



PHE-SICH-CT-IDS: A benchmark CT image dataset for evaluation semantic segmentation, object detection and radiomic feature extraction of perihematomal edema in spontaneous intracerebral hemorrhage

Deguo Ma^a, Chen Li^{a,*}, Tianming Du^a, Lin Qiao^b, Dechao Tang^a, Zhiyu Ma^a, Liyu Shi^a, Guotao Lu^a, Qingtao Meng^a, Zhihao Chen^a, Marcin Grzegorzec^c, Hongzan Sun^b

^a Microscopic Image and Medical Image Analysis Group, College of Medicine and Biological Information Engineering, Northeastern University, China

^b Shengjing Hospital, China Medical University, Shenyang, China

^c Institute of Medical Informatics, University of Luebeck, Luebeck, Germany

ARTICLE INFO

Dataset link: <https://figshare.com/articles/dataset/PHE-SICH-CT-IDS/23957937>

Keywords:

CT image dataset
Spontaneous intracerebral hemorrhage
Perihematomal edema segmentation
Object detection
Feature extraction

ABSTRACT

Background and objective: Intracerebral hemorrhage is one of the diseases with the highest mortality and poorest prognosis worldwide. Spontaneous intracerebral hemorrhage (SICH) typically presents acutely, prompt and expedited radiological examination is crucial for diagnosis, localization, and quantification of the hemorrhage. Early detection and accurate segmentation of perihematomal edema (PHE) play a critical role in guiding appropriate clinical intervention and enhancing patient prognosis. However, the progress and assessment of computer-aided diagnostic methods for PHE segmentation and detection face challenges due to the scarcity of publicly accessible brain CT image datasets.

Methods: This study establishes a publicly available CT dataset named PHE-SICH-CT-IDS for perihematomal edema in spontaneous intracerebral hemorrhage. The dataset comprises 120 brain CT scans and 7,022 CT images, along with corresponding medical information of the patients. To demonstrate its effectiveness, classical algorithms for semantic segmentation, object detection, and radiomic feature extraction are evaluated. The experimental results confirm the suitability of PHE-SICH-CT-IDS for assessing the performance of segmentation, detection and radiomic feature extraction methods.

Results: This study conducts numerous experiments using classical machine learning and deep learning methods, demonstrating the differences in various segmentation and detection methods on the PHE-SICH-CT-IDS. The highest precision achieved in semantic segmentation is 76.31%, while object detection attains a maximum precision of 97.62%. The experimental results on radiomic feature extraction and analysis prove the suitability of PHE-SICH-CT-IDS for evaluating image features and highlight the predictive value of these features for the prognosis of SICH patients.

Conclusion: To the best of our knowledge, this is the first publicly available dataset for PHE in SICH, comprising various data formats suitable for applications across diverse medical scenarios. We believe that PHE-SICH-CT-IDS will allure researchers to explore novel algorithms, providing valuable support for clinicians and patients in the clinical setting. PHE-SICH-CT-IDS is freely published for non-commercial purpose at <https://figshare.com/articles/dataset/PHE-SICH-CT-IDS/23957937>.

1. Introduction

1.1. Research background and motivation

Intracerebral hemorrhage (ICH) is a type of cerebrovascular accident resulting from bleeding within the brain tissue, leading to the accumulation of blood. This disorder is known to be caused by various factors, among which hypertension is the most common, accounting

for about 65% of spontaneous cases [1]. Other causes include amyloid angiopathy, brain tumors, aneurysms, arteriovenous malformations, cerebral cavernous malformations, and arteriovenous fistulae. These underlying causes may lead to the rupture of blood vessels, resulting in the formation of a hematoma within the brain parenchyma [2]. ICH remains a significant cause of morbidity and mortality worldwide, with an estimated incidence of 2.8 million cases annually. ICH accounts for approximately 10% to 15% of all strokes in the USA, Europe,

* Corresponding author.

E-mail addresses: lichen@bmie.neu.edu.cn (C. Li), sunhz@sj-hospital.org (H. Sun).

<https://doi.org/10.1016/j.combiomed.2024.108342>

Received 13 November 2023; Received in revised form 5 March 2024; Accepted 17 March 2024

Available online 20 March 2024

0010-4825/© 2024 Elsevier Ltd. All rights reserved.

and Australia, and up to 20% to 30% of strokes in Asia [3,4]. The incidence rate of ICH is 24.6 per 100,000 person-years, making it a considerable public health concern [5]. The consequences of ICH can be severe, including long-term disability and death. The financial burden of ICH is attributed in part to its high mortality, with up to 40% of patients succumbing to the condition within 30 days of onset, often after prolonged stays in the intensive care unit [6]. Moreover, the incidence of ICH is projected to rise due to the increasing use of anticoagulation, antiplatelet drugs, and an aging population.

Spontaneous intracerebral hemorrhage (SICH) is a frequent subtype of ICH, frequently occurring in the basal ganglia. Patients with SICH may experience early re-bleeding. SICH typically presents acutely, and timely imaging is crucial for accurate diagnosis, determination of the location and volume of bleeding.

The cranial computed tomography (CT) imaging is preferred in the stroke unit due to its prevalence, non-invasiveness, and affordability in nearly all hospitals, while offering good quality visual information on human organs. In CT images, hemorrhage appears as a hyper-intense bright region with sharp contrast against its surroundings and the perihematomal edema (PHE) as a low-density area around the hemorrhage. Numerous preclinical and clinical studies have demonstrated that the kinetics and peak volume of PHE have been shown to cause secondary brain injury (SBI) after SICH and is associated with a poor prognosis. Thus, PHE has been considered a promising therapeutic target for SICH. Furthermore, the CT images of PHE also provide several radiomic feature parameters that can predict hematoma expansion (HE). Patients with SICH may experience HE in the early stage, which can increase the mortality rate of the patient. HE is a critical determinant of disease progression and poor prognosis [7]. Practical scoring schemas have been developed based on these parameters and clinical criteria to predict HE accurately.

However, localization of PHE regions in CT images is extremely challenging due to significant overlap in CT values between HE and other brain tissues, such as cerebrospinal fluid and microvascular diseases. For experienced radiologists, detecting intracerebral hemorrhage using pixel labeling requires an average of 10 min per CT scan, while PHE labeling requires at least twice as much time. The present diagnosis process involves the examine of CT images by an expert radiologist to determine the presence of ICH and identify its type and location [8]. However, the diagnosis is dependent on the availability of subspecialty-trained neuroradiologists, leading to potential time inefficiencies and inaccuracies in remote areas where access to specialized care is limited. The current research status reveals several critical issues in the field: medical professionals face challenges in accurately diagnosing and localizing SICH due to the subjective nature of results, excessive workload, and extended working hours. Furthermore, medical doctors lack radiomic features parameters to predict HE. Therefore, there is a pressing need to address these relevant issues more effectively [9,10].

With the development of medical image processing technology, the primary goal is to achieve high accuracy and good performance in computer-assisted diagnosis. To accomplish this objective, several deep learning-based methods have been proposed and explored [11,12]. Specifically, medical image segmentation and detection are crucial for achieving this goal. The rapid development of object detection methods has enabled the efficient and rapid identification of hemorrhage and PHE locations [13]. In addition, the results of image segmentation methods can provide important reference for clinical doctors in predicting the risk of HE [14,15]. Overall, the medical image processing technology holds great potential for improving the accuracy and efficiency of medical diagnosis for SICH [16].

In this paper, a benchmark CT image dataset for evaluation semantic segmentation, object detection, and radiomic feature extraction of perihematomal edema in spontaneous intracerebral hemorrhage is introduced, namely PHE-SICH-CT-IDS, which is constructed 120 CT scans of patients with SICH. PHE-SICH-CT-IDS contains 3511 CT images of SICH occurring in the basal ganglia region, with associated labels for

the surrounding edematous zone around the hematoma. Additionally, PHE-SICH-CT-IDS provides SICH detection labels for the edematous zone and radiomic features of the edematous zone, which are valuable for clinical research and diagnosis. The evaluation results are obtained by using different classical machine learning and new deep learning methods for segmentation and detection on the images of the CT dataset. Dataset is available at <https://figshare.com/articles/dataset/PHE-SICH-CT-IDS/23957937>.

The main contributions of this paper are as follows:

- (1) The first open-source CT dataset of spontaneous intracerebral hemorrhage with perihematomal edema in the basal ganglia region (PHE-SICH-CT-IDS) is developed and released.
- (2) PHE-SICH-CT-IDS provides segmentation and detection capabilities for edematous areas, as well as radiomic features of the edematous areas around intracerebral hemorrhage, which can be used by medical researchers to study the correlation between PHE and HE within a short period of time.
- (3) The validation of the related edematous area segmentation and detection methods proposed by individuals highlights the distinguishability of PHE-SICH-CT-IDS for classical machine learning and deep learning methods.

1.2. Related work

This study analyzes and compares existing public CT datasets on ICH, and explores in-depth the currently known research results. Additionally, it highlights the limitations of the relevant datasets currently available.

