RESEARCH ARTICLE

Open-Set Tattoo Semantic Segmentation

ANDERSON BRILHADOR[ID]**, RODRIGO TCHALSKI DA SILVA**[ID]**,
CARLOS ROBERTO MODINEZ-JUNIOR**[ID]**, GABRIEL DE ALMEIDA SPADAFORA**[ID]**,
HEITOR SILVÉRIO LOPES**[ID]**, AND ANDRÉ EUGÊNIO LAZZARETTI**[ID]**, (Member, IEEE)**

Bioinformatics and Computational Intelligence Laboratory - LABIC/CPGEI, Federal University of Technology—Paraná, Curitiba, Paraná 80230-901, Brazil

Corresponding author: Anderson Brilhador (andersonbrilhador@gmail.com)

ABSTRACT Tattoos can serve as an essential source of biometric information for public security, aiding in identifying suspects and victims. In order to automate tattoo classification, tasks like classification require more detailed image content analysis, such as semantic segmentation. However, a dataset with appropriate semantic segmentation annotations is currently lacking. Also, there are countless ways to categorize tattoo classes, and many are not directly categorizable, either because they belong to a specific artistic trait or characterize an object with previously undefined semantics. An effective way to overcome these limitations is to build recognition systems based on open-set assumptions. Nevertheless, state-of-the-art open set approaches are not directly applicable in tattoo semantic segmentation, mainly due to the significant class imbalance (predominant background). To the best of our knowledge, this paper is the first to explore semantic segmentation in closed and open-set scenarios for tattoos. In this sense, this paper presents two key contributions: (i) a novel large-margin loss function and generalized open-set classifier approach and (ii) an open-set tattoo semantic segmentation dataset with a publicly accessible test set, enabling comparisons and future research in this area. The proposed approach outperforms other methods, achieving 0.8013 of AUROC, 0.6318 of Macro F1, 0.4900 of mIoU, and notably 0.2753 of IoU for the unknown class, demonstrating the feasibility of this approach for automatic tattoo analysis. The paper also highlights key limitations and open research areas in this challenging field. Dataset and codes are available at https://github.com/Brilhador/tssd2023.

INDEX TERMS Open-world, open-set, semantic segmentation, large-margin learning, tattoo classification.

I. INTRODUCTION

Tattoos are forms of human expression and are also considered an art. In their almost unique features, tattoos go beyond artistic expressions and can serve as essential sources of biometric information. Consequently, it can be useful in identifying their bearers, mainly for public security [1], [2] because tattoos can be used to identify not only suspects but also victims [3], [4]. In addition, the subject has raised studies on ethical and social issues that may encompass the topic [5].

Compared to other biometrics, tattoos bring a series of characteristics that make them very difficult to recognize. Other biometrics usually have well-defined standards, robust techniques, well-established methods for their treatment and recognition, standardized data capture and storage,

The associate editor coordinating the review of this manuscript and approving it for publication was Joewono Widjaja[ID].

and other factors that help their reliability and robustness. However, tattoos still need to have such characteristics and requirements. Apart from the issues related to processing and using general biometrics, tattoo recognition has a singular complexity because it can be divided into several sub-problems, each equally significant [6].

First, an image can be submitted to detect, locate, and segment (outline or instance) the tattoo contained therein. Subsequently, the image can be classified, de-identified, or re-identified (image-to-image, sketch-to-image, partial, or similar). After preprocessing an image, the best result could be a well-segmented tattoo without any pollution or background. Then, for all classification, de-identification, or re-identification tasks, these images contain only the most essential information to store and, later, process [6].

Nonetheless, tasks that consider the meaning of the content of images, such as classification, may require a more detailed

separation of objects in a tattoo image, called semantic segmentation [7]. At this point, a segmented tattoo could identify and detach each object in the tattoo, after which each could be analyzed separately. For instance, security and biometric recognition systems could benefit from tattoo semantic split at the pixel-level for fine-grained feature extraction by semantic classes, reduce false positives by precisely delineating the boundaries of biometrics features, and allow selective anonymization of regions of an image. Moreover, tattoo biometrics systems may require the identification of multiple semantic classes in multiple areas of the image, which can be performed accurately by tattoo semantic segmentations.

Although this topic has been widely explored in studies related to images and videos in many different areas, semantic segmentation is still underexplored in the context of tattoos. Several works only focus on tattoo classification and detection [6], [8], and only a few researches deal with segmentation [9], [10] without identifying the semantics of the components that compose the tattoos.

One of the reasons that make semantic segmentation difficult is related to the complexity that tattooing can have. As mentioned, tattoos are expressions of art, and their features can be as varied as possible and imaginable. In this way, objects can be positioned very closely, mixed, overlapped, and distorted, and abstract images can also be present, among many other hindering factors [11]. Additionally, the lack of public and comprehensive datasets makes it even more challenging to develop efficient methods for segmentation. Some recent works propose public datasets, such as [12] and [13]. However, in the case of [13], using semi-synthetic images without employing the semantics associated with each class makes it difficult to generalize the proposed – characteristic also observed in [12].

Still, in this context, it is essential to emphasize the complexity and semantic variability of tattoo classes. Tattoo categories can vary significantly, and some are not easily categorized due to their association with specific artistic styles or objects with undefined semantics. The tattoo recognition scenario, especially from a public safety perspective, is also somewhat challenging, especially given the circumstances in which the information is obtained and analyzed. It is not uncommon for tattoo information to be obtained partially, and, therefore, semantic segmentation has great relevance in the identification process, as in the following scenarios: (i) semantic segmentation can be used to create databases with automatic textual annotations, as it is common for a witness or victim to remember or have visual contact with only parts of a tattoo of a wanted person, and, in this way, from the description, it would be possible to identify tattoos with those parts visualized; (ii) semantic segmentation is important in the pre-processing of images in preparation for information recognition processes, such as for partial re-identification of tattoos, wherein an automatic process, the segmented parts can be recovered separately in cases of partial image collections [6].

An effective way to overcome these limitations is to build recognition systems based on dynamic and open-set perception. These systems are designed to handle objects from unknown classes commonly encountered in real-world applications. Open-set recognition has extensively studied the ability to recognize new classes [14]. Open-set semantic segmentation, in turn, is an approach that incorporates open-set perception into semantic segmentation. The main difference with closed-set semantic segmentation is that open-set semantic segmentation must correctly classify samples belonging to known classes while rejecting those belonging to unknown classes. In the context of this work, open-set semantic segmentation can be an ally in improving databases and models for identifying and classifying tattoo objects, allowing the improvement of annotations and descriptions of complex tattoos. Therefore, semantic segmentation must also be seen as a middle process, not just as an end process in the tattoo recognition roadmap.

Studies have explored the use of open-set semantic segmentation in different applications [15], [16]. These studies focus on adapting or building open-set classifiers to make closed-set semantic segmentation models capable of recognizing unknown classes. While the outcomes of these studies are promising, the performance of these approaches is limited due to the low representation of the obtained features, resulting in an "irregular" logit space with low discrimination among the classes. Recent research [17], [18] has demonstrated that incorporating metric learning techniques can enhance open-set recognition. Metric learning aids in obtaining more discriminative features and building a logit space that tightly clusters known classes while maintaining a considerable distance from unknown classes. However, it is essential to acknowledge that applying metric learning in the context of semantic segmentation can be impractical. This is primarily due to the exponential complexity of the task, as calculating pairwise distances among logit vectors of pixels becomes computationally expensive.

Recent studies have investigated the potential of large-margin learning to acquire more discriminative features, yielding improved outcomes in image classification tasks [19], [20], [21]. Hence, strategies established on large-margin learning present promising and viable alternatives for building a well-defined logit space that enhances the separation among decision boundaries of semantic classes.

Motivated by these results, our study explores the discriminative capabilities of large-margin learning to produce more distinctive features for tattoo semantic segmentation. This approach effectively increases the spatial separation among decision boundaries to different semantic classes, forcing the build of ideal logit space as illustrated in Figure 1b. Furthermore, this expanded separation among decision boundaries will set the stage for accommodating unknown classes in the future, enhancing the effectiveness of tattoo semantic segmentation within an open-set scenario. Visual comparisons depicting the differences in logit space
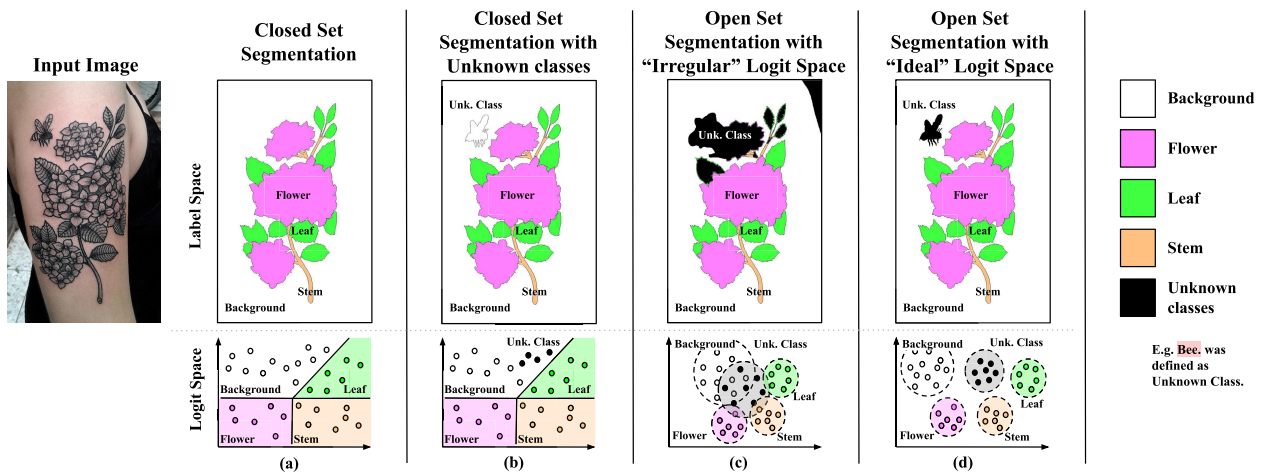
FIGURE 1. Closed set and open set in dense labeling scenarios. The label space represents the pixel-level predictions, and logit space refers to a subset of pixel samples in a 2D manifold separated by labels and decision boundaries for each class. a) Closed set without the presence of unknown classes. b) Closed set with the presence of unknown classes. These unknown classes result in misclassification, being segmented as known classes. c) Open set segmentation with "irregular" logit space. The term "irregular" suggests undefined or overlapping decision boundaries among classes. d) Open set segmentation with "ideal" logit space. The term "ideal" refers to well-defined decision regions between the classes and enough space (Open space) between them to include new classes [14], [18].

resulting from the presence of unknown classes can be observed in Figure 1.

Given the limitations presented so far and the fact that, to the extent of our knowledge, open-set classification has not been used in the context of tattoo recognition and is an open research gap, this work aims to propose a new large-margin-based loss function adapted to the context of open set semantic segmentation in tattoos. This novel loss function seeks to overeat the closed set segmentation results and enable and increase the open set segmentation results compared to other state-of-the-art semantic segmentation loss functions. Furthermore, we propose using a publicly available test dataset, with an annotation aimed at semantic segmentation in an open-set context, whose high complexity will serve as a benchmark for comparing methods in this area. The main contributions of this paper are then summarized as follows:

- Test set publicly available for the dataset, allowing comparisons and future work in the open-set and closed-set scenarios;
- Novel class semantic augmentation method to expand the tattoo samples;
- Novel large-margin loss function for open-set tattoo semantic segmentation to build more discriminative features and handle the class imbalance;
- A generalized open-set classifier approach based on open principal component scoring with incremental learning called G-OpenIPCS;
- Detailed and in-depth comparison with different state-of-the-art loss functions and open-set semantic segmentation methods;
- Statement of the main challenges for the open-set tattoo semantic segmentation.

This paper is organized as follows. Section II presents related works, mainly including related datasets, tattoo segmentation, and open-set semantic segmentation. Our proposed methods, particularly the novel tattoo semantic segmentation dataset, novel tattoo semantic augmentation method, and the novel loss function and open set classifier, are detailed in Section III. The experiment setup is discussed in Section IV, with results, discussions, and comparisons with state-of-the-art approaches presented in Section V. Finally, Section VI shows the conclusions, with the open challenges in open-set tattoo semantic segmentation discussed in Section VII.

