Universiteit Leiden

Master Computer Science

TRANSITION - Extraction of color features in Odonates with computer vision

Name:	Megan Mirnalini Sundaram Rajaraman
Student ID:	4018907
Date:	[23/01/2026]
Specialisation:	Artificial Intelligence
1st supervisor:	Rita Pucci
2nd supervisor:	Vincent Kalkman
3nd supervisor:	Fons J. Verbeek

Master's Thesis in Computer Science

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Contents

1 Introduction **1**

2 Related Work **3**
2.1 Annotation Tools and Dataset Preparation . . . 3
2.2 Instance and Semantic Segmentation . . . 4
2.3 Color Extraction . . . 5
2.4 Additional Reading on Odonates and *Sympetrum striolatum* . . . 6

3 Methodologies **7**
3.1 Dataset Acquisition and Curation . . . 7
3.1.1 Analysis of the Dataset . . . 7
3.1.2 Distribution of the dataset by country . . . 8
3.1.3 Distribution of the observed species - *Sympetrum striolatum* . . . 10
3.1.4 Annotation . . . 12
3.2 Models Used for Segmentation . . . 13
3.2.1 You Only Look Once - YOLO . . . 13
3.2.2 Mask R-CNN . . . 15
3.2.3 MaskDINO . . . 16
3.2.4 Mask2Former . . . 17
3.3 Color Extraction . . . 18
3.3.1 K-Means Clustering . . . 18
3.3.2 Extraction of HSV values . . . 18
3.4 Zero-Shot Learning and Fine-tuning . . . 19
3.5 Creation of Survey and Reliability of Model . . . 19
3.6 Metrics Used for Evaluation . . . 20
3.6.1 Evaluation of Performance of Models . . . 20
3.6.2 Evaluation of Reliability . . . 21
3.6.3 Validation of hypothesis . . . 21

4 Experiments **23**
4.1 Research Pipeline . . . 23
4.2 Model Training and Performance . . . 24
4.3 Reliability of initial findings . . . 25
4.4 Statistical analysis . . . 25
4.5 Color Extraction . . . 25

5 Results **27**
5.1 Zero shot learning . . . 27
5.2 Fine Tuning Experiment 1 . . . 29
5.3 Fine Tuning Experiment 2 . . . 31
5.4 Reliability of Initial Findings . . . 33
5.4.1 Results from Entomologists . . . 33
5.4.2 Results from Computer Scientists . . . 34
5.5 Statistical Analysis . . . 38
5.6 Color Analysis . . . 40
5.6.1 Color Palette using K-Means Clustering . . . 40
5.6.2 Mean HSV extraction and statistical analysis . . . 42

6 Conclusion **48**
6.1 Discussions and Future Work . . . 48
6.2 Conclusion . . . 49

A Additional Information **51**
A.1 Confusion Matrix . . . 51

B Training Results **52**
B.1 Hyperparameter Tuning on YOLO . . . 52
B.2 Inference Images . . . 55
B.2.1 Initial Training . . . 55
B.2.2 Refined Results . . . 59

C Color Analysis **60**
C.1 Month-wise Analysis . . . 60

Abstract

In this thesis, we propose a pipeline to identify and segment Odonates (dragonfly and damselfly) using deep neural networks, and trained on images from citizen science data. The goal of the project was to robustly segment the body parts of each subject visible in the image into head, thorax, abdomen and wings. Some of the challenges included segmenting the small body parts, which were camouflaged with their surroundings and similar in appearance to each other. The additional goal of the project was to develop a color palette based on each part of the dragonfly, to understand color distribution from open-source datasets and enable interesting studies on the correlation between color and climate, habitat, and more. Furthermore, a vital part of this project was to establish a segmentation that can work with limited available data and generalize to other species too. Recent works on instance segmentation in computer vision show impressive abilities in segmenting previously unseen objects. However, these models are not yet suitable for highly specialized, fine grained tasks such as insect body part segmentation where annotated data is severely limited (< 100 examples). The starting point for this project was curating a dataset from open-source datasets from GBIF for the subject (Order Odonata - dragonflies and damselflies). After curating the dataset, various techniques to adapt the foundational models to the sparse dataset was investigated. Based on the techniques, the model was trained and tested for image classification and segmentation into background and the subject of interest (Odonates). After that, the model was adapted for semantic segmentation of the dragonfly into the head, thorax and abdomen. Finally, based on the segmentation, a color palette of the colors found in the segmented part was built, which was used for further analysis and study. This is particularly useful in biological and ecological studies where color changes are an indicator of various environmental and geographical factors.

