BREAST CANCER DETECTION WITH EXPLAINABLE AI: GRAD-CAM AND BOUNDING BOX-BASED INTERPRETABILITY

Samrat Kumar Dev Sharma, Md. Yusuf Hossain Ador, Md. Rukonuzzaman Rukon, Futanta Chakma and Mahmud Hossain

Presented By:

Samrat Kumar Dev Sharma
Department of Statistics
Jagannath University, Dhaka

BBCC 5th
Bangladesh Breast
Cancer Conference

Introduction

RESEARCH BACKGROUND

- Breast cancer is the most common cancer in women worldwide.
- Early detection improves survival rates and treatment outcomes.

- Traditional mammography is time-consuming and varies by radiologist.
- AI models achieve high accuracy but face adoption challenges.
- Lack of interpretability limits clinical trust in AI.

- Explainable AI (XAI) enhances model transparency with Grad-CAM & bounding boxes.
- Improving interpretability bridges AI and medical professionals.
- Class imbalance & generalization must be addressed for reliable AI diagnostics.

RESEARCH OBJECTIVE

1. Develop an AI-based breast cancer detection model with high interpretability using XAI techniques.
2. Use Grad-CAM heatmaps and bounding box visualizations to highlight diagnostic features in mammograms.
3. Improve model performance by addressing class imbalance and incorporating metadata.

METHODOLOGY

Data Source

RSNA Screening Mammography Breast Cancer Detection dataset.

Dataset

There are a total of 54706 images in train across 11913 patients.

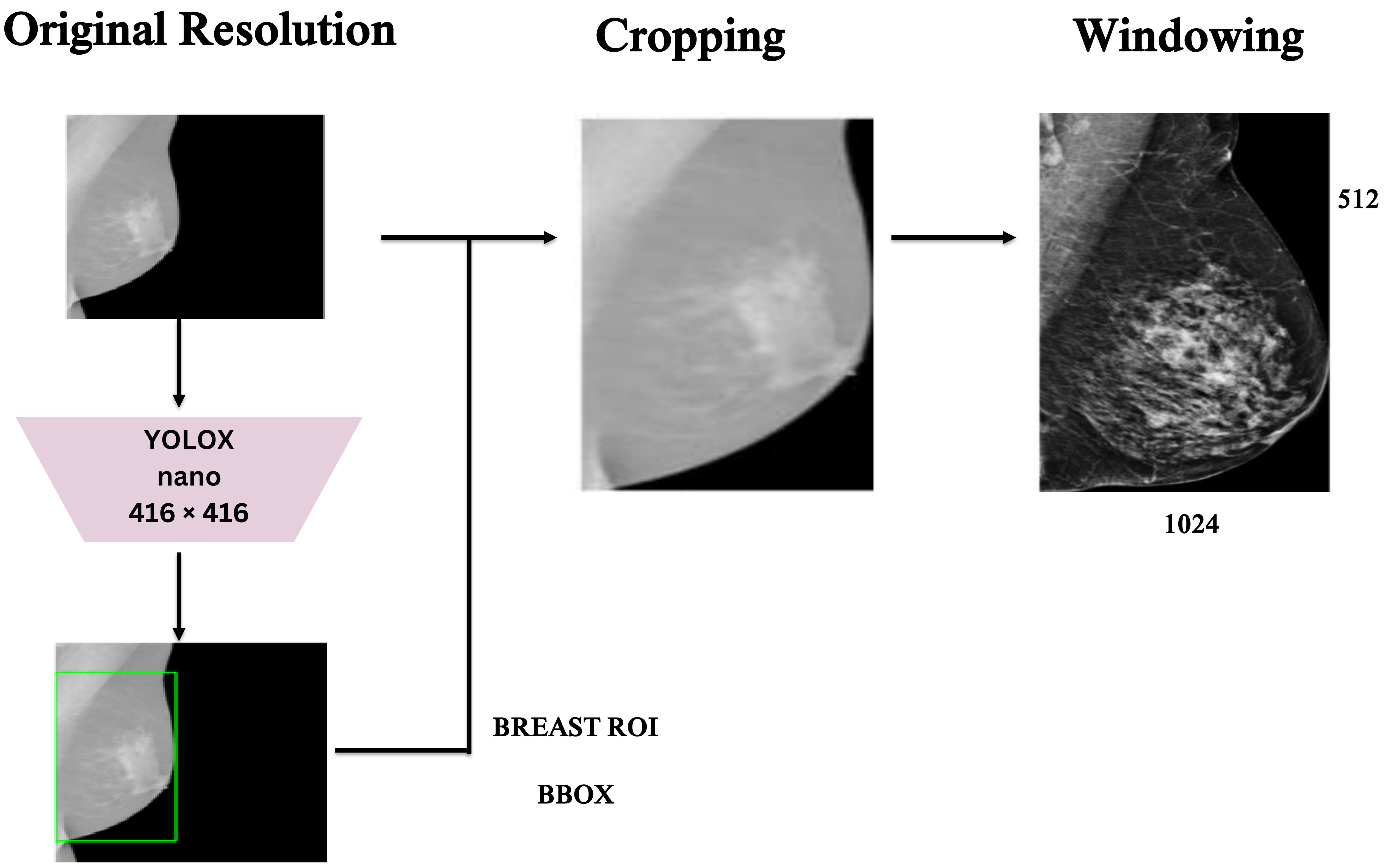
Metadata : Laterality, view, age, desnsity rating, cancer diagnosis

