

# Two-Stage Mammographic Density Classification: Unsupervised Hierarchical Gabor Segmentation with Vision Transformer and ConvNeXt

PARVANEH ALINIYA, University of Nevada, Reno, NV 89557, USA

MIRCEA NICOLESCU, University of Nevada, Reno, NV 89557, USA

MONICA NICOLESCU, University of Nevada, Reno, NV 89557, USA

GEORGE BEBIS, University of Nevada, Reno, NV 89557, USA

Mammographic breast density (MBD) is a key determinant in screening mammography, influencing cancer detection, risk assessment, and treatment planning. This necessitates the development of a robust, unified model for MBD classification. We propose a hybrid two-stage framework in which the segmentation stage is fully unsupervised. First, hierarchical Gabor-based texture segmentation produces coarse and fine density maps without manual labels. Second, the original mammogram is concatenated with the refined Gabor maps to form a three-channel input to a classifier. This design yields an interpretable by-product—a density-oriented segmentation—alongside the final category prediction. We evaluate Vision Transformer (ViT) and ConvNeXt backbones on INbreast and observe consistent gains over single-image baselines; ConvNeXt, with the Gabor-augmented input, achieves the best overall performance.

Additional Key Words and Phrases: medical image classification, mammography, mammographic breast density, ConvNeXt, Vision Transformer, Gabor filters, deep learning, computer-aided diagnosis, unsupervised segmentation

## 1 INTRODUCTION

Mammography is the primary imaging modality for population screening for early cancer detection [24]. Image brightness reflects tissue composition; bright regions can correspond to fibroglandular tissue, pectoral muscle, or abnormalities, so reliable interpretation depends on both texture and intensity. Mammographic Breast Density (MBD) is defined as "a comparison of relative amounts of fat versus fibroglandular tissue in the breast" according to Breast Imaging Reporting and Data System (BI-RADS) fifth edition [23]. The BI-RADS 5th edition defines four density categories (ACR 1–4): almost entirely fatty (ACR 1), scattered fibroglandular (ACR 2), heterogeneously dense (ACR 3), and extremely dense (ACR 4). The visual assessment by radiologists is the primary source of evaluation, introducing interobserver variability to the process.

These considerations underscore the need for a unified system that estimates MBD reliably across heterogeneous cases. Accurate categorization is critical for three reasons: (i) MBD has been shown to correlate with breast cancer risk [2, 10, 13, 24]; therefore, it informs screening and treatment planning; (ii) higher density can obscure lesions and complicate interpretation, delaying detection and care; and (iii) density is a key input to modern automated CAD pipelines, where errors in MBD estimation can propagate and degrade overall performance.

Mammographic density classification has attracted substantial interest, yet performance remains limited by several factors. Heterogeneous acquisition parameters alter brightness and contrast, complicating inter-class discrimination and increasing intra-class variability. Pixel-level density segmentations are rarely available because they are time-consuming to obtain, which constrains supervised approaches. To address these issues, we propose a two-stage pipeline that combines an unsupervised segmentation module with a classifier.

---

Authors' addresses: Parvaneh Aliniya, paliniya@unr.edu, University of Nevada, Reno, NV 89557, USA, Virginia Street, Reno, Nevada, 89557; Mircea Nicolescu, University of Nevada, Reno, NV 89557, USA, Virginia Street, Reno, Nevada, 89557, mircea@unr.edu; Monica Nicolescu, University of Nevada, Reno, NV 89557, USA, Virginia Street, Reno, Nevada, 89557, monica@unr.edu; George Bebis, University of Nevada, Reno, NV 89557, USA, Virginia Street, Reno, Nevada, 89557, bebis@unr.edu.

The contributions of this paper are threefold: (1) We introduce a framework that leverages density segmentation to guide downstream classification. (2) We incorporate hierarchical (coarse and fine) density maps alongside the raw mammogram as input, introducing explicit texture signals. (3) We evaluate on the INbreast dataset using ViT and ConvNeXt backbones and include an ablation study to quantify design choices.

## 2 RELATED WORK

Prior work on mammographic density estimation—spanning segmentation and classification—falls broadly into volume-based [6, 10] and area-based [14, 22] approaches. Both align with the BI-RADS notion of density as the proportion of fibroglandular to fatty tissue, but they mainly differ in required inputs. Volume-based methods exploit raw (for-processing) mammograms and acquisition metadata (e.g., attenuation parameters) to estimate fibroglandular volume; however, such metadata are often unavailable in public datasets [21]. In contrast, area-based methods rely solely on the mammogram content to estimate the fibroglandular area, making them more widely applicable.

