Article

Deep Learning Realizes Photoacoustic Imaging Artifact Removal

Ruonan He, Yi Chen, Yufei Jiang, Yuyang Lei, Shengxian Yan, Jing Zhang and Hui Cao *

School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710119, China;
ruonan0104@snnu.edu.cn (R.H.); chenyiyi@snnu.edu.cn (Y.C.); jyf0010@snnu.edu.cn (Y.J.);
lyy_weirong@snnu.edu.cn (Y.L.); 2023303971@snnu.edu.cn (S.Y.); zhanggale@snnu.edu.cn (J.Z.)
* Correspondence: caohui@snnu.edu.cn

Abstract: Photoacoustic imaging integrates the strengths of optics and ultrasound, offering high resolution, depth penetration, and multimodal imaging capabilities. Practical considerations with instrumentation and geometry limit the number of available acoustic sensors and their "view" of the imaging target, which result in image reconstruction artifacts degrading image quality. To address this problem, YOLOv8-Pix2Pix is proposed as a hybrid artifact-removal algorithm, which is advantageous in comprehensively eliminating various types of artifacts and effectively restoring image details compared to existing algorithms. The proposed algorithm demonstrates superior performance in artifact removal and segmentation of photoacoustic images of brain tumors. For the purpose of further expanding its application fields and aligning with actual clinical needs, an experimental system for photoacoustic detection is designed in this paper to be verified. The experimental results show that the processed images are better than the pre-processed images in terms of reconstruction metrics PSNR and SSIM, and also the segmentation performance is significantly improved, which provides an effective solution for the further development of photoacoustic imaging technology.

Keywords: photoacoustic imaging; artifact reduction; Pix2Pix; YOLOv8

1. Introduction

Photoacoustic imaging (PAI) has shown great potential in providing detailed views of human anatomy [1–5]. Photoacoustic imaging technology enables high-resolution imaging of human tissues by delivering pulsed laser beams. In this technique, light signals that are preferentially absorbed by the tissue cause the generation of acoustic waves, which are subsequently detected and imaged by a conventional ultrasound (US) transducer. Photoacoustic imaging is widely used for the detection and treatment of cancer, vascular monitoring, medication delivery, surgical navigation, and metal implant detection [6–10]. However, radioactivity artifacts are often present in photoacoustic images, which can pose diagnostic and therapeutic challenges for clinicians [11–13]. Radioactivity artifacts are primarily caused by strong reflection of acoustic waves, which is often not adequately accounted for by conventional beam imaging methods. Reflections [14–16] lead to signal localization errors and thus affect image accuracy. At the same time, inconsistencies in the acoustic environment, such as variations in sound speed, density, or attenuation, make it difficult to accurately model acoustic wave propagation. Although photoacoustic imaging technology has a promising future for clinical applications, the presence of artifacts limits its further development [17–19].

There are various methods to remove artifacts from photoacoustic images in current studies. Traditional methods such as singular value decomposition [20] and Gaussian filters [21] can only remove some artifacts caused by strong acoustic reflection [22]. Photoacoustically guided focused ultrasound (PAFUSion) technology uses an innovative method to eliminate reflection artifacts by simulating the photoacoustic(PA) field, but it needs to meet the requirements of ultrasound and photoacoustic image matching, which may reduce the frame rate and bring errors in the case of tissue movement [23]. Alternatively,

some studies have attempted to utilize neural networks [24–26] to estimate beamforming delay functions to reduce artifacts caused by sound velocity errors. Other studies have explored Wasserstein generative adversarial networks (WGAN-GP) to reduce finite-view and finite-bandwidth artifacts [27,28], but these approaches have not yet adequately considered the effects of light excitation and propagation on imaging quality. There are also studies using Pixel-DL [29] technology to improve PAT image reconstruction quality and reduce computational costs, but this method is not suitable for heterogeneous media (such as in vivo imaging).

In this paper, a new photoacoustic image processing model is proposed to cope with different types of artifacts in photoacoustic images by using the YOLOv8-pix2pix algorithm. The YOLOv8 algorithm handles the precise segmentation of brain tumors [30,31] and the removal of artifacts near the signal source. Meanwhile, the Pix2Pix conversion approach primarily addresses artifacts caused by signal aliasing. After the modeling process, not only are the artifacts effectively removed, but also the quality of the images is significantly improved. Meanwhile, the application of the YOLOv8 algorithm in image segmentation provides reliable image data support for the detection and treatment of brain tumors.

2. Materials and Methods

2.1. Overview of the Framework

Figure 1 illustrates the overall structure proposed in this paper. The framework includes five steps of photoacoustic image generation, removal of mixing artifacts, removal of artifacts outside the signal source, deep learning image segmentation, and model validation and evaluation, which are extensively utilized in medical image processing. First, photoacoustic images are generated by k-wave simulation; second, the generated photoacoustic images are de-artifacted using the Pix2Pix network to eliminate useless information. Subsequently, YOLOv8 is used to remove artifacts beyond the signal source. Moreover, the deep learning model (YOLOv8) is used to segment the images before and after artifact removal to aid in verifying the artifact-removal effect and assessing the accuracy of the removed images. Finally, various combinations of deep learning models are evaluated and validated.
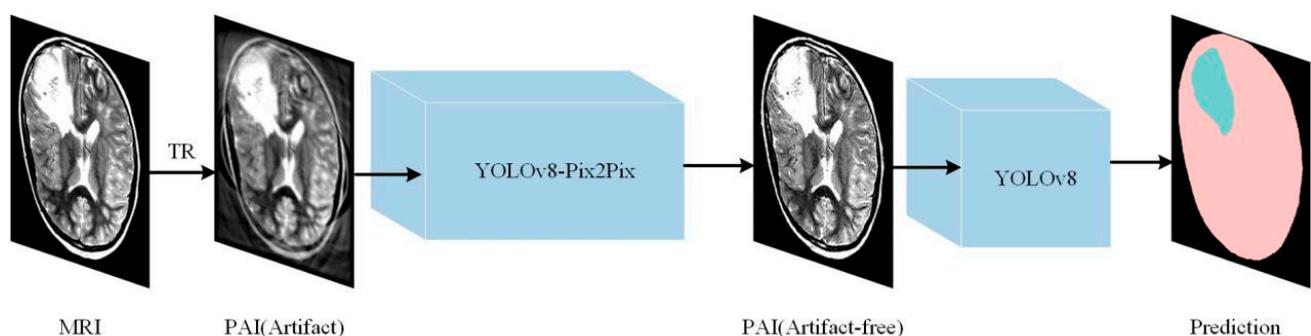

Figure 1. Schematic diagram of deep learning structure for photoacoustic images.

2.2. Deep Learning Algorithm for Photoacoustic Image Segmentation

The YOLOv8 algorithm [32,33] is a fast single-stage target detection method with excellent performance and efficient real-time performance. It is able to accurately capture target features in photoacoustic images and provide accurate segmentation results while maintaining a fast processing speed. Its network structure draws on the design principles of the YOLACT network to achieve real-time object–instance segmentation and maintain a high accuracy rate.