Chapter 1

Introduction

Odonata is a small group of predatory insects that are found in all the continents except Antarctica. They are very common and easy to spot. The order *Odonata* comprises of multiple sub-species that include dragonflies and damselflies. Dragonflies are small winged insects that first originated in the early Jurassic period (201.3 - 174.7 million years ago) and still exist to this day. They are characterized by compound eyes (made up of thousands of ommatidia) [1], two pairs of strong wings and an elongated body. They are closely related to damselflies, and are very similar in appearance. The defining characteristic is that the wings are away from the body for the dragonfly, whereas for the damselfly, they are close to the abdomen [1].

Such creatures are now being threatened due to habitat destruction, and clearance of forests [2]. As they are tropical in nature, temperature changes affect their physiology greatly, namely their developmental rate, immune function and the development of pigment for thermo-regulation [3]. This, in turn, affects the flight and movement of Odonates. As they are predatory insects, this could also negatively influence the ecosystem by disturbing the natural balance of insect species. Hence, there exists an urgent need to understand the change in coloration and influence of environmental and geographical factors, on the coloration.

In this thesis, a new Computer Vision based pipeline to support the monitoring of Odonata is proposed. In addition, this thesis focuses on the body coloration analyzing and the differences in colors among the body parts of the insects.

Identifying and segmenting the body parts of the Odonate is a non-trivial challenge, and one of the tasks for this project. There exists multiple machine learning segmentation models that are trained to identify dogs, cats and other animals, including butterflies and birds. However, there exists very few models that are trained on insects or work on insects. This is because existing models struggle to identify and segment the insects, as they are extremely small and well-camouflaged. Hence, there is a need for models trained specifically on insects.

In addition, there are very few annotated datasets that can be used to identify and train them. Although there are digitalized collections of insect records from the museums all over the world, they cannot be used, primarily because they are of dead specimens. Although these are, in most cases, well-preserved, the overall color and pigmentation can fade with time. Therefore, it is vital that a dataset consisting of records of live insects, captured close to the geographical point of interest is used to maintain the validity to the study.

The approaches listed in this thesis aim to address some of the aforementioned gaps by utilizing data from citizen science publicly available online. The thesis can therefore be divided into three main cores: annotation and preparation of the dataset, instance and semantic segmentation of the object and extraction of color from said object.

The models used for this thesis are built and trained on citizen science data. Citizen science data refers to data that was generated and contributed by the general public. This data exists in public websites such as GBIF, BioPortal and are free to use. As they are typically captured by the general public from a phone camera, they also contain metadata such as the location, date as well as the time of capture. This information proves useful in analyzing the color patterns that exist in a specific location. These images are then verified by entomologists and other information such as species, order, epithet, gender and life stage are added. This information is vital, as males and females exhibit different pigmentation.

As the datasets are not prepared to facilitate the segmentation of the parts of the body for color extraction, one of the tasks for the project is the annotation of a functional dataset for instance and semantic segmentation [1]. The annotation of datasets performed by humans is time consuming but it is a crucial step for applied machine learning. As there are no available datasets that contain the Odonates and required parts, a significant part of this thesis project focused on manual annotation to train the models.

After manual annotation and the first round of experiments, the best trained model is used to increase the number of annotated images by autonomous annotation. In order to obtain an unbiased opinion about the accuracy and shortcoming of the trained model, a survey campaign was performed to evaluate the reliability of the model as an annotator.

