

Overview of the SnakeCLEF 2020: Automatic Snake Species Identification Challenge

Lukáš Pícek^{1,4} , Isabelle Bolon² , Andrew M. Durso³ ,
Rafael Ruiz de Castañeda² 

¹ Department of Cybernetics, FAV, University of West Bohemia, Czechia

² Institute of Global Health, Department of Community Health and Medicine,
University of Geneva, Switzerland

³ Department of Biological Sciences, Florida Gulf Coast University, USA

⁴ PiVa AI s.r.o, Czechia

Abstract. Building a robust and accurate AI-driven system for automatic snake species identification is an important goal for biodiversity and global health. As the existence of such a system can potentially help to lower deaths and disabilities caused by snakebites, we have prepared SnakeCLEF2020: Automatic Snake Species Identification Challenge, which provides an evaluation platform and labeled data (including geographical information) for biodiversity and health research purposes. SnakeCLEF 2020 was designed to provide an evaluation platform that can help track the performance of end-to-end AI-driven snake species recognition systems. We have collected 287,632 images of 783 snake species from 145 countries. Here we report 1) a description of the provided data, 2) evaluation methodology and principles, 3) an overview of the systems submitted by the participating teams, and 4) a discussion of the obtained results.

Keywords: LifeCLEF, SnakeCLEF, fine grained visual categorization, global health, epidemiology, snake bite, snake, reptile, benchmark, biodiversity, species identification, machine learning, computer vision, classification

1 Introduction

Creating an automatic and robust system for snake species identification is an important goal for biodiversity, conservation, and global health. With recent estimates of 81,410-137,880 deaths and 435,000-580,000 victims of permanent disability and disfigurement (globally each year) [38,40] caused by venomous snakebite, understanding the geographical distribution of the more than 3,800 species of snakes and differentiating species from images (particularly images of low quality) can significantly improve epidemiology data and treatment

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CLEF 2020, 22-25 September 2020, Thessaloniki, Greece.



Fig. 1. Four observations of the same snake species (Boomslang, *Dispholidus typus*) with high visual dissimilarity related to sex and age (female top left and bottom right, male top right, juvenile bottom left). ©Mark Heystek, *iNaturalist*, CC-BY-NC; ©Daniel Rautenbach, *iNaturalist*, CC-BY-NC; ©Bart Wursten, *iNaturalist*, CC-BY-NC; ©Margriet B, *iNaturalist*, CC-BY-NC.

outcomes. The goals and usage of image-based snake identification are complementary with those of other LifeCLEF challenges: 1) classifying snake species in images, 2) predicting the list of species that are the most likely to be observed at a given location, and 3) eventually developing automated tools that can facilitate the integration of changing taxonomies and discoveries.

Having a system that is capable of recognizing or differentiating snake species from images could significantly improve snakebite eco-epidemiological data and snakebite clinical management (i.e., correct antivenom administration) and patient outcomes [4,6]. Although only about 20% of snake species worldwide are medically-important [21], our knowledge of their epidemiological importance (i.e., how many snakebites each are responsible for in different areas) is incomplete. Most people are not capable of accurately identifying snakes to species [7,11,42], and trained herpetologists are relatively few in number. Thus, automated snake species identification would have value in public health. There are only a few initiatives that seek to identify snakes using computer vision techniques. So far, a handful of computer vision and machine learning algorithms specific to snakes have been developed, but these can only identify a few hand-picked species in



Fig. 2. Randomly selected images from the SnakeCLEF 2020 training set. ©stewartb, iNaturalist, CC-BY-NC; ©Jennifer Linde, iNaturalist, CC-BY-NC; ©Gilberto Ponce Tejada, iNaturalist, CC-BY-NC; © Ryan van Huyssteen, iNaturalist, CC-BY-NC; ©Jessica Newbern, iNaturalist, CC-BY-NC.

the simplest cases [1,9,15,25]. Two larger-scale initiatives use computer vision and machine learning algorithms to identify species of reptiles and amphibians (HerpMapper’s Fitch bot¹) or animals and plants more generally (iNaturalist² [34]), both in connection with Visipedia³ [2], each include hundreds of species of snakes but have not been evaluated specifically with snakes in mind. None of these are yet usable in real-world situations where lives may be at stake.

Since snake species identification is a fine-grained visual categorization task, the main difficulty of this challenge is high intra-class and low inter-class variance [29]. In other words, certain classes could be highly variable in appearance depending on geographic location, sex, or age (Figure 1) and at the same time could be visually similar to other species (e.g., mimicry) (Figure 3).