Class Imbalance

The dataset exhibited a severe class imbalance, with 1,158 cancer cases compared to 53,548 non-cancer cases, potentially leading to model bias toward the majority class. To address this issue, focal loss was implemented to reduce the emphasis on easily classified samples while prioritizing harder-to-classify cancer cases. Additionally, an up-sampling factor of 10 was applied to the cancer class, ensuring a more balanced distribution and improving the model's ability to detect malignant cases accurately.

METHODOLOGY

Image Data Preprocessing



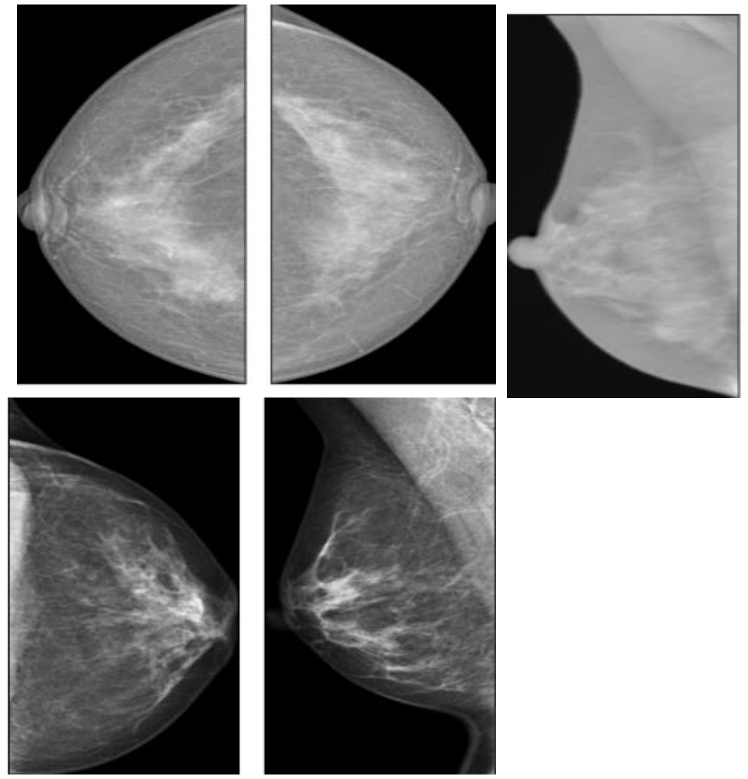
Images Augmentation

METHODOLOGY

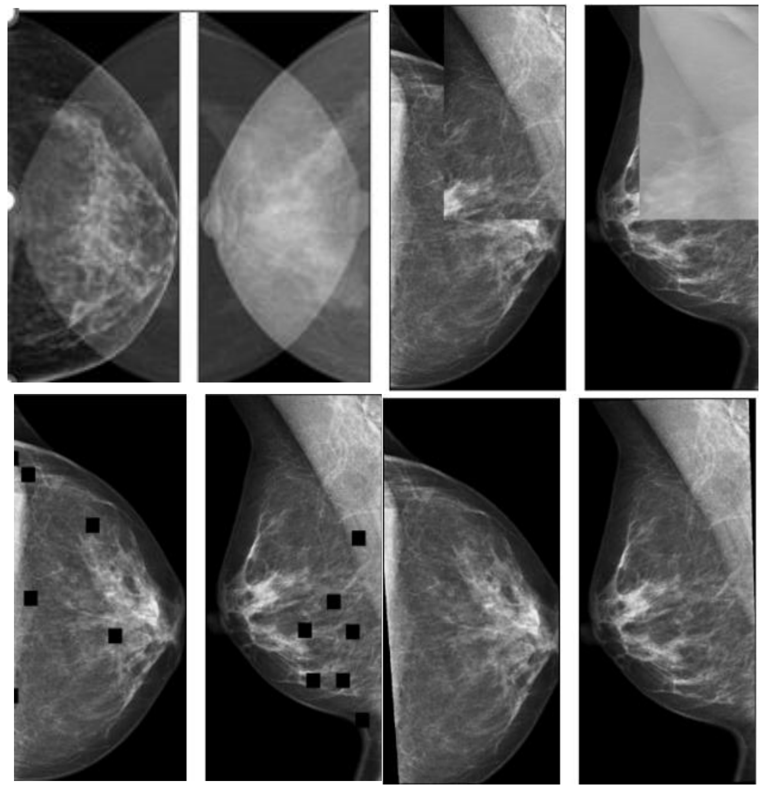
Images Augmentation

- **Random horizontal flips** → Simulates natural breast variations
- **Brightness/contrast adjustments** → Enhances robustness to imaging conditions
- **Hue/saturation modifications** → Improves tissue differentiation
- **Coarse dropout** → Prevents over-reliance on specific features
- **Mix-up augmentation** → Generates synthetic samples for better generalization