After isolation of the required body parts, the final task of this project focuses on extraction of color from each identified part, and the statistical analysis of color distribution based on geolocalization and daytime and the correlation between them.

[1] Semantic and instance segmentation are di□erent segmentation tasks, which help in recognizing di□erent classes of the object and di□erent objects. This thesis deals with both instance and semantic segmentation, as identification of the dragonfly and isolation of the di□erent parts of the Odonates is required.

Chapter 2

Related Work

This section presents existing literature and key papers that were instrumental in development of the project.

2.1 Annotation Tools and Dataset Preparation

Annotation is required, as the model requires ground truth labels for training and validation. Accurate annotation is crucial as it helps distinguish di erent parts, as well as reduce errors in boundaries of masks. For this, various papers in multiple domains for segmentation were analyzed, to understand how annotation and dataset preparation was tackled. Typically, annotation is done manually, as done in Microsoft Common Objects in Context [4], which is a benchmark dataset for object detection. The dataset contains commonly found objects of 91 categories, and around 2.5 million instances. However, annotation was done by means of crowd-sourcing and then, evaluating with the help of expert annotators. Similarly, another benchmark dataset for pixel-level and instance-level semantic labeling, Cityscapes Dataset [5], was created in-house with multiple annotators and quality control agents using LabelMe tool [6]. Another large scale dataset is Mapillary Vistas Dataset [7]. This is a dataset of 25000 street-level images for instance-level semantic segmentation, and was manually annotated with 69 professional annotators, using an in-house tool for the process. Similarly, other datasets such as the MPII Multi-person Dataset [8], was annotated by in-house workers and through Amazon Mechanical Turk, or OpenPose [9], which uses the ClickWorker platform [10] for annotation.

The tools discussed in most of the above mentioned paper utilize crowd-sourcing as part of the annotation. Even if some of the tools commonly used, such as LabelMe [6] or FiftyOne [11] seemed feasible for annotation, these uses polygons and did not seem to provide the granularity required for capturing the wing curvature of the dragon y. Another tool that is commonly used in histology and pathology processes is QuPath [12], which is open-source, and has multiple tools like the polygon selection, Magic Wand, and even a Pytorch plugin to run models for annotation.

Another approach is annotating a small portion of the dataset and using models or external tools to annotate larger portions of the dataset, such as MassID45 [13], which is the approach chosen for the project.

2.2 Instance and Semantic Segmentation

Instance and semantic segmentation is done extensively in this project to identify and segment the parts of theOdonates Semantic segmentation is a computer vision task, which identi es the object and distinguishes the boundaries between the object and the background. Instance segmentation is a computer vision task that predicts the di erent parts of an identi ed object and draws clear boundaries between each part. In contrast, semantic segmentation assigns a single class to every pixel of the object. Like annotation, multiple papers were analysed in various domains, as there is

Figure 2.1: An example image illustrating the di erence between instance and semantic segmentation from the Robotic Manipulation book taught in MIT. [14]

very few literature that look at dragon ies (or insects) in general.

One of the main papers is Segment Anything [15], which proposed the SAM model and the corresponding Segment Anything dataset for training. This paper produced a foundation model that can adapt to both semantic and instance segmentation tasks.

There are multiple papers that address segmentation outside the ecological domain, such as SiNGR [16] or this paper which focus on segmentation and provide a comprehensive overview of models and techniques that are applicable for various functions in the medical domain [17]. Other papers such as OpenPose [9] deal with human detection, particularly with the recognition of poses and estimation of multiple instances of people in the same image, and PoseTed [18] which deal with recognition of poses of the feet. Survey papers such as this [19] provides insight on datasets and models that can be utilized for instance segmentation. There are also multiple papers for videos, such as VideoCUTLER [20] which explores instance segmentation in images, and Low-Resource challenges [21] that explores the application of low-resource settings for vision models. These papers helped to gain an overview of models that are at the state of the art in instance and semantic segmentation.