2 Dataset and Evaluation Protocol

For this challenge, we prepared a large dataset with 287,632 photographs belonging to 783 snake species and taken in 145 countries. The majority of the data were gathered from online biodiversity platforms (i.e., iNaturalist⁴, HerpMapper⁵) and were further extended by data scraped from Flickr⁶. Furthermore, we have assembled a total of 28,418 images from private collections and museums. The final dataset has a heavy long-tailed class distribution, where the most frequent species (*Thamnophis sirtalis*) is represented by 14,433 images and the least frequent by just 20 (*Naja pallida*). Such a distribution with small inter-class variance, high intra-class variance, and a high number of species (classes) creates a challenging task even for current state-of-the-art classification approaches.

¹ <https://whattheherp.com/>

² https://www.inaturalist.org/computer_vision_demo

³ <https://vision.cornell.edu/se3/projects/visipedia/>

⁴ <https://www.inaturalist.org/>

⁵ <https://www.herpmapper.org/>

⁶ <https://www.flickr.com/>



Fig. 3. Medically important snake species (left) and similar-looking non-venomous species (right). [a] Bibron's Stiletto Snake (*Atractaspis bibronii*), [b] Common Purple-glossed Snake (*Amblyodipsas polylepis*), [c] Variable Coralsnake (*Micrurus diastema*), [d] Variegated False Coralsnake (*Pliocercus elapoides*), [e] Amazonian Palm Viper (*Bothrops bilineatus*), [f] Guianan Green Snake (*Xenodon weneri*), [g] Rhombic Night Adder (*Causus rhombeatus*), [h] African Egg-eating Snake (*Dasypeltis scabra*). ©Peter Vos, *iNaturalist*, CC-BY-NC; ©Alex Rebelo, *iNaturalist*, CC-BY-NC; ©schooten, *iNaturalist*, CC-BY-NC; ©Iris Melgar, *iNaturalist*, CC-BY-NC; ©Jason Fantuzzi, *iNaturalist*, CC-BY-NC; ©William van Niekerk, *iNaturalist*, CC-BY-NC; ©Josh Vandermeulen, *iNaturalist*, CC-BY-NC; ©Sébastien Sant, *iNaturalist*, CC-BY-NC.

Table 1: Details of the SnakeCLEF 2020 dataset split into the training, validation and testing sets.

Subset	# of images	% of data	min. # of images/class
Training	245,185	85.24%	17
Validation	14,029	4.88%	2
Testing	28,418	9.88%	1
Total	287,632	100	20

2.1 Training and validation split

To allow participants to validate their intermediate results easily, we have split the full dataset into a training subset with 245,185 images, and validation subset with 14,029 images. Both subsets have similar class distribution, while the minimum number of validation images per class is one. Example images from the training subset are depicted in Figure 2.

2.2 Testing set

Unlike other LifeCLEF challenges, the final testing set remains undisclosed as it is composed of private images from individuals and natural history museums who have not put those images online in any form. A brief description of this final testing set is as follows: twice as big as the validation set, contains all 783 classes, similar class distribution, and observations from almost all the countries presented in training and validation sets.

2.3 Geographical Information

Considering that all snake species have distinct, largely stable geographic ranges, with a maximum of more than 125 species of snakes occurring within the same $50 \times 50 \text{ km}^2$ area [24], geographical information probably plays a crucial role in correct snake species identification [41]. To evaluate this, we have gathered two levels of the geographical label (i.e., country and continent) for approximately 80% of the data. We have collected observations across 145 countries and all continents. A small proportion of images (ca. 1-2%), particularly from Flickr, contain captive snakes that are kept outside of their native range (e.g., North American *Pantherophis guttatus* in Europe or Australian *Morelia viridis* in the USA). We opted to retain these for three reasons:

1. Users of an automated identification system may wish to use it on captive snakes (e.g., in the case of customs seizures [12,19]).
2. Bites from captive snakes may occur (although the identity of the snake would normally be clear in this case; e.g. [28,37]).
3. Captive snakes sometimes escape and can found introduced populations outside their native range (e.g. [5,17]).