II. RELATED WORKS

In order to present a general overview of the state-of-the-art in the context of tattoo segmentation, we divided the related works into three parts where this paper presents main contributions: datasets, tattoo segmentation, and open set semantic segmentation. The following subsections detail each of these works, pointing out the innovative aspects of this work on each front.

A. AVAILABLE DATASETS

Regarding the datasets, we chose to organize in the Table 1 a summary of the main characteristics of the most common datasets in tattoo detection, classification, and segmentation problems. These criteria include the number of samples, public availability, and type of annotation (classification, detection of objects with bounding box – BB, or segmentation). Regarding the annotation focused on segmentation, we also included whether it contains semantic segmentation.

In the case of [12], despite the relatively large number of samples and an annotation focused on tattoo detection,

TABLE 1. Tattoo datasets.

Ref.	Total	Public?	Annot.	Semantic?
[12]	7,526	No	BB	-
[9]	890	Yes	Seg.	No
[22]	5,740	Yes	Class.	-
[23]	5,000	Yes	BB	-
[24]	210	Yes	Class.	-
[13]	5,500	Yes	Seg.	No
Ours	2,106	315	Seg.	Yes

TABLE 2. Tattoo segmentation models.

Ref.	Network	Semantic?	Open-Set?
[9]	ConvNet	No	No
[31]	AlexNet+VGG	No	No
[13]	ViT-based	No	No
Ours	SegFormer	Yes	Yes

the dataset is private, which makes comparisons, establishing benchmarks, and analyzing new methods for the dataset difficult. On the other hand, the dataset proposed in [23] is publicly available and annotated only with BB, i.e., it is not possible to use it in the context of segmentation. The datasets proposed in [22] and [24], in turn, are focused exclusively on classification and provide only one label for each image or image patch, restricting their use to multi-class classification problems, without the location of the tattoo in the image.

Hrkać et al. [9] were the first to address tattoo segmentation. The proposed method aimed to de-identify soft biometric identifiers (tattoos) by discriminating tattoo and non-tattoo image patches with a deep neural network. In this sense, the proposed dataset presents a pixel-level annotation of the presence or absence of a tattoo in the image. However, the authors did not individualize the tattoo classes in the proposed annotation. The same is observed in the dataset presented in [13]. Furthermore, the authors presented a proposal using semi-synthetic images. This characteristic can sometimes lead to an image far from a real tattoo, compromising the segmentation approach.

Our dataset, in turn, presents some original and innovative features that can complement currently available datasets: (i) Inclusion of a semantic segmentation annotation; (ii) Several classes of tattoos in the same image, increasing complexity; (iii) Quite varied sizes of tattoos and classes of tattoos in the same image; and (iv) Tattoos in different regions of the body, maintaining the variability that exists in real situations. As will be detailed later, only the test set is made publicly available since most of the images available in these scenarios contain public use restrictions. However, we believe that as it is the first dataset with semantic annotation in segmentation, it will allow the comparison and evaluation of different approaches to this problem, which is significantly challenging.

B. TATTOO SEGMENTATION

Tattoo segmentation methods were presented in many studies, but their results were suppressed, maybe because segmentation was not the main focus. Furthermore, as many of them were carried out a long time ago, the research did not address deep learning methods, for example, and the methods cannot be directly compared with the approach adopted here.

In such cases, authors have performed their researches using methods based on: (i) Content-Base Image Retrieval

(CBIR) and Edge Direction Coherence Vector (EDCV) [25]; (ii) 3×3 Sobel filter [26], Active Contour CBIR (ACCBIR), and Vector Field Convolution (VFC) [1]; (iii) a complex system combining bottom-up and top-down priorities that transfer tattoo segmentation to detection split-merge skin detection, followed by figure-ground tattoo segmentation [27]; (iv) LoG (Laplacian of Gaussian) and Sobel kernel filters called quasi-connected components (QCC), using the GrabCut algorithm to produce the final segmented tattoo image [28]; (v) a negative image method with HSV (hue, saturation, and value, or lighting) model [29]; (vi) identification of pixels of skin in regions close to the tattoos and a graph-cut model based on skin color and a visual bump map [30], and (vii) a k-means cluster used in LAB color space to detect the skin area with a morphology processing used to smooth the clear graphic of the tattoo image segment [10]. The main limitation of these hand-crafted-based methods is that the feature extractor may have adequate performance for some classes and datasets but significantly lower performance for others, with compromised generalization. This is accentuated for datasets with greater variability in tattoo images.

As far as our research has reached, only two studies have used deep learning methods for the problem of tattoo segmentation.

Based on the study on CNNs, [9] used the structure of a ConvNet network to train small pieces of images to learn to identify which ones have or do not have pieces of tattoos. After training the network, a sliding window was passed through the image to be tested, and each segment of the sliding window was tested as a piece with a tattoo or not, marking the positive pieces. Parts with possible tattoos would be segmented piece by piece at the end of the slide.

Hrkać et al. [31] proposed a continuation of work presented in [9], this time testing three different networks for tattoo segmentation: (i) an architecture consisting only of multiple fully connected layers, without convolutional layers; (ii) an architecture inspired by the AlexNet network; and (iii) an architecture inspired by the VGGNet network. On the other hand, state-of-the-art segmentation models based on Vision Transformer (ViT) were evaluated in [13]; however, the main idea of that work was the unsupervised tattoo generator that allowed the creation of many semi-synthetic images with tattooed subjects. Hence, as shown in Table 2, related approaches used still needed to follow a semantic segmentation methodology and did not use an open-set semantic segmentation view, which is the focus of our current study and detailed as follows.

C. OPEN-SET SEMANTIC SEGMENTATION

The success of the fully convolutional network (FCN) in closed-set semantic segmentation [32] has led to the successful implementation of various neural network models for closed-set semantic segmentation on different applications [7]. However, these methods are unsuitable for open-set scenarios, which are common in real-world computer vision. This is because the closed-set perception fails when unknown classes from training are found in the test phase [33]. Due to that, several proposals have been developed for the open-set context in different applications, mainly autonomous driving, remote sensing, and data collection [34].

For open-set semantic segmentation, the loss function selection to guide the optimization process is an important aspect of the achieved results [35]. In general, studies limit themselves to using the successful and widely employed cross-entropy loss (CE) [15], [16], [33], [36], [37], [38], [39], [40]. This loss function measures the disparity between the predicted values and the ground truth, guiding the model's learning process based on labeled data. Hence, the studies focus on adapting or building open-set classifiers to make closed-set semantic segmentation models capable of recognizing unknown classes, even in label noise scenarios [41].

In this sense, metric learning has recently gained significant attention in addressing the open-set problem. Metric learning is an approach based on learning a distance metric that reinforces similarity between objects in the latent space. It has been studied in various fields, such as image classification [14], semantic segmentation [42], [43], and zero-shot segmentation [44]. According to [42], the metric learning strategy aims to direct the feature extraction process to obtain a well-controlled latent space, maximizing inter-class spacing and minimizing intra-class spacing based on a distance metric. Thus, the unknown samples are repelled into open space, as can be seen in Figure 4.

Similarly, the large-margin-based loss (LM) functions [19], [20], [21] maximize the margins between classes by imposing a regularization on the logit vectors of pixels to induce an increased separation between the boundary regions of the semantic classes. This strategy improves the generalization of models and provides open space between classes that can be valuable in the context of open-set semantic segmentation, allowing unknown samples to be projected into these open spaces. Furthermore, the large-margin loss adopts a more efficient training strategy than other metric learning approaches, such as those involving cubic costs for computing pairwise distances between the logit vectors of pixels [45].

Recent studies [40], [46] have made improved open-set semantic segmentation for autonomous driving applications by employing negative auxiliary data. Unlike these studies, our approach uses no auxiliary data to enhance the performance of open-set semantic segmentation. Table 3 summarizes the approaches used for open-set semantic segmentation.

TABLE 3. Open-set semantic segmentation approaches.

Ref.	Loss	Context	Aux. data?
[15], [16], [36], [38], [39]	CE	Remote Sensing	No
[33], [37]	CE	General	No
[33]	CE	Synthetic Data	No
[40], [46]	CE	Aut. Driving	Yes
[42]	Metric	Aut. Driving	No
[44]	Metric	General	No
[43]	Metric	General	No
Ours	LM	Tattoo	No

In the context of tattoo semantic segmentation, as stated in the previous Section, to the extent of our knowledge, no works focus on the open set context—the most closely related works are [23] and [47]. Reference [23] built a tattoo search approach that can learn tattoo detection and compact representation jointly in a single CNN via multi-task learning is presented. However, the compactness proposed by the authors is more focused on the compressive yet discriminative feature learning for large-scale visual search and instance retrieval applications, i.e., the efficiency of the search procedure. Open-set classification is not presented, and discussions of the proposed multi-task procedure are not encouraged for other applications. While [47] presented a classification method based on the Extreme Value Theory for tattoo classification. However, the focus of the proposed approach was the mid-level representations as a tool to adjust the trade-off between accuracy and efficiency. Hence, the results are mainly dedicated to real-world computer vision systems, where high accuracy is maintained even on commodity hardware with a low computational budget. Details regarding open-set are also not addressed.

To the best of our knowledge, this is the first work to explore tattoo semantic segmentation in closed and open-set scenarios, establishing benchmarks for both conditions using our publicly available test set. Our approach introduces the large-margin loss function as an efficient learning strategy to build a well-defined logit space and handle class imbalance, using contemporary network architecture based on transformers and presenting the generalist approach to integrating a robust open-set classifier for semantic segmentation tasks.

III. PROPOSED METHODS

This section outlines the proposed methods. Firstly, in Section III-A, a novel-built TSSD2023 dataset is presented. Then, Section III-B introduces a novel tattoo semantic augmentation to expand tattoo samples of the TSSD2023 dataset. Subsequently, in Section III-C, a novel large-margin loss function is proposed to handle class imbalance and enhance the discriminative of the classes in the TSSD2023 dataset. Lastly, in Section III-D, a generalized approach for OpenIPCS is proposed for open-set semantic segmentation tasks.

A. TATTOO SEMANTIC SEGMENTATION DATASET

Identifying the various things that make up a tattoo can be defined as a semantic segmentation task. This process allows machines to comprehend the meaning behind a tattoo better. Its usefulness is particularly evident in security systems, where it assists in identifying and monitoring individuals through surveillance systems, for instance, locating a person based on a brief tattoo description. However, as previously presented, current tattoo datasets are limited to image classification [22], [24], object detection [12], [23], or tattoo segmentation [9], [13], which only separate the tattoo from the background and fail to provide a comprehensive understanding of the significance of the tattoos.

In this manner, we created the Tattoo Semantic Segmentation Dataset (TSSD2023), respecting the copyrights of image owners to the greatest extent possible. To build this dataset, web scraping was conducted on Flickr.[1] Numerous terms related to tattoos were used as search queries on the platform. Then, a visual inspection was performed by humans to confirm whether the images obtained contained tattoos and whether their content was suitable for sharing. Finally, the licenses users attribute to these images when sharing them on the platform were considered to define the training, validation, and test sets.

As a result, the test dataset exclusively comprises images for which sharing permissions had been granted. In contrast, the training and validation datasets consist only of images for which sharing was unauthorized. This division strategy was adopted due to the limited availability of images with public sharing licenses. Thus, the test sets will be publicly available for comparison and development of future work. However, the training and validation sets will be kept private to ensure that none of the authors' copyrights are infringed.

Figure 2 shows 33 classes[2] for Known Known Class (KKC) tattoos and 23 classes for Unknown Unknown Class (UUC) tattoos selected for annotation in TSSD2023.[3] The motivation for choosing this split for KKC and UUC classes was mainly based on the number of images available for each semantic class. Classes with reduced representation in the universe of available tattoos and, consequently, public samples were selected exclusively to compose the UUCs due to the impossibility of successfully training the segmentation models on these classes. Furthermore, the proposed division guarantees that the test semantic classes defined as UUC

---

[1]A photo and video hosting platform established in 2004. Available in: https://flickr.com/

[2]The dataset also includes annotations for the 'stem/branch' and 'rope' classes, which were omitted from the analysis due to the limited number of annotations.

[3]The closed-set approach assumes that the training and testing pixels belong to the same label space (defined as KKCs), meaning that the train and test sets contain the same classes. However, this assumption does not hold in real-world scenarios, especially in earth observation applications. During the prediction phase, the model may face pixels from classes not seen during the training phase (UUCs). We direct the reader to [14] for a deeper reading about these definitions.

