UDC 004.9:338

N. Volkova, PhD, Assoc. Prof.,
M. Shvandt
Odessa Polytechnic National University, 1 Shevchenko Ave., Odesa, Ukraine, 65044; e-mail: volkova.n.p@op.edu.ua

SEGMENTATION-BASED APPROACH FOR OBJECT DETECTION

Н. Волкова, М. Швандт. **Підхід на основі сегментації для виявлення об'єктів.** У цьому дослідженні запропоновано підхід виявлення об'єктів на основі сегментації, розроблений для аналізу поведінки водних організмів у контрольованих лабораторних умовах. Робота спрямована на вирішення труднощів, що виникають під час тривалих відеозаписів бичків у закритих акваріумах - зокрема, нестабільність фону, перенесення осаду та часткові перекриття особин, які ускладнюють застосування традиційних методів трекінгу. Для цього було запропоновано підхід на основі вдосконаленого в роботі методу сегментації SLIC Superpixel. Базовий метод SLIC було вдосконалено шляхом включення багатошарових контрастних ознак і перевірок однорідності пікселів у межах локального оточення. Запропонований підхід містить наступні етапи: попередня обробка, сегментація, кластеризація, постобробка. Етап попередньої обробки включає двосторонню та медіанну фільтрацію, нормалізацію контрасту й яскравості та за потреби – масштабування зображення для покращення чіткості. Подальше віднімання фону і порогове значення пікселів у межах сегментованих областей дозволяють усунути хибні спрацьовування, спричинені візуальними артефактами та оклюзіями. На етапі кластеризації застосовується уточнена метрика відстані для оцінки узгодженості пікселів у багатошаровому просторі ознак (LAB, нормалізоване зображення у відтінках сірого, результати субтракції), що підвищує точність сегментації. На етапі постобробки фрагментовані області об'єктів об'єднуються для покращення просторової когерентності. Експериментальна перевірка на відеокадрах з бичками продемонструвала підвищення точності виявлення об'єктів понад 6% у порівнянні з підходом на основі базового методу SLIC. Модульність і простота запропонованого підходу дозволяють легко розширювати його застосування на інші біологічні об'єкти – зокрема, на поведінковий аналіз гризунів - без потреби використання глибокого навчання чи ресурсоємних архітектур, що робить його придатним для задач етології, нейронауки та прецизійної аквакультури. Подальші дослідження будуть присвячені реалізації підходу в режимі реального часу та розширеному аналізу траєкторій.
Ключові слова: підхід, обробка зображень, обробка відео, сегментація зображень, відстеження об'єктів, виявлення об'єктів

N. Volkova, M. Shvandt. **Segmentation-based approach for object detection.** This study proposes a segmentation-based approach for object detection, developed for analyzing aquatic behavior in controlled laboratory environments. The research focuses on overcoming detection challenges in long-term video recordings of bullheads housed in enclosed aquariums, where sediment drift, background instability, and partial occlusions often confound traditional tracking techniques. To address these issues, an approach based on the improved SLIC Superpixel segmentation method was proposed. The basic SLIC method was modified to incorporate multi-layer contrast features and neighborhood-based pixel uniformity checks. The proposed approach includes the following stages: preprocessing, segmentation, clustering, and post-processing. The preprocessing stage includes bilateral and median filtering, contrast and brightness normalization, and optional image upscaling to improve clarity. Subsequent background subtraction and context-aware thresholding within segmented regions help eliminate false positives caused by floating debris and occluded contours. At the clustering stage, a refined distance metric is introduced to evaluate pixel coherence in a multilayered feature space, which include LAB components, subtraction results, and histogram-equalized grayscale representations, improving segmentation accuracy. Additionally, at the post-processing stage fragmented object blobs are merged to enhance spatial continuity. Empirical validation was conducted on a dataset of bullhead video frames recorded under realistic aquatic conditions. The approach based on the improved SLIC Superpixel segmentation method demonstrated an increase in object detection accuracy of more than 6% compared to the approach based on the basic SLIC method. The modularity and simplicity of the proposed approach allow it to be easily extended to other biological objects – in particular, for the behavioral analysis of rodents – without relying on deep neural networks or computationally intensive frameworks, making it suitable for tasks in ethology, neuroscience, and precision aquaculture. Further research will be devoted to implementing the approach in real-time and advanced trajectory analysis.
Keywords: approach, image processing, video processing, image segmentation, object tracking, object detection

Introduction

In recent decades, the study of animal behavior has become increasingly vital across disciplines such as neuroscience, ethology, environmental monitoring, and agricultural science. Understanding how animals respond to environmental conditions not only reveals critical insights into their welfare and health but also serves as a valuable indicator of broader ecological changes [1, 2]. Concurrently, modern computer technologies, particularly computer vision, have unlocked new possibilities for quantifying and analyzing these behaviors with high precision, reproducibility, and scalability [3].

Traditionally, behavioral studies relied on manual annotation or rudimentary sensors, approaches that were inherently time-consuming, limited in resolution, and often prone to observer bias [4]. In contrast, advances in imaging hardware, motion analysis, and object tracking algorithms now enable continuous, non-invasive monitoring of animal behavior across a wide variety of settings. While the role of neural networks and deep learning is certainly influential in this domain, many high-impact

DOI: 10.15276/opu.1.71.2025.17

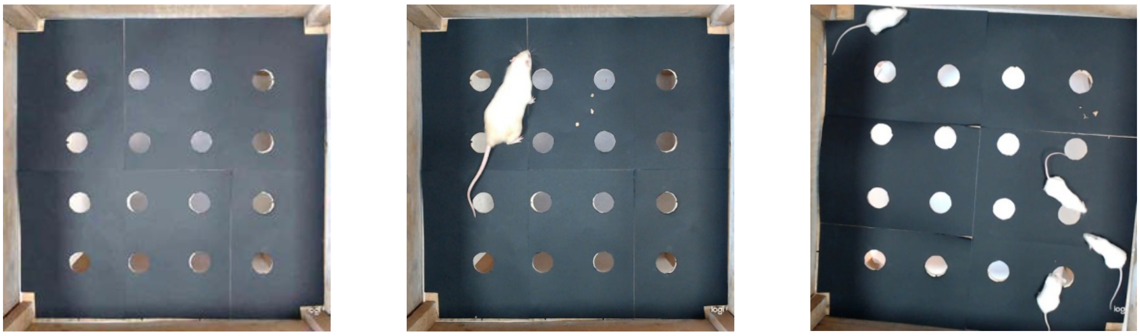
ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

behavioral studies successfully apply well-established computer vision techniques such as background subtraction, blob detection, and trajectory clustering without relying heavily on opaque models [5, 6].