Two commonly used public datasets for ICH detection are the CQ500 dataset [17,18] and the RSNA Intracranial Hemorrhage Detection Challenge dataset (RSNA ICHD dataset). The CQ500 dataset, originating from the Centre for Advanced Research in Imaging, Neurosciences, and Genomics in New Delhi, India, encompasses a diverse range of CT scanners with slices per rotation varying from 16 to 128. CQ500 includes 491 head CT scans with each of intraparenchymal, subdural, extradural, and subarachnoid hemorrhages, as well as calvarial fractures. RSNA ICHD dataset, which was released by the Radiological Society of North America (RSNA) in collaboration with members of the American Society of Neuroradiology and MD.ai for the RSNA Challenge on Kaggle, and contains over 25,000 CT scans.

For ICH segmentation, the PhysioNet ICH dataset contains 82 CT scans. Among these, 36 scans are from patients diagnosed with ICH by radiologists, including types such as intraventricular, intraparenchymal, subarachnoid, epidural, and subdural [19,20]. Each CT scan has around 30 slices with a 5 mm slice thickness. The number of images in this dataset is small and the variety is insufficient. ATLAS v2.0 is a dataset designed for the segmentation of stroke lesion regions in MR T1-weighted (T1 W) unimodal images. The dataset comprises a total of 1271 images, with 955 being publicly accessible. This set is further divided into 655 training images and 300 test images, each accompanied by concealed annotations. Table 3 lists the differences in the parameters of these datasets.

In addition to ICH detection and segmentation, many approaches have been proposed for PHE segmentation [21,22]. However, a lack of public or private datasets with PHE masks has prevented the validation of many of these approaches. Others have been validated on private datasets with different characteristics, such as the region of onset in the brain and the type of ICH diagnosed. Given these differences, it is not possible to objectively compare different approaches. Therefore, a dataset is required to benchmark and extend the work on PHE segmentation in ICH.

The existing datasets in the segmentation domain primarily focus on segmenting the hemorrhage region of ICH, neglecting PHE segmentation. Moreover, ICH occurrences vary across different brain regions, lacking uniqueness in spatial variables. These datasets also lack

Table 1
Patient demographics.

Statistical indicators		Percentage	Amount
Subsequent hematoma expansion	Hematoma expansion	38.33%	46
	non-hematoma expansion	61.66%	74
Gender	Male	60.83%	73
	Female	39.16%	47
Age	Over 50	70.83%	85
	No more than 50	29.16%	35
Hemorrhage area	Left basal ganglia region	50.00%	60
	Right basal ganglia region	49.16%	59
Past medical history	Hypertension	80.00%	96
	Diabetes	10.00%	12

quantitative analysis of ICH causes, failing to account for the variable of spontaneous ICH onset. Consequently, segmentation results from these datasets are limited in assessing risk of edema expansion for subsequent patient management. Although segmentation on MRI datasets tends to yield better results due to their high quality and multimodal nature, MRI imaging is more expensive and slower than CT imaging, and MRI machines are less accessible in underdeveloped regions.

In view of these factors, there is an urgent need for a dataset that has the following criteria: 1. Segmentation of PHE, 2. Selection of fixed brain regions, 3. Exclusive focus on spontaneous ICH, and 4. Increased accessibility of CT data. Therefore, we have developed the PHE-SICH-CT-IDS to fulfill these requirements. Additionally, PHE-SICH-CT-IDS offers object detection functionality to enhance its versatility and usability.

1.3. Structure of this paper

In this section, the background and motivation for the dataset preparation are presented along with a review of related research papers. Detailed methods for preparing and evaluating the dataset, encompassing information about each individual element. Section 3 presents the results of the dataset evaluation, followed by a discussion of these results in Section 4. Finally, a summary of the findings and potential future work are presented.

2. Materials and methods

2.1. Dataset preparation

PHE-SICH-CT-IDS includes 7022 CT images, consisting of 3511 CT slice images and 3511 ground truth images of PHE. Moreover, 3511 object detection labels in XML format are provided. Additionally, PHE-SICH-CT-IDS contains 120 CT scans and corresponding PHE labels in NIFTI format for eligible patients. Radiomic features of the PHE are also available for these 120 patients. The dataset includes medical information for these patients such as presence of the subsequent hematoma expansion, gender, age, diagnostic haemorrhagic area and previous medical history (e.g., hypertension and diabetes). The patient demographic of PHE-SICH-CT-IDS are shown in Table 1. In response to the different functions of the dataset and the diverse requirements of experiments, the dataset is divided into three sub-datasets, with specific information provided in Table 4. The subsequent section provides an introduction to the specifics of the applied datasets.

The complete process, starting from data solicitation and concluding with the compilation of the final dataset, is depicted in Fig. 2.

PHE-SICH-CT-IDS:

- **Data source:**
The head CT scan data of SICH was collected between December 2020 and April 2023 from Shengjing Hospital of China Medical University. The CT scans were rigorously screened by three senior radiologists with 15, 17, and 20 years of experience in cranial

CT interpretation, according to strict criteria(See Data selection criteria for details). The specific scanning parameters of CT are shown in Table 2. Each CT scan consisted of an average of approximately 35 slices.

- **Data selection criteria:**

Inclusion criteria:

- (1) Diagnosis of spontaneous intracranial hemorrhage.
- (2) Hemorrhage occurring in the basal ganglia region of the brain.
- (3) Time from SICH symptom onset to hospital admission and scanning within 12 h.
- (4) Age over 18 years.
- (5) Patients who did not undergo surgical treatment after admission and reexamined their head CT within 72 h.

Exclusion criteria:

- (1) Poor image quality.
- (2) History of neurosurgical procedures.
- (3) Coagulation disorders or a history of anticoagulant use.
- (4) Secondary ICH or bleeding into the ventricles

- **Data preprocessing:**

Filter and exclude data based on criteria, then crop the data to ensure that the brain remains centered in the images. Adjust the window level and width to enhance the brain texture for each data.

- **Perihematomal edema segmentation labels:**

During the data collection process, if the selection criteria are met, the open-source software LIFEx v5.10 is used for PHE segmentation. Using the cerebral window settings (Level=40, Width=120), radiologists manually outline the PHE range on each CT image layer. The outlining is based on the density of the PHE. Notably, this study exclusively focuses on outlining areas with low-density PHE. The specific segmentation labels are shown in Fig. 3.

- **Object detection labels:**

Annotations for object detection are created using the Labelme annotation tool and saved in PASCAL VOC format as XML files with the same name as the corresponding images. The annotated regions include not only the area of intracranial hemorrhage but also the surrounding PHE. The detection labels in the dataset are presented in Fig. 3.

- **Radiomic features:**

Based on the segmented PHE regions, the open-source software LIFEx v5.1 is used for radiomic feature extraction.

- **Dataset preparation workflow:**

- (1) Selecting data that meets the criteria.
- (2) Retaining patient-related medical information.
- (3) Creating segmentation labels. (Fig. 2-(b))
- (4) Extracting radiomic features. (Fig. 2-(c))

Table 2
The scanning parameters of CT are as follows.

Parameter	Specification
Scanner	PHILIPS Brilliance iCT
Tube voltage	120 kV
Tube current	300 mA
Field of view	25 cm
CT resolution	512 × 512
Slice thickness	5.00 mm
Slice interval	5.00 mm

- (5) Converting original data and segmentation label data in NIFIT format to PNG format slices with adjusted window level and width. (Fig. 2-(d))
- (6) Creating object detection labels. (Fig. 2-(f)(h))
- Sub-dataset and image format:
 - (1) Sub-dataset A: NIFIT format. 120 unique CT scans and 120 PHE segmentation masks of NIFTI format, Medical information of patients with 120 corresponding data, Radiomic feature of PHE with 120 corresponding data.
 - (2) Sub-dataset B: PNG format and 512 × 512 pixels. 3511 CT slice images and 3511 ground truth images of PNG format, A total of 3511 detection annotations were provided for all images, Medical information.
 - (3) Sub-dataset C: JPG format and 512 × 512 pixels. 3511 CT slice images and 3511 ground truth images of JPG format, A total of 3511 detection annotations were provided for all images, Medical information.

2.2. Dataset description

2.2.1. Normal CT image

A brain CT image of a healthy individual reveals normal brain structure and tissue characteristics. The cortex and white matter regions exhibit normal density and contrast, with well-defined gray-white matter boundaries. The ventricular system appears in a normal size and shape, and the sulcal patterns and cerebral gyri are intact and clearly visible. No signs of abnormal hemorrhage, masses, edema, or other pathological changes were observed. Overall, the image depicts a healthy brain with normal anatomical structure and tissue integrity. Some examples are shown in Fig. 1.

2.2.2. CT image of SICH

During CT examinations, ICH is characterized by uniformly high-density images in an elongated or circular shape, while PHE manifests as low-density images surrounding the site of hemorrhage. Examples of the related content can be observed in Fig. 1.

2.3. Methods of segmentation

Machine learning plays a central role in various fields such as cancer detection, prognosis and medical diagnostics. Its application in breast tumor and pulmonary lesion localization and diagnosis is notable [25,26]. By using machine learning to analyze gene expression in lung squamous cell carcinoma, eight critical genes involved in the onset and progression of the disease were identified. Machine learning-based classification predictions aid in accurately distinguishing between two distinct cellular states, further enhancing our understanding of these conditions [27]. Exploring “multi-omics” data through machine learning techniques demonstrates their broad applicability in uncovering new insights in marine biology [28]. In addition, in the area of adversarial generation, a GAN-based retinal vessel segmentation algorithm with a large receptive field effectively captures large-scale visual features in retinal images [29].