2.4 Evaluation Protocol

The main goal of this challenge was to build a system that is autonomously able to recognize 783 snake species based on the given image and geographical location input. Every participant had to submit their whole solution into the GitLab-based evaluation system (hosted on AICrowd platform⁷), which performed evaluation over the undisclosed testing set. Since data were secret, each participating team could submit up to 5 submissions per day. The primary evaluation metric for this challenge was the macro-averaged Dice Similarity Coefficient (DSC), also known as macro-averaged F1 score (F1), which is not biased by class frequencies. The Macro F1 score is defined as the mean of class-wise/species-wise F1 scores:

$$\text{Macro F}_1 = \frac{1}{N} \sum_{i=0}^N F_{1_i},$$

where i is the species index and N the number of classes/species. Final Macro F1 is performed by first computing the F1 score for each class/species as harmonic mean of the Precision and the Recall.

$$F_1 = 2 \times \frac{\textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}},$$

$$\textit{Precision} = \frac{TP}{TP + FN}, \quad \textit{Recall} = \frac{TP}{TP + FN},$$

The secondary metric was calculated as Multi-class Classification Logarithmic Loss e.g., Cross Entropy Loss:

$$L(p, q) = - \sum_x p(x) \cdot \log(q(x)),$$

where x is the index of the class, p is the true distribution (onehot vector) and q is the predicted distribution (softmax). This metric considers the uncertainty of a given prediction based on how much it differs from the actual label. This gives us a more subtle evaluation of the performance.

3 Participants and Methods

Out of 8 registered teams in the SnakeCLEF 2020 challenge, only two teams managed to submit a working version of their recognition system that takes an image and location as an input and returns softmax prediction values. Even though participants were able to evaluate their system five times a day, we registered only 27 successful submissions. All submissions and their achieved scores are reported in Table 1. Detailed description for each run, evaluated experiments and used methods, techniques and experiments are further developed in individual working notes (FHDO_BCSG [3], Gokuleloop [18]).

⁷ <https://www.aicrowd.com>

FHDO_BCSG, Germany, 25 runs, [3]: The FHDO_BCSG team approach combined two stages. Firstly, they used a Mask R-CNN [10] instance detection framework to extract regions where each snake species occurs. Secondly, the different EfficientNet [31] models were utilized as a classifier over the extracted regions resized to (128×128) . Lastly, they multiplied the softmax values for each image by the species *a priori* probability for a given geographic location, estimated from training and validation sets. Their best-submitted model was an EfficientNet-B4 fine-tuned for 50 epochs from the ImageNet-1k pre-trained checkpoint. This model achieved an F1 score of 0.404 and a Log-Loss of 6.650. The high Log-Loss was achieved due to the application of softmax normalization after the multiplication of the location data, which leads to small differences in the predictions.

After the deadline of this challenge, the FHDO_BCSG team evaluated a few more runs that were excluded from the competition but produced interesting results. They experimented with higher input sizes (196×196 , 224×224 and 380×380), another object detection model (trained on the extended dataset), a longer training period ($50 \rightarrow 109$ epochs) and a slightly different method for geographical information integration (not performed for "unknown" location). By utilizing all of the above, they were able to achieve an F1 score of 0.594 and a Log-Loss of 1.064.

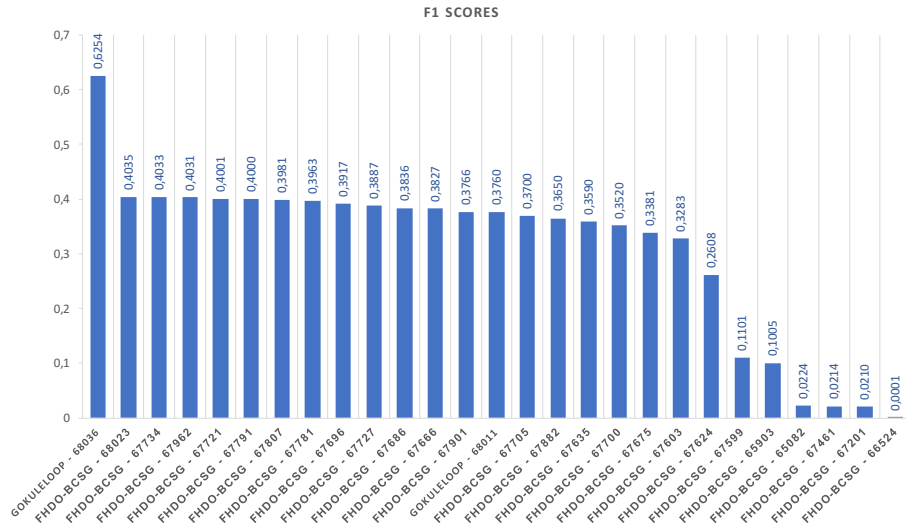


Fig. 4. Achieved F1 scores for all the submissions submitted into the SnakeCLEF 2020 Competition. Higher score is better.

