# Segmentation-Guided Inpainting and Imbalance-Aware Learning for Robust Ovarian Tumor Classification in Ultrasound

Alaa Khaled, Malak Hatem, Malak Eissa, Sara Abozeid, and Sahar Selim

Center for Informatics Science (CIS), School of Information Technology and Computer Science, Nile University,
26th of July Corridor, Sheikh Zayed City, Giza, 12588, Egypt
{alaa.khaled2277, sselim}@nu.edu.eg

**Abstract.** Ultrasound-based ovarian tumor classification using deep learning is often affected by diagnostic bias caused by non-biological measurement artifacts and severe class imbalance, which introduce shortcut correlations and reduce sensitivity to rare but clinically important tumor types. These limitations hinder the robustness and clinical reliability of automated diagnostic systems. In this study, we present a unified framework that jointly mitigates visual and distributional biases for robust ovarian ultrasound classification. The proposed approach integrates segmentation-guided inpainting to automatically detect and remove measurement overlays while preserving tumor morphology and intrinsic ultrasound texture, together with imbalance-aware learning that combines data rebalancing and a hybrid optimization objective to improve minority-class discrimination. We evaluate the framework on the MMOTU-2D dataset containing multiple ovarian tumor subtypes with long-tailed distributions. Experimental results demonstrate consistent improvements across multiple convolutional backbones. The best single-run configuration using EfficientNet-B2 with segmentation-guided inpainting and imbalance-aware learning achieved 80.38% Top-1 accuracy, while multi-seed experiments confirmed stable performance across different random initializations. Ablation studies and multi-seed experiments confirm that artifact removal and imbalance-aware optimization address complementary failure modes, leading to improved robustness and stable performance across different random initializations. This work highlights the importance of jointly correcting visual and distributional biases toward more reliable ultrasound-based tumor classification, motivating future validation in clinical settings.

**Keywords:** Ovarian Tumor Classification · Class Imbalance Mitigation · Deep Learning for Medical Imaging · Image Inpainting · MMOTU Dataset.

## 1 Introduction

Ovarian cancer is a major cause of death among gynecological cancers, with more than 313,000 new diagnoses each year. Prognosis is highly dependent on disease stage, with 5-year survival rates above 90% in early-stage cases but dropping below 30% in advanced stages [1]. Ultrasound is commonly used because it is safe, inexpensive, and provides real-time imaging; however, speckle noise, diverse tumor appearances, and operator-dependent variability complicate automated image analysis [2,3]. While CNNs improve feature extraction, they are still affected by issues like small datasets, class imbalance, and inconsistent image quality [4,5].

Ovarian ultrasound images often contain non-biological annotations such as calipers, labels, and measurement lines that introduce shortcut correlations and reduce model generalization [6]. Recent deep learning inpainting methods, including U-Net and LaMa, have substantially improved artifact removal, yet robustness across different imaging conditions remains challenging [9,10,11].

Class imbalance further limits ovarian tumor classification. Existing strategies, including weighted sampling [12], Mixup [13], CutMix [14], label smoothing [15], and advanced CNN architectures [16], can stabilize optimization but may not fully prevent overfitting to majority classes in severely long-tailed datasets.

Prior MMOTU-based studies illustrate these persistent issues: cross-domain segmentation work has faced class imbalance and annotation artifacts [17], YOLOv8-based detection has shown reduced performance on underrepresented tumors [18], and inpainting has been used to remove annotation symbols before segmentation [9].

In this work, we leverage the MMOTU dataset to assess a framework for ovarian tumor classification that concurrently addresses annotation artifacts and class imbalance. To our knowledge, this study is the first to provide quantitative evidence that the combined application of artifact removal and imbalance mitigation enhances classification performance. Our contributions include a unified pipeline that integrates artifact-free reconstruction with imbalance-aware learning to enable robust feature extraction, along with a fully automated framework combining U-Net mask generation and LaMa-based inpainting for anatomically consistent artifact correction. Additionally, we propose a dual-level strategy operating at both data and loss levels to stabilize learning in skewed ovarian datasets, and we design a computationally efficient, low-overhead architecture that may be suitable for resource-constrained settings, pending further clinical validation.