Five classical machine learning methods are utilized for segmenting PHE-SICH-CT-IDS, including *k*-means, Markov Random Field(MRF), Otsu, Watershed, and Sobel. Additionally, this study explores the segmentation of PHE-SICH-CT-IDS through six distinct deep learning methods. These methods, categorized as classical or novel, encompass U-Net, UNet++, SegNet, SwinUNet, TransUNet and Diffusion-based-Segmentation(Diffusion-Seg).

2.3.1. Classical machine learning methods

k-means clustering algorithm is a well-known partitioning clustering and segmentation method. It clustering for extracting shape features from images characterized by low contrast and backgrounds exhibiting multiple levels of variation brings forth several advantages [30]. MRF is a model described using an undirected graph [31]. In the context of image segmentation, it can be viewed as an image clustering problem. It involves grouping pixels with similar properties into the same class [32]. Otsu’s method, also known as the Otsu’s thresholding algorithm, was proposed in 1979 for determining the threshold value for image binarization. It maximizes between-class variance, effectively separating foreground and background [33]. It is optimal due to its simplicity and wide application. Watershed segmentation is a mathematical morphology-based method rooted in topological theory [34]. It treats images as topographic landscapes, delineating catchment basins and boundaries. Sobel segmentation is a method that utilizes local difference operators to detect edges in images, operating by computing the differences between two adjacent pixels along the diagonal direction [35].

2.3.2. Deep learning methods

U-Net is primarily used for addressing challenges in biomedical image analysis and has gained popularity in the field of medical image segmentation [11]. The left half represents the feature extraction part, while the right half represents the upsampling part, also known as the encoder-decoder structure. Additionally, the network employs skip connections by connecting the upsampling results with the outputs of the encoder submodules that have the same resolution [36].

UNet++ incorporates a series of short connections instead of the long connections that the original U-Net architecture includes [37]. By combining both long and short connections with varying receptive fields, UNet++ has the advantage of capturing features at different levels and integrating them through feature concatenation.

SegNet was introduced to address the challenges of semantic image segmentation in autonomous driving and intelligent robotics [38]. There exists a symmetric relationship between the Encoder and the Decoder. The Encoder undertakes feature extraction via convolutions, while pooling enlarges the receptive field, thereby diminishing the image dimensions. Contrarily, the Decoder encompasses deconvolution and upsampling that reconstruct features post image classification, restoring them to their initial dimensions.

TransUNet utilizes the encoder structure of the transformer in its encoder, allowing for better feature extraction [39]. The transformer is deficient in capturing low-level details, which can result in limited localization capabilities. Due to the limited receptive field, CNNs [40,41] struggle to effectively utilize global information, although they excel in extracting local details.

Swin-UNet is a pure transformer architecture similar to U-Net, specifically designed for medical image segmentation, comprising of an Encoder, Bottleneck, Decoder, and Skip Connections [42,43]. The tokenized image blocks are fed into the Transformer-based U-shaped En-Decoder architecture through skip connections to facilitate local and global semantic feature learning [44].

Diffusion-based-Segmentation (Diffusion-Seg) presents a novel approach rooted in the Denoising Diffusion Probability Model (DDPM) [45,46]. The method is trained on ground truth segmentation, utilizing images as priors at each step during both the training and sampling processes. Stochastic sampling yields a distribution of segmentation masks, allowing for the computation of pixel-wise uncertainty maps. This inherent feature enables an implicit ensemble of segmentations, thereby enhancing overall segmentation performance.

Table 3

Recent datasets for segmentation and detection of intracerebral hemorrhage.

Year	Name	Reference	Category	Amount	Modality
2018	CQ500	[17,18]	Detection, Classification	491 scans	CT
2019	RSNA Intracranial Hemorrhage Detection Challenge dataset	[23]	Detection	25,000 scans	CT
2020	PhysioNet ICH dataset	[19,20]	Segmentation	36 scans	CT
2022	ATLAS V2.0	[24]	Segmentation	955 scans	MRI
2023	Our PHE-SICH-CT-IDS		Segmentation, Detection	120 scans	CT

Table 4

Dataset content of PHE-SICH-CT-IDS.

Subdataset name	Format	Amount	Function	Others
Subdataset A	NIFTI	120 scans	Segmentation Feature extraction	Radiomic features Medical information
Subdataset B	PNG	120 scans/7,022 images	Segmentation Detection	Medical information
Subdataset C	JPG	120 scans/7,022 images	Segmentation Detection	Medical information

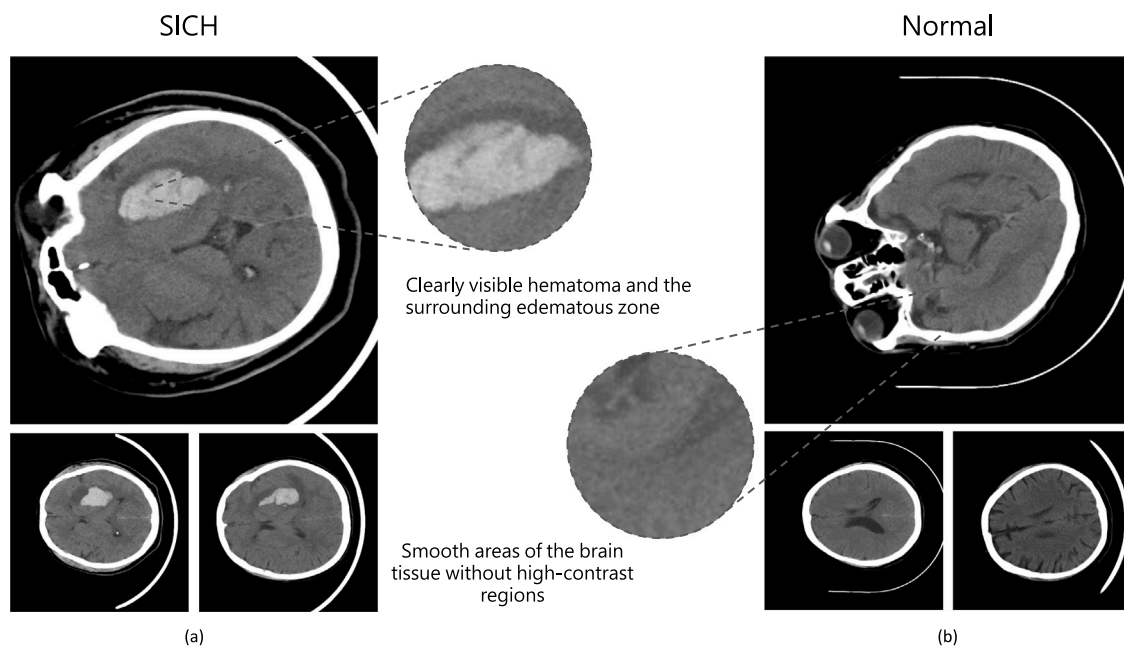


Fig. 1. Example of CT images: (a) CT images of spontaneous intracerebral hemorrhage (SICH), (b) normal CT images without SICH. CT slices of SICH show evident hemorrhagic areas and surrounding edematous zones, while normal CT slices appear smooth and without anomalies.

2.4. Methods of detection

Faster R-CNN is the pioneering framework that achieved end-to-end implementation of object detection tasks using deep learning models [47]. It inherits the technical trajectory of both R-CNN and Fast R-CNN [48] and further introduces the RPN network for efficient and batch-wise generation of region proposals.

SSD (Single Shot MultiBox Detector) enables multi-scale object detection and has achieved superior performance compared to Faster R-CNN on various datasets [49]. It introduces a single-stage detector that is more accurate and faster than previous algorithms such as YOLO, without employing RPN and pooling operations.

The first improvement in YOLOv3 [50] is the replacement of the backbone network with a more effective one, DarkNet53, which extracts relevant features from images to achieve our desired objectives. Compared to DarkNet19 used in YOLOv2 [51], the new network employs a greater number of convolutions, specifically 53 layers of convolutions. YOLOv4 [52] improves the Average Precision (AP) and Frames Per Second (FPS) of YOLOv3 by 10% and 12% respectively.

It incorporates the CSPDarkNet-53 network, which offers excellent trade-offs between speed and accuracy. YOLOv5 incorporates a novel backbone network architecture called CSPNet [53], which outperforms the DarkNet53 in terms of computational efficiency and accuracy. As a result, YOLOv5 exhibits improvements in both speed and precision [54, 55].