The YOLOv8 network [34] is mainly composed of a backbone network and a head network, as depicted in Figure 2. The backbone network utilizes 3×3 convolution, the C2f module, and the SPPF module, which are lighter than YOLOv5. The C2f module replaces the traditional C3 module and enhances the gradient flow through jump connections and

segmentation operations. Additionally, the CSP version of the network employs residual connections and direct connections to optimize the information flow. The neck structure is enhanced with the FPN+PAN configuration to facilitate feature fusion, which helps in detecting objects at multiple scales. During training, YOLOv8 improves the training data with image enhancement techniques such as mosaic enhancement, enhancing the model's adaptability to various scenes. The technique forces the model to learn how to detect objects that are partially obstructed and in different locations. Over the past 10 training cycles, the YOLOv8 network has deactivated mosaic enhancement, a method that has proven effective in improving network accuracy.
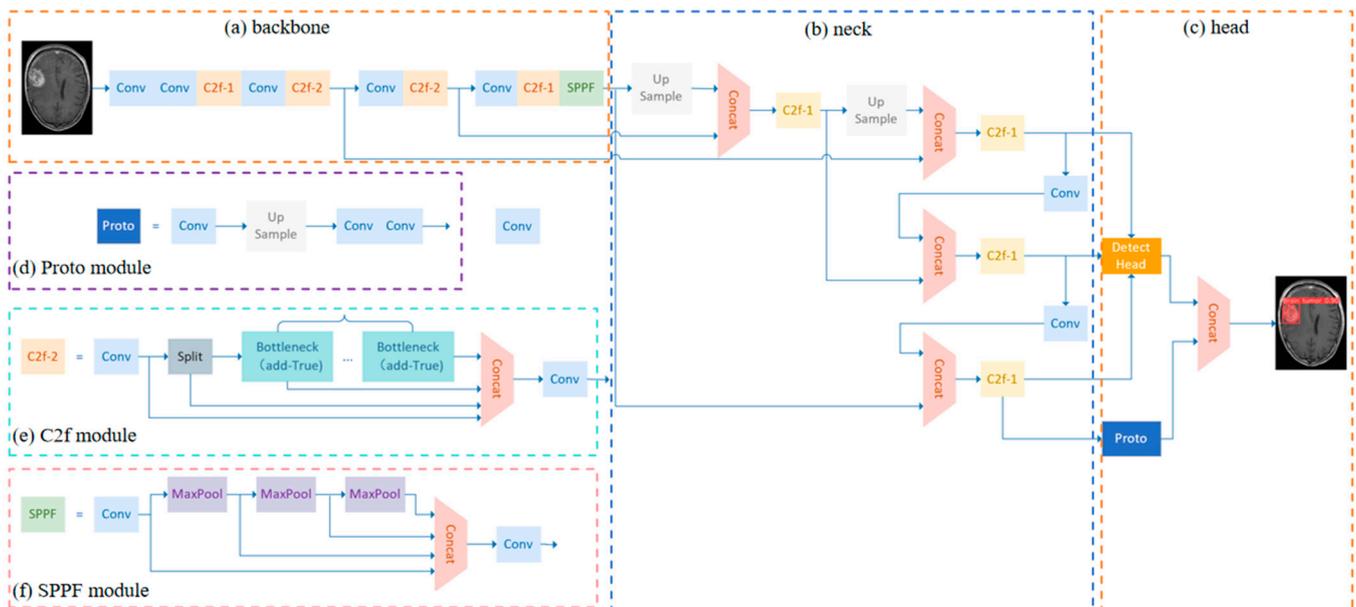

Figure 2. YOLOv8 architecture.

Applications of YOLOv8 to photoacoustic image processing of brain tumors include artifact removal and segmentation of brain tumors. The model was trained on labeled brain and brain tumor pictures, and it can successfully remove artifacts from non-signal sources and segment brain tumors, allowing brain tumor segmentation and artifact removal from photoacoustic images.

2.3. Deep Learning Algorithm for Photoacoustic Image Artifact Removal

Deep learning is widely used in removing artifacts, but existing algorithms can only remove bar and ring artifacts in photoacoustic images and cannot effectively deal with various complex artifacts. The YOLOv8-Pix2Pix model proposed in this paper (the model framework is shown in Figure 3) addresses the characteristics of photoacoustic images and can comprehensively handle different types of artifacts while effectively recovering image details. The model first utilizes the YOLOv8 algorithm (Figure 3a) to localize the photoacoustic image and accurately segment the photoacoustic source to remove artifacts outside the photoacoustic source. Then, the Pix2Pix algorithm (Figure 3b) is employed to process the artifacts in the photoacoustic image caused by radiation artifacts and the mixing of the photoacoustic source, enabling comprehensive artifact removal in the photoacoustic image.
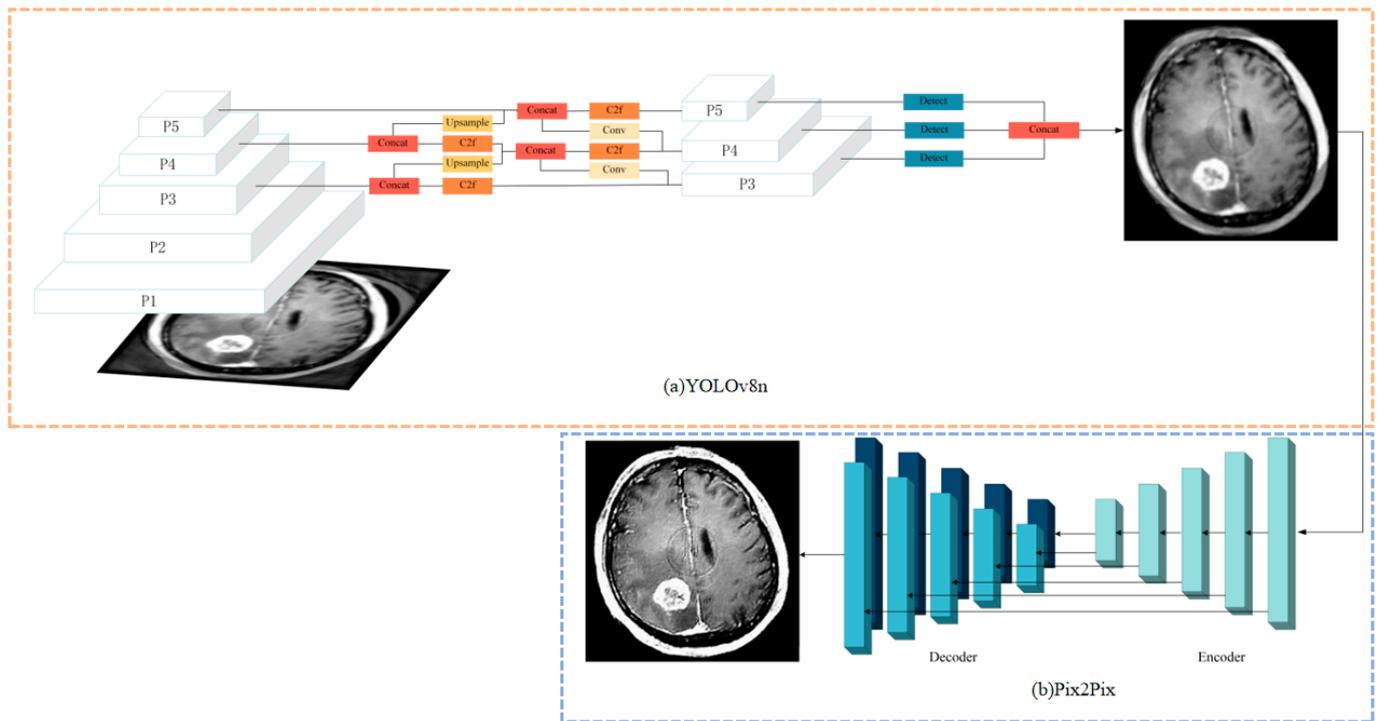
Figure 3. YOLOv8-Pix2Pix model flow for processing photoacoustic image artifacts.

Pix2Pix [35,36] is a Conditional Generative Adversarial Network (cGAN)-based image-to-image translation model designed for image transformation tasks that require explicit one-to-one correspondence, and can effectively remove aliasing artifacts within optical sound sources. The model consists of two core components, the generator (*G*) and the discriminator (*D*). The generator's task is to convert the input image into an output image, which is continuously optimized by means of adversarial training in order to deceive the discriminator and generate a sufficiently realistic image. The discriminator is responsible for distinguishing between the generated image and the real image, thus prompting the generator to produce a more realistic image. Pix2Pix GAN's objective function VPix2Pix GAN is defined as

$$argmin_G max_D V_{pix2pixGAN}(G, D) = E_{x \sim P_{low}(x), y \sim P_{full(y)}}[log D(x, y)] + E_{x \sim P_{low}(x)}[\log(1 - D(x, G(x)))] \tag{1}$$

where $x \sim Plow(x)$ denotes the brain tumor Magnetic Resonance Imaging (MRI) image and $y \sim Pfull(y)$ denotes the corresponding brain tumor photoacoustic image. The discriminator loss is defined as the cross-entropy, as shown in Equation (2), which is used to adjust the weights of the discriminators. Compared to the common L2 loss, the L1 loss enforces low-frequency correctness, as shown in Equation (3), and encourages the reduction of image blur. Therefore, the generator is trained by minimizing the difference between the de-artifacted image and the reference image, i.e., the mean absolute error (MAE).

$$V_{L1}(G) = E_{x \sim P_{low(x)}, y \sim P_{full(y)}}[\|y - G(x)\|_1] \tag{2}$$

Then, our final objective function can be expressed as

$$G^* = argmin_G max_D V_{pix2pixGAN}(G, D) + \lambda V_{L1}(G) \tag{3}$$

where $\lambda = 100$ is an adjustable parameter to control the balance between $V_{pix2pixGAN}(G, D)$ and $V_{L1}(G)$.