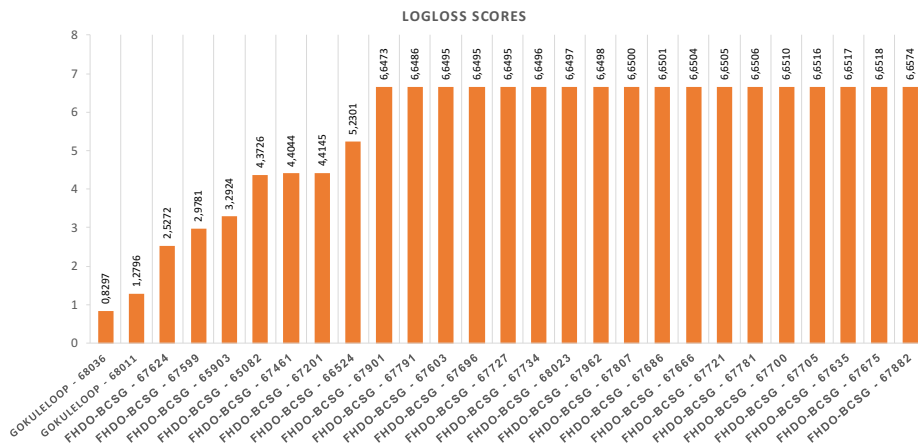


Fig. 5. Achieved Logarithmic Loss values for all the submissions submitted into the SnakeCLEF 2020 Competition. Smaller score is better.

gokuleloop, India, 2 runs, [18]: The Gokuleloop team approach was focused on domain-specific fine-tuning. This approach was inspired by an extensive study about the domain-specific image classification [16]. In a nutshell, Gokuleloop experimented with different pre-trained weights and their impact on the final performance. He fine-tuned the same CNN architecture (ResNet-50-V2) from different checkpoints i.e ImageNet-1k and ImageNet-21k. Input size (456×456), hyperparameter settings and augmentations were kept the same. Model comparison showed significant improvement in the case of ImageNet-21k checkpoint. Accuracy over the validation set was increased by 11.09% ($68.48\% \rightarrow 79.57\%$) and F1 score by 0.3113 ($0.27 \rightarrow 0.5813$). Finally, location metadata were incorporated via a naive probability weighting approach that increased the F1 score by 0.0206 over the validation set. The final system, which adopted a ResNet-50-V2 architecture fine-tuned from ImageNet-21k weights and a naive probability weighting approach, achieved a top F1 score of 0.625 while having a Log-Loss of 0.83.

4 Results and Discussion

All results achieved by the two successful participants of the SnakeCLEF 2020: Automatic Snake Species Identification Challenge - organized within the LifeCLEF 2020 Lab [14] - are reported in Table 2. Figure 4 and Figure 5 shows rearranged scores for both validation metrics, F1 and Logarithmic Loss. The best F1 score of 0.62536 achieved by gokuleloop shows an interesting performance and sets up the baseline for future research in this topic.

Table 2: Results of the SnakeCLEF 2020: Automatic Snake Species Identification Challenge, including macro-averaged F1 score and Cross Entropy Loss.

Rank	Team Name	Run ID	F1	Log Loss
#1	gokuleloop	68036	0.62536	0.82969
#2	FHDO-BCSG	68023	0.40354	6.64965
#3	FHDO-BCSG	67734	0.40334	6.64961
#4	FHDO-BCSG	67962	0.40313	6.64980
#5	FHDO-BCSG	67721	0.40005	6.65048
#6	FHDO-BCSG	67791	0.39997	6.64859
#7	FHDO-BCSG	67807	0.39810	6.64995
#8	FHDO-BCSG	67781	0.39631	6.65056
#9	FHDO-BCSG	67696	0.39167	6.65031
#10	FHDO-BCSG	67727	0.38870	6.64954
#11	FHDO-BCSG	67686	0.38363	6.65014
#12	FHDO-BCSG	67666	0.38266	6.65035
#13	FHDO-BCSG	67901	0.37658	6.64726
#14	gokuleloop	68011	0.37601	1.27955
#15	FHDO-BCSG	67705	0.36998	6.65164
#16	FHDO-BCSG	67882	0.36497	6.65741
#17	FHDO-BCSG	67635	0.35898	6.65168
#18	FHDO-BCSG	67700	0.35196	6.65104
#19	FHDO-BCSG	67675	0.33809	6.65178
#20	FHDO-BCSG	67703	0.32830	6.64953
#21	FHDO-BCSG	67624	0.26080	2.52722
#22	FHDO-BCSG	67599	0.11012	2.97812
#23	FHDO-BCSG	67903	0.10047	3.29237
#24	FHDO-BCSG	65082	0.02241	4.37263
#25	FHDO-BCSG	67461	0.02135	4.40436
#26	FHDO-BCSG	67201	0.02098	4.41450
#27	FHDO-BCSG	66524	0.00012	5.23012