2.5. Methods of feature extraction

A total of 1316 texture features were extracted in this study. These features were automatically extracted from the expert-segmented PHE region. They included first-order statistics features [24], shape features [56], gray-level co-occurrence matrix (GLCM) features [57], gray-level run-length matrix (GLRLM) features, gray-level size-zone matrix (GLSZM) features, gray-level dependence matrix (GLDM) features [58] and neighborhood gray-tone difference matrix (NGTDM) features [59]. A consistency test was performed on the extracted radiomics features using the intra-class correlation coefficient (ICC), and only features with ICC > 0.80 were retained for subsequent modeling. The purpose

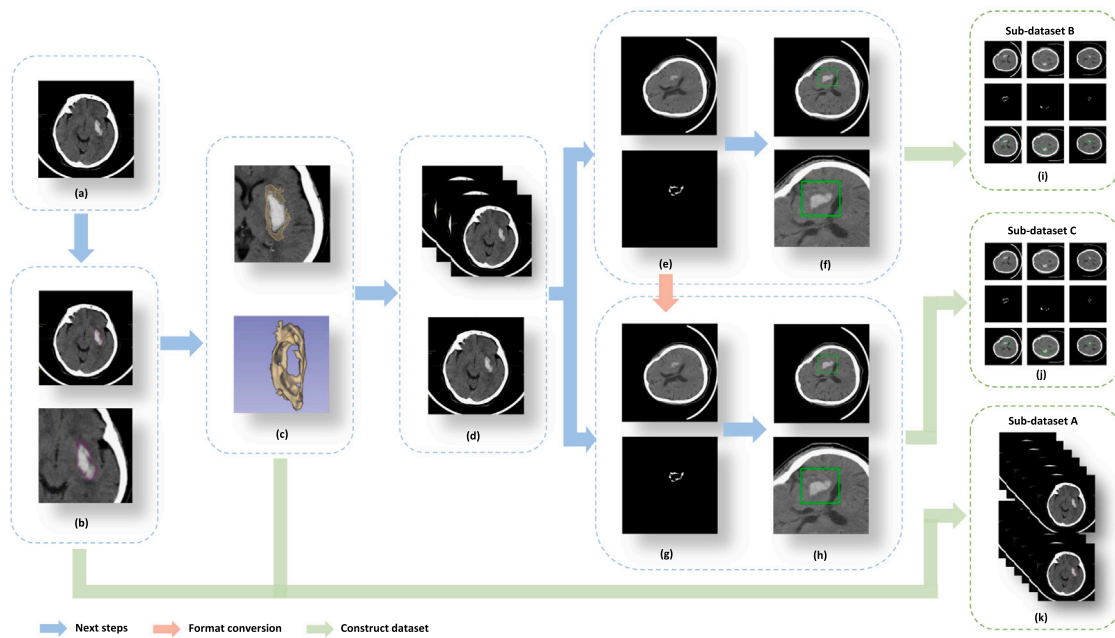


Fig. 2. Data preparation workflow of PHE-SICH-CT-IDS. (a) preprocessed data, (b) making segmentation label, (c) feature extracting, (d) data slicing and renaming, (e)(g) segmentation label, (f)(h) making detection label, (i) Sub-dataset B: 3511 CT slice images and 3511 ground truth images of PNG format, A total of 3511 detection annotations were provided for all images, (j) Sub-dataset C: 3511 CT slice images and 3511 ground truth images of JPG format, A total of 3511 detection annotations were provided for all images, (k) Sub-dataset A: 120 unique CT scans and 120 PHE segmentation masks of NIFTI format.

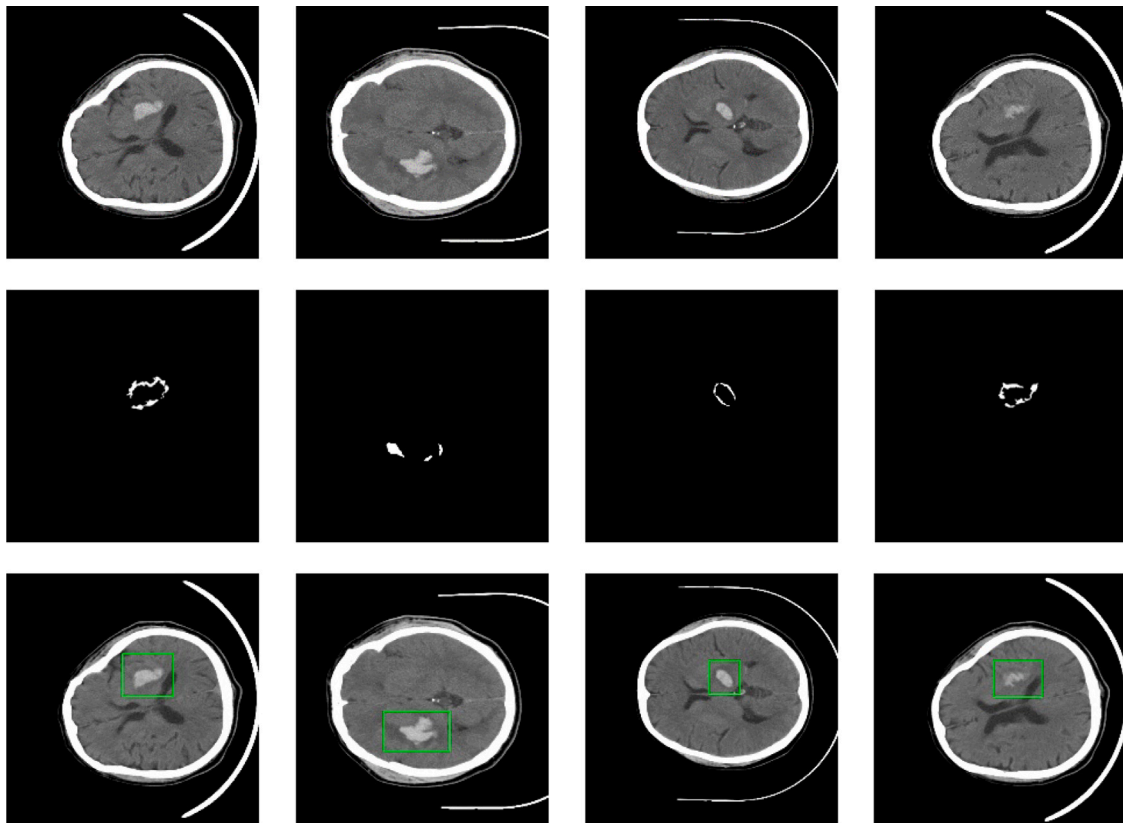


Fig. 3. The labels of SICH-CT-IDS. The first row represents the original CT images, the second row represents the segmentation labels, and the third row represents the detection labels.

of this study is to construct a model based on CT images, using the radiomics features of the PHE around hemorrhage, to predict the risk

of HE in the short term in patients with SICH, and to provide better individualized treatment options for the clinic.

First-order statistics describe the distribution of voxel intensities within an image region defined by a mask, utilizing commonly used and fundamental measures [24]. Shape features include descriptors for the three-dimensional size and shape of the ROI. These features are independent of the grayscale intensity distribution within the ROI and provide information about the shape, size, and spatial relationships of the objects or regions of interest [56]. GLCM is a statistical method used to describe image texture features [57]. It quantifies the spatial relationships between grayscale values of pixels at varying directions and distances within the image [60]. Similarly, GLRLM finds application as a statistical method to elucidate the distribution of consecutive gray levels within an image [60]. Furthermore, GLSZM [61] operates as a statistical tool aimed at capturing the distribution of gray levels within regions of diverse sizes in an image. GLDM [58] takes on the role of a statistical method employed to uncover the interdependence among different gray levels within an image. Lastly, NGTDM [59] emerges as a statistical approach used to illustrate the variations in gray levels among neighboring pixels within an image [62].

2.6. Evaluation of PHE-SICH-CT-IDS

The Dice ratio is a statistical metric employed to assess the similarity between two sets, commonly utilized in the performance evaluation of image segmentation. The Dice coefficient ranges between 0 and 1, where a higher Dice ratio, approaching 1, indicates a greater overlap between the predicted and actual segmentation results, signifying better performance. The Dice ratio is defined in Eq. (1).

$$Dice = \frac{2 \times |A \cap B|}{|A| + |B|} \quad (1)$$

In Eq. (1), for a segmentation task, A and B respectively represent the ground truth and the predicted segmentation mask.

The Jaccard Index, also known as Intersection-over-Union (IoU), is employed to measure the similarity between two sets, particularly useful for assessing the overlap between predicted segmentation results and the actual segmentation results. In object detection, it quantifies the degree of overlap between two bounding boxes, reflecting the proximity of the predicted box to the ground truth box. The Jaccard index is defined in Eq. (2).

$$Jaccard(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

Hausdorff distance considers the maximum matching distance between two sets and serves as a metric for assessing the consistency between image segmentation results and ground truth. A smaller Hausdorff distance between two sets indicates higher similarity between them.

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (3)$$

where,

$$h(A, B) = \max_{a \in A} \left\{ \min_{b \in B} d(a, b) \right\} \quad (4)$$

$$h(B, A) = \max_{b \in B} \left\{ \min_{a \in A} d(a, b) \right\} \quad (5)$$

Here, A and B represent the points of sets A and B respectively, and $d(a, b)$ is the distance from a point a in set A to a point b in set B .

Precision evaluates the probability of accurately classifying elements among all those classified as positives, measuring the extent of misclassification. Recall assesses the probability of correctly classifying positive instances among all elements that should have been classified as positives, indicating the extent of omissions. The calculated results for both metrics fall within the range of [0,1]. Recall and precision are defined in Eqs. (6) and (7).

