Intelligent Multimedia Processing & Communication Systems Journal

IMPCSJ

J IMPCS (2025) 20: 1-9
DOI 10.71856/IMPCS.2025.1208735

Research Paper

Breast Cancer Tumor Analysis with an Approach to Overcome the Overfitting Problem in Small Training Dataset by Combining Transfer Learning and Adversarial Generative Networks

Zeinab Delshad[1], Salman Karimi[2*]

1. Ph.D. Student, Department of Electronics Engineering, Faculty of Engineering, Lorestan University, Lorestan, Iran

2. Associate Professor, Department of Electronics Engineering, Faculty of Engineering, Lorestan University, Lorestan, Iran * *Corresponding Author,* karimi.salman@lu.ac.ir

Article Info

Article history:
Received: 3 Apr 2025
Accepted: 14 May 2025
DOR:

Keywords:
augmentation,
Breast cancer,
Data Neural network,
Overfitting,
Synthetic data.

ABSTRACT

This research pioneers an innovative and genuinely promising methodology to tackle the pervasive challenge of overfitting in the domain of breast cancer tumor analysis, a common obstacle, particularly when confronted with limited datasets. The central aim is to significantly enhance diagnostic accuracy and fortify overall model robustness. This is achieved through a synergistic integration of transfer learning principles with the sophisticated data augmentation capabilities afforded by Deep Convolutional Generative Adversarial Networks (DCGANs). The fundamental problem is the scarcity of training samples, a frequent predicament in medical imaging. This scarcity can lead to models that excel on familiar training data but falter significantly when encountering new, unseen patient cases, undermining their clinical utility. To surmount this, the researchers judiciously utilized the standard MIAS (Mammographic Image Analysis Society) database. A DCGAN architecture, known for its proficiency in generating realistic images by pitting a generator network against a discriminator network, was employed. This network learns the underlying patterns and distributions of the original data to produce synthetic mammographic images. This process effectively expanded the training pool, resulting in the creation of 10,000 high-quality synthetic data points. Crucially, these synthetic images were designed to realistically mimic the complex and often subtle characteristics inherent in actual breast tumor images found within the MIAS dataset, ensuring they contribute meaningfully to model training. These newly generated synthetic samples, combined with the limited original MIAS data, formed an augmented dataset. This enriched dataset was then used to train the YOLOV11m neural network architecture. The application of transfer learning was pivotal here. This technique allowed the YOLOV11m model to benefit from knowledge pre-acquired from training on larger, more general datasets, significantly enhancing its learning efficiency and overall performance. This is especially critical when the original domain-specific dataset (like MIAS) is small, as it provides a robust foundational understanding of visual features. The experimental results compellingly demonstrated the remarkable efficacy of this integrated methodology. The YOLOV11m model, when trained on this augmented dataset, achieved an impressive 99.1% accuracy in distinguishing between benign and malignant tumors.

I. Introduction

Breast cancer, as one of the most common cancers among women, claims millions of lives every year. Early diagnosis of this disease plays a key role in increasing patient survival rates, and mammography images are the main screening tool and early detection of breast tumors [1]. However, accurate and automated analysis of mammography images is currently facing serious challenges Including lack of high-quality and balanced data, and overfitting of deep learning models due to limited educational examples. Overfitting occurs when the model becomes dependent on specific details of the training data instead of learning general features, resulting in poor performance on new data [2]. In recent years, several methods have been proposed to improve the detection of breast tumors using deep learning. Some researchers have attempted to extract and classify important features of images using Convolutional networks (CNN). However, these models usually require a large amount of labeled data and suffer from overfitting when data is scarce. In some studies, traditional data augmentation techniques such as image rotation and translation have been used, but these methods do not provide enough diversity to overcome the lack of data [3]. also, there have been attempts to use transfer learning with pre-trained models, which, although it somewhat alleviates the problem of data scarcity, still do not perform well when faced with very limited or unbalanced data. One of the important advances in this field is the use of adversarial generative networks (GAN), especially DCGAN, to generate realistic synthetic data [4]. These networks are able to produce diverse and high-quality examples of medical images, thus expanding the training dataset. However, simply using synthetic data without combining it with other methods may not lead to sufficient improvement in model generalizability.

