Check for updates

RESEARCH ARTICLE

ADVANCED SENSOR RESEARCH Open Access

www.advsensorres.com

A Holographic Sensor-Integrated Deep Learning Framework for Noninvasive Assessment of Stored Red Blood Cell Quality

*Seonghwan Park, Hyunbin An, Abdur Rehman, and Inkyu Moon**

Prolonged storage of red blood cells (RBCs) induces morphological degradation that can compromise transfusion efficacy. Traditional quality assessment methods are often labor-intensive and time-consuming, limiting their utility in real-time settings. Although deep learning has been applied to RBC imaging, most approaches require large datasets and complex architectures, making them impractical for efficient deployment. This study introduces a holographic sensor-integrated deep learning framework for noninvasive RBC quality assessment using small datasets. A diffusion model is employed to synthetically generate phase images and segmentation masks, augmenting limited data. Self-supervised learning with pre-trained models further enhances classification performance while maintaining a streamlined model architecture. Compared to conventional segmentation methods, the proposed framework achieves higher accuracy and significantly faster inference. It also enables reliable detection of storage-induced morphological changes, providing proportional indicators of transfusion viability. Experimental results validate its effectiveness as a practical tool for real-time, sensor-driven monitoring of RBC quality.

1. Introduction

Red blood cells (RBCs) transport oxygen from the lungs to tissues and carbon dioxide from tissues to the lungs, thereby playing a crucial role in human physiology.[1] This essential function supports metabolism and energy production throughout the body, making the health and functionality of RBCs vital to overall health. Healthy RBCs are biconcave disks, which property allows them to deform as they pass through narrow capillaries and optimize their oxygen-carrying capacity. However, storage can compromise the shape and structural integrity of RBCs.[2–5] RBCs can be stored for up to 49 days, as is well-documented; but as storage time increases, significant morphological changes can occur. These changes include membrane loss, shape transformation to a more spherical form, and reduced deformability, all of which can impair the ability of RBCs to effectively perform their essential functions. Degradation of RBC morphology over time during storage is a critical factor to consider in transfusion medicine. Transfusing RBCs with abnormal shapes can lead to reduced oxygen delivery, increased hemolysis, and other complications. Therefore, research into technologies that can quickly and accurately assess RBC morphology is essential to improve transfusion outcomes and ensure the efficacy and safety of blood transfusions. However, RBC observation via conventional imaging systems such as bright-field microscopy cannot provide important biophysical cellular parameters and therefore does not offer sufficient detail to reliably distinguish between different RBC morphologies.[6,7]

Digital holographic sensing and imaging systems has emerged as a powerful imaging technique that enables quantitative, label-free analysis of transparent biological samples. Its ability to capture optical phase information has led to widespread applications in the morphological analysis of red blood cells,[8–11] cancer cells,[12,13] and cardiomyocytes.[14,15] Among the various holographic imaging techniques, digital in-line holographic microscopy (DIHM) has been employed for RBC classification by extracting geometric and optical features such as real and virtual focal lengths, projected area, and perimeter.[16,17] While DIHM offers advantages in terms of system simplicity and compactness, it inherently suffers from twin-image artifacts due to the inline configuration of the reference and object beams. These artifacts degrade the accuracy of phase reconstruction and typically require computationally demanding post-processing, limiting its suitability for high-throughput or real-time applications.

As an alternative configuration, off-axis digital holographic microscopy (DHM) has been developed, in which the reference and object beams are spatially separated. This spatial separation enables the resolution of twin-image artifacts in the frequency

S. Park, H. An, A. Rehman, I. Moon
Department of Robotics & Mechatronics Engineering
DGIST
Daegu 42988, South Korea
E-mail: inkyu.moon@dgist.ac.kr

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adsr.202500073

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adsr.202500073

domain, allowing for direct and artifact-free phase reconstruction. In off-axis DHM, holograms are recorded as spatial interference patterns generated by the coherent interaction of the reference and object waves transmitted through biological specimens.[18,19] These holograms are then numerically propagated using scalar diffraction theory under the Fresnel approximation, followed by phase retrieval to produce quantitative phase images (QPIs). These QPIs provide high-contrast, high-resolution optical thickness maps that support detailed morphological and biophysical analysis of transparent samples such as RBCs.

Deep learning has emerged as a powerful tool in recent years, bringing significant advancements in fields such as computer vision. Recognizing this potential, many researchers have applied deep learning models to the phase images of biological samples obtained from digital holography.[13,20–22] O'Connor et al.[23] demonstrated a successful deep learning strategy for cell identification and disease diagnosis that used spatio-temporal cell information recorded by a compact digital holographic microscopy system, while O'Connor et al.[24] demonstrated the use of a digital holographic deep learning system for rapid COVID−19 screening that used a compact and field-portable microscope to analyze the spatiotemporal behavior of red blood cells. Kim et al.[25] introduced a GAN-based approach for the phenotypic assessment of red blood cell that used digital holographic microscopy to achieve high classification accuracy and efficient segmentation. Kim et al.[26] proposed an AI-based digital holographic microscopy technique that accurately measured the 3D position and orientation of red blood cells to enhance the analysis of their dynamic behavior. However, previous studies have generally improved the performance of deep learning models by increasing the amount of data, or employing complex architectures for training. But in the fields of medicine and biology, it is extremely challenging to obtain large-scale datasets. Additionally, while making the model deeper and more complex can boost performance, the use of limited data comes with the risk of overfitting, and can also slow down the speed at which image results are produced.

We introduce a holographic sensor-integrated deep learning framework for generating RBC segmentation masks that classify four distinct cell types, utilizing phase images captured via digital holographic sensing (**Figure** 1a). We suggest an approach that uses a small amount of data to achieve high accuracy and rapid analysis. The model construction is based on two approaches (Figure 1b): data augmentation using a diffusion model, and model pretraining based on self-supervised learning. Diffusion models are generative AI techniques that progressively remove noise to reconstruct target images.[27,28] Compared to previously widely used generative adversarial networks (GANs),[29] diffusion models demonstrate superior accuracy. Self-supervised learning (SSL) provides a powerful technique that allows a model to learn from unlabeled data, thereby reducing dependence on annotated datasets.[30] SSL involves pretraining a model with unlabeled data, and fine-tuning it with a small amount of labeled data to perform such tasks as classification, segmentation, or object detection. By leveraging these two methods, it becomes unnecessary to invest additional time and resources in acquiring larger datasets, or to modify the structure of the segmentation model to increase its complexity. Consequently, the model achieves high accuracy even with limited data, while maintaining a more straightforward structure that facilitates faster analysis.

We obtained phase images of red blood cells classified into four types using digital holographic microscopy. The RBCs in the phase images were manually classified into these four types to create segmentation masks. A diffusion model was used to generate images that resembled the acquired RBC phase images and their corresponding segmentation masks. Additionally, a pretrained model based on ResUnet was employed as the backbone, leveraging self-supervised learning techniques, and training on a well-known blood cell dataset. Through pretraining based on self-supervised learning, and data augmentation using diffusion models, our proposed method improved model performance. In consequence, it outperformed other segmentation models in terms of accuracy, while also delivering faster performance. Furthermore, we evaluated the phenotype of RBC storage lesions, observing morphological changes in RBCs over the storage period (Figure 1c). This approach demonstrates the feasibility of training with a limited dataset while using a less complex model structure, compared to conventional modeling techniques. This represents an ideal training approach for real-time red blood cell storage evaluation analysis.

The contributions of the proposed scheme can be summarized as follows:

1) We propose a holographic sensor-integrated deep learning framework that achieves high accuracy and fast analysis using small amounts of data, eliminating the need for larger datasets or complex model modifications.
2) We use diffusion models to generate RBC phase images and segmentation masks to significantly improve data diversity, and address challenges associated with small datasets.
3) Self-supervised learning on large unlabeled datasets allows the development of lightweight yet highly effective segmentation models, which enable effective fine-tuning on small datasets.
4) Compared to existing models, the proposed method achieves superior segmentation accuracy and faster processing times, demonstrating its suitability for real-time applications in transfusion medicine.
5) This holographic sensing-driven approach enables the sensitive detection and quantitative evaluation of morphological changes in RBCs over storage periods, thereby ensuring cellular integrity and advancing transfusion safety.
6) To the best of our knowledge, this study is the first attempt to simultaneously employ data augmentation based on diffusion model training, and pretraining models based on self-supervised learning, to develop a high-performance and efficient model using a small amount of red blood cell data.