4.1 Outcomes

Based on closer review of the participants' solutions, we derived the following outcomes:

Input resolution is related to CNN performance. As expected, input resolution had a significant influence on CNN performance. In the case of snake species recognition, with just subtle differences between species, resizing images to smaller dimensions might remove important information. The experiment performed by the FHDO_BCSG team showed that only a relatively small image dimension increase ($128 \times 128 \rightarrow 196 \times 196$) boosts the model performance F1 score by 0.083. Furthermore, gokuleloop used an input size of 456×456 and outperformed the best FHDO_BCSG solution by 0.22182 in terms of F1 score.

Data cleaning did not improve the performance. FHDO_BCSG team performed an in-depth data analysis and revealed that the provided training dataset contains approximately 4,000 duplicate images and about 4,000 images without a snake species in it. Interestingly, duplicate removal experiments over the testing set were inconclusive. In one iteration the system performance drops significantly; in a second, the system performed slightly better [3]. As the testing dataset was rigorously curated and does not include duplicates or images without snake species, the different behaviour does not have a clear explanation. One could perhaps assume that a small number ($\sim 3\%$) of False Positive images helps to prevent overfitting. Others have also found that machine learning algorithms are surprisingly robust to annotation errors if the training set has sufficient size [35].

There exists a significant impact of the pre-trained model. The experiment evaluated by gokuleloop showed that CNN performance in the domain of snake species recognition depends strongly on the pre-trained weights. Fine-tuning the same ResNet-50-V2 architecture from ImageNet-21k pre-trained weights rather than from ImageNet-1k increased the system performance by 0.3113 (F1 score) to the final value of 0.5813). Such a significant increase is remarkable, and the impact of pre-trained weights should be studied in greater depth.

Overall performance is relatively poor. Considering that current state-of-the-art methods are already capable of automatically detecting and recognizing a large number of plant [36] and animal species with human-level accuracy [20,23,30,33,39], the final F1 score of 0.62536 showed that snake identification is a much harder task with a lot of room for improvement.

Usage of geographical information showed a performance boost. In contrast to previously published research related to automatic snake identification, usage of geographical information was an essential part of the SnakeCLEF 2020 competition. Both teams significantly improved the system's performance by utilizing various techniques that used the provided location metadata. Gokuleloop improved his F1 score by 0.0206 and FHDO_BCSG by 0.125 (based on their post-competition experiment).

Some interesting ideas were evaluated. Experiments performed by both teams showed that recent state-of-the-art regularisation techniques have the potential to improve the overall performance. Additionally, two interesting ideas were tested. Namely:

Mixup augmentation [43]: Mixing two ground truth samples by the linear interpolation of their images and labels (one-hot labels). The interpolation is managed via alpha compositing.

Binary image branding [22]: Integration of the location information directly into the image. This is done via 8 binary boxes that encode the location.

5 Conclusions and Perspectives

This paper presents an overview and results of the first edition of the SnakeCLEF 2020: Automated Snake Identification Challenge organized in conjunction with the Conference and Labs of the Evaluation Forum (CLEF⁸) and LifeCLEF⁹ research platform. For this competition, we have used the largest and most diverse snake image dataset to date, covering 783 snake species with 245,185 training images, 14,029 validation images and 28,418 testing images. This dataset represents the most challenging dataset for automated snake species recognition in existence to date.

The final results showed that even current state-of-the-art machine learning approaches with advanced regularisation techniques are not capable of recognizing many similar-looking species. Considering that the best system submitted into the competition achieved a maximum F1 score of 0.62536, and that the 783 snake species in our training dataset represent only about one fifth of the totality of currently-described snake biodiversity [32], there remains ample room for further improvement.