The rising popularity of automated behavioral analysis stems from a convergence of technological readiness and scientific necessity. Firstly, global challenges such as climate change, disease emergence, and food security demand real-time tools to assess animal responses to environmental pressures [7]. Secondly, behavior often reflects underlying physiological or cognitive states more rapidly than biochemical markers can detect. As such, automated behavioral metrics provide a non- invasive view into animal welfare and can even serve as early-warning systems for stress or illness [8, 9].

Analysis of recent publications and formulation of the problem

As stated previously, mice and fish are widely used in ecotoxicology and ethology research. A part of such research is based on animal behavior study in an enclosed conditions as an early stage of study. An example of such enclosed conditions for mice/rats behavior study is a box with holes in it. Above the box a hard-mounted camera is installed and test subjects are being filmed during a considerable period of time (from 30 minutes to several hours). The observations include counting the movements of animals between holes and their attempts to look into them as a sign of territory discovering (Fig. 1).



a *b* *c*

Fig. 1. Mice/rats behaviour study: *a* – test environment (box), *b* – screenshot of a video with a lab rat, *c* – screenshot of a video with lab mice

In our research, special attention is given to the behavior of fish, especially bullheads. The enclosed test environment is represented by a square aquarium filled with real sea water to simulate the real environmental conditions. The aquarium is also fed by oxygen that is pumped through the pipes. A hard-mounted camera is installed above the aquarium and is recording the bullheads during some period of time (it varies from several hours to 1 day). The filming is split into videos of 30 minutes long to make further analysis easier (Fig. 2).

a *b* *c*

Fig. 2. Fish (bullheads) behaviour study: *a*, *b* – screenshots from videos with bullheads; *c* – test environment (aquarium)

The aquarium can hold up to 10 test subjects during the experiment. The behaviour patterns that currently are of greatest interest to the researchers among other ones are:

– The general number of movements (position changes) for each subject;

– The general number of movements for complete test group;

– Number of attacks/conflicts between 2 subjects (a movement of subject A towards subject B that results in subject B escaping in other direction);

– Keeping each object's ID during the complete recording session. This is particularly important in case of placement of 2 or 3 different kinds of bullheads into the aquarium.

Currently the lab personnel is analysing the video sequences manually which is a very time consuming task and may lack accuracy as all patterns are being recognized 'by hand' and roughly. Thus an method that could automatically detect, track and analyse the subject movements is in order.

The task execution is also complicated by two aspects. Firstly, despite the fact that subject's position cannot change too drastically between 2 or 3 consequent frames, its movement direction change can be quite radical. And secondly, while the test environment is stable in general, the aquarium background remains being subject to noticeable changes: before and during test sessions food is being dispensed into the aquarium. Also the results of bullheads' life activities tend to accumulate on the aquarium bottom. Both these sediments tend to migrate in the form of clots under the influence of water flows from moving fish and oxygen fed from the tubes. As they can be of the similar colour as some of bullheads and get quite big in area, it makes sometimes difficult to distinguish the fish position if it is located above such clot. It also eliminates the ability to use such straightforward approaches as tracking by colour. All these points are important to be taken into consideration during the development of the general algorithm.

An importance to study the aspects of life and behavior of fish has lead to a number of studies related to attempts to create methods for fish classification, detection or tracking under specific conditions. In [10] a set of review of fish tracking technologies tailored for aquaculture, aiming to solve the limitations of manual monitoring and delayed intervention is presented. The authors categorize tracking systems into acoustic telemetry, RFID tagging, and computer vision methods. Of particular focus is the integration of deep learning models like YOLOv5 and Faster R-CNN with underwater video feeds. The study discusses architectural modifications for low-light video processing, including real-time detection of fish biomass and abnormal behavior. Challenges such as occlusion, fish overlap, and image noise are tackled through multi-sensor fusion, combining sonar data with visual inputs. The authors propose a hybrid vision-acoustic system that allows scalable monitoring in offshore environments. Benchmarked models show improved detection accuracies (>90%) under murky conditions. The paper also suggests future directions: drone-assisted fish tracking, federated learning-based behavior modeling, and ethical implications of using AI in live-animal monitoring.

Targeting fish movement studies in ecology, another research [11] presents a modular open-source pipeline for automatic detection and tracking. Researchers address the problem of scaling fish motion tracking using cost-efficient cameras in noisy marine habitats. The pipeline begins with motion-based frame differencing and blob detection, followed by morphological operations to isolate fish contours. A Kalman filter-based tracking module enables identification of trajectories over multiple frames. Validation was performed on field-collected videos from estuarine ecosystems, where traditional methods fail due to turbidity and background clutter. The system was capable of detecting small fish with minimal latency. The authors benchmark detection against manual annotations and report precision scores exceeding 0.89. Key contributions include modularity, real-time adaptability, and compatibility with ecological databases. Suggested improvements involve using deep learning models to replace handcrafted motion filters and incorporating stereo cameras for depth-aware tracking in complex fish schools.

Another study [12] addresses the issue of free-moving fish detection in unconstrained underwater video environments. A hybrid architecture is developed combining optical flow and Gaussian Mixture Models (GMM) with convolutional neural networks, specifically YOLO. The system handles occlusions and background noise by applying motion segmentation before feeding candidates to the detection model. A notable innovation is temporal modeling of fish trajectories, allowing detection across frames even when fish visibility drops. The pipeline achieves significant gains in recall (over 85%) under challenging conditions such as poor lighting and fish clustering. Field validation was conducted using videos from baited underwater stations and natural reef footage. Results show that motion-based pre-filtering improves deep model precision and reduces false positives. The authors propose scaling the approach for behavioral studies, integrating fish length estimation and depth mapping. They emphasize open-source implementation and hardware-efficiency, making the tool deployable on edge devices for marine conservation tasks.

 ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

Focusing on underwater species identification, the following research [13] introduces a ResNet architecture enhanced by cross-layer pooling and transfer learning. The problem addressed is poor classification performance in noisy, low-resolution underwater imagery. A tailored CNN with residual blocks extracts features robustly across varied aquatic backgrounds. Transfer learning is employed to fine-tune pre-trained models from ImageNet on fish datasets with limited annotations. Cross-layer pooling refines feature selection from deeper layers, mitigating feature dilution. The authors implement class-balanced focal loss to address dataset imbalance between frequent and rare species. Experiments on in site footage achieved classification accuracies upwards of 94%. Key technical innovations include using temporal context from consecutive frames and classifier fusion. The system is positioned as an assistive tool for fishery surveys and ecological monitoring where automated annotation is impractical.

This comprehensive survey [14] studies image-based fish classification methods and evaluates their strengths across various aquatic datasets. A comparison of traditional classifiers, such as SVM, ANN, KNN, with deep learning models such as AlexNet and ResNet is conducted. Image preprocessing techniques including grayscale conversion, denoising (Gaussian and median filters), and segmentation (Sobel, GrabCut) are discussed in terms of their effectiveness on underwater images. A notable contribution is their emphasis on dataset quality and annotation, highlighting common issues such as occlusion, shadow artifacts, and camera-induced blur. The paper underscores the relevance of shape, texture, and color features, stressing that hybrid descriptors outperform single-feature approaches. Performance benchmarking is presented using multiple datasets, and ensemble models combining SVM and PCA were noted for reliability. The study concludes that while CNNs outperform others on large datasets, smaller setups benefit from simpler classifiers with robust preprocessing.

The research [15] focuses on detection of Epizootic Ulcerative Syndrome (EUS), a serious fish disease, using image classification powered by MobileNet and transfer learning. The authors curated a dataset of infected fish images captured under field conditions, with visible lesions as diagnostic markers. Preprocessing techniques like cropping, normalization, and CLAHE were applied to improve contrast and eliminate lighting inconsistencies. The researchers opted for MobileNet due to its lightweight nature, ideal for deployment on portable diagnostic tools. Transfer learning from ImageNet allowed rapid convergence and better generalization on small datasets. The model was trained on annotated samples with softmax classification and tested using accuracy, precision, recall, and F1 score. It achieved over 95% classification accuracy, proving MobileNet's efficiency in real-world veterinary applications. Moreover, the paper explores on-device deployment scenarios, aiming to assist aquaculture workers in remote areas. This study bridges ML classification with field-level disease intervention and illustrates how compact neural networks can support mobile bio-surveillance efforts.

In the conference article [16] a machine learning system for fish recognition in offshore aquaculture cages is presented. A novel combination of vertical multi-camera installations and image enhancement techniques is proposed to handle low-light, high-turbidity underwater footage. Retinex-based preprocessing and contour-based filtering are used to highlight fish silhouettes against noisy backgrounds. YOLOv4 is customized to account for vertical swimming motion and overlapping fish instances, allowing multi-object detection with bounding boxes that adapt to camera tilt. Data labeling was semi-automated using synthetic fish images generated via GANs to augment the sparse real-world footage. The researchers also propose bounding box smoothing over consecutive frames to prevent flickering predictions, essential in fish counting scenarios. The approach demonstrates robustness, real-time capability, and potential for integration into automated feeding and health monitoring systems.

One more article [17] introduces the Atrous Pyramid GAN Segmentation Network (APGSN), designed to segment fish in complex underwater environments. Traditional segmentation models struggle with occlusions and scale variance, so this architecture leverages dilated convolutions across pyramid layers to capture multi-scale features without losing resolution. The generator synthesizes fish masks while the discriminator enforces spatial realism. Pre-training was performed using synthetic fish masks, and the network was fine-tuned on real aquatic datasets. Compared to U-Net and FCN benchmarks, APGSN achieves superior pixel-level precision and boundary adherence. The model maintains segmentation accuracy across fish types, environmental conditions, and video frames. Experiments involved footage from Chinese offshore aquaculture sites and validated the model's robustness against lighting noise and water distortion. Key outcomes include over 95% Intersection-over-Union scores and low false-positive rates. The study concludes with deployment suggestions for real-time segmentation on aquaculture robots and underwater drones, pushing the model toward practical use.

The research in [18] explores a morphological fish detection pipeline built around unsupervised segmentation using *K*-means clustering. Researchers address the challenge of separating fish from noisy underwater backgrounds where lighting and water turbidity degrade image clarity. CLAHE is used for contrast enhancement, followed by *K*-means to isolate dominant intensity regions. Morphological operators, such as dilation, erosion, and edge tracing, are used to extract fish contours robustly from cluttered scenes. The researchers also implemented heuristic post-processing to remove overlapping false positives. Although relatively simple, the algorithm achieves high accuracy (about 92%) and low latency, making it suitable for deployment on edge devices. The authors suggest integrating this pipeline with motion filters or optical sensors to improve real-time tracking. Its relevance lies in providing a no-deep-learning alternative for small farms lacking computational infrastructure.

In [19] authors propose an image enhancement method for offshore cage inspections under murky water conditions. Using Retinex filtering combined with guided image filtering, the method enhances luminance and reduces scattering artifacts common in deep marine environments. Researchers analyzed the clarity of fish contours pre- and post-processing using entropy and PSNR metrics. Their benchmark dataset includes images from Autonomous Underwater Vehicles (AUVs) in Yellow Sea cage farms. The improved images allow for better edge segmentation and fish density estimation. The study demonstrates a 40% improvement in visibility across sample frames. Also, this approach is hardware-agnostic and does not rely on CNNs or large training sets, making it compatible with inspection robots and drone-based systems. It's positioned as a low-complexity preprocessing stage to support downstream object detection modules.