In the present study, a hybrid approach based on the integration of transfer learning and adversarial generative networks is presented to simultaneously overfitting and data shortage problems in breast tumor analysis should be addressed. In this method, various synthetic data are first

generated using DCGAN and added to the original dataset. Then, using transfer learning, the knowledge of pre-trained models is transferred to the target model to reduce the need for large dataset and increase the model's accuracy in diagnosing benign and malignant tumors. Experimental results on the MIAS dataset show that this approach achieves 99.1% accuracy in tumor classification and effectively reduce the overfitting problem.

Overall, this research, by presenting a novel and hybrid approach, has taken an effective step towards improving the automatic diagnosis of breast cancer and developing intelligent medical systems, and can be used as a model for other areas of medical imaging with limited data.

To address the aforementioned challenges, this paper proposes a hybrid methodology that synergizes Deep Convolutional Generative Adversarial Networks (DCGANs) with the state-of-the-art YOLOv11m architecture. While previous works have combined GANs with other models, our research introduces specific novel contributions:

A Novel Architectural Pairing: This is one of the first studies to specifically pair the powerful YOLOv11m object detection model with DCGANs for breast cancer analysis. This combination leverages the advanced feature extraction capabilities of YOLOv11m for detecting subtle tumor characteristics, which are enhanced by the high-fidelity synthetic images generated by the DCGAN.

Targeted Data Augmentation Strategy: We present a structured data generation process, creating a large-scale, balanced dataset of 10,000 synthetic images. Unlike generic augmentation, our approach involves training class-specific DCGANs (one for benign and one for malignant tumors) to ensure accurate labeling and high relevance of the synthetic data, directly tackling the class imbalance problem in the original MIAS dataset.

Detailed Implementation and Reproducibility: We provide clear architectural details, including flowcharts and preprocessing steps, to ensure the reproducibility of our results. This includes a transparent explanation of how synthetic images are generated, labeled, and integrated into the training pipeline, a detail often overlooked in similar research.

This integrated approach not only overcomes the overfitting problem but also achieves a remarkable 99.1% accuracy, setting a new benchmark for classification in small medical imaging datasets.

This integrated approach not only overcomes the overfitting problem but also achieves a remarkable 99.1% accuracy, setting a new benchmark for classification in small medical imaging datasets.

II. A Review of Related Research

In recent years, researchers have made extensive efforts to improve breast tumor detection using deep learning and have solved the problem of overfitting in small dataset. For example, Dhungel et al [5] achieved 94.3% accuracy using Convolutional networks (CNN) and traditional data augmentation on the IN-breast

dataset, but their model was still dependent on large amounts of data. Al-antari et al [6] presented a fast and accurate approach for breast lesion detection using the YOLOv3 architecture on MIAS and DDSM data and reported an accuracy of 96.7% However, these models suffer from performance degradation when faced with limited data.

To overcome this limitation, Jiang et al [7] combined generative adversarial networks (GAN) and CNN to generate realistic synthetic images and achieved 97.2% accuracy on MIAS data. Singh et al [8] also, succeeded in increasing the generalizability of the model and achieving an accuracy of 98.1% by combining transfer learning (ResNet50model) and GAN-based data augmentation. Zhu et al [9] reported an accuracy of 97.8% using DCGAN to generate synthetic data and train a classifier, and showed that increasing data diversity can greatly reduce the overfitting problem.

In other studies, such as Ragab et al [10], combining CNN and transfer learning with the VGG16 model on MIAS and BCDR data achieved an accuracy of 96.4%. Zhang et al [11] gained accuracy of 98.5% with a combination of GAN and CNN and emphasized the importance of deep data augmentation. Alom et al [12] reported an accuracy of 97.3% with the Inception-ResNet model and data augmentation. Salama et al [13] gained an accuracy of 98.7% with a combination of GAN and DenseNet.

Shen et al [14] recorded an accuracy of 98.9% using EfficientNet and GAN-based data augmentation.

These studies show that using generative adversarial networks to generate synthetic data significantly increases the accuracy and generalizability of the adopted models. However, it is still difficult to achieve higher accuracy and completely eliminate overfitting in very limited data. Therefore, this study attempts to provide a more appropriate answer than previous articles.

III. Types of breast cancer classes

Tumors come in shapes, including: round, oval, lobulated, nodular, stellate, and irregular. Tumors that have more irregular shapes are malignant, while tumors with round, regular, and smooth borders are in the benign stage [15]. Therefore, a mammogram image that contains a benign tumor is not cancerous, while an image of a malignant tumor is cancerous. The difference between benign and malignant tumors [15] is shown in Figure 1.