Despite recent advancements in semantic and instance segmentation, very few papers and current implementation focus on insects and animals. An example of one such paper is Fantastic Animals [22], which explores segmentation for marine animals using DualSAM. This paper deals with multiple re nements required to deal with the occlusion and lighting that marine images contain. This paper shows how important prior knowledge is in segmentation. Similarly, another key paper is Learning Part Segmentation from Synthetic Animals [23], which deals with semantic part segmentation in synthetic animals, and uses Skinned Multi-Animal Linear (SMAL) models for segmentation. There are some papers that look at insect segmentation, such as [24], which explores the use of an U-Net

model to count and segment the stink bug, and the paper discussed previously [13], which looks at multiple classes of insects.

2.3 Color Extraction

Color Models The nal part of the project is the extraction of color from the identi ed parts of the dragon y. Color information can be processed in multiple scales. The most commonly used scales are RGB, HSI/HSV and CMYK. In RGB color space, the image is represented as its spectral components on a Cartesian coordinates system and is a triplet of values (R, G, B). The RGB/CMYK model works well for hardware and model representations, but is not suited for the human interpretation. In order to interpret it easily, the HSV model was proposed. The H represents the hue of the object, and is typically used for a pure color. The S represents the saturation of the object, and denotes how much of the color has been diluted by white light. The V represents the Value of the object, and is an indicator of how much lightness/blackness is in the region. The higher the value, the brighter and closer to white or a fully saturated color it is, and the lower the value, the darker and closer to black it is. [25].

While these color models exists for representation and understanding of the color of the objects, color extraction is usually done in the hyperspectral space. This processes the color for each pixel on the electromagnetic spectrum, and is an accurate and preferred method of color extraction. This has been done in other elds such as astronomy for mapping lunar mineralogy [26], or in dentistry to map the color change in simulated teeth [27], [28]. Even within insects, hyperspectral imaging has been the primary tool of choice [29] to identify di erent species in insects [30], [31] or analyze development [32]. However, hyperspectral imaging is laborious, and requires an expensive hyperspectral camera as well as specimens to record the samples. As the goal of the project was to devise a method that works for large-scale, readily available datasets, and it is not feasible to record 300,000 specimens, another method was implemented.

When looking at how to analyse the color for insects and other organisms, biological models seem to be the popular choice. Color analysis done within insects such as butter ies [33] or this analysis [34] look at the colors and utilize biological models to extract and understand the di erence in colors. There has been extensive research done to map the colors in birds [35], which uses biological models to extract and analyse the color di erences with respect to the geographic location.

There are papers which utilize machine learning or code to process color information in CIELAB-space, such as this analysis on Papilionidae butter ies [36], or on birds [37].

There are few existing papers that look at color analysis and extraction on RGB or HSV/HSI scale. Such papers are on the green anole (Anolis carolinensis), which uses K-Means Clustering as a method to cluster the dominant colors in RGB space [38] and on ant color [39], which extracts the dominant HSV values for each species, and looks at the variation over the climate and location. Both these papers proved instrumental for the color extraction and analysis of this project.

2.4 Additional Reading on Odonates and Sympetrum striolatum

Dragon y belong to the infra-order Anisoptera, because the Greek word anisos mean unequal, and their hindwings are broader than their forewings. In contrast, Damsel ies belong to the infra-order Zygoptera, as it stands for paired-wings (meaning all the wings are equal in shape). As mentioned in chapter 1, dragon ies and damsel ies are predatory insects that are present in mostly tropical regions. They feed on a variety of insects ranging from mosquitoes to ies [40]. Their presence and feeding habits have greatly in uenced both aquatic and terrestrial ecosystems, and are good indicators of the quality of aquatic habitats [41].

They also exhibit color polymorphism (which means that there are two or more color forms (morphs) within the same species). This allows the males to `hide' from ghter males, by camou aging, while still being near females, and the females to avoid harassment [42]. This behavior is greatly in uenced and a ected by the temperature of the location. Higher temperatures a ect wing coloration, which a ects the ight and performance of the males [43]. In addition to this, there are also evidence to suggest that the color changes with respect to the latitude [3].