Original images



Augmentation Images



5TH BBCC BANGLADESH BREAST CANCER CONFERENCE 2025

METHODOLOGY

Models

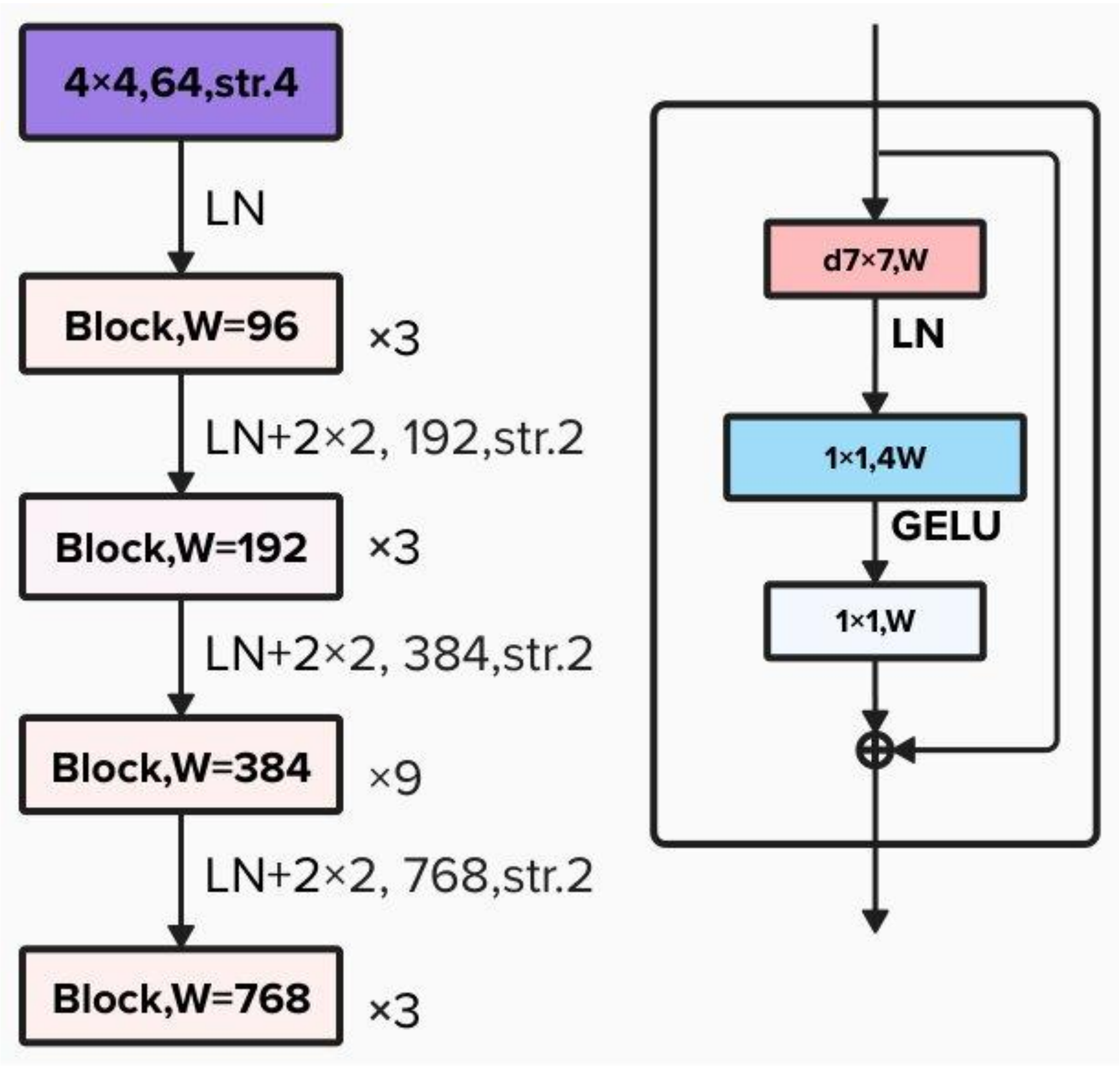


Fig 1:ConvNeXtV1-small

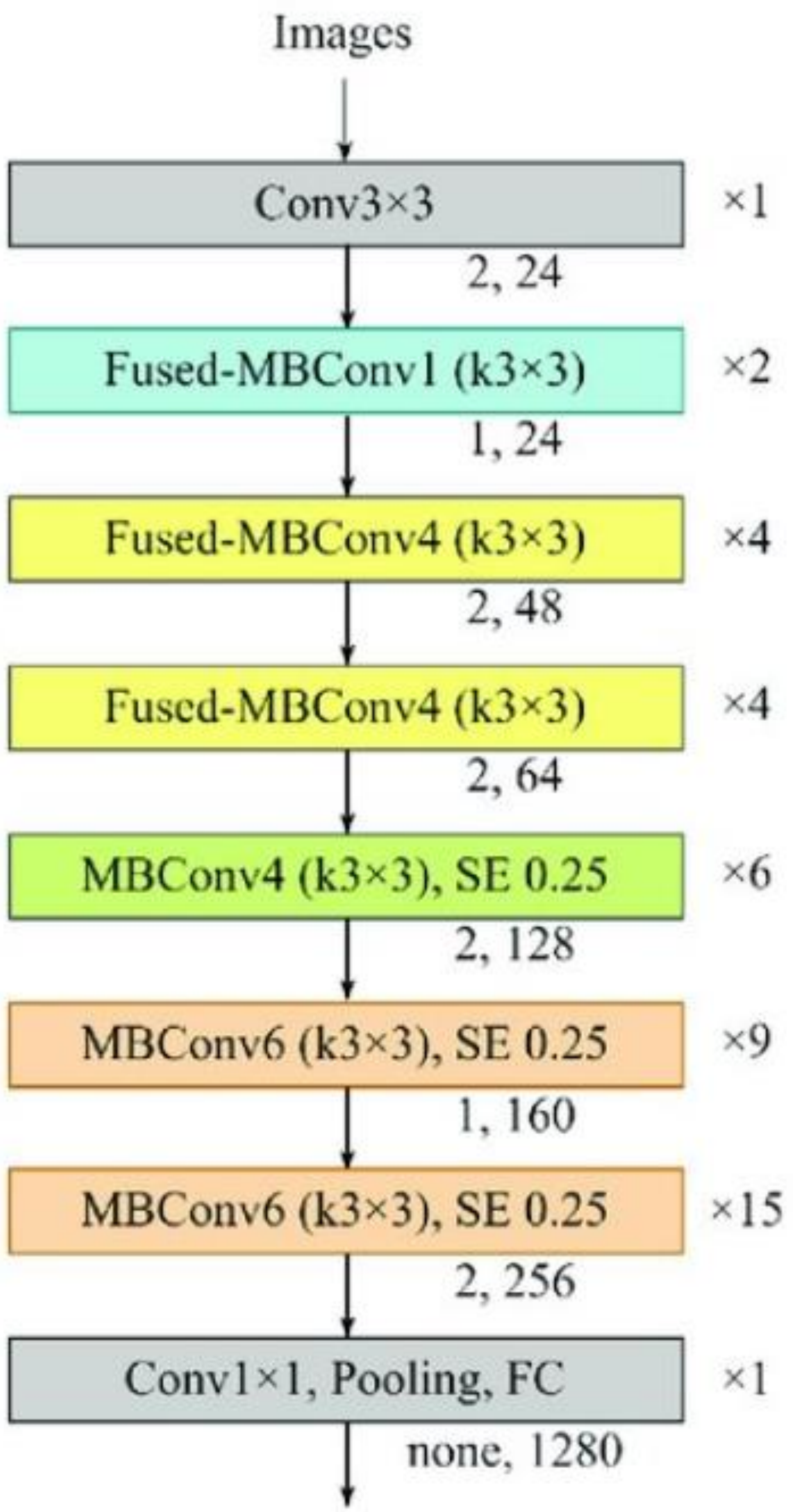


Fig 2: EfficientNetV2

METHODOLOGY