Traditional area-based methods [5, 26] estimate mammographic breast density (MBD) using handcrafted cues. Byng et al. [5] thresholded the intensity histogram to compute density, but dependence on brightness alone proved limiting; a later variant incorporated texture via fractal analysis [4], yet remained vulnerable to artifacts and lacked generalizability. Oliver et al. [19] first removed the pectoral muscle and then applied fuzzy C-means to a smoothed image to segment the breast, extracting morphological and texture features for classification. Overall, such hand-engineered pipelines struggle with acquisition variability and do not generalize reliably due to limited feature extraction.

Deep learning has substantially mitigated these limitations by providing more generalizable feature extractors [8]. Early work largely adapted CNNs to the task: Mohamed et al. [17] used a CNN [15] to classify two BI-RADS categories (ACR 2 and 3). A parallel line of research leveraged multiple views per exam. Ciritsis et al. [7] trained a customized CNN separately on MLO and CC views; Wu et al. [27] extracted features from all four standard images with a vanilla CNN and fused them before classification; and Busaleh et al. [3] proposed a two-view network (MLO and CC) with a fusion stage similar to [27].

Beyond CNN classifiers, several works explore representation learning with autoencoders and GANs. To enhance feature extraction, Kallenberg et al. [12] used a sparse autoencoder for *unsupervised* feature learning, followed by *supervised* density segmentation—thus still requiring ground-truth masks. Saffari et al. [21] employed a conditional GAN to generate MBD segmentations and then performed classification with traditional/CNN models; notably, the generated masks served as training “ground truth.”

Our approach is an area-based method that explicitly models texture via Gabor analysis. We pair an unsupervised hierarchical texture segmentation—capturing regional patterns at multiple granularity levels—with a classifier. The former produces density-sensitive maps; the latter provides strong feature extraction for the final prediction.

## 3 PROPOSED METHOD

The visual signal for MBD is expressed through subtle brightness and texture patterns in mammograms. Traditional area-based methods struggle to generalize, while modern deep models can lose fine-grained cues during preprocessing and downsampling— a costly trade-off in medical imaging with severe class imbalance. This challenge is acute for MBD, where the breast is largely unstructured and diagnostic cues reside in minute intensity variations. To bridge these gaps, we propose a hybrid framework: an unsupervised, texture-driven enhancement that injects density cues into the input, paired with a deep classifier that supplies robust representation learning. This design preserves fine detail, focuses attention on density-relevant regions, and mitigates information loss from image resizing.

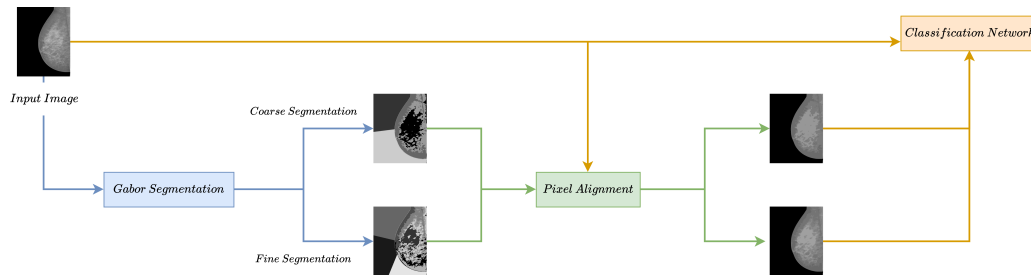


Fig. 1. Overview of the proposed method. Hierarchical Gabor segmentation produces coarse and fine density maps from the input.

To this end, we adopt a two-step pipeline—unsupervised segmentation followed by supervised classification. First, we apply Gabor filters [11] to derive texture-based segmentations that serve as proxies for density patterns (Fig. 1). To capture structure at multiple granularity levels, we produce coarse and fine segmentations mainly by varying the number of clusters in the Gabor-based method. Because raw density maps are not intensity-aware, a *pixel-refinement* module converts each map into an intensity-consistent image by assigning, to every pixel in a segment, the mean intensity of that segment in the original mammogram. This alignment preserves meaningful brightness cues and suppresses label artifacts. Finally, we concatenate the original image with the refined coarse and fine maps channel-wise and feed this three-channel input to the classifier.

## 4 METHOD EVALUATION

The proposed framework is classifier-agnostic and compatible with any image classifier that accepts three-channel input. For evaluation, the Vision Transformer [9] and ConvNeXt [16] methods, two baseline models in the transformer-based and convolution-based categories of models, have been selected as the classifiers. For both backbones, we use implementations from TIMM [25]. For segmentation, we adapt an enhanced Gabor segmentation implementation [20].

### 4.1 Dataset and Experimental Setup

For evaluation, we use the INbreast dataset [18], which contains 410 images; one image lacking an ACR category is excluded, yielding 409 images. During preprocessing, pixel intensities are normalized and images are resized. The data are split into training, validation, and test sets in a 70/10/20 ratio. The exact data split is provided in the project repository for reproducibility. All models were trained using AdamW optimizer with learning rate  $5 \times 10^{-5}$  and batch size 16. Images were resized to  $224 \times 224$ . We employed standard data augmentation including random horizontal flips, and rotations.