In future editions, we would like to specifically target medically important scenarios [4], i.e., venomous species that are easily confused with non-venomous species. Additionally, more in-depth performance evaluation should be done to understand better which body parts of a snake (e.g., head, body, tail, eye) or visual features contribute the most to the system’s decision [26]. Moreover, comparing AI-driven systems with human experts will reveal how far automated systems are from human expertise [8]; note that humans should remain in the loop for health applications [13]. Finally, the current dataset will be further extended, with priority for new species as well as additional images for those species represented by the fewest images, with help from citizen scientists and experts [27,35].

Acknowledgements

We thank S.P. Mohanty for hosting and helping run the Snake ID Challenge on AICrowd and all the participants in all rounds of the AICrowd Snake ID Challenge. We thank the users and admins of open citizen science initiatives (iNaturalist, HerpMapper, and Flickr) for their efforts building, curating, and sharing their image libraries. C. Montalcini supported the gathering of images and metadata. S. Khandelwal provided assistance testing and summarizing datasets. P. Uetz provided support for snake taxonomy through the Reptile Database. D. Becker, C. Smith, and M. Pingleton provided support for and access to HerpMapper data. A. Abegg, J. Akuboy, G. Alcoba, A. Baskette, M. Di Bernardo, M. Bodio, O. Entiauspe, P. Freed, M. Frietas, X. Glaudas, R. Gray, T. Huang, I. Bolon, Y. Kalki, A. Laudisoit, K. P. Limbu, J. Louies, J. Martinez, K. Mebert, C. Merminod, D. Pandey, D. Raju, S. Ruane, M. Ruedi, A. Schmitz, S. Tatum, A. Visvanathan, and W. Wüster donated snake photos and helped A.M. Durso

⁸ <http://www.clef-initiative.eu/>

⁹ <http://www.lifeclef.org/>

verify identifications. This research was supported by the Fondation privée des Hôpitaux Universitaires de Genève (award QS04-20). We thank A. Flahault and the Fondation Louis-Jeantet, and F. Chappuis for supporting R. Ruiz de Castañeda and this research at the Institute of Global Health and at the Department of Community Health and Medicine of the University of Geneva. L. Pícek was supported by the Ministry of Education, Youth and Sports of the Czech Republic project No. LO1506, and by the grant of the UWB project No. SGS-2019-027. A. M. Durso was supported by the Fondation privée des Hôpitaux Universitaires de Genève (award QS04-20).

References

1. Amir, A., Zahri, N.A.H., Yaakob, N., Ahmad, R.B.: Image classification for snake species using machine learning techniques. In: International Conference on Computational Intelligence in Information System. pp. 52–59. Springer (2016)
2. Barry, J.: Identifying biodiversity using citizen science and computer vision: Introducing visipedia. In: TDWG 2016 ANNUAL CONFERENCE (2016)
3. Bloch, L., Boketta, A., Keibel, C., Mense, E., Michailutschenko, A., Pelka, O., Rückert, J., Willemeit, L., Friedrich, C.M.: Combination of image and location information for snake species identification using object detection and efficientnets. In: CLEF working notes 2020, CLEF: Conference and Labs of the Evaluation Forum, Sep. 2020, Thessaloniki, Greece. (2020)
4. Bolon, I., Durso, A.M., Botero Mesa, S., Ray, N., Alcoba, G., Chappuis, F., Ruiz de Castañeda, R.: Identifying the snake: First scoping review on practices of communities and healthcare providers confronted with snakebite across the world. *PLoS one* 15(3), e0229989 (2020)
5. Cabrera-Pérez, M.Á., Gallo-Barneto, R., Esteve, I., Patiño-Martínez, C., López-Jurado, L.F., et al.: The management and control of the california kingsnake in gran canaria (canary islands): project life+ lampropeltis. *Aliens: The Invasive Species Bulletin* 32, 20–28
6. Ruiz de Castañeda, R., Durso, A.M., Ray, N., Fernández, J.L., Williams, D.J., Alcoba, G., Chappuis, F., Salathé, M., Bolon, I.: Snakebite and snake identification: empowering neglected communities and health-care providers with ai. *The Lancet Digital Health* 1(5), e202–e203 (2019)
7. Cox, R.D., Parker, C.S., Cox, E.C., Marlin, M.B., Galli, R.L.: Misidentification of copperhead and cottonmouth snakes following snakebites. *Clinical toxicology* 56(12), 1195–1199 (2018)
8. Durso, A.M., Bolon, I., Kleinhesselink, A.R., Mondardini, M.R., Fernandez-Marquez, J.L., Gutsche-Jones, F., Gwilliams, C., Tanner, M., Smith, C.E., Wüster, W., Grey, F., Ruiz de Castañeda, R.: Crowdsourcing snake identification with online communities of professionals and avocational enthusiasts (in review)
9. Durso, A.M., Moorthy, G.K., Mohanty, S.P., Bolon, I., Salathé, M., Ruiz De Castañeda, R.: Supervised learning computer vision benchmark for snake species identification from photographs: Implications for herpetology and global health (in review)
10. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn. In: The IEEE International Conference on Computer Vision (ICCV) (Oct 2017)