2. Diffusion Model for Data Augmentation

The diffusion model is a generative model that progressively transforms simple noise into complex data distributions through iterative steps and consists of a forward process and a reverse process. The forward process gradually adds Gaussian noise to the data in a series of time steps, effectively transforming the data into a standard normal distribution. The noisy images generated

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

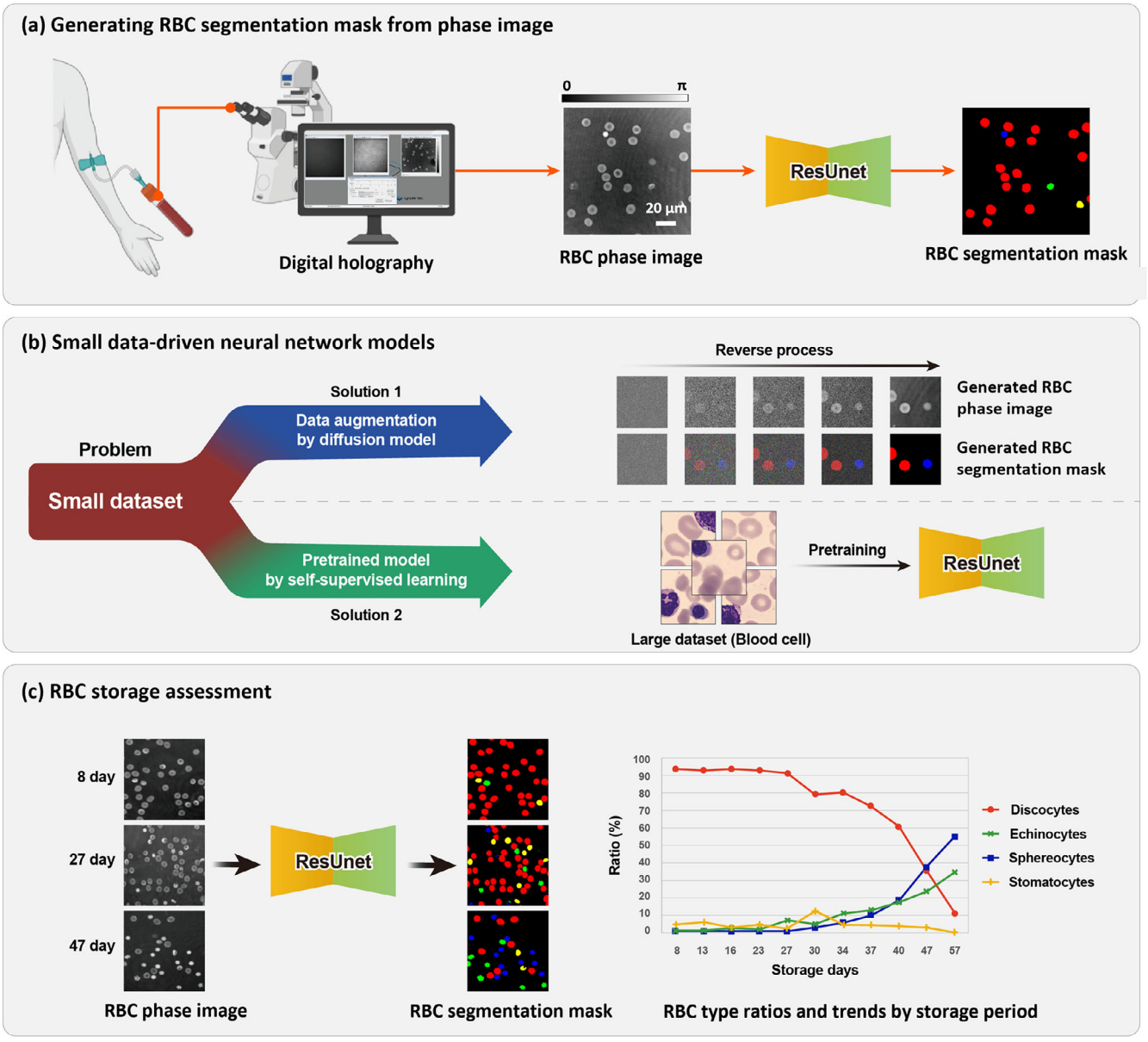

Figure 1. Overview of the proposed method for red blood cell segmentation and storage assessment. a) RBC segmentation process: Digital holography is used to capture RBC phase images, which are segmented into masks using a ResUnet model to identify different RBC types. b) Small data-driven neural network training: To address the challenges of limited datasets, two solutions are employed: data augmentation using a diffusion model to generate synthetic RBC phase images and segmentation masks, and pretraining the ResUnet model on a large blood cell dataset using self-supervised learning. c) RBC storage assessment: RBC phase images at different storage durations are processed by the ResUnet model to generate segmentation masks, enabling calculation of RBC type ratios (discocytes, echinocytes, spherocytes, and stomatocytes) and their trends over time.

by the forward process are used to train the model in the reverse process. The reverse process aims to denoise the data step-by-step, recovering the original data distribution from pure noise. In the reverse process, the model is trained to iteratively denoise the noisy data by learning the mapping between the noisy and clean data distributions, thus enabling efficient sampling with competitive quality and diversity. Our goal is to enhance segmentation performance through data augmentation by using diffusion models to generate a large number of similar samples from a limited dataset. **Figure 2** shows the three-step process that we employ to generate new datasets.

2.1. Training a Diffusion Model to Generate RBC Phase Images

Figure 2a aims to train a diffusion model that generates fake RBC phase images that are similar to real RBC phase images, using a denoising diffusion GAN model that integrates the diffusion

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

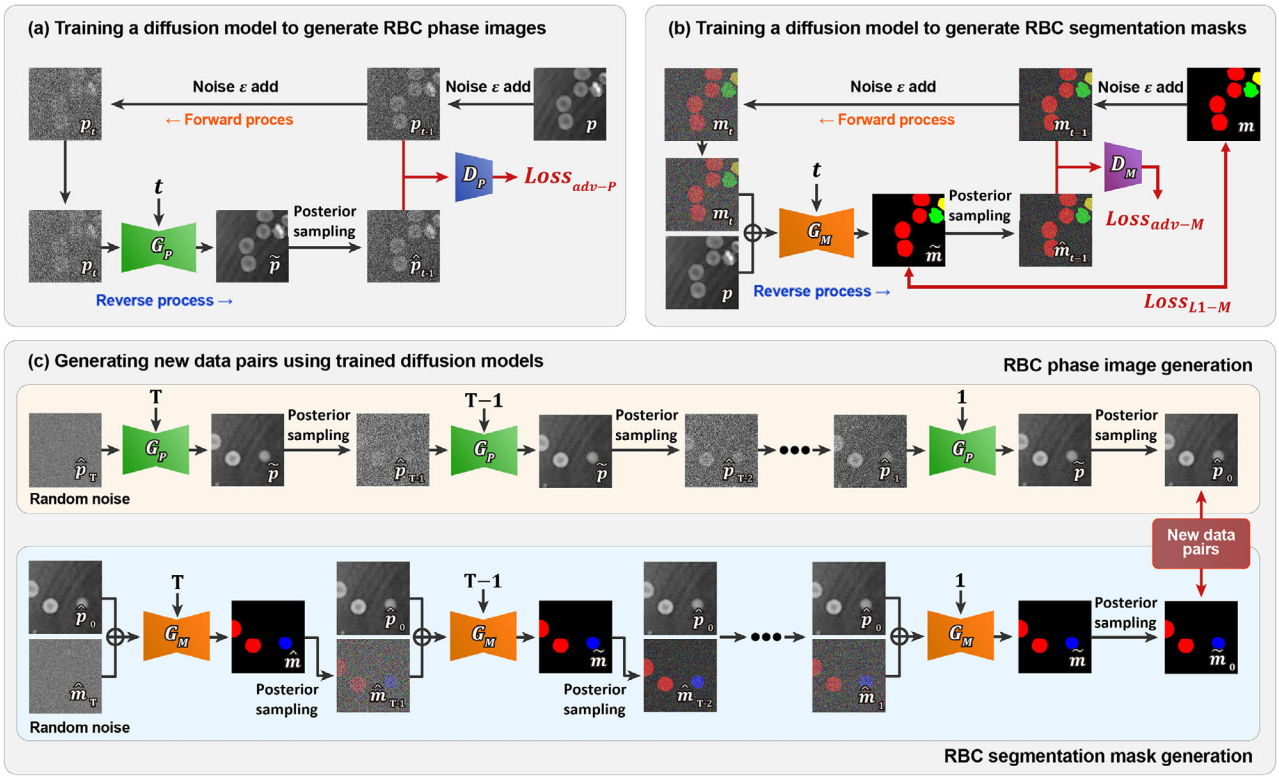

Figure 2. Overview of the data augmentation technique using diffusion model. This consists of a three-step process: a) training a diffusion model to generate RBC phase images, b) training a diffusion model to generate RBC segmentation masks corresponding to the RBC phase images, and c) generating new data pairs using the two trained diffusion models. The diffusion model consists of a generator and a discriminator based on the denoising diffusion GAN. Each image is 256 × 256 pixels, corresponding to a physical size of ≈36.02 μm × 36.02 μm.

process with the GAN framework to improve the quality of image generation. This includes two processes in model training: the forward process, and the backward process.

