Faces extracted from Time Magazine 1923-2014

Ana Jofre, Vincent Berardi, Carl Bennett, Michael Reale, Josh Cole

ARTICLE INFO

ABSTRACT

We present metadata of labeled faces extracted from a *Time* magazine archive that contains 3,389 issues ranging from 1923 to 2012. The data we are publishing consists of three subsets: *Dataset 1*) the gender labels and image characteristics for each of the 327,322 faces that were automatically-extracted from the entire *Time* archive, *Dataset 2*) a subset of 8,789 faces from a sample of 100 issues that were labeled by Amazon Mechanical Turk (AMT) workers according to ten dimensions (including gender) and used as training data to produce *Dataset 1*, and *Dataset 3*) the raw data collected from the AMT workers before being processed to produce *Dataset 2*.

Introduction

Time magazine is an American weekly news magazine that has been published continuously since 1923, when it was founded by Henry R. Luce and Briton Hadden with the goal of concisely summarizing current events in a manner that could be digested in around one hour.[1] Its news coverage was (and remains) characterized by a person-centric perspective, which often takes the form of photographs of featured individuals. After its inception, *Time* quickly became the most popular news magazine in the United States and, due to its presence as a staple of mainstream news consumption, it has both reflected and influenced American popular attitudes, perhaps more than any comparable publication.

Time's continuous publication over the last nearly one hundred years, along with its position as a powerful popular cultural influence, makes it a unique vehicle for exploring trends in American society, including the centrality of images to twentieth and twenty-first culture, economics, and politics.[2] In our study, we focused upon how *Time* reflected this visual sensibility through images of individual faces. Tracking and analyzing the faces that appear in the magazine, through both digital-analytic and traditional humanistic means, offers a sense of what type of people – as represented through photographic depictions of their faces -- were considered culturally significant throughout the different eras of the magazine's publication.

A number of scientific and humanistic resources support the cultural importance of the human face – the object of so much of *Time*'s visual archive. Psychologist Cathy Mondloch has demonstrated that a child recognizes their mother's face -- so crucial to survival – as soon as seven seconds after birth.[3] In sociologist Anthony Giddens' notion of "facework commitments," modern people unconsciously exhibit facial "cues and signals" to others in order to ensure passersbys in urban settings that they are trustworthy and not a threat.[4] This psychological and sociological emphasis on the human face extends to artistic and photographic imagery of faces. the art historian Ernst Gombrich challenges his readers to tear up a photograph of the faces of loved ones to demonstrate the subjective power that they hold over us,[5] while the media historian Tessa Morris-Suzuki insists that our imagination – and thus our empathy – is charged by faces and facial expressions in a manner that transcends space and time.[6] The human face, in short, has unmatched cultural power.

In this work, we present a summary of human faces extracted from a *Time* magazine archive that contains 3,389 issues ranging from 1923 to 2014. The data is presented in three parts. *Dataset 1* consists of 327,322 faces that were computationally extracted from the entire archive and classified by gender. In *Dataset 2*, 8,789 faces from 100 selected issues in the archive were extracted and classified by human labor according to ten dimensions including gender, race, age, and facial expression. To increase classification accuracy, multiple individuals evaluated each image and the results were aggregated into one definitive classification. In the interest of full transparency, we include *Dataset 3*, which consists of the individual classifications prior to aggregation into *Dataset 2*.

Our dataset of classified faces is derived from a specific cultural source, as opposed to face datasets targeted toward face-recognition applications, such as *Cohn-Kanade+*[7] or *Labeled Faces in the Wild,*[8] which provide users with a variety of faces with different expressions under different lighting conditions. It bears a greater similarity to projects such as *Selfie-City,*[9] which consists of extracted 'selfies' from Instagram, in that it allows researchers to understand visual patterns from a particular media source. Like *Selfie-City*, this data could lend itself to an interactive web-based data visualization, but also allows researchers to gain theoretical insights into a specific media source.

Our datasets are particularly amenable to applying distant viewing[10] techniques to the archive. For example, researchers may be able to use this data to study how visual elements in faces reveal the historical context of *Time* magazine, analogous to Arnold *et al* 's use of visual elements from two network era sitcoms to reveal two different perspectives of feminism in the 1960s.[11] In our own work with these datasets, we investigated society's changing attitudes concerning women through the lens of *Time* magazine.[12] In publishing this dataset, our aim is to allow other researchers to fully explore this data in order to generate results that complement our own findings.