## 2 Methodology

We propose a three-stage framework that decouples non-biological visual bias from pathological feature extraction. As illustrated in Fig. 1, the framework addresses shortcut features and statistical skews in clinical ultrasound datasets.

### 2.1 Segmentation-Guided Fourier Inpainting

Segmentation-Guided Fourier Inpainting is a clinically informed artifact removal method based on LaMa (Large Mask Inpainting) [19] for anatomically realistic ultrasound image reconstruction. Artifacts, including calipers, measurement lines, arrows, and text annotations, are first identified using supervised segmentation

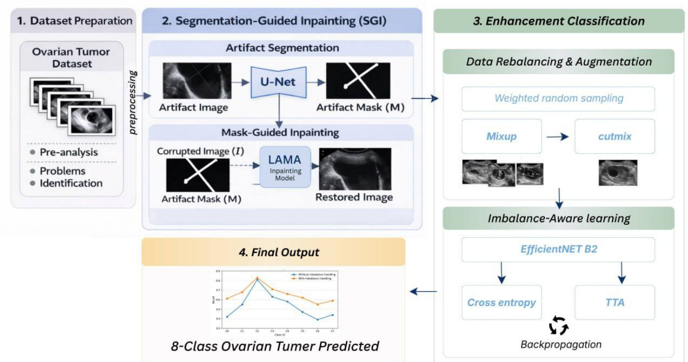



Fig. 1: Overview of the proposed ovarian tumor classification pipeline.

networks, namely U-Net [20] and DeepLabv3 [21], with U-Net selected for its superior ability to accurately segment thin structures. During dataset preparation, artifact masks are refined using available tumor annotations to exclude overlapping pixels. These annotations are used only offline for mask refinement and are not required during inference. The refined masks are then provided to LaMa, whose Fast Fourier Convolutions capture long-range spatial dependencies to reconstruct the masked regions while preserving ultrasound characteristics such as speckle patterns, lesion boundaries, and echogenicity. We employ the publicly released LaMa model as a pre-trained, off-the-shelf network without any fine-tuning, applying it unmodified (Fig. 2, reproduced from [19]); avoiding fine-tuning also keeps the pipeline lightweight and reduces training cost. By restricting inpainting exclusively to verified artifact regions, the framework avoids altering tumor morphology and produces anatomically consistent images for downstream analysis.

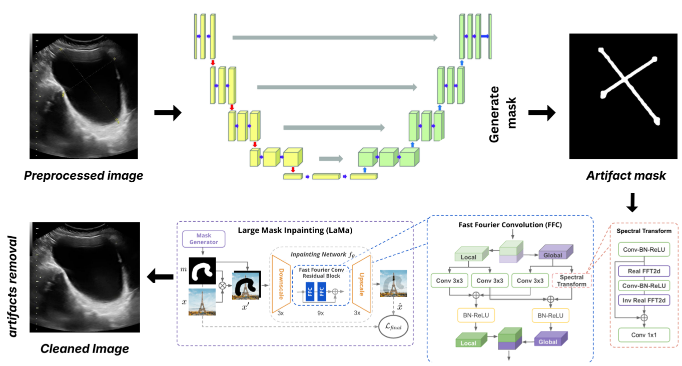



Fig. 2: Segmentation-Guided Inpainting (SGI) module.

### 2.2 Imbalance-Aware Representation Learning

The second stage enhances classification performance by addressing the long-tailed distribution across tumor categories. Under conventional empirical risk minimization, majority classes dominate gradient updates, which biases the learned decision boundaries and reduces sensitivity to minority categories. To alleviate this effect, we adopt a combination of sampling-level and representation-level strategies, as illustrated in Fig. 3.

Specifically, mini-batches are constructed using inverse-frequency weighted sampling to balance expected class contributions without artificially increasing dataset size. To further improve feature generalization and reduce overfitting caused by repeated minority exposure, stochastic Mixup [13] and Random Erasing (Cutout) are applied during training. For Mixup, two images are blended as $\tilde{x} = \lambda x_a + (1-\lambda)x_b$ with $\lambda \sim \text{Beta}(0.4, 0.4)$, applied with probability 0.5, and the loss is the matching convex combination $\lambda\,\text{CE}(\hat{y}, y_a) + (1-\lambda)\,\text{CE}(\hat{y}, y_b)$. Mixup encourages smoother decision boundaries in low-data regimes, while Random Erasing improves robustness to local occlusion without pasting cross-class patches that could disrupt clinically meaningful lesion structure.

