Науковий вісник НЛТУ України
Scientific Bulletin of UNFU
https://doi.org/10.36930/40340813
Article received 18.11.2024
Article accepted 23.12.2024
UDC 004.[048+89]
Correspondence author: B. S. Lukashchuk bohdan.lukashchuk@gmail.com
ISSN 1994-7836 (print)
ISSN 2519-2477 (online)

B. S. Lukashchuk[1], Yu. V. Shabatura[2]

[1] *Ukrainian National Forestry University, Lviv, Ukraine*

[2] *Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine*

METHODOLOGY FOR EVALUATING COMPLEX OBJECT CONTOUR DETECTION ACCURACY IN SLIC-BASED IMAGE SEGMENTATION

This paper investigates the application of the Simple Linear Iterative Clustering (SLIC) algorithm for complex object image segmentation, on the example of images of human body injuries. The study solves the problem of the lack of quantitative evidence regarding SLIC's algorithm performance in high-precision area and boundary assessment of the lesion on a digital image of a human body with a wound injury on it. A comprehensive methodology is developed to evaluate SLIC's algorithm efficacy across various complex images and image resolutions. The research utilizes a combined dataset of 3696 wound images from the Foot Ulcer Segmentation Challenge (FUSeg) and WoundSeg datasets. Bayesian optimization is utilized to fine-tune SLIC algorithm hyperparameters, focusing on the number of segments and compactness. Results indicate that SLIC algorithm demonstrates consistent performance across different implementations, achieving Dice scores around 0.84 and Soft Boundary F1 scores around 0.55. The study reveals that the optimal number of segments for SLIC algorithm can be defined relative to the spatial dimensions of the input image, with maximal image dimension *2 being the most effective value. A thorough analysis of various segmentation metrics is conducted, including IoU, Dice Score, and Boundary F1 Score. The research introduces and employs the Soft Boundary F1 Score – modification of Boundary F1 Score, a novel metric designed to provide a more nuanced evaluation of boundary detection accuracy while offering a smoother optimization landscape. This metric proves particularly valuable in assessing the performance of SLIC algorithm in image with complex objects on them segmentation tasks. Importantly, this research presents an idealized SLIC-based segmentation approach, where superpixels are optimally combined using ground-truth masks to establish an upper bound of performance. This idealized SLIC algorithm segmentation is compared with a pre-trained on the FUSeg dataset UNet model, showcasing superior generalization capability across diverse wound types. On the WoundSeg dataset, the idealized SLIC algorithm approach achieved a Dice score of 0.84, significantly outperforming the UNet model (0.12 Dice score). As a result, this study provides valuable insights for improving complex objects segmentation methods and highlights the need for further research on developing effective methods for superpixel classification in real-world scenarios. The findings also highlight the potential of SLIC-based approaches in addressing the challenges of diverse data types and limited training data.

Keywords: image processing; semantic segmentation; algorithm; medical imaging; Bayesian optimization; deep learning.

Introduction / Вступ

Automatic segmentation of complex objects in digital images, classification of these objects and their components are important areas of research in the field of computer image analysis (computer vision). These approaches can be effectively applied in medicine for segmentation of wound lesions on human skin and classification of affected tissues. This contributes greatly to the automation of patient examination, treatment planning, and monitoring of healing progress [5, 7, 26]. Research on this topic mainly concentrates on utilizing supervised machine learning models, like different variants of artificial neural networks (ANN). The drawback of these approaches is the lack of generalization on new data. Thus, some of the works [4, 8, 23] concentrate on combining ANNs with unsupervised machine learning techniques, the most frequently used and the most effective of which is the k-means-based superpixel segmentation algorithm – Simple Linear Iterative Clustering (SLIC) [29]. For example, authors of the QTDU method [4] utilized SLIC for initial image segmentation before employing deep learning models for tissue classification. Chang et al. [8] integrated SLIC algorithm to facilitate the annotation process of pressure ulcer images, highlighting its use in outlining tissue boundaries – the annotator must label not every pixel, but every superpixel (groups of pixels grouped together by certain similarity into clusters). Another study used SLIC algorithm to segment pre-isolated wound areas for subsequent tissue classification [23].

However, despite its widespread use, the literature lacks clear quantitative evidence regarding SLIC's algorithm effectiveness for precise segmentation of complex objects and especially their boundaries on digital images, such as wounds on human body. There is also a lack of systematic evaluation of SLIC algorithm hyperparameters (parameters

Інформація про авторів:

Лукащук Богдан Сергійович, аспірант, кафедра інформаційних технологій. **Email:** bohdan.lukashchuk@gmail.com; https://orcid.org/0009-0002-2365-4923

Шабатура Юрій Васильович, д-р техн. наук, професор, кафедра електромеханіки та електроніки. **Email:** shabaturayuriy@gmail.com; https://orcid.org/0000-0002-9961-1244

Цитування за ДСТУ: Лукащук Б. С., Шабатура Ю. В. Методика оцінювання точності ідентифікації контурів зображень складних об'єктів, виділених у процесі сегментації алгоритмом SLIC. Науковий вісник НЛТУ України. 2024, т. 34, № 8. С. 109–119.

Citation APA: Lukashchuk, B. S., Shabatura, Yu. V. (2024). Methodology for evaluating complex object contour detection accuracy in SLIC-based image segmentation. *Scientific Bulletin of UNFU*, 34(8), 109–119. https://doi.org/10.36930/40340813

that remain constant during algorithm execution and are set to control the learning process), such as the optimal number of superpixels or compactness values, for various wound image characteristics.

Given this context, a comprehensive study to evaluate SLIC's algorithm complex object segmentation effectiveness is required, as well as establishing a methodology for such evaluation. Such research would fill a gap in the current literature and provide valuable insights for improving segmentation methods.

The research task is to develop a methodology for performing complex object image segmentation (demonstrated on human body wound images) based on the SLIC algorithm, as well as multi-criteria evaluation of its effectiveness.

Object of research – the complex digital image analysis and their segmentation using the SLIC algorithm.

Subject of research – the comparative evaluation of the efficiency of using the SLIC algorithm in solving complex digital image segmentation tasks, which allows to identify optimal segmentation parameters and evaluate the algorithm's execution performance across diverse datasets, thereby ensuring adaptability and robustness in segmentation task of images with complex boundaries.

The purpose *of the* research – to develop a methodology for performing analysis and optimal segmentation of complex digital images based on the SLIC algorithm application, enabling maximum effectiveness and accuracy in the boundary detection of specified image elements and the assessment of their area.

To achieve this *purpose*, the following main *research objectives are identified*:

1. To conduct an analytical review of modern metrics, applied in assessing segmentation boundary detection accuracy digital image, as well as approaches to SLIC algorithm application for segmentation of objects with complex shapes on the images, which enables to identify the segmentation evaluation metric most suitable, for assessing SLIC algorithm performance, particularly for the tasks involving complex object boundaries detection.
2. To develop a methodology for evaluating the effectiveness of the SLIC algorithm in complex object segmentation tasks with emphasis on boundary detection accuracy (demonstrated on the task of human skin wound segmentation on the images), which will allow for systematic assessment of SLIC algorithm performance across diverse data domains, providing insights into its adaptability and robustness for various segmentation challenges.
3. To assess SLIC's algorithm segmentation effectiveness across different hyperparameter settings using Bayesian optimization, algorithm implementations and image resolutions using publicly available datasets, which will provide a setup of the optimal configurations needed to achieve high-quality segmentation of complex objects using the SLIC algorithm.
4. To conduct a comparative analysis of segmentation results obtained using the SLIC algorithm and Convolutional Neural Network (CNN), commonly used in the medical imaging, which will help to determine scenarios and parameter settings where the SLIC algorithm offers advantages over CNNs in terms of segmentation quality and adaptability to limited training data.

By completing this task, we provide the scientific community with not only an assessment of how effectively superpixels, generated by SLIC algorithm segment complex object (like wounds) edges but also a methodology for reproducing them with other input data. This should have a great positive effect on future digital image automated analysis research.

