Spatiotemporal Features Improve Fine-Grained Butterfly Image Classification

Marta Skreta
Department of Computer Science
University of Toronto
Toronto, ON
martaskreta@cs.toronto.edu

Alexandra Luccioni
Department of Computer Science and Operations Research
Université de Montréal
Montreal, QC
sasha.luccioni@mila.quebec

David Rolnick
School of Computer Science
McGill University
Montreal, QC
drolnick@cs.mcgill.ca

Abstract

Understanding the changing distributions of butterflies gives insight into the impacts of climate change across ecosystems and is a prerequisite for conservation efforts. eButterfly is a citizen science website created to allow people to track the butterfly species around them and use these observations to contribute to research. However, correctly identifying butterfly species is a challenging task for non-specialists and currently requires the involvement of entomologists to verify the labels of novice users on the website. We have developed a computer vision model to label butterfly images from eButterfly automatically, decreasing the need for human experts. We employ a model that incorporates geographic and temporal information of where and when the image was taken, in addition to the image itself. We show that we can successfully apply this spatiotemporal model for fine-grained image recognition, significantly improving the accuracy of our classification model compared to a baseline image recognition system trained on the same dataset.

1 Introduction

Insect populations are plummeting on a global scale, including butterflies. One of the reasons for this is that butterflies are highly susceptible to changes in their surrounding environment, such as climate change, habitat loss, and pesticide use [1]. In fact, some species of butterflies are directly impacted by climate change through their growth and reproduction cycles, such as the Karner Blue butterfly (*Lycaeides melissa samuelis*), which is an endangered species. This is in part because warmer temperatures can induce butterfly eggs to hatch more quickly while their food source remains dormant, resulting in population decline [2]. Butterflies can also be indirectly impacted by climate change if their habitat becomes altered. For example, the Rocky Mountain Apollo (*Parnassius smintheus*) and the Monarch (*Danaus plexippus*) have both incurred ecosystem reduction due to indirect effects of climate change on their host plants [3, 4].

Apart from the loss of beautiful and culturally meaningful animals, a reduction in butterfly biodiversity could have dire consequences, since butterflies play a key role in a multitude of ecosystems. For instance, butterflies are important pollinators of many species of flowering plants. A reduction in butterfly populations, or even a change in the timing of butterfly life cycles, could mean that certain flower species experience a significant drop in pollination [5]. Butterflies are also an important food

Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

source for many animals, including birds, lizards, and other insects [6, 7, 8]. Thus, any changes in butterfly populations can trigger a “butterfly effect" and impact other parts of ecosystems, potentially resulting in far-reaching repercussions.

In order to understand changes in butterfly populations and employ appropriate conservation measures, it is important to track the distribution and abundance of butterflies. eButterfly is a website developed by a coalition of researchers across the United States and Canada, to which individuals are able to upload images of butterflies from any location and identify the species to which they belong [9]. Analyzing these butterfly images and the observational data that accompanies them can offer valuable information on changes in geographical and temporal distribution. However, it is important to first obtain accurate labels for individual butterfly species. While any user of the eButterfly website can upload an image, they may not necessarily know the species, since fine-grained butterfly identification can be challenging for novice entomologists. We leverage machine learning to automate the labelling process, allowing experts to focus only on the most challenging identifications and alleviating a bottleneck that has held back the eButterfly project.

2 Related Work

There has been extensive work on using images taken by photographers or camera traps for automatic identification of animal species, including birds, mammals, ground-dwelling insects, bees, and more (see e.g. [10, 11, 12, 13, 14, 15, 16]). Recent work has proposed to improve such image-based identification systems using complementary information, such as spatiotemporal distribution [17, 18]. The motivation behind this is that visually similar species may be present in different geographic regions and at different periods of the year, and therefore knowing where and when a picture was taken may be useful information for fine-grained classification. The authors of [17] developed a spatiotemporal prior that estimates the probability of a species being present based on where and when the image was taken, and applied this model especially to bird identification.

However, little work has applied such ideas to the problem of butterfly classification, which to date has been carried out largely based on pure image data [19, 20]. Our approach uses features extracted from both the image and its accompanying information, such as geolocation and date. To our knowledge, ours is the first approach that utilizes both image data and auxiliary features for fine-grained butterfly species monitoring. This approach is highly relevant as many butterfly species have highly restricted distributions, even within North America. For example, Figure 1 shows images of two species belonging to the genus *Speyeria*. Although they are visually similar, one is found primarily in Eastern North America and the other exclusively in the West.