In a detection task, precision indicates the number of detected targets that are true targets, while recall represents the proportion of true targets successfully identified by the model.

$$Precision = \frac{|TP|}{|TP| + |FP|} \quad (6)$$

$$Recall = \frac{|TP|}{|TP| + |FN|} \quad (7)$$

where TN: True Negative, FP: False Positive, FN: False Negative, TP: True Positive.

Average Precision (AP) is the area under the Precision-Recall curve, providing a comprehensive assessment of the model's performance across various thresholds. AP values range from 0 to 1, with 1 denoting perfect detection performance and 0 representing the poorest performance.

The F1-score, also known as the F1 value, serves as a metric for a comprehensive evaluation of model performance, especially beneficial in scenarios involving imbalanced datasets or where achieving a balance between Precision and Recall is crucial. The formula for calculating the F1-score is expressed in Eq. (8):

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (8)$$

3. Experiment results and analysis

3.1. Results of segmentation

3.1.1. Results of classical machine learning methods

Five classical machine learning segmentation methods are experimentally evaluated on the PHE-SICH-CT-IDS, comparing and analyzing the segmentation images and performance under different machine learning methods. The experimental segmentation results are depicted in Fig. 4. The distinct variations in results obtained by applying different classical machine learning segmentation methods demonstrate the effectiveness of PHE-SICH-CT-IDS in evaluating the performance of various segmentation methods. Additionally, the challenging segmentation of PHE highlights the urgent need for a dedicated edema segmentation dataset.

3.1.2. Results of deep learning methods

The segmentation performance is evaluated on the current dataset using six different deep learning models. In the experiments, each model is trained using a ratio of 4:4:2 for the training, validation, and test sets, respectively. 1404 images are allocated for training, an additional 1404 for validation, and 703 for testing. The learning rate is set to 0.000 05, the number of epochs is set to 100, and the batch size is set to 4. Fig. 5 presents the segmentation results obtained using three different models. The evaluation employs five commonly used metrics, including Dice coefficient, Jaccard index, Hausdorff distance, precision, and recall. The evaluation metrics for the segmentation experiments are shown in Table 5.

The experiments were conducted on an NVIDIA GeForce RTX 2080 GPU with 8 GB of memory. In terms of software, the programming was done in Python 3.8, and the PyTorch framework version 1.7.0 was utilized.

The Swin-UNet achieves the highest Dice of 0.3512, indicating superior similarity in image segmentation tasks. SegNet and UNet++ exhibit lower Dice indices, at 0.1863 and 0.2178, respectively. Similarly, Swin-UNet attains the highest Jaccard index of 0.2285, while SegNet records the lowest Jaccard index at 0.1208. Swin-UNet also demonstrates the lowest Hausdorff distance at 47.87, signifying minimal inconsistency between its segmentation results and ground truth. TransUNet achieves the highest precision at 0.7631, suggesting accurate predictions of positive instances. Swin-UNet achieves the highest recall at 0.3321, indicating its ability to capture more positive instances. SegNet performs relatively lower across multiple metrics, including Dice, Jaccard index, and precision. Its higher Hausdorff distance implies less consistency in segmentation results compared to other methods. The precision of Diffusion-Seg is 0.2566, and the Recall is 0.3125. The relatively lower precision and recall values suggest potential errors in the segmentation of certain positive classes.

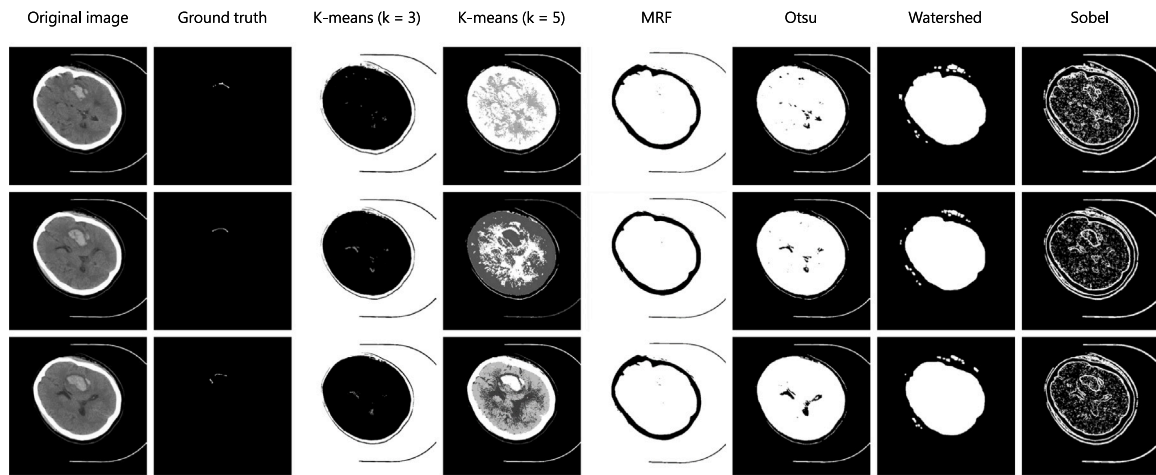


Fig. 4. Segmentation results of classical machine learning methods on the PHE-SICH-CT-IDS. In *K*-means, the number of clusters for partitioning the dataset was tested with values $k=3$ and $k=5$.

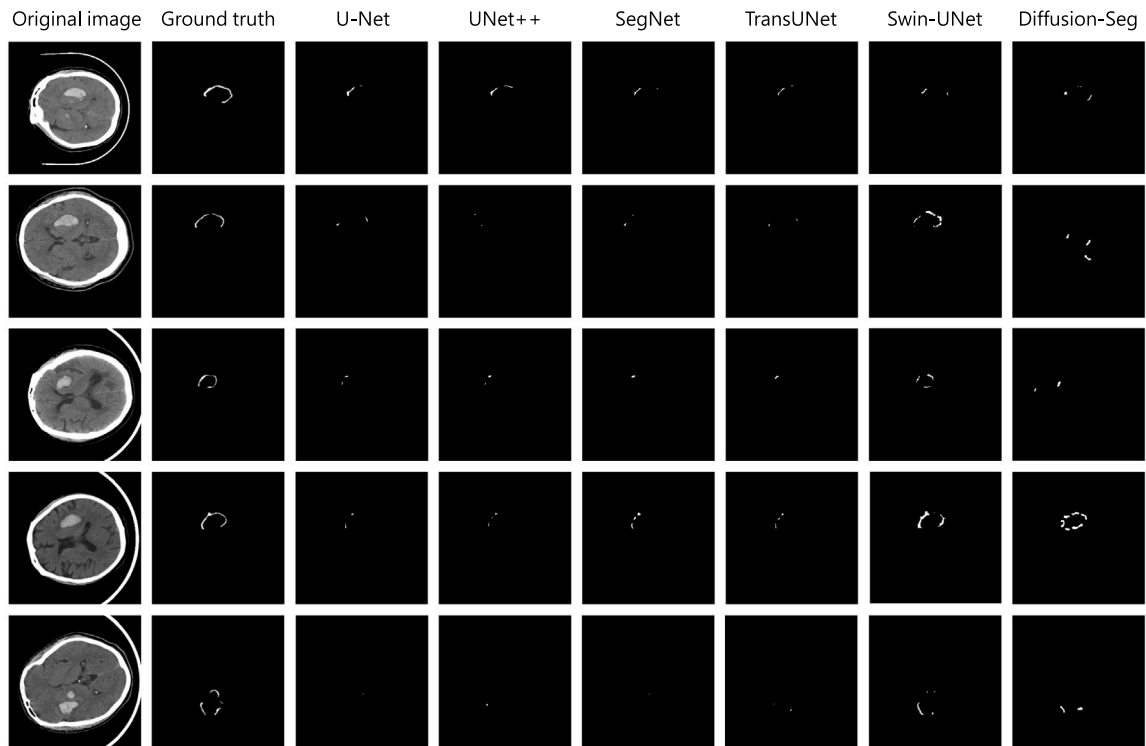


Fig. 5. Segmentation results of deep learning methods on the PHE-SICH-CT-IDS.

In summary, different methods exhibit varying strengths and weaknesses across different metrics. The choice of an appropriate method depends on specific task requirements, such as prioritizing precision or recall, as well as the need for consistency in segmentation results.

3.2. Results of detection

In this study, five commonly used object detection models with different architectures are employed to test the feasibility of using the PHE-SICH-CT-IDS for object detection. The targets of detection include not only the hemorrhagic regions but also the edematous zones. The detection results are depicted in Fig. 6. The performance is evaluated using five widely adopted metrics: AP, F1-score, Precision, Recall and

Table 5

Evaluation metrics for different segmentation methods based on deep learning.

Model	Dice	Jaccard index	Hausdorff distance	Precision	Recall
U-Net	0.2269	0.1488	68.25	0.6332	0.1582
UNet++	0.2178	0.1456	54.05	0.6086	0.1519
SegNet	0.1863	0.1208	48.03	0.5887	0.1224
TransUNet	0.2451	0.1552	52.40	0.7631	0.1591
Swin-UNet	0.3512	0.2285	47.87	0.4720	0.3321
Diffusion-Seg	0.2232	0.1569	45.43	0.2566	0.3125

IoU, as shown in Table 6. For this experiment, the models are trained for 300 epochs with a learning rate of 0.0001 and a batch size of 4.