The forward process gradually adds Gaussian noise to the RBC phase image over time steps from (0 to T):

$$p_t \xleftarrow{noise\ \epsilon\ add} p \tag{1}$$

where p is an original phase image and p_t is a phase image with noise added to p by t timesteps. The reverse process uses the RBC phase image p_t with added noise from the forward process as input. This reverse process uses a trainable model G_p to generate denoised image $\tilde{p}$ from the noisy RBC phase image p_t. The generated denoised image $\tilde{p}$, together with the input noisy image p_t, undergoes posterior sampling to produce image $\hat{p}_{t-1}$ with one step of noise removed:

$$\tilde{p} = G_p\left(p_t, t\right) \tag{2}$$

$$\hat{p}_{t-1} \xleftarrow{Posterior\ Sampling} \left(p_t, \tilde{p}, \epsilon\right) \tag{3}$$

The RBC phase image generator G_p is trained with an adversarial loss where at each timestep, the discriminator D_p differentiates between real and generated RBC phase images:

$$L_{adv-P} = \mathbb{E}\left[-\log\left(D_p\left(t, p_t, \hat{p}_{t-1}\right)\right)\right] \tag{4}$$

The trained model can be used to generate various types of RBC phase images from random noise. Section SA1 (Supporting Information) describes the detailed calculation of training a diffusion model to generate RBC phase images.

2.2. Training a Diffusion Model to Generate RBC Segmentation Masks from RBC Phase Images

Figure 2b aims to train a diffusion model to generate segmentation masks that correspond to real RBC phase images using manually created segmentation masks paired with the real RBC phase images. A denoising diffusion GAN model is used in the same way as the process of training the model to generate the previous phase image, while the model training includes forward and backward processes.

The forward process gradually adds Gaussian noise to the RBC segmentation mask image over the time steps (0 to T).

$$m_t \xleftarrow{noise\ \epsilon\ add} m \tag{5}$$

where m is an original phase image and m_t is a phase image with noise added to m by t timesteps. In the reverse process, the trainable model G_M takes two concatenated images as input: one is the noisy RBC segmentation mask m_t from the forward process, while the other is the RBC phase image p corresponding to the original RBC segmentation mask m. The model G_M generates a

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

denoised RBC segmentation mask $\tilde{m}$ based on these inputs. The generated denoised mask $\tilde{m}$ undergoes posterior sampling along with the input noisy mask m_t, to produce a mask $\hat{m}_{t-1}$ with one step of noise removed:

$$\tilde{m} = G_M\left(m_t, p, t\right) \tag{6}$$

$$\hat{m}_{t-1} \xleftarrow{\text{Posterior Sampling}} \left(m_t, \tilde{m}, \varepsilon\right) \tag{7}$$

The RBC segmentation mask generator G_M is trained using two losses. One is an adversarial loss L_{adv-M}, where the discriminator D_M learns to distinguish real segmentation masks from generated segmentation masks. The other is an L1 loss L_{L1-M}, where the segmentation mask generated by G_M is trained to be the segmentation mask corresponding to the p used in the input.

$$L_{adv-M} = \mathbb{E}\left[-\log\left(D_M\left(t, m_t, \hat{m}_{t-1}\right)\right)\right] \tag{8}$$

$$L_{L1-M} = \mathbb{E}\left[\left\|m_0 - \tilde{m}\right\|_1\right] \tag{9}$$

Using the RBC phase image and random noise as inputs, the trained model enables the generation of RBC segmentation masks that correspond to RBC phase images. Section SA2 (Supporting Information) describes the detailed calculation of training a diffusion model to generate RBC segmentation masks from RBC phase images.

2.3. Generating New Data Pairs Using Trained Diffusion Models

As the last step (Figure 2c), we use the two trained models to generate new data pairs. First, the diffusion model trained in the first step is used to generate an RBC phase image from the random noise. RBC phase images are generated by gradually removing noise from timestep T to 0. During this process, phase images are selected that contain various types of RBCs to enhance the performance of the final model. Next, the diffusion model trained in the second step is used to generate the RBC segmentation mask map corresponding to the previously generated RBC phase image. This process is also performed by gradually removing noise from timestep T to 0. These generated image pairs $(\hat{p}_0, \hat{m}_0)$ are then added to the existing dataset.

3. Self-Supervised Learning for the Pretrained Model

Self-supervised learning (SSL) harnesses large-scale unlabeled datasets by formulating auxiliary tasks, commonly referred to as pretext tasks, to enable models to learn meaningful representations without extensive reliance on labeled data. The primary objective of SSL is to facilitate the transfer of the trained model to downstream tasks, particularly in scenarios where labeled datasets are scarce.

In line with this goal, we adopted the Blood Cell Dataset[31] for the pretraining phase of SSL. This dataset contains 17092 images and is widely used in biomedical image analysis due to its high resolution and diversity. Although it was primarily constructed for white blood cell (WBC) classification, the dataset also includes a substantial number of red blood cells (RBCs) appearing in the background of the images. Therefore, the dataset is not WBC-exclusive but includes both WBC and RBC structures, making it relevant to our target domain.

Furthermore, it is important to note that SSL does not require the pretraining dataset to be strictly task-specific. Rather, it benefits from a large volume of visually similar biomedical images, which allows the model to learn generalizable low-level features such as edge contours, cell morphology, and spatial distribution patterns. In our case, the Blood Cell Dataset provided sufficient morphological variability and image volume to train a robust feature extractor through SSL. These learned representations were subsequently fine-tuned using our limited labeled RBC phase image dataset.

Two key learning paradigms were employed in our pretext task: momentum contrastive learning,[32] to train the encoder model to extract deep features from the cell images; and reconstruction learning, to train the decoder model to reconstruct the image. **Figure 3** provides an overview of the proposed SSL model. Prior to inputting the white blood cell image into the model, the image is randomly cropped into two regions, c and c'. These cropped images are then passed through the encoder module of the model to extract the corresponding feature vectors, z and z'. To update the encoder weights, contrastive loss is used to compare the feature vectors.[33] This contrastive learning framework employs the Momentum Contrast approach,[34] where the features of different images that are stored in the momentum queue are maintained, while only features of the same images are updated in each training batch. The contrastive loss is expressed as follows:

$$L_{contrastive} = -\log \frac{\exp\left(\frac{sim\left(z_i, z'_i\right)}{\tau}\right)}{\exp\left(\frac{sim\left(z_i, z'_i\right)}{\tau}\right) + \sum_{j=1}^{k} \exp\left(\frac{sim\left(z_i, z'_j\right)}{\tau}\right)} \tag{10}$$

where, τ is a temperature hyperparameter,[35] while the *sim* function represents the dot product between two vectors. z_i represents the features extracted from the encoder that processes the randomly cropped image region c, and z'_i represents the feature vector extracted from the momentum encoder that processes another crop c' from the same image, as shown in Figure 3. The momentum encoder shares the same architecture as the main encoder but is updated via a moving average of the main encoder's weights for stability. These two vectors (z_i and z'_i) form a positive pair. In contrast, z'_j represents feature vectors extracted from other images, also processed through the momentum encoder, and stored in the feature queue as negative samples. The number of such negative samples in the queue is denoted as k. The contrastive loss is designed to maximize the similarity between the positive pair (z_i and z'_i), while minimizing the similarity between z_i and each negative sample z'_j. This training strategy encourages the model to learn more discriminative and generalizable representations across a diverse set of image instances.

To train the decoder of the SSL model, the feature vectors z and z' extracted from the encoders are passed to the decoder to reconstruct the original cropped images. The reconstruction per-

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

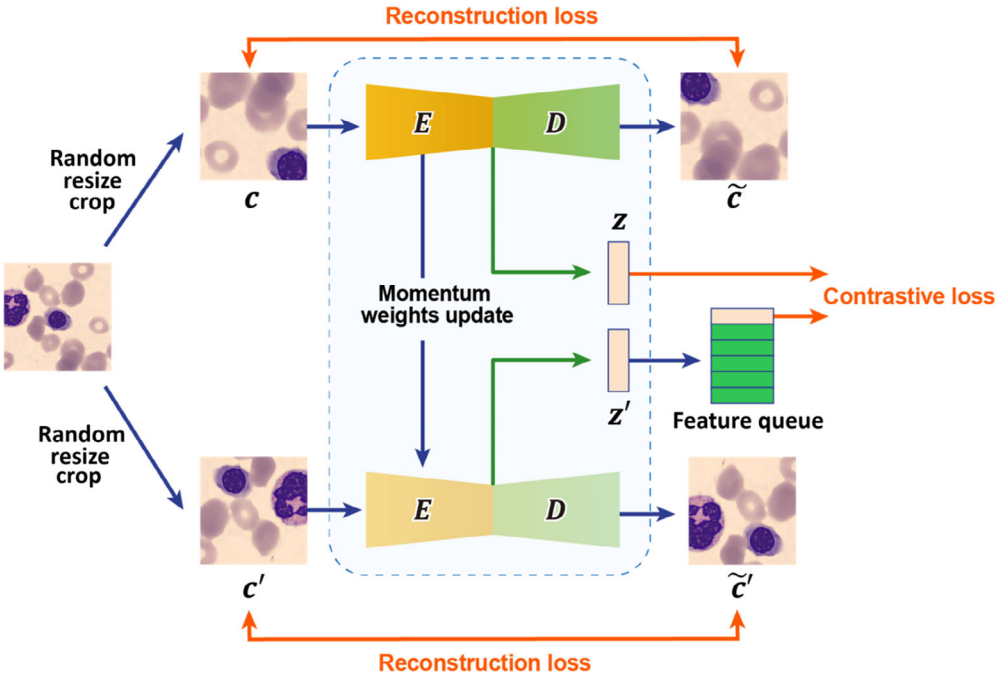

Figure 3. The architecture for self-supervised segmentation pretraining. The encoder is trained using contrastive learning, while the decoder simultaneously uses reconstruction loss. Both encoder and decoder are trained to be augmentation-invariant. These pretrained model weights are initialized for the segmentation task.