YOLOv8 is a target detection algorithm for automatic identification and removal of artifacts in photoacoustic images, with efficient real-time target detection and good adapt-

ability to complex scenes. The source of photoacoustic images usually has a unique shape and texture characteristics. YOLOv8 can quickly and accurately analyze the photoacoustic images, effectively locate and identify the source of photoacoustic images, and then remove other contents in the photoacoustic images, so as to improve the quality and clarity of the photoacoustic images. The high efficiency and accuracy of YOLOv8 make it an important tool for photoacoustic image processing, which provides reliable image support for medical diagnosis. reliable image support.

The YOLOv8-Pix2Pix model proposed in this paper is applied to the task of converting brain tumor MRI images to brain tumor photoacoustic images. In the training process, brain tumor MRI images are paired with k-waves to generate corresponding brain tumor photoacoustic images, and our model, which is trained with appropriate loss functions, is able to generate photoacoustic images similar to the real images but with artifacts removed.

3. Results

3.1. Experimental Setup

In this study, the image was preprocessed. The original dataset is divided into the training set, validation set and test set in the ratio of 6:2:1. The image is scaled to 640×640 pixels, while the pixel values are normalized to the range [0, 1]. In the training process, we use a variety of data enhancement strategies to expand the data set, such as random cropping, random scaling and random flipping.

The software and hardware environment configurations used for algorithm training in this paper are shown in Table 1.

Table 1. Experimental environment.

Type	Parameter
CPU	Intel(R) Core(TM) i7-10700
RAM	32 GB
GPU	NVIDIA Quadro P2200
Programming language	Python 3.10
Deep learning framework	PyTorch 2.2.1, CUDA 12.1
Dependency library	Numpy 1.26.4, tqdm 4.66.2, tensorboard 2.16.2, opencv-python 4.9.0.80

The hyperparameter settings for the algorithm are presented in Table 2.

Table 2. The main hyperparameters of the network.

Network	Main Hyperparameters	Specific Values
YOLOv8	Batch size	16
	Image size	640
	Optimiser	Adam
	Momentum	0.937
	Initial learning rate	0.01
	epoch	300
Pix2Pix	Batch_size	1
	Momentum	0.5
	Optimiser	Adam
	Loss function	cGAN
	Initial learning rate	0.0002
	epoch	200

3.2. Evaluation Metrics

In this study, we comprehensively evaluate the performance of deep learning image reconstruction and segmentation models. These commonly used indicators include peak

signal-to-noise ratio (PSNR), structural similarity index (SSIM), accuracy, recall rate, accuracy, intersection ratio (IoU), mean BFScore, and Dice coefficient. During the evaluation, the calculation of PSNR is based on the actual range of reconstructed images, and takes into account thermal noise, electronic noise and background noise in photoacoustic imaging. SSIM evaluation covers the brightness, contrast, structural fidelity of photoacoustic images and the accuracy of the spatial distribution of photoacoustic sources. These dimensions of comprehensive assessment enable us to more fully understand the deep learning model in the performance of image reconstruction and segmentation task.

The model evaluation measures are calculated according to Formulas (4)–(12).

$$PSNR = 10 \times log_{10}\left(\frac{MAX^2}{MSE}\right) \qquad (4)$$

$$MSE = \frac{1}{N}\sum_{i=0}^{N}\left(I_i - \hat{I}_i\right)^2 \qquad (5)$$

MAX is the maximum pixel value of the reconstructed image and *MSE* is the mean square error, which reflects the degree of difference between the reconstructed image and the original image. In photoacoustic imaging, the value of PSNR is affected by thermal, electronic, and background noise. A higher PSNR value usually means that the reconstructed image shows little difference from the original image and the reconstruction is of high quality.

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \qquad (6)$$

Among them, μ_x and μ_y are the mean values of the original image x and the reconstructed image y, σ_x^2 and σ_y^2 are their variances and σ_{xy} is their covariance. c_1 and c_2 are two constants used for stability calculation. SSIM takes into account brightness, contrast and fidelity of structure, and comprehensively evaluates the characteristics of photoacoustic imaging images.

$$precision = \frac{TP}{TP + FP} \qquad (7)$$

$$recall = \frac{TP}{TP + FN} \qquad (8)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (9)$$

$$IoU = \frac{TP}{TP + FP + FN} \qquad (10)$$

$$Mean\ BFScore = \frac{Area\ of\ Overlap}{Total\ number\ of\ Pixels} \qquad (11)$$

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \qquad (12)$$

Among them, *TP*, *FP*, *TN*, and *FN* represent 'True Positive', 'False Positive', 'True Negative', and 'False Negative', respectively. N denotes the number of categories. In Equation (11), the "overlap region" represents the overlap between the predicted segmentation and the ground truth, and the "total number of pixels" represents the number of pixels present in the image.

These formulas can be used for quantitative evaluation of the performance of deep learning image reconstruction and segmentation models.

3.3. Brain Tumor Dataset

The Br35H dataset is a medical imaging dataset for the detection of brain tumors. The dataset contains brain MRI tumor images and non-tumor brain MRI images. The Br35H dataset was used to define an initial photoacoustic stressor in k-waves to create

simulated PAT images. In total, 801 of these MRI images with annotated brain tumors were selected. The medium was assumed to be homogeneous, with a sound velocity of 1500 m/s and an attenuation coefficient of 0.75 dB/(MHz·cm), similar to human soft tissues. The sensor array has 64, 128, and 256 equally spaced detectors on a circle with a radius of 100 pixels to receive the photoacoustic waves. The reconstructed acoustic field images were post-processed with filtering, denoising, and interpolation operations to create a simulated PAT image dataset for deep learning reconstruction of photoacoustic images.

During the experiment, the dataset is divided into a training set of 500 images, a validation set of 201 images, and a test set of 100 images. All the results are tested on the test set. The YOLOv8 network is trained to localize and classify the sources and artifacts of photoacoustic images and remove the artifacts outside the sources. Then, the artifacts overlapping the signal sources are removed using the Pix2Pix algorithm. Finally, the artifact-removed image is used to segment the brain tumor using the YOLOv8 network to obtain an optoacoustic image that is artifact-free and accurately segments the brain tumor.

3.3.1. The Experimental Results and Analysis of the Removal of Artifacts

In the experiment, the traditional Gaussian filter and deep learning models (including CycleGAN, YOLOv8, and Pix2Pix and their combinations) were compared to remove artifacts in photoacoustic images. PSNR and SSIM are used as quantitative indicators of image reconstruction quality.