TABLE 4. Samples of the images and annotations of the TSSD2023 dataset.

are not found on the training and validation sets. Hence, we consider the proposed division sufficient to evaluate the open-set methods proposed in this work. Once the data were divided, UUCs were chosen to form part of the test set for open-set evaluation. This approach enabled the representation of similar and dissimilar semantic classes compared to KKCs, allowing for evaluation in straightforward and complex scenarios.

Each KKC class was meticulously labeled with unique identifiers to enable the model to distinguish each object semantically. In contrast, all UUC tattoos were assigned the same "unknown" class label. Notably, all KKC classes are represented in the training and validation sets and the test set. However, UUC tattoos are exclusively found in the test open-set. A human inspection process was also conducted to ensure these open-set recognition conditions [14].

Table 4 presents four possible tattoo composition situations found in TSSD2023: (i) Single tattoo: only one tattoo class is present in the image; (ii) Multiple tattoos: multiple tattoo classes can be found in a single image; (iii) Multiple tattoos (with overlap): similar to the previous scenario, but the tattoos are overlapped with each other. Due to this overlap, this scenario is more challenging than the previous one [33]; and (iv) Tattoos (unlabeled) as background: Tattoos (unlabeled) as background: Specific tattoo classes were categorized as background classes due to two primary reasons. First, this decision was necessary because, at times, it was impossible to determine their specific semantic tattoo class. Second, due to a limited number of available samples, assigning individual semantic labels to these classes was not feasible. These situations provide a comprehensive representation of tattoo compositions within the dataset.

In conclusion, TSSD2023 contains 2,106 tattoo images without specific image resolution standards that have been annotated at the pixel level. These images are divided into the following subsets: 1,404 for training, 387 for validation, and 254 for the closed-set test. Additionally, 61 images with UUC tattoos are included to form the test open-set with 315 images. Figure 3 illustrates the distribution of pixel percentages for each class within TSSD2023, demonstrating that the dataset
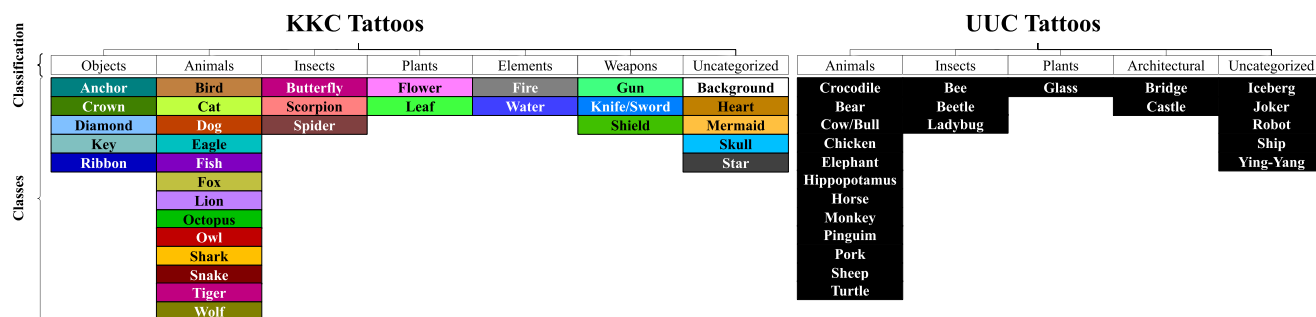
FIGURE 2. TSSD2023 Classes. KKCs tattoos correspond to the 33 labeled classes, while the UUCs tattoos are represented by the 23 classes defined as "unknown" (black label). These classes can be viewed within a conceptual taxonomy, facilitating an understanding of the domain coverage provided by the dataset.
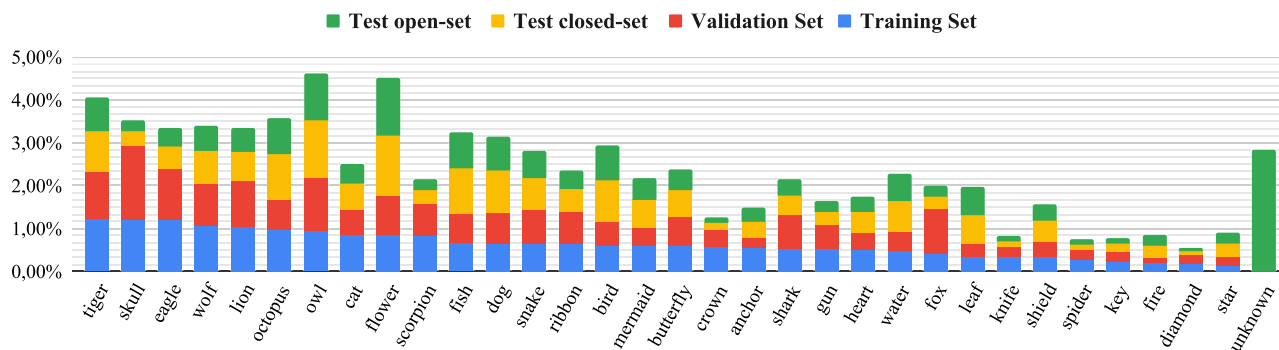

FIGURE 3. The distribution of pixels per class and set. These values are organized based on their quantities in the training set. The values of background classes were suppressed due to discrepancies with the other classes. The reference values for the background are 79.8% in training, 79.4% validation images, 81.4% test closed-set images, and 81.7% test open-set images.

TABLE 5. Samples of the data augmentations on only the heart class, except the mix column that applies the data augmentations on all semantic classes.
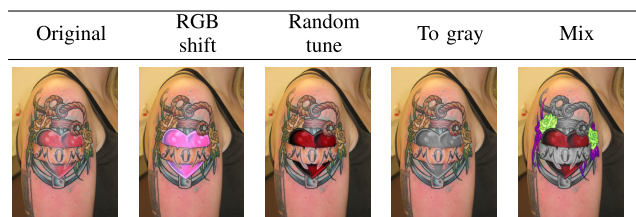
is notably imbalanced, particularly in the background class, which accounts for an average of $\approx 80\%$ of pixels across the subsets. It is also important to highlight that the unknown class represents approximately 3% of the pixels present in the test open-set. Samples of images from TSSD2023 can be observed in Table 4, showcasing tattoos presented in diverse scenarios, sizes, styles, positions, and combinations. These variations aim to maintain the variability of real-world situations.

B. CLASS SEMANTIC AUGMENTATION

Motivated by the limited number of samples within TSSD2023 compared to the extensive diversity of real-world tattoos, this work introduces a novel data augmentation named class semantic augmentation (CSA) to increase the variety of tattoos contained in TSSD2023. This method applies pixel-level transformations to distinct classes, enabling unique augmentation for different classes within an image, as illustrated in Table 5. In this instance, the pixels of the heart class go through different transformations of colorations, gray styles, and color tones, trying to approximate the infinite possibilities of representing tattoos in the real world from limited semantic data.

Algorithm 1 details the proposed implementation of tattoo semantic augmentation. The algorithm takes five parameters: the input image X, the ground truth Y, a list of pixel-level transformations to increase the variability of semantic classes t, a list containing the probabilities of executing each transformation p, and a list of semantic class indices ic that remain unchanged during data augmentation. It is important to note that the parameters p and ic are optional. The algorithm first identifies the indices of semantic classes within the image. Subsequently, it iterates through each index, applying one of the transformations specified in t. As shown in Table 5, a heart can be represented in various ways while other classes remain unchanged. It demonstrates that this semantic augmentation allows specific adjustments

Algorithm 1 Class Semantic Augmentation

Input: Image (X)
Input: Ground truth (Y)
Input: List of transforms (t)
Input: List of probability of choice (p)
Input: List of ignore classes (ic)
Output: Transformed image ($\hat{X}$)

1: $seg\_ids = unique(Y)$
2: $seg\_ids = remove\_ic(seg\_ids, ic)$
3: $\hat{X} = copy(X)$
4: **Foreach:** $i \in seg\_ids$ **do**
5:     $y = new\_array\_zeros(Y.shape)$
6:     $y[Y == i] = 1$
7:     $T = random\_choice(t, p)$
8:     $\hat{X} = apply\_transform(T, \hat{X}, y)$
9: **end foreach**
10: **return** $\hat{X}$

---

for each class in the dataset, making it useful in various application domains, such as autonomous driving [42] and fashion images [48]. In addition, different pixel-level transformations can be selected for more suitable application contexts. For instance, the color of a flower may vary. At the same time, the leaf and stem could be confined to modifying the original color tones, preserving the real-world patterns. On the other hand, all semantic classes in an image can be changed without restrictions, as observed in the mix column of Table 5 and as employed in this study. It is worth noting that this augmentation technique is limited to pixel-level transformations.

C. PROPOSED LOSS FUNCTION

Classical convolutional neural networks (CNNs) based semantic segmentation networks [32] can be divided into two parts: a feature extractor that includes several convolution layers followed by max-pooling and activation function, and the linear classifier $f = W^\top x + b$ in the last fully-connected layer applied on the feature vector x of the penultimate layer for obtaining the logit vector $f \in \mathbb{R}^C$ of each pixel of the input image, in which C represents the number of classes. To solve the overfitting problem and produce more discriminative logit vectors f in training, classifier margin has been exploited [19], [21], [49]. Following [50], the classification margin is the difference between the predicted score f_{c*} and target score f_y, where y indicates the ground truth class labels and $c^* = \operatorname{argmax}_{c \neq y} f_c$. Based on the margin $f_y - f_{c*}$, the traditional classifier margin loss function can be expressed as follows:

$$\mathcal{L}_{margin} = L\left(\max_{c \neq y} f_c - f_y + \rho\right), \quad (1)$$

in which ρ is a boundary control parameter, usually $\rho \geq 0$. Thus, increasing the value of ρ results in a larger classification distance between labels.

According to [20], the cross-entropy loss ($\mathcal{L}_{CE}$) can partially encourage the development of a large-margin classifier within the CNNs. Based on this analysis, a symmetric Kullback-Leibler (KL) divergence term $\mathcal{L}_{LM}$ was introduced as a regularization component for $\mathcal{L}_{CE}$, inducing a more large-margin classifier in the original $\mathcal{L}_{CE}$. The combination was termed the large-margin cross-entropy loss ($\mathcal{L}_{CE+LM}$), and it is the formula is presented as follows:

$$\mathcal{L}_{CE+LM} = \mathcal{L}_{CE} + \mathcal{L}_{LM}, \quad (2)$$

in which $\mathcal{L}_{CE}$ is defined as:

$$\mathcal{L}_{CE} = -\log p_y(f) = -\log \frac{\exp(f_y)}{\sum_{c=1}^{C} \exp(f_c)}, \quad (3)$$

and $\mathcal{L}_{LM}$ is expressed as:

$$\mathcal{L}_{LM} = \frac{\lambda}{2} \sum_{c \neq y} \left\{ \frac{\exp(f_c)}{\sum_{c' \neq y} \exp(f_{c'})} - \frac{1}{C-1} \right\}$$
$$\times \log \left\{ \frac{\exp(f_c)}{\sum_{c' \neq y} \exp(f_{c'})} \right\}, \quad (4)$$

where λ is a regularization parameter. Increasing the value of the λ enlarges the space between classes, thereby more resistance faced by the learning objectives. The detailed derivation has been explored in [19].

Considering the class imbalance between the background and the foreground classes in the tattoo semantic segmentation dataset, we modify the $\mathcal{L}_{CE+LM}$ replacing the $\mathcal{L}_{CE}$ by the focal loss $\mathcal{L}_{FCL}$. This choice is based on the fact that $\mathcal{L}_{FCL}$ is a variant of the $\mathcal{L}_{CE}$ that preserves the discriminative capacity of the original loss while dealing with the imbalance among the classes [51]. Thus, our proposed large-margin focal loss ($\mathcal{L}_{FCL+LM}$) is described by:

$$\mathcal{L}_{FCL+LM}(\text{ours}) = \mathcal{L}_{FCL} + \mathcal{L}_{LM}, \quad (5)$$

$\mathcal{L}_{FCL}$ is defined as:

$$\mathcal{L}_{FCL} = \alpha(1 - p_y(f))^\gamma \times \log p_y(f). \quad (6)$$

The $\mathcal{L}_{FCL}$ applies a modulating term to the $\mathcal{L}_{CE}$ to focus learning on hard examples and down-weight the numerous easy examples, where α control the class weights and γ reduce the loss contribution from easy examples. Thus, we obtain a loss that deals with imbalances between classes while increasing the classification margin to improve the discriminability of the trained model.

