VizPol: Real-Time Symbol Recognition for Field Reporting

Ishaan Jhaveri (iaj2107*), Guangxing Han (gh2561*),
Nina Berman (njb22*), Shih-Fu Chang (sc250*), Susan McGregor (sem2196*); *@columbia.edu

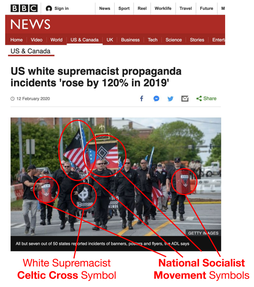
Introducing the Problem

Flyer from Portland Identifying Political Symbols

Since 2015, journalists covering political movements have had increasing difficulty identifying many of the graphical symbols being used by political groups, such as the many "alt-right" groups that use logos and symbols that originate and evolve in niche online or extremist communities. These are often unfamiliar to mainstream political journalists.

1488 Tattoo on Unite the Right 2 rallier in DC © Nina Berman

Nina Berman, a photojournalist who covers American political movements had the idea for this project in August 2018 when she was photographing the "Unite the Right 2" rally in Washington, DC. She saw the woman pictured on the right telling a reporter that she was protesting "threats to her first amendment rights" and photographed her. Only later, she realized that the tattoo on the woman's forearm spelled out "1488"—a white supremacist meme. 14 represents the white nationalist slogan, "We must secure the existence of our people and a future for white children," and 88 represents "Heil Hitler," "H" being the 8th letter of the alphabet. If Berman had known this in the moment she would have had more questions about the woman's true intentions at the rally and taken more pictures. She decided to build a tool to help journalists identify symbols with obscure meanings directly in the field.

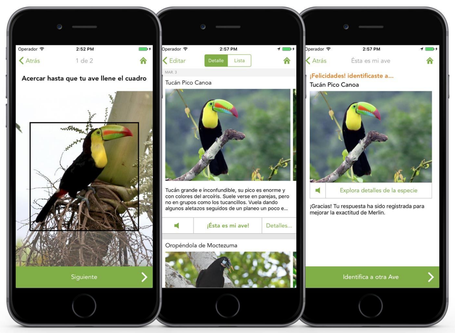


White Supremacist **Celtic Cross** Symbol

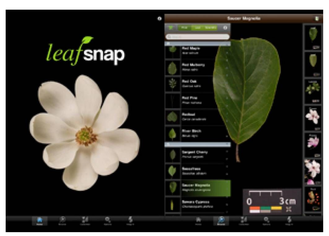
National Socialist Movement Symbols

News reports often conflate or confuse political groups because of a lack of knowledge of their symbols. Here's an article about the political groups **Patriot Front**, **American Identity Movement** and **New Jersey European Heritage Association,** with a picture of a different group, **National Socialist Movement**. The picture is not even labeled. Though the caption of the photo does mention **National Socialist Movement**, of the 20 labels this image has on Getty's database like Racism, Neo-nazism, Political Rally, etc., not 1 mentions **National Socialist Movement** or **Celtic Cross**.

Inspiration

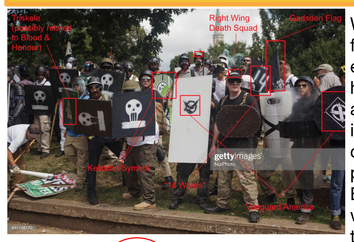


Cornell Lab of Ornithology's Merlin Bird ID App that allows users to take pictures of birds and identifies the species from the picture



Leafsnap, developed by researchers from Columbia University, the University of Maryland, and the Smithsonian Institution which identifies tree species from photographs of their leaves.

Methodology: Knowledge Base Construction



We selected an initial symbol set for the system by reviewing existing news images of recent, high-profile political events, such as the 2017 Unite the Right rally in Charlottesville. We also conducted pilot interviews with photojournalists like Nina Berman herself and reporters who routinely cover these events to complement our initial list

Odal Rune © Ishaan Jhaveri

III Percenters Flag © Nina Berman

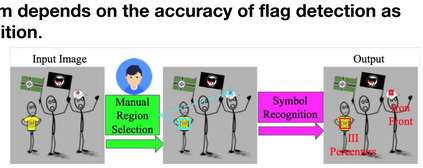
American Identity Movement © Nina Berman

We built the recognition tool using the initial set of expert-curated symbols. But the app has to stay current with rapidly evolving political groups and corresponding symbols to really be of use to photographers and reporters so we have designed a method to continuously augment the database through user input. This is discussed further in the TK section of this poster.

