WORLDWIDE
DRAGONFLY
ASSOCIATION

Are wing contours good classifiers for automatic identification in Odonata? A view from the Targeted Odonata Wing Digitization (TOWD) project

Mayra A. Sáenz Oviedo [1], William R. Kuhn [2], Martin A. Rondon Sepulveda [1], John Abbott [3], Jessica L. Ware [4] & Melissa Sanchez-Herrera [4,5]*

[1] Department of Epidemiology and Biostatistics. Pontificia Universidad Javeriana, Ak. 7 # 40-62, 10231, Bogotá, Colombia

[2] Discover life in America, Gatlinburg, TN, USA

[3] Department of Museum Research & Collections. University of Alabama Museums. Tuscaloosa, AL 35487, USA

[4] Division of Invertebrate Zoology. American Museum of Natural History. 200 Central Park West at West 79th Street, New York, NY 10024, USA

[5] Faculty of Natural Sciences, Biology Department, Universidad del Rosario, Sede Quinta Mutis. Ak. 26 # 63C-48, 11122, Bogotá, Colombia

* Corresponding author. Email: melsanc@gmail.com, melissa.sanchezh@urosario.edu.co

Abstract. In recent decades, a lack of available knowledge about the magnitude, identity and distribution of biodiversity has given way to a taxonomic impediment where species are not being described as fast as the rate of extinction. Using Machine Learning methods based on seven different algorithms (LR, CART, KNN, GNB, LDA, SVM and RFC) we have created an automatic identification approach for odonate genera, through images of wing contours. The training population is composed of the collected specimens that have been digitized in the framework of the NSF funded Odomatic and TOWD projects. Each contour was pre-processed, and 80 coefficients were extracted for each specimen. These form a database with 4656 rows and 80 columns, which was divided into 70% for training and 30% for testing the classifiers. The classifier with the best performance was a Linear Discriminant Analysis (LDA), which discriminated the highest number of classes (100) with an accuracy value of 0.7337, precision of 0.75, recall of 0.73 and a F1 score of 0.73. Additionally, two main confusion groups are reported, among genera within the suborders of Anisoptera and Zygoptera. These confusion groups suggest a need to include other morphological characters that complement the wing information used for the classification of these groups thereby improving accuracy of classification. Likewise, the findings of this work open the door to the application of machine learning methods for the identification of species in Odonata and in insects more broadly which would potentially reduce the impact of the taxonomic impediment.

Key words. Classification, Machine Learning, supervised, wings

Introduction

Dragonflies and damselflies (Odonata) are one of the most charismatic insect groups, due to their relatively big size, flight patterns and beautiful colorations. Their association with aquatic environments means they serve as excellent bioindicators of water quality, given their high susceptibility to environmental changes (Córdoba-Aguilar, 2008; Moore, 1997; Samways & Steytler, 1996). Global odonate richness is estimated to comprise around 6.323 species (Paulson et al., 2021). This is a relatively small number of species in comparison with other insect orders like

Coleoptera, which includes approximately 300,000 species (Lorenzo-Carballa & Cordero Rivera, 2012; Paulson personal communication, May 14, 2021). However, researchers suspect that ~20% of species remain to be discovered (Kalkman et al., 2008). The tangible and relatively low diversity in odonates makes them an ideal scenario to address the "taxonomic impediment"—a lack of available knowledge about and trained expertise to determine the magnitude, identity, and distribution of biodiversity (González, 2009). This phenomenon is of particular interest given the current worldwide "biodiversity crisis" (i.e., rapid declining of populations as a result of massive habitat destruction and climate change), in which, there are estimates that ~50% of the living species will face extinction in the next 50 years (Koh et al., 2004). Maximizing efforts to gather and learn the taxonomy and biology of species is more relevant now than ever (Ceballos et al., 2015; Kuhn, 2016; La Salle et al., 2016).

There are several morphological characteristics that define the odonate suborders Zygoptera and Anisoptera, including: characters regarding the shape of the wings, head, thorax, abdomen and genitalia (Garrison et al., 2006). In particular, wing shape and venation patterns are one of the most commonly used traits to classify dragonflies and damselflies to family or genus level. For example, without the aid of a microscope one can easily differentiate them because anisopterans have different shapes of the fore- and hindwings, which remain perpendicular to the body when in rest, while zygopterans' fore- and hindwings are similar in shape and are usually folded in line with the body (Heckman, 2006, 2008). Recent contributions by Appel & Gorb (2014) proposed detailed micro-morphological characteristics of the wings such as the types of vein joints and combinations among them (i.e., four types of vein joints and five combinations), spine distribution across the wings (i.e., located on transversal veins, possibly involved in movement limitation), and the distribution of patches of the flexible protein resilin in the wings (e.g., on the joints, and/or along the veins). These new morphological traits have been discussed in the classification for both suborders, and are used to infer function and flight behavior.

Recently, Kuhn (2016) developed an automatic classification system for 26 dragonfly genera, using standardized image scans of specimen wings. He trained and classified them using a random forest algorithm by extracting feature vectors to describe texture and patterning through Gabor Wavelet Filters and a color assessment with a chromaticity standardization sampling within the images. Here we assessed the classification power to genera of a novel classifier trait for wings—their contour. By using standardized wing images from the Targeted Odonata Wing Digitization project, we tested multiple Machine Learning classification algorithms (e.g., Linear Discriminant Analysis—LDA, Logistic Regression—LR, Classification and Regression Trees— CART, K-Nearest Neighbors—KNN, Naive Bayes—NB, Support Vector Machines—SVM, and Random Forest Classifier—RFC) to establish the potential use of the wing contour within automated classification systems for odonates.