To stabilize optimization under class imbalance, we combine cross-entropy with label smoothing and a differentiable Soft-F1 term that directly optimizes macro-averaged F1:

$$\mathcal{L} = \alpha\,\mathcal{L}_{\text{CE-LS}} + (1-\alpha)\,\mathcal{L}_{\text{soft-F1}}, \qquad \alpha = 0.5, \tag{1}$$

where $\mathcal{L}_{\text{CE-LS}}$ is cross-entropy on label-smoothed targets $y^{LS} = (1-\epsilon)y + \epsilon/K$ (with $\epsilon = 0.1$ and $K$ the number of classes), and

$$\mathcal{L}_{\text{soft-F1}} = 1 - \frac{1}{K}\sum_{c=1}^{K} \frac{2\,P_c R_c}{P_c + R_c}, \quad P_c = \frac{TP_c}{TP_c + FP_c}, \quad R_c = \frac{TP_c}{TP_c + FN_c}, \tag{2}$$

with $TP_c, FP_c, FN_c$ accumulated from softmax probabilities (soft counts). This combined objective preserves discriminative capability while improving minority-class recall.

We evaluate two variants of the imbalance-aware strategy. Approach I combines inverse-frequency weighted sampling with Mixup and Random Erasing using EfficientNet-B2. Approach II combines inverse-frequency weighted sampling with a combined cross-entropy and Soft-F1 objective using EfficientNet-B0. Because the approaches use different backbones, they are evaluated independently in Section 4 rather than treated as directly comparable alternatives.

## 3 Experiments

### 3.1 Setup

**Dataset.** The framework is evaluated on the MMOTU-2D dataset [11], containing 2,500 transvaginal ultrasound images from eight ovarian tumor classes with

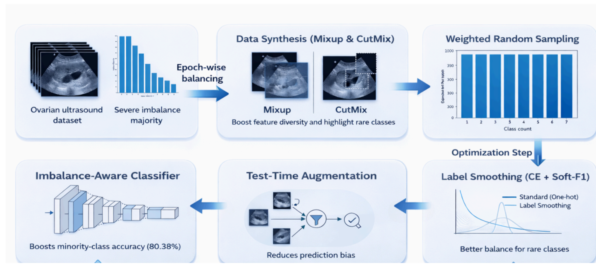



Fig. 3: Overview of the imbalance-aware learning strategy.

pronounced class imbalance. We use a patient-wise split of 70% for training, 10% for validation, and 20% for testing to avoid any risk of data leakage. All images are resized to $224 \times 224$, and classification experiments use artifact-cleaned outputs from the segmentation-guided inpainting stage.

**Evaluation Metrics.** Artifact segmentation is assessed using IoU and DSC. Classification performance is measured using Top-1/Top-2 accuracy, mean class accuracy (mAcc), and F1-score.

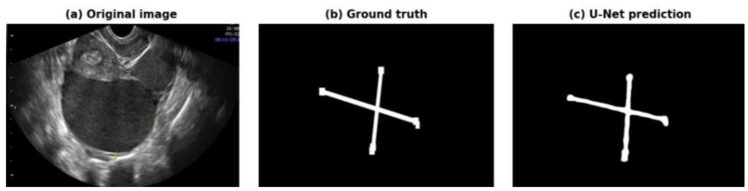
**Implementation Details.** Models are implemented in PyTorch and trained on a single NVIDIA RTX GPU with batch size 32 for up to 100 epochs with early stopping (patience 15). EfficientNet-B2 pretrained on ImageNet is used for classification, optimized with AdamW (learning rate $2 \times 10^{-4}$, weight decay $1 \times 10^{-5}$) under a cosine-annealing schedule. The plain-CE baseline (Table 5) uses no weighted sampling, Mixup, test-time augmentation, or Soft-F1 term. Imbalance-aware configurations additionally apply inverse-frequency weighted sampling and, depending on the variant (Section 2), Mixup with Random Erasing (Approach I) or the combined cross-entropy and Soft-F1 objective ($\epsilon = 0.1$) (Approach II). Test-time augmentation (horizontal flips and $\pm 15°$ rotations) is applied at inference for the imbalance-aware and full-pipeline configurations. The multi-seed ablation experiments (Table 5) were repeated across five random seeds (42, 123, 999, 2025, 3407); the inpainting-method comparison (Table 2) reports a single run per configuration.