### 4.2 Experimental Results

We evaluate using Accuracy, Precision, Recall, F1, and Cohen’s kappa  $\kappa$ . We compare the method against a *single-image* baseline that replicates the mammogram across three channels (listed as ViT and ConvNeXt in Table 1). Our proposed approach, which augments the input with refined Gabor maps, is denoted ViT-G and ConvNeXt-G. As summarized in Table 1, the proposed method consistently outperforms the baseline across all metrics for both backbones. While our approach exceeds the baseline for both architectures, the gains are larger with ConvNeXt. The confusion matrices in Fig. 2 corroborate this: the proposed method (middle column) achieves the strongest overall performance, with the ConvNeXt variant showing the greatest robustness.

Table 1. Comparison of the proposed method with baseline using ViT and ConvNeXt. Suffix -G denotes the proposed Gabor-augmented method, while -G-W indicates the Gabor method with pectoral muscle removal.

Method	Accuracy	Cohen’s $\kappa$	Precision	Recall	F1	ROC-AUC
ViT	67.07	0.5256	70.80	67.07	66.20	0.8555
<b>ViT-G (ours)</b>	68.29	0.5549	69.71	68.29	67.03	0.8872
ViT-G-W	65.85	0.5223	70.13	65.85	65.28	0.8721
ConvNeXt	69.51	0.5606	68.44	69.51	68.50	0.8749
<b>ConvNeXt-G (ours)</b>	<b>74.39</b>	<b>0.6304</b>	<b>74.20</b>	<b>74.39</b>	<b>73.98</b>	<b>0.9044</b>
ConvNeXt-G-W	71.95	0.5972	72.18	71.95	70.70	0.8800

Pectoral muscle and dense fibroglandular tissue both appear as bright regions in mammograms, raising concerns that networks might confuse these structures. To test this, we removed the pectoral muscle from density maps using ground-truth segmentation masks [1]. Unexpectedly, removal degraded performance for both architectures (listed as ViT-G-W and ConvNeXt-G-W in Table 1). A likely explanation could be loss of information as the pectoral muscle provides valuable anatomical context.



Fig. 2. Comparison of the confusion matrices of the proposed method with baseline using ViT and ConvNeXt. Confusion matrices for (left) baseline, (middle) proposed method (Gabor), and (right) pectoral muscle removed variant (Gabor-W).

## 5 DISCUSSION

This work proposes and evaluates a two-stage classification pipeline in which *unsupervised* Gabor segmentation extracts density cues at two granularity levels to aid the final classifier. We deliberately keep the overall design simple to isolate and evaluate the contribution of the Gabor prior. As shown in our experiments, the proposed approach yields substantial gains—especially with the ConvNeXt backbone—demonstrating that a simple, texture-aware augmentation can be highly effective. The idea is generic and can be readily adapted to related tasks in medical imaging and beyond.