Materials and methods

We analyzed data from the Targeted Odonata Wing Digitization Project (TOWD; https://digitizingdragonflies.org), which aims to digitize the wings of all North America species of Odonata and to develop tools for automatically extracting useful characters from odonate wings to facilitate comparative studies and automatic species classification. We analyzed a dataset comprising 2,328 dragonfly and damselflies specimens from 111 genera, which were digitized through the TOWD Project. The dataset consisted of the contour (outline) of the fore- and hindwings of each specimen. These data were extracted from digital scans of the specimens using an edge-finding algorithm to recover a series of points (x,y-coordinates) representing the location of each pixel along the edge of a wing. In most cases, the contours represented the right wings, which were excised from the specimen's body and scanned on a flatbed scanner, except in some cases where the left-side wings were scanned when the right ones were damaged (see Supplementary Table 1 for a list of specimens). In the latter case, wing contours were reflected left-to-right to match up with right-side wing contours. As part of the TOWD preprocessing, each contour was rotated so that the upper side (costal margin) is approximately horizontal, translated so that the upper-left corner is at (0,0) and scaled to millimeters. The edges of some wings were damaged, which was also apparent in their respective contours; such damage was used as an exclusion criterion.

Contours data were preprocessed and analyzed in Python (van Rossum & Drake Jr, 2009; v. 3.9.2) using the Spyder Integrated Development Environment (IDE) Spyder (Raybaut, 2009; v. 4.2.1), which is part of the Anaconda Software Distribution (2016). Data treatment was divided into four main steps (Available code: https://doi.org/10.5281/zenodo.6614239):

(i) Preprocessing and Fourier's descriptors extraction: Standardization was performed on every contour, to ensure the comparability of data and improve the classification accuracy (Pal & Sudeep, 2016). This was accomplished by following a series of functions that returned a slightly modified set of coordinates that fulfill common main characteristics: The contour was closed by appending the first coordinate to the last one, in case these didn't coincide; the direction of the coordinates of every contour was verified and changed to be on a clockwise orientation; in case the contour contained less than 200 points, some points were interpolated. Next, the apex of the wing is located, and the coordinates are rotated to make it the starting point. Finally, the contour was checked again to ensure it had been closed.

After the preprocessing, the Fourier descriptor's coefficients were extracted using the Python implementation for approximating contours with a Fourier series, PyEFD (Blidh, 2016). This process allowed the extraction of the same number of coefficients for each wing, regardless of their size. The normalized coefficients were kept in a separate database, with each specimen's unique identifier (uniq-id).

(ii) Database division in training and test datasets: A train-test division was performed, following a 70/30 proportion: 70% to train the model and 30% for testing/validation.

(iii) Definition, training, and testing of classifier algorithms: Seven classifiers were chosen to be trained and tested for classification from the Scikit-learn distribution (Pedregosa et al., 2011):

- Logistic Regression (LR) is a binary linear classifier, which is the simplest and is used as a baseline model. To adjust LR to a multiclass problem, where the classification is done with a one vs rest method, the option multi_class = 'ovr' was set.
- Classification and Regression Trees (CART) is a multiclass classifier that uses recursive partitioning following the Gini Impurity Index to build a decision tree.
- K-Nearest Neighbors (KNN) is a multiclass classifier that assumes similarity depending on class proximity, calculated as an Euclidean distance.
- Naïve Bayes (NB) is a multiclass classifier that assumes conditional independence between every pair of classes.
- Linear Discriminant Analysis (LDA): is a linear classifier for a multiclass problem. It ensures the maximum separability of classes by reinforcing the proportion of intra and inter class variance (Narayan, 2020; Tharwat et al., 2017).
- Support Vector Machines (SVM) build a hyperplane or group of hyperplanes on a higher dimensionality space that allow the separation of nonlinear problems (Gandhi, 2018). The option StandardScaler was used to normalize and scale the data; and the option SVC, is used to specify the classification task.
- Random Forest Classifier (RFC) fits several decision trees on different sub-samples of the data. To set the number of trees in the forest, the option n_estimators = 200 was set.

In addition, for each classifier, a cross validation score and a classification report was obtained with five items: Accuracy (number of correct predictions from the total number of predictions), Precision (number of true positives from all the positive predictions), Recall (number of positive predictions from the total number of positive classes), F1 score (following equation:

2 (True Positives (TP) × False Positives (FP) ÷ 2 TP + FP + False Negatives (FN))

and Support (number of individuals in each class).

(iv) Confusion matrices: Confusion matrices were plotted for each classifier to obtain a detailed visualization of the classification errors: On them, the predicted and real classes are found on the x- and y-axis, respectively. The correct predictions of the classifier are found on the diagonal where the predicted and true labels coincide. In consequence, the predictions that lay outside of this diagonal, correspond to classification errors that inform about the performance of the classifiers, as well as possible confusion patterns.

Finally, we performed ANOVA and Tukey tests in order to compare the accuracy and F1 scores from the classification report, along with box plots calculated from the data.
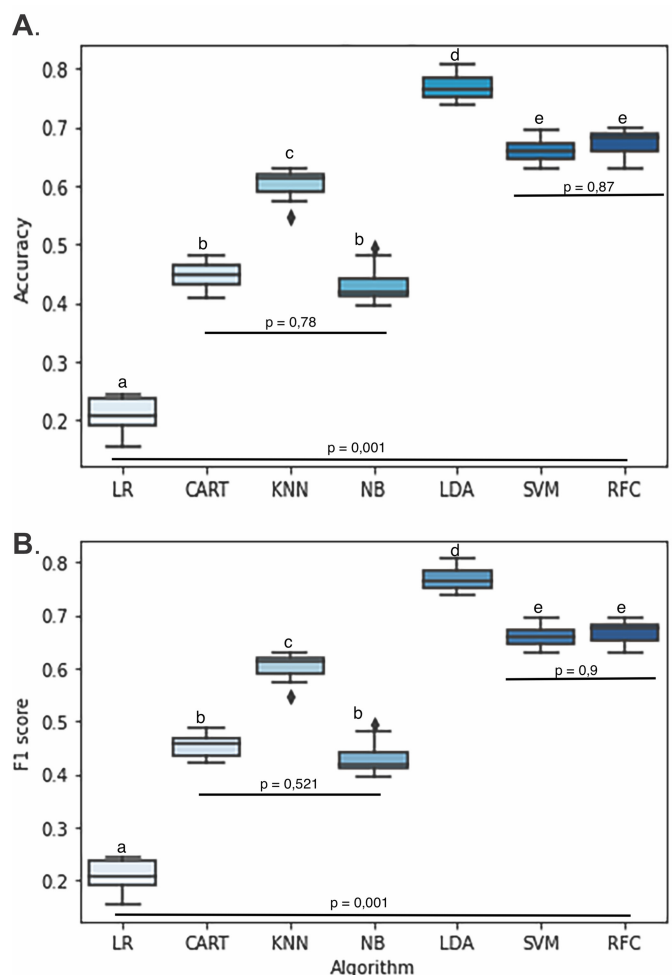
Figure 1. (A) Accuracy (number of correct predictions from the total number of predictions) and (B) F1 Score (a measure of a model's accuracy on a dataset that follows the formula: *((2 × Precision × Recall) ÷ (Precision + Recall))* boxplots of 3-fold cross validation for each of the seven classifiers tested. A total of 1397 individuals for each testing dataset per classifier was used; same letters indicate non-significant comparisons, *p*-values are shown for the CART—NB and SVM—RFC comparisons which were non-significant for both scores.