The research [20] presents a more granular comparative framework built around classifier efficiency and preprocessing methods. The authors examine more than 20 fish classification papers, dissecting how combinations of preprocessing (grayscale, filtering, segmentation) with classifiers (SVM, ANN, RF) affect accuracy. A meta-analysis identifies image quality, occlusion levels, and annotation fidelity as key factors influencing model performance. The survey gives detailed recommendations on choosing classifiers based on dataset size: ANN for large annotated sets, SVM for balanced mid-sized datasets, and ensemble classifiers for noisy underwater scenes. Additionally, the paper calls for the creation of standardized aquatic datasets featuring turbidity annotations and occlusion metadata. The review also highlights shortcomings in hyperparameter tuning and a lack of validation on real-world hardware deployments.

Considering their [21] prior segmentation network (APGSN), Zhou et al. expand the research into mobile deployment. They simulate GAN-generated fish masks on portable GPU rigs to validate APGSN's edge-readiness. Tested in real-time using drone footage, the system processes fish segmentation with a latency of under 0.8 seconds/frame. The team proposes an attention mechanism to refine mask edges further, improving segmentation on partially occluded targets. This iteration emphasizes robustness to motion blur and makes deployment suggestions for cage-integrity inspection, feeding pattern monitoring, and multi-species tagging.

This extensive review [22] spans 20 years of fish classification research, evaluating models from traditional handcrafted feature extractors to modern CNNs and GANs. Li et al. divide the pipeline into acquisition, preprocessing, and classification stages. Acquisition challenges include water distortion and multi-angle tracking; preprocessing techniques include CLAHE, denoising, and segmentation using GrabCut, Otsu's method, and pyramid shifting. For classification, they compare ANN, SVM, AdaBoost, and state-of-the-art deep learning models (AlexNet, VGG, ResNet, YOLO). Their key finding is that composite architectures outperform single classifiers, particularly under occlusion and environmental variability. The review includes performance benchmarks, dataset summaries, and future directions, such as hierarchical partial classifiers, LSTM-based temporal models, and edge-efficient segmentation networks.

In [23] one tests CNN variants on underwater datasets to resolve issues with low-resolution fish imagery. The researchers experiment with encoding-decoding CNNs (U-Net style) and shallow VGG-based classifiers, measuring accuracy on blurred and distorted samples. Preprocessing includes Gaussian blurring, morphological operations, and pyramid mean shifting. Training is performed on a Fish4Knowledge subset with fish masks manually refined. Performance metrics highlight trade-offs between network depth and runtime, with deep models yielding better generalization but slower inference. Their final framework is a hybrid CNN-KNN model which achieved ~98% accuracy in low-light samples and demonstrated reduced overfitting via dropout layers. Suggested future work includes integrating behavioral classification and real-time inference optimization.

ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

Together, these articles paint a comprehensive picture of how artificial intelligence is changing aquaculture and fish ecology research. They address many tasks, ranging from species classification and disease detection to motion tracking and semantic segmentation. But despite their variety and effectiveness, most of these methods are aimed at test conditions that differ from those observed in our case. In particular, they either read data from multiple sources (cameras or additional sensors), or the test conditions are more refined (objects are filmed from closer angles and more object details are visible).

Thus, solving the actual gobies detection and tracking problem will require a specially tailored approach that accounts for background noise, variable aquarium conditions, and the absence of additional sensor data.

The purpose and task of the research

Based on a review of current works, the purpose and task of the research were determined.

The aim of this study is to develop an approach for automatic detection and tracking of gobies in a closed test environment under conditions that complicate the analysis: unstable background, presence of sediments and lack of additional sensor data sources.

To achieve this, the following tasks were set:

– To analyse modern methods of fish detection and tracking based on video data;

– To identify the limitations of existing approaches in the context of a test environment for gobies;

– To develop an object detection approach for analysing the behavior of aquatic organisms in controlled laboratory conditions, capable of providing stable detection and tracking of objects against a noisy and changing background without the use of additional sensors;

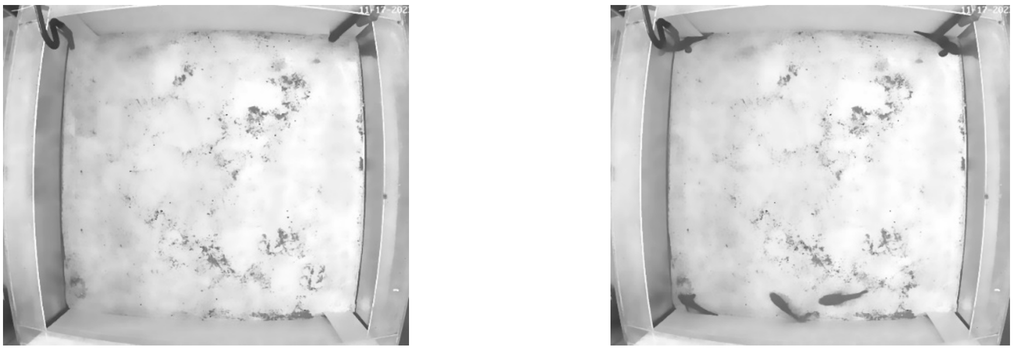
– To conduct experimental verification of the proposed approach on video recordings of aquarium experiments and to evaluate its accuracy.

Presenting main material

Considering the present test conditions of the enclosed environment an image preprocessing algorithm for object detection and tracking has been developed in [24], which is based on the method of image subtraction, in particular, the background [25, 26]. The algorithm proposed in [24] consists of two stages: image preprocessing and object localization.

A modified SLIC image segmentation method has become the basis for the proposed object detection approach. Before performing of image subtraction operation, both each frame and background image a required to be preprocessed to enhance their color characteristics and remove noise. Image preprocessing in the suggested approach includes the following stages:

1. Bilateral Filter for noise removal [27];
2. Brightness and contrast adjustment;
3. Additional use of a median filter for noise removal;
4. Accurate image upscaling to obtain a clearer representation of objects in the frame may be also applied optionally if the original video frame size is too small [28]. Figure 3 shows the results of the application of the proposed preprocessing algorithm to an empty image and a video frame.



a *b*

Fig. 3. Pre-processing results: *a* – empty aquarium image, *b* – a frame of a video with fish

After image color preprocessing object localization is taking place. But before object bounding box calculation, an additional image processing step is required. As mentioned earlier, during the filming of the video sequences the bottom of the aquarium is getting filled with the scattered food material and products of fish life activities. These image parts can cause partial occlusions or objects’ real posi-

```
/* Initialization */
Initialize cluster centers C_k = [l_k, a_k, b_k, x_k, y_k]^T by sampling pixels at regular grid steps S.
Move cluster centers to the lowest gradient position in 3×3 neighborhood.
Set label l(i) = −1 for each pixel i.
Set distance d(i) = ∞ for each pixel i.
repeat:
    /* Assignment */
    for each cluster center C_k do
        for each pixel I in a 2S×2S region around C_k do
            Compute the distance D between C_k and i.
            If D < d(i) then
                set d(i) = D
                set l(i) = k
            end if
        end for
    end for
    /* Update */
    Compute new cluster centers.
    Compute residual error E.
until E ≤ threshold
```

Fig. 4. Basic SLIC Superpixel segmentation algorithm [29, 30]

tion distortions. Since these particles are often of the same colors as the fish itself they cannot be removed by just applying color thresholding. Also mostly due to the fact that those sediments are moving, they tend to get onto the resulting image of background subtraction operation.