D. OPEN-SET CLASSIFIER FOR SEMANTIC SEGMENTATION

In semantic segmentation networks, the logit vector $f \in \mathbb{R}^C$ are commonly normalized using the softmax function into a probability distribution for each class $y \in \{1, \ldots, C\}$ to perform the final classification of each pixel. Therefore, the final classification of each pixel is defined as $\hat{Y}_f^{close} = \operatorname{argmax} p_y(f)$.

This learnable classifier cannot recognize UUC, making it unsuitable for open-set recognition as it assigns all features to

KKCs [42]. Thus, open-set classifiers must be developed to classify KKCs while recognizing UUCs accurately. Previous studies have investigated this development in semantic segmentation task [16], [38], [42], [52]. Based on those studies, one can present the general open set classifier as follows:

$$\hat{Y}_f^{open} = \begin{cases} C_{UUC} & \max(p_y(\boldsymbol{f})) \leqslant \lambda_{out}, \\ \hat{Y}_f^{close} & \max(p_y(\boldsymbol{f})) > \lambda_{out}. \end{cases} \quad (7)$$

where C_{UUC} denotes the UUC, and λ_{out} the cutoff threshold to determinate the UUC pixels.

A recent study [38] presented an open-set classifier method based on principal components analysis (PCA) on the internal logic vectors of CNNs to provide an open-set semantic segmentation, named Open Principal Component Scoring with Incremental Learning (OpenIPCS). Its training approach is efficient, which is crucial due to the exponential nature of pixel-level classification compared to image classification. Furthermore, the study notes that OpenIPCS outperforms open-set semantic segmentation strategies based on probability maps by a significant margin.

OpenIPCS was inspired by the Conditional Gaussian Distribution Learning (CGDL) proposed in [53], which is a Variational Autoencoder (VAE) model for conditional Gaussian distribution estimation, capable of learning conditional distributions of KKC and rejecting UUC examples. In distinction, the OpenIPCS replaced the VAE with PCA and uses multiple internal activation layers to adjust the generative model with validation samples only.

Reference [54] showed that logit vectors get closer to the label space as CNN layers deepen. Thus, in addition to the last layer f, OpenIPCS considers the enabling aspects of the previous layers $f, f^1, \ldots, f^L$, in which L denotes the total number of the CNNs layers. This approach combines low-level and high-level semantic information, thus enhancing the discrimination capability of the model.

Then, for each pixel, a corresponding feature vector $\hat{f}_{train}$ is built by concatenating the network's internal logit vectors $f, f^1, \ldots, f^L$. Such a concatenation produces high-dimensional and redundant features due to the hundreds or thousands of activation channels present in the CNN and the FCN layers [55], [56].

As described by [57], PCA can serve a dual purpose. Apart from its primary role in reducing dimensionality, it can also act as a probability density estimator with Gaussian priors. These features allow OpenIPCs to use PCA as a generative model (G) for UUC recognition while solving the high-dimensionality problem of feature vectors.

This G is incrementally adjusted using only the validation images. This process consists of adjusting the PCA with a batch of samples from the validation set. The classification step using the G consists of projecting the feature vector $\hat{f}_{test}$ of the test images in the latent space obtained by adjusting the PCA on the validation images and performing the inverse process to obtaining the $\hat{f}_{test}^G$. The difference between the original feature vector $\hat{f}_{test}$ and the $\hat{f}_{test}^G$ is calculated.

Consequently, pixels of KKCs have low difference values, while pixels belonging to the UUCs have high difference values. Thus, in accordance with [38], open-set recognition from OpenIPCS can be achieved as follows:

$$\hat{Y}_f^{openipcs} = \begin{cases} C_{UUC} & abs(\hat{f}_{test} - \hat{f}_{test}^G) > \lambda_{out}, \\ \hat{Y}_f^{close} & abs(\hat{f}_{test} - \hat{f}_{test}^G) \leqslant \lambda_{out}. \end{cases} \quad (8)$$

The λ_{out} do not represent equal statistical entities with the Equation 7. Thus, as in [38], the λ_{out} value was defined from preset values of True Positive Rate (TPR).

As outlined in [54], the deeper layers are closer to the label space. Based on this analysis, this work proposes a generalized version of OpenIPCS (G-OpenIPCS) that focuses only on the last and penultimate layers of CNNs. These layers contain more substantial semantic information, which proves highly valuable for training OpenIPCS. Consequently, it becomes possible to disregard the other network layers within the CNNs. This aspect simplifies our approach and is easy to incorporate into different network architectures, unlike the original OpenIPCS building exclusively on the FCN decoder.

IV. EXPERIMENTS SETUP

Due to the open-set recognition process being still dependent on models built in a closed-set [33], [38], our experiments are divided into three parts. First, in Section IV-A, we compare our proposed loss function, described in Section III-C, with other significant loss functions widely used in semantic segmentation or recently introduced to improve discriminability in deep neural networks. Subsequently, in Section IV-B, we employ the proposed G-OpenIPCS approach detailed in Section III-D to evaluate the performance of models trained with the loss functions from the previous experiment in an open-set tattoo semantic segmentation scenario. Finally, we compare the performance of our open-set semantic segmentation approach with other state-of-the-art open-set semantic segmentation methods. Figure 4 provides an overview of the proposed method. Each stage of Figure 4 is detailed as follows.

A. CLOSED-SET SEMANTIC SEGMENTATION

1) DATASETS
A total of 2,045 images from TSSD2023 were used to train and evaluate the closed-set models. These images follow the following division: 1,404 for training, 387 for validation, and 254 for closed set testing, as described in Section III-A.

2) PRE-PROCESSING AND DATA AUGMENTATIONS
This step involves a sequence of transformations to train and improve the semantic segmentation network, motivated by the limited dataset of training images available in TSSD2023 around the high diversity of tattoos that virtually have no limits in the real world. This process involves the following steps: a) Implement a resizing operation, which can be either random or center crop, or resize to get an image
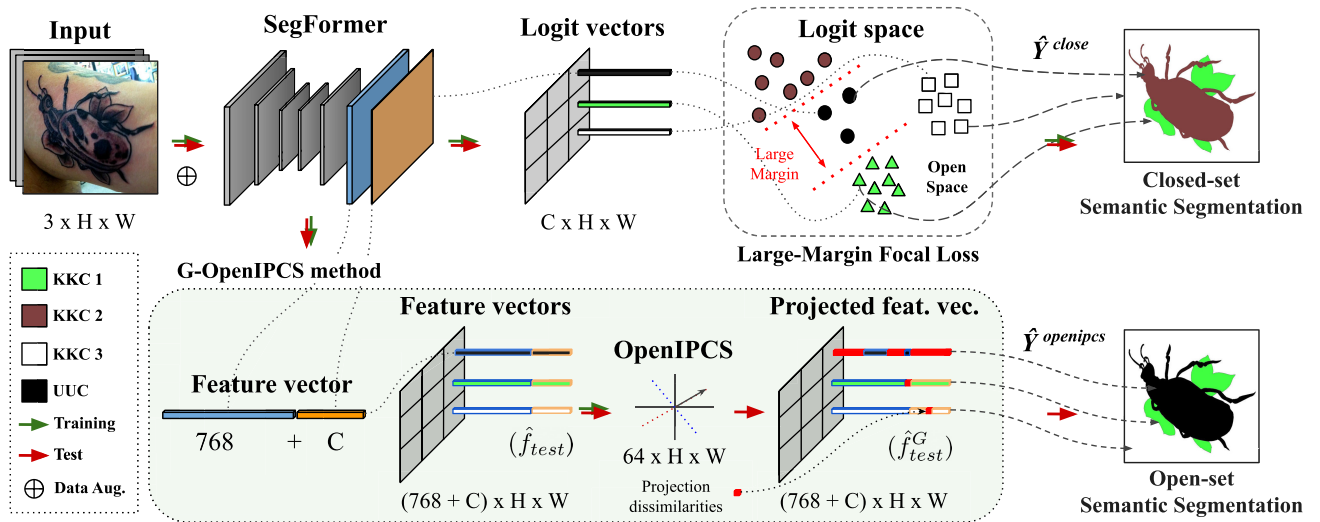
FIGURE 4. Overview of the proposed methods. Closed-set segmentations are obtained from models trained by different loss functions. The proposed Large-Margin Focal Loss expands the decision boundaries between classes to improve class discrimination while dealing with the problem of imbalance between classes. Increasing the spacing between classes allows unknown samples to be projected into open space. The open-set segmentation process is highlighted within the green block. The proposed G-OpenIPCS recognizes UUC pixels, leaning on discrepancies between feature vectors. This approach uses only the last two layers of the segmentation model. The penultimate layer equals 768 channels for SegFormer and 512 channels for Swin+UPNet.

with a resolution of 224 × 224; b) Next, there is a 50% probability of applying one of the geometric transformations to the image, including horizontal and vertical flipping, shifting, scaling, or rotation. These transformations have a scale and shift limit of 0.2, and rotation is limited to a maximum of positive or negative 45 degrees; c) Executing the tattoo semantic augmentation method proposed by this work, as described in more detail below. This augmentation involves RGB shifting, conversion to grayscale, and the application of random tone curve transformations. Each of these transformations has a 25% probability of being applied to each class within an image; d) Subsequently, there is a 50% chance of applying certain global modifications to the image. These modifications can consist of random adjustments to brightness and contrast with a limit of 0.3, Gaussian blur with a 5-neighbor mask, or fancy PCA transformation; e) Next, coarse dropout transformation [58] with a maximum of 8 plots per image with the maximum resolution of 32 × 32 is applied; f) Next, the images are normalized using the mean and standard deviation values obtained from the ImageNet [59]; g) Finally, there is a 50% probability of applying the CutMix transformation [60] at the batch level to blend the images. All of these transformations were applied or developed with the on of the Albumentations library.[4]

3) IMPLEMENTATION DETAILS
Due to recent advances presented by the semantic segmentation networks based on transformers [61], [62], [63], this work uses the SegFormer [63], which consists of a hierarchical Transformer encoder with lightweight multilayer

perceptron (MLP) decoders. The model was pre-trained on the ImageNet dataset [59] and was acquired through the Hugging Face library.[5] During closed-set training, a batch size of 16 images on an RTX 3090 with 500 epochs was utilized, employing a patience factor of 10 epochs. The Nadam optimizer [64] was used, with an initial learning rate set at 0.0001 and a weight decay coefficient of 0.00001.

4) BASELINE LOSS FUNCTIONS
In order to investigate the performance of large-margin focal loss ($\mathcal{L}_{LMFCL}$) in both closed-set and open-set semantic segmentation, we conducted a comparative study with five other loss functions: i) cross-entropy loss ($\mathcal{L}_{CE}$) – it was selected because it is widely used and a standard choice for semantic segmentation; ii) focal loss ($\mathcal{L}_{FCL}$) [51] – this is a commonly used variation of cross-entropy that is effective in addressing class imbalance; iii) Distance-based cross-entropy loss ($\mathcal{L}_{DCE}$) [65] and iv) Distance-based cross-entropy loss combined with Variance Loss ($\mathcal{L}_{DCE+VL}$) [42] were selected as they have been proposed to enhance the results of open-set recognition by increasing the inter-class distance and decreasing the intra-class distance to produce more discriminative features; v) label-distribution-aware margin loss ($\mathcal{L}_{LDAM}$) [66] was chosen as a key loss for long-tailed recognition. It increases the margin for tail classes and decreases for main classes based on class frequency; vi) dice similarity coefficient loss ($\mathcal{L}_{DSC++}$) [67] recently introduced loss that combines the robustness of the $\mathcal{L}_{DSC}$ to class imbalance while penalizing overconfident predictions, hence improving the performance of semantic segmentation on out-of-distribution

---

[4]https://albumentations.ai/

[5]https://huggingface.co/

TABLE 6. Parameters per each loss function utilized for the model training.