Results

We extracted a dataset of 4656 rows and 81 columns of Fourier coefficients, after the preprocessing images and the Fourier extraction loop we defined. Each row of this dataset belongs to an individual organism and each column to one of the coefficients. In total, we obtained 39 descriptors for each wing (hindwing and forewing) per individual, to be used later in the classification.

To the latter database, we tagged the genus label to each individual (row) in order to create a training and a testing set following a 70:30 proportion, respectively. As a result we generated 3259 individuals (70%) for the training, and 1397 individuals (30%) for the testing sets. The accuracy scores were similar enough in all seven classifiers defined (LDA, SVC, LR, CART, NB, RFC, KNN) between the two sets, which rules out possible overfitting of the classification models (see Supplementary Tables 2, 3). Furthermore, using the testing set the classification report obtained showed that the LDA classifier had the best performance in terms of: (1) accuracy (0.7337); (2) precision (0.75); (3) recall (0.73) and F1 score (0.73); in comparison with the other six classifiers tested (Supplementary Table 3). The ANOVAs performed for the F1 score and accuracy were significant (Fig. 1; Supplementary Tables 4 + 6), across the models.
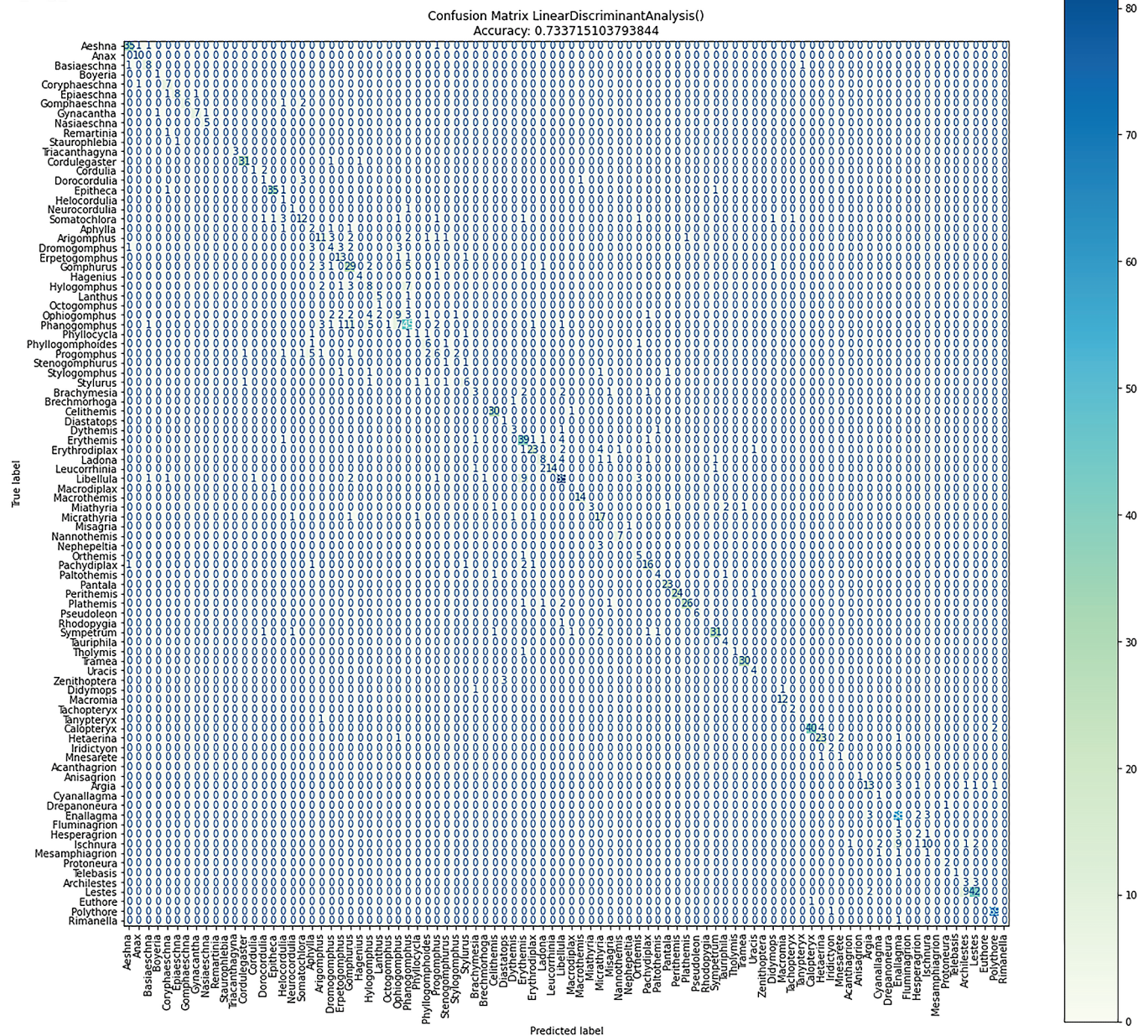
Figure 2. Confusion matrix. (A) LDA Confusion Matrix. **next page.** (B) Confusion Matrix showing misclassification zones distributed mainly on four families: Gomphidae (blue), Libellulidae (red), Coenagrionidae (green), Lestidae (orange). Each cell of the matrix corresponds to every possible true label and predicted label pairing. The color bar on the side of each plot, shows the code for the number of coincidences on each cell (from 0 = white, to 80 = dark blue).