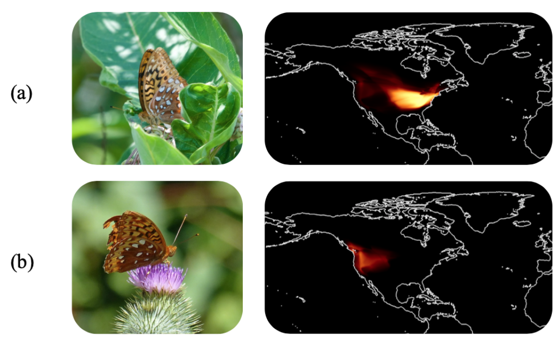

Figure 1: Images and geographic ranges of (a) *Speyeria cybele* and (b) *Speyeria zerene* in the month of June from the eButterfly database. Although they are visually similar, one is found primarily in Eastern North America and the other exclusively in the West.

3 Dataset

Our work uses images collected via the eButterfly website by citizen scientists across North America [9]. As of October 2020, eButterfly contains over 100,000 hand-labelled images and 400,000 geolocation observations for over 600 butterfly species, all verified by expert entomologists. (While we have partnered directly with eButterfly, this data is also available for use by other researchers on request.) As different butterflies have different abundances, the dataset is imbalanced: over 400 species contain less than 100 observations, while the remaining species have up to 2,700 images.

In some of our experiments, we increase the representation of rare species slightly, so that the model has enough data to learn but remains less likely to predict them than more common species. To do this, we augment the eButterfly dataset with observation data from the iNaturalist project, a publicly-available dataset of images of animals, plants, and other organisms [21]. We use the iNaturalist data to ensure that all species in the dataset have at least 100 observations. If there are not enough unique images available in the combined iNaturalist and eButterfly datasets, we sample the images with replacement until 100 images are collected per species.

4 Methods

The baseline for our model is an image-only deep-learning-based identification model developed by Kantor et al [22]. The model uses the ResNet-50 architecture [23]. To train the model, we remove the final two linear layers of ResNet-50 and learn them from scratch to take into account the different number of classes in eButterfly.

The goal of our approach is to build on the image-based model by learning a spatiotemporal prior that encodes the presence of a species given geographic and temporal data associated with the images. As in [17], we train two encoder models (Figure 2): the first learns the probability of a given species being present in the image with the form $P(y|\mathbf{I})$ where y is the class and $\mathbf{I}$ is the image. The second model is trained to estimate the species from spatiotemporal features with $P(y|\mathbf{x})$, where $\mathbf{x}$ is the concatenation of the image's longitude, latitude, and capture date. Our model is a neural network containing 9 fully connected layers with skip connections between them, and the loss function is a variation of the cross-entropy loss where we multiply the penalty of false negatives by a large constant (empirically we find that 10 gives best performance). Upweighting false negatives encourages the model to extrapolate the presence of species between discrete observation locations.

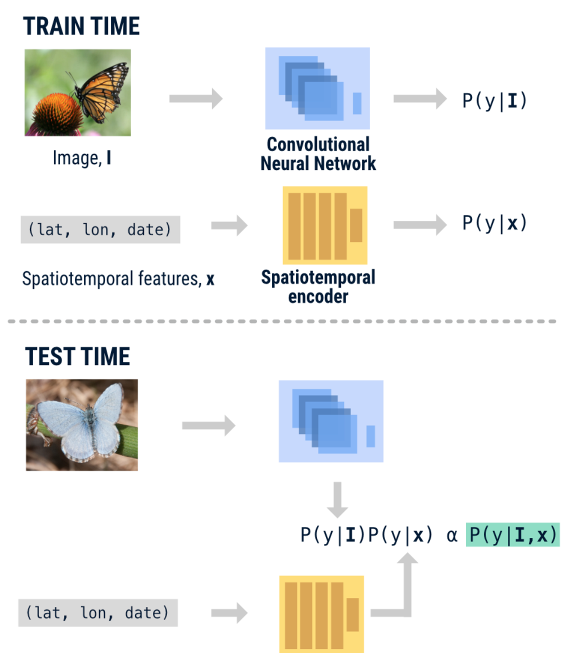

Figure 2: Schematic of incorporating spatiotemporal features. During training, we predict the butterfly species from the corresponding image and geo-spatiotemporal data independently. At test time, we use the output from the spatiotemporal model as a Bayesian prior.

