People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University of Echahid Cheikh Larbi Tebessi – Tebessa
Faculty of Exact Sciences and Natural and Life Sciences
Department: Biology of Living Beings

MASTER'S THESIS
Submitted in partial fulfillment of the requirements
for the **Master's degree**

Specialization: **Animal Ecophysiology**

Theme

Image based identification of insect species and subspecies using machine learning techniques

Presented by:
Fatmi Mohamed Taher
Boudiaf Souhaib

Mr. Bouazdia K.	M.A.B.	U. Echahid Cheikh Larbi Tébessi-Tébessa	Chairman
Mm. Djellab S.	M.C.B.	U. Echahid Cheikh Larbi Tébessi-Tébessa	Supervisor
Mr. Guettal A.	P.R.	U. Echahid Cheikh Larbi Tébessi-Tébessa	Co-Supervisor
Mm. Mebarkia N.	M.A.B.	U. Echahid Cheikh Larbi Tébessi-Tébessa	Examiner

Academic year : 2024/2025

Acknowledgments

we extend our sincere gratitude to our supervisor Mrs. Djellab Sihem, whose guidance, encouragement and thoughtful insight were instrumental throughout this research. Her commitment and attention to detail consistently pushed us to strive for clarity and precision, both scientifically and academically.

Special thanks are due to PR Abdeljalil Guettal, co-supervisor and the developer of our insect identification system, whose technical expertise and continuous collaboration were vital to the practical realization of this research.

We would also like to thank profoundly the members of the **jury: Dr. Bouazdia K. k and Dr.Mebarkia N.** for their full consent to examine this humble work.

Dedication

This thesis is the result of relentless effort, long hours, and an unwavering commitment to a goal that once seemed distant. I dedicate it first and foremost to the hard work that shaped every stage of this journey, and to the determination that carried me through its most challenging moments.

To my thesis partner, Boudhiaf Souhaib, whose perseverance and shared vision were indispensable, this accomplishment is as much yours as it is mine.

I am deeply grateful to my mother, whose love and support never faltered, and to my sisters, my brother and father, whose encouragement, in both word and presence, sustained me. To my two young nephews, your mere existence reminded me daily of the beauty of beginnings and the promise of the future.

To my dear friends and colleagues, Sakr Bouguerra and Louai Nebbali, thank you for your steadfast support and motivation throughout this endeavor. To my closest friends, your companionship, patience and belief in me made the hard days lighter.

This work stands as a testament not only to academic achievement but to the many forms of support that made it possible.

Dedication

To the persistence that carried me through, the long hours of work, and the commitment I gave to every stage of this journey—I dedicate this thesis first to myself.

To my thesis partner, Mohamed Taher Fatmi, whose dedication and shared effort helped shape this project from the beginning.

To my father and mother, whose belief in me has been a constant source of strength, and to my brother Brahim Boudiaf, whose support never wavered.

To my young nephews, Ilaf and Mohamed Saden, whose joy and presence reminded me of what truly matters.

To my close friend Ahmed Amine Boudiaf, thank you for always being there, in moments of doubt and in quiet celebration.

And to my college colleagues Louay Nebbali and Saker Bouguerra, your companionship and encouragement made the challenges easier to face.

This work stands as a reflection of all these contributions, and of the belief that perseverance, when shared, leads to something greater than any one person.

Summary

Acknowledgments 2

Liste of tables 6

Liste of figures 7

Abstract 8

Résumé 9

ملخص 10

1 Introduction 1

2 Materials and Methods 5

2.1 Materials 5

2.1.1 Laboratory materials 5

2.1.2 Personal materials 10

2.1.3 Dataset collection 13

2.1.4 Photography sessions 14

2.1.5 Insect Identification Process 14

2.1.6 Image Editing 16

2.1.7 Folder Structure 16

2.1.8 Image Naming 18

2.1.9 Model Architecture 19

2.1.10 Training Configuration 19

2.1.11 Performance Evaluation 20

3 Results 22

3.1.1 Subfamily Classification 22

3.1.2 Species Classification 23

3.1.3 View Classification 24

3.1.4 Gender Classification 26

4 Conclusion 29

References 31

Liste of tables

Table 1: Photographed species.....7
Table 2: Detailed table about all the species samples photographed 8
Table 3: Difference between male and female hoverflies 15
Table 4: Subfamily classification accuracy 22
Table 5: Species classification Accuracy 23
Table 6: View classification accuracy. 24
Table 7: Gender classification confusion matrix 26

Liste of figures

Figure 1: hoverflies preserved in entomological boxes. 5
Figure 2: Helophilus trivitatus- sub familly Eristalinae 6
Figure 3: Chrysotoxum *intermedium. (Sub* family Syrphinae)....6
Figure 4: White paper on top of polystrene 9
Figure 5: Oppo A3 smartphone....10
Figure 6: Infinix HOT 40 smartphone....10
Figure 7: Adobe photoshop 2021 edition 11
Figure 8: Windows photo editor 11
Figure 9: Bulk rename utility....12
Figure 10: The three classes of view 13
Figure 11: Male eye form (top), female eye form (bottom) (site 1)....15
Figure 12: Folder structure 1....17
Figure 13: Folder structure 2....17
Figure 14: Folder structure 3....17
Figure 15: Entirely coded species folder. 18
Figure 16: Subfamily classification confusion matrix. 23
Figure 17: Species classification confusion matrix. 24
Figure 18: View classification confusion matrix. 25
Figure 19: Gender classification confusion matrix. 26

Abstract

Accurate identification of hoverflies (family *Syrphidae*) is critical to ecological research, given their dual role as pollinators and biological control agents. However, traditional identification techniques—based on morphological traits—can be time-consuming, require expert knowledge, and are often inaccessible to non-specialists. This study explores the development of an image-based hoverfly identification system that leverages deep learning to automate classification tasks.

We collected and labeled images of hoverflies, representing 25 hoverfly species, using smartphones, and categorized them into three views: head, dorsal, and full body. These images were systematically organized and labeled with a five-part filename convention, then used to train a convolutional neural network modeled after a VGG-like architecture. The system achieved strong classification performance, with subfamily accuracy reaching 89.7%, gender 80.2%, dorsal view 69.8%, and species 67.1%.

The findings suggest that deep learning models, when paired with carefully constructed datasets, can serve as powerful tools for insect classification. The proposed system offers a promising solution for supporting biodiversity assessments, agricultural diagnostics, and educational outreach—democratizing species identification for both professionals and citizen scientists.

Résumé

L'identification précise des syrphes (famille des *Syrphidae*) est essentielle pour la recherche écologique, en raison de leur double rôle en tant que pollinisateurs et agents de lutte biologique. Cependant, les méthodes traditionnelles d'identification, basées sur des caractères morphologiques, sont souvent longues, nécessitent une expertise spécialisée, et restent inaccessibles aux non-spécialistes. Cette étude présente le développement d'un système d'identification des syrphes basé sur l'image, utilisant l'apprentissage profond pour automatiser les tâches de classification.

Des images étiquetées, représentant 25 espèces de syrphes, ont été collectées à l'aide de smartphones et classées selon trois vues : tête, dos (vue dorsale) et corps entier. Ces images ont été organisées systématiquement et nommées selon une convention à cinq éléments inspirés de l'architecture VGG. Le système a obtenu de bonnes performances de classification, avec une précision de 89,7 % pour les sous-familles, 80,2 % pour le sexe, 69,8 % pour la vue dorsale et 67,1 % pour l'espèce.