Ref.	Loss	Parameters
-	$\mathcal{L}_{CE}$	-
[51]	$\mathcal{L}_{FCL}$	α: 1.0, γ: 2.0
[65]	$\mathcal{L}_{DCE}$	T: 3
[66]	$\mathcal{L}_{LDAM}$	s: 30, λ_{max}: 0.5
[20]	$\mathcal{L}_{LMCE}$	λ: 0.3
[42]	$\mathcal{L}_{DCE+VL}$	T: 3, λ_{vl}: 0.01
[67]	$\mathcal{L}_{DSC++}$	γ: 2.0
Ours	$\mathcal{L}_{LMFCL}$	λ: 0.3, α: 1.0, γ: 2.0

data; vii) large-margin cross-entropy loss ($\mathcal{L}_{LMCE}$) [20] was chosen because it has recently been proposed to improve the results of closed-set classifiers by increasing the space between the decision boundaries of each class. Moreover, we utilized the $\mathcal{L}_{LMCE}$ as a foundation to develop our proposed loss function for the open-set tattoo semantic segmentation on a scenario that presents various challenges, such as high intra-class variability, class imbalance, and small tattoos, while enhanced discriminative capabilities in closed-set and open-set semantic segmentation. Table 6 presents the loss function parameters used for model training.

5) METRICS

In order to assess the results of closed-set semantic segmentation, followed by open-set segmentation, we employed the following metrics. Firstly, we used the Area Under the ROC Curve (AUROC) because it has been recently utilized to assess open-set semantic segmentation [38], [42]. This metric aids in measuring the model's ability to distinguish between classes, particularly with respect to UUC. Secondly, the macro-averaged F1-Score (Macro F1), recommended in previous works for open-class classification [68], [69]. It helps evaluate the model's precision and recall of the segmented pixels. Lastly, the Mean Intersection over Union (mIoU) is a standard metric for semantic segmentation evaluation that provides an overview of the model's performance across all classes, including both KKCs and UUC.

B. OPEN-SET SEMANTIC SEGMENTATION

1) DATASETS

G-OpenIPCS was trained using just the 387 validation images. The test open-set with 315 images was used to evaluate the proposed method, as detailed in Section III-A.

2) IMPLEMENTATION DETAILS

For each model trained using the evaluated loss functions in closed-set semantic segmentation, a G-OpenIPCS model was trained to perform semantic segmentation in an open-set scenario. As depicted in Figure 4, our approach utilizes the last two connected layers of the segmentation network to construct feature vectors for each pixel. In the case of SegFormer, this feature vector is equal to the number of

KKCs plus the logit vector of the inner layer, with a length of 768. Subsequently, we defined 64 principal components for the adjustment in G-OpenIPCS. It is worth noting that the feature vector size and the number of principal components may vary depending on the semantic segmentation model and the application domain. Once G-OpenIPCS is trained, it becomes possible to recognize UUCs by considering the discrepancy between the raw feature vector and the projected feature vector, as outlined in Equation 8.

3) BASELINE

Initially, a comparative analysis of the open-set semantic segmentation results was conducted using G-OpenIPCS trained for each evaluated loss function. This step serves to validate the performance of the proposed loss function in the open-set scenario. Subsequently, our approach was compared with other significant open-set segmentation methods proposed in the state-of-the-art literature. These methods include OpenIPCS [38], Anomalous Probability Map (APM) [42], Maximum Unnormalized Logit (MaxLogit) [39], and Maximum Softmax Probability (MSP) [43]. For these methods, training followed the same configurations as described in Section IV-A, except that only geometric transformations were applied. All networks were pre-trained on the ImageNet dataset [59]. Specifically, for DRN50+PSPNet [42] and RN101+PSPNet [39], images were resized to dimensions of 250, 300, 350, 400, and 450, with a down-sampling factor of 8 for DRN50. Additionally, we implemented SoftMax-Thresholding (SoftMax-T) [38] and OpenMax [36] as the baselines for our implementation. As this is the first study on tattoo semantic segmentation and without other tattoo semantic datasets, we only compare our approach with state-of-the-art methods that do not use auxiliary data to supply their open-set recognition models. Finally, to demonstrate the generalization of our approach with baselines, we replicated the experiments, replacing the SegFormer with the Swin+UperNet model [70].

4) CUTOFF THRESHOLDS

All the methods compared in Section V-C require a cutoff threshold to distinguish between KKC and UUC pixels. To realistically undertake an open-set recognition task, these limits are defined empirically based on the available KKCs during the validation data following the conditions of the original papers. No information about the UUCs is used to select the cutoff thresholds for the open-set classifiers. For the proposed G-OpenIPCS, the preset values of the cutoff thresholds (λ_{out}) are determined based on TPR quantiles, as outlined in [38].

5) METRICS

To evaluate the semantic segmentation results in an open set, we kept the AUROC, Macro F1, and mIoU used in the experiment in a closed set, described in Section IV-A. Additionally, we highlight the results involving UUCs.

TABLE 7. Comparison of the closed-set segmentation results achieved per each loss function. Bold values indicate the best overall results, including all loss functions. The results were obtained using SegFormer.

Ref.	Loss	AUROC	Macro F1	mIoU
-	$\mathcal{L}_{CE}$	.7966±.122	.6179±.224	.4805±.216
[51]	$\mathcal{L}_{FCL}$	.8024±.111	.6433±.213	.5053±.209
[65]	$\mathcal{L}_{DCE}$	.8030±.118	.6108±.219	.4724±.218
[66]	$\mathcal{L}_{LDAM}$	**.8295**±.105	.6497±.181	.5057±.191
[20]	$\mathcal{L}_{LMCE}$	.8110±.112	.6457±.217	**.5097**±.216
[42]	$\mathcal{L}_{DCE+VL}$	.7843±.120	.5933±.217	.4519±.205
[67]	$\mathcal{L}_{DSC++}$	.8108±.108	.6303±.200	.4874±.196
Ours	$\mathcal{L}_{LMFCL}$	.8142±.102	**.6514**±.189	.5090±.194

V. RESULTS AND DISCUSSIONS

This section presents the results and discussions obtained from our experiments. The purpose is to validate the impact of the proposed Large-Margin Focal Loss on closed and open-set tattoo semantic segmentation, in addition to demonstrating the effectiveness of the proposed G-OpenIPCS for open-set semantic segmentation. Initially, in Section V-A, we present the results of the loss functions evaluated in this work in a closed set scenario. Next, in Section V-B, we evaluate the performance of models trained using the proposed G-OpenIPCS method. Finally, in Section V-C, we compare our proposed approach with other state-of-the-art techniques.

A. CLOSED-SET SEMANTIC SEGMENTATION

This section focuses on evaluating the performance impact of the proposed $\mathcal{L}_{LMFCL}$ in a closed-set semantic segmentation scenario. To demonstrate this, several loss functions from the literature were compared, as described in Section IV-A. Table 7 presents the overall results for all evaluated loss functions in terms of AUROC, Macro F1, and mIoU. When evaluating the performance of loss functions, $\mathcal{L}_{LMFCL}$ outperforms all other metrics in terms of Macro F1 and maintains consistent results across AUROC and mIoU. This makes $\mathcal{L}_{LMFCL}$ a suitable choice for closed-set semantic segmentation. However, $\mathcal{L}_{LDAM}$ and $\mathcal{L}_{LMCE}$ exhibit an outperform over $\mathcal{L}_{LMFCL}$ in terms of AUROC and mIoU, respectively. Additionally, $\mathcal{L}_{FCL}$ stood out by producing results that closely resemble those of $\mathcal{L}_{LMFCL}$ and $\mathcal{L}_{LMFCL}$, despite not incorporating a large margin in its formulation. This emphasizes its capability to address class imbalance encountered in TSSD2023.

On the other hand, $\mathcal{L}_{CE}$, $\mathcal{L}_{DSC++}$, and particularly $\mathcal{L}_{DCE}$ and $\mathcal{L}_{DCE+VL}$ exhibit lower effectiveness compared to the other loss functions. Among these less favorable outcomes, $\mathcal{L}_{DCE+VL}$ yielded the poorest results. This could indicate that the significant variability present in TSSD2023 posed challenges for the metric learning process. Thus, functions based on large-margin, such as proposed $\mathcal{L}_{LMFCL}$ and $\mathcal{L}_{LMCE}$, were demonstrated to be more effective for closed-set semantic tattoo segmentations.

Table 8 presents the closed-set segmentation results categorized by class obtained from proposed $\mathcal{L}_{LMFCL}$.

TABLE 8. Closed-set segmentation results separated by classes obtained from the SegFormer trained using the proposed $\mathcal{L}_{LMFCL}$. The results were sorted in descending order of mIoU↑ values. The top-5 best classes are highlighted in green, while the top-5 worst classes are highlighted in red.

Class	AUROC	Macro F1	mIoU ↑
background	0.8886	0.9669	0.9359
tiger	0.9632	0.8656	0.7631
octopus	0.9102	0.8552	0.7470
snake	0.9270	0.8451	0.7318
key	0.9136	0.8290	0.7079
owl	0.8813	0.8287	0.7076
butterfly	0.8729	0.8164	0.6897
lion	0.8714	0.8106	0.6816
star	0.8787	0.8028	0.6706
dog	0.8863	0.7868	0.6485
scorpion	0.9149	0.7586	0.6110
flower	0.8212	0.7412	0.5888
fox	0.8302	0.7083	0.5483
fish	0.8278	0.7042	0.5435
shark	0.8604	0.6978	0.5359
gun	0.8507	0.6914	0.5284
cat	0.7772	0.6878	0.5241
anchor	0.8192	0.6565	0.4886
bird	0.7641	0.6490	0.4804
diamond	0.8530	0.6280	0.4577
mermaid	0.9150	0.6256	0.4551
eagle	0.8788	0.6180	0.4472
spide	0.8606	0.5998	0.4283
heart	0.7696	0.5809	0.4093
ribbon	0.7572	0.5496	0.3789
wolf	0.8008	0.5403	0.3702
skull	0.7562	0.5393	0.3693
shield	0.7086	0.5375	0.3675
crown	0.7241	0.4718	0.3087
leaf	0.6564	0.4310	0.2747
water	0.6409	0.3939	0.2452
knife	0.5633	0.1796	0.0987
fire	0.5262	0.0989	0.0520
all	0.8142	0.6514	0.5090

It emphasizes the top 5 best and top 5 worst results, excluding the background class. The top-5 best have high IoU values, indicating that the model's segmentation performance for these categories is particularly accurate. Additionally, they exhibit relatively high AUROC and Macro F1 scores. Notably, the classes within the top 5 best results do not necessarily constitute the majority of pixel quantity in TSSD2023, as illustrated in Figure 3. This suggests that the excellent performance comes from other aspects, such as the

TABLE 9. Visual result samples of open-set semantic segmentation for each loss function. UUCs are depicted as black pixels. The respective colors for other semantic classes can be found in Figure 2. The UUC has been highlighted in yellow, while the top-5 best classes are highlighted in green, and the top-5 worst classes are highlighted in red. The results were obtained using SegFormer and proposed G-OpenIPCS.

semantic dissimilarity of this set of tattoos to other classes, and mainly due to the recurrence of these tattoos being tattooed individually, without overlapping with other semantic classes, as can be briefly observed in Table 9. In contrast, the top 5 worst classes can be categorized as complementary tattoos, meaning they are rarely encountered in isolation but are typically accompanied by other predominantly dominant tattoos in terms of pixel quantity. For instance, a crown is almost always found with another element, such as a human

or animal face. Leaves often complement flowers and stems, and water is nearly always associated with aquatic creatures like fishes, sharks, and octopuses.

In outline, $\mathcal{L}_{LMFCL}$ showed considerable impact in the closed-set semantic segmentation, presenting the best results overall. However, aspects can still be considered to improve performance in closed-set segmentation. In addition to dealing with class imbalance, it is still necessary to deal with the overlap between semantic classes, improving the accuracy

TABLE 10. Comparison of the open-set segmentation results achieved per each loss function. All results are obtained on SegFormer and the proposed G-OpenIPCS method. Bold values indicate the best overall results, including all loss functions.

Ref.	Loss	AUROC	AUROC $_{UUC}$	Macro F1	F1 Score $_{UUC}$	mIoU	IoU $_{UUC}$
-	$\mathcal{L}_{CE}$	.7737±.122	**.8204**	.5892±.234	.3759	.4525±.220	.2314
[51]	$\mathcal{L}_{FCL}$	.7865±.112	.7212	.5919±.216	.2886	.4505±.207	.1686
[65]	$\mathcal{L}_{DCE}$	.7827±.119	.7333	.5770±.215	.3782	.4357±.209	.2332
[66]	$\mathcal{L}_{LDAM}$	**.8077**±.118	.4977	.5851±.202	.0428	.4399±.193	.0219
[20]	$\mathcal{L}_{LMCE}$	.7975±.111	.6916	.6070±.209	.4052	.4650±.204	.2541
[42]	$\mathcal{L}_{DCE+VL}$	.7693±.116	.7345	.5681±.214	.3903	.4254±.201	.2425
[67]	$\mathcal{L}_{DSC++}$	.7939±.110	.6353	.5822±.203	.2605	.4375±.197	.1497
Ours	$\mathcal{L}_{LMFCL}$	.8013±.104	.7827	**.6318**±.201	**.4318**	**.4900**±.201	**.2753**

of segmentations, especially in complementary tattoo classes. This will be detailed in Section VII.