We assume that $\mathbf{x}$ and $\mathbf{I}$ are conditionally independent given y; in this case, we have the relationship $P(y|\mathbf{I};\mathbf{x}) \propto P(y|\mathbf{I})P(y|\mathbf{x})$, which allows us to use our spatiotemporal model as a Bayesian prior at test time. To obtain the prediction for a test sample, we pass a butterfly image through the image classifier and its corresponding spatiotemporal information through the spatiotemporal model separately. We then multiply the outputs from both models to obtain the final class prediction.

5 Results

The metrics that we track and report are *macro accuracy*, which is the total number of correct observations divided by the total observations, and *micro accuracy*, which is the average of model performance on each class. The latter metric is more challenging when considering imbalanced datasets such as ours, since it treats all classes equally, while the former is a better representation of performance on data "in the wild" as it reflects the natural abundances of different species.

We find that using geographic and temporal information increases our model's macro and micro accuracies by 2% and 6%, respectively (Table 1). This demonstrates that incorporating additional features to the image classification model has the potential to significantly improve final classification performance, especially for underrepresented classes, and calls for further exploration.

Table 1: Model performance on butterfly classification model using only images (Image only), incorporating coordinates of image (Image + (Lat, Lon)) and incorporating coordinates and date of image (Image + (Lat, Lon, Date))

Accuracy	Image only	Image + (Lat, Lon)	Image + (Lat, Lon, Date)
Top 1 Macro	84.56	86.16	**86.53**
Top 1 Micro	59.87	64.47	**65.65**
Top 3 Macro	93.84	95.06	**95.38**
Top 3 Micro	77.53	83.14	**83.74**

As can be seen in Table 1, there is a large difference between macro and micro accuracies, indicating that our model does significantly worse on species that are less common in the wild or otherwise underrepresented in the data. To palliate this, we augment our dataset with images from the iNaturalist dataset, as detailed in §3. Table 2 shows the impact of augmenting the eButterfly dataset with images from iNaturalist and ensuring each species has at least 100 images.

Table 2: Performance of butterfly image classification model on eButterfly dataset, combined eButterfly + iNaturalist (iNat) datasets, and combined eButterfly + iNat datasets with coordinate and date of image (Lat, Lon, Date), where iNaturalist is used to augment rare species.

Accuracy	eButterfly	eButterfly + iNat	eButterfly + iNat + (Lat, Lon, Date)
Top 1 Macro	84.56	84.94	**87.90**
Top 1 Micro	59.87	69.51	**75.73**
Top 3 Macro	93.84	93.94	**95.86**
Top 3 Micro	77.53	83.59	**89.38**

We find that augmenting rare classes with iNaturalist data improves the performance by 6-10% for the micro accuracy, indicating that rare species benefit from having more training data. The gain in macro accuracy is more modest since this metric is heavily influenced by species that are common, for which performance is not significantly altered by this method. When combining our spatiotemporal model with data augmentation, we see a performance boost of up to 15% compared to the baseline.

6 Conclusion & Future Work

In this work, we demonstrate that incorporating spatiotemporal information regarding where and when an image was taken can greatly improve the fine-grained classification of North American butterflies. We are working with the eButterfly project to deploy our model as a tool to recommend species identifications to eButterfly users, thereby reducing the time and effort needed for expert entomologists to vet all incoming observations.

In future work, we intend to improve our spatiotemporal prior by incorporating information from satellite images, which will allow us to directly model butterfly habitats (e.g. urban zones, forests and wetland areas). We are also planning to incorporate information from datasets on bird distribution, as there are correlations between co-occurrences of various bird and butterfly species and there is considerable information about birds that is widely available. Finally, we intend to leverage our spatiotemporal prior to redistribute probabilities between visually similar images, as captured by the confusion matrix of the image model. For example, if our image model predicts species A with high likelihood but the probability from the spatiotemporal prior is low, we can redistribute some of the probability mass on species A to species B, which is visually similar to A but has a higher value for the spatiotemporal prior. We believe these various innovations will further increase performance and lead to more effective tools for advancing citizen science.

References

[1] Francisco Sánchez-Bayo and Kris A.G. Wyckhuys. Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, 232:8–27, April 2019.

[2] Tamatha A. Patterson, Ralph Grundel, Jason D. K. Dzurisin, Randy L. Knutson, and Jessica J. Hellmann. Evidence of an extreme weather-induced phenological mismatch and a local extirpation of the endangered Karner blue butterfly. *Conservation Science and Practice*, 2(1), December 2019.