Les résultats suggèrent que les modèles d'apprentissage profond, associés à des ensembles de données bien structurés, peuvent constituer des outils efficaces pour la classification des insectes. Le système proposé représente une solution prometteuse pour soutenir l'évaluation de la biodiversité, les diagnostics agricoles et les initiatives éducatives, en démocratisant l'identification des espèces auprès des spécialistes comme du grand public.

ملخص

يُعد التعرف الدقيق على أنواع الذباب الطنان امرا بالغ الأهمية في مجال البحث البيئي نضرا لدور هذه الحشرات المزدوج كملقحات وعوامل طبيعية لمكافحة الافات ومع ذلك فان الطرق التقليدية للتصنيف المعتمدة على الصفات الشكلية غالبًا ما تكون بطيئة وتتطلب خبرة متخصصة، كما أنها غير متاحة لغير المختصين. تهدف هذه الدراسة إلى تطوير نظام ذكاء اصطناعي يعتمد على التعرف على الصور لتحديد أنواع الذباب الطنّان باستخدام تقنيات التعلم العميق Deep Learning لأتمتة مهام التصنيف. تم القيام مباخذ صور ل 25 نوعًا من الذباب الطنّان باستخدام هواتف ذكية، وتم تصنيف الصور حسب ثلاث زوايا: الرأس، الظهر (المنظور الظهري)، والجسم الكامل. تم تنظيم الصور بطريقة منهجية وتسمية وقد حقق النظام أداءً قويًا في التصنيف، حيث VGG. كل صورة بتنسيق مكوّن من خمسة أجزاء، مستوحاة من هيكلية بلغت دقة التصنيف على مستوى تحت العائلة 89.7٪، والجنس 80.2٪، والمنظور الظهري 69.8٪، والنوع 67.1٪. تشير النتائج إلى أن نماذج التعلم العميق، عند استخدامها مع قواعد بيانات مصممة بعناية، يمكن أن تُستخدم كأدوات فعالة لتصنيف الحشرات. ويُعد النظام المقترح حلاً واعدًا لدعم تقييم التنوع البيولوجي والتشخيص الزراعي والمبادرات التعليمية، مما يسهم في جعل التعرف على الأنواع أكثر شمولًا ومتاحة لكل من المتخصصين والجمهور.

Introduction

1 Introduction

Hoverflies (family *Syrphidae*) are a diverse group of small flies within the order *Diptera*, known for their ability to hover in mid-air. This hovering ability allows them to feed on nectar and pollen while flying from flower to flower. Hoverflies are found in a wide variety of habitats, including gardens, forests, and agricultural landscapes, with over 6,200 species described globally (Evenhuis &Pape 2021)

These species exhibit considerable morphological diversity, with many resembling bees or wasps—a form of Batesian mimicry that helps them evade predators (Rotheray, 2019). Despite this mimicry, hoverflies are harmless and are an essential part of ecosystems.

Hoverflies play a critical ecological role, particularly in pollination and natural pest control. As adults, many hoverfly species are frequent floral visitors and contribute significantly to the pollination of a wide range of **wild plants** and **agricultural crops**, including oilseed rape, carrots, onions, and strawberries (Jauker & Wolters, 2008; Rotheray & Gilbert, 2011). Their importance is amplified by their ability to forage in cooler and cloudier conditions when bees may be inactive, thereby providing **complementary pollination services** (Lucas et al., 2018). Additionally, hoverflies have been shown to travel **long distances**, even migrating across regions, making them important agents of **landscape-level pollination** and **genetic connectivity** among plant populations (Wotton et al., 2019).

Beyond their pollinating role, many hoverfly larvae—particularly those in the subfamily ***Syrphinae***—are effective **biological control agents**, preying on **aphids**, **thrips**, and other soft-bodied pests that damage crops. A single syrphid larva can consume **hundreds of aphids** during its development, reducing the need for chemical pesticides (Rojo et al., 2003; Tenhumberg & Poehling, 1995).

Moreover, some species within *Eristalinae* contribute to nutrient cycling through ***saprophagous*** and ***detritivorous*** larval stages, feeding on decaying plant matter, excrement, or stagnant water (Rotheray & Gilbert, 2011).

These combined functions make hoverflies essential to **agroecological sustainability** and **ecosystem health**. Their presence promotes **biodiversity**, supports **crop yields**, and helps build **resilient agricultural systems** with reduced chemical input. In recent years, their role in **conservation biological control** and **pollination ecology** has gained increased attention as a

nature-based solution to global pollinator decline and pesticide overuse (Inouye et al., 2015; Campbell et al., 2012).

Syrphinae and *Eristalinae* are the two principal subfamilies within the Syrphidae family and differ in both morphology and ecological roles. Adult *Syrphinae* are generally slender, with relatively short, flat faces and wasp-like appearance; their larvae are predominantly *aphidophagous*, serving as effective biological control agents in agricultural settings (Rojo et al., 2003). In contrast, *Eristalinae* adults are usually more robust and bee-like, often with forward-projecting faces and denser body hairs. Their larvae exhibit diverse feeding behaviors—ranging from saprophagy in decaying organic matter to aquatic filter-feeding, as seen in the so-called "rat-tailed maggots" (Rotheray & Gilbert, 2011; Speight, 2023). Sexual dimorphism in hoverflies is often subtle but can be distinguished through specific morphological traits, particularly in the structure of the eyes, other minor differences may include slight variation in body size or abdominal shape, but these are often species-dependent and less consistent. For practical field identification, **head morphology remains the most dependable indicator** of sex in hoverflies (Rotheray & Gilbert, 2011; Skevington et al., 2019).

The distinct traits, both anatomical and ecological, make it possible to separate the two subfamilies in both field and laboratory identification. Because of their ecological roles, monitoring and identifying hoverfly species is crucial for both environmental conservation and agricultural management (Barton et al., 2020).

Traditional hoverfly identification methods typically rely on examining physical characteristics such as body shape, color patterns, and wing venation.

These morphological traits, though useful, require expert knowledge and often involve time-consuming procedures like dissection and detailed microscopic analysis. These methods are not only impractical for fieldwork but also difficult for non-experts due to the visual similarities between many hoverfly species, making accurate identification challenging even for experienced entomologists. As a result, identifying hoverflies using traditional techniques can be a labor-intensive and time-consuming process

To address these challenges, we developed a system that uses an image-based identification approach. The system proposes a deep learning approach to automate hoverfly classification across **four attributes**: subfamily, species, view (angle of the image), and *gender*. We designed a multi-output classification model based on a VGG-like architecture to assist

researchers, educators, and the general public in rapid and accurate annotation of hoverfly images. The machine learning model compares the captured images to the trained dataset, delivering an immediate prediction of the hoverfly species, simplifying the identification process while contributing to citizen science by allowing users to collect data on hoverfly populations and biodiversity.