The post-hoc Tukey multiple comparisons test showed differences for the accuracy and the F1 score comparison among all the classifiers, with the exception of the CART and NB comparison, and the SVM and RFC comparison (Fig. 1; Supplementary Tables 5 + 7). The latter performance metrics relies on the calculated confusion matrix per model, the LDA classifier confusion matrix shows the highest number of individuals on the diagonal, meaning these are true positives (Fig. 2A). Despite its better performance we noticed consistency in particular taxa that create misclassification in almost all classifiers, that we call confusion groups (Fig. 2B, Supplementary Figures 1–6). Particular genera within the following four families—Gomphidae, Libellulidae, Coenagrionidae and Lestidae—seem to be responsible for the misclassification observed (Fig. 3).

Discussion

Image preprocessing functions allowed a standardization of the coordinates on the contour dataset. This process has been found to guarantee data comparability and improve classification accuracy when compared with non-preprocessed images (Pal & Sudeep, 2016; Shahriar & Li, 2020; Sharma et al., 2020). The similarity of accuracy scores for all the classifiers in both training and testing sets, suggest that there were not overfitting issues in the models tested (Brownlee, 2017). Furthermore, we detected different numbers of classes (genera) for each of the seven classifiers. For example, the classifier with the best performance, LDA, created and recognized a total of 100 (classes proxy of genera) from 111 genera we included in the taxon sampling.
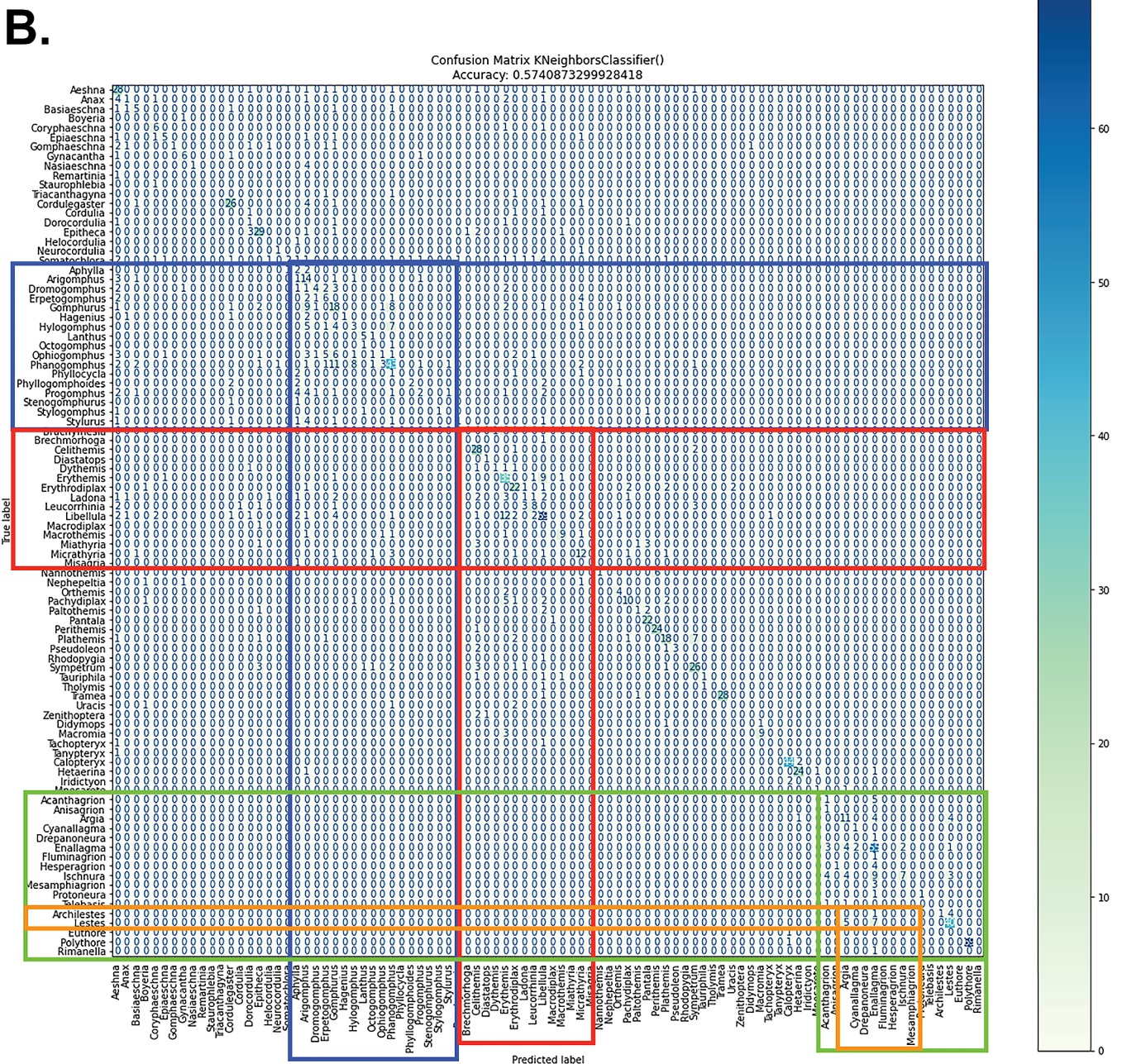
Figure 2. Continued (see page before).
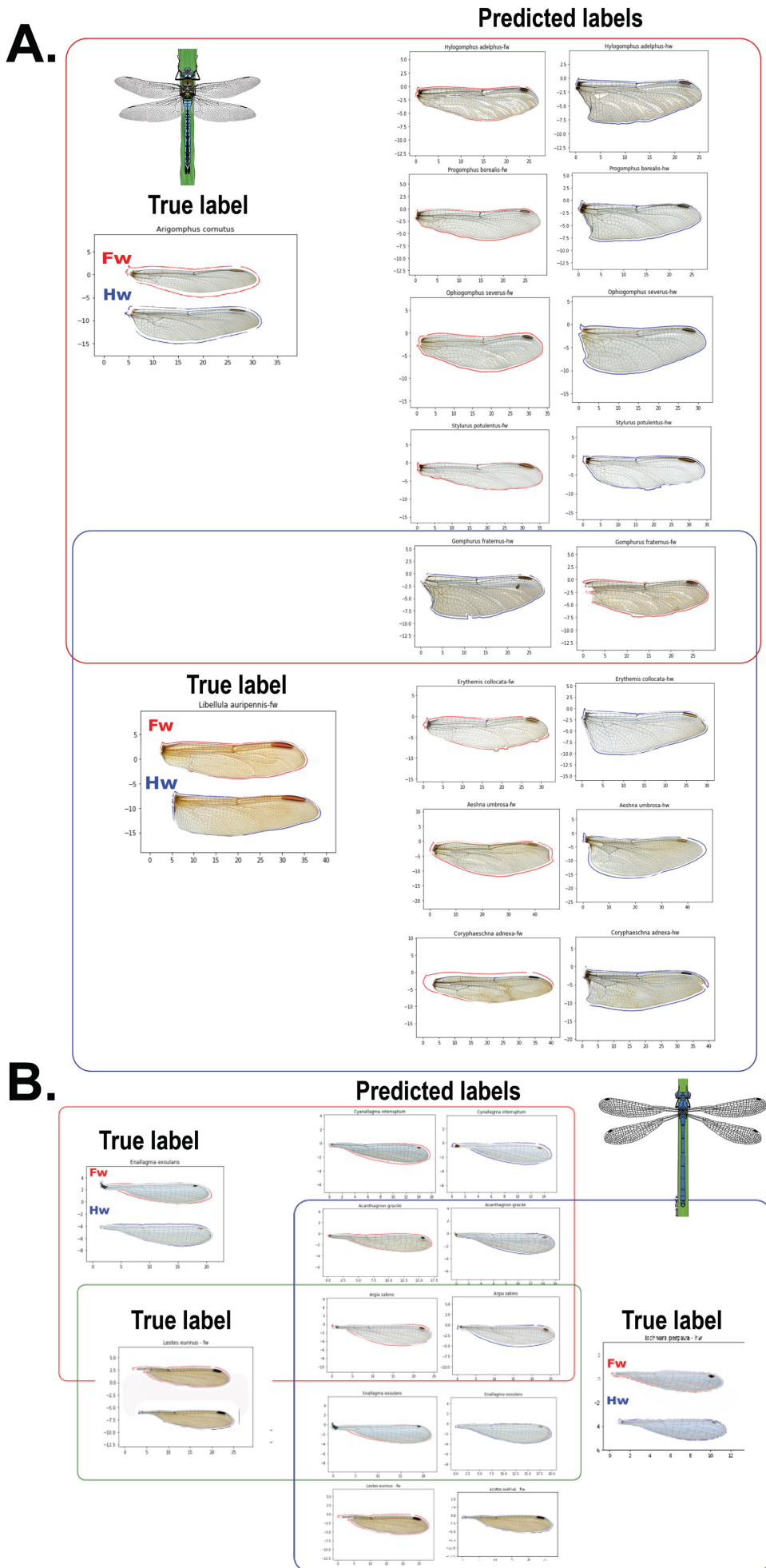
Figure 3. Confusion groups. (A) Anisoptera: True (real) label on left column and Predicted label on the right column. Surrounded by a red square (top): the Gomphidae genus *Arigomphus* (True label) was predicted as *Hylogomphus, Progomphus, Ophiogomphus, Stylurus* and *Gomphurus* (also from the Gomphidae family). At the bottom of the figure, surrounded by a blue square: the genus *Libellula* (Libellulidae), was confused with *Gomphurus* (Gomphidae), *Erythemis* (Libellulidae), *Aeshna* (Aeshnidae) and *Coryphaeshna* (Aeshnidae). (B) Zygoptera: True (real) label on left and right column and Predicted label on the center column. Surrounded by a red square (top left) Coenagrionidae genus *Enallagma* was predicted as *Argia*, *Acanthagrion* and *Cyanallagma* (Also Coenagrionidae genera). Surrounded by a green square, the genus *Lestes* (Lestidae) was predicted as Coenagrionidae genera *Argia* and *Enallagma*. At the bottom of the figure, surrounded by a blue square, genus *Ischnura* (Coenagrionidae), was predicted as *Acanthagrion*, *Argia*, *Enallagma* (all Coenagrionidae) and *Lestes* (Lestidae). Illustrations from Amanda Whispell.