B. OPEN-SET SEMANTIC SEGMENTATION

This section evaluates the impact on the performance of the proposed $\mathcal{L}_{LMFCL}$ in an open-set scenario. Table 10 compares the open-set segmentation results obtained using various loss functions, all evaluated using the proposed G-OpenIPCS method. Among the loss functions, $\mathcal{L}_{LMFCL}$ achieves the highest AUROC score of 0.8013, Macro F1 with 0.6318, and also outperforms the mIoU score with 0.4900. These results indicate that $\mathcal{L}_{LMFCL}$ is the most effective loss function in terms of overall performance. Considering only UUC, although $\mathcal{L}_{CE}$ has the best performance in terms of AUROC (UUC). $\mathcal{L}_{LMFCL}$ remains competitive in this metric while providing superior performance in terms of the F1-Score (UUC) and mIoU (UUC).

Table 11 presents an evaluation of class-level performance from open-set semantic segmentation results from proposed $\mathcal{L}_{LMFCL}$. The top-5 best practically remained the same as the closed-set results, which is natural given the dependence of the closed-set model of the proposed approach on open-set recognition. However, the 'star' class was an exception, replacing the 'snake' class among the top 5 in terms of IoU. This is probably due to the low similarity of the 'star' class to the other semantic classes in TSSD2023, which practically maintained the performance obtained in the closed set. Classes that show more significant similarity to other classes, regardless of whether a KKC or a UUC, become more challenging for open-set segmentation, as they generate considerable uncertainty for the segmentation model.

This case of declining performance owing to class similarity is also evident among the top 5 worst classes. Classes such as 'leaf', 'water', 'knife', and 'fire' remained in this category, underscoring their significant reliance on the closed-set model. However, there was an exception in the 'wolf' class, which, despite its poor performance in the closed set, experienced a deterioration after open-set segmentation. This decline is attributed to its semantic similarity with other KKC animals, such as 'dog' and 'fox', and its resemblance to UUCs, particularly in bear tattoos. This provides misclassifications, as depicted in Figure 2.

TABLE 11. Open-set segmentation results are separated by classes obtained from the SegFormer trained using the proposed $\mathcal{L}_{LMFCL}$. All results are obtained on the proposed G-OpenIPCS method. The results were sorted in descending order of mIoU↑ values. The UUC has been highlighted in yellow, while the top-5 best classes are highlighted in green, and the top-5 worst classes are highlighted in red.

Class	AUROC	Macro F1	mIoU ↑
background	0.9025	0.9643	0.9310
tiger	0.9599	0.8828	0.7901
key	0.9118	0.8589	0.7527
octopus	0.8978	0.8417	0.7267
owl	0.8777	0.8205	0.6956
star	0.8594	0.8188	0.6932
snake	0.9177	0.8080	0.6779
butterfly	0.8714	0.8064	0.6756
dog	0.8584	0.7690	0.6247
scorpion	0.8940	0.7424	0.5903
diamond	0.8420	0.7195	0.5619
lion	0.8449	0.7192	0.5615
fox	0.8279	0.7124	0.5533
flower	0.7971	0.7043	0.5435
fish	0.8221	0.6991	0.5374
gun	0.8430	0.6972	0.5351
cat	0.7758	0.6965	0.5344
shark	0.8604	0.6798	0.5150
mermaid	0.9158	0.6754	0.5099
bird	0.7631	0.6544	0.4863
anchor	0.8075	0.6431	0.4740
eagle	0.8780	0.5941	0.4226
skull	0.7502	0.5733	0.4019
shield	0.7063	0.5429	0.3725
spide	0.8202	0.5420	0.3718
ribbon	0.7594	0.5402	0.3700
crown	0.7121	0.4761	0.3124
heart	0.6827	0.4573	0.2964
unknown	**0.7827**	**0.4318**	**0.2753**
wolf	0.7600	0.4178	0.2641
leaf	0.6462	0.4068	0.2553
water	0.6416	0.4017	0.2513
knife	0.5426	0.1322	0.0708
fire	0.5134	0.0520	0.0267
all	0.8013	0.6318	0.4900

Furthermore, KKCs can face challenges due to the high semantic variability exhibited within tattoos of the same class, arising from the infinite possibilities for representing an object through tattoos. For illustration, consider the representation of a 'cat' in various styles, such as cartoonish, realistic, stick-figure, geometric, and so on. It becomes exceedingly challenging to group all these diverse representations to

TABLE 12. Comparison between the proposed approach and state-of-the-art open-set semantic segmentation techniques. Bold values indicate the best overall results, including all methods.

Ref.	Loss	Network	Open Set Classifier	AUROC	Macro F1	mIoU	IoU UUC
Baseline	$\mathcal{L}_{CE}$	SegFormer	SoftMax-T	.7855±.126	.5782±.229	.4398±.216	.1350
Baseline	$\mathcal{L}_{CE}$	SegFormer	OpenMax	.7785±.123	.5781±.227	.4389±.212	.1611
Baseline	$\mathcal{L}_{CE}$	Swin+UPerNet	SoftMax-T	.7234±.095	.4852±.198	.3428±.188	.1145
Baseline	$\mathcal{L}_{CE}$	Swin+UPerNet	OpenMax	.7134±.095	.4736±.197	.3325±.181	.1258
[38]	$\mathcal{L}_{CE}$	DN121+FCN	OpenIPCS	.6523±.098	.3596±.217	.2422±.183	.0632
[38]	$\mathcal{L}_{CE}$	WRN50+FCN	OpenIPCS	.6695±.098	.3983±.214	.2722±.185	.0470
[42]	$\mathcal{L}_{DCE+VL}$	DRN50+PSPNet	APM	.7123±.139	.4290±.284	.3152±.244	.1657
[39]	$\mathcal{L}_{CE}$	RN101+PSPNet	MaxLogit	.7358±.139	.5173±.293	.4031±.268	.1439
[43]	$\mathcal{L}_{CE+PC}$	RN50+DeepLabV3+	MaxLogit	.7048±.105	.4471±.211	.3126±.191	.1327
[43]	$\mathcal{L}_{CE+PC}$	RN50+DeepLabV3+	MSP	.7401±.097	.4694±.207	.3157±.190	.1354
Ours	$\mathcal{L}_{LMFCL}$	SegFormer	G-OpenIPCS	**.8013**±.104	**.6318**±.201	**.4900**±.201	**.2753**
Ours	$\mathcal{L}_{LMFCL}$	Swin+UPerNet*	G-OpenIPCS	.7921±.116	.5927±.211	.4518±.213	.1981

* Evaluation of the generalization of the proposed method using Swin+UPerNet [62], [70].

ensure no shape is mistakenly isolated as an outlier and segmented as part of the unknown-unknown class (UUC). We endeavored to construct a robust data augmentation pipeline to address this limitation, as we believe it offers a potential solution (detailed in Section VII).

Regarding the UUC, the results obtained using $\mathcal{L}_{LMFCL}$ remain somewhat limited but show great promise. Notably, the UUC exhibited superior results in terms of IoU and F1-Score when compared to the top 5 worst classes. However, in terms of AUROC, 12 KKCs are surpassed.

In conclusion, utilizing the proposed $\mathcal{L}_{LMFCL}$ in conjunction with the G-OpenIPCS achieves the best overall results in open-set segmentation tasks. Its robust segmentation performance, capable of effectively distinguishing between UUCs and KKCs, makes it a promising choice for challenging scenarios with high semantic variability between classes. However, it is essential to note that there are still tough situations and areas where further improvements are needed in the context of open-set segmentation for TSSD2023.

C. COMPARISON WITH OTHERS OPEN-SET SEMANTIC SEGMENTATION METHODS

Table 12 comprehensively compares different state-of-the-art open-set semantic segmentation methods, including the proposed approach, and two more baselines: SoftMax-T and OpenMax. The proposed approach using $\mathcal{L}_{LMFCL}$ and G-OpenIPCS stands out as the best-performing method. It achieves the highest AUROC, Macro F1, mIoU, and IoU (UUC) values, indicating its superiority in open-set tattoo semantic segmentation. The scores obtained for AUROC of 0.8013, Macro F1 of 0.6318, mIoU of 0.4900, and IoU (UUC) of 0.2753 are significant compared to other approaches in the literature. However, the performance difference is less to the baselines, except for the values obtained from IoU to UUC. This indicates that the proposed approach produces more accurate and visually coherent segmentation results.

In comparison to the original OpenIPCS, it is essential to emphasize the strengths that make our approach superior. The original OpenIPCS relies on an FCN decoder that combines multiple layers from various levels of the network to construct the feature vector. In contrast, our G-OpenIPCS approach follows a more straightforward and intuitive path, considering that the critical features for class discrimination are primarily situated in the latter layers of the segmentation network. This approach leads to improved results and avoids the potential scrambling of high-level and low-level information that can occur when using multiple layers, as in the original OpenIPCS.

Due to this more streamlined design choice, our approach integrates with other modern segmentation architectures, such as transform-based structures. For instance, we employed the SegFormer and Swin+UperNet networks, outperforming FCN-based models in open-set tattoo semantic segmentation. Notably, the combination using the SegFormer outperforms other methods by a significant margin. This underscores that our design choice directly enhances performance, especially considering the high reliance on features in open-set segmentation models [42].

Moreover, it is worth noting that the original OpenIPCS is built upon $\mathcal{L}_{CE}$. While this loss function consistently delivers results in closed-set semantic segmentation, the challenge of open-set segmentation demands segmentation models with a heightened discriminative capacity, which is precisely what our proposed $\mathcal{L}_{LMFCL}$ aims to enhance.

D. ABLATION STUDIES OF THE DATA AUGMENTATION COMPONENTS

We performed an ablation study to examine the impact of various data augmentation techniques on our approach to open-set recognition for tattoo semantic segmentation. These techniques were divided into four categories, as explained in Section IV-A: geometric, image adjustments, dropout, and our proposed CSA method. We also included a no-augmentation experiment as a baseline for comparison.

The results in Table 13 indicate that our approach performs better when all data augmentation methods are

TABLE 13. Evaluation of the effect of data augmentation components. The results were obtained using proposed $\mathcal{L}_{LMFCL}$, SegFormer, and proposed G-OPenIPCS.

No Aug.	Geometric	Image Adjustments	Dropout	CSA (Ours)	mIoU	IoU $_{UUC}$
✓					0.4556±.186	0.1547
	✓				0.4320±.182	0.1797
		✓			0.4814±.203	0.1884
			✓		**0.4959**±.189	0.1538
				✓	0.4576±.203	0.1770
	✓	✓	✓		0.4733±.196	0.2303
	✓	✓	✓	✓	0.4900±.201	**0.2753**

combined. Among these methods, image adjustments consistently showed performance gains in mIoU and IoU for UUC. The other approaches, including geometric and proposed CSA, improved the values of IoU for UUC, while the dropout method improved the mIoU. Notably, the proposed CSA significantly improved performance when combined with the other three methods. In summary, our approach is appropriate for effective open-set tattoo semantic segmentation.

E. EVALUATION OF THE PERFORMANCE IN THE CLOSED-SET TATTOO SEGMENTATION SCENARIO

Semantic segmentation focuses on recognizing specific classes of tattoos, such as cats, dogs, stars, and others. Tattoo segmentation only aims to isolate the tattoo region from the background. This distinction is beneficial in situations that do not require detailed semantic information about tattoos. However, it may be interesting that our approach also behaves adequately in closed-set tattoo segmentation scenarios. Thus, we evaluated the performance of our proposed approach in a closed tattoo segmentation scenario using the DeMSI dataset [9], where tattoos were manually annotated without any semantic differentiation. In addition to measuring the IoU for tattoos, we also evaluate the False Positive Rate (FPR), which indicates the proportion of background pixels incorrectly classified as tattoos, and the False Negative Rate (FNR), which means the proportion of tattoo pixels incorrectly classified as background.