Figure 4 demonstrates the effectiveness of ground truth images with various methods for removing artifacts in photoacoustic images. In the comparison experiments, the effectiveness of different methods in removing artifacts in photoacoustic images was observed and evaluated in detail. Through the comparison of the image effect graphs, it is found that the various methods present different characteristics in dealing with artifacts. It is noticed that the photoacoustic images directly reconstructed by the traditional TR method suffer from obvious blurring and information loss, and the artifacts are significantly present at the internal and external boundaries of the images. In addition, the artifact-removal effect of Gaussian filter is worse, which indicates that it is not suitable to be applied for photoacoustic image de-artifacting. The CycleGAN algorithm performs well in removing internal artifacts with improved clarity of the internal structure, but there are still challenges in removing artifacts at the external and brain boundaries. The YOLOv8 algorithm performs better in dealing with the external artifacts; however, the internal image is still subject to blurring and lack of information, and the artifacts are more noticeable. The Pix2Pix algorithm improves the internal information with the Pix2Pix algorithm, which has achieved some success, but the handling of external artifacts still needs to be improved. The method of combining the YOLOv8 algorithm with CycleGAN can more effectively remove internal and external artifacts, but there are limitations in artifact removal at the brain boundary. However, the method combining the YOLOv8 algorithm with Pix2Pix shows much better results, by not only effectively removing internal and external artifacts, but also better preserving the detailed structure and texture of the image. In addition to the visual observation of the image effect, the evaluation metrics, including PSNR and SSIM, indicate that the artifact-removal performance of the YOLOv8-Pix2Pix method is superior to other methods, achieving the highest values. Comprehensive experimental results show that the method combining the YOLOv8 algorithm and Pix2Pix exhibits significant advantages in both the artifact-removal effect and image quality evaluation, providing a feasible deep learning solution for photoacoustic image processing.

Table 3 shows the average numerical results of the reconstruction of each method on the test set. The TR method is the most primitive direct reconstruction method with a PSNR of 28.918 dB and an SSIM of 0.566. The Gaussian filter performs poorly, which may be due to the fact that its simple filtering method fails to remove the artifacts efficiently. The YOLOv8 algorithm may be more concerned with the removal of artifacts by focusing more on the noise elimination, but neglected the preservation of image details and structures, resulting in the lowest SSIM. Among the comparison algorithms, YOLOv8-Pix2Pix shows

the best effect with a PSNR of 31.459 dB and an SSIM of 0.650, which indicates that this method can significantly improve the reconstruction quality while maintaining the image details, and proves that it is an effective photoacoustic image artifact-removal method.
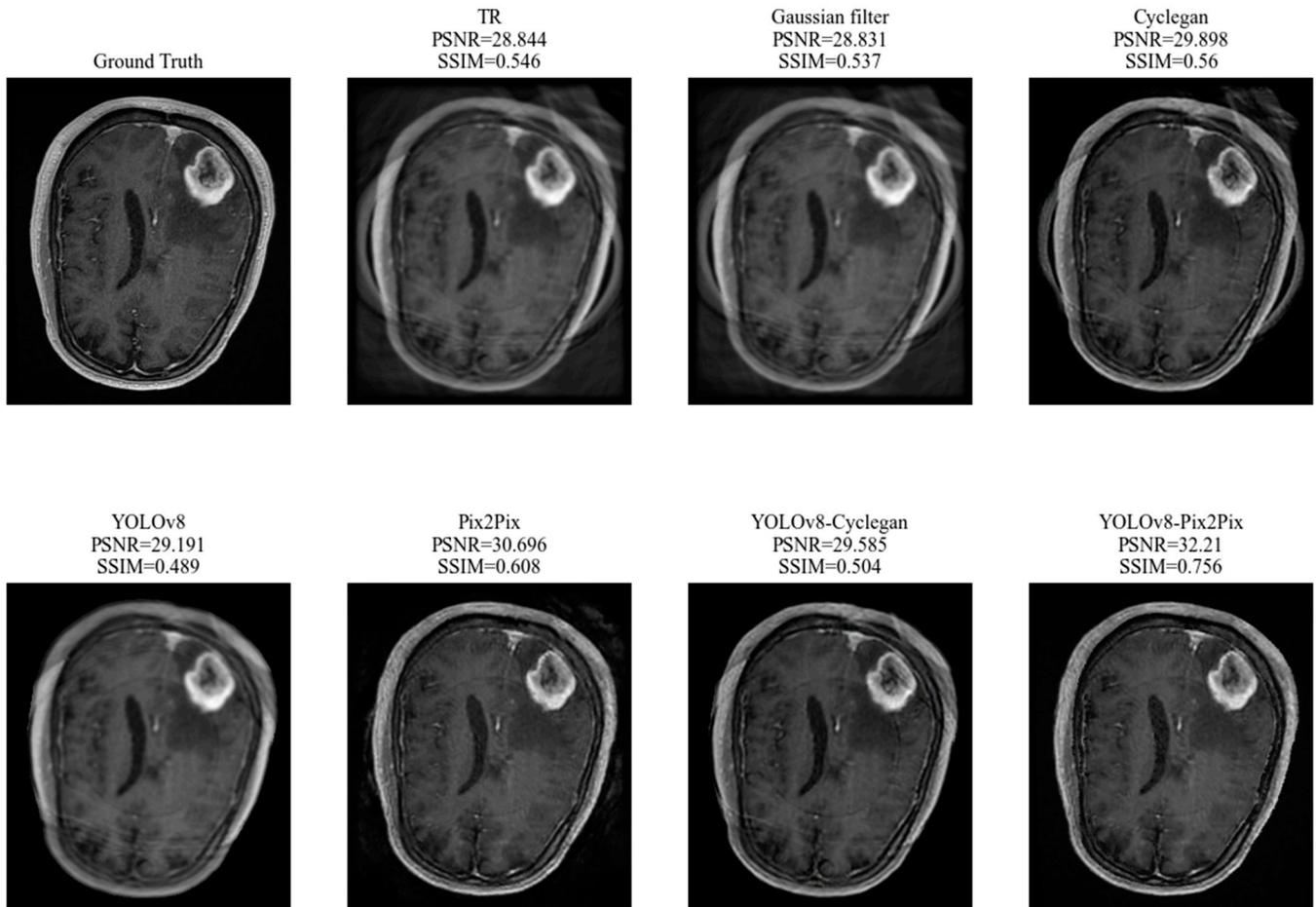

Figure 4. Removal of artifacts in photoacoustic images using different methods.

Table 3. Mean PSNR and SSIM for different reconstruction methods.

	TR	Gaussian Filter	Cyclegan	Pix2Pix	YOLOv8	YOLOv8-Cyclegan	YOLOv8-Pix2Pix
PSNR	28.918	28.821	29.998	30.421	29.113	29.568	31.459
SSIM	0.566	0.526	0.571	0.612	0.491	0.513	0.650