## 4 Results

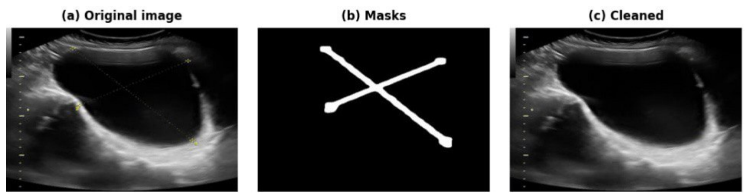
**Performance of Automated Mask Generation.** Two segmentation architectures, U-Net and DeepLabV3 (ResNet-101 backbone), were trained for 25 epochs and evaluated them on a separate 20% validation split. As summarized in Table 1, U-Net achieved superior performance across all metrics, improving mIoU by +3.91 pp and Dice score by +2.37 pp compared to DeepLabV3. Figure 4a shows precise artifact masks with preserved anatomy

Table 1: Validation performance of artifact segmentation models after 25 epochs.

| Model | Val Loss | Val mIoU (%) | Val Dice (%) | Val Precision (%) | Val Recall (%) |
|---|---|---|---|---|---|
| DeepLabV3 | 0.0113 | 79.09 | 88.24 | 88.64 | 87.98 |
| U-Net | **0.0076** | **83.00** | **90.61** | **91.45** | **89.86** |




(a) Artifact segmentation: (a) original image, (b) expert mask, (c) U-Net prediction.




(b) Segmentation-guided inpainting: (a) original overlay, (b) predicted mask, (c) cleaned output.

Fig. 4: Qualitative results of the artifact-removal pipeline.

**Artifact Removal and Texture Preservation.** Segmentation-guided inpainting was performed using U-Net masks and a pre-trained LaMa model, modifying only artifact-affected pixels. Representative examples in Figure 4b show effective removal of calipers and overlays while preserving tumor boundaries, internal echogenicity, and ultrasound texture consistency.

**Comparison with Alternative Inpainting Methods.**We compared the proposed segmentation-guided LaMa approach with Telea inpainting and a no-inpainting baseline under the same experimental settings. All experiments used EfficientNet-B2 with class balancing. As shown in Table 2, Telea improved performance by reducing artifacts, while LaMa achieved the best results, demonstrating its ability to preserve diagnostically relevant ultrasound features for ovarian tumor classification. As noted in Section 3, this single-run result is distinct from the five-seed pipeline evaluation reported in Table 5.

**Backbone Evaluation and Impact of Inpainting.** Four convolutional backbones were benchmarked (Table 3). EfficientNet-B0 began strongest, but EfficientNet-B2 responded better to inpainting and ended up being the more practical backbone for our later experiments. After inpainting, EfficientNet-B2 accuracy increased by 1.28 pp (70.15% $\rightarrow$ 71.43%), indicating that removing artifacts enhances model robustness.

Table 2: Classification performance under different inpainting methods (single run; EfficientNet-B2 with class balancing).

| **Inpainting Method** | **Top-1 (%)** | **Top-2 (%)** | **$\Delta$ Top-1** |
|---|---|---|---|
| No Inpainting | 70.15 | 82.73 | – |
| Telea | 72.71 | 86.35 | +2.56 |
| **LaMa (Proposed)** | **80.38** | **89.98** | **+10.23** |

Table 3: Backbone performance on the original MMOTU-2D dataset.