In the evaluation of object detection methods on the PHE-SICH-CT-IDS, Faster R-CNN and YOLOv5 demonstrate outstanding performance,

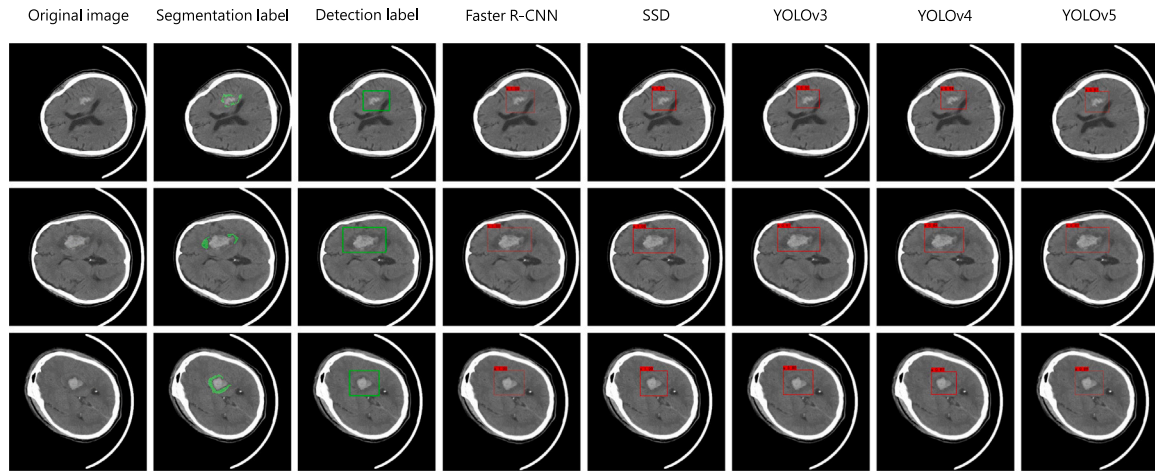


Fig. 6. The object detection experiment results on the PHE-SICH-CT-IDS are presented. The first column shows the original CT images, the second column displays the overlay of the segmented edematous regions on the original images, and the third column shows the ground truth bounding boxes of both edematous regions and hemorrhagic lesions. The fourth to eighth columns illustrate the detection results of Faster R-CNN, SSD, YOLOv3, YOLOv4, and YOLOv5, respectively. The numerical values in the top-left corner of each bounding box represent the confidence scores predicted by the models for the detected objects. The first row displays the confidence scores as follows: 1.00, 1.00, 0.55, 0.94, 0.87. The second row shows: 1.00, 1.00, 0.91, 0.84, 0.87. The third row exhibits: 1.00, 1.00, 0.80, 0.97, 0.89.

Table 6

The evaluation results of the object detection experiments on the PHE-SICH-CT-IDS.

Model	AP	F1-score	Precision	Recall	IoU
Faster R-CNN	0.9656	0.9398	0.9155	0.9655	0.8148
SSD	0.9544	0.9378	0.9189	0.9577	0.8169
YOLOv3	0.9273	0.7609	0.9762	0.6235	0.7793
YOLOv4	0.9433	0.9130	0.9403	0.8873	0.8228
YOLOv5	0.9637	0.9571	0.9710	0.9437	0.8559

achieving the highest AP scores of 0.9656 and 0.9637, respectively. This indicates their excellence in addressing the challenges posed by object detection tasks. YOLOv5 attains the highest F1 score at 0.9571, showcasing a balanced trade-off between precision and recall. YOLOv3 excels in precision at 0.9762, while Faster R-CNN achieves the highest recall at 0.9655. YOLOv5 effectively strikes a balance between these two metrics. In terms of IoU, YOLOv5 outperforms other methods with a superior IoU score of 0.8559, while YOLOv3 lags behind at 0.7793. Faster R-CNN and SSD exhibit comparable IoU scores, albeit lower than YOLOv5. The lower F1 score of YOLOv3 at 0.7609 suggests challenges in simultaneously optimizing precision and recall.

3.3. Results of feature extraction

The sub-dataset A was randomly partitioned into a training set and a validation set in a 7:3 ratio. The features were normalized using the Z-score algorithm, dimensionality reduction was performed using pearson correlation coefficient (PCC) algorithm. Eight radiomic features were selected using analysis of variance (ANOVA), recursive feature elimination (RFE), and relief algorithms, with their weighted coefficients shown in Table 7. The machine learning model employed logistic regression and incorporated clinical parameters for modeling analysis. The training set achieved an area under curve (AUC) value of 0.835. The validation set yielded an AUC value of 0.745.

4. Discussion

4.1. Segmentation results discussion

According to Fig. 4, the PHE-SICH-CT-IDS exhibits noticeable variations in segmentation results when different classic machine learning segmentation methods are employed. However, it is challenging to

Table 7

The refined features and their corresponding weights after the selection process.

Feature	Weight
CT_gradient_firstorder_Skewness	-0.745
CT_square_firstorder_RobustMeanAbsoluteDeviation	1.366
CT_squareroot_gldm_ClusterShade	-0.601
CT_squareroot_gldm_LargeDependenceLowGrayLevelEmphasis	0.651
CT_squareroot_glszm_LowGrayLevelZoneEmphasis	0.092
CT_wavelet-HLL_glszm_SmallAreaHighGrayLevelEmphasis	0.229
CT_wavelet-HLL_gldm_Correlation	-0.707
CT_wavelet-LLL_firstorder_Maximum	-1.069

accurately segment PHE in CT images due to several reasons. Firstly, the edematous region associated with cerebral hemorrhage exhibits similarity in morphology and density with surrounding tissues, making it difficult to achieve clear segmentation from the surrounding structures. This similarity can result in errors in machine learning models, particularly in boundary regions or in edematous areas with complex shapes. Secondly, the edematous region displays significant variability across different CT images, including differences in size, shape, and location. This variability poses a challenge in building a universal model capable of accommodating different patients, scanners, and scanning parameters, thereby limiting the model's generalizability. As a result, classical machine learning methods face significant difficulties in accurately segmenting the edematous region. However, distinct segmentation results were achieved for CT images. There are noticeable differences among different classic machine learning segmentation methods. Hence, the PHE-SICH-CT-IDS proves to be effective in assessing the segmentation performance of different classic machine learning segmentation methods.

Overall, deep learning methods outperform classical machine learning methods. Comparing the results in Figs. 4 and 5, it is evident that deep learning methods accurately locate and segment the objects.

In PHE-SICH-CT-IDS, Swin-UNet demonstrates outstanding performance in Dice, Jaccard index, and Recall, although it exhibits a relatively lower precision. The lowest Hausdorff distance indicates that Swin-UNet possesses robustness in image segmentation tasks. It achieves a Dice of 0.3512, suggesting its superior ability to preserve target shapes in image segmentation tasks. TransUNet excels in precision but shows relatively average performance in other metrics, indicating accurate predictions of positive instances but lesser similarity to actual segmentation results. It can be observed that the two models combining

Transformer [63] with the U-Net architecture perform the best. However, these models have larger sizes and longer training times. Among them, SegNet has the shortest training time and the smallest model size, but it also has the poorest performance among the models. U-Net and UNet++ exhibit comparable performance. However, they demonstrate significant dissimilarity in terms of Hausdorff distance. In comparison to other models, Diffusion-Seg exhibits some disparities in performance, particularly in Precision, Recall, and Hausdorff distance. This could imply challenges in capturing specific details or edge information under certain circumstances, resulting in suboptimal segmentation outcomes.

We have chosen diverse methods considering their strengths in handling specific image features and shapes. For instance, U-Net is selected for its superior capability in retaining details. Future research could focus on enhancing the performance of existing methods, particularly in terms of robustness when dealing with large-scale datasets. Additionally, considering the introduction of deep reinforcement learning and diffusion model may enhance the model's learning capabilities.

4.2. Detection results discussion

Based on the observation in Fig. 6, the object detection algorithms demonstrate excellent performance on the PHE-SICH-CT-IDS, accurately detecting almost all targets, including hemorrhagic regions and PHE zones. Table 6 presents significant variations in various evaluation metrics among the different algorithms.

In object detection tasks, precision and recall evaluate the model's accuracy and recall rate, respectively. A low recall value indicates that the model misses some actual positive instances, while a high precision value signifies the model's high accuracy in predicting positive samples.

Notably, YOLOv5 exhibits the best overall performance in object detection, achieving the highest precision of around 0.9710. The high precision but lower recall and F1 score of YOLOv3 may suggest a tendency for the model to generate more false negatives, instances where actual targets are not detected. This inclination could stem from the model's conservatism, being more sensitive to thresholds, choosing a high threshold to enhance precision at the expense of recall. The model prioritizes ensuring the accuracy of its detection results, being reluctant to miss any potential targets. Additionally, the uneven distribution of samples between positive and negative classes in the dataset may contribute to this phenomenon, leading to dataset imbalance. With our dataset having a lower proportion of positive instances, the model may more easily achieve high precision, but recall may be compromised. Therefore, caution is advised in scenarios requiring a balance between precision and recall when selecting YOLOv3. Additionally, SSD exhibits the fastest detection speed among the tested methods. Therefore, the PHE-SICH-CT-IDS proves to be effectively utilized for image object detection.