Analysis of recent research and publications. Simple linear image clustering (SLIC) is a superpixel generation algorithm proposed by Radhakrishna Achanta and colleagues that adapts the k-means clustering approach in the image domain for efficient superpixel generation. This method has become an essential tool of computer vision due to its simplicity, efficiency, and quality of contour selection for the image. SLIC algorithm works by clustering pixels by combining color properties in CIELAB and spatial coordinates in 5-dimensional space. The advantage of this method is that unlike classical k-means, which check the distance from each pixel to each other, in SLIC algorithm, the search space is limited to areas proportional to the size of the superpixel. In the context of complex image analysis (like the ones, with wounds on human skin), according to the already mentioned papers [4, 8, 23] – SLIC algorithm has proven particularly useful due to its ability to generate superpixels that fit well with seeking object boundaries. Nevertheless, there are some peculiarities of algorithm use in these methods – they do not expect to cover a full object or some specific wound tissue area with one superpixel; the authors tackle this problem by generating multiple superpixels and then merging them, based on some deep-learning based similarity measure (like superpixel classification to one of the wound tissue types or wound/not wound).

Authors of [4] and [23] do not state how many superpixels they have divided their images into. Authors of [8] state that they have tested different numbers of superpixel hyperparameter values, from 500 to 2000, on the fixed images with the size of 1000×750 px. and while an increasing number of superpixels causes better wound outline detection (no quantitative result provided), it becomes hard for the annotator to annotate them as well as for deep learning model to perform classification, because of texture information lost effectively. Thus, the authors conclude that 800-1000 superpixels are optimal. The authors of [14] chose another approach – they tend to segment images with wounds on them into a relatively small number of segments (superpixels) (2, 3, and 4) utilizing ordinary k-means clustering instead of SLIC algorithm. However, the authors utilize a small dataset of only nine images of three classes (three wound types). They do not show any statistically significant results but conclude that segmentation into a larger number of segments produces better results. Thus, in our research, we consider experimentally exploring the selection of many superpixels relative to the image resolution (more on this in the Materials and methods of research section) and merging them to cover the whole object area.

The other problem is selecting an efficiency evaluation metric. Traditionally used metrics for image semantic segmentation assessment are Dice Score and Intersection over Union (IoU) [22]. While being performant for training segmentation models, they are not designed to pay extra attention to the segmentation quality at the boundaries.

One of the famous metrics used in medical imaging for boundary segmentation accuracy evaluation is Hausdorff Distance (HD) [10], which measures the maximum distance between the points on the predicted boundary and the closest points on the ground truth boundary. A lower HD indicates a more accurate boundary detection. It is especially useful for identifying significant localized errors in segmentation [21]. The average Hausdorff distance is widely used to assess segmentation performance by comparing ground truth images with segmentations. However, the average HD can lead to ranking errors. Authors of [2] proposed a modified version, which balances the average Hausdorff distance, making it more suitable for quality assessment in medical imaging. The main weakness of this metric arises from its main strength – it is highly sensitive to outliers and does not present information on the overall boundary accuracy.

The other quite similar metric is the average symmetric surface distance (ASSD), which is the average of all the distances from points on the boundary of a machine-segmented region to the boundary of the ground truth and vice versa [32]. ASSD provides a more balanced and comprehensive view of boundary accuracy than HD. However, ASSD is still sensitive to outliers, which can significantly affect the results. This sensitivity appears because ASSD calculates the mean of all point-to-surface distances, and a few significant errors can affect the average result. In regions with considerable dissimilarities, such as those with complex or irregular shapes, ASSD may not accurately reflect the true segmentation quality. The metric does not account for the distribution of errors across the surface, potentially masking areas with poor segmentation performance [15].

The Boundary F1 Score, a harmonic mean between Precision and Recall of the boundary pixels, has gained popularity in evaluating segmentation boundaries, particularly in tasks where precise edge alignment is crucial. It is designed to eliminate the limitations of region-based metrics like the Dice Score, which may achieve high scores even when boundary errors are present [6]. This metric is less sensitive to outliers than the Hausdorff distance, making it more robust in scenarios where precise boundary detection is crucial [2]. However, it is essential to remember that the Boundary F1 Score can be sensitive to slight variations in boundary placement and does not account for the severity of misalignments, treating all errors equally. Despite these limitations, its focus on boundary accuracy and its balanced nature makes it particularly valuable for evaluating SLIC algorithm performance in wound segmentation tasks.

While SLIC algorithm has shown promise in various medical imaging applications, there is a clear gap in systematically evaluating its effectiveness specifically for complex object border detection. The existing metrics, such as IoU, Dice Score, HD, ASSD and Boundary F1 Score, each have strengths and limitations in assessing segmentation quality. Based on the review of the most popular metrics, used to assess boundary detection precision, we decided not to use HD and ASSD metrics for the further methodology development, as they are much less descriptive in terms of overall wound boundary detection precision and concentrate on exploring IoU, Dice and Boundary F1 scores.

Materials and methods of research. This section focuses on establishing the foundation to solve the research tasks. We develop methodology of SLIC algorithm segmentation evaluation by applying supervised segmentation masks to merge the superpixels, generated by SLIC algorithm into one, and evaluate the segmentation quality, paying particular attention to the seeking object border detection correctness through advanced metrics. The proposed methodology also incorporates cross-validation and Bayesian optimization [3] to fine-tune the SLIC algorithm hyperparameter selection and achieve optimal segmentation performance. Finally, we compare the most performant SLIC algorithm variant with the ANN, specifically trained for complex object (like human skin wounds) segmentation. Each component is described in detail to ensure reproducibility and clarity of understanding.

Research results and their discussion / Результати дослідження та їх обговорення

Overview of the SLIC algorithm segmentation evaluation methodology. The methodology consists of the several key steps:

1. Generating superpixels.
2. Applying supervised segmentation masks.
3. Creating superpixel masks.
4. Evaluating segmentation quality using specialized metrics.
5. Optimizing the process of SLIC algorithm hyperparameter selection through cross-validation and Bayesian search.

Figure 1 shows the flowchart representation of these steps:

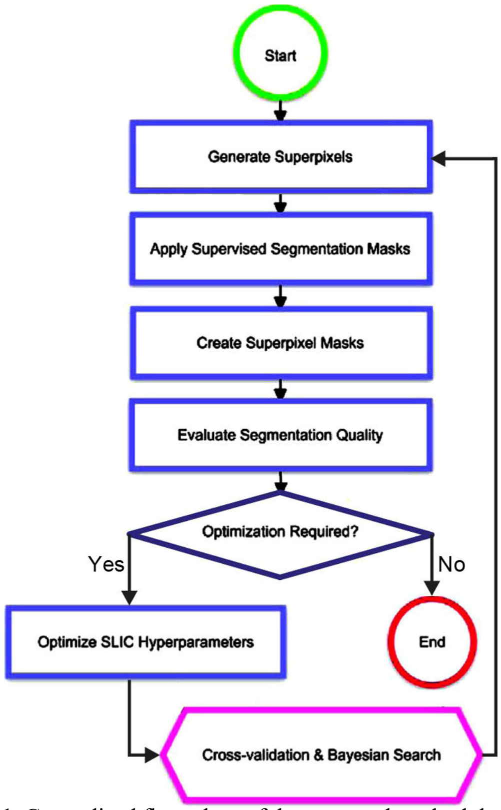

Fig. 1. Generalized flow-chart of the proposed methodology of complex object segmentation quality evaluation / Узагальнена блок-схема запропонованої методології оцінювання якості сегментації зображень складних об'єктів

Superpixel generation as an abstraction. The first step is generating superpixels from the wound images using the SLIC algorithm. The SLIC algorithm implementation is treated as an abstraction, meaning it can be easily replaced or modified without changing the workflow. This abstraction allows for the flexibility to experiment with different SLIC algorithm implementation options and hyperparameter settings, providing the ability to explore the optimal configuration for the specific functional and non-functional requirements of the specific wound-segmentation task.

The superpixel generation process begins with the input image, followed by selecting the SLIC algorithm variant and parameters. The algorithm then processes the image and outputs a label map, assigning each pixel to a superpixel cluster. This label map is used in the following steps to apply the supervised segmentation mask (ground truth) and generate the superpixel segmentation mask.