11. Henke, S.E., Kahl, S.S., Wester, D.B., Perry, G., Britton, D.: Efficacy of an online native snake identification search engine for public use. *Human–Wildlife Interactions* 13(2), 14 (2019)
12. Hierink, F., Bolon, I., Durso, A.M., de Castañeda, R.R., Zambrana-Torrelío, C., Eskew, E.A., Ray, N.: Forty-four years of global trade in cites-listed snakes: Trends and implications for conservation and public health. *Biological Conservation* 248, 108601 (2020)
13. Holzinger, A., Valdez, A.C., Zieffe, M.: Towards interactive recommender systems with the doctor-in-the-loop. *Mensch und Computer 2016–Workshopband* (2016)
14. Joly, A., Goëau, H., Kahl, S., Deneu, B., Servajean, M., Cole, E., Picek, L., Ruiz De Castañeda, R., Bolon, I., Lorieul, T., Botella, C., Glotin, H., Champ, J., Vellinga, W.P., Stöter, F.R., Dorso, A., Bonnet, P., Eggel, I., Müller, H.: Overview of lifeclef 2020: a system-oriented evaluation of automated species identification and species distribution prediction. In: *Proceedings of CLEF 2020, CLEF: Conference and Labs of the Evaluation Forum, Sep. 2020, Thessaloniki, Greece.* (2020)
15. Joshi, P., Sarpale, D., Sapkal, R., Rajput, A.: A survey on snake species identification using image processing technique. *International Journal of Computer Applications* 975, 8887 (2018)
16. Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., Houlby, N.: Big transfer (bit): General visual representation learning. *arXiv preprint arXiv:1912.11370* (2019)
17. Kraus, F.: *Alien reptiles and amphibians: a scientific compendium and analysis, vol. 4.* Springer Science & Business Media (2008)
18. Moorthy, G.K.: Impact of pretrained networks for snake species classification. In: *CLEF working notes 2020, CLEF: Conference and Labs of the Evaluation Forum, Sep. 2020, Thessaloniki, Greece.* (2020)
19. Natusch, D.J., Carter, J.F., Aust, P.W., Van Tri, N., Tinggi, U., Riyanto, A., Lyons, J.A., et al.: Serpent’s source: Determining the source and geographic origin of traded python skins using isotopic and elemental markers. *Biological Conservation* 209, 406–414 (2017)
20. Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J.: 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 (2018), <https://www.pnas.org/content/115/25/E5716>
21. Organization, W.H.: *Who guidelines for the production, control and regulation of snake antivenom immunoglobulins* (2016)
22. Pelka, O., Nensa, F., Friedrich, C.M.: Variations on branding with text occurrence for optimized body parts classification. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* pp. 890–894. IEEE (2019)
23. Picek, L., Sulc, M., Matas, J.: Recognition of the amazonian flora by inception networks with test-time class prior estimation. In: *CLEF working notes 2019, CLEF: Conference and Labs of the Evaluation Forum, Sep. 2019, Lugano, Switzerland.* (2019)
24. Roll, U., Feldman, A., Novosolov, M., Allison, A., Bauer, A.M., Bernard, R., Böhm, M., Castro-Herrera, F., Chirio, L., Collen, B., et al.: The global distribution of tetrapods reveals a need for targeted reptile conservation. *Nature Ecology & Evolution* 1(11), 1677–1682 (2017)