formance of the decoder is evaluated using mean squared error as the reconstruction loss:

$$L_{rec} = \frac{1}{n}\sum_{i=1}^{n}(c - \tilde{c})^2 + (c' - \tilde{c}')^2 \tag{11}$$

where, c and c' are the original cropped images, and $\tilde{c}$ and $\tilde{c}'$ are the images reconstructed by the decoder of the SSL model. By integrating both contrastive loss and reconstruction loss, the overall loss function for our SSL network is defined as:

$$L_{ssl} = L_{contrastive} + L_{rec} \tag{12}$$

For the downstream task, pretrained weights of models are transferred to segment the RBC images.

4. ResUnet Training and Inference

Our proposed architecture replaces the standard ResNet50[36] encoder within the U−Net,[37] and establishes skip connections to the decoder module. Compared to the original U−Net, this modified model, ResUnet, enhances feature extraction, while also providing greater flexibility when transferring the model to downstream tasks. **Figure** 4 shows the process of training and inference from small data using ResUnet. The ResUnet model was first pretrained using self-supervised learning with a white blood cell dataset, following the strategy described in Section 3. This approach enabled the model to learn meaningful representations from unlabeled data prior to downstream fine-tuning. The pretrained ResUnet model was then fine-tuned using the original RBC phase images and manually annotated segmentation masks, combined with newly generated RBC phase images and their corresponding segmentation masks synthesized through the diffusion-based data augmentation method detailed in Section 2.3. This augmentation process involved sequentially applying two denoising diffusion GANs: one to generate realistic RBC phase images from random noise, and another to generate matching segmentation masks conditioned on those images. To train the ResUnet model, $L_{CE-Dice}$ was computed, which combines cross entropy loss and dice loss. After training the ResUnet model to perform the RBC segmentation task, the inference process was conducted using an unseen dataset. During inference, only the trained ResUnet was utilized, without involving processes such as the diffusion model or self-supervised learning. As a result, the RBC segmentation task could be quickly and efficiently performed. The generated masks were merged to fit the entire size to obtain the final output.

5. Results and Discussion

5.1. Generated Phase Image by Diffusion Model

RBC phase images were generated using the diffusion model with various random noises, and then segmentation masks corresponding to the generated RBC phase images were generated. **Figure** 5 shows the samples generated by training the diffusion model on the data that was actually obtained. The following observations may be made: First, the generated RBC samples look realistic; the patterns and textures of the generated RBC morphology are similar to those of the real samples. Second, the generated RBC morphology is very diverse; this diversity plays a critical role in mitigating the imbalance observed in the real dataset,

 © 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

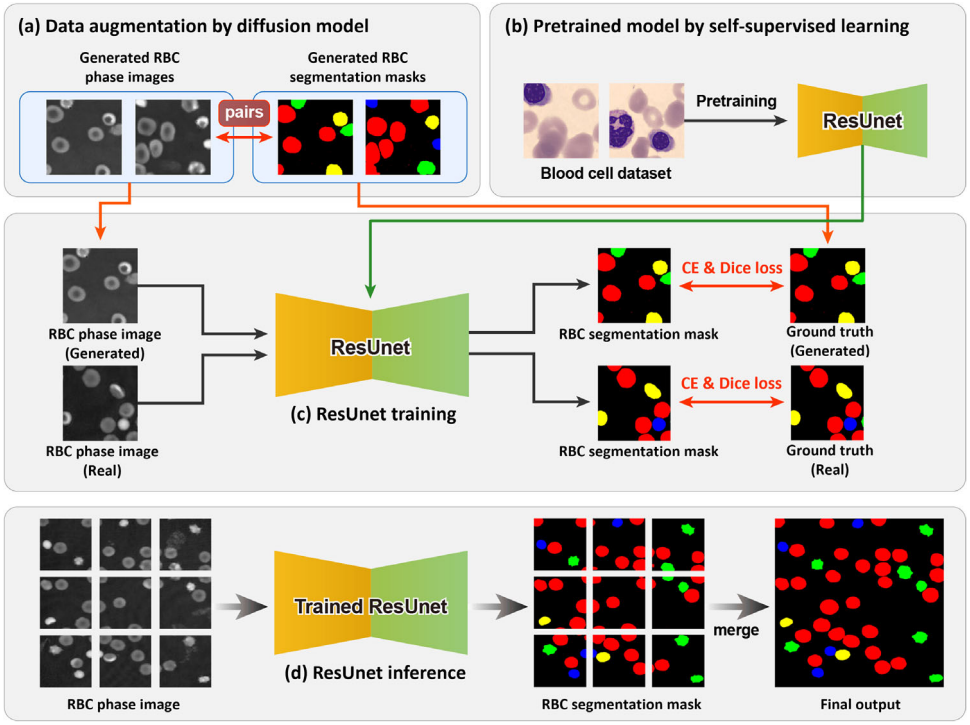

Figure 4. The process of training and inference from small data using ResUnet. ResUnet is trained to generate RBC segmentation masks from RBC phase images using data augmented with diffusion models and a pre-trained model using self-supervised learning. During training, 256 × 256 pixel image patches (corresponding to 36.02 μm × 36.02 μm) are used as input. During inference, the full-size RBC phase image (900 × 900 pixels, corresponding to 129.68 μm × 129.68 μm) is divided into non-overlapping 256 × 256 patches, and segmentation masks are predicted for each patch. The individual patch-level masks are then merged to reconstruct the final output mask for the entire image.

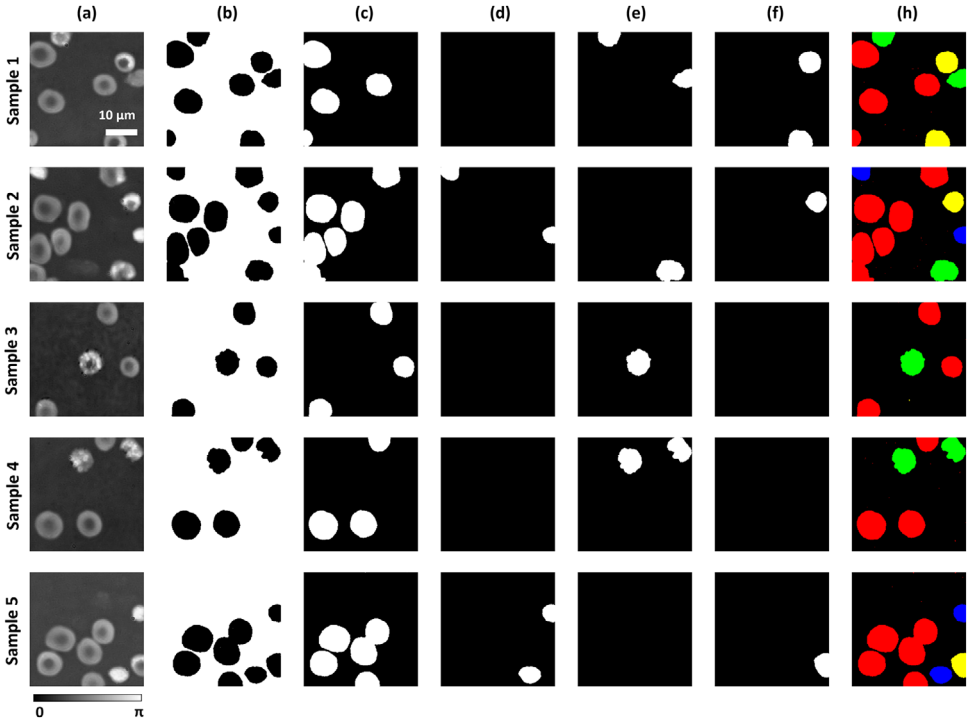

Figure 5. RBC phase image and mask image generated from trained diffusion model. Each image is 256 × 256 pixels, corresponding to a physical size of ≈36.02 μm × 36.02 μm. a) Generated phase image. b) Generated background mask. c) Generated discocyte mask. d) Generated spherocyte mask. (e) Generated echinocyte mask. f) Generated stomatocyte mask. h) Mask image colored by RBC type (red, discocyte; blue, spherocyte; green, echinocyte; yellow, stomatocyte).