Chapter 3

Methodologies

This section introduces the methods and techniques that were used throughout the project. It also covers the annotation techniques and reasons for annotation, the models used in the project and a general overview of their architecture and losses, the metrics used for evaluating the performance of the models, and the metrics used for calculating the reliability of the results.

The methodologies can be divided into the three main cores (as mentioned earlier) - annotation and preparation of the dataset, segmentation of the object of interest (Odonates) and the extraction of color.

3.1 Dataset Acquisition and Curation

As the objective of the project was to use images that were collected from open source datasets and sites, the dataset was prepared from scratch using citizen science data. The dataset comprises of insects of the order *Odonata* observed in Europe. In order to extract the color from the Odonates, only adults i.e., *Imago* lifestage were considered. This also ensures that the wings are functional and fully developed. After filtering, the dataset consists of 759,423 records from 73 published datasets. The DOI of the dataset is given here [44].

After downloading the images, the dataset was further pruned. The records were filtered only for IDs that contain a corresponding image. This reduced the dataset to 254940. For proper segmentation of the dragonfly, it is important that the dragonfly is clearly visible. Hence, images that contain the full view of the dragonfly or nearly the full view were considered. To reduce the complexity, images that had more than one dragonfly were discarded. Images where the dragonfly was camouflaged or hardly visible were also discarded. Some samples of the discarded images are provided in Figure 3.1.

3.1.1 Analysis of the Dataset

Multiple analyses were done on the dataset to determine if there are any underlying patterns or themes. A simple elevation map of Europe is provided in Figure 3.2, as this will be referenced later on.

Figure 3.1: Examples of discarded images. In the first image, the dragonfly is out of focus and blurry. In the second image, the dragonfly is not clearly visible and well-camouflaged (it is sitting on the crops). The third image consists of a dragonfly held by a hand, with the thumb hiding the body of the dragonfly. Only the wings and part of the head is visible. These types of images are removed to the best of abilities from the dataset.

Figure 3.2: European Environment Agency (EEA). Elevation map of Europe (9 km grid) . Available at: https://sdi.eea.europa.eu/data/374f566b-6d4b-4a10-9c72-7d06579e724e.

3.1.2 Distribution of the dataset by country

Firstly, the dataset was plotted on a map of Europe to see if there are any visible patterns and to ensure it is distributed fairly. This was done by converting the dataset into a GeoDataFrame. This

was done with the help of a python package called GeoPandas, which is a subclass of Pandas, and supports geographical data. This was also made simpler by the fact that all of the images in the dataset had the decimalLatitude and decimalLongitude of the image tied to it.

Using this, the dataset was reformatted to a GeoDataFrame and plotted as points on top of a map of Europe. The plot is provided in Figure 3.3.

Figure 3.3: Distribution of records of the dataset in Europe. The underlying map is a map of EU in 2016, which includes the UK. Each point indicates an occurrence of the Odonate at that specific latitude/longitude.

From the plot, it is clear that the dataset has a high portion of occurrences in three countries - Belgium, the Netherlands and Germany. This is understandable as they are lower in elevation and relatively flat in topography (as seen from Figure 3.2). The occurrences are lower in the borders of France, Switzerland and Italy, which is also understandable, as the Alps are present there and are higher in elevation. Similarly, there also seems to be a high percentage of observation in Sweden, particularly in the coast. This corresponds to the areas of lower elevation (under 162m). It is also seen that there seems to proportionally higher observations in the coastal areas (where the elevation is much lower) and borders.

3.1.3 Distribution of the observed species - Sympetrum striolatum

After the pointplot of the entire dataset was generated, the occurrences of the species of interest - Sympetrum striolatum was examined. The distribution was plotted throughout Europe in a similar fashion (as that of Figure 3.3).

Figure 3.4: Distribution of observations of Sympetrum striolatum in Europe. The underlying map is a map of EU in 2016, which includes the UK. Each point indicates an occurrence of the species at that speci c latitude/longitude.

As seen from Figure 3.4, the species is found in abundance in the Netherlands, with a fair count of observations in Belgium and Germany. There seems to be a few observations in Sweden as well, but only in the lower parts of Sweden, where the climate and topography is similar to that of the Netherlands.

