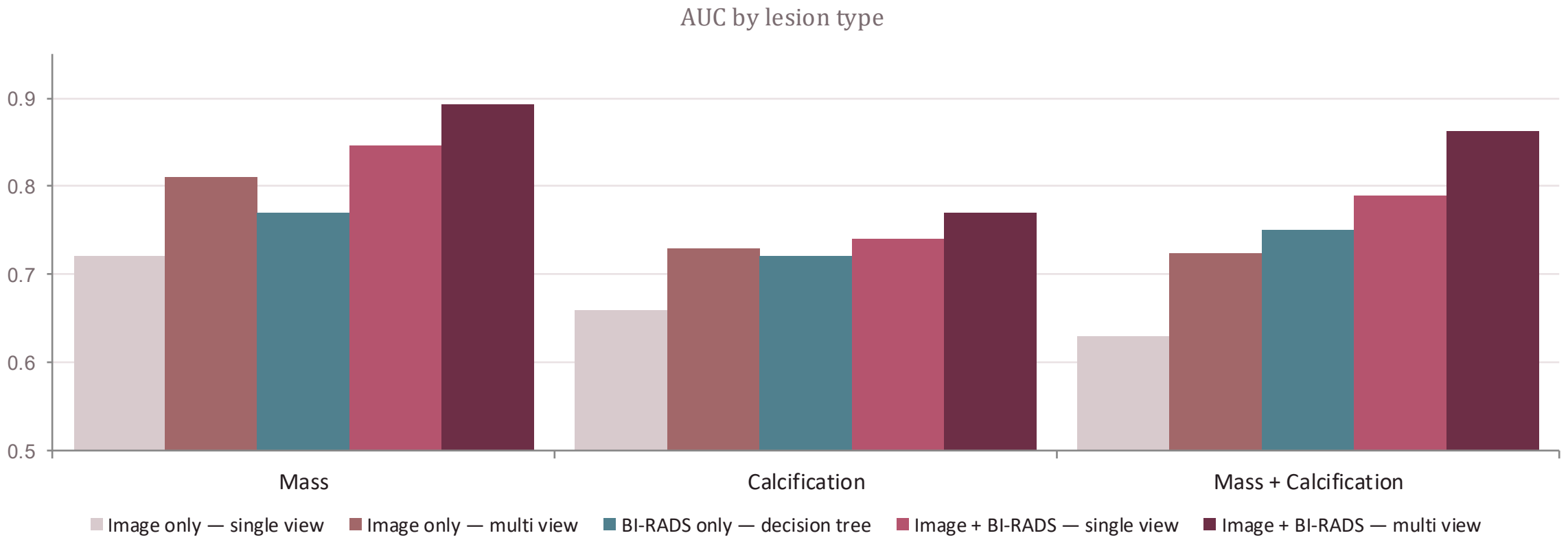
- One of the few public mammography sets shipping BI-RADS descriptors as metadata
- CC and MLO views, radiologist annotations, ROI crops and masks
- Task: benign vs. malignant

Training protocol

- 1024 × 1024 inputs; flips and elastic deformation
- SGD, momentum 0.9, lr $10^{-3} \rightarrow 10^{-4}$ after 500 iterations, batch 16
- Stratified five-fold cross-validation for every configuration

RESULTS

Every task, every model



In all three tasks the two tallest bars are the two BI-RADS-augmented models.

FINDING 1

Knowledge beats complexity

Single-view image-only baseline · Mass lesions · AUC gains shown

Baseline
0.720
Single view image only

+0.090
Multi view image only

spend architecture

+0.127
Single view + BI-RADS

spend knowledge

The cheaper model wins

A single-view network with descriptors beats a full multi-view network without them.

Calcifications, same pattern: +0.081 by adding descriptors to a single-view model, versus +0.070 for the multi-view image-only model.

FINDING 2

Complementary, not compensatory

Multi-view models — AUC without vs. with BI-RADS descriptors

Mass	+0.082
Calcification	+0.040
Mass + Calc.	+0.140

Not just for weak models

Descriptors still help the strongest architecture we have — so they carry information the pixels do not deliver.

Balanced, not traded off: sensitivity improves without inflating false positives.

FINDING 3

One model for both lesion types

Unified model · masses + calcifications

0.863

AUC · same architecture, same parameter count as the lesion-specific models

0.723	Unified model · without BI-RADS
0.810	Dedicated mass model · image only
0.730	Dedicated calc. model · image only

the unified BI-RADS model beats all three

Why this matters

Merging heterogeneous lesion types normally costs accuracy unless you add capacity. Here the descriptors supply the shared semantic structure instead — generalisation from knowledge, not from parameters.