Models

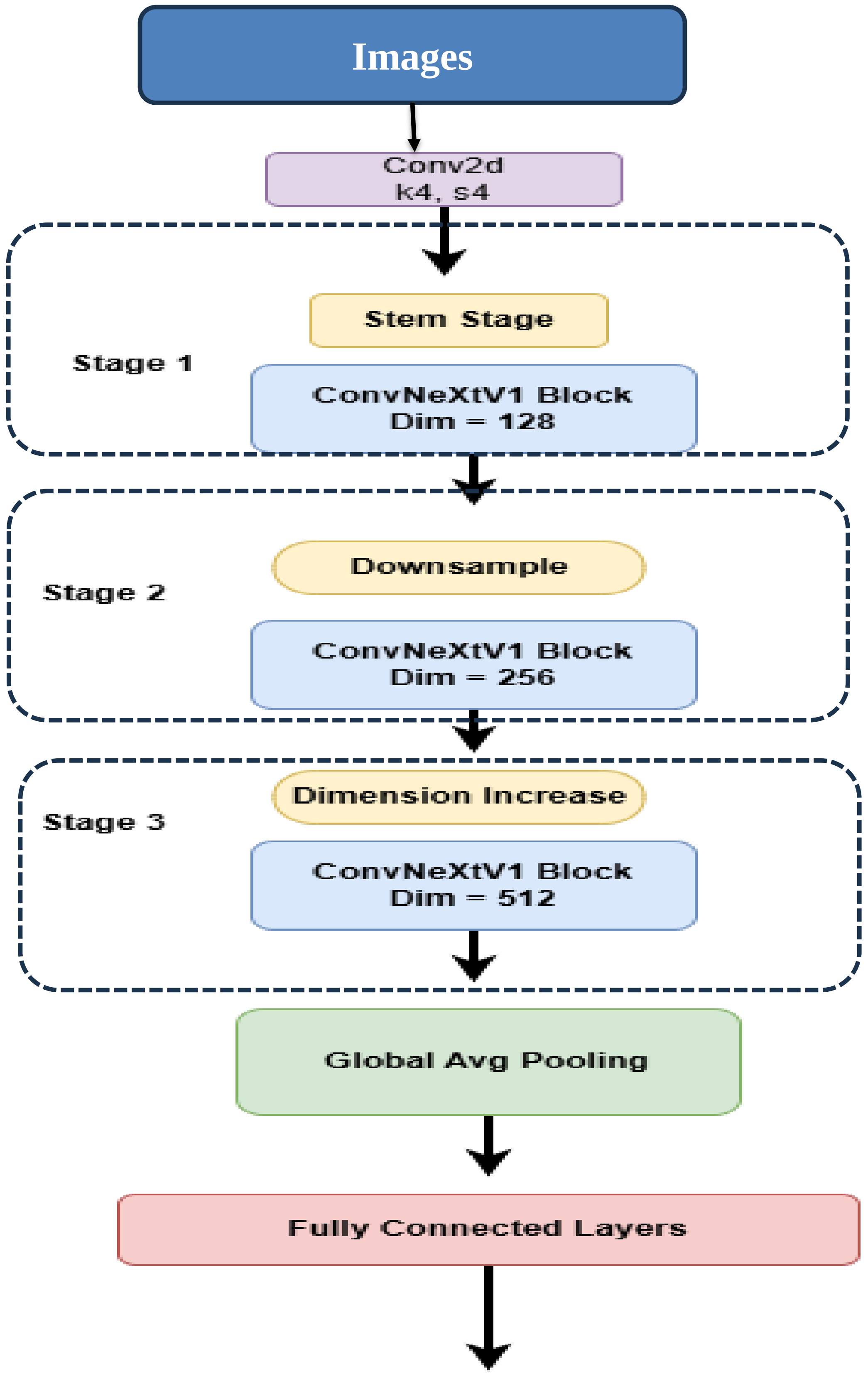


Fig. 1: Implementing ConvNeXtV1- small for advanced breast cancer classification.