(see Table 1, Supplementary Table 3). These differences may be due to class imbalance, meaning that there is unequal representation of genera in the dataset, with some of them having only one individual in the dataset (Table 1). Therefore, it is possible that during the data partitioning some groups were not included in the training dataset, which prevents the label from being created and in consequence, it would then not be included in the classification report. Likewise, if any of the groups were not represented in the testing dataset then its label would still be created, but the values for the metrics would be zero.

Furthermore, since machine learning algorithms depend on the distribution of classes in the training set to estimate the probability of observing examples in each class, class imbalance causes algorithms to learn that less well represented classes are not as important as the majority classes, so the performance will be better in the latter (Brownlee, 2017). To solve this inconvenience, an alternative could be to partition the data set in a stratified way, to ensure that all classes are balanced in the training and test sets. Moreover, it is necessary to increase the number of individuals in the less represented genera.

According to Kuhn (2016), the accuracy values are not strongly affected by the number of classes. In his study, a comparison was made between models with different numbers of classes, which ranged from three to 26. The results of the research suggest that a greater number of classes does not have a significant effect on accuracy, which slightly decreased its variation, as the number of classes increased, staying around 80%. Thus, it is possible to infer that the influence on the accuracy of the number of classes in the present study is also low.

The qualitative assessment of the confusion matrix (Fig. 2, Supplementary Figures 1–6), reveals classification mistakes in particular taxa just by looking at wing contours. The diagonal of the matrices shows the coincidences between the real and the predicted labels: if there individuals appear along this diagonal, that means the performance of the classifier is better, since on this diagonal the coincidences between the true labels and the predictions (true positives) will be found (we expect a 1:1 relationship if so). Consequently, on either side of these true positives diagonals, classification errors (false positives and false negatives) are found. True negatives, meanwhile, correspond to all the true instances found on the diagonal, different from the one of interest (Harrington, 2012).