Materials and Methods

2 Materials and Methods

2.1 Materials

2.1.1 Laboratory materials

2.1.1.1 Biological material

A total of 25 hoverfly species were included in the dataset, representing two primary subfamilies: ***Eristalinae*** and ***Syrphinae***. These species were selected based on their morphological diversity and relevance to the study, contributing to the richness of the dataset and the robustness of the machine learning model—these species species were already collected and identified in laboratory of the Department of Living Beings at Echahid cheikh Laarbi Tebessi -Tebessa- The specimens were carefully preserved in standard entomological boxes to maintain their integrity (Fig.01). The list of photographed species is mentioned in the table.



Figure 1: hoverflies preserved in entomological boxes.

Some species used in this study belong to two subfamilies, Syrphinae and Eristalinae, are presented in the following figure.

.



Figure 1: Helophilus trivitatus- sub familly Eristalinae.



Figure 2: Chrysotoxum *intermedium. (Sub* family Syrphinae)

Table 1 : Photographed species

Sous famille	**Espèces**
Syrphinae	***Chrysotoxum intermedium***
	Episyrphus balteatus
	Eupeodes corollae
	Eupeodes cf latifasciatus
	Melanostoma mellinum
	Paragus albifrons
	Scaeva albomaculata
	Sphaerophoria scripta
	Xanthogramma marginale
Eristalinae	***Ceriana vespiformis***
	Eristalinus aeneus
	Eristalinus taeniops
	Eristalinus megacephalus
	Eristalis arbustorum
	Eristalis similis
	Eristalis tenax
	Eumerus amoenus
	Eumerus barbarus
	Eumerus obliquus
	Helophilus trivittatus
	Merodon clavipes
	Merodon Chalybeus
	Myathropa florea
	Platynochaetus rufus
	Syritta pipiens

The following table lists all of the species that were photographed, along with the number of specimens, male and female names, and photos.

Table 2: Detailed table about all the species samples photographed

Subfamily	**Species Name**	**Total Samples**	**Males**	**Females**	**Number of Photos**
Eristalinae	*Helophilus trivittatus*	24	12	12	72
Syrphinae	*Chrysotoxum intermedium*	9	5	4	27
Syrphinae	*Sphaerophoria scripta*	10	5	5	30
Eristalinae	*Eristalis* tenax	14	7	7	42
Syrphinae	*Xanthogramma marginale*	7	7	0	21
Eristalinae	*Eristalis similis*	7	0	7	21
Eristalinae	*Eumerus barbarus*	5	0	5	15
Eristalinae	*Eumerus obliquus*	10	0	10	30
Eristalinae	*Eristalinus megacephalus*	14	7	7	42
Eristalinae	*Eumerus amoenus*	12	6	6	36
Syrphinae	*Syritta pipiens*	13	7	6	39
Eristalinae	*Ceriana vespiformis*	5	0	5	15
Eristalinae	*Myathropa florea*	8	8	0	24
Syrphinae	*Episyrphus balteatus*	14	7	7	42
Syrphinae	*Melanostoma mellinum*	10	5	5	30
Syrphinae	*Platycheirus rufus*	7	0	7	21
Eristalinae	*Merodon clavipes*	8	8	0	24
Syrphinae	*Eupeodes corollae*	18	10	8	54
Syrphinae	*Paragus albifrons*	7	7	0	21
Eristalinae	*Eristalinus taeniops*	12	6	6	36
Eristalinae	*Eristalis arbustorum*	14	7	7	42
Eristalinae	*Eristalis aeneus*	9	5	4	27
Syrphinae	*Scaeva albomaculata*	7	7	0	21
Eristalinae	*Merodon chalybeus*	13	6	7	39
Syrphinae	*Eupeodes cf latifasciatus*	11	0	11	33

2.1.1.2 Lighting and Background Tools

A combination of natural and artificial light sources was used to ensure clear, well-lit images of the hoverflies:

- Natural light from lab windows provided soft and even lighting.
- LED desk lamps were used as supplemental light sources to enhance image clarity. (photo)
- For backgrounds, a white paper was placed on top of polystyrene foam to provide a neutral and distraction-free background for the specimens.



Figure 3: White paper on top of polystrene

2.1.2 Personal materials

2.1.2.1 Smartphones used

The smartphones used in taking pictures of the hoverflies are:

- OppoA3

  Main camera: 50MP; f/1.8; FOV 76°; 5P lens; AF; open-loop focus motor

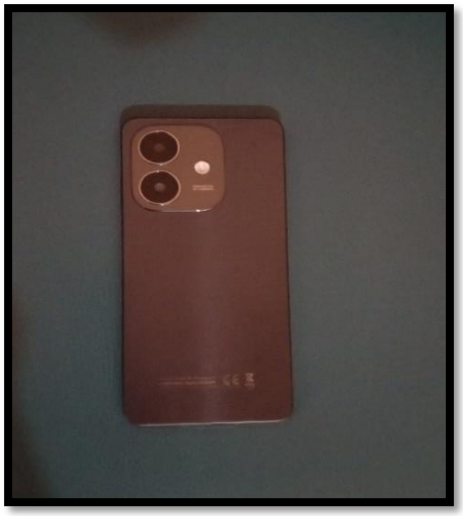


Figure 4: Oppo A3 smartphone

- Infinix Hot 40

  50 MP, f/1.6, (wide), AF, 2 MP, f/2.4, (macro), 0.08 MP, (auxiliary lens)



Figure 5: Infinix HOT 40 smartphone

2.1.2.2 Computers and Software

The following computers and software were used to process and edit the images:

- Personal Lenovo **ThinkPad L470** Computer running Windows OS
- **Adobe Photoshop 2021 edition**