[3] Alessandro Filazzola, Stephen F. Matter, and Jens Roland. Inclusion of trophic interactions increases the vulnerability of an alpine butterfly species to climate change. *Global Change Biology*, 26(5):2867–2877, March 2020.

[4] Jessica J Hellmann, Ralph Grundel, Chris Hoving, and Gregor W Schuurman. A call to insect scientists: challenges and opportunities of managing insect communities under climate change. *Current Opinion in Insect Science*, 17:92–97, October 2016.

[5] Josef Settele, Jacob Bishop, and Simon G. Potts. Climate change impacts on pollination. *Nature Plants*, 2(7), July 2016.

[6] Stephen B. Malcolm. Anthropogenic impacts on mortality and population viability of the monarch butterfly. *Annual Review of Entomology*, 63(1):277–302, January 2018.

[7] Dheeraj Halali, Athul Krishna, Ullasa Kodandaramaiah, and Freerk Molleman. Lizards as predators of butterflies: Shape of wing damage and effects of eyespots. *The Journal of the Lepidopterists' Society*, 73(2):78, December 2019.

[8] Sara L. Hermann, Carissa Blackledge, Nathan L. Haan, Andrew T. Myers, and Douglas A. Landis. Predators of monarch butterfly eggs and neonate larvae are more diverse than previously recognised. *Scientific Reports*, 9(1), October 2019.

[9] Kathleen Prudic, Kent McFarland, Jeffrey Oliver, Rebecca Hutchinson, Elizabeth Long, Jeremy Kerr, and Maxim Larrivée. eButterfly: Leveraging massive online citizen science for butterfly conservation. *Insects*, 8(2):53, May 2017.

[10] Thomas Berg, Jiongxin Liu, Seung Woo Lee, Michelle L Alexander, David W Jacobs, and Peter N Belhumeur. Birdsnap: Large-scale fine-grained visual categorization of birds. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2011–2018, 2014.

[11] Grant Van Horn, Steve Branson, Ryan Farrell, Scott Haber, Jessie Barry, Panos Ipeirotis, Pietro Perona, and Serge Belongie. Building a bird recognition app and large scale dataset with citizen scientists: The fine print in fine-grained dataset collection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 595–604, 2015.

[12] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. The iNaturalist species classification and detection dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8769–8778, 2018.

[13] Stefan Schneider, Saul Greenberg, Graham W. Taylor, and Stefan C. Kremer. Three critical factors affecting automated image species recognition performance for camera traps. *Ecology and Evolution*, 10(7):3503–3517, March 2020.

[14] Mohammad Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, and Jeff Clune. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, 115(25):E5716–E5725, June 2018.

[15] Oskar L. P. Hansen, Jens-Christian Svenning, Kent Olsen, Steen Dupont, Beulah H. Garner, Alexandros Iosifidis, Benjamin W. Price, and Toke T. Høye. Species-level image classification with convolutional neural network enables insect identification from habitus images. *Ecology and Evolution*, 10(2):737–747, December 2019.

[16] Ivan F. Rodriguez, Remi Megret, Edgar Acuna, Jose L. Agosto-Rivera, and Tugrul Giray. Recognition of pollen-bearing bees from video using convolutional neural network. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, March 2018.

[17] Oisin Mac Aodha, Elijah Cole, and Pietro Perona. Presence-only geographical priors for fine-grained image classification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.

[18] Grace Chu, Brian Potetz, Weijun Wang, Andrew Howard, Yang Song, Fernando Brucher, Thomas Leung, and Hartwig Adam. Geo-aware networks for fine-grained recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, Oct 2019.

[19] Yılmaz Kaya, Lokman Kayci, Ramazan Tekin, and Ö Faruk Ertuğrul. Evaluation of texture features for automatic detecting butterfly species using extreme learning machine. *Journal of Experimental & Theoretical Artificial Intelligence*, 26(2):267–281, 2014.

[20] Lin Nie, Keze Wang, Xiaoling Fan, and Yuefang Gao. Fine-grained butterfly recognition with deep residual networks: A new baseline and benchmark. In *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–7. IEEE, 2017.

[21] Ken-Ichi Ueda. iNaturalist research-grade observations, 2020.

[22] C. Kantor, B. Rauby, E. Jehanno, L. Boussioux, A. Luccioni, H. Talbot. Guided attention for fine-grained and hierarchical classification, 2020.

[23] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2016.