Unlike the present investigation, on which the shape of the contour of the wings from 111 genera of dragonflies was exclusively evaluated, and seven classifiers tested, Kuhn (2016) made a classification of 26 genera of dragonflies, in which characteristics such as color, texture and shape of the wings were included, reaching a maximum accuracy of 91%, using only the Random Forest algorithm classifier. We suggest a possible explanation for the difference found in accuracy between Kuhn's (2016) and our data is due to additional characters assessed for the differentiation of species (texture, coloration and wing proportions). Our data suggests that the contour used here by itself does not provide enough information to obtain the accuracy found in Kuhn (2016). Thus, we suggest that the combination of the wing contours and the wing attributes previously assessed by Kuhn (2016) (including a morphometric analysis using 15 measurements, a chromatic analysis and, finally, the use of the Gabor wavelet transformation on the images with different rotations and scales) might increase the accuracy of the automatic identification for these taxa. In addition, we did find that the LDA classifier has better performance, suggesting the need to assess other classifiers than RF, which include all the possible wing attributes to test their performance in the classification. We expect to combine our results with the previous wing attributes tested by Kuhn (2016) for the automatic identification to keep decreasing the taxonomic impediment in the current biodiversity crisis.

The largest number of misclassifications of our data are centered on the tested genera within the anisopteran families Gomphidae and Libellulidae and the zygopteran Coenagrionidae and Lestidae families. This is interesting as Gomphidae, Libellulidae and Coenagrionidae are the most species rich families in the Odonata. Our results suggest that most of the confusion and classification errors are distributed among particular groups within families belonging to the same suborder (Fig. 3). In particular, there are two confusion groups that belong to the Anisoptera suborder (Fig. 3). In the first group (Fig. 3A, red square), six genera of the Gomphidae family are included, while in the second group (Fig. 3A, blue square), there are two genera that belong to the Libellulidae family, two genera of the Aeshnidae family and one of the Gomphidae family. Likewise, Kuhn's (2016) confusion matrix has similar classification mistakes to the ones we observed here (Fig. 3). For example, the genus *Erythemis* with the classifier and attributes tested by Kuhn (2016) was confused with species of the genera *Pachydiplax* and *Libellula;* in our results it was also confused with *Libellula* and a couple of aeshnids (Fig. 3A). For Zygoptera, our observed confusion occurs mainly between the Lestidae and Coenagrionidae families (Fig. 3B). The occurrence of greater confusion within this suborder may be a consequence of the low level of variation in their shape between families. This fact, in turn, underscores the need for identification of the Zygoptera facilitated by characteristics such as coloration, types of joints of the veins in the wings, patterns of venation, presence of spines and distribution of resilin patches (Appel & Gorb, 2014; Hassall, 2014).

Interestingly, our data suggest that these confusion groups have similar wing contours, which can lead us to look for possible hypotheses that explain these similarities among these taxa. Some explanation can be due to their ecology: for example, within the Anisoptera there

Table 1. Genera found by each classifier and number of individuals in the dataset. The first column ("Genus") has the names of the 111 genera included in the dataset. The Xs mark where the class was found and the dark gray empty cells show the classes (genera) that were absent in the classification report, for each of the classifiers.

Genus	Count	LDA	LR	NB	CART	KNN	SVM	RFC
Acanthagrion	16	x	x	x	x	x	x	x
Aeshna	131	x	x	x	x	x	x	x
Amphiagrion	1					x		
Amphipteryx	1					x		
Anax	39	x	x	x	x	x	x	x
Anisagrion	2	x	x	x	x	x	x	x
Aphylla	33	x	x	x	x	x	x	x
Archilestes	20	x	x	x	x	x	x	x
Argia	85	x	x	x	x	x	x	x
Arigomphus	81	x	x	x	x	x	x	x
Basiaeschna	27	x	x	x	x	x	x	x
Boyeria	11	x	x	x	x	x	x	x
Brachymesia	22	x	x	x	x	x	x	x
Brechmorhoga	4	x	x	x	x	x	x	x
Calopteryx	149	x	x	x	x	x	x	x
Cannaphila	4	x			x			
Castoraeschna	1	x						
Celithemis	149	x	x	x	x	x	x	x
Cordulegaster	90	x	x	x	x	x	x	x
Cordulia	8	x	x	x	x	x	x	x
Coryphaeschna	25	x	x	x	x	x	x	x
Crocothemis	2	x			x			
Cyanallagma	3	x	x	x	x	x	x	x
Diastatops	13	x	x	x	x	x	x	x
Didymops	6	x	x	x	x	x	x	x
Dorocordulia	17	x	x	x	x	x	x	x
Drepanoneura	1	x	x	x	x	x	x	x
Dromogomphus	45	x	x	x	x	x	x	x
Dythemis	14	x	x	x	x	x	x	x
Enallagma	264	x	x	x	x	x	x	x
Epiaeschna	22	x	x	x	x	x	x	x
Epipleoneura	2	x						
Epitheca	130	x	x	x	x	x	x	x
Erpetogomphus	57	x	x	x	x	x	x	x
Erythemis	174	x	x	x	x	x	x	x
Erythrodiplax	107	x	x	x	x	x	x	x
Euthore	1	x	x	x	x	x	x	x
Fluminagrion	1	x	x	x	x	x	x	x
Gomphaeschna	17	x	x	x	x	x	x	x
Gomphurus	160	x	x	x	x	x	x	x
Gynacantha	29	x	x	x	x	x	x	x
Hagenius	17	x	x	x	x	x	x	x
Helocordulia	13	x	x	x	x	x	x	x
Hesperagrion	15	x	x	x	x	x	x	x
Hetaerina	75	x	x	x	x	x	x	x
Heteragrion	2							
Hylogomphus	62	x	x	x	x	x	x	x
Idiataphe	1	x	x	x	x	x	x	x
Iridictyon	3	x	x	x	x	x	x	x
Ischnura	83	x	x	x	x	x	x	x
Ladona	55	x	x	x	x	x	x	x
Lanthus	21	x	x	x	x	x	x	x
Leptobasis	2				x			
Lestes	187	x			x			