Application of supervised segmentation mask and superpixel segmentation mask generation. Once the superpixels are generated, the next step is to apply a supervised segmentation mask – ground-truth annotation of the object of interest. This step is crucial for identifying which superpixels intersect with the annotated wound region, thereby isolating the relevant superpixels that will form the basis of the final segmentation. The ground-truth segmentation mask is overlaid on the superpixel map. Superpixels that intersect with the mask for at least 10 percent of the superpixel area are identified and marked as relevant. They are merged to form a unified "superpixel mask". This mask represents the final predicted segmented object of interest (wound) area. It is used to evaluate the accuracy of the segmentation, as we desire to have it and its borders be as accurate as possible:

$$\alpha(l) = \frac{1}{|l|}\sum_{(x,y\in l)} M(x,y)\,, \tag{1}$$

where $\alpha(l)$ is an overlap ratio, $M(x, y)$ is the ground-truth segmentation mask function, $M(x, y) = 1$ indicates that the pixel (x, y) belongs to the wound region, $|l|$ is the total number of pixels in superpixel l,

$$R(l) = \left\{\begin{matrix} 1,\ if\,\alpha(l) \geq 0 \\ 0,\ otherwise \end{matrix}\right\}, \tag{2}$$

where $R(l)$ is the relevance function for the superpixel l, returning 1 if the superpixel is considered relevant (overlaps with the wound region by at least 10 %) and 0 otherwise.

Metric Calculation and Evaluation. Evaluating the accuracy of the superpixel mask is a critical aspect of the framework, as it determines the overall effectiveness of the segmentation process. We review several metrics for this, including the already mentioned, widely used semantic segmentation tasks, such as the IoU and Dice Score, but the essential one is the Boundary F1 Score.

Intersection over Union (IoU) is a standard metric in segmentation tasks; it measures the overlap between the predicted superpixel mask and the ground truth mask. IoU is particularly useful for assessing the overall accuracy of the segmentation in terms of area coverage.

The Dice Score is another standard metric for segmentation evaluation. It focuses on the similarity between the predicted and ground truth masks. It is sensitive to false positives and false negatives, making it a robust metric for general segmentation tasks.

Boundary F1 Score is a more specific metric. In its heart, it is a Dice Score, but it is calculated only for the pixels on the boundaries of the predicted mask compared with the ground truth one. This metric is useful in medical image analysis, where, as mentioned earlier, precise boundary detection is critical.

Let us review Figure 2 to illustrate the effect of the different segmentation experiment outcomes on these metric values.

Figure 2 shows a black polygon outline as an illustrative representation of the ground-truth mask and a colored dashed outline as the predicted mask. Ten different examples are used to calculate three aforementioned metrics and illustrate their behavior based on the predicted/ground-truth relation. Figure 3 shows the bar chart of the values of these metrics for each of the made-of examples.

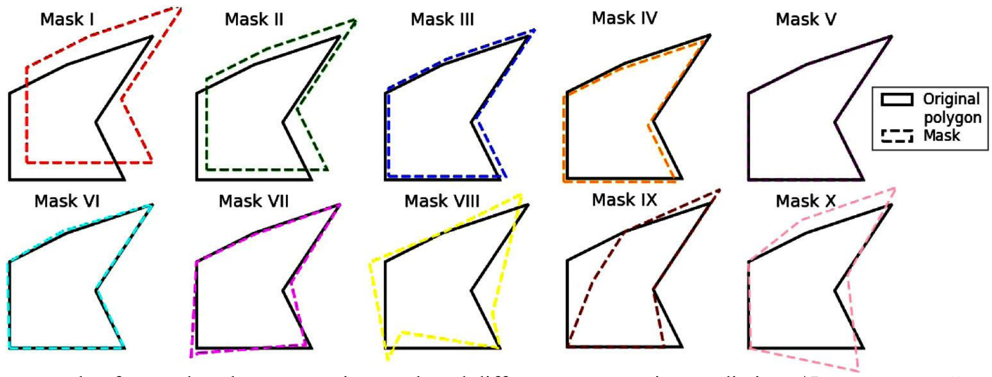

Fig. 2. Illustrative example of ground truth segmentation mask and different segmentation predictions / Ілюстративний приклад еталонної маски сегментації та різних прогнозів сегментації

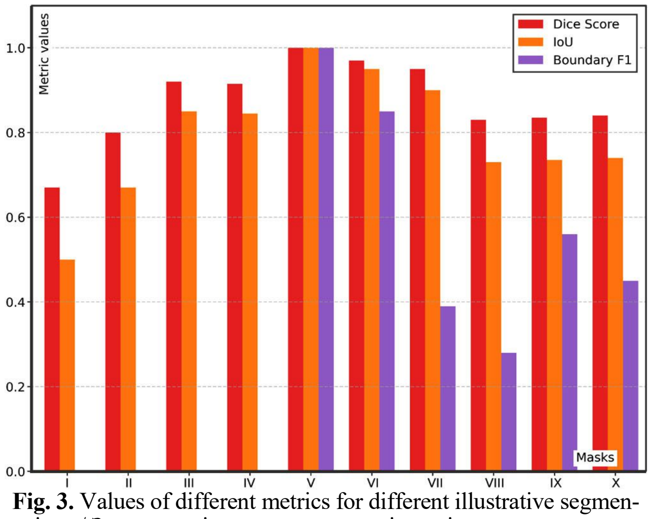

Fig. 3. Values of different metrics for different illustrative segmentations / Значення різних метрик для різних ілюстративних сегментацій

The metrics calculated for the masks reveal a main difference between traditional overlap measures (IoU, F1 Score). For masks, 1 to 4, both IoU and Dice show moderate to high overlap, with scores ranging from 0.50 to 0.92. However, the Boundary F1 Score remains 0 for these masks, highlighting a significant limitation: while the overall area overlap might be considerable, these metrics fail to capture the essential aspect for some research areas – boundary alignment. This issue is evident when the mask is misaligned or "shifted", leading to zero boundary precision despite significant area overlap.

In contrast, masks VI to X demonstrate that even when IoU and Dice are constantly high with smooth changes, the Boundary F1 Score can vary significantly (from 0.26 to 0.86), reflecting differences in how well the boundaries align. The Boundary F1 Score's sensitivity to edge alignment is a key strength, making it indispensable in applications where precise delineation of object edges is crucial, such as in medical imaging.

The tolerance parameter in the Boundary F1 Score plays a significant role in the result. It defines the maximal allowed distance within which boundary points are considered aligned. A smaller tolerance increases the strictness of the metric, lowering the Boundary F1 Score even if the overall segmentation is relatively accurate. A larger tolerance mitigates minor boundary misalignments, leading to a higher score. Therefore, the choice of tolerance directly affects the balance between penalizing boundary errors and recognizing overall segmentation quality.

This analysis highlights the limitations of using only IoU or Dice Score metrics for segmentation, as they do not account for boundary precision. Boudnary F1 addresses this issue, as illustrated by the steep decline in its values compared to other metrics when segmentation quality changes. However, its high sensitivity to minor mask variations may be impractical during the early stages of SLIC algorithm hyperparameter optimization, particularly for images with complex object boundaries. IoU and Dice may provide smoother landscapes but fail to capture crucial boundary information, while Boundary F1 may create overly steep gradients, potentially leading to suboptimal local minimum that don't accurately reflect wound segmentation quality and is too strict.

We propose to utilize Soft Boundary F1 Score in Bayesian optimization to eliminate the aforementioned issues. The metric is calculated using the same general approach, as in the already mentioned Boundary F1 Score [6], but with the addition of exponential distance-based tolerance, allowing for small misalignments in boundary detection:

$$SoftBF1(P,G,\tau)=\frac{2\cdot(SP\cdot SR)}{SP+SR}. \quad (3)$$

Where P is the predicted segmentation mask, G is the ground-truth mask, τ – distance tolerance, SP (Soft Precision) (4), and SR (Soft Recall) (5) are calculated using distance transforms.

$$SP=\frac{\sum \exp(-Dist(G/\tau)\cdot P_{boundary})}{\sum P_{boundary}}, \quad (4)$$

$$SR=\frac{\sum \exp(-Dist(P/\tau)\cdot G_{boundary})}{\sum G_{boundary}}, \quad (5)$$

where $Dist(X)$ is the distance transform of mask X (X represents here either P or G mask), $X_{boundary}$ – boundary pixels of mask X, τ – is the distance tolerance parameter that controls how quickly the exponential term decreases when distance increases. Smaller τ causes metric to be more sensitive to differences of the border between predicted and ground-truth. In our experiments, we use 2–4-pixel values for it. This exponential term creates a smoother distance-based weighting that decreases as the distance between predicted and ground truth boundaries increases in the. This allows for some tolerance in boundary detection and thus smoother gradient during optimization, while still prioritizing boundary accuracy.