Description of the data

The archive from which this published data was extracted was retrieved from *Time* magazine's *Time Vault.*[13] It consists of approximately 500,000 .jpg files, with one file for each page of 3,389 issues with dates ranging from 1923 to 2014. Most images have a size of 1095 x 1550 pixels, but those from issues published after January 19, 2004 have a size of 1163×1550 pixels.

The archive we used is not an exhaustive collection of all *Time* magazine issues. To determine the completeness of this data, we compared our archive to a master list of all *Time* issues published between 1923 and 2009. As summarized in Figure 1, the archive did not contain issues from 1981 and 1989. There appear to be two distinct trends in the data: 1.) 1923 through 1968, with an average of 15.3 missing issues each year and 2.) 1969-2004, with 11.4 issues missing each year (not including 1981 and 1989). Beyond 2004, the coverage of our archive markedly decreases with nearly the same number of issues in and missing from the archive in 2009. The dates of the missing issues were equally distributed among months, but were strongly biased towards being from the first week of whichever month it was missing from. Despite the incomplete coverage of the entire *Time* catalog, the nearly 3,400 issues contained in our archive should be more than sufficient to allow for an exploration of *Time*'s visual culture.
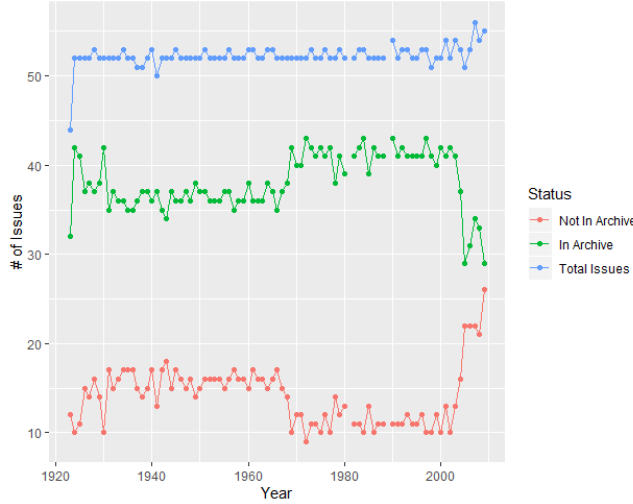
Figure 1: Summary of the number of issues in our archive. The green line indicates the number of issues from a given year contained in the archive while the red line indicates the number of issues not in the archive. The blue line is the sum of the green and red lines and represents the total number of issues for each year. Our archive contained zero issues from 1981 and 1989, so these years were eliminated from the figure.

Our data consists of three parts, which we refer to here as *Dataset 1*, *Dataset 2*, and *Dataset 3*. Each of these data sets is presented in plain .csv tables and can be used for computational analysis. The values in each column are integers, Boolean, or well-defined discrete classifications. A description of each data set plus a data dictionary is provided below.

Dataset 1

To form this dataset, we extracted 327,322 faces from all 3,389 issues in our collection and automatically classified each face as male or female. We present this data as a single table with columns identifying the date, issue, page number, the coordinates identifying the position of the face on the page, and gender classification. Table 1 lists the values presented in this dataset. Note that the coordinates identifying the position of the face on the page are based on the size and resolution of the pages as found in the *Time Vault*, which are noted above.

Column Name	Possible value
Year	Integer: publication year (1923 – 2014)
Month	Integer: publication month (1 – 12)
Day	Integer: publication day (1 – 31)
Page	Integer: page number (1 – 250)
x1	Integer: x coordinate of top left corner of face crop
y1	Integer: y coordinate of top left corner of face crop
x2	Integer: x coordinate of bottom right corner of face crop
y2	Integer: y coordinate of bottom right corner of face crop
Crop accuracy	Proportion: Estimated accuracy of face crop
Gender	Categorical: Female, Male, or Unknown
Gender accuracy	Proportion: Estimated accuracy of gender classification

Table 1: Description of columns in Dataset 1 table.