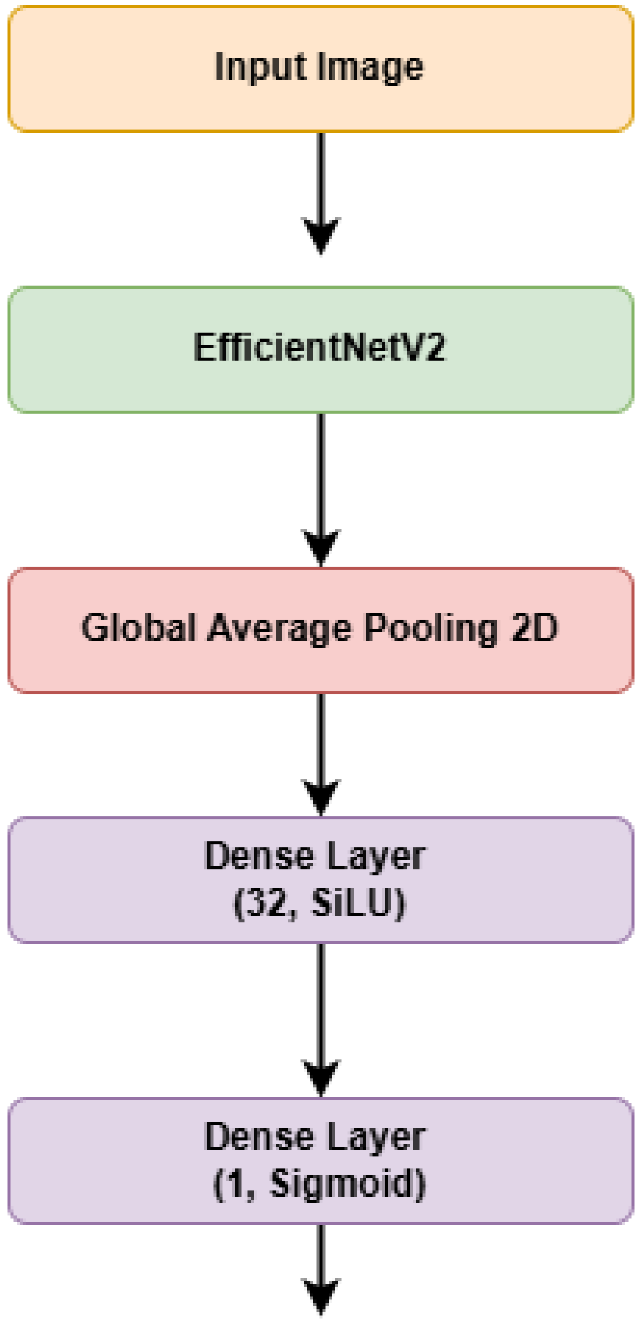
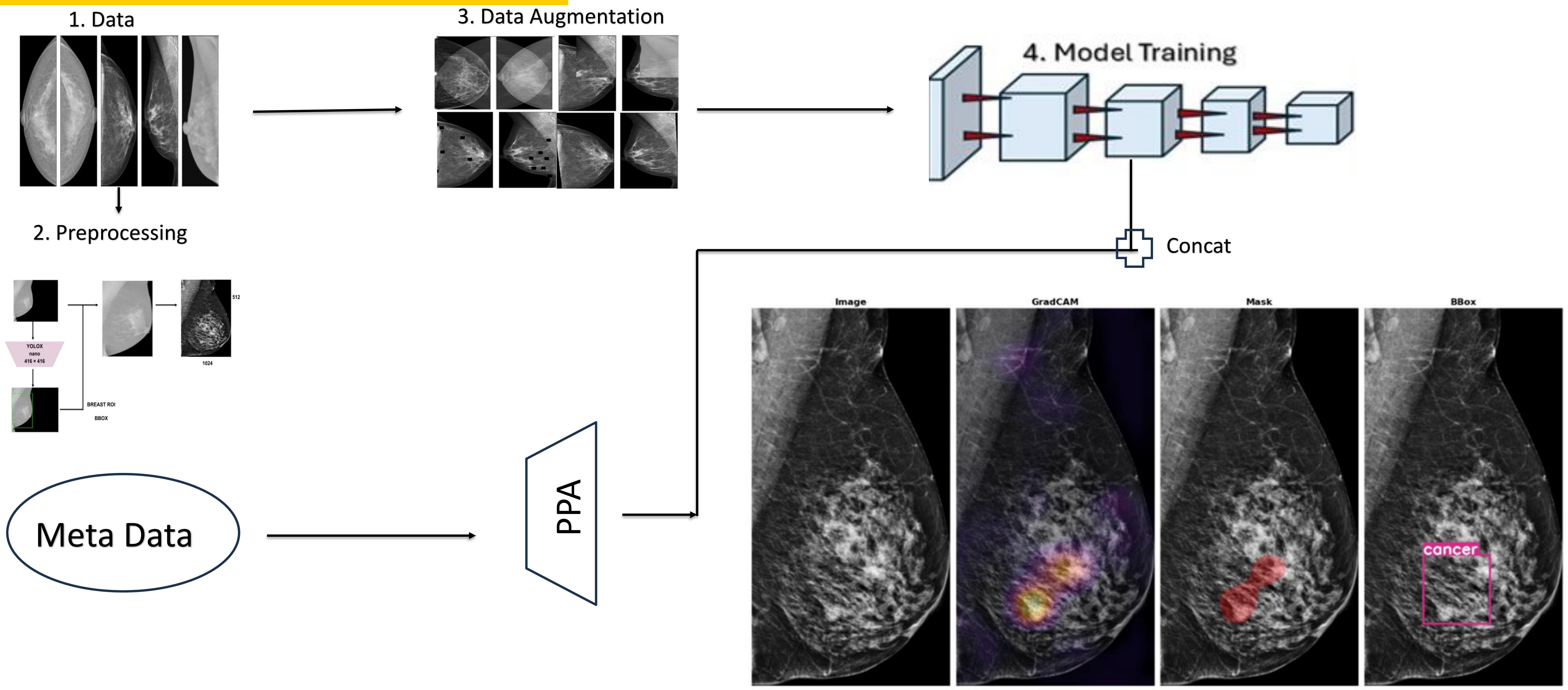