Just converting the result from grayscale to binary image and thresholding the blobs by size may not be the most accurate way since some of those separated small white blobs can be parts of fish itself being separated from the subject by the difference in grayscale color during subtraction that is still remaining after the preprocessing step.

To solve this problem, the following solution was proposed. The subtraction result image is being divided into segments and within each segment we calculate the amount of pixels that have value higher than the preset threshold. If this amount is lower than, for instance, a 1/3 of all segment pixels (ratio depends on image size and size of object of interest), then all pixel values within this pixel are set to 0 (black).

The construction of the grid for background image splitting is based on the result of image segmentation method. The SLIC method was chosen as the base segmentation algorithm [29, 30] for its accuracy. This method is aimed to be used with images converted to LAB color space. Its algorithm is shown on Figure 4.

Thus cluster centers for Superpixels are defined as:

$$C_k = [l_k, a_k, b_k, x_k, y_k]^T , \tag{1}$$

where: l_k, a_k, b_k – color component, a x_k, y_k – spatial component. Correspondently the distance metrics are:

– for color component:

$$d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} ; \tag{2}$$

– for spatial component:

$$d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} . \tag{3}$$

The combined metric is:

$$D' = \sqrt{\left(\frac{d_c}{N_c}\right)^2 + \left(\frac{d_s}{N_s}\right)^2} , \tag{4}$$

where N_c and N_s are color coefficient and interval of division into Superpixels.

In our case, the color representation of the image is not critical since the test environment is in general an image of contrasts (white background and dark foreground objects), especially when recorded at night using camera`s IR mode. Thus, it was decided to propose an extension of the basic SLIC [29] for usage of multiple image layers instead of 3. Each new layer will represent the image features in terms of different contrast representations, for instance: the background subtraction image itself, the *L*-layer from LAB color space with color thresholding, or the grayscale frame with histogram equalization [31].

Considering this fact, the cluster centers from (1) are now presented as:

$$C_k = [l_{k,1}, ..., l_{k,m}, x_k, y_k]^T , \; m = \overline{3, n} , \tag{5}$$

where n is the image layer count. Also, thus the d_c distance (2) now is:

ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

$$d_c = \sqrt{\sum_{m=1}^{n} (l_{j,m} - l_i)^2} \, . \quad (6)$$

The space distance (3) is being kept unchanged.

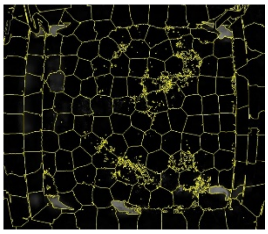
Also another addition was made to the basic method. On step of computing the distance D between C_k and i an additional check was added in order to ignore the pixels that are not surrounded by neighborhood with similar color characteristics. This is aimed on keeping the clusters uniform in order for the subject to be enclosed into a single Segment as much as possible eliminating the interspersion of parts of other segments in it:

– For each candidate pixel a r-neighborhood of pixels is selected, then the within this surrounding neighborhood the average color is being calculated across all n image feature channels;

– Then one compares the average to the pixel's actual color. If the Euclidean color difference (across all channels) exceeds the set neighborhood color threshold, the pixel is skipped in the clustering step.

Thus this approach should help suppressing labeling of edge/noisy pixels that deviate sharply from their local context, reinforcing perceptual consistency, which benefits applications like object boundary detection or segment smoothing and preventing cluster contamination by outlier pixels early in iteration, especially in high-frequency texture zones.

a



b

Fig. 5. Post-processing with the segmentation extension: *a* – result of background subtraction, *b* – segmentation result as grid of segments (Superpixels) applied over the subtraction result

Research results

The test was conducted on a set of 75 images (frames selected from the video sequence). For testing purposes the images (including an image of an empty background) were scaled down by 65% from original size of 1784×1520. For such resolution the initial Superpixel grid was set to contain grid cells of size was 80 pixels. For both basic and modified SLIC the number of clustering iterations is 10, $N_c = 40$, the size of uniformity check neighborhood is 7, the corresponding color threshold is 130. It is worth noticing, that as an temporary extra post-processing step the blob union was applied (small blobs within a small range are being united since they are likely to be parts of a single object separated by, for instance, an oxygen tube).

Also for this particular case in addition to standart LAB layers a result of background subtraction (Fig. 5) was utilized as an additional image feature layer. Tables 1 and 2 show the confusion matrices for basic and modified methods. Here in the confusion matrices, the percentage of pixels that belong to the area that is the object and are segmented as object pixels are denoted TP (true positive). TN (true negative) – the percentage of pixels that belong to the background and are segmented as background pixels. FP (false positive) – the percentage of pixels that belong to the background but are segmented as an object. FN (false negative) – the percentage of pixels that belong to the area that is the object but are segmented as background pixels. In addition the total number of correct detections versus the total number of real objects across the entire image set was calculated. The results are shown on Table 3.