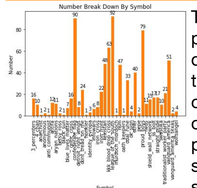
Methodology: Computer Vision Overview

Many symbols occur on flags. The first Machine Learning "pipeline" we built just takes an image as input. The system first automatically detects regions of the image that contain flags, then identifies the symbols on these flags. **The accuracy of this system depends on the accuracy of flag detection as well as symbol recognition**.

Equally, many symbols occur on other surfaces like hats, shirts, tattoos, etc. Instead of designing a custom pipeline for each of an arbitrary number of surfaces we built a generalized pipeline that relies on user input to select a region of the input image that contains a symbol the user is interested in identifying. This system takes an image and a region of an image as input and identifies the symbol in that region. **The accuracy of this system depends only on the accuracy of symbol recognition**.



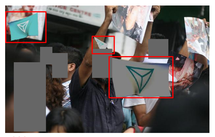
Methodology: Synthetic Training Data



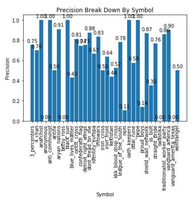
Statistics about photographs from expert photographers

Though we had some photos from individual photographers who are familiar with the domain as discussed above, we didn't have nearly enough to train and test an entire model. Moreover, we wanted our model to recognize some symbols that didn't occur at all in the set of photos we had from photographers. We experimented with Google Image searches for the name of each symbol, but many symbols had no results, and the images corresponding to symbols with results were too low-quality (low-resolution or computer-modified) to train our model effectively. The approach also did not allow us to control for potential intellectual property or privacy concerns. Rights challenges also prevented us from using imagery from editorial databases such as a Getty Images and Redux. So we decided to use synthetic data to train both the flag detector and the symbol recognizer. The photographs we had from photographers were reserved as test images. We used GoogLeNet as our base network.

We used Blender to generate the synthetic copies of each symbol

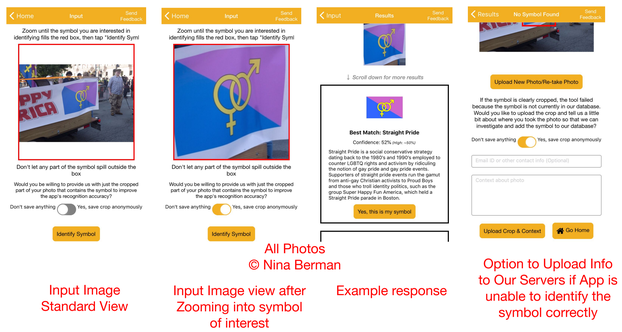


We pasted the synthetic copies onto images from the UCLA Protest Images dataset



Using Blender generated synthetic images far outperformed (81% accuracy to 63% accuracy) a model trained on Google Images exclusively. The success of this experiment lead to an experiment with synthetic images where we applied random image transformations (brightness, contrast, transparency, etc.) to images of the symbols and pasted them onto varied backgrounds. Here are the results of this experiment.

Mobile App



Input Image Standard View

Input Image view after Zooming into symbol of interest

Example response

Option to Upload Info to Our Servers if App is unable to identify the symbol correctly

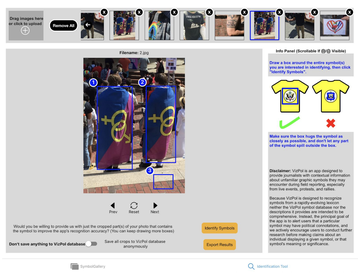
All Photos © Nina Berman

The mobile app is targeted for political beat photojournalists and reporters wanting to gain knowledge about a symbol they have just observed in-the-field.

The app allows a user to select an existing photo from their phone, and zoom in on the relevant symbol until it fills the in-app viewer. The user then clicks a button to send their selection to a server for identification. If the server finds a match or matches, it responds with a list of at most 3 symbols, organized in decreasing order of confidence. The user can then indicate the symbol that matches their image, or indicates that none are correct. In either case, the user can opt to upload their photo to improve or expand the image recognition model. The threshold for returning a match is set to 1/N where N is the number of symbols the model is currently able to recognize (for the moment, N = 56), to ensure that the model returns only reasonable-quality matches. In either case the user can choose to upload metadata about the photo's location, date and event. This metadata may be useful for providing additional symbol context in later versions of the app.