## REFERENCES

- [1] Parvaneh Aliniya, Mircea Nicolescu, Monica Nicolescu, and George Bebis. 2024. Towards Robust Supervised Pectoral Muscle Segmentation in Mammography Images. *Journal of Imaging* 10, 12 (2024), 331.
- [2] Norman F Boyd, Lisa J Martin, Michael Bronskill, Martin J Yaffe, Neb Duric, and Salomon Minkin. 2010. Breast tissue composition and susceptibility to breast cancer. *Journal of the National Cancer Institute* 102, 16 (2010), 1224–1237.
- [3] Mariam Busaleh, Muhammad Hussain, Hatim A Aboalsamh, and Sarah A Al Sultan. 2022. TwoViewDensityNet: two-view mammographic breast density classification based on deep convolutional neural network. *Mathematics* 10, 23 (2022), 4610.
- [4] JW Byng, NF Boyd, E Fishell, RA Jong, and MJ Yaffe. 1996. Automated analysis of mammographic densities. *Physics in Medicine & Biology* 41, 5 (1996), 909.
- [5] Jeffrey William Byng, NF Boyd, E Fishell, RA Jong, and Martin J Yaffe. 1994. The quantitative analysis of mammographic densities. *Physics in Medicine & Biology* 39, 10 (1994), 1629.
- [6] Stefano Ciatto, Daniela Bernardi, Massimo Calabrese, Manuela Durando, Maria Adalgisa Gentilini, Giovanna Mariscotti, Francesco Monetti, Enrica Moriconi, Barbara Pesce, Antonella Roselli, et al. 2012. A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification. *The Breast* 21, 4 (2012), 503–506.
- [7] Alexander Ciritis, Cristina Rossi, Ilaria Vittoria De Martini, Matthias Eberhard, Magda Marcon, Anton S Becker, Nicole Berger, and Andreas Boss. 2019. Determination of mammographic breast density using a deep convolutional neural network. *The British journal of radiology* 92, 1093 (2019), 20180691.
- [8] Jian Deng, Yanyun Ma, Deng-ao Li, Jumin Zhao, Yi Liu, and Hui Zhang. 2020. Classification of breast density categories based on SE-Attention neural networks. *Computer Methods and Programs in Biomedicine* 193 (2020), 105489.
- [9] Alexey Dosovitskiy. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020).
- [10] Ralph Highnam, Sir Michael Brady, Martin J Yaffe, Nico Karssemeijer, and Jennifer Harvey. 2010. Robust breast composition measurement-VolparaTM. In *International workshop on digital mammography*. Springer, 342–349.
- [11] Anil K Jain and Farshid Farrokhnia. 1991. Unsupervised texture segmentation using Gabor filters. *Pattern recognition* 24, 12 (1991), 1167–1186.
- [12] Michiel Kallenberg, Kersten Petersen, Mads Nielsen, Andrew Y Ng, Pengfei Diao, Christian Igel, Celine M Vachon, Katharina Holland, Rikke Rass Winkel, Nico Karssemeijer, et al. 2016. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE transactions on medical imaging* 35, 5 (2016), 1322–1331.
- [13] Marzena Kamińska, Tomasz Ciszewski, Karolina Łopacka-Szatan, Paweł Miotła, and Elżbieta Starosławska. 2015. Breast cancer risk factors. *Menopause Review/Przegląd Menopauzalny* 14, 3 (2015), 196–202.
- [14] Brad M Keller, Diane L Nathan, Yan Wang, Yuanjie Zheng, James C Gee, Emily F Conant, and Despina Kontos. 2012. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. *Medical physics* 39, 8 (2012), 4903–4917.
- [15] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *nature* 521, 7553 (2015), 436–444.
- [16] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. 2022. A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 11976–11986.
- [17] Aly A Mohamed, Wendie A Berg, Hong Peng, Yahong Luo, Rachel C Jankowitz, and Shandong Wu. 2018. A deep learning method for classifying mammographic breast density categories. *Medical physics* 45, 1 (2018), 314–321.
- [18] Inês C Moreira, Igor Amaral, Inês Domingues, António Cardoso, Maria Joao Cardoso, and Jaime S Cardoso. 2012. Inbreast: toward a full-field digital mammographic database. *Academic radiology* 19, 2 (2012), 236–248.
- [19] Arnau Oliver, Jordi Freixenet, Robert Marti, Josep Pont, Elsa Pérez, Erika RE Denton, and Reyer Zwiggelaar. 2008. A novel breast tissue density classification methodology. *Ieee transactions on information technology in biomedicine* 12, 1 (2008), 55–65.
- [20] <Owner or Org>. [n.d.]. *Unsupervised-Texture-Segmentation-Using-Gabor-Filter*. <https://github.com/sayanmandal/Unsupervised-Texture-Segmentation-Using-Gabor-Filter>
- [21] Nasibeh Saffari, Hatem A Rashwan, Mohamed Abdel-Nasser, Vivek Kumar Singh, Meritxell Arenas, Eleni Mangina, Blas Herrera, and Domenech Puig. 2020. Fully automated breast density segmentation and classification using deep learning. *Diagnostics* 10, 11 (2020), 988.
- [22] Radhika Sivaramakrishna, Nancy A Obuchowski, William A Chilcote, and Kimerly A Powell. 2001. Automatic segmentation of mammographic density. *Academic radiology* 8, 3 (2001), 250–256.
- [23] David Allen Spak, JS Plaxco, L Santiago, MJ Dryden, and BE Dogan. 2017. BI-RADS® fifth edition: A summary of changes. *Diagnostic and interventional imaging* 98, 3 (2017), 179–190.
- [24] Bolette Mikela Vilmun, Ilse Vejborg, Elsebeth Lynge, Martin Lillholm, Mads Nielsen, Michael Bachmann Nielsen, and Jonathan Frederik Carlsen. 2020. Impact of adding breast density to breast cancer risk models: a systematic review. *European journal of radiology* 127 (2020), 109019.
- [25] Ross Wightman. 2019. PyTorch Image Models. <https://github.com/rwightman/pytorch-image-models>. <https://doi.org/10.5281/zenodo.4414861>
- [26] John N Wolfe. 1976. Breast patterns as an index of risk for developing breast cancer. *American Journal of Roentgenology* 126, 6 (1976), 1130–1137.
- [27] Nan Wu, Krzysztof J Geras, Yiqiu Shen, Jingyi Su, S Gene Kim, Eric Kim, Stacey Wolfson, Linda Moy, and Kyunghyun Cho. 2018. Breast density classification with deep convolutional neural networks. In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 6682–6686.