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

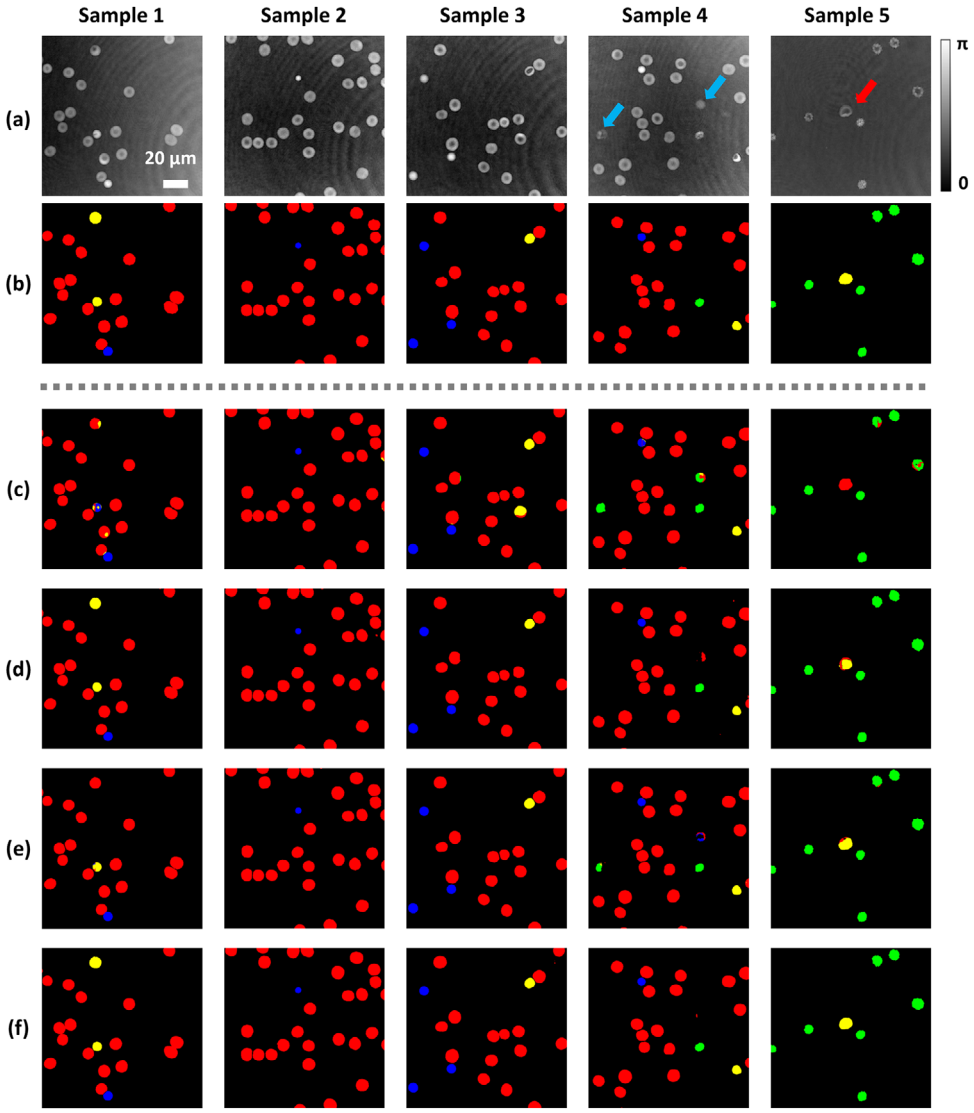

Figure 6. Representative segmentation results from the proposed model under different training settings. a) Input RBC phase image. b) Manually annotated ground truth. c) Result of ResUnet trained without pretraining or augmentation. d) Result after applying data augmentation using a diffusion model. e) Result after applying self-supervised pretraining. f) Result after applying both self-supervised pretraining and data augmentation. All models were evaluated on a test set of 40 images ($n = 40$).

where discocytes are disproportionately represented compared to spherocytes and stomatocytes. To specifically address this issue, we adopted a targeted sampling strategy during the data augmentation process.

From 5000 generated phase-mask image pairs, we manually selected 500 samples that included a balanced and diverse representation of all four RBC types, with particular attention given to increasing the presence of underrepresented cell types such as spherocytes and stomatocytes. These selected samples were then augmented through rotation-based transformations to produce a final dataset of 2000 images. This process not only increased the overall quantity of training data, but also enhanced the morphological diversity required for robust learning across all RBC categories.

Further, the generated samples retained structural consistency with real-world RBCs while introducing sufficient variation, improving the model's ability to generalize. The segmentation masks corresponding to these synthetic phase images were also generated with high accuracy, ensuring reliable ground truth labels for training. This class-balanced data augmentation approach ultimately improved model performance, as demonstrated in later evaluation results.

5.2. Segmentation Results of the Proposed Model

Figure 6 illustrates the impact of incorporating data augmentation and self-supervised learning on the performance of ResUnet.

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1. Quantitative performance analysis of the dice score in multi-class RBC segmentation. Each model was evaluated on 40 test images (n = 40).

Method	Dice Score				
	Discocyte	Spherocyte	Echinocyte	Stomatocyte	Overall
U−Net	0.7548	0.9654	0.9297	0.9385	0.8971
Deeplab v3+	0.651	0.8914	0.7661	0.8707	0.7948
FCN	0.5348	0.9543	0.9332	0.6425	0.7662
PSPNet	0.9251	0.9802	0.9402	0.9748	0.955
GAN + Watershed	0.9548	0.9575	0.9466	0.9064	0.9413
Denoising diffusion GAN	0.9531	0.9801	0.9626	0.9776	0.9683
ResUnet *(backbone)*	0.8762	0.9646	0.9467	0.9295	0.9292
ResUnet + DA	0.891	0.9653	0.9056	0.9813	0.9358
ResUnet + SSL	0.8794	0.9603	0.9536	0.9589	0.9381
ResUnet + DA&SSL	0.948	0.9879	0.981	0.9839	0.9752

Table 2. Quantitative performance analysis of the Aggregated Jaccard Index in multi-class RBC segmentation. Each model was evaluated on 40 test images (n = 40).

Method	Aggregated Jaccard Index				
	Discocyte	Spherocyte	Echinocyte	Stomatocyte	Overall
U−Net	0.6433	0.9399	0.8784	0.8968	0.8396
Deeplab v3+	0.5181	0.844	0.6019	0.8127	0.6941
FCN	0.4499	0.9238	0.8831	0.6344	0.7228
PSPNet	0.8419	0.9539	0.8724	0.9404	0.9021
GAN + Watershed	0.8932	0.9169	0.8915	0.8288	0.8826
Denoising diffusion GAN	0.9049	0.9628	0.9305	0.9577	0.9389
ResUnet *(backbone)*	0.7872	0.9394	0.9044	0.8931	0.881
ResUnet + DA	0.8517	0.9476	0.8772	0.9648	0.9103
ResUnet + SSL	0.7932	0.9321	0.9154	0.9255	0.8968
ResUnet + DA&SSL	0.9214	0.977	0.9636	0.9699	0.9579

When ResUnet was trained using only the acquired data, there were several instances where a single RBC was split into multiple types or misclassified, as can be seen in the red section of sample 5 in Figure 6. Additionally, non-RBC areas, such as the blue section of sample 4 in Figure 6, were incorrectly identified as the RBC. When data augmentation based on the diffusion model was employed, the results improved, with segmentation and classification being more accurate, compared to using only the acquired data. However, issues, such as incorrect classification and segmentation into multiple types, persisted. Similar results were observed when using a pre-trained ResUnet model with self-supervised learning. But when data augmentation and self-supervised learning were simultaneously applied, these problems were nearly eliminated. Non-RBC areas were accurately identified as background, and the RBC classification results closely matched the labels. This demonstrates that applying the two processes together can dramatically improve model performance to achieve more accurate segmentation and classification.

Furthermore, although the generated dataset was constructed using a sampling strategy that intentionally minimized overlapping cells through controlled spatial dispersion, there were still cases in which multiple RBCs appeared in close proximity. Even in such challenging scenarios, the model was able to correctly segment and classify adjacent RBCs without confusion, indicating its robustness to partial contact or touching boundaries. This suggests that the proposed method is not only effective under ideal, isolated conditions, but also generalizes well to realistic imaging situations where cells may be closely packed but not fully overlapping.