Figure 5 shows the results of artifact removal for Pix2Pix and YOLOv8-Pix2Pix at different detectors. As expected, reducing the number of detectors used to sample the sound pressure leads to more severe artifacts as well as lower average PSNR and SSIM. Compared to Pix2Pix, YOLOv8-Pix2Pix has a higher average PSNR and SSIM at all sample sparsity levels tested. Both algorithms effectively remove internal and external artifacts from the reconstructed images when using 256 detectors, but the Pix2Pix algorithm suffers from some external artifacts. At the sampling level using 128 detectors, YOLOv8-Pix2Pix recovered images of higher quality than Pix2Pix. Artifacts are indicated by the red boxes in Figure 5. Blurring is present in Pix2Pix but can be clearly removed in the YOLOv8-Pix2Pix reconstruction. Neither algorithm was able to reliably reconstruct the internal information of the photoacoustic image at the sparse level using 64 detectors.
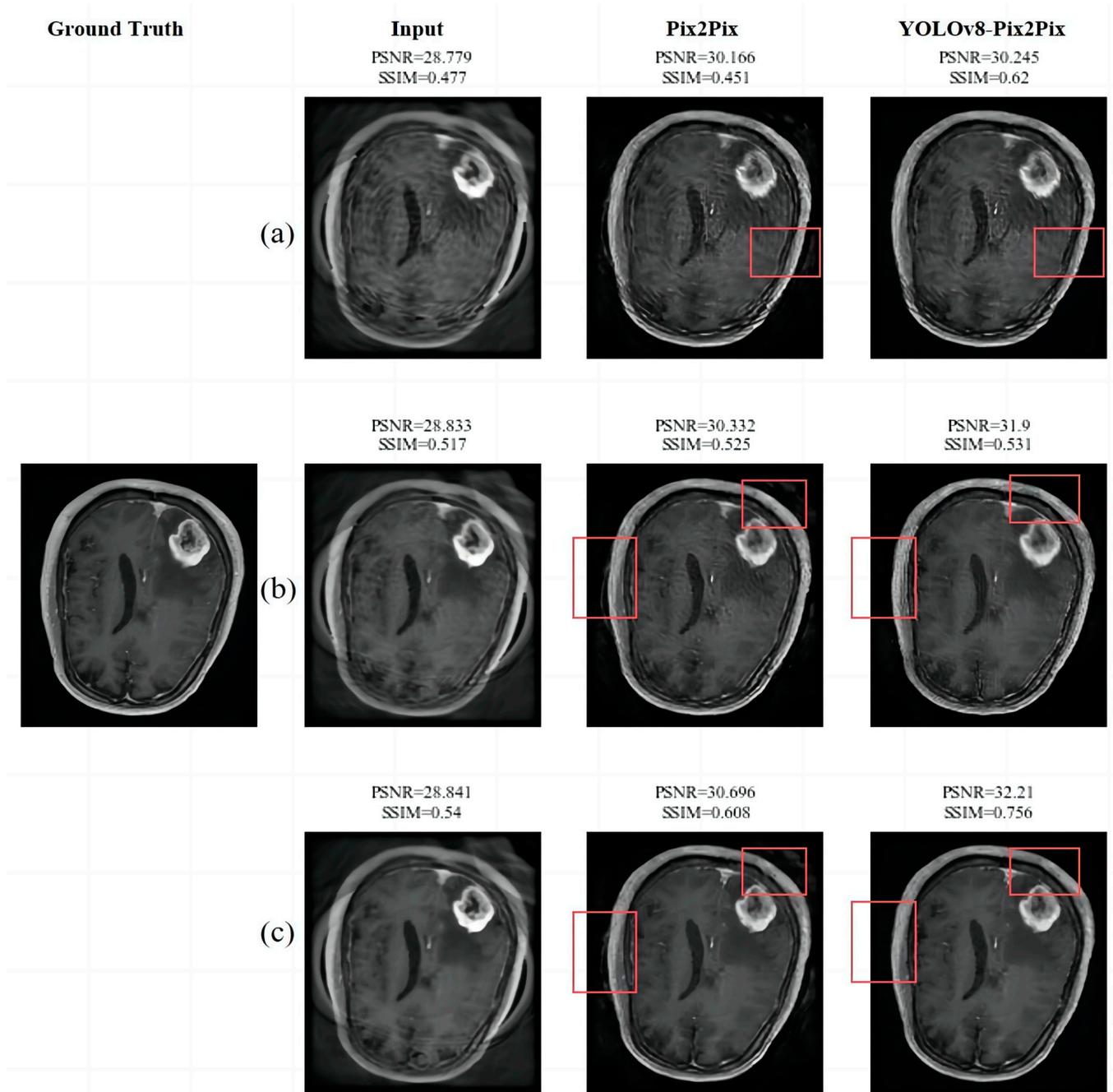
Figure 5. Removal of artifact images using (**a**) 64, (**b**) 128 and (**c**) 256 sensors in different algorithms. The red box is used to focus the artifact removal effect.

The time required to process a single brain tumor image using YOLOv8-Pix2Pix is as follows: the preprocessing time was 0.5 ms, inference time was 11.7 ms, loss calculation time was 0.0 ms, and post-processing time was 1.8 ms. These metrics show the efficiency of the model in all processing stages.

3.3.2. Photoacoustic Image Segmentation Experimental Results and Analysis

The segmentation effects of removing artifacts' photoacoustic images and photoacoustic images using YOLOv8n under the same training set are compared in photoacoustic image segmentation experiments. The predicted segmentation labels are compared with ground truth images using pixel accuracy, intersection and union ratios, and Dice coefficients as quantitative metrics of image segmentation quality.

Figure 6 demonstrates the segmentation effect of the original image MRI, photoacoustic image and the photoacoustic image with artifacts removed on a single image. Evaluating the segmentation results, it can be observed that the segmentation effect of the artifact-removed image is significantly better than that of the photoacoustic image with no artifacts removed. The segmentation labels of the artifact-removed image are highly consistent with the original image labels, with clear morphological edges, presenting excellent segmentation accuracy and precision. On the contrary, the segmentation labels of the image without artifact removal showed missing information and blurred edges, which could not correctly capture the subtle features in the image.
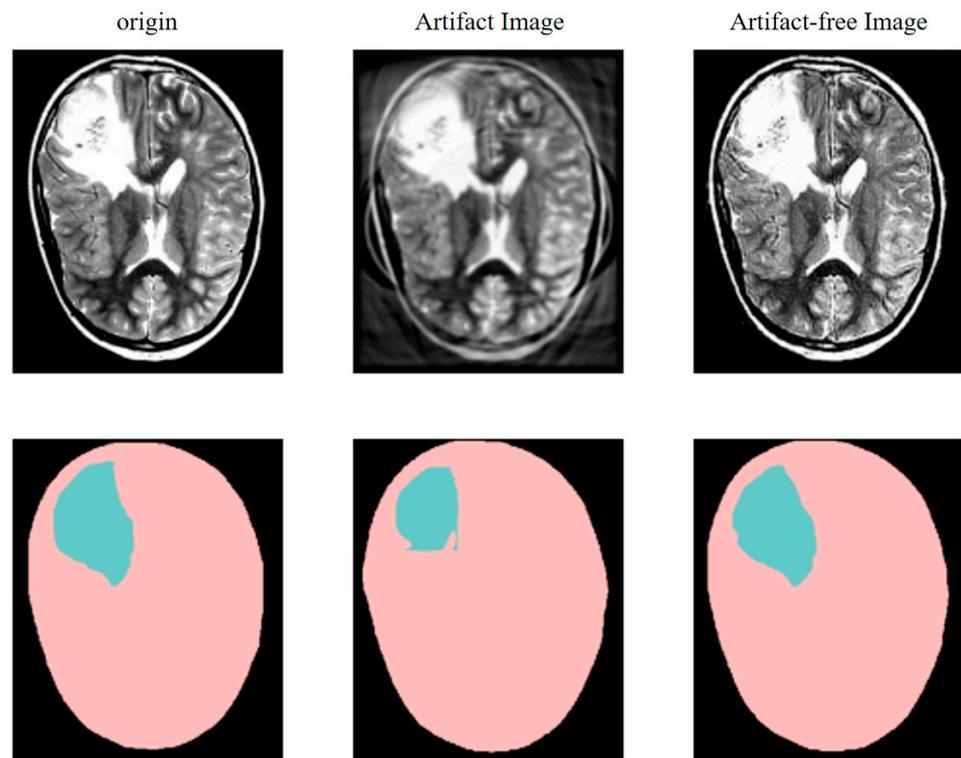

Figure 6. Segmentation results of the original image, photoacoustic image and removal of artifacts.

Table 4 demonstrates the segmentation effect of the brain and brain tumors in photoacoustic images with artifacts versus those with artifacts removed in the test dataset. The results show that the segmentation of the photoacoustic image with artifacts removed is significantly more effective compared to the image with artifacts. Specifically, the artifact-removed image exhibited an overall precision of 0.955, a recall of 0.932, a pixel-level accuracy of 0.986, an intersection and merger ratio of 0.946, an average boundary F1 score of 0.972, and a Dice coefficient of 0.973. In comparison, the corresponding metrics for the image with artifacts were 0.915, 0.852, 0.961, 0.773, 0.863, and 0. 864. In addition, the segmentation assessment metrics for the brain and brain tumors are also broken down in Table 1, and the results show that the photoacoustic images with artifacts removed have improved overall precision, recall, and boundary F1 scores across all categories compared to the images with artifacts.

Table 4. Evaluation metrics for brain and brain tumor segmentation in photoacoustic images with artifacts and with artifacts removed.

		Precision	Recall	Accuracy	IoU	Mean BFScore	Dice
Artifact Image	All	0.915	0.852	0.961	0.773	0.863	0.864
	Brain tumor	0.843	0.704	0.970	0.621	0.766	0.766
	Brain	0.986	1	0.952	0.924	0.961	0.961
Artifact-free Image	All	0.955	0.932	0.986	0.946	0.972	0.973
	Brain tumor	0.918	0.864	0.994	0.928	0.963	0.963
	Brain	0.992	1	0.978	0.964	0.982	0.982

3.4. Photoacoustic Detection Experiment

In previous experiments, open brain tumor MRI data were converted to photoacoustic images where YOLOv8-Pix2Pix was used to effectively eliminate artifacts, achieving remarkable results on a standardized dataset. To enhance the diversity of the dataset and the practicality of the proposed algorithm, this experiment was designed as a photoacoustic detection experimental system, and its experimental verification of the photoacoustic imaging-based artifact-removal system was carried out using YOLOv8-Pix2Pix.