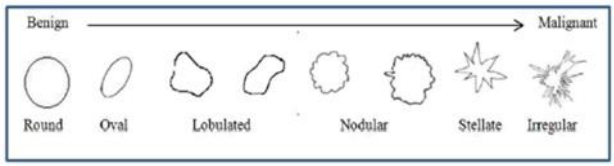

Fig 1: Types of breast cancer classes and their division into benign and malignant [15]

IV. Introducing the MIAS dataset

MIAS (Mammographic Image Analysis Society) dataset is one of the most reliable and widely used databases in the field of mammography image analysis for breast cancer diagnosis. This collection was produced by the British Mammography Image Analysis Society and it is widely used in scientific researches to develop and evaluate automated breast tumors detection algorithms [16]. The MIAS dataset contains 322 high-quality digital mammography images from 161 female patients. Each image is saved in PGM format and provides researchers with information such as the location, type and size of the lesion (mass or classification), along with labels of whether the lesion is benign or malignant. The images in this collection are also categorized by breast tissue density, covering examples of breasts with depleted fat, medium density, and high density.

In this collection, tumors are divided into different categories based on their appearance, such as well-defined masses, ill-defined masses, spinous masses, classifications, and tissue abnormalities. This diversity allows for accurate evaluation of detection algorithms. Also, images of normal samples, benign and malignant tumors are separately present in the dataset so that machine learning models can learn key differences between these groups [17].

V. Comparison of GAN types and the advantage of DCGAN in data generation.

Generative adversarial networks (GANS) are a family of deep learning models used to generate realistic synthetic data. Various types of GAN architectures have been developed, each with its own characteristics and applications. Some of the most important types of GANs include:

Vanilla GAN: The simplest type of GAN with two networks: generator, and discriminator, both of which consist of fully connected layers. This type is not suitable for complex images [18].

(Deep Convolutional GAN) DCGAN: It uses deep convolutional layers in both the generator and discriminator sections and is very effective for producing high-quality and realistic images, especially medical images [19].

Conditional GAN (CGAN): Allows for conditional data generation based on specific labels or features (e. g., generating benign or malignant tumor images) [20].

Wasserstein GAN (WGAN): By changing the cost function, it increases the stability of training and the quality of the produced images, and reduces the instability problem in conventional GANs [21].

CycleGAN: Used to convert image styles (e. g. converting day to night images or MRI to CT scans) [22].

StyleGAN: Used to produce highly realistic images with precise control over image style and characteristics [23].

Pix2Pix: Used for image-to-image translation (such as converting a sketch to a real photo) [24].

In the Breast Tumor Analysis Project, the goal is to generate realistic synthetic mammography images to increase the volume and diversity of training data.

Due to its convolutional structure, DCGAN is particularly suitable for producing images with fine details and preserving spatial features. This architecture has higher training stability than Vanilla GAN and its produced images are visually closer to real examples. DCGAN is also simpler and more efficient to implement than more advanced models such as StyleGAN for medical applications and relatively limited data. This model reduces the risk of instability and model collapse by using BatchNorm layers and appropriate activation functions [25] and for this reason, it has been selected in the present study for data augmentation and combating overfitting.

VI. The benefits of YOLO in cancer detection and its difference with traditional CNNs

The You Only Look Once (YOLO) network has become one of the most popular models for detecting and diagnosing cancerous tumors in medical images due to its specific architecture and unique advantages. Unlike traditional convolutional networks (CNN) that are typically used for image classification or feature extraction, YOLO is an object detection model. This means it can simultaneously determine the exact location of the tumor (Bounding Box) and its type (classification) in the image [26].

Advantages of YOLO over other CNNs

1.2 simultaneous diagnosis of tumor location and type

YOLO divides the image into small segments and assigns box coordinates and class to each segment at the same time. This feature is crucial for mammography and pathology images where tumors may appear anywhere.

1.3 very high speed

The YOLO architecture is designed to perform the entire identification and classification process in a single step. This makes YOLO very suitable for real-time applications such as rapid screening of medical images.

1.4 Remarkable accuracy in detecting small objects

New versions of YOLO, such as YOLOv11, have improved the architecture, increasing the ability to detect small tumors and fine details, while many conventional CNNs are weak in this area.

1.5 High generalizability and flexibility

YOLO is easily adaptable to medical data with different classes (e. g. benign and malignant tumor types) and can identify multiple tumor types in a single image.