5.3. Comparison Results with Evaluation Metrics

The results were compared across various model architectures, including U−Net,[37] Deeplab v3+,[38] FCN,[39] PSPNet,[40] GAN + Watershed[25] and Denoising diffusion GAN.[41] **Tables 1** and **2** present the segmentation evaluation metrics for each model. Metrics were calculated per type, with the "Overall" metric representing the average. ResUnet serves as the backbone model. ResUnet + DA uses a dataset created through a diffusion-based data augmentation process. ResUnet + SSL involves fine-tuning the ResUnet model trained with self-supervised learning for RBC segmentation. ResUnet + SSL & DA combines both processes. The denoising diffusion GAN was the model used for data generation.

In the case of our proposed backbone, ResUnet, when trained without DA or SSL, demonstrated respectable performance, though with slightly lower results for discocytes and stomatocytes. This was interpreted as an increase in error, because although these two sample types are visually similar, they contain significantly more discocytes and significantly fewer stomatocytes. Notably, the relatively lower performance for discocytes is a common issue observed across all models. However, when DA or SSL is applied to ResUnet, the performance for both discocytes and stomatocytes improves. With the application of DA, the improvement is more pronounced, as the diffusion model-based DA selectively includes various types in the generated data, facilitating more effective learning on the previously limited stomatocyte data, thus significantly enhancing stomatocyte performance. ResUnet with both DA and SSL applied shows a substantial performance boost, achieving an overall dice score of 0.97, with AJI scores exceeding 0.9 for all types. This demonstrates that combining DA and SSL leads to a more refined model.

Among the models compared, U−Net, Deeplab v3+, and FCN showed dice scores below 0.9 and AJI scores below 0.85, indicating suboptimal performance. GAN + watershed model showed dice scores over 0.9 but AJI scores below 0.9. However, both PSPNet and Denoising diffusion GAN performed exceptionally well, with dice scores and AJI scores exceeding 0.9. In particular, apart from our proposed ResUnet + DA&SSL, Denoising Diffusion GAN was the only model to achieve AJI scores above 0.9 across all types. Therefore, in terms of accuracy, Denoising diffusion GAN can also be considered a highly suitable model.

Table 3 presents an analysis of whether the models can be used in real-time applications by quantifying the time required to test 10 images, the number of model parameters, and the computational complexity. Since DA and SSL do not impact the testing environment, the values for ResUnet with these enhancements remain the same as the base ResUnet model. Simple models such

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

Table 3. Quantitative computation performance analysis in multi-class RBC segmentation. The time column indicates the total time required to process 10 test images (n = 10) during inference. (s: seconds, G: 10^9 MACs).

Method	Time (s)	Parameters	MACs (G)
U−Net	1.4517	54 424 581	18.42
Deeplab v3+	1.7595	39 757 733	9.28
FCN	1.1224	18 643 845	25.51
PSPNet	2.7531	65 700 293	65.72
GAN + Watershed	1.6574	13 043 269	81.00
Denoising diffusion GAN	24.9052	135 891 461	279.03
ResUnet *(backbone)*	1.7472	73 688 197	50.17
ResUnet + DA	1.7472	73 688 197	50.17
ResUnet + SSL	1.7472	73 688 197	50.17
ResUnet + DA&SSL	1.7472	73 688 197	50.17

as U-Net, DeepLab v3+, FCN, and GAN + watershed have relatively fewer parameters and require less time for testing. However, this simplicity often comes at the cost of reduced performance. The most noteworthy result is the Denoising Diffusion GAN model. Despite achieving high performance on segmentation evaluation metrics with real data, its testing time is nearly 15 times longer than that of the other models. This extended time is due to the diffusion model requiring step-by-step sampling from noise, making it challenging to use in real-time applications. PSPNet also showed good performance in the accuracy evaluation, but took more than 1s longer than the other simple models. In contrast, ResUnet demonstrates relatively short testing times, fewer parameters, and reduced computational load. Despite its simple structure, the ResUnet model with DA and SSL offers high performance, and can segment in a short amount of time. Therefore, the proposed model is suitable for real-time applications; and due to its minimal parameters and computational requirements, can be deployed on lower-performance systems.

5.4. RBCs Storage Assessment by the Proposed Model

We used the dataset reported in Ref. [8] to evaluate the model performance in observing RBC morphology changes due to RBC storage lesions. This dataset contains RBC samples stored for different periods of (8, 13, 16, 23, 27, 30, 34, 37, 40, 47, and 57). There are 6 images for each storage period, for a total of 66 images. **Figure 7** shows the predictions of the trained model for the RBC dataset, which consists of RBCs obtained from different storage periods. As the storage period increased, we observed changes in cell types, with discocytes being the dominant type at the beginning, but spherocytes and echinocytes increasing over time. **Figure 8** shows the changes in the ratio of RBC types as the storage period increases. Consistent with the image results, most RBCs at the beginning of the storage period are discoid RBCs, which account for more than 90% of the total cells. After 40 d, we observed a drastic change in the RBC ratio, where RBCs transformed from discocytes to transient echinocytes, and finally into spherocytes. We manually counted cells in each class, to calculate the accuracy of each RBC type. **Figure 9** visualizes the classification confusion matrix for each RBC class in the storage lesion dataset. In the case of discocytes, there were more than 2400 cells, but they were well classified as discocytes. All other types of RBCs also showed an accuracy of more than 95%. However, there were cases where cells that were echinocytes and stomatocytes were classified as discocytes.

This can be used as a reference to evaluate the quality of stored blood. The proposed model results are consistent with previous studies that show that with increasing storage periods, the main shape of RBCs changes. The goal was to develop a learning method that can accurately segment and classify RBCs from digital holographic images, while obtaining results quickly. In addition, the method aims to efficiently measure the distribution of RBCs within the cell images. This method can help monitor the cell distribution, and determine whether the blood is safe for transfusion.

6. Conclusion

In transfusion medicine, the ability to accurately assess the morphological integrity of RBCs during storage is paramount. The degradation of RBCs over time can lead to compromised cell function and decreased transfusion efficacy, posing significant challenges to patient safety. Traditional methods of RBC analysis are often limited by their reliance on extensive manual evaluation, which can be both time-consuming, and susceptible to human error. To overcome these limitations, this study introduces an innovative approach using digital holographic microscopy in conjunction with advanced deep learning techniques.

Our label-free holographic sensing approach integrates diffusion models and self-supervised learning to improve RBC segmentation and morphological analysis, even when only limited datasets are available. The diffusion model significantly aids in data augmentation by generating high-quality generated images from noise, which enriches the training dataset, and improves model robustness. Self-supervised learning enables the model to effectively learn from a limited amount of labeled data, addressing the common issue of data scarcity in medical imaging. The combination of these technologies results in a model that achieves higher accuracy, while also reducing computational time, making it suitable for real-time applications. Our results demonstrate a marked improvement over traditional methods, suggesting that this technology can facilitate more accurate and efficient monitoring of RBC quality during storage. Ultimately, the implementation of this technique could lead to safer and more effective blood transfusions, as it ensures the delivery to patients of functionally intact RBCs. Future research will focus on extending this approach to other complex biological samples and exploring broader applications in medical diagnostics, potentially paving the way for widespread improvements in healthcare delivery.

7. Experimental Section

Data Preparation: To construct the RBC dataset, 2 mL of peripheral blood was drawn from three healthy donors using sterile syringes. All procedures followed established ethical and biosafety guidelines, with prior approval from the national biosafety committee at DGIST and relevant institutional review boards in South Korea (DGIST-180713-BR-012-01). Fol-

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

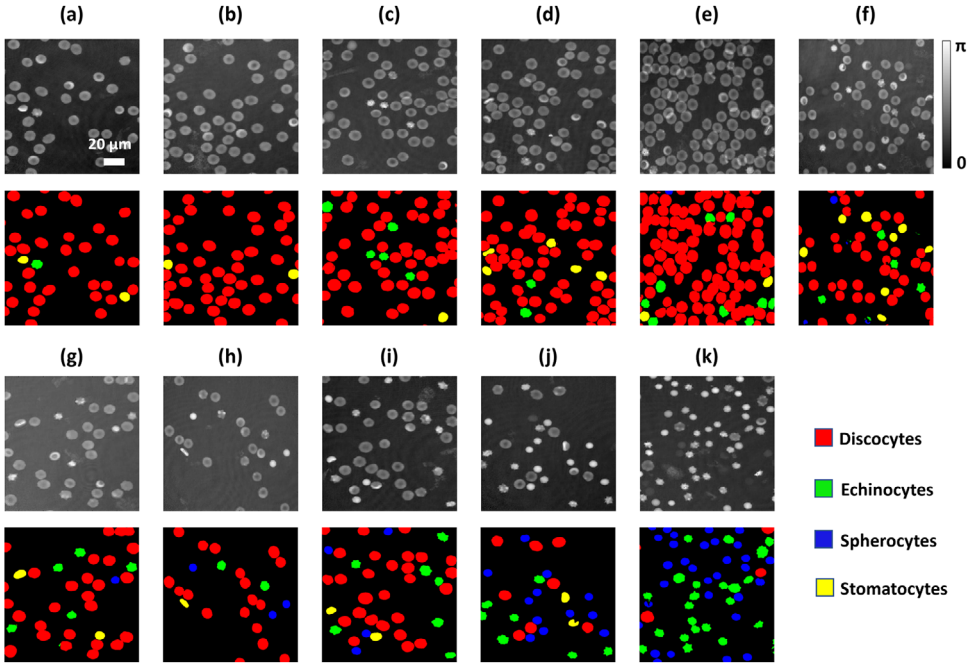

Figure 7. Model classification and segmentation results on the storage lesion dataset. (a–k) The original RBC phase image along with the corresponding segmentation map generated by our model for (8, 13, 16, 23, 27, 30, 34, 37, 40, 47, and 57 days) of storage, respectively. For each storage period, six test images (n = 6) were evaluated, and segmentation masks were generated by the trained model without manual correction.

lowing collection, the samples were centrifuged at 5000×g for 5 min at 4 °C to separate the blood components, and the buffy coat was carefully removed to obtain purified RBCs.