3.4.1. Data Acquisition

In the experiments, a homemade imaging platform shown in Figure 7 was used to acquire photoacoustic signals. A pulsed laser (MC Inc. Real Light Technology Co. Ltd., Beijing, China) emitting a 532 nm laser and a multichannel ultrasound system (Vantage 256, Verasonics, Kirkland, WA, USA) were used to send synchronized signals, while photoacoustic signals were acquired using an ultrasound array (L11-5, Verasonics). Subsequently, the acquired photoacoustic images were used for the re-construction of PA imaging using both a conventional and an optimized reconstruction procedure, respectively. The experimental dataset comprised 201 photoacoustic images with different numbers of photoacoustic sources (PAS), and the image size was set to 500×500 pixels ($N_x = 500$, $N_y = 500$). The modeled light sources for photoacoustic data were included. Finally, the images after artifact removal were compared with the original images by qualitative and quantitative methods.
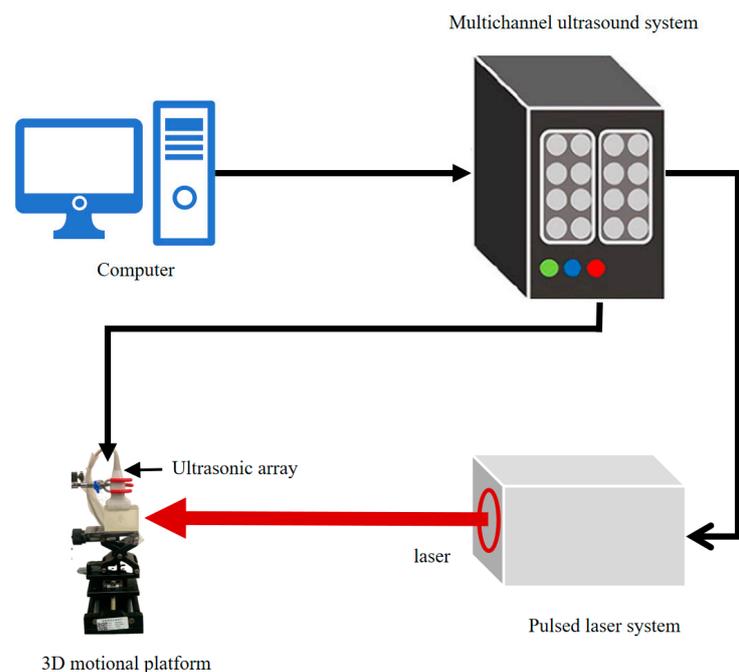
Figure 7. The experiment setup of the PA imaging system.

The experiment involves four mimetic models, each employing a different photoacoustic source configuration. Model I use a pencil core as the photoacoustic source due to its good optical absorption properties. Model II uses a combination of copper and pencil cores to achieve richer photoacoustic effects. Models III and IV use a copper core, a pencil core, and a steel wire as photoacoustic sources, respectively. Figure 8 shows in detail the configurations of Models I to IV and their photoacoustic sources.
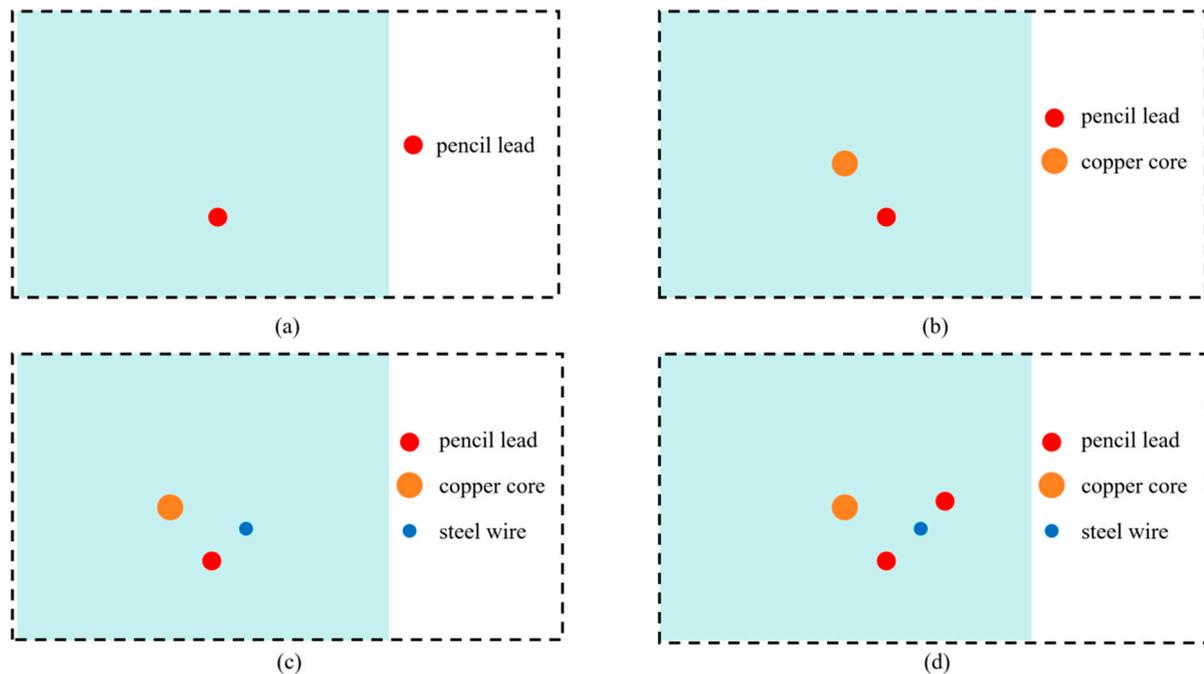

Figure 8. The schematics of the experiment samples. (**a**) Phantom I. (**b**) Phantom II. (**c**) Phantom III. (**d**) Phantom IV.

3.4.2. The Experimental Results and Analysis of the Removal of Artifacts

In the artifact-removal experiments, the processing results of the original photoacoustic image were compared with the photoacoustic image processed using the YOLOv8-Pix2Pix model across different numbers of photoacoustic sources. The reconstructed photoacoustic images were compared with their standard maps using PSNR and SSIM as quantitative indicators of image reconstruction quality.