Genus	Count	LDA	LR	NB	CART	KNN	SVM	RFC
Leucorrhinia	70	x	x	x	x	x	x	x
Libellula	342	x	x	x	x	x	x	x
Macrodiplax	9	x	x	x	x	x	x	x
Macromia	40	x	x	x	x	x	x	x
Macrothemis	43	x	x	x	x	x	x	x
Mecistogaster	1							
Mesamphiagrion	4	x	x	x	x	x	x	x
Miathyria	17	x	x	x	x	x	x	x
Micrathyria	84	x	x	x	x	x	x	x
Misagria	2	x	x	x	x	x	x	x
Mnesarete	5	x	x	x	x	x	x	x
Nannothemis	18	x	x	x	x	x	x	x
Nasiaeschna	14	x	x	x	x	x	x	x
Nehalennia	1				x			
Neoerythromma	1							
Neoneura	1	x						
Nephepeltia	6	x	x	x	x	x	x	x
Neurocordulia	7	x	x	x	x	x	x	x
Octogomphus	9	x	x	x	x	x	x	x
Oligoclada	1	x						
Ophiogomphus	88	x	x	x	x	x	x	x
Oplonaeschna	1							
Orthemis	35	x	x	x	x	x	x	x
Pachydiplax	68	x	x	x	x	x	x	x
Palaemnema	3				x			
Paltothemis	14	x	x	x	x	x	x	x
Pantala	85	x	x	x	x	x	x	x
Perithemis	81	x	x	x	x	x	x	x
Phanogomphus	233	x	x	x	x	x	x	x
Phyllocycla	17	x	x	x	x	x	x	x
Phyllogomphoides	29	x	x	x	x	x	x	x
Plathemis	88	x	x	x	x	x	x	x
Polythore	220	x	x	x	x	x	x	x
Progomphus	62	x	x	x	x	x	x	x
Protoneura	4	x	x	x	x	x	x	x
Pseudoleon	18	x	x	x	x	x	x	x
Remartinia	1	x	x	x	x	x	x	x
Rhionaeschna	3							
Rhodopygia	2	x	x	x	x	x	x	x
Rimanella	1	x	x	x	x	x	x	x
Somatochlora	55	x	x	x	x	x	x	x
Staurophlebia	1	x	x	x	x	x	x	x
Stenocora	1	x			x			
Stenogomphurus	7	x	x	x	x	x	x	x
Stylogomphus	18	x	x	x	x	x	x	x
Stylurus	38	x	x	x	x	x	x	x
Sympetrum	146	x	x	x	x	x	x	x
Tachopteryx	10	x	x	x	x	x	x	x
Tanypteryx	3	x	x	x	x	x	x	x
Tauriphila	21	x	x	x	x	x	x	x
Telebasis	6	x	x	x	x	x	x	x
Tholymis	9	x	x	x	x	x	x	x
Tramea	86	x	x	x	x	x	x	x
Triacanthagyna	13	x	x	x	x	x	x	x
Tuberculobasis	1							
Uracis	9	x	x	x	x	x	x	x
Zenithoptera	4	x	x	x	x	x	x	x
Total Not found		11	19	19	12	17	19	19

Sáenz Oviedo, Kuhn, ... & Sanchez-Herrera

Are wing contours good classifiers for automatic identification in Odonata?

are an array of flight behaviors (fliers, gliders and perchers; Corbett & May, 2008) and these flight styles can be reflected in the similarities found in wing contours within both our observed confusion groups. For example, in migratory species of libellulids' hindwings can show convergence towards a wing planform that favors the gliding flight as an energy saving strategy (Suarez-Tovar & Sarmiento; 2016). For zygopterans, their flight is more passive, and their ability to disperse might be associated with slow flight or overflight (Bomphrey et al., 2016), which would explain any similarities in wing contours for coenagrionids and lestids. Comparisons of the damping ratios and natural frequencies of two dragonfly and two damselfly species, shows that for the anisopterans damping properties between fore- and hindwings were significantly different, while in zygopterans there were no or very weak differences in the damping ratios between both wings, suggesting that the structural design and wing shape can influence the aerodynamics of their flight behaviors (Lietz et al; 2021). In addition, functional morphology traits of the wings, like types of joints of the wing veins, spines and presence of resilin, a protein that gives certain flexibility to the wings of insects can be evaluated in this groups, like previously done by Appel and Gorb (2014) to understand the wing contour similarities in these taxa.

Overall, our results suggest that the wing contours by themselves can discriminate with a moderate accuracy and precision, in comparison with other wing attributes obtained using high resolution images. In addition, we tested multiple classifying algorithms for the contours, where LDA had the best performance.

Acknowledgements

References

Anaconda Software Distribution. (2016). https://anaconda.com

Appel, E. & Gorb, S. N. (2014). Zoologica Comparative functional morphology of vein joints in Odonata. *Zoologica, 159*.

Blidh, H. (2016). *Python implementation of "Elliptic Fourier Features of a Closed Contour."* https://pyefd.readthedocs.io/en/latest/

Bomphrey, R. J., Nakata, T., Henningsson, P. & Lin, H. T. (2016). Flight of the dragonflies and damselflies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1704). doi:10.1098/rstb.2015.0389

Brownlee, J. (2017). *Master Machine Learning Algorithms* (1.12). Machine Learning Mastery. https://machinelearningmastery.com/master-machine-learning-algorithms/

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M. & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*, 5. doi:10.1126/sciadv.1400253

Corbett, P. S. & May, M. L. (2008). Fliers and perchers among Odonata: dichotomy or multidimensional continuum? A provisional reappraisal. *International Journal of Odonatology*, *11*(2), 155–171. doi:10.1080/13887890.2008.9748320

Córdoba-Aguilar, A. (2008). Dragonflies and Damselflies: Model Organisms for Ecological and Evolutionary Research. In *Dragonflies and Damselflies: Model Organisms for Ecological and Evolutionary Research*. doi:10.1093/acprof:oso/ 9780199230693.001.0001

Gandhi, R. (2018). *Support Vector Machine—Introduction to Machine Learning Algorithms*. Towards Data Science. https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47

Garrison, Rosser W., von Ellenrieder, N. & Louton, J. A. (2006). Dragonfly genera of the New World: an illustrated and annotated key to the Anisoptera. In *Choice Reviews Online*. Johns Hopkins University Press.