To ensure consistent cell dispersion and minimize overlapping in the field of view, the purified RBCs were resuspended in 1 mL of HEPA buffer, which included HEPES, NaCl, glucose, KCl, $MgCl_2$, and bovine serum albumin to maintain osmotic stability and prevent cell aggregation. From this suspension, two drops of the mixture were placed onto a clean slide glass using a 20 μL pipette, and a cover glass was gently applied. This preparation method was designed to produce a suitable cell density for phase imaging while avoiding significant cell overlap, which could impair phase reconstruction and segmentation accuracy. Although hematocrit levels were not measured directly, the consistency of this protocol ensured well-separated RBCs in the acquired images, providing reliable input for model training and evaluation.

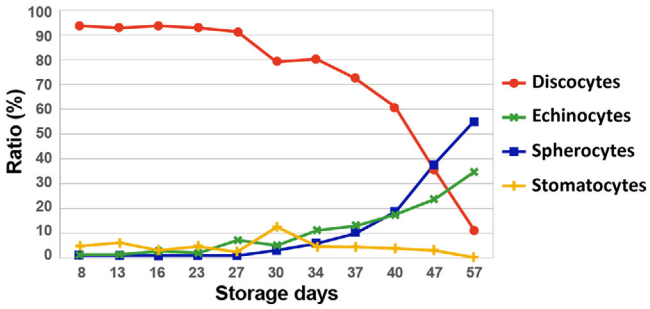

Figure 8. Comparison of trends between the ratio of different RBC types and storage period. As the storage period increases, RBCs change from discocytes to spherocytes. For each storage period, six test images (n = 6) were analyzed. The total number of RBCs per time point was as follows: day 8: 254, day 13: 313, day 16: 363, day 23: 421, day 27: 490, day 30: 262, day 34: 188, day 37: 156, day 40: 309, day 47: 213, and day 57: 290.

Figure 10 presents a representative data sample used in this study. Holographic imaging was performed using an off-axis Mach–Zehnder interferometric setup. The captured holograms were numerically propagated to reconstruct quantitative phase images, based on Fresnel diffraction calculations. Additional algorithmic details are provided in Section SB (Supporting Information). Both holograms and reconstructed images had a resolution of 900 × 900 pixels, corresponding to a field of view of

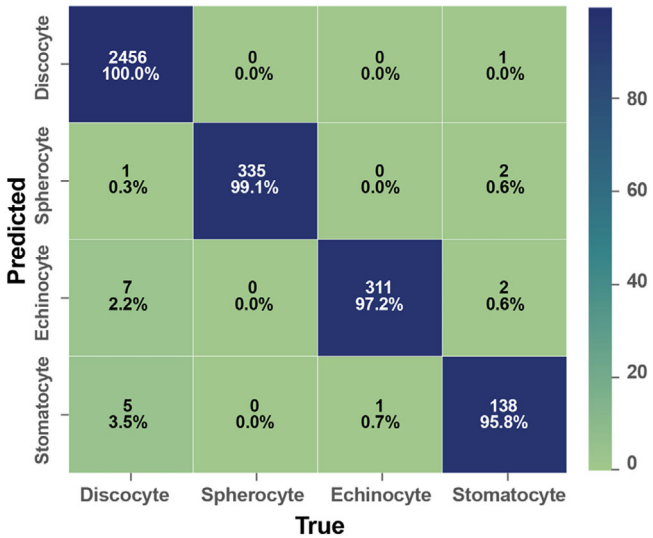

Figure 9. Confusion matrix of the classification accuracy of each type of RBC for the different storage periods dataset. The number of RBCs per class was as follows: discocyte: 2469, spherocyte: 335, echinocyte: 312, and stomatocyte: 143. Accuracy was calculated as the number of correctly predicted RBCs divided by the total number of RBCs in each class.

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

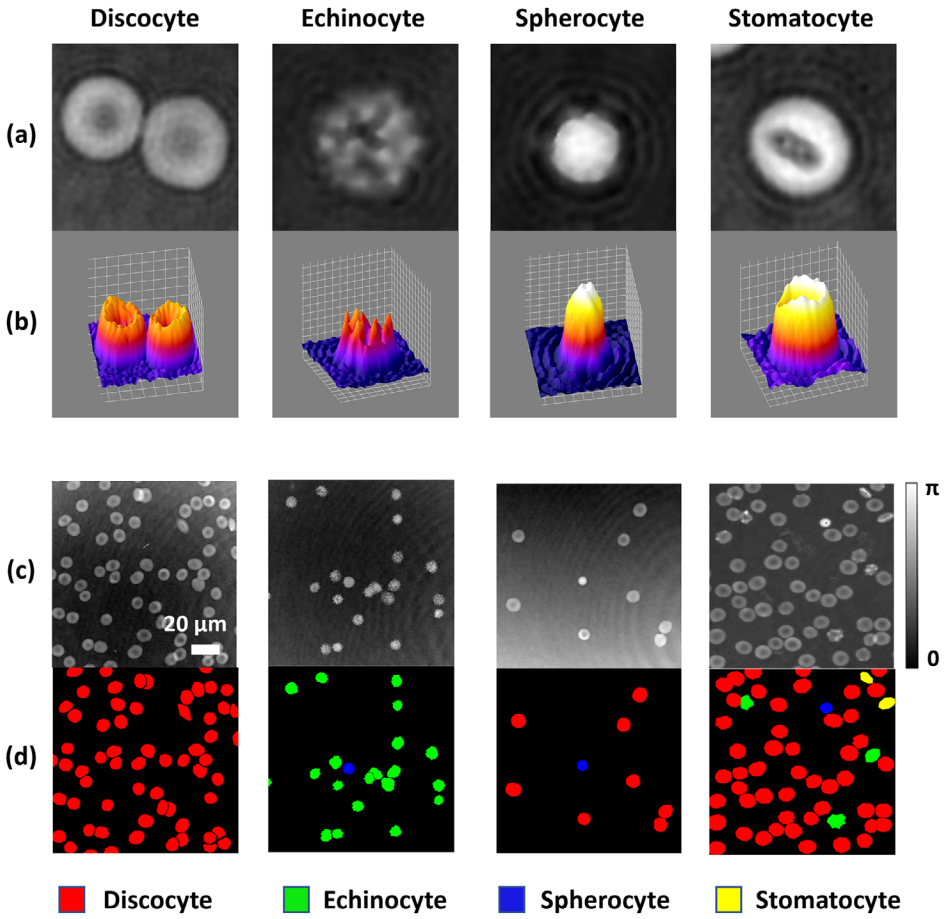

Figure 10. The RBC dataset used for training (Number of datasets: 179 images). a) Four different types of RBCs. b) Their respective 3D visualization. c) Phase images obtained using digital holography. Each full-size image is 900 × 900 pixels, corresponding to a field of view of ≈129.68 μm × 129.68 μm. d) Their corresponding labels.

≈129.68 μm × 129.68 μm. In total, 219 holograms of RBCs were recorded and annotated by an expert biologist. For training 179 and for test data 40 were used. Labels were encoded as five-channel one-hot masks: background (0), discocytes (1), spherocytes (2), echinocytes (3), and stomatocytes (4), with distinct color coding (black, red, blue, green, yellow, respectively). This labeling scheme enabled accurate identification of individual RBC types by deep learning models. To optimize memory usage and augment the dataset, each image was divided into nine non-overlapping patches of 256 × 256 pixels.