Figure 9 illustrates the comparison of the original PAI image with the algorithmically processed image and the validation map after fusion with ultrasound imaging from a single photoacoustic source to a multi-photoacoustic source scenario, with the photoacoustic sources labeled PAS1, PAS2, and PAS3 in Figure 9. The artifacts in the original PAI image increase with the number of photoacoustic sources, which is especially significant in Figure 9d. This increase in artifacts is mainly due to the boundary reflection signal (BRS) interactions between multiple photoacoustic sources as well as due to the fact that a single photoacoustic source cannot uniformly illuminate all the photoacoustic sources, which leads to more artifacts and degradation of the imaging quality. The improved algorithm showed significant results in removing these artifacts, as shown in the middle figure in Figure 9d, clearly removing these complex artifacts. By fusing the PAI image with the US image, it is verified that the improved algorithm is able to accurately identify and localize the photoacoustic source even in the presence of multiple BRS reflections and irradiations.
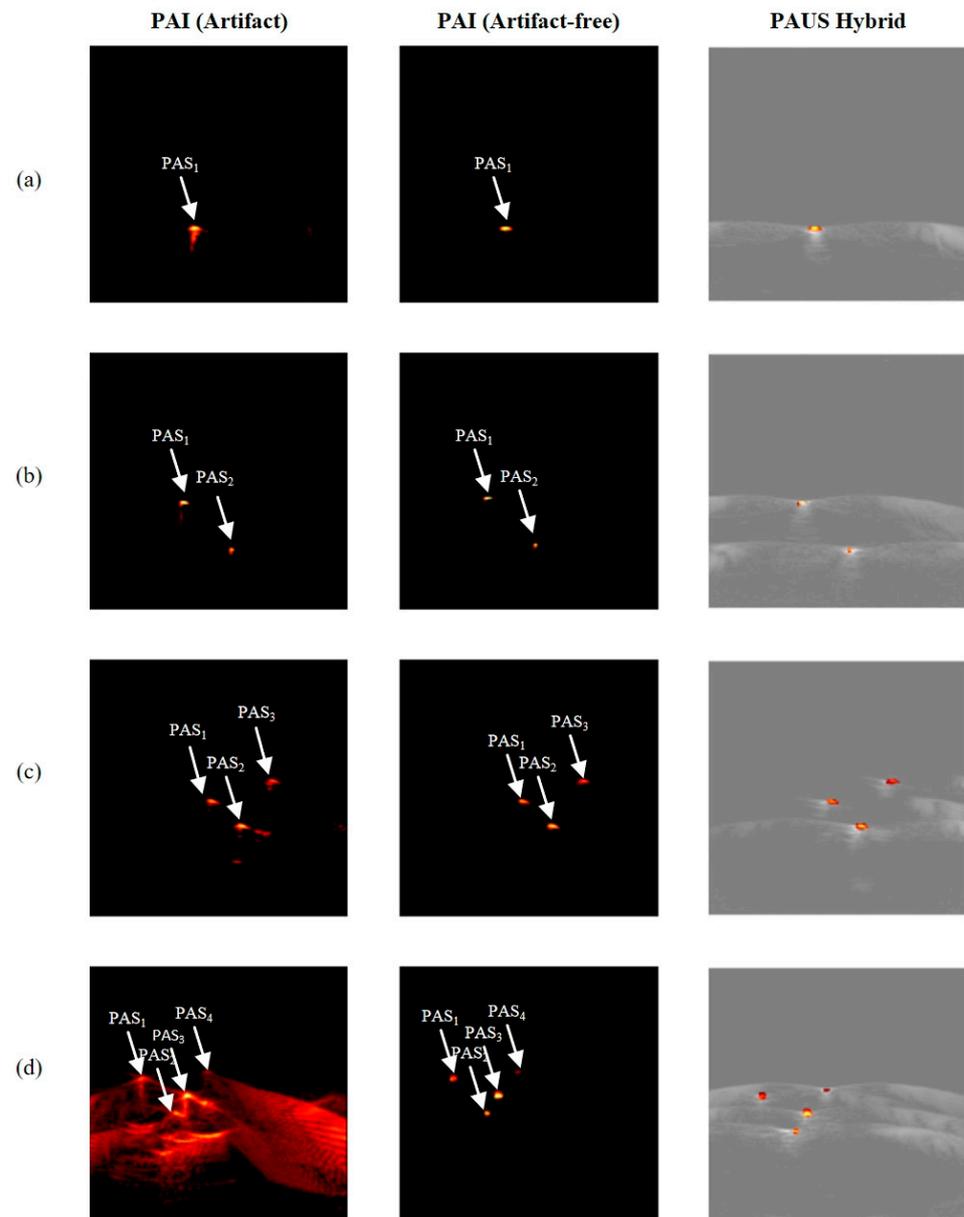
Figure 9. Experimental imaging results using (**a**) 1, (**b**) 2, (**c**) 3, (**d**) 4 photoacoustic sources.

Table 5 demonstrates the evaluation metrics of the original photoacoustic image compared with the photoacoustic image after artifact removal. Clearly, the PSNR and SSIM of the mimic photoacoustic images processed using the YOLOv8-Pix2Pix model are significantly improved compared with those before artifact removal. Even in the Phantom IV with complex artifacts and low evaluation indexes, the use of this model can still effectively recover the image information and improve the values of PSNR and SSIM.

Table 5. Evaluation metrics for de-artifacting in photoacoustic images with and without artifacts.

		Phantom I	Phantom II	Phantom III	Phantom IV
Artifact Image	PSNR	74.089	78.531	71.162	54.164
	SSIM	0.993	0.995	0.985	0.646
Artifact-free Image	PSNR	79.988	78.933	74.514	73.616
	SSIM	0.998	0.997	0.996	0.992

According to the experimental results, the time required to process a single image in a photoacoustic detection experiment using YOLOv8-Pix2Pix is as follows: the preprocessing time is 2.0 ms, inference time is 22.6 ms, loss computation time is 0.0 ms, and post-processing time is 2.2 ms.

4. Discussion

In this paper, we propose a deep learning-based photoacoustic imaging artifact-removal and segmentation algorithm to solve the problems of single means and poor adaptability in the existing technology. The proposed YOLOv8-Pix2Pix model is able to effectively remove artifacts outside and inside the signal source, and segmentation is performed by YOLOv8 for processing and analyzing the simulated photoacoustic data of brain tumors and experimentally acquired photoacoustic data. The experimental results show that the proposed algorithm exhibits significant advantages over other methods on the brain tumor dataset, not only effectively removing the internal and external artifacts, but also better preserving the detailed structure and texture of the image. In the segmentation experiments, the segmentation labels of the images after artifact removal are highly consistent with the original image labels, and the morphological edges are clear, presenting excellent segmentation precision and accuracy. In contrast, the image segmentation labels without removing artifacts showed missing information and blurred edges, and could not accurately capture the subtle features in the image. In order to verify the effectiveness of the proposed algorithm and further explore its potential in practical applications, we designed an experimental system for photoacoustic detection. The proposed algorithm is able to maintain stable performance under complex backgrounds and artifacts, accurately identify and localize photoacoustic sources, and clearly remove complex artifacts. The deep learning-based artifact-removal and segmentation algorithm proposed in this paper has a broad application prospect in the field of photoacoustic imaging, which can provide more accurate and clearer image information for medical image processing and is expected to play an important role in clinical diagnosis and treatment. Future research will focus on optimizing the model and expanding the range of tests to verify its performance under different conditions, thereby driving its broader application in the field of medical image processing.

Author Contributions: Conceptualization, R.H.; methodology, R.H.; software, R.H.; validation, R.H., Y.C., Y.J., S.Y., Y.L. and J.Z.; data curation, R.H.; writing—original draft preparation, R.H.; writing—review and editing, R.H., Y.C. and Y.J.; visualization, Y.C.; supervision, H.C.; funding acquisition, H.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, Y.; Glorieux, C.; Yang, S.; Gu, K.; Xia, Z.; Hou, R.; Hou, L.; Liu, X.; Xiong, J. Adaptive polarization photoacoustic computed tomography for biological anisotropic tissue imaging. *Photoacoustics* **2023**, *32*, 100543. [CrossRef]
2. Hauptmann, A.; Cox, B. Deep learning in photoacoustic tomography: Current approaches and future directions. *J. Biomed. Opt.* **2020**, *25*, 112903. [CrossRef]
3. Dadkhah, A.; Jiao, S. Integrating photoacoustic microscopy with other imaging technologies for multimodal imaging. *Exp. Biol. Med.* **2021**, *246*, 771–777. [CrossRef] [PubMed]