Fig. 2: Implementing EfficientNetV2 for advanced breast cancer classification.

Workflow



Grad cam and BBOX Detecting cancer with Interpretability

Results & Model Performance

Evaluation Metric: Probabilistic F1 (pF1) score

$$pF_1 = 2\frac{pPrecision \cdot pRecall}{pPrecision + pRecall}$$

$$pPrecision = \frac{pTP}{pTP + pFP}$$

$$pRecall = \frac{pTP}{TP + FN}$$

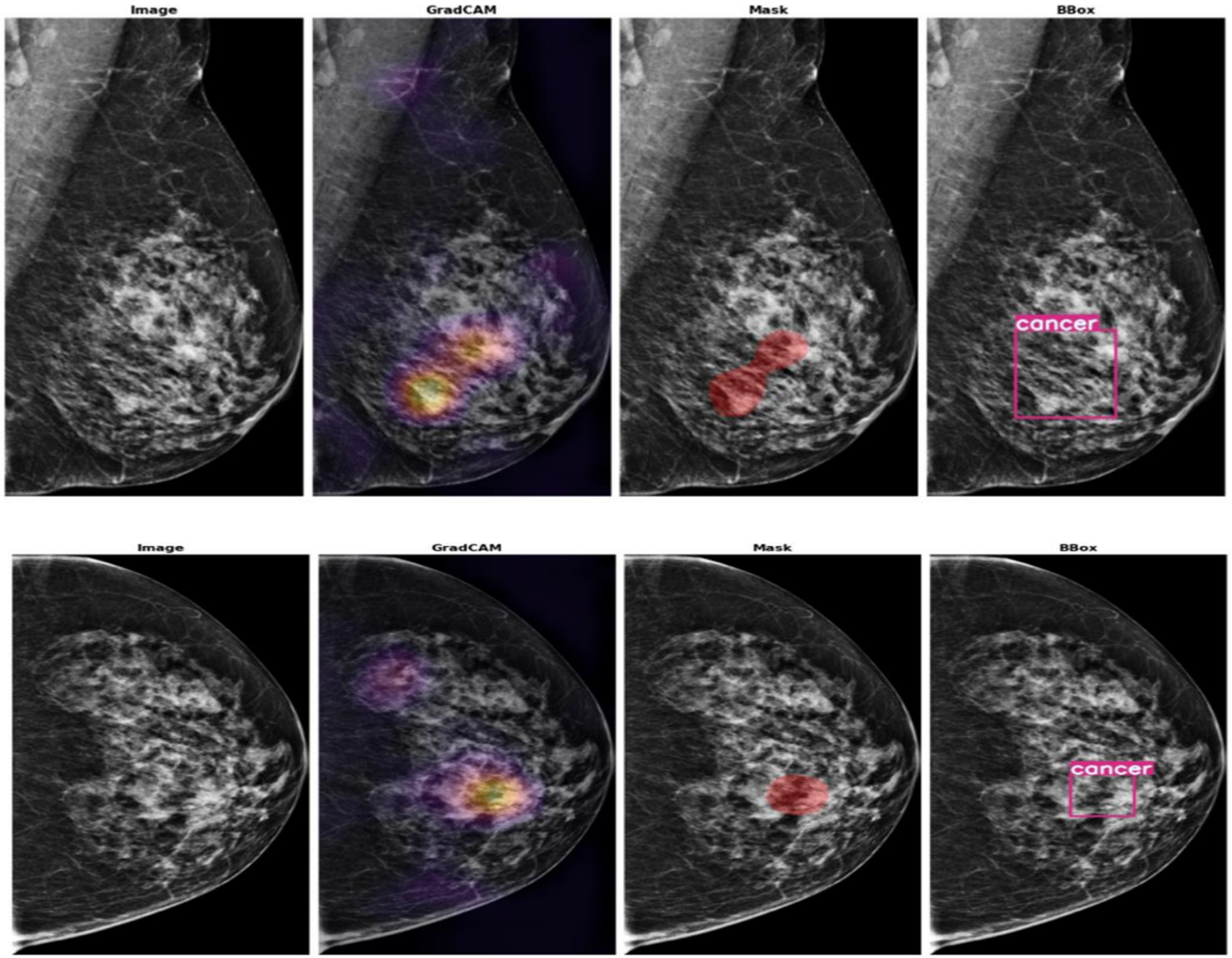
Results Table:

Model	pF1
ConvNeXtV1-small	0.558
EfficientNetV2	0.331

ConvNeXtV1-small **outperformed EfficientNetV2**, showing better feature extraction capabilities.

5TH BBCC BANGLADESH BREAST CANCER CONFERENCE 2025

Visual Results - Grad-CAM & Bounding Boxes



Findings:

1. **Malignant cases:** Strong Grad-CAM activations in tumor regions.
2. **Benign cases:** Highlighted dense breast tissue
3. Bounding boxes **refined lesion localization**, reducing false positives.

Challenges & Limitations

•**Class Imbalance:** Required advanced sampling and label smoothing to mitigate prediction bias.

•**Computational Constraints:** Limited TPU/GPU access hindered larger model experimentation.

•**Dataset Bias:** Potential geographic and ethnic variations may affect generalizability.

•**High-Configuration Needs:** Demanded powerful TPUs/multi-GPU systems for efficient training.

CONCLUSION

•Explainable AI techniques **enhance transparency and trust** in AI-driven breast cancer detection.

•Grad-CAM & bounding boxes **improve lesion localization** and align with radiological assessments.

•The study underscores **the necessity of interpretability in AI-based diagnostics**.

Future Work

- Explore **Transformer-based architectures** (Swin Transformer, Vision Transformer).

- Improve **data augmentation** techniques for robust feature learning.

- Collaborate with radiologists for **real-world validation** of model predictions.

REFERENCES

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 770–778.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2818–2826.
- Zhang, Q., Wu, Y. N., & Zhu, S. C. (2018). Interpretable convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 8827–8836.

5TH BBCC BANGLADESH BREAST CANCER CONFERENCE 2025

REFERENCES

- Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., & Zang, Y. (2019). Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognition, 88,* 385–398.
- RSNA Screening Mammography Breast Cancer Detection Dataset. (2023). Available at: https://www.kaggle.com/competitions/rsna-breast-cancer-detection
- Wu, J., Hicks, C., & Wang, H. (2019). Interpretable deep learning for automatic diagnosis in breast cancer. *Nature Machine Intelligence, 1(5),* 236–245.

THANK YOU