The quantitative results are presented in Table 14. We observed a degradation in the performance of the proposed approach when directly applying the model trained on TSSD2023 and converting the multiclass output to a binary output. This degradation is primarily reflected in the FNR values, indicating that many tattoo pixels were misclassified as background. We attribute this behavior to the labeling design used in TSSD2023. The primary factor is the detailed annotations in TSSD2023, designed to accurately classify the pixel classes, as seen in Table 4. For example, the shark's mouth was separated from the shadow of the tattoo. In contrast, DeMSI does not provide refined annotations for tattoos, leading to multiple background pixels being

TABLE 14. Evaluation of the performance on the DeMSI dataset in the closed-set tattoo segmentation scenario. The results were obtained using the proposed $\mathcal{L}_{LMFCL}$ and SegFormer.

Fine-tuned	IoU $_{Tattoo}$	FPR	FNR
	0.3878	0.0234	0.5784
✓	0.8232	0.0278	0.0945

TABLE 15. Visual result samples of closed-set tattoo segmentation on the DeMSI dataset. The tattoo and background pixels are depicted as black and white, respectively. The results were obtained using the proposed $\mathcal{L}_{LMFCL}$ and SegFormer.
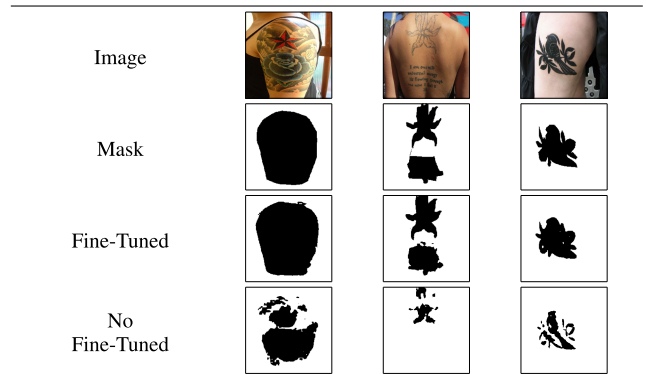
incorrectly annotated as part of a tattoo, as shown in Table 15. The second factor relates to tattoo classes with insufficient semantic samples, such as dragons, letters, and spider webs, which were excluded in TSSD2023 and are ignored by our approach, as seen by the text tattoo in Table 15. Due to these differences in annotation design between DeMSI and TSSD2023 datasets, performance degradation occurs when directly applying our approach.

To alleviate the performance degradation caused by annotation differences between the datasets, we fine-tuned the SegFormer network's last layer using the DeMSI dataset. For the fit, was used 60% of the images for training and 40% for testing as proposed by [13]. After this adjustment, our approach proved effective for performing closed-set tattoo segmentation, as evidenced by the significantly reduced FNR values in Table 14. Visually, the improvement in segmentation quality before and after fine-tuning can be observed in Table 15.

VI. CONCLUSION

This paper built a novo tattoo semantic segmentation dataset called TSSD2023, introducing an unexplored and challenging problem in semantic tattoo recognition. This dataset can serve as a valuable basis for future research.

Furthermore, the paper has presented the proposed Large-Margin Focal Loss ($\mathcal{L}_{LMFCL}$) to enhance tattoo semantic segmentation outcomes in both closed and open-set scenarios. In the closed-set scenario, $\mathcal{L}_{LMFCL}$ performed competitively and outperformed other evaluated loss functions in terms of Macro F1, demonstrating its suitability for closed-set semantic segmentation. An in-depth analysis at the class level revealed that superior performance is generally observed in classes depicted by high semantic dissimilarity and minimal overlap with other tattoos within an image. Conversely, the most challenging classes rarely appear in isolation, often serving as complementary components to other tattoos. In the open-set scenario, this paper proposes a generalized approach for the OpenIPCS method named G-OpenIPCS that facilitates the integration of this open-set classifier with more modern segmentation network architectures, such as transform-based networks. Using G-OpenIPCS, we compared the performance of different loss functions with the proposed $\mathcal{L}_{LMFCL}$, which showed higher scores regarding AUROC, Macro F1, and overall mIoU. When considering only the UUCs, $\mathcal{L}_{LMFCL}$ performed competitively in AUROC while outperforming F1-Score and IoU. This highlights its effectiveness in the domain of open-set tattoo semantic segmentation. A more thorough examination revealed that classes bearing high similarity to other classes, overlapped, and with limited samples, it represented the significant challenges for open-set tattoo semantic segmentation.

It is worth noting that when a semantic tattoo class has high variability, segmentation errors can occur due to false outliers. To address this issue, we propose a new data augmentation technique named Class Semantic Augmentation (CSA) that increases the available semantic information of classes for better model generalization. Nevertheless, there are opportunities for further advancements in this research area.

When comparing our approach, which combines $\mathcal{L}_{LMFCL}$, SegFormer, and G-OpenIPCS, with other state-of-the-art methods for open-set tattoo semantic segmentation, our approach consistently attains the highest overall results. It surpasses the performance of other state-of-the-art methods, underscoring its potential for addressing challenging scenarios in open-set semantic tattoo segmentation.

Furthermore, our approach applicability extends beyond the domain of tattoo segmentation, and it holds the potential to enhance open-set semantic segmentation in various other application domains. Finally, it is essential to note that while our approach demonstrates promise, there remains room for further improvements in specific areas. This includes improving the handling of class imbalance, class overlap, high intra-class semantic variability, and, in some cases, high inter-class similarity, detailed as follows.

VII. CHALLENGES IN OPEN SET TATTOO SEMANTIC SEGMENTATION

Although the approach discussed here is promising and with results superior to state-of-the-art methods for open-set semantic segmentation, in the context of tattoos, we can observe numerous difficulties and open points for this line of research. Below, we detail some of those that we consider most relevant in the proposed context:

A. SEMANTIC SHIFT

Semantic shift is caused by classes that influence model predictions due to their semantic similarity to other classes, causing segmentation errors [33], [37]. This semantic shift can be caused in three different ways in the tattoo dataset. The first is the presence of invisible objects in the background class that are similar to known and unknown classes. The background class is a large grouping of objects relatively irrelevant to the application. However, tattoos are semantically very similar to illustrations projected onto different surfaces. For example, a skull on a tattoo artist's chair can be easily identified as a tattoo, generating a misclassification. Furthermore, due to the great difficulty in annotating a large volume of tattoos for semantic segmentation, some tattoos are "ignored" in the labeling process and noted as belonging to the background class, which can cause misclassifications.

The second comprises objects of known classes are similar to each other. Some semantic classes are similar to others within the set of known classes. For example, the 'wolf', 'dog', and 'fox' classes belong to the same animal family, thus presenting similar characteristics that even humans may find difficult to distinguish. This high similarity makes the process of recognition by the segmentation model difficult, which can, in some cases, generate classification errors. Furthermore, this semantic similarity can extend to small parts of the tattoos, which can cause small segmentation errors. Finally, the third includes objects of unknown classes similar to known classes. This challenge is similar to the previous one but more demanding. The segmentation model must be trained to produce robust decision boundaries to separate known data from each other and separate it from unknown data. While in the previous problem, it is only necessary to distinguish among known classes.

A possible way to alleviate these factors' influence is to make a descriptive note of the tattoo using natural language. This can help identify tattoos more effectively, avoiding reducing the artistic feature of a tattoo to a set of labels.

B. INTRA-CLASS VARIABILITY

An open challange is also related to objects from known classes with different characteristics from the group of objects from the same semantic class. Because tattoos are a form of artistic expression based on drawings, the number of

designs that can be thought of to create a tattoo is practically unlimited. This infinite universe of possibilities makes each tattoo unique, with the variability of tattoos present in the real world being practically immeasurable. When defining a semantic class for a set of objects, we intuitively state that these objects are similar, which is generally true. However, concerning tattoos, it is possible that despite belonging to the same semantic class, practically none of the patterns are identically replicated between tattoos. In some cases, the patterns are so different between objects of the known class that the segmentation model segments the known tattoo as belonging to the unknown class.

C. OBJECTS OF KNOWN AND UNKNOWN CLASSES SEGMENTED WITH LOW PRECISION

This challenge incorporates several subproblems encountered in closed-set segmentation and also in open-set segmentation. In segmenting closed sets, the models suffer from the subproblem of overlap between objects, making it difficult to separate the boundaries between objects accurately. Another common subproblem is small objects present in images, generally classified as objects from other classes, such as the background class. The open-set semantic segmentation task inherits all these subproblems. However, in open-set segmentation, we still have the challenge of accurately classifying regions of completely unknown objects while dealing with the subproblems derived from closed-set segmentation.

D. DATA AUGMENTATION FOR SEMANTIC SEGMENTATION

The data augmentation proposal presented here may be promising for incorporating more data into tattoo datasets, the annotation of which is costly and time-consuming. A possible path is to combine the approach presented in [13] with the data augmentation ideas proposed in this paper in order to create greater variability of classes and examples, including data in an open set, to enable the training of more complex models and, consequently, increase final performance.

E. COMPUTATIONAL COMPLEXITY

We understand that a detailed analysis of computational complexity must be conducted for certain applications, mainly involving embedded systems. This analysis must include the training and deployment of the model according to the target device. Specific architectures for this type of application can also be evaluated, as demonstrated in [71].

REFERENCES

[1] S. T. Acton and A. Rossi, "Matching and retrieval of tattoo images: Active contour CBIR and glocal image features," in *Proc. IEEE Southwest Symp. Image Anal. Interpretation*, Mar. 2008, pp. 21–24.

[2] T. Harbert, "FBI wants better automated image analysis for tattoos [news]," *IEEE Spectr.*, vol. 52, no. 9, pp. 13–16, Sep. 2015.

[3] J.-E. Lee, R. Jin, and A. K. Jain, "Rank-based distance metric learning: An application to image retrieval," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[4] S. Fang, J. Coverdale, P. Nguyen, and M. Gordon, "Tattoo recognition in screening for victims of human trafficking," *J. Nervous Mental Disease*, vol. 206, no. 10, pp. 824–827, 2018.

[5] F. Bacchini and L. Lorusso, "A tattoo is not a face. Ethical aspects of tattoo-based biometrics," *J. Inf., Commun. Ethics Soc.*, vol. 16, no. 2, pp. 110–122, May 2018.

[6] R. T. da Silva and H. Silvério Lopes, "A transfer learning approach for the tattoo classification problem," in *Proc. IEEE Latin Amer. Conf. Comput. Intell. (LA-CCI)*, Nov. 2022, pp. 1–6.

[7] Y. Mo, Y. Wu, X. Yang, F. Liu, and Y. Liao, "Review the state-of-the-art technologies of semantic segmentation based on deep learning," *Neurocomputing*, vol. 493, pp. 626–646, Jul. 2022.

[8] C. Jiawang and Z. Yuan, "Tattoo recognition based on triplet GAN," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 9595–9597.

[9] T. Hrkać, K. Brkić, and Z. Kalafatić, "Tattoo detection for soft biometric de-identification based on convolutional neural networks," in *Proc. OAGM & ARW Joint Workshop*, 2016, pp. 1–23.

[10] J. Dong, X. Qu, and H. Li, "Color tattoo segmentation based on skin color space and K-mean clustering," in *Proc. 4th Int. Conf. Inf., Cybern. Comput. Social Syst. (ICCSS)*, Jul. 2017, pp. 53–56.

[11] M. Nicolás-Díaz, A. Morales-González, and H. Méndez-Vázquez, "Weighted average pooling of deep features for tattoo identification," *Multimedia Tools Appl.*, vol. 81, no. 18, pp. 25853–25875, Jul. 2022.

[12] M. Ngan and P. Grother, "Tattoo recognition technology–challenge (Tatt-C): An open tattoo database for developing tattoo recognition research," in *Proc. IEEE Int. Conf. Identity, Secur. Behav. Anal.*, Mar. 2015, pp. 1–6.

[13] L. J. Gonzalez-Soler, C. Rathgeb, and D. Fischer, "Semi-synthetic data generation for tattoo segmentation," in *Proc. 11th Int. Workshop Biometrics Forensics (IWBF)*, Apr. 2023, pp. 1–6.

[14] C. Geng, S.-J. Huang, and S. Chen, "Recent advances in open set recognition: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3614–3631, Oct. 2021.