FINDING 4

Descriptors alone: strong prior, hard ceiling

Decision tree · no pixels

Mass	0.770
Calcification	0.720
Mass + Calc.	0.750

Above every single-view CNN

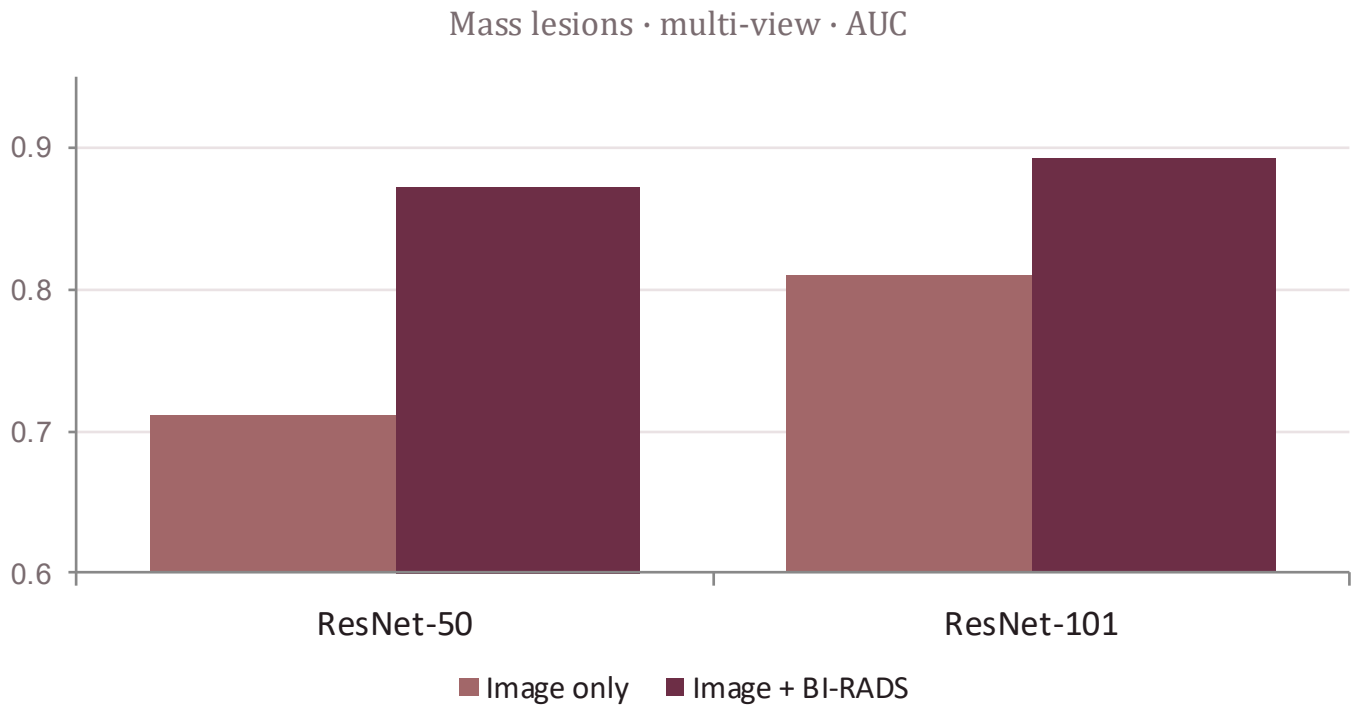
- Beats the single-view image-only model on all three tasks (e.g. 0.770 vs 0.720 for masses)
- Approaches it on calcifications (0.720 vs 0.730) and beats it on the combined task (0.750 vs 0.723)
- A few categorical attributes rival a deep network trained on 1024 × 1024 images

...but it is beaten by every fused model

0.770 vs 0.892 on masses · 0.720 vs 0.770 on calcifications · 0.750 vs 0.863 combined.
The descriptors compress a full mammogram into a handful of categorical attributes — that compression is exactly where the ceiling comes from. Neither modality is redundant.

FINDING 5

Knowledge survives a stronger encoder



Knowledge is not redundant

Descriptors help at both encoder capacities — a deeper backbone does not learn them away.

But the gap does narrow

Gain from BI-RADS: +0.161 with ResNet-50, +0.082 with ResNet-101. Better vision closes part — not all — of the gap.

Best of both: the strongest encoder combined with descriptors gives the best result overall (AUC 0.892, accuracy 0.790).

CONTRIBUTIONS

Five findings

1	**Knowledge beats complexity**	Adding descriptors lifts AUC more than adding a second view.
2	**Gains survive strong models**	Multi-view networks improve too — the signal is complementary.
3	**One model, both lesion types**	A single unified model handles masses and calcifications.
4	**Descriptors alone: strong, but capped**	A decision tree on BI-RADS beats single-view CNNs — and plateaus.
5	**Robust to encoder strength**	A deeper backbone does not learn the descriptors away.

LIMITATIONS

What this study does not yet show — and what is next

THE SETTING TODAY

One dataset

CBIS-DDSM only — digitised film, and a modest number of lesions.

Descriptors come from a reader

At inference a radiologist supplies the descriptors, so the model assists a reading rather than producing one.

Lesion-level setting

Annotated ROIs, not whole-image screening — one step away from the deployment task.

WHAT WE ARE DOING ABOUT IT

Generalisation across datasets — DONE

Already extended and validated beyond CBIS-DDSM in follow-up work (To appear in MICCAI 2026).

Descriptors predicted from pixels — ON-GOING

Predicted from the image and re-injected into the model — removing the annotation dependency.

Automatic ROI recovery — ON-GOING

Accurate automatic lesion localisation makes the pipeline fully end-to-end.

Thank you

Questions welcome

SEE THE FOLLOW-UP WORK

Semantic Feature Modulation for Mammographic Lesion Classification

Shahar Mahpod · Gil Ben-Artzi

MICCAI 2026 · Oral

LOOKING FOR COLLABORATION

Breast radiology — and neuro-radiology

If your group has mammographic data, radiologist readings or a clinical question worth modelling, please come and talk to me.

Gil Ben-Artzi · Department of Computer Science, Ariel University

gilba@g.ariel.ac.il

https://gil-ba.com

Computing resources provided by the Ariel HPC Center, Ariel University.