1.6 Advanced evaluation criteria

YOLO uses metrics such as mAP (mean accuracy), IOU (interaction to unity ratio), and confusion matrix to evaluate the model performance, which are very important for medical applications [27].

Limitations of traditional CNNs:

Conventional CNNs are mainly used for whole image classification and do not specify the exact location of the tumor. To identify tumor location, more complex architectures such as R-CNN or Faster R-CNN are required which have a slower speed than YOLO.

They are more difficult to implement and train for real-time applications.

The use of YOLO in the diagnosis of various types of cancer, especially in mammography and pathology images, has a significant advantage over traditional CNNs due to its speed, accuracy, ability to simultaneously identify multiple tumors, and determine their precise location. These features make YOLO the first choice for automated cancer detection projects in medical images, especially when large data volumes and processing speed are of high importance.

VII. Proposed Methodology

1.7 Overall Framework

The overall framework of our proposed methodology is illustrated in Figure 2. The process begins with the MIAS dataset, which undergoes initial preprocessing. To address data scarcity, we employ a class-specific DCGAN to generate a large volume of synthetic mammogram images. These synthetic images are then labeled and combined with the original dataset to form an

augmented training set. Finally, this enriched dataset is used to train a YOLOv11m model, which leverages transfer learning for robust tumor detection and classification. Each step of this framework is detailed in the following sections.

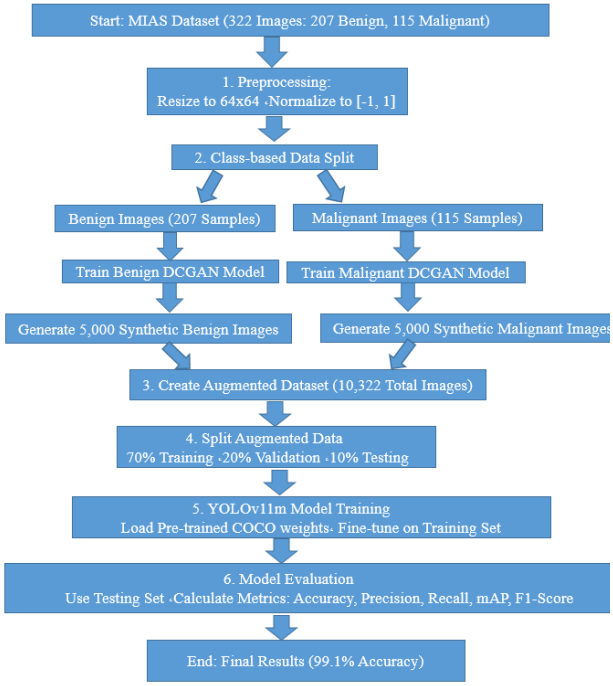

Fig 2: Framework of our proposed methodology

7.2 MIAS Dataset and Preprocessing

This study utilizes the Mammographic Image Analysis Society (MIAS) dataset, a standard benchmark for breast cancer research. The dataset contains 322 digital mammograms from 161 patients. Each image is provided with information about the location and class of the abnormality, categorized as either benign or malignant. The original dataset consists of 207 benign and 115 malignant cases, exhibiting a notable class imbalance.

Prior to training, all images (both original and synthetic) underwent a standardized preprocessing pipeline:

Resizing: All images were resized to 64x64 pixels to ensure consistent input dimensions for the DCGAN architecture.

Normalization: Pixel values were normalized to the range [-1, 1] by the formula (pixel - 127.5) / 127.5. This step is crucial for stabilizing the training of the GAN.

Tensor Conversion: The processed images were converted into PyTorch tensors for efficient processing on the GPU.

3.3. Synthetic Data Generation using DCGAN

To overcome the limitations of the small and imbalanced MIAS dataset, we employed a Deep Convolutional Generative Adversarial Network (DCGAN) to generate high-quality synthetic images.

1.8 7,2,1 DCGAN Architecture

Our DCGAN consists of two competing networks: a Generator and a Discriminator.

The Generator: Takes a 100-dimensional random noise vector as input and upsamples it through a series of ConvTranspose2d layers to produce a 64x64 pixel image. Each convolutional layer is followed by a BatchNorm layer and a ReLU activation function, except for the output layer which uses a Tanh function to scale the output to the [-1, 1] range.