When the occurrence is grouped by gender (plotted in Figure 3.5), it is observed the there seems to be nearly equal distributions in all other countries. However, in the Netherlands, there seems to be a larger female population (8113) which accounts for 53% of the total population, while the male population accounts for 47% (7076).

Figure 3.5: A countplot of the observed species - Sympetrum striolatum throughout Europe and grouped by gender. The males are represented by the teal color and the females are represented by the orange color. The graph shows very high counts for Netherlands, and a relatively lower count for Germany and Sweden. Other countries have a population distribution in the single digits

As the occurrences are much higher in the Netherlands, further analysis was done locally (for the Netherlands) which is explained in detail in Section 5.5. However, a count-plot was done for the observations by year. This is shown in Figure 3.6.

Figure 3.6: A countplot of the observed species - Sympetrum striolatum by year and gender, in the Netherlands. The counts are grouped by gender, with the males represented by the teal color and the females represented by the orange color. The graph shows a gradual upward trend till 2018 and a steep increase from that year onward.

As seen from the gure, there do not seem to be observations of Sympetrum striolatum before the year 2000. The rst decade seems to be a gradual increase in the observations. However, after 2019,

there seems to be a migration towards the Netherlands, as the observations seem to have nearly doubled for the female population (531 in 2018 followed by 852 in 2019). Consecutively, there seems to be a growth in the male population as well in the next year (596 in 2019 followed by 872 in 2020).

3.1.4 Annotation

After initial analysis and pruning of the dataset, 70 images were selected for manual annotation. Annotation was done for the head, thorax, abdomen and wings of the dragon y.

In this project, the annotation was done manually on the images for training the model. This decision was supported by the ablation study conducted with the automatic annotation using RoboFlow and the zero-shot learning performed with the selected models used for segmentation, which is recorded in Section 5.1.

Since annotation is crucial, a small trial was carried out to see if automated annotation could be done, as it would save time and e ort. Two images were uploaded to RoboFlow, which is a standard tool for annotation. The platform consists of tools as well as models for annotation - both manual and automated. Grounding DINO [45] was used as the model for automated annotation. As seen in Figure 3.7, the con dence levels are set quite low (25% for the wings, 35% for the abdomen, 30% for the thorax, and 15% for the head), to identify most of the parts of the dragon y. These values were chosen at random, as to identify most of the required parts. While two of the four wings are not identi ed, the model does manage to nd the rest of the parts. When the same con dence levels are retained, it struggles to adapt to newer images. This is seen in Figure 3.8. In this image, the model is inaccurate and misidenti es the thorax and abdomen. This illustrated the need for manual annotation.

For the rst round of experiments, annotation was done using QuPath [12] (explained in detail in Figure 3.1.4) and converted to YOLO-compatible format using Python code. However, the other three models use COCO-style annotations[1], and the conversion proved to be tricky. Hence, the annotations were uploaded and validated again using RoboFlow [46] and downloaded in both YOLO and COCO formats.

For the second round of experiments, as the annotations were generated from the YOLOv1 model, an extension called YOLO-Label [47] was used to validate the generated annotations and make changes. However, as the conversion was still tricky, RoboFlow [46] was used again to validate and export to the required formats.

QuPath

QuPath is an open-source annotation tool mainly used for bio-image analysis, such as histological and pathological data. However, it can also be used for other datasets. The Wand provided in QuPath helped draw accurate annotations for the wings and helped capture the contours of the wings accurately. In this manner, 70 images were manually annotated and the labels were exported

[1] YOLO uses annotations structured as images and labels in the form of text les, and arranged in a speci c folder format. Each line in the text le corresponds to an annotated object in the image and is written as class_no x_center y_center width height . However, COCO-style annotations consist of all the images and annotations structured as a JSON le, with the class ID associated with each annotation at the end. The aforementioned structure is used only by YOLO, whereas COCO-style annotations are used by models that are built on Detectron2 library. The other three models use this.