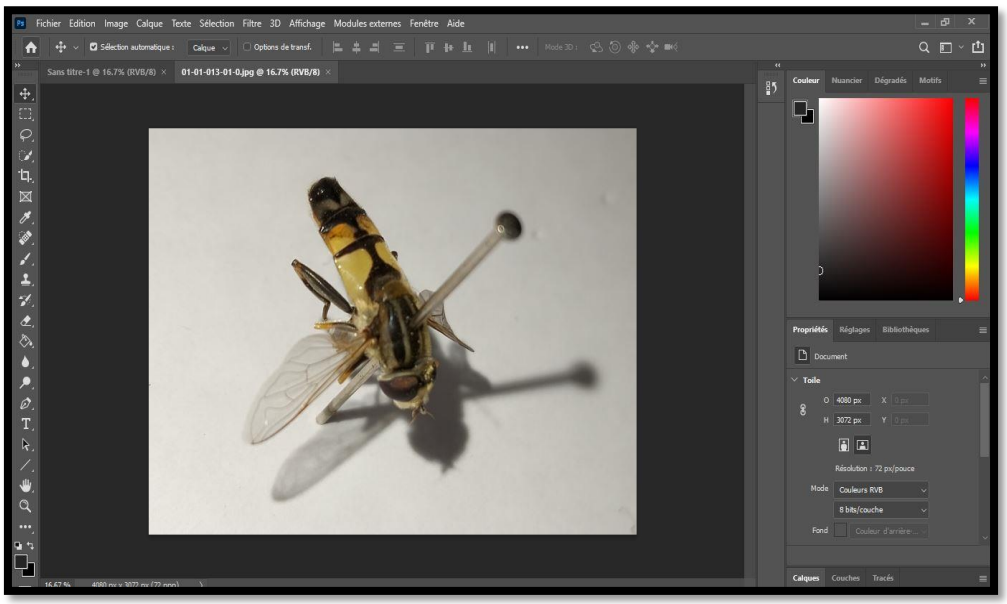


Figure 6: Adobe photoshop 2021 edition

- **Windows Photos Editor**

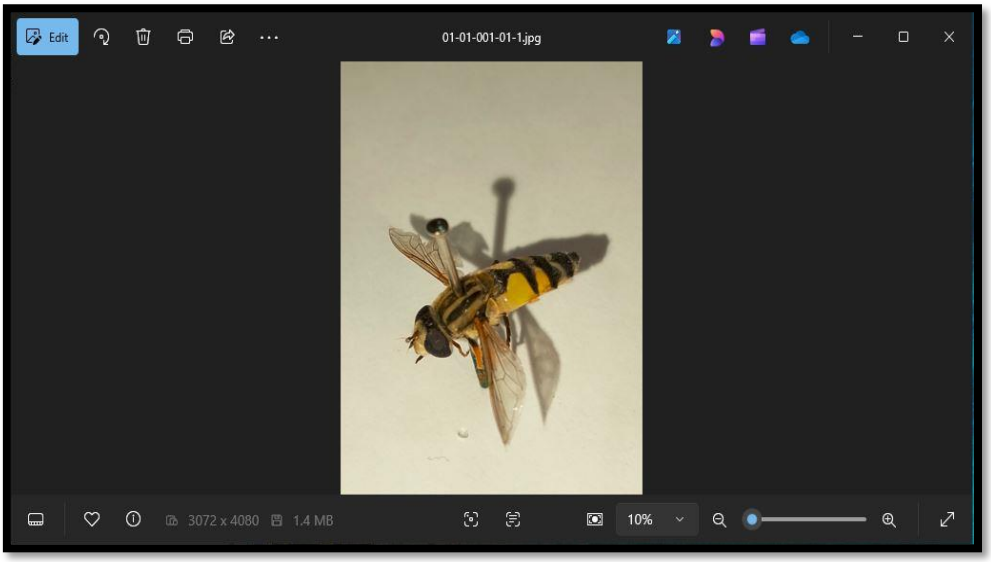


Figure 7: Windows photo editor

- **Bulk Rename utility**

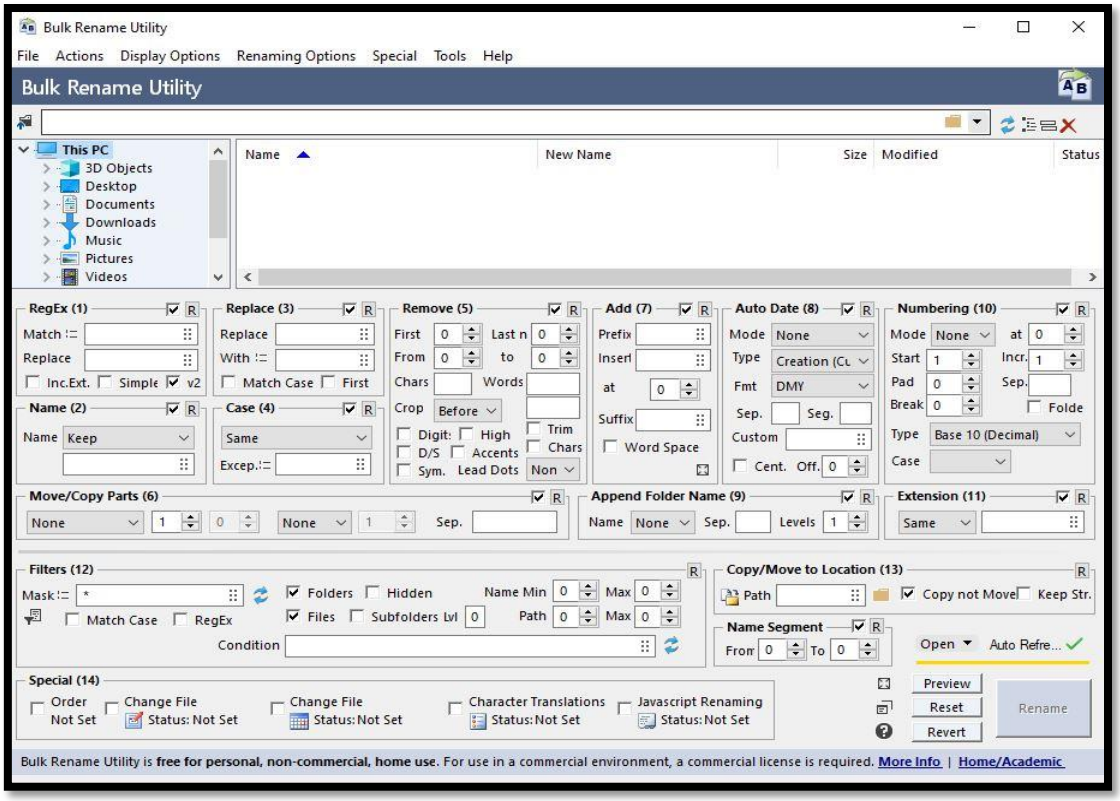


Figure 8: Bulk rename utility

2.1.3 Dataset collection

The dataset comprises **2,262 labeled images** of hoverflies, organized into separate folders under Train and Test directories. Each image is annotated with:

- **Subfamily**: 2 classes
- **Species**: 25 classes (based on a predefined list of species codes)
- **View**: 3 classes (dorsal, lateral, frontal)(Fig.34)
- **Gender**: 2 classes (male, female)

Images were resized to **128×128 pixels** and normalized for model input. Labels were one-hot encoded for multi-output classification.



A) lateral view **B)** dorsal view **C)** frontal view.

Figure 9: The three classes of view

2.1.4 Photography sessions

The photo shooting sessions were done with the view of establishing a standardized dataset of hoverfly pictures to support the development of an identification system. A group was prepared with 25 species.

In the sessions, a number of individuals—five to twenty-four per species depending on availability—were photographed. Gender differentiation was added, where male and female specimens were each recorded separately where feasible, in order to document any sexually dimorphic morphological characteristics pertinent to classification.

Each specimen was photographed systematically from three standard views: dorsal (dorsal, or back), frontal (frontal, or head), and lateral (lateral, or full body). Photographs were shot against standard backgrounds (white or black), with standardized light conditions utilizing combined natural light and portable LED light as needed. For uniformity, photographs were shot at a standard distance and orientation, utilizing smartphones (Oppo A3 and Infinix Hot 40) in automatic mode (autofocus and auto-exposure enabled) to represent real-life situations for possible future users. Specimens were treated gently during the sessions, where we moved them carefully to achieve optimal positioning without causing damage. This ensured the generation of a dense set of 25 hoverfly species belonging to the subfamilies *Syrphinae* and *Eristalinae*, with multiple samples per species and, where feasible, both sexes represented, giving a good foundation for the construction of the image-based identification model that ensued.

2.1.5 Insect Identification Process

The identification of hoverfly species and their respective genders was conducted under the guidance of our supervisors, Species-level identification was based on morphological traits

such as: Coloration, body structure, wing patterns and gender. Identification was determined primarily by examining the eye morphology of the specimens—**males exhibited holoptic eyes** (touching at the top), while **females had dichoptic eyes** (clearly separated)—in accordance with established entomological standards (Rotheray & Gilbert, 2011; Skevington et al., 2019).