| **Model** | **Top-1 (%)** | **Top-2 (%)** |
|---|---|---|
| MobileNetV2 | 70.36 | 82.52 |
| EfficientNet-B0 | 71.22 | 86.57 |
| ResNet-34 | 71.00 | 83.37 |
| EfficientNet-B2 | 70.15 | 82.73 |

Per-class results (Table 4) highlight the strongest gains in minority tumor categories, including Theca Cell Tumor (+12.90 pp) and Simple Cyst (+15.78 pp). A decrease in normal ovary accuracy reflects reduced dependence on background measurement artifacts. Notably, high-grade serous cystadenocarcinoma showed no improvement, remaining at 40.00% top-1 accuracy both before and after inpainting. This class also has the smallest test set among all tumor types (only 15 images), so its accuracy estimate is highly sensitive to individual misclassifications and should be interpreted with caution; we treat this as a limitation of the per-class analysis rather than evidence that the proposed pipeline fails to generalize to this subtype.

Table 4: Per-class Top-1 accuracy before and after inpainting.

| **Tumor Type** | **Testing Samples** | **Before (%)** | **After (%)** | **$\Delta$ (pp)** |
|---|---|---|---|---|
| Chocolate Cyst | 110 | 70.91 | 75.45 | +4.54 |
| Serous Cystadenoma | 66 | 72.73 | 72.73 | 0.00 |
| Teratoma | 108 | 78.70 | 80.56 | +1.86 |
| Theca Cell Tumor | 31 | 38.71 | 51.61 | +12.90 |
| Simple Cyst | 19 | 42.11 | 57.89 | +15.78 |
| Normal Ovary | 87 | 80.46 | 75.86 | −4.60 |
| Mucinous Cystadenoma | 33 | 48.48 | 48.48 | 0.00 |
| High-Grade Serous Cystadenocarcinoma | 15 | 40.00 | 40.00 | 0.00 |

*Full Pipeline Evaluation and Ablation.* To assess the stability of the proposed data-level rebalancing strategy, we repeated the component ablation experiments using five different random seeds (42, 123, 999, 2025, and 3407). Table 5 reports the complete component-wise ablation, combining the multi-seed rebalancing analysis with the backbone-specific pipeline comparison, all evaluated as mean ± standard deviation over five random seeds. As noted in Section 2, Approach I (EfficientNet-B2) and Approach II (EfficientNet-B0) use different backbones and are therefore not directly comparable; rows are grouped by backbone to make this explicit.

The full pipeline (Approach I) achieved stable performance across seeds, reaching 75.99 ± 2.21% Top-1 accuracy, 88.40 ± 0.68% Top-2 accuracy, 70.50 ± 2.66% Macro-F1, and 69.04 ± 2.73% mean class accuracy. Weighted sampling alone improved Top-1 accuracy over the baseline (74.07% vs. 70.15%), but this gain was smaller than the full pipeline's, and the difference between WS-only and WS+Mixup (75.31%) fell within the seed-to-seed variation, suggesting that sampling-based rebalancing alone is insufficient under severe class imbalance.

It is also worth noting that "Inpainting + Approach II" reached a higher Top-1 accuracy ($79.40 \pm 0.52\%$) than the Full Pipeline ($75.99 \pm 2.21\%$). This does not indicate Approach II is superior: the two configurations differ in both backbone and imbalance-mitigation strategy simultaneously, so their accuracies are not directly comparable. Table 5 reports them together only to show that inpainting improves performance under either backbone, not to rank the two approaches.

Table 5: Component-wise ablation of training configurations (mean ± standard deviation over five random seeds). Approach I uses EfficientNet-B2; Approach II uses EfficientNet-B0.