Evaluation Metrics: Two evaluation metrics were calculated to evaluate the segmentation results generated by deep learning. First, the Dice score, often referred to as the Dice coefficient, was a crucial metric that was used to evaluate segmentation models. The Dice score was calculated by taking twice the area of overlap between the predicted and ground truth segmentations, divided by the sum of the areas of the predicted and ground truth segmentations:

$$DiceScore = \frac{2 \times |G \cap S|}{|G| + |S|} \tag{13}$$

where, $|G|$ and $|S|$ are pixels of ground truth and the corresponding segmented image, respectively, while $|G \cap S|$ represents the intersection of the two images. These models were also evaluated on the Aggregated Jaccard Index (AJI). AJI was useful to evaluate the segmentation model at the object level, as compared to the pixel level, and is defined as:[42]

$$AJI = \frac{\sum_{i=1}^{N} |G_i \cap S_i|}{\sum_{i=1}^{N} |G_i \cup S_i| + \sum_{i \in R} |S_i|} \tag{14}$$

where, S_i is the predicted segmentation map, G_i is the ground truth, and R is the set of predicted segmentation results that do not match with the ground truth. As these inaccuracies will decrease the AJI score, this is useful for detecting over- and under-segmentation results.

This study also measured the time required to process 10 images to evaluate the model's potential for real-time application, using the same system conditions as the model. The tests were conducted on a Quadro RTX 6000 GPU. The number of parameters and multiply–accumulate operations (MACs) used in the model were also assessed. MACs refer to the fundamental computational steps in neural networks, where each operation involves multiplying two numbers, then adding the result to an accumulator, which serves as a key measure of model computational complexity.

Statistical Analysis: All input RBC phase images were normalized to a range of [−1, 1] prior to model training. To enhance data diversity and alleviate class imbalance, rotation-based data augmentation was applied to both acquired and synthetically generated RBC phase–mask pairs. Segmentation performance was evaluated using two quantitative metrics: Dice coefficient and AJI, both calculated by comparing the predicted

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

segmentation masks with the corresponding manually annotated ground truth. For the comparison of model performance, evaluation was conducted on a test set comprising 40 RBC images, and the reported values represent the average scores across all test images for each RBC subtype (discocyte, spherocyte, echinocyte, and stomatocyte).

To assess storage-dependent morphological changes, a separate RBC dataset was tested that included samples stored over 11 different time periods (8, 13,16, 23, 27, 30, 34, 37, 40, 47, and 57 days). For each storage period, six images were evaluated, and the ratio of each RBC type was calculated by counting the number of segmented RBC per class in the model-generated mask and dividing by the total RBC count. For RBC classification accuracy analysis, class-wise correctness was determined by comparing the model's predicted label with the manually annotated label for each individual RBC instance. The accuracy was then computed as the number of correctly classified RBCs divided by the total number of RBCs per class.

All data processing, model training, and quantitative analyses were performed using Python and PyTorch.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (No. RS-2023-00253613).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

deep learning, diffusion model, holographic image sensor, red blood cells, self-supervised learning, storage lesions

Received: May 23, 2025
Revised: August 5, 2025
Published online: October 2, 2025

[1] J. R. Hess, R. L. Sparrow, P. F. Van Der Meer, J. P. Acker, R. A. Cardigan, D. V. Devine, *Transfusion* **2009**, *49*, 2599.
[2] J. R. Hess, *Transfus. Apher. Sci.* **2010**, *43*, 51.
[3] B. Blasi, A. D'Alessandro, N. Ramundo, L. Zolla, *Transfus. Med.* **2012**, *22*, 90.
[4] C. Roussel, M. Dussiot, M. Marin, A. Morel, P. A. Ndour, J. Duez, C. Le Van Kim, O. Hermine, Y. Colin, P. A. Buffet, P. Amireault, *Transfusion* **2017**, *57*, 1007.
[5] T. Yoshida, M. Prudent, A. D'Alessandro, *Blood Transfus.* **2019**, *17*, 27.
[6] C. Di Ruberto, A. Loddo, L. Putzu, *Comput. Biol. Med.* **2020**, *116*, 103530.
[7] A. Molina, J. Rodellar, L. Boldú, A. Acevedo, S. Alférez, A. Merino, *Comput. Biol. Med.* **2021**, *136*, 104680.
[8] I. Moon, F. Yi, Y. H. Lee, B. Javidi, D. Boss, P. Marquet, *Opt. Express* **2013**, *21*, 30947.
[9] F. Yi, I. Moon, Y. H. Lee, *J. Biomed. Opt.* **2015**, *20*, 016005.
[10] K. Jaferzadeh, I. Moon, M. Bardyn, M. Prudent, J.-D. Tissot, B. Rappaz, B. Javidi, G. Turcatti, P. Marquet, *Biomed. Opt. Express* **2018**, *9*, 4714.
[11] I. Moon, K. Jaferzadeh, *Proc. SPIE* **2019**, 10997.
[12] V. K. Lam, T. C. Nguyen, B. M. Chung, G. Nehmetallah, C. B. Raub, *Cytometry A* **2018**, *93*, 334.
[13] K. Jaferzadeh, S. Son, A. Rehman, S. Park, I. Moon, *Adv. Photon. Res.* **2022**, *4*, 2200043.
[14] E. Ahamadzadeh, K. Jaferzadeh, S. Park, S. Son, I. Moon, *Biosens. Bioelectron.* **2022**, *195*, 113570.
[15] K. Jaferzadeh, B. Rappaz, Y. Kim, B.-K. Kim, I. Moon, P. Marquet, G. Turcatti, *ACS Sens.* **2023**, *8*, 2533.
[16] J. Park, T. Go, S. Lee, *Ann. Biomed. Eng.* **2017**, *45*, 2563.
[17] T. Go, H. Byeon, S. Lee, *Biosens. Bioelectron.* **2018**, *103*, 12.
[18] E. Cuche, P. Marquet, C. Depeursinge, *Appl. Opt.* **1999**, *38*, 6994.
[19] B. Kemper, G. Von Bally, *Appl. Opt.* **2008**, *47*, A52.
[20] Y. Rivenson, Y. C. Wu, A. Ozcan, *Light Sci. Appl.* **2019**, *8*.
[21] G. Barbastathis, A. Ozcan, G. Situ, *Optica* **2019**, *6*, 921.
[22] S. Park, Y. Kim, I. Moon, *Opt. Laser Technol.* **2024**, *176*, 111030.
[23] T. O'Connor, A. Anand, B. Andemariam, B. Javidi, *Biomed. Opt. Express* **2020**, *11*, 4491.
[24] T. O'Connor, J. B. Shen, B. T. Liang, B. Javidi, *Opt. Lett.* **2021**, *46*, 2344.
[25] E. Kim, S. Park, S. Hwang, I. Moon, B. Javidi, *IEEE J. Biomed. Health Inform.* **2022**, *26*, 1318.
[26] Y. Kim, J. Kim, E. Seo, S. J. Lee, *Biosens. Bioelectron.* **2023**, *229*, 115232.
[27] J. Ho, A. Jain, P. Abbeel, *Adv. Neural. Inf. Process Syst.* **2020**, *33*, 6840.
[28] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, B. Poole, *Proc. ICLR* **2020**, *2020*, 1.
[29] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, *Adv. Neural Inf. Process Syst.* **2014**, *27*, 2672.
[30] J. Gui, T. Chen, J. Zhang, Q. Cao, Z. Sun, H. Luo, D. Tao, *IEEE Trans. Pattern Anal. Mach. Intell.* **2024**, *46*, 9052.
[31] A. Acevedo, S. Alférez, A. Merino, L. Puigví, J. Rodellar, *Comput. Methods Programs Biomed.* **2019**, *180*, 105020.
[32] X. Chen, H. Fan, R. Girshick, K. He, arXiv preprint, arXiv:2003.04297 **2020**.
[33] A.v.d. Oord, Y. Li, O. Vinyals, arXiv preprint, arXiv:1807.03748 **2018**.
[34] K. He, H. Fan, Y. Wu, S. Xie, R. Girshick, presented at *2020 Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, June **2020**.
[35] Z. Wu, Y. Xiong, S. X. Yu, D. Lin, presented at 2018 IEEE Conf. on Computer Vision and Pattern Recognition Workshops, CVPR, Salt Lake City, UT, USA, June **2018**.
[36] K. M. He, X. Y. Zhang, S. Q. Ren, J. Sun, presented at 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, June **2016**.
[37] O. Ronneberger, P. Fischer, T. Brox, *Proc. Med. Image. Comput. Comput. Assist. Interv.* **2015**, *9351*, 234.
[38] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, presented at *European Conf. on Computer Vision*, Munich, Germany, September **2018**.
[39] J. Long, E. Shelhamer, T. Darrell, presented at 2015 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW), MA, USA, June **2015**.
[40] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, presented at 2017 *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July **2017**.
[41] Z. Xiao, K. Kreis, A. Vahdat, arXiv preprint, arXiv:2112.07804 **2021**.
[42] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, A. Sethi, *IEEE Trans. Med. Imaging* **2017**, *36*, 1550.

© 2025 The Author(s). Advanced Sensor Research published by Wiley-VCH GmbH

27511219, 2025, 12, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adsr.202500073 by Daegu Gyeongbuk Institute Of, Wiley Online Library on [02/02/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License