25. Rusli, N.L.I., Amir, A., Zahri, N.A.H., Ahmad, R.B.: Snake species identification by using natural language processing. *Indonesian Journal of Electrical Engineering and Computer Science* 13(3), 999–1006 (2019)
26. Rzanny, M., Mäder, P., Deggelmann, A., Chen, M., Wäldchen, J.: Flowers, leaves or both? how to obtain suitable images for automated plant identification. *Plant methods* 15(1), 77 (2019)
27. Rzanny, M., Seeland, M., Wäldchen, J., Mäder, P.: Acquiring and preprocessing leaf images for automated plant identification: understanding the tradeoff between effort and information gain. *Plant methods* 13(1), 1–11 (2017)
28. Schaper, A., Desel, H., Ebbecke, M., Haro, L.D., Deters, M., Hentschel, H., Hermanns-Clausen, M., Langer, C.: Bites and stings by exotic pets in europe: An 11 year analysis of 404 cases from northeastern germany and southeastern france. *Clinical Toxicology* 47(1), 39–43 (2009)
29. Seeland, M., Rzanny, M., Alaqraa, N., Wäldchen, J., Mäder, P.: Plant species classification using flower images—a comparative study of local feature representations. *PLoS one* 12(2), e0170629 (2017)
30. Sulc, M., Pícek, L., Matas, J.: Plant recognition by inception networks with test-time class prior estimation. In: CLEF working notes 2018, CLEF: Conference and Labs of the Evaluation Forum, Sep. 2018, Avignon, France. (2018)
31. Tan, M., Le, Q.V.: Efficientnet: Rethinking model scaling for convolutional neural networks. *CoRR* abs/1905.11946 (2019), <http://arxiv.org/abs/1905.11946>
32. Uetz, P., Hošek, J., Hallermann, J.: The reptile database (2020), <http://reptile-database.reptarium.cz>
33. Valan, M., Pícek, L.: Mastering large scale multi-label image recognition with high efficiency over camera trap images. *arXiv preprint arXiv:2008.07828* (2020)
34. Van Horn, G., Aodha, O.M., Song, Y., Cui, Y., Sun, C., Shepard, A., Adam, H., Perona, P., Belongie, S.: The inaturalist species classification and detection dataset. In: *Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, UT (2018), <https://vision.cornell.edu/se3/wp-content/uploads/2018/03/1916.pdf>, <https://vision.cornell.edu/se3/wp-content/uploads/2018/03/1916-suppl.pdf>
35. Van Horn, G., Branson, S., Farrell, R., Haber, S., Barry, J., Ipeirotsis, P., Perona, P., Belongie, S.: Building a bird recognition app and large scale dataset with citizen scientists: The fine print in fine-grained dataset collection. In: *Computer Vision and Pattern Recognition (CVPR)*. Boston, MA (2015), http://vision.cornell.edu/se3/wp-content/uploads/2015/05/Horn_Building_a_Bird_2015_CVPR_paper.pdf
36. Wäldchen, J., Rzanny, M., Seeland, M., Mäder, P.: Automated plant species identification—trends and future directions. *PLoS computational biology* 14(4), e1005993 (2018)
37. Warrick, B.J., Boyer, L.V., Seifert, S.A.: Non-native (exotic) snake envenomations in the us, 2005–2011. *Toxins* 6(10), 2899–2911 (2014)
38. (WHO), W.H.O.: Snakebite envenoming: Global situation. (2019 (accessed August 28, 2020)), <https://www.who.int/news-room/fact-sheets/detail/snakebite-envenoming>
39. Willi, M., Pitman, R.T., Cardoso, A.W., Locke, C., Swanson, A., Boyer, A., Veldhuis, M., Fortson, L.: Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution* 10(1), 80–91 (2019), <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13099>
40. Williams, D.J., Faiz, M.A., Abela-Ridder, B., Ainsworth, S., Bulfone, T.C., Nickerson, A.D., Habib, A.G., Junghanss, T., Fan, H.W., Turner, M., et al.: Strategy for

- a globally coordinated response to a priority neglected tropical disease: Snakebite envenoming. *PLoS neglected tropical diseases* 13(2), e0007059 (2019)
41. Wittich, H.C., Seeland, M., Wäldchen, J., Rzanny, M., Mäder, P.: Recommending plant taxa for supporting on-site species identification. *BMC bioinformatics* 19(1), 190 (2018)
 42. Wolfe, A.K., Fleming, P.A., Bateman, P.W.: What snake is that? common australian snake species are frequently misidentified or unidentified. *Human Dimensions of Wildlife* pp. 1–14 (2020)
 43. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412* (2017)