The Soft Boundary F1 score has several advantages:

1. It provides a smoother optimization gradient, even with significant boundary misalignments, due to its tolerance for small displacements (this parameter should be carefully selected based on the level of edge detection accuracy, you want to preserve and your input image size).
2. It remains sensitive to incremental improvements in boundary accuracy, allowing for fine-tuning of the segmentation.
3. It balances the assessment of overall wound region segmentation with boundary detection.

The Soft Boundary F1 score achieves this by incorporating a distance-based tolerance in its calculation, allowing for small misalignments in boundary detection. This approach is particularly suitable for wound segmentation, where precise boundary delineation is crucial, but minor deviations should not be overly penalized.

Figure 4 shows the behavior of SoftBF1 compared to previously mentioned metrics across different segmentation scenarios (same illustrative polygons and masks, as in the previous example).

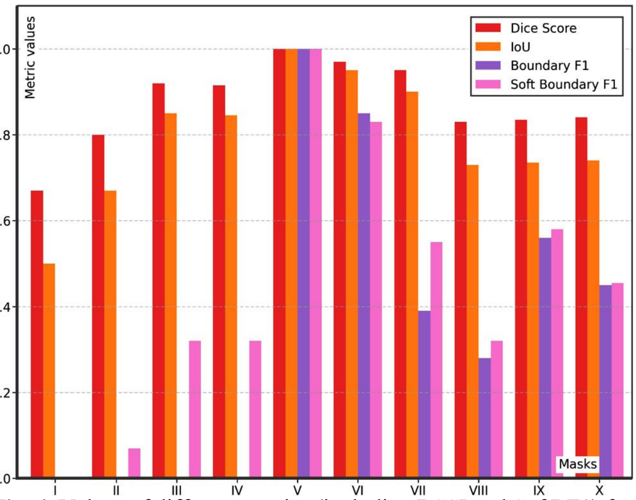

Fig. 4. Values of different metrics (including BSSI and SoftBF1) for different illustrative segmentations / Значення різних метрик (враховуючи BSSI та SoftBF1) для різних ілюстративних сегментацій

The Soft Boundary F1 Score, a modification of the standard Boundary F1, it resolves the issue of zero metric values for minor mask shifts, ensuring smoother gradients in the optimization process. As shown on the Figure 4, SoftBF1 avoids the steep penalties, observed in Boudnary F1 (e.g., masks I-III), while maintaining comparable performance across scenarios, enabling more robust optimization is segmentation tasks. However, as this metric is a result of preliminary research, further experimentation and validation on diverse datasets and segmentation tasks are required to fully assess its effectiveness and generalizability.

Data. Two datasets in wound image segmentation are publicly available: the Foot Ulcer Segmentation Challenge Dataset (FUSeg) [31] and the WoundSeg Dataset [24]. Each contains images of wounds on human skin and corresponding segmentation masks annotated by professionals. We decided to use both to evaluate the proposed approach of SLIC algorithm hyperparameter tuning and the ability of superpixels to outline areas of the objects of interest.

The FUSeg dataset, initially created for the Foot Ulcer Segmentation Challenge, comprises approximately 1210 images of diabetic foot ulcers, with around 1110 of them being publicly available. Each image in this dataset has a resolution of 512×512 pixels, providing high-detail representations of diabetic ulcers. The obvious drawback is that this dataset contains only one type of wound image, but until recently, it was the only public dataset in this area.

WoundSeg Dataset offers a broader spectrum of wound types. It contains 2686 images with corresponding segmentation masks, with a resolution of 384×384 pixels. This dataset contains images of wounds of different types: diabetic ulcers, pressure ulcers, traumatic wounds, venous ulcers, surgical wounds, arterial ulcers, and phlegmon.

We merged these datasets, resulting in a combined set of approximately 3696 wound images and their corresponding ground truth segmentation masks. The combined dataset allows us to assess our SLIC algorithm validation and hyperparameter tuning method not only on multiple images but also considering the great variety of wound types and two different resolutions. The latter allows us to evaluate the robustness of our SLIC algorithm hyperparameter optimization across different image scales, which is crucial for developing a wound segmentation system applicable to various clinical imaging setups.

Experiment Setup. This section explains the proposed methodology for optimizing algorithm hyperparameters in wound segmentation. We use the combined dataset described in the Data section, comprising 3696 wound images from FUSeg and WoundSeg datasets, with resolutions of 512×512 and 384×384 pixels, respectively.

Experiment steps:

1. Data preparation:
   a) Load and preprocess images as described in the Data section.
   b) Sample 300 images from the whole dataset, preserving equal proportion of images from each of the datasets (150 FUSeg, 150 WoundSeg)
   c) Prepare the dataset for 3-fold cross-validation; each fold has the same proportion of different image resolutions.
2. SLIC algorithm implementation setup: we evaluate three SLIC implementation options (first two – CPU-optimized, last one – GPU version of the algorithm):
   a) Scikit-image SLIC algorithm implementation [30].
   b) FastSLIC implementation [1].
   c) cuda-SLIC implementation [28].
3. Hyperparameter optimization: for each SLIC implementation and each fold, use scikit-learn's Bayesian optimization to find optimal values for the following parameters:
   a) Number of segments (*n_segments*): we explore three options based on the maximum image dimension (*max_dim* = *max* (*width, height*)): *max_dim* *2; *max_dim*; *max_dim* / 2; *max_dim* / 4. Using the maximum image width/height value for *n_segments* came as an intuitive insight during initial experimentation on the superpixel generation methods [12]. Here, we want to prove or discard this idea. The logic for using the maximum dimension (*max_dim*) and its fractions for *n_segments* is hidden in maintaining a balance between the segmentation quality of boundary detection and computational efficiency. Using *max_dim* ensures that even the longest edge of the image can be segmented into separate superpixels, potentially capturing details of wound boundaries. The fractions (*max_dim*/2 and *max_dim*/4) allow rougher segmentations, which may work better for larger wound areas or for reducing computational resource use.
   b) Compactness: value range between 1 and 30.
   c) Objective function – the Soft Boundary F1 Score metric, introduced earlier.
4. Optimization process: for each SLIC algorithm implementation and each fold:
   a) Perform Bayesian optimization for 20 iterations or until convergence on the training set (2 folds). With the 30 iterations we get 10 pairs of hyperparameters being tested.
   b) At each iteration:
      – Generate superpixels using current hyperparameters.
      – Apply the process described in "Application of supervised segmentation mask and superpixel segmentation mask generation".
      – Calculate the Soft Boundary F1 Score metric.
      – Evaluate the best hyperparameters on the validation fold.
5. Cross-validation aggregation:
   a) Average the best Soft Boundary F1 Score scores and optimal hyperparameters across all folds.
   b) Calculate the standard deviation of Soft Boundary F1 Score values.
6. Final Evaluation:
   a) Compute average Soft Boundary F1 Score metric, and average execution time on the entire data set using the SLIC algorithm implementation and hyperparameter values, obtained in the optimization process.

For the Bayesian optimization of SLIC algorithm hyperparameters, we use the Gaussian Process Regressor [3] as the surrogate model, implemented in the scikit-optimize package. The objective function is set to maximize the Soft Boundary F1 Score metric. We use the Expected Improvement (EI) [3] acquisition function to balance exploration and exploitation.

Google Colab Pro with Nvidia A100 GPU was used to conduct experiments.

As a final evaluation step, we compare SLIC algorithm segmentation with a deep-learning approach to validate our findings within the broader field of wound segmentation. It is important to note that in this comparison, SLIC algorithm is not used as a standalone segmentation method but rather as a preprocessing step to generate superpixels. These superpixels are then combined using ground-truth masks to simulate an ideal superpixel classification, which is the goal for future research. This approach allows us to assess the upper bound of SLIC's algorithm performance and its potential as a preprocessing tool for wound segmentation.