The Discriminator: Takes a 64×64 image (either real or synthetic) as input and processes it through a series of Conv2d layers. Each layer uses a LeakyReLU activation function and BatchNorm to extract features and stabilize training. The final layer is a fully connected layer with a Sigmoid activation function, which outputs a single probability score indicating whether the input image is real or fake.

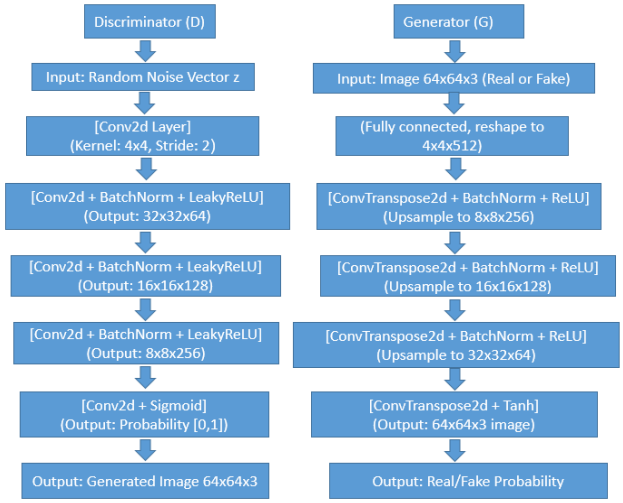

Fig 3: DCGAN Architecture [18]

7.2.2. Training and Labeling Process

A critical aspect of our methodology is the generation of accurately labeled synthetic data. To achieve this, we trained two separate DCGAN models:

- One DCGAN was trained exclusively on the 207 benign images from the MIAS dataset.
- A second DCGAN was trained exclusively on the 115 malignant images.

This class-specific training ensures that all images produced by the first GAN are inherently benign, and all

images from the second are malignant. The networks were trained for 50 epochs with a learning rate of 0.0002. After training, we used the trained generators to produce 5,000 synthetic benign images and 5,000 synthetic malignant images. This process resulted in a new, balanced, and significantly larger dataset, as detailed in Table 1.

Table 1: Composition of the Augmented Dataset

Data Type	Benign Class	Malignant Class	Total
Original MIAS	207	115	322
Synthetic (DCGAN)	5,000	5,000	10,000
Total Augmented	5,207	5,115	10,322

7.3.2. Transfer Learning and Training

To accelerate learning and improve generalization, we employed transfer learning. The YOLOv11m model was initialized with weights pre-trained on the large-scale COCO dataset. This allows the model to leverage general visual knowledge before fine-tuning on our specific medical imaging task.

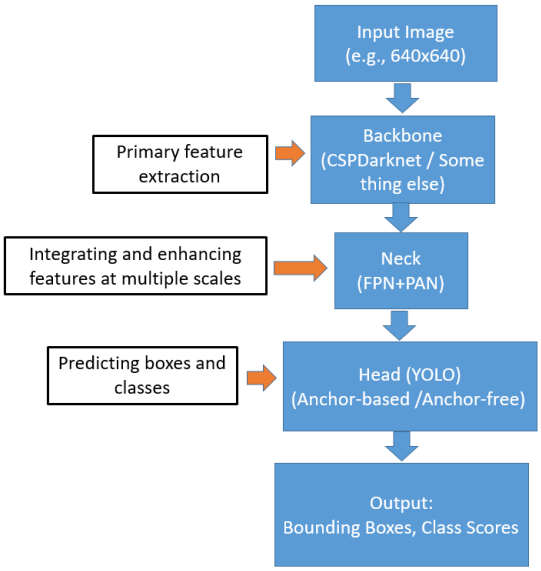

Fig 4**:** YOLOV11m Architecture [29]

The model was then trained on our augmented breast cancer dataset for 100 epochs. We used an image size of 1280 pixels and a batch size of 4. The training process optimized the model's ability to distinguish between benign and malignant tumors, and the final performance was evaluated on the held-out test set.

At the end of the training, the evaluation criteria are displayed in the form of graphs in Figures 5,6, and 7.