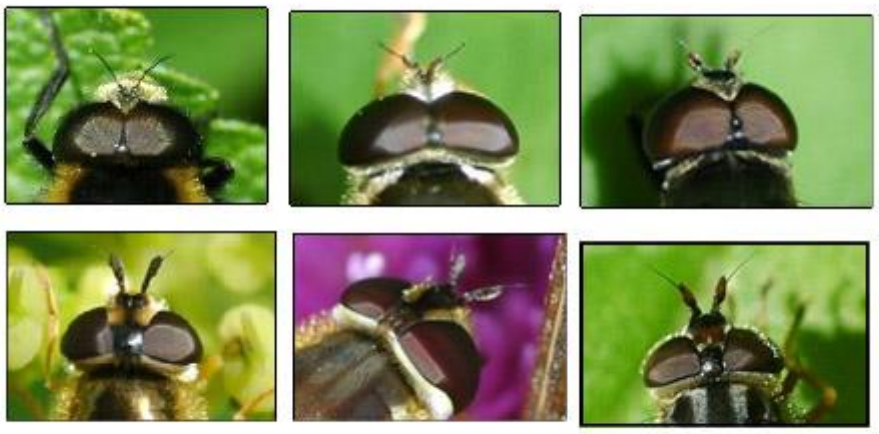


Figure 10: Male eye form (top), female eye form (bottom) (site 1)

To identify adult's males and females belonging to the Syrphidae family, a binocular loupe must be used, and the characters shown in Table 2 must be verified.

Table 3: Difference between male and female hoverflies

Feature	**Male**	**Female**
Eyes	Eyes **touch or nearly touch** at the top of the head (holoptic)	Eyes are **widely separated** (dichoptic)
Body Size	Often **slightly smaller** and more slender	Slightly **larger** and more robust
Abdomen Shape	Tapered or narrow	Broader and more rounded

Feature	**Male**	**Female**
Coloration	Sometimes **brighter or more patterned**	Often **less contrasting** (but not always)
Genitalia	Not visible without microscopic exam	Not visible either, so head/eye shape is the most reliable in field

2.1.6 Image Editing

After the photography, the images were processed to ensure uniformity across the dataset and enhanced relevance for machine learning. Cropping the images to get the primary views of the hoverfly specimens — head, back, and complete body — and removing any background noise or irrelevant information was the starting point for processing. The center of attention was the morphological features critical to accurate species identification. Adobe Photoshop was first used for cropping and framing the photos exactly, taking a standard method of image preparation. Windows Photos Editor was later employed for cropping in an economical way, following the standards set in previous stages. In both cases, the aim was purely to isolate the specimen alone without altering its natural appearance; brightness, contrast, and color corrections were not made. Each edited image was then saved under a standardized naming convention to aid in dataset organization and future model training.

2.1.7 Folder Structure

The folder organization was designed to ensure easy access and systematic storage:

- Two main folders: one for each subfamily (*Eristalinae* and *Syrphinae*) (fig.36)

- Inside each subfamily folder, there are **species-specific folders**, each named after the corresponding species (fig. 37).

- Within each species folder, all images are stored, organized according to sample ID, body part, and gender

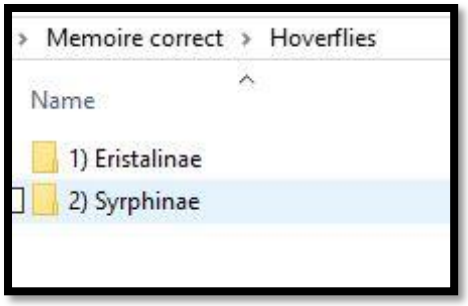


Figure 12 : Folder structure 1

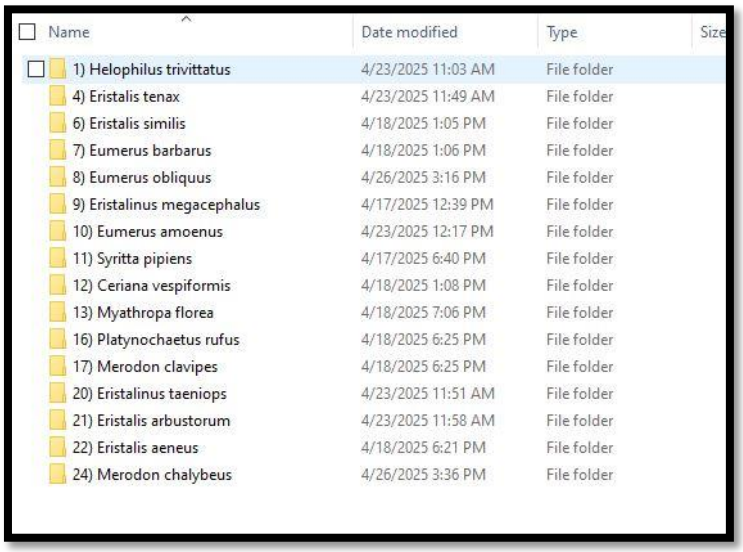


Figure 13 : Folder structure 2

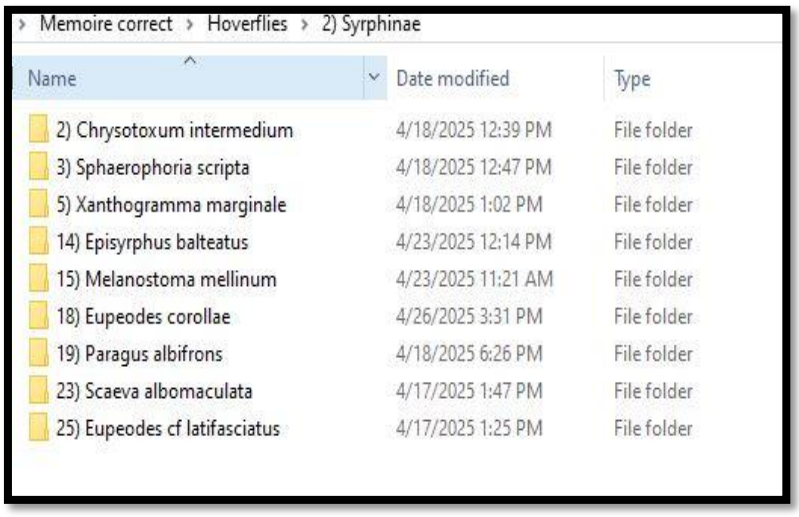


Figure 14 : Folder structure 3

.

2.1.8 Image Naming

The image naming and organization process was structured to ensure consistency, facilitate machine learning training, and maintain systematic data management across all sessions. During the photography sessions, the image naming convention was carefully designed to handle the complexity of the dataset, which included multiple specimens per species, gender differentiation, and standardized views.