Due to its widespread use in medical image segmentation tasks, we chose the UNet [27] model as a representative deep learning method. The UNet model was pre-trained on the Foot Ulcer segmentation dataset, achieving a Dice score of 0.86 on the test dataset. The model architecture follows the original UNet design with minor modifications: it expects 512×512 resolution RGB images as input, uses a 32-channel initial convolution layer, and incorporates BatchNormalization [17] layers after each convolution layer to reduce internal covariate shift.

The primary goal of this comparison is to assess the potential of SLIC-based approaches against a deep learning method for complex object area segmentation when applied to diverse wound types. We evaluate both methods on the entire WoundSeg dataset, which contains various wound types, not present in the Foot Ulcer dataset used for UNet training. This comparison is intended to provide insights into the robustness and adaptability of SLIC algorithm for wound segmentation across different wound types and image characteristics.

SLIC algorithm implementations evaluation. The experiment results provide valuable information on the SLIC algorithm performance for wound image segmentation. Table 1 below shows the SLIC algorithm segmentation Dice Score, Soft Boundary F1 Score, assnd Execution time (in seconds) for the three mentioned earlier SLIC algorithm implementations and the corresponding best sets of SLIC algorithm hyperparameters found during Bayesian optimization.

Table 1. Algorithm segmentation quality and execution time comparison for the three SLIC implementations / Порівняння якості сегментації та тривалості виконання алгоритму для трьох реалізацій SLIC

Implementation	Dice Score (mean$^{\pm std}$)	Soft Boundary F1 (mean$^{\pm std}$)	Execution Time (s) (mean$^{\pm std}$)
cuda-slic	$0.8430^{\pm 0.1023}$	$0.5537^{\pm 0.1239}$	$0.0303^{\pm 0.0116}$
fast-slic	$0.8480^{\pm 0.0964}$	$0.5649^{\pm 0.1264}$	$0.0223^{\pm 0.0706}$
skimage	$0.8426^{\pm 0.1014}$	$0.5506^{\pm 0.1264}$	$0.1934^{\pm 0.0922}$

The Bayesian optimization for each SLIC algorithm implementation converged on very similar hyperparameter values. All three implementations received the best results at *n_segments* multiplier of 2, corresponding to *max_dim* *2 segments. The results support our initial hypothesis that using the maximum image dimension to determine the number of segments is effective for wound segmentation on images with different dimensions.

The optimal compactness values differ across implementations, with skimage having the highest compactness (19), cuda-slic (12) and fast-slic (8). This variation suggests that either different implementations may benefit from different levels of spatial compactness in the superpixel generation process or that there is no considerable difference in the result for the approximate range of compactness [8, 21]. The last conclusion is supported by the chart, demonstrating Bayesian optimization convergence curve (Soft Boundary F1) maximization progression (Figure 5).

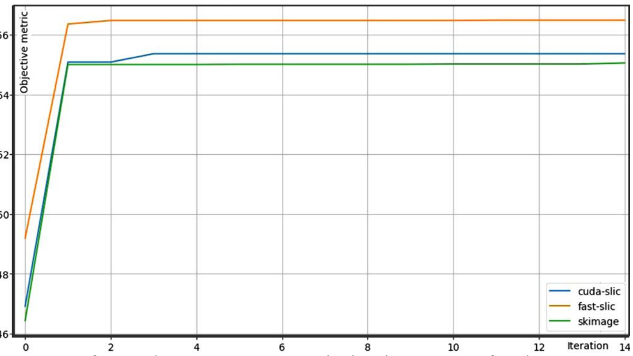

Fig. 5. Soft Boudnary F1 Score optimization curves for three SLIC algorithm implementations / Криві оптимізації метрики Soft Boundnary F1 для трьох реалізацій алгоритму SLIC

Optimization curve demonstrates that while fast-slic convering the fastest, all three implementations reach their best values within the first two iterations and later only cuda-slic implementation has optimization objective improvement on the third iteration.

All implementations show almost identical Dice and Soft Boudnary F1 Score values, with very significant differences between these metrics. Dice values are around 0.84, and Soft Boundary F1 values are around 0.55. Results indicate that while SLIC algorithm superpixels provide good segmentation quality in traditional segmentation metrics, improvements must be made to boundary detection. However, only final step of evaluation can compare algorithm and UNet model boundary detection correctness.

In terms of execution time, results provide unexpected insights. Fast-slick implementation appears to be the fastest on average (0.0223 seconds), but the standard deviation value (0.0706), which is larger than the average itself, is a clear descriptor of significant outliers. Skimage implementation ($0.1934^{\pm 0.0922}$) has the same problem with the outliers in the execution time, but it is also the slowest implementation among the tested. Cuda-slic implementation shows mean execution time being close to the fast-slic implementation with much greater stability in the standard deviation ($0.0303^{\pm 0.0116}$). As segmentation metric results are almost similar across the implementations, execution time indicates that cuda-slic implementation is the best option for further research.

To better examine Dice and Soft Boudary F1 scores, we present Figures 6-7, which illustrate this metric values distribution for the FUSeg and WoundSeg datasets for the cuda-slic implementation of the SLIC algorithm.

The two datasets have no notable differences regarding the Dice Score distribution. Both are right-skewed distributions, with FUSeg being considerably less steep. This right-skewness suggests that while most samples achieve high scores, there is a "tail" of specific cases that are mis segmented and require further attention. For both datasets, central values fall within the approximate range of [0.8; 0.9]. Both datasets contain few samples with small amounts of zero or near-zero Dice Score metric values; notably, the FUSeg dataset includes more outliers, an important aspect for further examination. We could also see that some of the FUSEG samples have perfect Dice Score values.

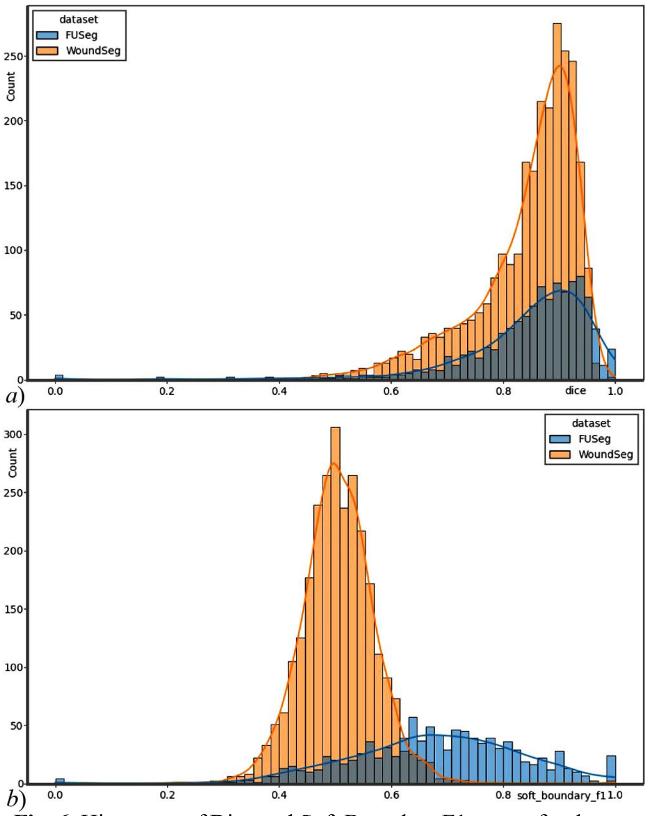

Fig. 6. Histogram of Dice and Soft Boundary F1 scores for the cuda-slic implementation of the SLIC algorithm / Гістограми метрик під час використання cuda-slic реалізації алгоритму SLIC