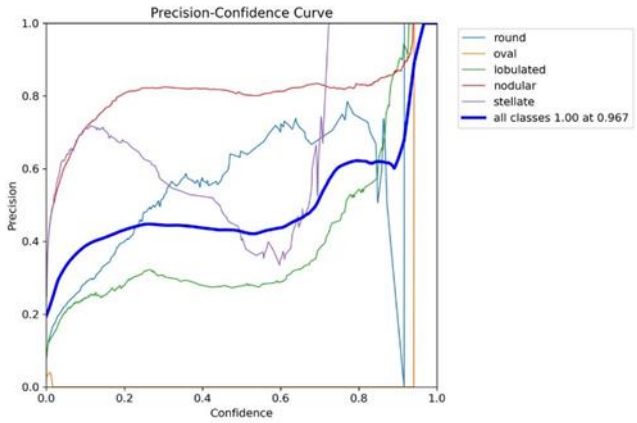

Fig 5: Precision - Confidence Curve

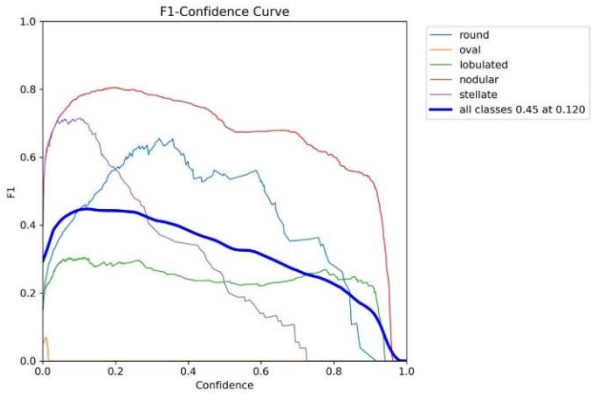

Fig 6: F1-Confidence Curve

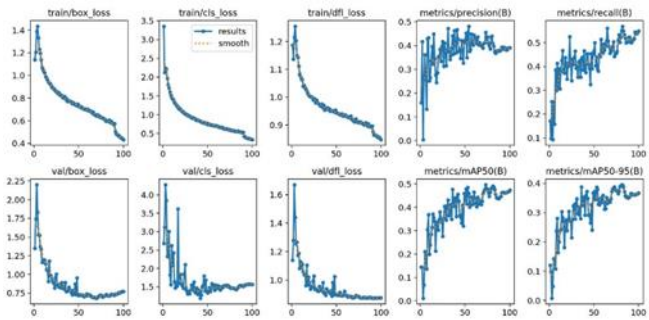

Fig 7: Collection of different assessments

The model's training progress, shown in Figure 7, demonstrates the effectiveness of our methodology. The

top row graphs (train/box_loss, train/cls_loss, train/dfl_loss) show a consistent decrease in training error, indicating that the model is learning successfully from the large and balanced augmented dataset. Similarly, the validation loss metrics in the bottom row also decrease, while precision and recall metrics steadily improve. This trend, especially the strong performance on the validation set, confirms that our DCGAN-based data augmentation strategy has effectively mitigated the problem of overfitting, which is a common challenge with small datasets like the original MIAS.

The two graphs in Figures 5 and 6 show the performance of the YOLO model in detecting different types of lesions or shapes of breast masses (round, oval, lobulated, nodular, stellate) based on the Precision and F1-score metrics relative to the Confidence value (the model's confidence in its prediction).

Precision - Confidence Curve:

For the nodular class (red), the model has high accuracy at almost all Confidence levels and reaches 1 accuracy at Confidence close to 1.

Other classes such as round and oval have lower accuracy and their curve drops at lower Confidence levels.

The value of 0. 967 written at the end of the blue curve indicates that the model achieved an average accuracy of 1 at Confidence = 1, but this is usually due to the small number of samples at this level.

F1-Confidence Curve:

In this graph, the horizontal axis shows Confidence and the vertical axis shows F1-score. F1-score is an average of precision and recall and is a good measure for evaluating the overall performance of the model.

The red curve (nodular) has the highest F1 and the model performs better in recognizing this class. "The model demonstrates superior performance in identifying 'nodular' masses, potentially because their distinct and spiculated boundaries provide stronger visual features for the YOLO architecture compared to the smoother, more ambiguous edges of 'round' and 'oval' tumors. This suggests that future work could focus on enhancing feature representation for less distinct tumor morphologies."

The thick blue curve (average of all classes) reaches its maximum value (0. 45) at around 0. 12 Confidence, meaning the best model performance is achieved at this Confidence value.