4.3. Feature extraction results discussion

This study demonstrates the value of a combined model constructed using texture features extracted from CT images for prognostic predictions in patients with SICH. By applying this combined model, frontline clinicians can self-assess the risk of hematoma expansion in SICH patients and develop personalized treatment strategies. Based on these findings, it is concluded that PHE-SICH-CT-IDS can be utilized to evaluate image features.

4.4. Discussion of method limitations

Using PHE-SICH-CT-IDS, the experimental results demonstrate the performance of different segmentation methods. While the results obtained show promising performance metrics, it is important to acknowledge the limitations and potential challenges associated with the applied methods.

A notable limitation is the relatively low Dice and Jaccard indices for all models, indicating suboptimal overlap between the predicted and ground truth segmentation masks. This suggests that accurate delineation of PHE in CT images may be difficult for the segmentation methods. In addition, the Hausdorff distance values, especially for U-Net and UNet++, indicate relatively large deviations between the predicted and ground truth segmentation boundaries.

CNN-based models such as U-Net and SegNet have shown competitive performance. However, they may struggle to capture long-range dependencies and spatial relationships in complex medical images. Similarly, although transformer-based models such as TransUNet offer advantages in handling sequential data, their applicability to medical image segmentation may be limited due to a lack of spatial understanding. Finally, diffusion-based methods such as Diffusion-Seg, while promising in their use of probabilistic modeling, may face challenges in terms of computational complexity and interpretability.

Addressing these limitations is critical to comprehensively evaluate and improve the robustness of the methods, and to guide future research efforts to improve the effectiveness.

5. Conclusion and future work

This study developed PHE-SICH-CT-IDS, the first open-source CT dataset tailored for segmenting perihematomal edema (PHE) in patients with spontaneous basal ganglia intracerebral hemorrhage. In addition to performing PHE segmentation, the dataset facilitates hemorrhage detection and extraction of radiomic features. PHE-SICH-CT-IDS comprises three sub-datasets with diverse data types, catering to the specific needs of users. To evaluate the segmentation performance, we employ classical machine learning and deep learning methods, such as U-Net and Swin-UNet. The experimental results demonstrated the ability of PHE-SICH-CT-IDS to distinguish between segmentation methods, supported by multiple metrics. Object detection methods, including Faster R-CNN and YOLOv5, accurately identified hemorrhage and surrounding edema. PHE-SICH-CT-IDS shows promise for effective image-based object detection. Radiomic features were extracted, and a joint model combining logistic regression and clinical parameters significantly improved classification performance, aiding clinicians in assessing HE risk in SICH patients.

In the forthcoming future, the dataset will undergo expansion to encompass a broader scope of medical scenarios. Concurrently, efforts will focus on refining and advancing the segmentation methods to provide a more sophisticated and robust resource for the medical community. We will employ advanced text feature extraction and selection techniques [64], including the latent Dirichlet allocation feature extraction model and the RFINCA feature selector [65,66]. This ongoing pursuit of excellence aims to empower medical professionals with cutting-edge tools and foster groundbreaking advancements in the realm of medical research.

CRediT authorship contribution statement

Deguo Ma: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chen Li:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Tianming Du:** Validation, Data curation, Conceptualization. **Lin Qiao:** Data curation. **Dechao Tang:** Data curation. **Zhiyu Ma:** Data curation. **Liyu Shi:** Data curation. **Guotao Lu:** Data curation. **Qingtao Meng:** Data curation. **Zhihao Chen:** Software, Data curation. **Marcin Grzegorzczek:** Methodology. **Hongzan Sun:** Supervision, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

We state that there is no Conflict of Interest in this paper.

Data availability

Ethics is proved by China Medical University, China: No. 2022PS433K. The datasets presented in this study can be found in online repositories. The names of the repository and accession number can be found below: <https://figshare.com/articles/dataset/PHE-SICH-CT-IDS/23957937>.

Acknowledgments

This work is supported by the “National Natural Science Foundation of China” (No. 82220108007). We thank B.A. Qi Qiu, from Foreign Studies College in Northeastern University, China, for her professional English proofreading in this paper. We also thank Miss. Zixian Li and Mr. Guoxian Li for their important discussion in this work. Chen Li and Hongzan Sun have the same contributions to this paper as corresponding authors.

References

- [1] M.I. Aguilar, W.D. Freeman, Spontaneous intracerebral hemorrhage, *Seminars Neurol.* 30 (5) (2010).
- [2] R.F. Keep, Y. Hua, G. Xi, Intracerebral haemorrhage: mechanisms of injury and therapeutic targets, *Lancet Neurol.* 11 (8) (2012) 720–731.
- [3] E.J. Benjamin, P. Muntner, A. Alonso, Heart disease and stroke statistics-2019 update: A report from the American heart association, *Circulation* 139 (10) (2019) e56–528.
- [4] O. Adeoye, J.P. Broderick, Advances in the management of intracerebral hemorrhage, *Nat. Rev. Neurol.* 6 (11) (2010) 593–601.
- [5] E.C. Jauch, J.A. Pineda, J.C. Hemphill, Emergency neurological life support: Intracerebral hemorrhage, *Neurocritical Care* 23 (2 Supplement) (2015) 83–93.
- [6] C. Asch, M.J. Luitse, G.J. Rinkel, I. Tweel, C.J. Klijn, Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis, *Lancet Neurol.* 9 (2) (2010) 167–176.
- [7] D. Dowlatabadi, A.M. Demchuk, M.L. Flaherty, M. Ali, P.L. Lyden, E.E. Smith, Defining hematoma expansion in intracerebral hemorrhage: Relationship with patient outcomes, *Neurology* 76 (14) (2011) 1238–1244.
- [8] C. Li, H. Chen, X. Li, N. Xu, Z. Hu, D. Xue, S. Qi, H. Ma, L. Zhang, H. Sun, A review for cervical histopathology image analysis using machine vision approaches, *Artif. Intell. Rev.* 53 (2020) 4821–4862.
- [9] M.M. Rahaman, C. Li, Y. Yao, F. Kulwa, X. Wu, X. Li, Q. Wang, DeepCervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques, *Comput. Biol. Med.* 136 (2021) 104649.
- [10] S. Kosov, C. Shirahama, C. Li, M. Grzegorzec, Environmental microorganism classification using conditional random fields and deep convolutional neural networks, *Pattern Recogn.* 77 (2018) 248–261.
- [11] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, 2015, arXiv:1505.04597.
- [12] Y. Wang, T. Bai, T. Li, L. Huang, Osteoporotic vertebral fracture classification in X-rays based on a multi-modal semantic consistency network, *J. Bionic Eng.* 19 (6) (2022) 1816–1829.
- [13] H. Chen, C. Li, G. Wang, X. Li, M.M. Rahaman, H. Sun, W. Hu, Y. Li, W. Liu, C. Sun, et al., GasHis-Transformer: A multi-scale visual transformer approach for gastric histopathological image detection, *Pattern Recognit.* 130 (2022) 108827.
- [14] Q. Gao, M. Almekkawy, ASU-Net++: A nested U-Net with adaptive feature extractions for liver tumor segmentation, *Comput. Biol. Med.* (2021).
- [15] J. Zhang, C. Li, S. Kosov, M. Grzegorzec, K. Shirahama, T. Jiang, C. Sun, Z. Li, H. Li, LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation, *Pattern Recognit.* 115 (2021) 107885.
- [16] M.M. Rahaman, C. Li, Y. Yao, F. Kulwa, M.A. Rahman, Q. Wang, S. Qi, F. Kong, X. Zhu, X. Zhao, Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches, *J. X-ray Sci. Technol.* 28 (5) (2020) 821–839.
- [17] S. Chilamkurthy, R. Ghosh, S. Tanamala, M. Biviji, N.G. Campeau, V.K. Venugopal, V. Mahajan, P. Rao, P. Warier, Development and validation of deep learning algorithms for detection of critical findings in head CT scans, 2018, CoRR, arXiv:1803.05854.
- [18] S. Chilamkurthy, R. Ghosh, S. Tanamala, M. Biviji, Campeau, Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study, *Lancet* 392 (10162) (2018) 2388–2396.
- [19] M.D. Hssayeni, M.S. Croock, A.D. Salman, H.F. Al-khafaji, Z.A. Yahya, B. Ghorani, Intracranial hemorrhage segmentation using a deep convolutional model, *Data* 5 (1) (2020) 14.
- [20] M. Hssayeni, M. Croock, A. Salman, H. Al-khafaji, Z. Yahya, B. Ghorani, Computed tomography images for intracranial hemorrhage detection and segmentation, in: *Intracranial Hemorrhage Segmentation using A Deep Convolutional Model. Data*, Vol. 5, (1) 2020, p. 14.
- [21] Y. Chen, C. Qin, J. Chang, Y. Liu, Q. Zhang, Z. Ye, Z. Li, F. Tian, W. Ma, J. Wei, et al., Defining delayed perihematomal edema expansion in intracerebral hemorrhage: segmentation, time course, risk factors and clinical outcome, *Front. Immunol.* 13 (2022).
- [22] Z. Kuang, Z. Yan, L. Yu, X. Deng, Y. Hua, S. Li, Uncertainty-aware deep learning with cross-task supervision for PHE segmentation on CT images, *IEEE J. Biomed. Health Inf.* 26 (6) (2022) 2615–2626.
- [23] A.E. Flanders, L.M. Prevedello, G. Shih, S.S. Halabi, J. Kalpathy-Cramer, R. Ball, J.T. Mongan, A. Stein, F.C. Kitamura, M.P. Lungren, et al., Construction of a machine learning dataset through collaboration: the RSNA 2019 brain CT hemorrhage challenge, *Radiol. Artif. Intell.* 2 (3) (2020) e190211.
- [24] S.-L. Liew, B.P. Lo, M.R. Donnelly, A. Zavaliangos-Petropulu, J.N. Jeong, G. Barisano, A. Hutton, J.P. Simon, J.M. Juliano, A. Suri, et al., A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms, *Sci. data* 9 (1) (2022) 320.
- [25] R. Ranjbarzadeh, S. Dorosti, S.J. Ghouschi, A. Caputo, E.B. Tirkolaee, S.S. Ali, Z. Arshadi, M. Bendechache, Breast tumor localization and segmentation using machine learning techniques: Overview of datasets, findings, and methods, *Comput. Biol. Med.* 152 (2023) 106443.
- [26] Y. Chen, L. Feng, C. Zheng, T. Zhou, L. Liu, P. Liu, Y. Chen, LDANet: Automatic lung parenchyma segmentation from CT images, *Comput. Biol. Med.* 155 (2023) 106659.
- [27] H.K. Joon, A. Thaler, D. Gupta, Machine learning analysis of lung squamous cell carcinoma gene expression datasets reveals novel prognostic signatures, *Comput. Biol. Med.* 165 (2023) 107430.
- [28] J. Manochkumar, A.K. Cherukuri, R.S. Kumar, A.I. Almansour, S. Ramamoorthy, T. Efferth, A critical review of machine-learning for “multi-omics” marine metabolite datasets, *Comput. Biol. Med.* (2023) 107425.
- [29] H. Zhao, X. Qiu, W. Lu, H. Huang, X. Jin, High-quality retinal vessel segmentation using generative adversarial network with a large receptive field, *Int. J. Imaging Syst. Technol.* 30 (3) (2020) 828–842.
- [30] N. Dhanachandra, K. Manglem, Y.J. Chanu, Image segmentation using K-means clustering algorithm and subtractive clustering algorithm, *Procedia Comput. Sci.* 54 (2015) 764–771.
- [31] H. Deng, D.A. Clausi, Unsupervised image segmentation using a simple MRF model with a new implementation scheme, 2004, pp. 2323–2335.
- [32] X. Li, C. Li, M.M. Rahaman, H. Sun, X. Li, J. Wu, Y. Yao, M. Grzegorzec, A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches, *Artif. Intell. Rev.* 55 (6) (2022) 4809–4878.
- [33] D. Liu, J. Yu, Otsu method and K-means, in: 2009 Ninth International Conference on Hybrid Intelligent Systems, Vol. 1, IEEE, 2009, pp. 344–349.
- [34] R. Shojaii, J. Alirezaie, P. Babyn, Automatic lung segmentation in CT images using watershed transform, in: *IEEE International Conference on Image Processing* 2005, Vol. 2, IEEE, 2005, pp. II–1270.
- [35] J. Zhang, C. Li, M.M. Rahaman, Y. Yao, P. Ma, J. Zhang, X. Zhao, T. Jiang, M. Grzegorzec, A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches, *Artif. Intell. Rev.* (2022) 1–70.
- [36] C. Li, K. Wang, N. Xu, A survey for the applications of content-based microscopic image analysis in microorganism classification domains, *Artif. Intell. Rev.* 51 (2019) 577–646.
- [37] Z. Zhou, M.M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation, in: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings*, Vol. 4, Springer, 2018, pp. 3–11.
- [38] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, 2016, arXiv:1511.00561.
- [39] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A.L. Yuille, Y. Zhou, TransUNet: Transformers make strong encoders for medical image segmentation, 2021, arXiv:2102.04306.
- [40] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE* 86 (11) (1998) 2278–2324.
- [41] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, *Commun. ACM* 60 (6) (2017) 84–90.
- [42] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, M. Wang, Swin-unet: Unet-like pure transformer for medical image segmentation, in: *European Conference on Computer Vision*, Springer, 2022, pp. 205–218.
- [43] W. Liu, C. Li, N. Xu, T. Jiang, M.M. Rahaman, H. Sun, X. Wu, W. Hu, H. Chen, C. Sun, et al., CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN, visual transformer and multilayer perceptron, *Pattern Recognit.* 130 (2022) 108829.