The iOS App is developed and being used by about 50 external users at the moment. The Android App is currently under development.

Web App



Identify Multiple Bounding Boxes Request

All Photos © Nina Berman

Identify Multiple Bounding Boxes Response

The web app is designed for photo editors or journalists working at their desk or doing research in online chat rooms where symbols are often used.

Given the larger screen space, the web app allows users to draw one or more bounding boxes on a single image to obtain match results for all of them. Any of these recognition processes can be run on a picture uploaded by the user, and by clicking on each bounding box (either a flag detected or a bounding box drawn by the user) estimated most likely symbols and their short explanation are given in a sidebar.

Discussion & Future Work

Feature Updates

We have several feature updates planned to extend the current app's functionality.

- We need to implement **user accounts** and **query history** pages.
- We need to make the symbol gallery on the Web App searchable

Project Future

- We are exploring a potential collaboration with the Associated Press which will give us many more images with symbols in them to find new symbols for our database
- We are talking to independent researcher Adam Harvey about potentially including VizPol in his project VFrame, a suite of tools to help journalists and researchers covering protests and warzones.

International Expansion

- We are exploring expanding the database to include overseas political symbols for overseas use.

Adding Symbols

- We are considering ways to expand our base of committed users that will help us build the database.
- We are considering potential partnerships with students to help us look through images and build the database.

References

[1] [n.d.]. Deconstructing the symbols and slogans spotted in Charlottesville. https://www.washingtonpost.com/graphics/2017/local/charlottesville-videos/
[2] [n.d.]. Demo Expo. https://summit.nycmedialab.org/demo-expo
[3] [n.d.]. Four Corners Project. https://fourcornersproject.org/en/
[4] 2020. US white supremacist propaganda incidents 'rose by 120% in 2019'. https://www.bbc.com/news/world-us-canada-51480500
[5] Simone Bianco, Marco Buzzelli, Davide Mazzini, and Raimondo Schettini. 2015. Logo recognition using cnn features. In International Conference on Image Analysis and Processing. Springer, 438–448.
[6] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi. 2014. Describing Textures in the Wild. In Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR).
[7] Christian Eggert, Dan Zecha, Stephan Brehm, and Rainer Lienhart. 2017. Improving small object proposals for company logo detection. In Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval. ACM, 167–174.
[8] Ishaan Jhaveri [n.d.]. When is a frog not a frog? Building a new digital tool to track political symbols. https://www.cjr.org/tow_center/political-symbols-to-w-center.php
[9] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).
[10] Sasha Koren. 2019. Introducing the News Provenance Project. https://open.nytimes.com/introducing-the-news-provenance-project-723dbaf07c44
[11] Neeraj Kumar, Peter N. Belhumeur, Arijit Biswas, David W. Jacobs, W. John Kress, Ida C. Lopez, and João V. B. Soares. 2012. Leafsnap: A Computer Vision System for Automatic Plant Species Identification.. In ECCV (Lecture Notes in Computer Science), Andrew W. Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid (Eds.), Vol. 7573. Springer, 502–516. http://dblp.uni-trier.de/db/conf/eccv/eccv2012-2.html#KumarBBJKLS12
[12] Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Malloci, Tom Duerig, and Vittorio Ferrari. 2018. The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale. arXiv:1811.00982 (2018).
[13] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems. 91–99.
[14] Hawes Spencer. 2017. A Far-Right Gathering Bursts Into Brawls. https://www.nytimes.com/2017/08/13/us/charlottesville-protests-unite-the-right.html
[15] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition. 1–9.
[16] Grant Van Horn, Steve Branson, Ryan Farrell,Scott Haber, Jessie Barry, Panos Ipeirotis, Pietro Perona, and Serge Belongie. 2015. 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. Boston, MA, 595–604.
[17] Donghyeon Won, Zachary C Steinert-Threlkeld, and Jungseock Joo. 2017. Protest activity detection and perceived violence estimation from social media images. In Proceedings of the 25th ACM international conference on Multimedia. ACM, 786–794.
[18] Xu Zhang, Felix Xinnan Yu, Svebor Karaman, Wei Zhang, and Shih- Fu Chang. 2018. Heated-up softmax embedding. arXiv preprint arXiv:1809.04157 (2018).