González, A. (2009). El conocimiento sistemático impedimento taxonómico la biodiversidad y. *Revista de La Sociedad Española de Biologia Evolutiva*, *4*(1), 19–32.

Harrington, P. (2012). *Machine Learning in Action Ill MANNING Shelter Island*. Manning Publications Co.

Hassall, C. (2014). Continental variation in wing pigmentation in calopteryx damselflies is related to the presence of heterospecifics. *PeerJ*, *2014*(1), e438. doi:10.7717/peerj.438

Heckman, C. W. (2006). Encyclopedia of South American Aquatic Insects: Odonata – Anisoptera. In *Encyclopedia of South American Aquatic Insects: Odonata – Anisoptera*. The Netherlands: Springer. doi:10.1007/978-1-4020-4802-5

Heckman, C. W. (2008). Encyclopedia of South American Aquatic Insects: Odonata – Zygoptera. In *Encyclopedia of South American Aquatic Insects: Odonata – Zygoptera*. The Netherlands: Springer. doi:10.1007/978-1-4020-8176-7

Kalkman, V. J., Clausnitzer, V., Dijkstra, K. D. B., Orr, A. G., Paulson, D. R. & van Tol, J. (2008). Global diversity of dragonflies (Odonata) in freshwater. *Hydrobiologia*, *595*(1), 351–363. doi:10.1007/s10750-007-9029-x

Koh, L. P., Dunn, R. R., Sodhi, N. S., Colwell, R. K., Proctor, H. C. & Smith, V. S. (2004). Species Coextinctions and the Biodiversity Crisis. *Science*, *305*(September), 1632–1635. doi:10.1126/science.1101101

Kuhn, W. R. (2016). *Three approaches to automating taxonomy, with emphasis on the Odonata (dragonflies and damselflies)*. (Thesis). Rutgers, The State University of New Jersey.

la Salle, J., Williams, K. J. & Moritz, C. (2016). Biodiversity analysis in the digital era. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702). doi:10.1098/rstb.2015.0337

Lietz, C., Schaber, C.F., Gorb, S.N. et al. (2021) The damping and structural properties of dragonfly and damselfly wings during dynamic movement. *Commun Biol 4*, 737. doi:10.1038/s42003-021-02263-2

Lorenzo-Carballa, M. O. & Cordero Rivera, A. (2012). Odonatos. In P. Vargas & R. Zardoya (Eds.), *El árbol de la Vida: sistemática y evolución de los seres vivos.* pp. 293–301.

Moore, N. W. (1997). *Dragonflies: Status Survey and Conservation Action Plan*. Gland, Switzerland, and Cambridge, UK: IUCN.

Narayan, Y. (2021). Hb vsEMG signal classification with time domain and Frequency domain features using LDA and ANN classifier. *Materials Today: Proceedings*, *37,* 3226–3230. doi:10.1016/j.matpr.2020.09.091

Pal, K. K. & Sudeep, K. S. (2016). Preprocessing for image classification by convolutional neural networks. *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 1778–1781. doi:10.1109/RTEICT.2016.7808140

Paulson, D. R., Schorr, M. & Deliry, C. (2021). *World Odonata List · University of Puget Sound*. https://www2.pugetsound.edu/academics/academic-resources/slater-museum/biodiversity-resources/dragonflies/world-odonata-list2/

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*(85), 2825–2830.

Raybaut, P. (2009). *Spyder IDE* (4.2.1). Pythonhosted. https://www.spyder-ide.org/

Samways, M. J. & Steytler, N. S. (1996). Dragonfly (Odonata) distribution patterns in urban and forest landscapes, and recommendations for riparian management. *Biological Conservation*. doi:10.1016/S0006-3207(96)00032-8

Shahriar, M. T. & Li, H. (2020). A Study of Image Pre-processing for Faster Object Recognition. *ArXiv*, *October 2020*. https://arxiv.org/abs/2011.06928

Sharma, P., Hans, P. & Gupta, S. C. (2020). Classification of plant leaf diseases using machine learning and image preprocessing techniques. *Proceedings of the Confluence 2020 – 10th International Conference on Cloud Computing, Data Science and Engineering*, 480–484. doi:10.1109/Confluence47617.2020.9057889

Suárez-Tovar, C. M. & Sarmiento, C. E. (2016), Beyond the wing planform: morphological differentiation between migratory and non-migratory dragonfly species. *Journal of Evolutionary Biology, 29*, 690-703. doi:10.1111/jeb.12830

Tharwat, A., Gaber, T., Ibrahim, A. & Hassanien, A. E. (2017). Linear discriminant analysis: A detailed tutorial. *AI Communications, 30*(2), 169–190. doi:10.3233/AIC-170729

van Rossum, G., & Drake Jr, F. L. (2009). Python 3 Reference Manual. CreateSpace. https://dl.acm.org/doi/book/10.5555/1593511

Supplementary material

Supplementary Figure 1. Random Forest Classifier Confusion Matrix.

Supplementary Figure 2. Support Vector Machines Confusion Matrix.

Supplementary Figure 3. K-Nearest Neighbors Confusion Matrix.

Supplementary Figure 4. Classification and Regression Trees Confusion Matrix.

Supplementary Figure 5. Naïve Bayes Confusion Matrix.

Supplementary Figure 6. Logistic Regression Confusion Matrix.

Supplementary Table 1. Taxonomic information of specimens included in the analysis.

Supplementary Table 2. Training accuracy scores.

Supplementary Table 3. Summary of classification report: Number of classes found, accuracy, precision, recall, F1 score and support values of the classifiers tested.

Supplementary Table 4. ANOVA results for accuracy scores comparison.

Supplementary Table 5. Tukey multiple comparisons test for accuracy scores.

Supplementary Table 6. ANOVA results for F1 scores comparison.

Supplementary Table 7. Tukey multiple comparisons test for F1 scores.