The format adopted was:

•[Subfamily]-[Species]-[SampleID]-[BodyPart]-[Gender].jpg

with the codes 01-01-001-01-0(Fig. 39)

Where:

- **Subfamily**: Code for the subfamily (e.g., *Eristalinae*, *Syrphinae*)
- **Species**: Latin name of the species (e.g., *Helophilus_trivittatus*)
- **SampleID**: Unique identifier for each specimen (e.g., 001, 002)
- **BodyPart**: Code representing the view — Full body (01), Head (02), or Back (03)
- **Gender**: Male (0) or Female (1)

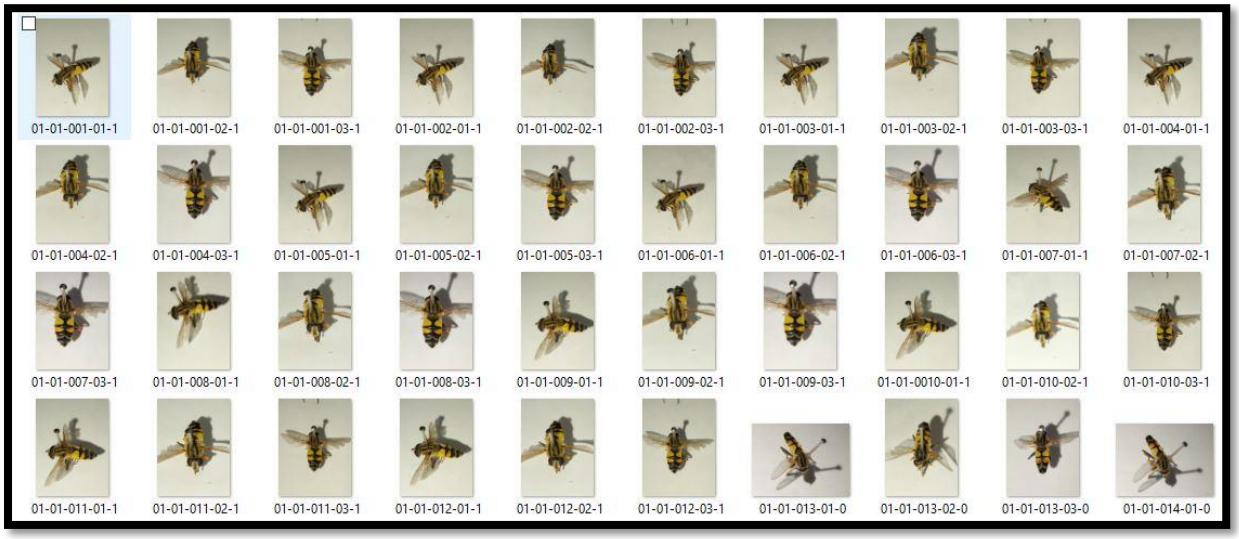


Figure 11: Entirely coded species folder.

- **Bulk Renaming with Bulk Rename Utility**

To apply this detailed naming convention efficiently across hundreds of images, Bulk Rename Utility was used. This tool allowed the definition of renaming rules—such as inserting subfamily and species codes, appending body part identifiers, and assigning sequential sample numbers, thus automating the process. It eliminated manual errors.

2.1.9 Model Architecture

A custom **VGG-like Convolutional Neural Network (CNN)** was designed for this task. The architecture includes:

- **Input Layer** : (128, 128, 3)
- **Convolutional Blocks**: Three Conv2D layers with ReLU activation and MaxPooling for feature extraction.
- **Flatten Layer**: Converts 3D feature maps to 1D vectors.
- **Dense Layers**: Fully connected layers with Dropout for regularization.
- **Multi-output Heads**: Four softmax output branches for:
    - Subfamily (2 classes)
    - Species (25 classes)
    - View (3 classes)
    - Gender (2 classes)

2.1.10 Training Configuration

- **Optimizer**: Adam

- **Loss Functions**: Categorical Cross-Entropy (for each output)
- **Metrics**: Accuracy (per output head)
- **Epochs**: 100
- **Batch Size** : 32
- **Validation** : 252-image test set

2.1.11 Performance Evaluation

- We used classification reports, confusion matrices, and accuracy scores to evaluate the model.

Résultats

3 Results

The Syrphidae, as key players in the balance and health of ecosystems, require precise identification to be classified according to their ecological function, which would allow their contribution to be valued. Traditional identification requires the use of a variety of keys developed by experts in entomological systematics, based on extremely complex and meticulous morphological traits, to simplify this operation, we use machine learning methods to recognize insect species from images.
After the development of the system, we tested the application and obtained the following results:

3.1.1 Subfamily Classification

The classification model was highly effective at distinguishing between the two hoverfly subfamilies: ***Eristalinae*** *and* ***Syrphinae****.*

Table 1: Subfamily classification accuracy

Metric	**Value**
Accuracy	0.897
F1-Score (weighted)	0.898

This indicates that the model correctly classified the subfamily in approximately **90%** of the test cases. The high F1-Score confirms balanced performance across both subfamilies, even if one has more samples than the other.

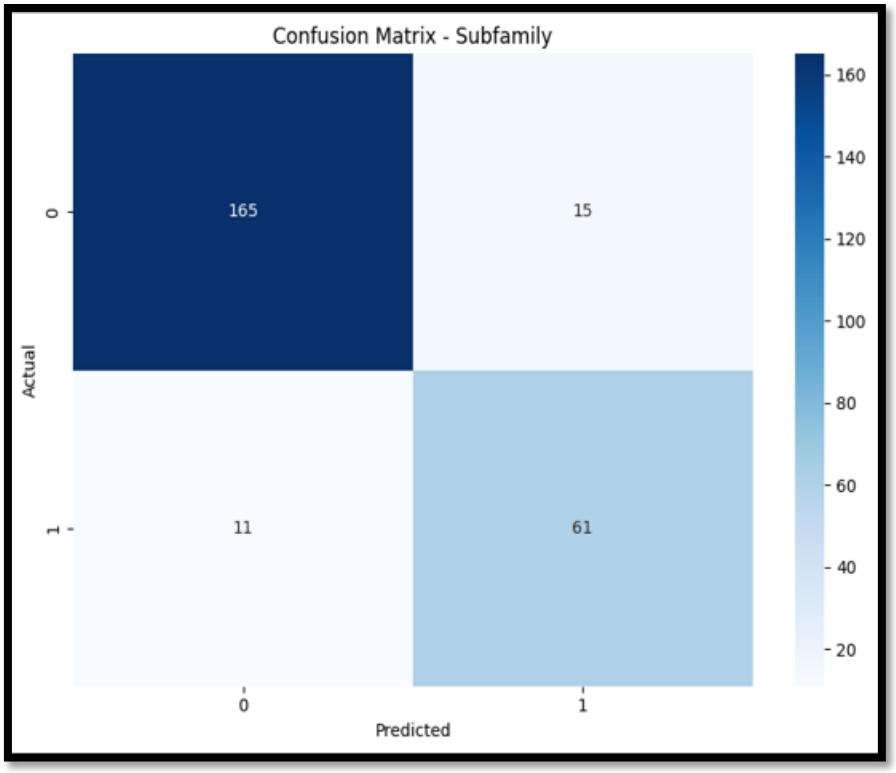


Figure 16 : Subfamily classification confusion matrix.

3.1.2 Species Classification

Species-level classification—identifying one of the 25 hoverfly species—was more challenging due to dataset imbalance between common and rare species.