4.   Singh, M.K.A.; Xia, W. Portable and Affordable Light Source-Based Photoacoustic Tomography. *Sensors* **2020**, *20*, 6173. [CrossRef]
5.   Chen, Q.; Qin, W.; Qi, W.; Xi, L. Progress of clinical translation of handheld and semi-handheld photoacoustic imaging. *Photoacoustics* **2021**, *22*, 100264. [CrossRef] [PubMed]
6.   Zheng, Y.; Liu, M.; Jiang, L. Progress of photoacoustic imaging combined with targeted photoacoustic contrast agents in tumor molecular imaging. *Front. Chem.* **2022**, *10*, 1077937. [CrossRef] [PubMed]
7.   Han, S.; Ninjbadgar, T.; Kang, M.; Kim, C.; Kim, J. Recent advances in photoacoustic agents for theranostic applications. *Nanomaterials* **2023**, *13*, 695. [CrossRef] [PubMed]
8.   Yu, Y.; Feng, T.; Qiu, H.; Gu, Y.; Chen, Q.; Zuo, C.; Ma, H. Simultaneous photoacoustic and ultrasound imaging: A review. *Ultrasonics* **2024**, *139*, 107277. [CrossRef] [PubMed]
9.   Zhang, C.; Liu, G.; Xie, Z.; Shu, Z.; Ren, Z.; Yao, Q. Vascular recognition system based on photoacoustic detection. *J. Laser Appl.* **2021**, *33*, 012051. [CrossRef]
10.   Meng, J.; Yu, J.; Wu, Z.; Ma, F.; Zhang, Y.; Liu, C. WSA-MP-Net: Weak-signal-attention and multi-scale perception network for microvascular extraction in optical-resolution photoacoustic microcopy. *Photoacoustics* **2024**, *37*, 100600. [CrossRef]
11.   Guan, S.; Khan, A.A.; Sikdar, S.; Chitnis, P.V. Fully Dense UNet for 2-D Sparse Photoacoustic Tomography Artifact Removal. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 568–576. [CrossRef] [PubMed]
12.   Shen, K.; Liu, S.; Feng, T.; Yuan, J.; Zhu, B.; Tian, C. Negativity artifacts in back-projection based photoacoustic tomography. *J. Phys. D-Appl. Phys.* **2021**, *54*, 074001. [CrossRef]
13.   Qi, Y.; Cao, H.; Yin, G.; Zhang, B.; Guo, J. Optimized Reconstruction Procedure of Photoacoustic Imaging for Reflection Artifacts Reduction. *Ultrason. Imaging* **2022**, *44*, 204–212. [CrossRef] [PubMed]
14.   Guo, M.; Lan, H.; Yang, C.; Liu, J.; Gao, F. AS-Net: Fast Photoacoustic Reconstruction with Multi-Feature Fusion from Sparse Data. *IEEE Trans. Comput. Imaging* **2022**, *8*, 215–223. [CrossRef]
15.   Nguyen, H.N.Y.; Steenbergen, W. Feasibility of identifying reflection artifacts in photoacoustic imaging using two-wavelength excitation. *Biomed. Opt. Express* **2020**, *11*, 5745–5759. [CrossRef] [PubMed]
16.   Huang, M.; Liu, W.; Sun, G.; Shi, C.; Liu, X.; Han, K.; Liu, S.; Wang, Z.; Xie, Z.; Guo, Q. Unveiling precision: A data-driven approach to enhance *Photoacoust.* imaging with sparse data. *Biomed. Opt. Express* **2024**, *15*, 28–43. [CrossRef] [PubMed]
17.   Lu, M.; Liu, X.; Liu, C.; Li, B.; Gu, W.; Jiang, J.; Ta, D. Artifact removal in photoacoustic tomography with an unsupervised method. *Biomed. Opt. Express* **2021**, *12*, 6284–6299. [CrossRef]
18.   Hakakzadeh, S.; Amjadian, M.; Zhang, Y.; Mostafavi, S.M.; Kavehvash, Z.; Wang, L. Signal restoration algorithm for photoacoustic imaging systems. *Biomed. Opt. Express* **2023**, *14*, 651–666. [CrossRef]
19.   Hakakzadeh, S.; Rajendran, P.; Nili, V.A.; Kavehvash, Z.; Pramanik, M. A Spatial-Domain Factor for Sparse-Sampling Circular-View Photoacoustic Tomography. *IEEE J. Sel. Top. Quantum Electron.* **2023**, *29*, 6800409. [CrossRef]
20.   Judith, A.M.; Priya, S.B.; Mahendran, R.K. Artifact Removal from EEG signals using Regenerative Multi-Dimensional Singular Value Decomposition and Independent Component Analysis. *Biomed. Signal Process. Control* **2022**, *74*, 103452. [CrossRef]
21.   Yang, Y.; Hui, H.J.; Zeng, L.L.; Zhao, Y.; Zhan, Y.Z.; Yan, T. Edge-Preserving Image Filtering Based on Soft Clustering. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 4150–4162. [CrossRef]
22.   Chen, W.; Tao, C.; Hu, Z.; Yuan, S.; Liu, Q.; Liu, X. Non-invasive and low-artifactin vivo brain imaging by using a scanning acoustic-photoacoustic dual mode microscopy. *Chin. Phys. B* **2022**, *31*, 044304. [CrossRef]
23.   Singh, M.K.A.; Jaeger, M.; Frenz, M.; Steenbergen, W. In vivo demonstration of reflection artifact reduction in photoacoustic imaging using synthetic aperture photoacoustic-guided focused ultrasound (PAFUSion). *Biomed. Opt. Express* **2016**, *7*, 2955–2972. [CrossRef]
24.   Godefroy, G.; Arnal, B.; Bossy, E. Compensating for visibility artefacts in photoacoustic imaging with a deep learning approach providing prediction uncertainties. *Photoacoustics* **2021**, *21*, 100218. [CrossRef]
25.   Shahid, H.; Khalid, A.; Yue, Y.; Liu, X.; Ta, D. Feasibility of a generative adversarial network for artifact removal in experimental photoacoustic imaging. *Ultrasound Med. Biol.* **2022**, *48*, 1628–1643. [CrossRef] [PubMed]
26.   Lan, H.; Jiang, D.; Yang, C.; Gao, F.; Gao, F. Y-Net: Hybrid deep learning image reconstruction for photoacoustic tomography in vivo. *Photoacoustics* **2020**, *20*, 100197. [CrossRef] [PubMed]
27.   Vu, T.; Li, M.; Humayun, H.; Zhou, Y.; Yao, J. A generative adversarial network for artifact removal in photoacoustic computed tomography with a linear-array transducer. *Exp. Biol. Med.* **2020**, *245*, 597–605. [CrossRef] [PubMed]
28.   Zhang, H.; Li, H.; Nyayapathi, N.; Wang, D.; Le, A.; Ying, L.; Xia, J. A New Deep Learning Network for Mitigating Limited-view and Under-sampling Artifacts in Ring-shaped Photoacoustic Tomography. *Comput. Med. Imaging Graph.* **2020**, *84*, 101720. [CrossRef]
29.   Guan, S.; Khan, A.A.; Sikdar, S.; Chitnis, P.V. Limited-View and Sparse Photoacoustic Tomography for Neuroimaging with Deep Learning. *Sci. Rep.* **2020**, *10*, 8510. [CrossRef]
30.   Liang, B.; Wang, S.; Shen, F.; Liu, Q.H.; Gong, Y.; Yao, J. Acoustic impact of the human skull on transcranial photoacoustic imaging. *Biomed. Opt. Express* **2021**, *12*, 1512–1528. [CrossRef]
31.   Kang, M.; Ting, C.-M.; Ting, F.F.; Phan, R.C.-W. Bgf-yolo: Enhanced yolov8 with multiscale attentional feature fusion for brain tumor detection. *arXiv* **2023**, arXiv:2309.12585.
32.   Dong, C.; Du, G. An enhanced real-time human pose estimation method based on modified YOLOv8 framework. *Sci. Rep.* **2024**, *14*, 8012. [CrossRef]

33. Reis, D.; Kupec, J.; Hong, J.; Daoudi, A. Real-time flying object detection with YOLOv8. *arXiv* **2023**, arXiv:2305.09972.
34. Zhang, J.; Gong, Q.; Zhang, H.; Wang, Y.; Wang, Y. A Novel Pix2Pix Enabled Traveling Wave-Based Fault Location Method. *Sensors* **2021**, *21*, 1633. [CrossRef]
35. Xu, T.; Yan, H.; Yu, H.; Zhang, Z. Removing Time Dispersion from Elastic Wave Modeling with the pix2pix Algorithm Based on cGAN. *Remote Sens.* **2023**, *15*, 3120. [CrossRef]
36. Karimipourfard, M.; Sina, S.; Mahani, H.; Karimkhani, S.; Sadeghi, M.; Alavi, M.; Faghihi, R. A Taguchi-optimized Pix2pix generative adversarial network for internal dosimetry in 18F-FDG PET/CT. *Radiat. Phys. Chem.* **2024**, *218*, 111532. [CrossRef]