- [44] J. Zhang, C. Li, Y. Yin, J. Zhang, M. Grzegorzec, Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer, *Artif. Intell. Rev.* 56 (2) (2023) 1013–1070.
- [45] J. Wolleb, R. Sandkühler, F. Bieder, P. Valmaggia, P.C. Cattin, Diffusion models for implicit image segmentation ensembles, in: *International Conference on Medical Imaging with Deep Learning*, PMLR, 2022, pp. 1336–1348.
- [46] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, in: *Advances in Neural Information Processing Systems*, Vol. 33, 2020, pp. 6840–6851.
- [47] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, in: *Advances in Neural Information Processing Systems*, Vol. 28, 2015.
- [48] R. Girshick, Fast R-CNN, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440–1448.
- [49] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, Ssd: Single shot multibox detector, in: *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, the Netherlands, October 11–14, 2016, Proceedings, Part I*, Vol. 14, Springer, 2016, pp. 21–37.
- [50] J. Redmon, A. Farhadi, Yolov3: An incremental improvement, 2018, *arXiv preprint arXiv:1804.02767*.
- [51] J. Redmon, A. Farhadi, YOLO9000: better, faster, stronger, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7263–7271.
- [52] A. Bochkovskiy, C.-Y. Wang, H.-Y.M. Liao, Yolov4: Optimal speed and accuracy of object detection, 2020, *arXiv preprint arXiv:2004.10934*.
- [53] C.-Y. Wang, H.-Y.M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, I.-H. Yeh, CSPNet: A new backbone that can enhance learning capability of CNN, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 390–391.
- [54] A. Chen, C. Li, S. Zou, M.M. Rahaman, Y. Yao, H. Chen, H. Yang, P. Zhao, W. Hu, W. Liu, et al., SVIA dataset: A new dataset of microscopic videos and images for computer-aided sperm analysis, *Biocybern. Biomed. Eng.* 42 (1) (2022) 204–214.
- [55] P. Ma, C. Li, M.M. Rahaman, Y. Yao, J. Zhang, S. Zou, X. Zhao, M. Grzegorzec, A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches, *Artif. Intell. Rev.* 56 (2) (2023) 1627–1698.
- [56] Y. Peng, P. Luan, H. Tu, X. Li, P. Zhou, Pulmonary fissure segmentation in CT images based on ODoS filter and shape features, *Multimedia Tools Appl.* (2023) 1–22.
- [57] Q. Firdaus, R. Sigit, T. Harsono, A. Anwar, Lung cancer detection based on CT-scan images with detection features using gray level co-occurrence matrix (GLCM) and support vector machine (SVM) methods, in: *2020 International Electronics Symposium, IES, IEEE*, 2020, pp. 643–648.
- [58] N. Ahmadi, G. Akbarizadeh, Iris tissue recognition based on GLDM feature extraction and hybrid MLPNN-ICA classifier, *Neural Comput. Appl.* 32 (2020) 2267–2281.
- [59] B. Pathak, D. Barooah, Texture analysis based on the gray-level co-occurrence matrix considering possible orientations, *Int. J. Adv. Res. Electr., Electron. Instrum. Eng.* 2 (9) (2013) 4206–4212.
- [60] Ş. Öztürk, B. Akdemir, Application of feature extraction and classification methods for histopathological image using GLCM, LBP, LBGLCM, GLRLM and SFTA, *Procedia Comput. Sci.* 132 (2018) 40–46.
- [61] G. Thibault, J. Angulo, F. Meyer, Advanced statistical matrices for texture characterization: application to cell classification, *IEEE Trans. Biomed. Eng.* 61 (3) (2013) 630–637.
- [62] F. Kulwa, C. Li, J. Zhang, K. Shirahama, S. Kosov, X. Zhao, T. Jiang, M. Grzegorzec, A new pairwise deep learning feature for environmental microorganism image analysis, *Environ. Sci. Pollut. Res.* 29 (34) (2022) 51909–51926.
- [63] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, in: I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc., 2017.
- [64] B. Sun, Y. Xu, S. Xie, D. Xu, Y. Liang, Data processing methods of flow field based on artificial lateral line pressure sensors, *J. Bionic Eng.* 19 (6) (2022) 1797–1815.
- [65] A. Gupta, R. Katarya, PAN-LDA: A latent Dirichlet allocation based novel feature extraction model for COVID-19 data using machine learning, *Comput. Biol. Med.* 138 (2021) 104920.
- [66] F. Ozyurt, T. Tuncer, A. Subasi, An automated COVID-19 detection based on fused dynamic exemplar pyramid feature extraction and hybrid feature selection using deep learning, *Comput. Biol. Med.* 132 (2021) 104356.