[15] C. C. V. da Silva, K. Nogueira, H. N. Oliveira, and J. A. d. Santos, "Towards open-set semantic segmentation of aerial images," in *Proc. IEEE Latin Amer. GRSS ISPRS Remote Sens. Conf. (LAGIRS)*, Mar. 2020, pp. 16–21.

[16] I. Nunes, M. B. Pereira, H. Oliveira, J. A. dos Santos, and M. Poggi, "Conditional reconstruction for open-set semantic segmentation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2022, pp. 946–950.

[17] M. Gutoski, A. E. Lazzaretti, and H. S. Lopes, "Deep metric learning for open-set human action recognition in videos," *Neural Comput. Appl.*, vol. 33, no. 4, pp. 1207–1220, Feb. 2021.

[18] D. Miller, N. Sunderhauf, M. Milford, and F. Dayoub, "Class anchor clustering: A loss for distance-based open set recognition," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 3569–3577.

[19] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-margin softmax loss for convolutional neural networks," 2016, *arXiv:1612.02295*.

[20] T. Kobayashi, "Large margin in softmax cross-entropy loss," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, 2019, pp. 1–20.

[21] X. Li, D. Chang, T. Tian, and J. Cao, "Large-margin regularized softmax cross-entropy loss," *IEEE Access*, vol. 7, pp. 19572–19578, 2019.

[22] Q. Xu, S. Ghosh, X. Xu, Y. Huang, and A. W. K. Kong, "Tattoo detection based on CNN and remarks on the NIST database," in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2016, pp. 1–7.

[23] H. Han, J. Li, A. K. Jain, S. Shan, and X. Chen, "Tattoo image search at scale: Joint detection and compact representation learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 10, pp. 2333–2348, Oct. 2019.

[24] M. Nicolás-Díaz, A. Morales-González, and H. Méndez-Vázquez, "Deep generic features for tattoo identification," in *Proc. 24th Iberoamerican Congr. Prog. Pattern Recognit., Image Anal., Comput. Vis., Appl.*, 2019, pp. 272–282.

[25] A. Jain, Y. Chen, and U. Park, "Scars, marks & tattoos (SMT): Physical attributes for person identification," Dept. Comput. Sci., Michigan State Univ., East Lansing, MI, USA, 2007.

[26] A. K. Jain, J.-E. Lee, and R. Jin, "Tattoo-ID: Automatic tattoo image retrieval for suspect and victim identification," in *Advances in Multimedia Information Processing*. Springer, 2007, pp. 256–265.

[27] J. D. Allen, N. Zhao, J. Yuan, and X. Liu, "Unsupervised tattoo segmentation combining bottom-up and top-down cues," in *SPIE Proc.*, May 2011, pp. 1–7.

[28] B. Heflin, W. Scheirer, and T. E. Boult, "Detecting and classifying scars, marks, and tattoos found in the wild," in *Proc. IEEE 5th Int. Conf. Biometrics: Theory, Appl. Syst. (BTAS)*, Sep. 2012, pp. 31–38.

[29] P. Duangphasuk and W. Kurutach, "Tattoo skin detection and segmentation using image negative method," in *Proc. 13th Int. Symp. Commun. Inf. Technol. (ISCIT)*, Sep. 2013, pp. 354–359.
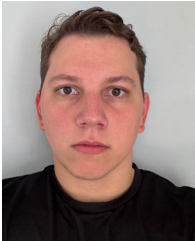
[30] J. Kim, A. Parra, H. Li, and E. J. Delp, "Efficient graph-cut tattoo segmentation," in *Visual Information Processing and Communication*. Bellingham, WA, USA: SPIE, 2015, pp. 1–8.

[31] T. Hrkac, K. Brkic, S. Ribaric, and D. Marcetic, "Deep learning architectures for tattoo detection and de-identification," in *Proc. 1st Int. Workshop Sens., Process. Learn. Intell. Mach. (SPLINE)*, Jul. 2016, pp. 1–5.

[32] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[33] A. Brilhador, A. E. Lazzaretti, and H. S. Lopes, "A comparative study for open set semantic segmentation methods," in *Anais Do 15 Congresso Brasileiro De Inteligencia Computacional*. Maharashtra, India: SBIC, 2021, pp. 1–8.

[34] I. Nunes, C. Laranjeira, H. Oliveira, and J. A. dos Santos, "A systematic review on open-set segmentation," *Comput. Graph.*, vol. 115, pp. 296–308, Oct. 2023.

[35] Y. Tian, D. Su, S. Lauria, and X. Liu, "Recent advances on loss functions in deep learning for computer vision," *Neurocomputing*, vol. 497, pp. 129–158, Aug. 2022.

[36] Y. Liu, Y. Tang, L. Zhang, L. Liu, M. Song, K. Gong, Y. Peng, J. Hou, and T. Jiang, "Hyperspectral open set classification with unknown classes rejection towards deep networks," *Int. J. Remote Sens.*, vol. 41, no. 16, pp. 6355–6383, Aug. 2020.

[37] Z. Cui, W. Longshi, and R. Wang, "Open set semantic segmentation with statistical test and adaptive threshold," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2020, pp. 1–6.

[38] H. Oliveira, C. Silva, G. L. S. Machado, K. Nogueira, and J. A. dos Santos, "Fully convolutional open set segmentation," *Mach. Learn.*, vol. 112, no. 5, pp. 1733–1784, May 2023.

[39] D. Hendrycks, S. Basart, M. Mazeika, A. Zou, J. Kwon, M. Mostajabi, J. Steinhardt, and D. Song, "Scaling out-of-distribution detection for real-world settings," in *Proc. 39th Int. Conf. Mach. Learn. (ICML)*, 2022, pp. 8759–8773.

[40] M. Grcić, P. Bevandić, and S. Šegvić, "DenseHybrid: Hybrid anomaly detection for dense open-set recognition," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 500–517.

[41] X. Guo, J. Liu, T. Liu, and Y. Yuan, "Handling open-set noise and novel target recognition in domain adaptive semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 1–16, Oct. 2023.

[42] J. Cen, P. Yun, J. Cai, M. Yu Wang, and M. Liu, "Deep metric learning for open world semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 15313–15322.

[43] J. Hong, W. Li, J. Han, J. Zheng, P. Fang, M. Harandi, and L. Petersson, "GOSS: Towards generalized open-set semantic segmentation," *Vis. Comput.*, vol. 40, no. 4, pp. 2391–2404, Apr. 2024.

[44] H. Zhang and H. Ding, "Prototypical matching and open set rejection for zero-shot semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 6954–6963.

[45] B. Yu, T. Liu, M. Gong, C. Ding, and D. Tao, "Correcting the triplet selection bias for triplet loss," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 71–87.

[46] S. Nandan Rai, F. Cermelli, B. Caputo, and C. Masone, "Mask2Anomaly: Mask transformer for universal open-set segmentation," 2023, *arXiv:2309.04573*.

[47] M. J. Wilber, E. Rudd, B. Heflin, Y.-M. Lui, and T. E. Boult, "Exemplar codes for facial attributes and tattoo recognition," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2014, pp. 205–212.

[48] W. He, J. Wang, L. Wang, R. Pan, and W. Gao, "A semantic segmentation algorithm for fashion images based on modified mask RCNN," *Multimedia Tools Appl.*, vol. 82, no. 18, pp. 28427–28444, Jul. 2023.

[49] W. Wan, Y. Zhong, T. Li, and J. Chen, "Rethinking feature distribution for loss functions in image classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9117–9126.

[50] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," *J. Mach. Learn. Res.*, vol. 2, no. 1, pp. 265–292, 2001.

[51] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[52] H.-M. Yang, X.-Y. Zhang, F. Yin, Q. Yang, and C.-L. Liu, "Convolutional prototype network for open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 5, pp. 2358–2370, May 2022.

[53] X. Sun, Z. Yang, C. Zhang, K.-V. Ling, and G. Peng, "Conditional Gaussian distribution learning for open set recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 13477–13486.

[54] R. Shwartz-Ziv and N. Tishby, "Opening the black box of deep neural networks via information," 2017, *arXiv:1703.00810*.

[55] R. Kumar Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," 2015, *arXiv:1505.00387*.

[56] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*. Honolulu, USA: IEEE, Jul. 2017, pp. 2261–2269.

[57] M. E. Tipping and C. M. Bishop, "Mixtures of probabilistic principal component analyzers," *Neural Comput.*, vol. 11, no. 2, pp. 443–482, Feb. 1999.

[58] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," 2017, *arXiv:1708.04552*.

[59] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[60] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo, and J. Choe, "CutMix: Regularization strategy to train strong classifiers with localizable features," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6022–6031.

[61] S. Zheng, J. Lu, H. Zhao, X. Zhu, Z. Luo, Y. Wang, Y. Fu, J. Feng, T. Xiang, P. H. S. Torr, and L. Zhang, "Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 6877–6886.

[62] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 9992–10002.

[63] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "SegFormer: Simple and efficient design for semantic segmentation with transformers," in *Proc. Adv. Neural Inf. Process. Sys. (NIPS)*, vol. 34, Dec. 2021, pp. 12077–12090.

[64] T. Dozat, "Incorporating Nesterov momentum into Adam," in *Proc. 4th Int. Conf. Learn. Represent.*, 2016, pp. 1–4.

[65] H.-M. Yang, X.-Y. Zhang, F. Yin, and C.-L. Liu, "Robust classification with convolutional prototype learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3474–3482.

[66] K. Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma, "Learning imbalanced datasets with label-distribution-aware margin loss," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–20.

[67] M. Yeung, L. Rundo, Y. Nan, E. Sala, C.-B. Schönlieb, and G. Yang, "Calibrating the dice loss to handle neural network overconfidence for biomedical image segmentation," *J. Digit. Imag.*, vol. 36, no. 2, pp. 739–752, Dec. 2022.

[68] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1757–1772, Jul. 2013.

[69] A. Bendale and T. E. Boult, "Towards open set deep networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1563–1572.

[70] Z. Wang, J. Li, Z. Tan, X. Liu, and M. Li, "Swin-UperNet: A semantic segmentation model for mangroves and spartina alterniflora loisel based on UperNet," *Electronics*, vol. 12, no. 5, p. 1111, Feb. 2023.

[71] B. Olimov, J. Kim, and A. Paul, "REF-Net: Robust, efficient, and fast network for semantic segmentation applications using devices with limited computational resources," *IEEE Access*, vol. 9, pp. 15084–15098, 2021.

ANDERSON BRILHADOR received the M.Sc. degree in computer science from the Federal University of Technology—Paraná (UTFPR), Brazil, in 2015, where he is currently pursuing the Ph.D. degree in electrical engineering and industrial informatics. He is a Professor with the Computer Science Program, UTFPR. His research interests include computer vision, machine learning, deep learning, and data mining.

RODRIGO TCHALSKI DA SILVA received the M.Sc. degree in industrial computing from the Federal University of Technology—Paraná (UTFPR), Brazil, in 2022, where he is currently pursuing the Ph.D. degree in industrial computing. His research interests include computer vision, tattoo recognition, machine learning, and complex networks.
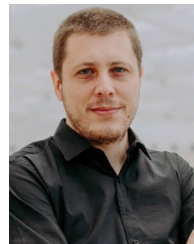
HEITOR SILVÉRIO LOPES received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Technology—Paraná (UTFPR), Curitiba, in 1984 and 1990, respectively, and the Ph.D. degree from the Federal University of Santa Catarina, in 1996. Currently, he is a tenured Full Professor with the Department of Electronics and the Graduate Program on Electrical Engineering and Applied Computer Science (CPGEI), UTFPR. His research interests include computer vision, deep learning, evolutionary computation, and data mining.

CARLOS ROBERTO MODINEZ-JUNIOR is currently pursuing the degree in electronic engineering with the Federal University of Technology—Paraná (UTFPR), Brazil. He has been working with image processing. His research interest includes semantic segmentation.

GABRIEL DE ALMEIDA SPADAFORA is currently pursuing the degree in computer engineering with the Federal University of Technology—Paraná (UTFPR), Brazil. His research interests include image processing and neural networks.

ANDRÉ EUGÊNIO LAZZARETTI (Member, IEEE) received the B.Sc., M.Sc., and D.Sc. degrees in electrical engineering from the Federal University of Technology—Paraná, in 2007, 2010, and 2015, respectively. He is currently a Professor with the Department of Electronics, Federal University of Technology—Paraná. His research interests include machine learning, deep learning, and digital signal processing.

● ● ●