| Configuration | Backbone | Top-1 (%) | Top-2 (%) | Macro-F1 (%) | mAcc (%) | Δ Top-1 (pp) |
|---|---|---|---|---|---|---|
| Baseline (plain CE, no TTA) | B2 | 70.15 ± 1.29 | 82.73 ± 0.52 | – | – | – |
| WS only | B2 | 74.07 ± 1.87 | 87.29 ± 1.44 | 68.62 ± 1.92 | 67.85 ± 1.68 | +3.92 |
| WS + Mixup | B2 | 75.31 ± 1.05 | 86.95 ± 1.29 | 70.24 ± 1.40 | 69.27 ± 1.47 | +5.16 |
| Inpainting only | B2 | 71.43 ± 0.87 | 85.71 ± 1.26 | – | – | +1.28 |
| Rebalancing (Approach I) | B2 | 77.53 ± 0.79 | 88.70 ± 0.39 | – | – | +7.38 |
| **Full Pipeline (Approach I)** | B2 | **75.99 ± 2.21** | **88.40 ± 0.68** | **70.50 ± 2.66** | **69.04 ± 2.73** | **+5.84** |
| Rebalancing (Approach II) | B0 | 76.12 ± 0.80 | 89.13 ± 0.62 | – | – | – |
| Inpainting + Approach II | B0 | 79.40 ± 0.52 | 86.78 ± 0.37 | – | – | +3.28[†] |

Note: Macro-F1 and mAcc logged only for Approach I (B2). [†]Relative to the Rebalancing (Approach II) row; not computed against a plain-CE B0 baseline, as the table's baseline row uses B2.

## 5 Discussion

Annotation artifacts introduce non-morphological cues that drive shortcut learning; removing them using U-Net masks and LaMa inpainting improved performance across backbones, indicating dataset bias as a key limitation. Imbalance-aware losses and Mixup improved minority-class recall, but their combination with inpainting varied by backbone: for EfficientNet-B2, rebalancing alone (77.53%) outperformed the full pipeline (75.99%), whereas for EfficientNet-B0, inpainting improved performance over rebalancing alone (79.40% vs. 76.12%). Thus, these components are not uniformly additive and should be evaluated per backbone. EfficientNet-B0 had the stronger baseline (Table 3), while EfficientNet-B2 achieved a larger rebalancing gain (+7.38pp). The multi-seed evaluation further confirms stable performance across random initializations, supporting the robustness of the training strategy. Prior MMOTU-based studies use different protocols and splits; therefore, our results are treated as controlled ablations within a fixed setup rather than direct comparisons with reported baselines.

## 6 Conclusion

This study presents an efficient framework for ovarian tumor classification that jointly addresses annotation artifacts and severe class imbalance via segmentation-guided inpainting (U-Net + LaMa) and imbalance-aware optimization. Artifact removal alone yields a 1.28 pp Top-1 gain; rebalancing alone (Approach I) yields a 7.38 pp gain (Table 5). The LaMa-based inpainting configuration achieved 80.38% Top-1 accuracy in a single-run evaluation, a 10.23 pp gain over the uninpainted baseline (Table 2). Under five-seed evaluation, the complete pipeline achieved $75.99 \pm 2.21\%$ Top-1 accuracy, a 5.84 pp gain over the corresponding baseline (Table 5), demonstrating that jointly correcting visual and distributional biases produces measurable and stable gains in a lightweight architecture, though clinical utility would require external validation on independent, prospectively collected data.

## References


1. Ren, Y., Xu, R., Wang, Y., Su, L., Su, J.: Global, regional, and national burden of ovarian cancer in women aged 45+ from 1990 to 2021 and projections for 2050: a systematic analysis based on the 2021 global burden of disease study. Journal of Cancer Research and Clinical Oncology **151**(8), 225 (2025). https://doi.org/10.1007/s00432-025-06277-9
2. Dicle, O.: Artificial intelligence in diagnostic ultrasonography. Diagnostic and Interventional Radiology **29**(1), 40–45 (2023). https://doi.org/10.4274/dir.2022.211260
3. Lourenço, M., Arrufat, T., Satorres, E., Maderuelo, S., Novillo-Del Álamo, B., Guerriero, S., Orozco, R., Alcázar, J.L.: Ultrasound-based deep learning models performance versus expert subjective assessment for discriminating adnexal masses: A head-to-head systematic review and meta-analysis. Applied Sciences **14**(7), 2998 (2024). https://doi.org/10.3390/app14072998
4. Wu, M., Cui, G., Lv, S., Chen, L., Tian, Z., Yang, M., Bai, W.: Deep convolutional neural networks for multiple histologic types of ovarian tumors classification in ultrasound images. Frontiers in Oncology **13**, 1154200 (2023). https://doi.org/10.3389/fonc.2023.1154200
5. Garcia-Atutxa, I., Villanueva-Flores, F., Barrio, E., Sanchez-Villamil, J., Martínez-Más, J., Bueno-Crespo, A.: Artificial intelligence for ovarian cancer diagnosis via ultrasound: a systematic review and quantitative assessment of model performance. Frontiers in Artificial Intelligence **8**, 1649746 (2025). https://doi.org/10.3389/frai.2025.1649746