Table 1

Confusion matrix for detections performed with help of basic SLIC method

Ground Truth marks (marked by hand), %	Results obtained with Basic SLIC, %	
	Fish	Background
Fish	83.008	16.992
Background	0.322	99.678

Table 2

Confusion matrix for detections performed with help of modifed SLIC method

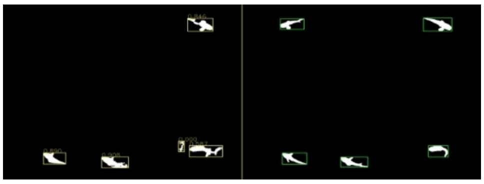
Ground Truth marks (marked by hand), %	Results obtained with modifed SLIC, %	
	Fish	Background
Fish	91.245	8.755
Background	0.509	99.491

Table 3

Correct detections comparison per image set with different segmentation methods as an extension to detection approach

Method used	Correct Detections	Ground Truth Objects (Blobs)
Basic SLIC	316	389
Modified SLIC	340	

The modified SLIC method allowed to increase the number by ~6.8% on this image set (Fig. 6) which is important for correct object identification on each frame.


a

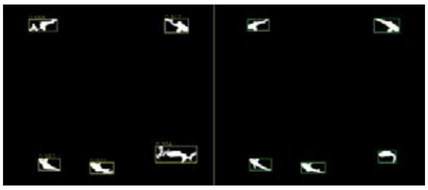

b

Fig. 6. Increased number of correctly detected objects: *a* – usage of basic SLIC, *b* – usage of modifed SLIC

Conclusions

This research proposes a practical and scalable solution for object detection in controlled biological observation environments, specifically tailored for aquatic behavior analysis. The proposed approach, through the integration of an extended SLIC superpixel segmentation algorithm with multilayer contrast evaluation and pixel uniformity checks, addresses challenges such as background clutter, partial occlusions, and sediment interference in long-term fish monitoring. Testing on bullhead video sequences demonstrated improved detection and tracking quality. The method ensures accurate segmentation, more reliable bounding box generation, and enhanced continuity of object tracking. Quantitative behavioral data were obtained, including conflict occurrence and movement metrics. Comparative analysis with baseline segmentation methods confirmed the advantages of the proposed approach in terms of accuracy and robustness under unstable background conditions. The modularity of the approach also enables adaptation to other enclosed environments, such as rodent observation chambers, expanding its potential use in neuroscience and agricultural research. Further research will focus on developing and refining methods and algorithms to improve the detection accuracy of bullheads and the overall quality of image processing for behavioral analysis.

Akwnoledgments. Special thanks for the research assistance and provided test videos and images of lab animals to Faculty of Biology of Odesa I.I. Mechnikov National University.

Література

1. Tinbergen N. On aims and methods of ethology. *Zeitschrift für Tierpsychologie*. 1963. Vol. 20, No. 4. P. 410–433. DOI: https://doi.org/10.1111/j.1439-0310.1963.tb01161.x.
2. Dawkins M. Behavior as a tool in welfare assessment. *Applied Animal Behaviour Science*. 2004. Vol. 86, No. 3-4. P. 227–233. DOI:https://doi.org/10.1016/j.applanim.2004.02.001.
3. Automated image-based tracking and its application in ecology / Dell A., Bender J., Branson K., Couzin I., de Polavieja G., et al. *Trends in Ecology & Evolution*. 2014. Vol. 29, No. 7. P. 417–428. DOI:https://doi.org/10.1016/j.tree.2014.05.004.
4. Anderson D., Perona P. Toward a science of computational ethology. *Neuron*. 2014. Vol. 84, No. 1. P. 18–31. DOI: https://doi.org/10.1016/j.neuron.2014.09.005.
5. Detecting abnormal animal behaviors using optical flow and background subtraction / Yin Z., Xiao L., Ma R., Han Z., Li Y., et al. *Computers and Electronics in Agriculture*. 2020. Vol. 174. Art. 105471. DOI: https://doi.org/10.1016/j.compag.2020.105471.
6. Beyan C., Fisher R. Animal behavior recognition using spatio-temporal features. *Pattern Recognition*. 2018. Vol. 76. P. 12–22. DOI:https://doi.org/10.1016/j.patcog.2017.10.008.
7. Sensor-based analysis of animal behavior / Manteuffel G., Puppe B., Schön P., Bruckmaier R., Janssen D., et al. *Animal*. 2009. Vol. 3, No. 9. P. 1197–1204. DOI:https://doi.org/10.1017/S1751731109004526.
8. Neethirajan S. Recent advances in wearable sensors for animal health management. *Sensors and Biosensors Research*. 2017. Vol. 20. P. 1–11. DOI:https://doi.org/10.1016/j.sbsr.2018.02.004.
9. Understanding fish behavior during typhoon events in real-life underwater environments / Spampinato C., Palazzo S., Boom B., Lin H., Wei J., et al. *Multimedia Tools and Applications*. 2014. Vol. 70, No. 1. P. 199–236. DOI:https://doi.org/10.1007/s11042-012-1101-5.
10. Khan S. U., Ghazali K. H., Khan S. N. Advances in fish tracking technologies for aquaculture: Overcoming challenges and shaping future research. *International Journal of Innovation and Industrial Revolution*. 2025. Vol. 7, No. 20. P. 29–61. DOI:https://doi.org/10.35631/IJIREV.720003.
11. Automatic detection of fish and tracking of movement for ecology / Lopez-Marcano S., Jinks E. L., Buelow C. A., Brown C. J., Wang D., et al. *Ecology and Evolution*. 2021. Vol. 11, No. 12. P. 8254–8263. DOI:https://doi.org/10.1002/ece3.7656.

ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

12. Automatic fish detection in underwater videos by a deep neural network-based hybrid motion learning system / Salman A., Siddiqui S. A., Shafait F., Mian A., Shortis M. R., et al. *ICES Journal of Marine Science*. 2020. Vol. 77, No. 4. P. 1295–1307. DOI:https://doi.org/10.1093/icesjms/fsz025.
13. Automatic fish species classification in underwater videos. *ICES Journal of Marine Science* / Siddiqui S. A., Salman A., Malik M. I., Shafait F., Mian A., et al. 2018. Vol. 75, No. 2. P. 374–389. DOI: https://doi.org/10.1093/icesjms/fsx109.
14. Alsmadi M. K., Almarashdeh I. A survey on fish classification techniques. *Journal of King Saud University – Computer and Information Sciences*. 2022. Vol. 34. P. 1625–1638. DOI: https://doi.org/ 10.1016/j.jksuci.2020.07.005.
15. Rachman F., Akbar M. N. S., Putera E. Fish disease detection of Epizootic Ulcerative Syndrome using deep learning image processing technique. *Proceedings of the 9th International Conference on Fisheries and Aquaculture*. 2023. Vol. 8, No. 1. P. 23–34. DOI: https://doi.org/10.17501/23861282. 2023.8102.
16. Recognition of fish in aqua cage by machine learning with image enhancement / Li Z., Alraie H., Solpico D., Nishida Y., Ishii K., et al. *2024 IEEE/SICE International Symposium on System Integration (SII)*. 2024. P. 637–643. DOI: https://doi.org/10.1109/SII58957.2024.10417229.
17. Atrous Pyramid GAN segmentation network for fish images with high performance / Zhou X., Chen S., Ren Y., Zhang Y., Fu, J., et al. *Electronics*. 2022. Vol. 11, No. 911. P. 1–21. DOI: https://doi.org/10.3390/electronics11060911.
18. Saifullah S., Suryotomo A. P., Yuwono B. Fish detection using morphological approach based on k-means segmentation. *Universitas Pembangunan Nasional Veteran Yogyakarta*. 2021. DOI: https://doi.org/10.48550/arXiv.2101.06352.
19. Gao T., Jin, J., Xu, X. Study on detection image processing method of offshore cage. *Journal of Physics: Conference Series*. 2021. Vol. 1769. Art. 012070. DOI: https://doi.org/10.1088/1742-6596/1769/1/012070.
20. Recent advances of machine vision technology in fish classification / Li D., Wang Q., Li X., Niu M., Wang H., Liu C. *ICES Journal of Marine Science*. 2022. Vol. 79, No. 2. P. 263–284. DOI: https://doi.org/10.1093/icesjms/fsab264.
21. Deep B. V., Dash R. Underwater fish species recognition using deep learning techniques. *Proceedings of the 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)*. 2019. P. 665–669. DOI: https://doi.org/10.1109/SPIN.2019.8711584.
22. Alsmadi M. K., Almarashdeh I. A survey on fish classification techniques. *Journal of King Saud University – Computer and Information Sciences*. 2022. Vol. 34. P. 1625–1638. DOI:https://doi.org/10.1016/j.jksuci.2020.07.005.
23. Temperate fish detection and classification: A deep learning-based approach / Knausgård K. M., Wiklund A., Sørdalen T. K., Halvorsen K. T., Kleiven A. R., Jiao L., et al. *Applied Intelligence*. 2021. P. 1–14. DOI: https://doi.org/10.1007/s10489-021-02527-7.
24. Volkova N., Shvandt M. Image preprocessing algorithm for object detection & tracking. *Інформаційні управляючі системи та технології (ІУСТ-2024)*. 2024. September 23-25. P. 194–198.
25. OpenCV Developers. How to use background subtraction methods. OpenCV Documentation. URL: https://docs.opencv.org/3.4/d1/dc5/tutorial_background_subtraction.html (дата звернення: 15.05.2025).
26. Paris S., Kornprobst P., Tumblin J., Durand F. Bilateral filtering: Theory and applications. *Foundations and Trends® in Computer Graphics and Vision*. 2009. Vol. 4, No. 1. P. 1–73. https://doi.org/10.1561/0600000020.
27. Raid A. M., Khedr W. M., El-dosuky M. A., Aoud M. Image restoration based on morphological operations. *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*. 2014. Vol. 4, No. 3. P. 9–21. https://doi.org/10.5121/ijcseit.2014.4302.
28. Achanta R., Shaji A., Smith K., Lucchi A., Fua P., et al. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2012. Vol. 34, No. 11. P. 2274–2282. https://doi.org/10.1109/TPAMI.2012.120.
29. Zhu A., Mei J., Qiao S., Yan H., Zhu Y., et al. *Superpixel transformers for efficient semantic segmentation*. 2023. arXiv Preprint. https://arxiv.org/abs/2309.16889.
30. Hum Y. C., Lai K. W., Salim M. I. M. Multiobjectives bihistogram equalization for image contrast enhancement. *Complexity*. 2014. Vol. 20, No. 1. P. 22–36. https://doi.org/10.1002/cplx.21499.
31. Joshi P., Escrivá D. M., Godoy V. *OpenCV by example*. Packt Publishing Ltd., 2016.

References

1. Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*(4), 410-433. DOI: https://doi.org/10.1111/j.1439-0310.1963.tb01161.x