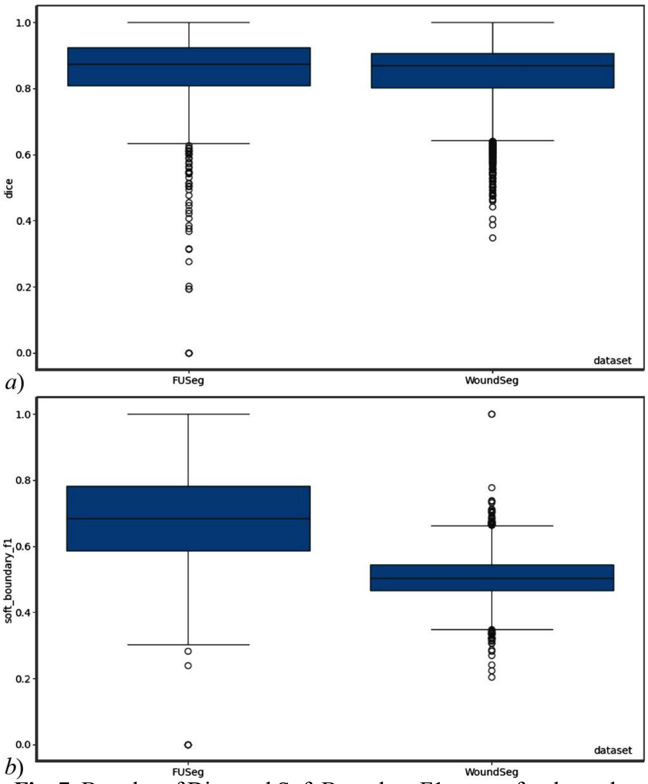

Fig. 7. Boxplot of Dice and Soft Boundary F1 scores for the cuda-slic implementation of the SLIC algorithm / Коробчаста діаграма результатів Dice і Soft Boundary F1 під час використання cuda-slic реалізації алгоритму SLIC

The Soft Boundary F1 Score distributions reveal a difference between WoundSeg and FUSeg. The WoundSeg metric distribution is much steeper than FUSeg and is centered in the range [0.45; 0.55], while the FUSeg dataset distribution center lies more in the range [0.6; 0.8] but is much smoother. As with the Dice Score distribution, samples have zero metric values. The difference between WoundSeg and FUSeg distributions indicates that different wound types possess complex and diverse outlines. While we can detect boundaries with better precision for one wound type, such as foot ulcer wounds in the FUSeg dataset, results for a mix of wound types are comparatively worse despite maintaining similar Dice Scores.

In conclusion, this distribution analysis reveals that while segmentation, based on the SLIC algorithm, performs consistently across implementations, it shows varying efficacy for different wound types and metrics. The algorithm generally achieves a high overall segmentation Dice Score of 0.84, which indicates the applicability of SLIC algorithm superpixels for wound segmentation (as examples of complex objects), but it demonstrates more variability in boundary precision, particularly for diverse object types.

After analyzing samples with extremely poor segmentation quality (zero or almost zero SoftBFF1 and Dice scores), it was discovered this this is mostly caused by one of two reasons:

- Object (wound) is not present on the image at all (very rare case, few samples per whole dataset).
- Object (wound) segment is very small, so even the smallest SLIC algorithm superpixel cannot capture its boundaries.

The second case is especially complex to be solved by the further SLIC algorithm tuning, as forcing superpixels to be smaller till they capture the smallest possible regions may bring them back to being just ordinary pixels, thus eliminating all benefits of utilizing superpixels.

Figure 8 shows an example of good SLIC algorithm segmentation (cropped near the wound to emphasize the segmentation), where the ground-truth mask is overlayed with Green, and the SLIC algorithm segmentation mask with Red and Yellow indicates the area of their intersection.

Fig. 8. Example of an ideal SLIC algorithm segmentation / Приклад досконалої сегментації алгоритмом SLIC

SLIC algorithm and UNet wound segmentation comparison. The comparison between SLIC algorithm and UNet model for wound segmentation on the WoundSeg dataset reveals striking differences in performance and generalization capabilities (Table 2). For SLIC algorithm we use cuda-slic implementation with compactness hyperparameter 12 and *n_segments* multiplier equals 2.

Table 2. Comparison of cuda-slic SLIC algorithm implementation with the UNet for segmentation on the WoundSeg dataset / Порівняння cuda-slic реалізації алгоритму SLIC та моделі UNet для сегментації ран із набору даних WoundSeg

Model	Dice Score (mean$^{\pm std}$)	Soft Boundary F1 (mean$^{\pm std}$)
cuda-slic SLIC	$0.8416^{\pm 0.0936}$	$0.5057^{\pm 0.0653}$
Unet	$0.1181^{\pm 0.1947}$	$0.0759^{\pm 0.1223}$

The approaches, based on the SLIC algorithm, where superpixels are optimally combined using ground-truth masks, demonstrate consistent performance with a Dice score of $0.8416^{\pm 0.0936}$ and a Soft Boundary F1 score of $0.5057^{\pm 0.0653}$. The small standard deviations indicate stable performance across various object types. In contrast, Unet shows significantly lower performance, with a Dice score of $0.1181^{\pm 0.1947}$ and a Soft Boundary F1 score of $0.0759^{\pm 0.1223}$. The large standard deviations for Unet suggest highly variable performance across different images.

These results highlight the potential of the aforementioned approaches for adapting to diverse wound types. The UNet model, despite its sophistication and high performance on the Foot Ulcer dataset, fails to generalize effectively to other wound types. This contrast underscores a key advantage of the approach, that utilizes SLIC algorithm: its ability to adapt to various wound characteristics without requiring training data or computational resources. SLIC algorithm segmentation could be a useful preprocessing tool for wound segmentation in clinical settings where wound diversity is high and annotated training data may be limited.

However, it is important to note that while the approach, based on the SLIC algorithm shows promising results, its Soft Boundary F1 score (0.5057) is far from ideal, however, this value is still larger than that of U-Net (0.0759). Additionally, the current results assume perfect superpixel classification, which is an idealized scenario. The next crucial step in this research is to develop effective methods for classifying superpixels, generated by SLIC algorithm as wound or non-wound regions, which would be capable of learning on a small amount of data and/or generalize well on a variety of data sets.

Discussion of research results. While being not a last-year algorithm, SLIC continues to grow its usage in a combination with different other methods in the field of medical image analysis.

For example, authors of [11] utilize SLIC algorithm for breast cancer classification using ultrasound images. They perform RoI (Region of Interest) extraction using SLIC algorithm with the consequential image feature extraction and classification with the Support Vector Machine machine learning method. The authors state that the accuracy of the segmentation is 88 % and sensitivity of the method up to 92.05 % – which are impressive results.

The other research related to breast cancer diagnostics using Image analysis is presented by Mathias Öttl and others [25]. The authors of [8] in this research utilize SLIC algorithm to improve data annotation of HER2 protein (specific protein related to breast cancer). Authors investigate different variants of the SLIC algorithm and achieve the best Boundary F1 Score (with a margin of 15 pixels) of 0.46 when combining domain-adapted SLIC algorithm with hierarchical clustering, which is an intriguing result, as it is close to the results of our research, – we achieve Soft Boundary F1 Score – 0.50 for the task of diverse wound segmentation.

However, SLIC algorithm is utilized not only in the field of medical imaging but also in other complex mediums. The authors of [16] perform analysis of complex geological structures – sandstone thin section images. They improve the SLIC algorithm, providing it with the features extracted by Gabor filters and Local Binary Patterns. This results in the semi-supervised method that achieves good segmentation quality and various rock-type classification accuracy of 96.3 % on the test dataset.

The other medium that utilizes SLIC algorithm is the aerospace industry. The authors of [13] solve the problem of fatigue image identification of turbine blades and overall microscopic crack and damage detection in aero-engine components. They use SLIC algorithm to extract image boundaries, a UNet feature migration model to obtain multidimensional superpixel features, and a Random Forest classifier to predict and update image segmentation results and achieve the best mean IoU metric of 0.92.

SLIC algorithm in a combination with deep learning models is used inf [20], where SLIC algorithm is used for the UAV (unmanned aerial vehicles) image processing, segmenting forest regions on the UAV images. After segmentation, the corresponding segments are classified with the ResNet-50, fine-tuned based on the available forest/background data, achieving 89 % accuracy. While also a combination of SLIC algorithm and deep learning, authors of [9] modify the algorithm so it uses features extracted by deep learning to cluster the image into superpixels.