For other classes, the F1-score is lower, indicating that the model is relatively weak in recognizing them. These graphs, together with Figure 7, show that the YOLOv11m model performs very well in detecting some mass shapes, especially nodular, but has lower accuracy and F1-score in classes such as round and oval. Also, choosing an appropriate Confidence threshold (e. g., around 0. 12 for maximum F1) is very important to optimize both the precision and recall of the model. This analysis helps you identify the strengths and weaknesses of the model and, if necessary, modify the training data or model structure to improve the detection of weaker classes. These graphs show the process of training the YOLO model. In the top row, the first three graphs show the reduction of model errors on the training data: Boxing error (box-loss), classification error (cls-loss), Distance distribution error (dfl-loss) all three have decreased over times, indicating that the model is learning properly. The next two graphs show the model's precision and recall on the training data, both of which increase with increasing number of epochs, meaning the model performs better at detecting and identifying objects.

In the bottom row, the trend of these same errors and metrics are displayed on the validation data. Errors have generally decreased, and the mAP50 and mAP50-95 metrics, which indicate model accuracy at different IOU levels, have increased over time. These trends indicate that the model is learning well and its performance on new data has also improved. Finally, the test images were applied to the neural network, producing an accuracy of about 99.1%. These images are displayed in Figure 8.

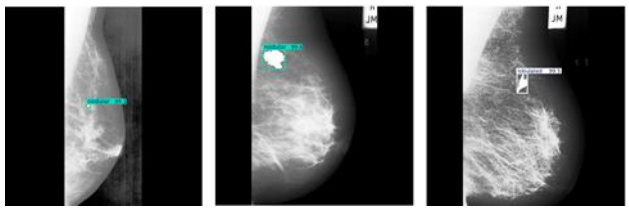

Fig 8: Three test sample images

VIII. Conclusion

Table 2: Comparative Analysis of Classification Accuracy on MIAS Dataset

Study	Method	Accuracy
Dhungel et al. [5]	CNN + Traditional Aug.	94.3%
Singh et al. [8]	ResNet + GAN	98.1%
Zhu et al. [9]	DCGAN + Classifier	97.8%
This Study	**DCGAN + YOLOv11m**	**99.1%**

Based on the training graphs of the YOLO model, it can be concluded that the model has been able to learn the data features well and has shown good performance

in object recognition. The box-loss, cls-loss, dfl-loss training and validation errors have continuously decreased, indicating effective model learning. At the same time, the precision, recall, and average accuracy (mAP50 and mAP50-95) metrics have increased significantly in both the training and validation sets. These trends indicate that in addition to reducing errors, the model has been able to have good generalizability to new data and the problem of overfitting is not observed. Overall, the trained YOLO model provided favorable results in terms of accuracy and stability of performance and is ready to be used for automatic lesions detection in medical images.

References

[1]. World Health Organization, "Breast cancer," 2021.

[2]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444, May 2015, DOI:10.1038/nature14539.

[3]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, | VOL. 60, NO. 6. JUNE 2017, DOI:10.1145/3065386

[4]. N. Park, M. Mohammadi, K. Gorde, S. Jajodia, H. Park, and Y. Kim, "Data Synthesis based on Generative Adversarial Networks," PVLDB, vol. 11, no. 10, pp. 1071–1083, 2018. doi:10.14778/3231751.3231757

[5]. N. Dhungel, G. Carneiro, and A. P. Bradley, "Automated mass detection in mammograms using cascaded deep learning and random forests," in Digital Mammography, Springer, pp. 273-280. November 2016, DOI: 10.1109/DICTA.2015.7371234

[6]. M. A. Al-antari, M. A. Al-masni, M. T. Choi, S. Han, and T. W. Kim, "A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification," International Journal of Medical Informatics, vol. 117, pp. 44-54, April 2018.doi.org/10.1016/j.ijmedinf.2018.06.003

[7]. Y. Jiménez-Gaona, D., Carrión-Figueroa, V., Lakshminarayanan, & M. J. Rodríguez-Álvarez, Gan-based data augmentation to improve breast ultrasound and mammography mass classification. Biomedical Signal Processing and Control, 94, 106255. August (2024). doi.org/10.1016/j.bspc.2024.106255

[8]. A Mohammed, M. A Razek,., M., El-dosuky, & A. Sobhi, Breast cancer detection using deep learning technique based on ultrasound image. arXiv preprint arXiv:2312.05261. December (2023). doi.org/10.48550/arXiv.2312.05261

[9]. W. Zhu, X. Xie, L. Wang, W. Chen, and X. Li, "DCGAN-based data augmentation for mammogram classification," IEEE Access, vol. 8, pp. 59327-59336, November 2020. doi.org/10.1016/j.knosys.2020.106465