2. Dawkins, M. (2004). Behavior as a tool in welfare assessment. *Applied Animal Behaviour Science, 86*(3-4), 227-233. DOI:https://doi.org/10.1016/j.applanim.2004.02.001
3. Dell, A., Bender, J., Branson, K., Couzin, I., & de Polavieja, G., et al. (2014). Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution*, *29*(7), 417-428. DOI:https://doi.org/10.1016/j.tree.2014.05.004
4. Anderson, D., & Perona, P. (2014). Toward a science of computational ethology. *Neuron*, *84*(1), 18-31. DOI: https://doi.org/10.1016/j.neuron.2014.09.005
5. Yin, Z., Xiao, L., Ma, R., Han, Z., Li, Y., et al. (2020). Detecting abnormal animal behaviors using optical flow and background subtraction. *Computers and Electronics in Agriculture, 174,* 105471. DOI: https://doi.org/10.1016/j.compag.2020.105471
6. Beyan, C., & Fisher, R. (2018). Animal behavior recognition using spatio-temporal features. *Pattern Recognition*, *76,* 12-22. DOI:https://doi.org/10.1016/j.patcog.2017.10.008
7. Manteuffel, G., Puppe, B., Schön, P., Bruckmaier, R., Janssen, D., et al. (2009). Sensor-based analysis of animal behavior. *Animal*, *3*(9), 1197-1204. DOI: https://doi.org/10.1017/S1751731109004526
8. Neethirajan, S. (2017). Recent advances in wearable sensors for animal health management. *Sensors and Biosensors Research*, *20,* 1-11. DOI: https://doi.org/10.1016/j.sbsr.2018.02.004
9. Spampinato, C., Palazzo, S., Boom, B., Lin, H., Wei, J., et al. (2014). Understanding fish behavior during typhoon events in real-life underwater environments. *Multimedia Tools and Applications*, *70* (1), 199-236. DOI: https://doi.org/10.1007/s11042-012-1101-5
10. Khan, S. U., Ghazali, K. H., & Khan, S. N. (2025). Advances in fish tracking technologies for aquaculture: Overcoming challenges and shaping future research. *International Journal of Innovation and Industrial Revolution*, *7*(20), 29-61. DOI: https://doi.org/10.35631/IJIREV.720003
11. Lopez-Marcano, S., Jinks, E. L., Buelow, C. A., Brown, C. J., Wang, D., et al. (2021). Automatic detection of fish and tracking of movement for ecology. *Ecology and Evolution*, *11*(12), 8254-8263. DOI:https://doi.org/10.1002/ece3.7656
12. Salman, A., Siddiqui, S. A., Shafait, F., Mian, A., Shortis, M. R., et al. (2020). Automatic fish detection in underwater videos by a deep neural network-based hybrid motion learning system. *ICES Journal of Marine Science*, *77*(4), 1295-1307. DOI:https://doi.org/10.1093/icesjms/fsz025
13. Siddiqui, S. A., Salman, A., Malik, M. I., Shafait, F., Mian, A., et al. (2018). Automatic fish species classification in underwater videos. *ICES Journal of Marine Science*, *75*(2), 374-389. DOI:https://doi.org/10.1093/icesjms/fsx109
14. Alsmadi, M. K., & Almarashdeh, I. (2022). A survey on fish classification techniques. *Journal of King Saud University – Computer and Information Sciences*, *34*, 1625-1638. DOI:https://doi.org/10.1016/j.jksuci.2020.07.005
15. Rachman, F., Akbar, M. N. S., & Putera, E. (2023). Fish disease detection of Epizootic Ulcerative Syndrome using deep learning image processing technique. In *Proceedings of the 9th International Conference on Fisheries and Aquaculture* (Vol. 8, No. 1, pp. 23-34). DOI:https://doi.org/10.17501/23861282.2023.8102
16. Li, Z., Alraie, H., Solpico, D., Nishida, Y., Ishii, K., et al. (2024). Recognition of fish in aqua cage by machine learning with image enhancement. In *2024 IEEE/SICE International Symposium on System Integration (SII)* (pp. 637-643). DOI:https://doi.org/10.1109/SII58957.2024.10417229
17. Zhou, X., Chen, S., Ren, Y., Zhang, Y., Fu, J., et al. (2022). Atrous Pyramid GAN segmentation network for fish images with high performance. *Electronics*, *11*(911), 1-21. DOI:https://doi.org/10.3390/electronics11060911
18. Saifullah, S., Suryotomo, A. P., & Yuwono, B. (2021). Fish detection using morphological approach based on k-means segmentation. *Universitas Pembangunan Nasional Veteran Yogyakarta*. DOI:https://doi.org/10.48550/arXiv.2101.06352
19. Gao, T., Jin, J., & Xu, X. (2021). Study on detection image processing method of offshore cage. *Journal of Physics: Conference Series*, *1769,* 012070. DOI:https://doi.org/10.1088/1742-6596/1769/1/012070
20. Li, D., Wang, Q., Li, X., Niu, M., Wang, H., & Liu, C. (2022). Recent advances of machine vision technology in fish classification. *ICES Journal of Marine Science*, *79*(2), 263-284. DOI:https://doi.org/10.1093/icesjms/fsab264
21. Deep, B. V., & Dash, R. (2019). Underwater fish species recognition using deep learning techniques. In *Proceedings of the 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 665-669). DOI:https://doi.org/10.1109/SPIN.2019.8711584
22. Alsmadi, M. K., & Almarashdeh, I. (2022). A survey on fish classification techniques. *Journal of King Saud University – Computer and Information Sciences*, *34*, 1625-1638. DOI:https://doi.org/10.1016/j.jksuci.2020.07.005

ISSN 2076-2429 (print)
ISSN 2223-3814 (online)

23. Knausgård, K. M., Wiklund, A., Sørdalen, T. K., Halvorsen, K. T., Kleiven, A. R., Jiao, L., et al. (2021). Temperate fish detection and classification: A deep learning-based approach. *Applied Intelligence*, 1-14. DOI:https://doi.org/10.1007/s10489-021-02527-7
24. Volkova, N., & Shvandt, M. (2024). Image preprocessing algorithm for object detection & tracking. *Information management systems and technologies (IMST-2024)*, September 23-25, 194-198.
25. OpenCV Developers. (n.d.). How to use background subtraction methods. *OpenCV Documentation*. Retrieved from https://docs.opencv.org/3.4/d1/dc5/tutorial_background_subtraction.html (Accessed: May 2025)
26. Paris, S., Kornprobst, P., Tumblin, J., & Durand, F. (2009). Bilateral filtering: Theory and applications. *Foundations and Trends® in Computer Graphics and Vision*, *4*(1), 1-73. DOI:https://doi.org/10.1561/0600000020
27. Raid, A. M., Khedr, W. M., El-dosuky, M. A., & Aoud, M. (2014). Image restoration based on morphological operations. *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*, *4*(3), 9-21. DOI:https://doi.org/10.5121/ijcseit.2014.4302
28. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., et al. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(11), 2274-2282. DOI:https://doi.org/10.1109/TPAMI.2012.120
29. Zhu, A., Mei, J., Qiao, S., Yan, H., & Zhu, Y., et al. (2023). Superpixel transformers for efficient semantic segmentation. *arXiv Preprint*. Retrieved from https://arxiv.org/abs/2309.16889
30. Hum, Y. C., Lai, K. W., & Salim, M. I. M. (2014). Multiobjectives bihistogram equalization for image contrast enhancement. *Complexity*, *20*(1), 22-36. DOI:https://doi.org/10.1002/cplx.21499
31. Joshi, P., Escrivá, D. M., & Godoy, V. (2016). *OpenCV by example*. Packt Publishing Ltd.

Волкова Наталія Павлівна; Natalia Volkova, ORCID: https://orcid.org/0000-0003-3175-2179
Швандт Максим Альбертович; Maksym Shvandt, ORCID: https://orcid.org/0000-0002-4580-3961

Received May 07, 2025

Accepted June 12, 2025