The SLIC algorithm is utilized across various domains where complex object segmentation is required, including medical imaging, geological analysis, aerospace industries, and UAV image processing. Despite its versatility, none of the reviewed studies provide clear guidelines for determining optimal of the SLIC algorithm hyperparameters for datasets with different characteristics from the ones used in the research. This highlights the necessity for a methodology introduced in this research to adapt SLIC algorithm for diverse tasks and datasets by effectively tuning its hyperparameters to achieve reliable segmentation results.

So, based on the work performed, the following scientific novelty and practical significance of the research results can be formulated.

The scientific novelty of the obtained research results – for the first time, a methodology has been developed for evaluating the effectiveness of using the SLIC algorithm in complex image segmentation tasks, incorporating the introduction of the Soft Boundary F1 metric and the optimization of hyperparameters through Bayesian methods, enabling improved accuracy in identifying objects on the image and calculating their areas.

The practical significance of the research results – the study's results may be applied to solve practical tasks of analyzing and processing complex images, particularly in medical applications, thus enhancing the efficiency and accuracy of doctors' diagnostic decision-making.

Conclusions / Висновки

In conclusion, this study successfully developed and implemented a comprehensive methodology for assessing the efficacy of SLIC algorithm in complex object segmentation task, on the example of wound segmentation, with the accent on boundary detection accuracy and determining algorithms optimal hyperparameters across various object types and two image resolutions. Our research, which analyzed 3696 wound images from the FUSeg and WoundSeg datasets, provided useful insights into the performance and potential of the SLIC algorithm for wound segmentation tasks. Based on the results of the research the following main conclusions can be drawn.

1. The SLIC algorithm demonstrates consistent performance across different implementations (cuda-slic, fast-slic, and skimage), with Dice scores around 0.84 and Soft Boundary F1 scores around 0.55. This indicates that the SLIC algorithm can effectively segment complex object areas, such as wounds, but has room for improvement in precise boundary detection.
2. Based on the conducted optimization we conclude that the optimal number of segments for SLIC algorithm, for the given task, can be defined relative to the spatial dimensions of the image to be segmented, experiments have shown that the most effective value can be defined as *max_dim* *2, where *max_dim* is the maximum size of the input image. This provides a valuable insight on SLIC setup in the given data domain – maximal number of segments can be defined relatively to the input image size, eliminating the need to select this parameter for each different input image size.
3. Based on the conducted optimization, the optimal compactness hyperparameter value, for the given task, lies in the range of [8, 20] with no great difference between the values in the range. This allows to use the defined value while integrating SLIC algorithm into compound image analysis methods on similar data, without the need for additional hyperparameter selection.
4. Execution speed evaluation shows that cuda-slic implementation of the SLIC algorithm appeared the most efficient, offering a balance of speed ($0.0303^{\pm 0.0116}$ seconds per image) and stability in execution time, which shows that among the others tested (skimage and fast-slic implementations) cuda-slic is the optimal choice, especially in a scenario where SLIC algorithm is combined with the deep learning models which utilize GPU, as cuda-slic implementation utilizes GPU as well.
5. The SLIC algorithm showed stable performance throughout different wound types, maintaining consistent Dice scores. However, the Soft Boundary F1 scores show variations in boundary detection accuracy between different wound types, particularly between the FUSeg (foot ulcers) and WoundSeg (diverse wound types) datasets.
6. Comparison of UNet and SLIC algorithm (in as setting of idealized superpixel classification) revealed, that compared to the UNet model pre-trained on foot ulcer images, the SLIC algorithm demonstrated better generalization capability across diverse wound types both in terms of overall segmentation quality and boundary detection accuracy. On the WoundSeg dataset, the SLIC algorithm achieved a Dice score of $0.8416^{\pm 0.0936}$, significantly outperforming the UNet model ($0.1181^{\pm 0.1947}$); Soft Boundary F1 Score $0.5057^{\pm 0.0653}$ for SLIC algorithm and $0.0759^{\pm 0.1223}$ for UNet, which indicates that superpixels, generated by the SLIC algorithm are able to detect complex object boundaries in a diverse data better, than the UNet model.

Based on the experiment conclusions, future research is to be focused on developing a hybrid approach that combines the stxrengths of the SLIC algorithm as a preprocessing step with deep learning techniques for superpixel classification in a way that emphasizes deep learning models' ability to learn on a small dataset and generalize on diverse data.