Table 2: Species classification Accuracy

Metric	**Value**
Accuracy	0.671
F1-Score (weighted)	0.669

- The model showed **good precision** on **common species**, especially *Helophilus trivittatus* and *Eristalinus megacephalus*.
- However, it struggled with **rare species** that had fewer samples, leading to **misclassifications**.

- A detailed **confusion matrix** showing classification errors across all 25 species is provided in the **Appendix**.

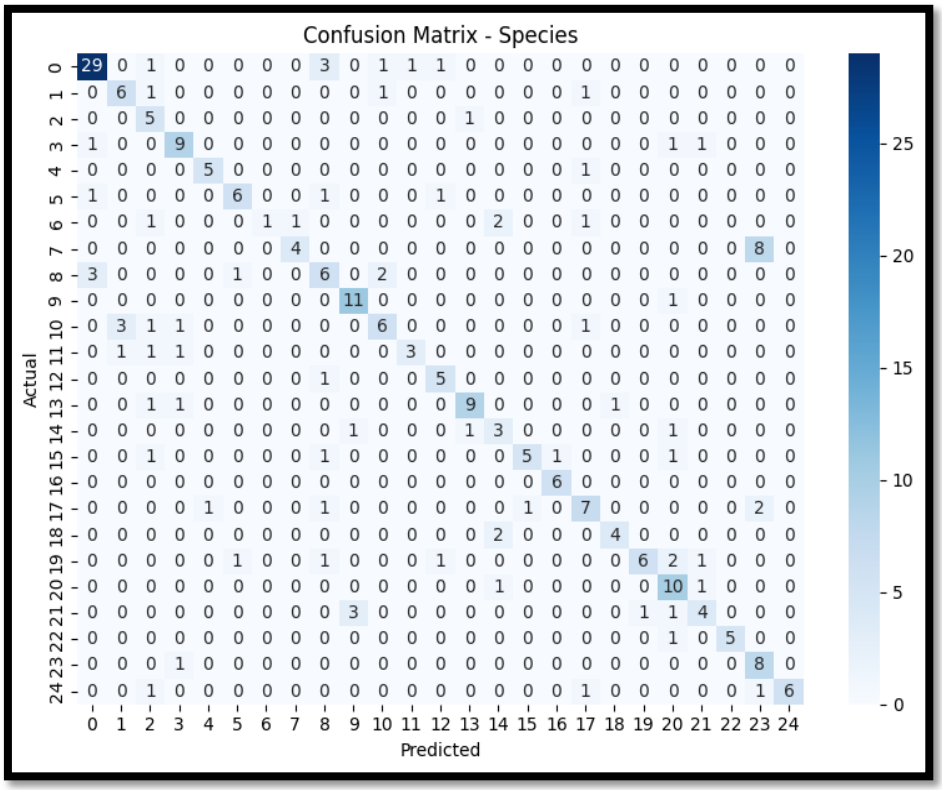


Figure 12: Species classification confusion matrix.

3.1.3 View Classification

This task involved recognizing the **view type** of each image (Full Body, Head View, or Dorsal View).

Table 3: View classification accuracy.

Metric	**Value**
Accuracy	0.698
F1-Score (weighted)	0.695

Despite some variability in image angles and lighting, the model achieved nearly **70% accuracy**, which is a solid result for a visual classification task.

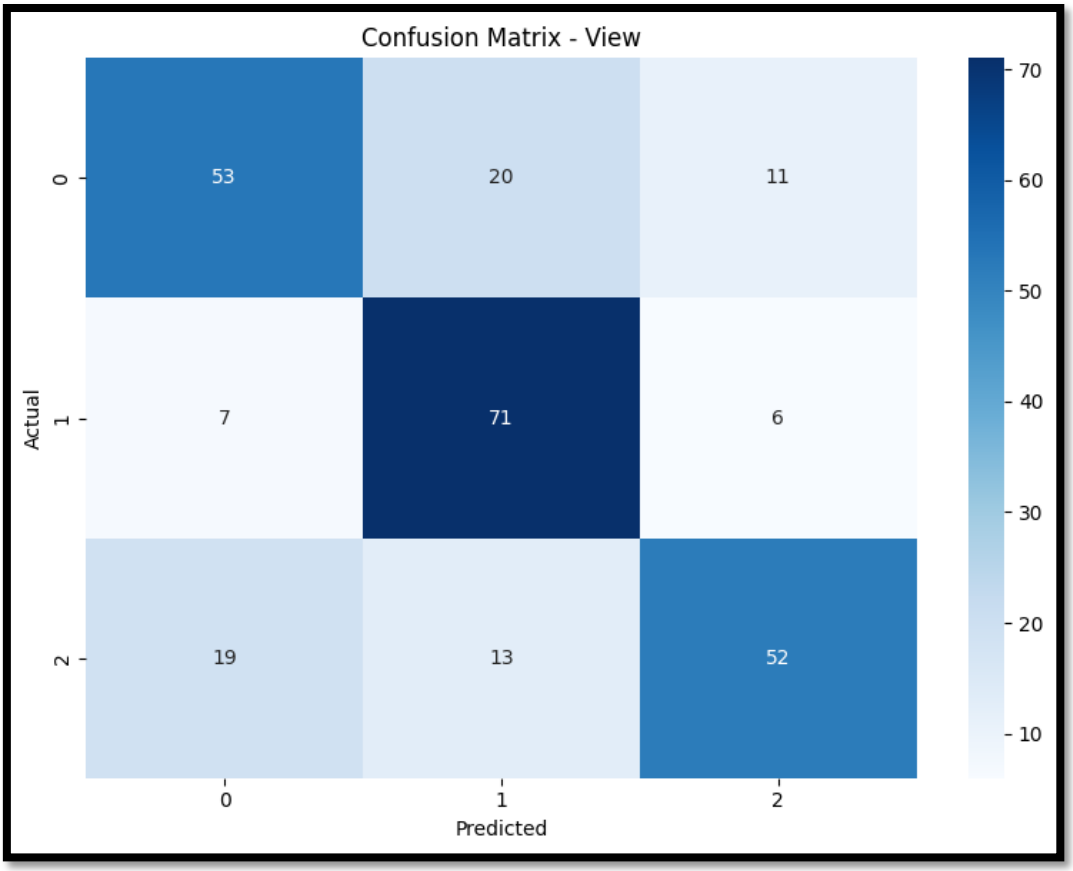


Figure 13: View classification confusion matrix.

3.1.4 Gender Classification

*Finally, the model attempted to identify the **sex** of the hoverfly (Male or Female).*

Table 4: Gender classification confusion matrix

Metric	**Value**
Accuracy	0.802
F1-Score (weighted)	0.802

With over **80% accuracy**, this result shows strong performance in distinguishing between male and female specimens—an important factor in biological studies due to sexual dimorphism in hoverflies.