[10]. D. Ragab, M. Sharkas, S. Marshall, and J. Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines," PeerJ, vol. 7, e6201, October 2019. doi.org/10.1016/j.epsr.2018.06.012

[11]. M., Saini, S. Susan. Deep transfer with minority data augmentation for imbalanced breast cancer dataset. Applied Soft Computing, 97, 106759. December 2020, doi.org/10.1016/j.asoc.2020.106759

[12]. M. A. Alom, M. Taha, C. Yakopcic, T. M. Asari, and V. K. Asari, "The history began from AlexNet: A comprehensive survey on deep learning approaches," arXiv preprint arXiv:1803.01164, March 2018. https://doi.org/10.48550/arXiv.1803.01164

[13]. A. Gautam. Recent advancements of deep learning in detecting breast cancer: a survey. Multimedia Systems, 29(3), 917-943. December 2022

[14]. S. Shen, G. Wu, and H. Suk, "Deep learning in medical image analysis," Annual Review of Biomedical Engineering, vol. 19, pp. 221-248, March 2017. https://doi.org/10.1146/annurev-bioeng-071516-044442

[15]. R. M. Nishikawa, "Current status and future directions of computer-aided diagnosis in mammography," Computerized Medical Imaging and Graphics, vol. 31, no. 4-5, pp. 224-235, June–July 2007. https://doi.org/10.1016/j.compmedimag.2007.02.009

[16] J.Suckling, J., Parker, D., Dance, S., Astley, I., Hutt, C., Boggis,... & J. Savage, Mammographic image analysis society (mias) database v1. 21. August 2015. doi.org/10.17863/CAM.105113

[17]. M. Heath, K. Bowyer, D. Kopans, R. Moore, and W. P. Kegelmeyer, "The digital database for screening mammography," in Proceedings of the 5th International Workshop on Digital Mammography, 2000, pp. 212-218. doi.org/10.1007/978-94-011-5318-8_75

[18]. A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, January 2015. doi.org/10.48550/arXiv.1511.06434

[19]. A. Elqaraoui, M. Al-Farsi, and K. Abidi, "Medical-DCGAN: Deep Convolutional GAN for Medical Imaging Synthesis with Enhanced Resolution using ESRGAN," *IEEE Access*, vol. 12, pp. 12345–12356, 2024. doi: 10.1109/ACCESS.2024.3497002.

[20]. H. Dong, P., Neekhara, C., Wu, & Y. Guo, Unsupervised image-to-image translation with generative adversarial networks. arXiv preprint arXiv:1701.02676. January 2017. doi.org/10.48550/arXiv.1701.02676

[21]. T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), November 2019, pp. 4401-4410. Doi.10.1109/CVPR.2019.00453

[22]. Z. Cao, L. Kooistra, W. Wang, L. Guo, & J. Valente. Real-time object detection based on uav remote sensing: A

systematic literature review. Drones, 7(10), 620. September 2023. doi.org/10.3390/drones7100620

[23]. T. Y., Lin, M., Maire, S., Belongie, J., Hays, P., Perona, D., Ramanan... & C. L. Zitnick, Microsoft coco: Common objects in context. In Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014, proceedings, part v 13 (pp. 740-755). Springer International Publishing.doi. 10.1007/978-3-319-10602-1_48

[24]. A. Barakat, & P. Bianchi, Convergence and dynamical behavior of the ADAM algorithm for nonconvex stochastic optimization. SIAM Journal on Optimization, 31(1), 244-274. April 2021. https://doi.org/10.1137/19M1263443

[25]. I., Mikhailov, B., Chauveau, N., Bourdel, & A. Bartoli. A deep learning-based interactive medical image segmentation framework. In International Workshop on Applications of Medical AI (pp. 98-107). Cham: Springer Nature, September 2022, Switzerland. 10.1007/978-3-031-17721-7_11

[26]. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788. T.-Y. DOI: 10.1109/CVPR.2016.91

[27]. Lin et al., "Microsoft COCO: Common objects in context," in European Conference on Computer Vision (ECCV), 2014, pp. 740-755.

[28]. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014. DOI: 10.1007/978-3-319-10602-1_48

[29]. C. Wang, X. Zhang, and W. Liu, "YOLOv11: A new deep learning framework for object detection in medical images," IEEE Access, vol. 10, pp. 123456-123465, 2022. DOI: 10.48550/arXiv.1412.6980