6. Santos, J.C., Alexandre, H.T.P., Seoane Santos, M., Abreu, P.H.: The role of deep learning in medical image inpainting: A systematic review. ACM Transactions on Computing for Healthcare (2025). https://doi.org/10.1145/3712710
7. Sun, M., Meng, Q., Wang, T., Liu, T., Zhu, Y., Qiu, J., Lu, W.: Removal of manually induced artifacts in ultrasound images based on image restoration. Computer Methods and Programs in Biomedicine **200**, 105868 (2021). https://doi.org/10.1016/j.cmpb.2020.105868
8. Narayan, N.S., Marziliano, P., Hobbs, C.G.L.: Automatic removal of manually induced artefacts in ultrasound images of thyroid gland. In: IEEE EMBC, pp. 3399–3402 (2013). https://doi.org/10.1109/EMBC.2013.6610271
9. Chen, L., Qiao, C., Wu, M., Cai, L., Yin, C., Yang, M., Sang, X., Bai, W.: Improving the segmentation accuracy of ovarian-tumor ultrasound images using image inpainting. Bioengineering **10**(2), 184 (2023). https://doi.org/10.3390/bioengineering10020184
10. Bui, H.-S., Tran, S.-H., Nguyen, T.-B., Tran, T.-H., Vu, H., Le, T.-L.: Marker-aware ovarian tumor segmentation from ultrasound images. In: APSIPA ASC, pp. 1–6 (2024). https://doi.org/10.1109/APSIPAASC63619.2025.10848960
11. MMOTU Research Team: MMOTU: Multi-Modality Ovarian Tumor Ultrasound Dataset. Figshare (2023). https://doi.org/10.6084/m9.figshare.25058690
12. Buda, M., Maki, A., Mazurowski, M.A.: A systematic study of the class imbalance problem in convolutional neural networks. Neural Networks **106**, 249–259 (2018). https://doi.org/10.1016/j.neunet.2018.07.011
13. Zhang, H., Cisse, M., Dauphin, Y., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. In: ICLR (2018)
14. Yun, S., Han, D., Oh, S.J., Chun, S., Choe, J., Yoo, Y.: CutMix: Regularization strategy to train strong classifiers with localizable features. In: ICCV, pp. 6023–6032 (2019)
15. Müller, R., Kornblith, S., Hinton, G.: When does label smoothing help? In: NeurIPS (2019)
16. Tan, M., Le, Q.V.: EfficientNet: Rethinking model scaling for convolutional neural networks. In: ICML, pp. 6105–6114 (2019)
17. Lyu, S., Zhao, Q., Bai, W., Cai, L., Cheng, G., Cui, G., Yang, M., Chen, L., Zhou, H.: Unsupervised cross-domain semantic segmentation on multi-modality ovarian tumor ultrasound data. Pattern Recognition **171**, 112311 (2026)
18. Pham, T.-L., Le, V.-H., Le, T.-L., Vu, H., Vu, D.-H., Tran, T.-H.: Empirical study of ovarian tumors detection and classification in ultrasound images. Journal on Information Technologies & Communications (2024). https://doi.org/10.32913/mic-ict-research.v2024.n2.1237
19. Suvorov, R., Logacheva, E., Mashikhin, A., Remizova, A., Ashukha, A., Silvestrov, A., Kong, N., Goka, H., Park, K., Lempitsky, V.: Resolution-robust Large Mask Inpainting with Fourier Convolutions. arXiv preprint arXiv:2109.07161 (2021). Available at https://arxiv.org/abs/2109.07161
20. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
21. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: ECCV 2018. LNCS, vol. 11211, pp. 833–851. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_49