References

1. Algy. (n.d.). GitHub – Algy/fast-slic: 20 x Real-time superpixel SLIC Implementation with CPU. GitHub. URL: https://github.com/Algy/fast-slic? tab=readme-ov-file
2. Aydin, O. U., Taha, A. A., Hilbert, A., Khalil, A. A., Galinovic, I., Fiebach, J. B., Frey, D., & Madai, V. I. (2021). On the usage of average Hausdorff distance for segmentation performance assessment: hidden error when used for ranking. *European Radiology Experimental*, *5*(1). https://doi.org/10.1186/s41747-020-00200-2
3. Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). *Algorithms for Hyper-Parameter Optimization*. URL: https://www.semanticscholar.org/paper/Algorithms-for-Hyper-Parameter-Optimization-Bergstra-Bardenet/03911c85305d42aa2eeb02be82ef6fb7da644dd0
4. Blanco, G., Traina, A. J., Traina, C., Jr, Azevedo-Marques, P. M., Jorge, A. E., De Oliveira, D., & Bedo, M. V. (2019). A superpixel-driven deep learning approach for the analysis of dermatological wounds. *Computer Methods and Programs in Biomedicine*, *183*, article ID 105079. https://doi.org/10.1016/j.cmpb.2019.105079
5. Carrión, H., Jafari, M., Yang, H., Isseroff, R. R., Rolandi, M., Gomez, M., & Norouzi, N. (2022). HEALNET – Self-supervised Acute Wound Heal-Stage Classification. In *Lecture notes in computer science* (pp. 446–455). https://doi.org/10.1007/978-3-031-21014-3_46
6. Celaya, A., Riviere, B., & Fuentes, D. (2023, February 8). *A Generalized Surface Loss for Reducing the Hausdorff Distance in Medical Imaging Segmentation*. https://doi.org/10.48550/arXiv.2302.03868
7. Chairat, S., Chaichulee, S., Dissaneewate, T., Wangkulangkul, P., & Kongpanichakul, L. (2023). AI-Assisted Assessment of Wound Tissue with Automatic Color and Measurement Calibration on Images Taken with a Smartphone. *Healthcare*, *11*(2), 273 p. https://doi.org/10.3390/healthcare11020273
8. Chang, C. W., Christian, M., Chang, D. H., Lai, F., Liu, T. J., Chen, Y. S., & Chen, W. J. (2022). Deep learning approach based on superpixel segmentation assisted labeling for automatic pressure ulcer diagnosis. *PLoS ONE*, *17*(2), article ID e0264139. https://doi.org/10.1371/journal.pone.0264139
9. Chang, C.-S., Ding, J.-J., & Lin, H.-S. (2019). *Learning based SLIC superpixel generation and image segmentation* (Vols. 1–2). https://doi.org/10.1109/ISPACS48206.2019.8986326
10. *Comparing images using the Hausdorff distance*. (1993, September 1). IEEE Journals & Magazine | IEEE Xplore. URL: https://ieeexplore.ieee.org/document/232073
11. Fang, Z., Zhang, W., & Ma, H. (2019, April 25). *Breast Cancer Classification with Ultrasound Images Based on SLIC*. arXiv.org. https://doi.org/10.48550/arXiv.1904.11322
12. Farmaha, I., Banaś, M., Savchyn, V., Lukashchuk, B., & Farmaha, T. (2019). Wound image segmentation using clustering based algorithms. *New Trends in Production Engineering*, *2*(1), 570–578. https://doi.org/10.2478/ntpe-2019-0062
13. Fei, C., Wen, J., Han, L., Huang, B., & Yan, C. (2022). Optimizable Image Segmentation Method with Superpixels and Feature Migration for Aerospace Structures. *Aerospace*, *9*(8), article ID 465. https://doi.org/10.3390/aerospace9080465
14. Furizal, F., Mawarni, S. S., Akbar, S. A., Yudhana, A., & Kusno, M. (2023). Analysis of the Influence of Number of Segments on Similarity Level in Wound Image Segmentation Using K-Means Clustering Algorithm. *Control Systems and Optimization Letters*, *1*(3), 132–138. https://doi.org/10.59247/csol.v1i3.33
15. Getto, R., Kuijper, A., & Von Landesberger, T. (2015). Extended surface distance for local evaluation of 3D medical image segmentations. *The Visual Computer*, *31*(6–8), 989–999. https://doi.org/10.1007/s00371-015-1113-z
16. Han, Y., & Liu, Y. (2024). Intelligent Classification and Segmentation of Sandstone Thin Section Image Using a Semi-Supervised Framework and GL-SLIC. *Minerals*, *14*(8), article ID 799. https://doi.org/10.3390/min14080799
17. Ioffe, S., & Szegedy, C. (2015, February 11). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. https://doi.org/10.48550/arXiv.1502.03167
18. Liu, Hongjin, Wang, Niannian, Fang, Hongyuan, Yu, Xiang, & Du, Weiyi. (2025, February). Identification of the number of leaks in water supply pipes based on wavelet scattering network and Bi-LSTM model with Bayesian optimization. *Measurement*, Vol. 243, article ID 116348. https://doi.org/10.1016/j.measurement.2024.116348
19. Mahboubi, Negareh, Xie, Junyao, & Huang, Biao. (2025, March). Point-by-point transfer learning for Bayesian optimization: An accelerated search strategy. *Computers & Chemical Engineering*, Vol. 194, article ID 108952. https://doi.org/10.1016/j.compchemeng.2024.108952
20. Martins, J., Marcato Junior, J., Menezes, G., Pistori, H., Diego, S., & Gonçalves, W. (2019). Image segmentation and classification with SLIC superpixel and convolutional neural network in forest context. https://doi.org/10.1109/IGARSS.2019.8898969
21. Mitura, J., Chrapko, B. E., & Bachanek-Mitura, O. (2024). MedVoxelHD: Improved CUDA-accelerated morphological Hausdorff distances in medical image analysis. *SoftwareX*, *26*, article ID 101744. https://doi.org/10.1016/j.softx.2024.101744
22. Müller, D., Soto-Rey, I., & Kramer, F. (2022). Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes*, *15*(1). https://doi.org/10.1186/s13104-022-06096-y
23. Niri, R., Douzi, H., Lucas, Y., & Treuillet, S. (2021). A Superpixel-Wise Fully Convolutional Neural Network Approach for Diabetic Foot Ulcer Tissue Classification. In *Lecture notes in computer science* (pp. 308–320). https://doi.org/10.1007/978-3-030-68763-2_23
24. Oota, S. R., Rowtula, V., Mohammed, S., Liu, M., & Gupta, M. (2023). WSNet: Towards An Effective Method for Wound Image Segmentation. *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. https://doi.org/10.1109/wacv56688.2023.00325
25. Öttl, M., Mönius, J., Marzahl, C., Rübner, M., Geppert, C., I., Hartmann, A., Beckmann, M. W., Fasching, P., Maier, A., Erber, R., & Breininger, K. (2022, January 19). *Superpixel Pre-Segmentation of HER2 Slides for Efficient Annotation*. https://doi.org/10.48550/arXiv.2201.07572
26. Ramachandram, D., Ramirez-GarciaLuna, J. L., Fraser, R. D. J., Martinez-Jiménez, M. A., Arriaga-Caballero, J. E., & Allport, J. (2022). Fully Automated Wound Tissue Segmentation Using Deep Learning on Mobile Devices: Cohort Study. *JMIR Mhealth and Uhealth*, *10*(4), article ID e36977. https://doi.org/10.2196/36977
27. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Lecture notes in computer science* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28
28. rosalindfranklininstitute. (n.d.). *GitHub – rosalindfranklininstitute/cuda-slic: A CUDA implementation of the SLIC Superpixel algorithm*. GitHub. URL: https://github.com/rosalindfranklininstitute/cuda-slic
29. *SLIC Superpixels Compared to State-of-the-Art Superpixel Methods*. (2012, November 1). IEEE Journals & Magazine | IEEE Xplore. URL: https://ieeexplore.ieee.org/document/6205760
30. Van Der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: image processing in Python. *PeerJ*, *2*, article ID e453. https://doi.org/10.7717/peerj.453
31. Wang, C., Mahbod, A., Ellinger, I., Galdran, A., Gopalakrishnan, S., Niezgoda, J., & Yu, Z. (2024). FUSeg: The Foot Ulcer Segmentation Challenge. *Information*, *15*(3), 140 p. https://doi.org/10.3390/info15030140
32. Yeghiazaryan, V., & Voiculescu, I. (2018). Family of boundary overlap metrics for the evaluation of medical image segmentation. *Journal of Medical Imaging*, *5*(01), article ID 1. https://doi.org/10.1117/1.jmi.5.1.015006

Б. С. Лукащук[1], Ю. В. Шабатура[2]
[1] *Національний Лісотехнічний Університет України, м. Львів, Україна*
[2] *Національна академія сухопутних військ імені гетьмана Петра Сагайдачного, м. Львів, Україна*

МЕТОДИКА ОЦІНЮВАННЯ ТОЧНОСТІ ІДЕНТИФІКАЦІЇ КОНТУРІВ ЗОБРАЖЕНЬ СКЛАДНИХ ОБ'ЄКТІВ, ВИДІЛЕНИХ У ПРОЦЕСІ СЕГМЕНТАЦІЇ АЛГОРИТМОМ SLIC

Досліджено особливості застосування алгоритму SLIC (англ. *Simple Linear Iterative Clustering*) для сегментації складних цифрових зображень, зокрема – зображень ушкоджень тіла людини. Встановлено, що цей алгоритм на цифрових зображеннях дає змогу ефективно оцінювати площу ураження тіла та виявляти межі ран. Проаналізовано відсутність кількісних метрик ефективності використання алгоритму в завданнях високоточного оцінювання ушкоджень тіла людини, що стало підставою для розроблення комплексної методології оцінювання його застосування для зображень з різною роздільною здатністю. Розроблено комбінований набір даних, який містить 3696 зображень із FUSeg (англ. *Foot Ulcer Segmentation Challenge*) та WoundSeg наборів даних. Проведено оптимізацію гіперпараметрів алгоритму SLIC-оптимізації за методом Байєса задля максимізації метрики Soft Boundary F1. Оптимізовані гіперпараметри містять кількість сегментів та компактність, які впливають на точність поділу об'єктів і коректність визначення меж. З'ясовано, що алгоритм SLIC демонструє стабільну ефективність, досягаючи Dice Score близько 0,84 та Soft Boundary F1 Score близько 0,55. Встановлено, що оптимальна кількість сегментів для алгоритму SLIC пропорційна просторовим розмірам вхідного зображення, а найефективнішим значенням є максимальний розмір зображення *2. Запропоновано модифікацію метрики Boundary F1 – Soft Boundary F1 Score для використання у процесі оптимізації гіперпараметрів алгоритму SLIC. Попередні результати свідчать, що вона потенційно може забезпечувати більш плавний ландшафт оптимізації порівняно зі стандартною Boundary F1 та є перспективною для задач сегментації складних об'єктів. Наведено ідеалізований підхід до сегментації складних об'єктів на прикладі сегментації пошкоджень на шкірі людини на цифрових зображеннях на підставі алгоритму SLIC, де суперпікселі комбінуються з еталонною маскою, встановлюючи верхню межу ефективності застосування алгоритму. Здійснено порівняння з сегментацією за допомогою моделі UNet, попередньо натренованої на наборі даних FUSeg, і встановлено, що ідеалізований підхід до сегментації з використанням алгоритму SLIC досяг значення Dice Score 0,84 на наборі даних WoundSeg, перевершивши UNet (0,12 Dice Score). Зроблено висновок, що підходи на підставі SLIC є ефективними для сегментації зображень складних об'єктів за обмеженої кількості навчальних даних. Наголошено на потребі проведення подальших досліджень, спрямованих на вдосконалення підходів до класифікації суперпікселів відомими методами, що базуватимуться на результатах генерації суперпікселів, наведених у цій роботі.

Ключові слова: оброблення зображень; семантична сегментація; алгоритм; медична візуалізація; оптимізація Байєса; глибоке навчання.