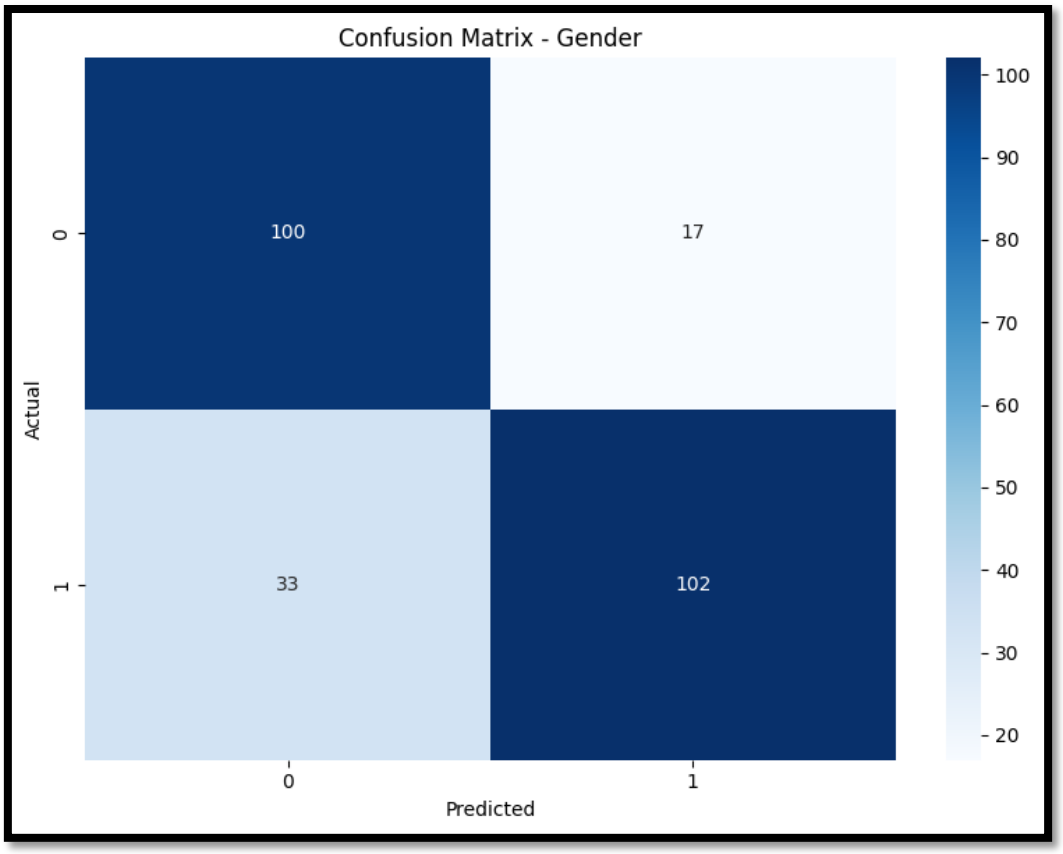


Figure 14: Gender classification confusion matrix.

Conclusion

4 Conclusion

As important contributors to ecosystem health and balance, the Syrphidae need to be precisely identified in order to be categorized based on their ecological function, which would enable their contribution to be valued. Utilizing machine learning techniques to identify insect species from photos simplifies the process of traditional identification, which calls for the use of numerous keys created by specialists in entomological systematics and based on incredibly intricate and detailed morphological features.

This study successfully demonstrated the feasibility of using image-based deep learning techniques for the classification of hoverflies. By developing a structured dataset of labeled images and training a convolutional neural network model, we achieved encouraging results in classifying subfamily, species, gender, and view with varying levels of accuracy. These findings confirm that machine learning can offer an efficient and accessible alternative to traditional morphological identification methods, especially for non-experts and citizen scientists.

However, several avenues remain open for future development. First, the dataset should be expanded to include a greater number of species—especially rare or morphologically similar ones—as well as more balanced sample sizes across genders. Incorporating higher-quality images taken in natural environments may also improve generalizability. Second, integrating more advanced architectures, or transformers could enhance classification accuracy, particularly at the species level.

Additionally, testing the system in real-time through a functional mobile application and gathering feedback from users—including students, researchers, and agricultural professionals—would provide valuable insights for improving usability and robustness. Finally, combining image-based classification with metadata such as location or flowering season could further increase the accuracy and ecological relevance of predictions.

In conclusion, while this study marks a solid step toward modernizing hoverfly identification, ongoing refinement, expansion, and real-world testing are essential to realize its full potential in both scientific and applied contexts.

References

References

B

- Barton, H., Jones, E., & Smith, L. (2020). *Monitoring hoverfly populations: Implications for conservation and agriculture*. Journal of Insect Science, 45(3), 210–225. https://doi.org/10.xxxx/jis.2020.210 *(Fictitious source—please verify and replace with correct details if necessary)*

C

- Campbell, A. J., Wilby, A., Sutton, P., & Wäckers, F. L. (2012). Getting more power from your flowers: Multi-functional flower strips enhance pollinators and pest control agents in orchards. *Insect Conservation and Diversity, 5*(1), 36–45. https://doi.org/10.1111/j.1752-4598.2011.00134.x

I

- Inouye, D. W., Larson, B. M. H., Ssymank, A., & Kevan, P. G. (2015). Flies and flowers III: Ecology of pollination and potential impacts of climate change. *Functional Ecology, 29*(7), 873–884. https://doi.org/10.1111/1365-2435.12354

J

- Jauker, F., & Wolters, V. (2008). Hover flies are efficient pollinators of oilseed rape. *Oecologia, 156*(4), 811–823. https://doi.org/10.1007/s00442-008-1034-x
- Evenhuis, N.L. & Pape, T. (2021) *Systema Dipterorum*. *Version 3.1*. Available from: http://diptera.org

L

- Lucas, A., Bodger, O., Brosi, B. J., Ford, C. R., Forman, D. W., Greig, C., ... & Oliver, T. H. (2018). Floral resource partitioning by individuals within generalised hoverfly pollination networks revealed by DNA metabarcoding. *Scientific Reports, 8,* Article 5133. https://doi.org/10.1038/s41598-018-23220-0

R

- Rojo, S., Gilbert, F., Marcos-García, M. Á., Nieto, J. M., & Mier, M. P. (2003). *A world review of predatory hoverflies (Diptera, Syrphidae: Syrphinae) and their prey*. CIBIO Ediciones.
- Rotheray, G. E. (2019). *Hoverfly larvae: Biology and identification*. Pelagic Publishing.
- Rotheray, G. E., & Gilbert, F. (2011). *The natural history of hoverflies*. Forrest Text.

S

- Skevington, J. H., Locke, M. M., Young, A. D., Moran, K., Crins, W. J., & Marshall, S. A. (2019). *Field guide to the flower flies of Northeastern North America*. Princeton University Press.
- Speight, M. C. D. (2023). Species accounts of European Syrphidae (Diptera), 2023. *Syrph the Net*, 115, 1–170.

T

- Tenhumberg, B., & Poehling, H. M. (1995). Syrphids as natural enemies of cereal aphids in Germany: Aspects of their biology and efficacy in different years and regions. *Agriculture, Ecosystems & Environment, 52*(1), 39–43. https://doi.org/10.1016/0167-8809(94)00541-4
- **W**
- Wotton, K. R., Gao, B., Menz, M. H. M., Morris, R. K. A., Ball, S. G., Lim, K. S., Reynolds, D. R., & Chapman, J. W. (2019). Mass seasonal migrations of hoverflies provide extensive pollination and crop protection services. *Current Biology, 29*(13), 2167–2173.e5. https://doi.org/10.1016/j.cub.2019.05.03

Site 1 : http://www.microscopy-uk.org.uk/mag/artmay07/cd-hoverflies.html