This data set was created by using deep learning to detect and extract face bounding boxes as well as classify the gender of the faces. To build a training set for the machine learning model, we employed crowdsourcing through a custom-built web application deployed via Amazon MTurk (AMT). AMT coders examined 3,772 pages and extracted 4,685 face bounding boxes, which formed the training set for the face extractor. A different set of 342 AMT coders annotated the extracted face images by gender. Faces marked as "Unknown" gender were removed from the training set for the gender classifier, resulting in 4,405 faces with gender ground truth.

After the AMT training set was created, RetinaNet detector[14] was used to extract faces from the remainder of the archive. RetinaNet was originally designed to detect objects in a scene. It outputs candidate bounding boxes classified as belonging to a certain category, in our case "face" versus "non-face," with each classification accompanied a confidence metric. Candidate faces were filtered using a threshold of 90% confidence to produce the 327,322 faces in *Dataset 1*.

To classify gender, we fine-tuned the VGGFace[15] architecture, whose original purpose was identity recognition (that is, given a face image, who is the person

in question). We altered the output of the network to classify each image as male or female and fine-tuned the network to increase accuracy.

An initial model was trained using the 4,405 AMT gender-marked face images. We used a 70%/30% train/test split and achieved an accuracy of 87% on the test set as summarized by the confusion matrix shown in Table 2.

Pred →	Male	Female
Male	909	74
Female	99	241

Table 2: Gender classification results, round 1 (testing set, AMT data only).

Our training data exhibited a class imbalance with approximately 75% of all faces classified as Male. Therefore, we employed a bootstrapping technique to acquire additional, more balanced, training data and thus improve our classifier. The model trained on the AMT data was used to classify all 327,322 faces from the archive. From these faces, we randomly selected images and manually verified the classification results. These new images plus the AMT data yielded a new dataset of 17,698 faces for the second round of training. The dataset was again split 70%/30% for train/test and the accuracy on the test portion was 95%, as summarized by the confusion matrix shown in Table 3. We note that these results also verify the inherent quality of the face images extracted: images with lower prediction accuracy tend to be lower resolution images.

Pred →	Male	Female
Male	2677	113
Female	154	2366

Table 3: Gender classification results, round 2 (testing set, AMT + verified classifier-predicted data).

This more accurate model was then used to classify the entirety of the archive, resulting in the predicted gender of the 327,322 faces of *Dataset 1*.

Dataset 2

Dataset 2 is smaller than *Dataset 1*, consisting of 100 issues compared to 3,389 issues, but is much richer in scope.

For this data set, we had human coders classify faces extracted from 100 issues in the *Time Vault* archive. Coders were presented with magazine pages with an overlaid rectangle encompassing one face and asked the following questions:

Does the face belongs to an adult or a child? What is the angle of the face (in profile or looking forward)? What is the gender of the face? Is the image in color or in monochrome; What is the image quality? Is it a photo or an illustration? In what context does the face appear (advertisement, feature story, or cover)? Is there more than one face in the image or is it a portrait containing a single face? What is the race of the face (select from U.S. census categories: American Indian, Asian, Black, Pacific Islander, White, or Unknown)? Is the face smiling? We present this data in tabular form with columns identifying the date, issue, page number, coordinates identifying the position of the face on the page, and each of the classification categories. The complete list of variables is summarized in Table 4 below.

Dataset 2 consists of 8,789 classified faces from 100 selected issues. Human labor was used to identify and extract 3,299 face images from 39 issues, which were later classified by another set of workers. This selection of 39 issues contains one issue per decade spanned by the archive plus one issue per year between 1961 and 1991, and the extracted face images were used to train the face extraction algorithm. The remaining 5,490 faces from 61 issues were extracted via machine learning before being classified by human coders. These 61 issues were chosen to complement the first selection of 39 issues: one issue per year for all years in the archive excluding those between 1961 and 1991. Thus, *Dataset 2* contains fully-labelled faces from at least one issue per year.

Column Name	Possible value
Year	Integer: publication year (1923 – 2014)
Month	Integer: publication month (1 – 12)
Day	Integer: publication day (1 – 31)
Page	Integer: page number (1 – 250)
x1	Integer: x coordinate of top left corner of face crop
x2	Integer: x coordinate of bottom right corner of face crop
y1	Integer: y coordinate of top left corner of face crop
y2	Integer: y coordinate of bottom right corner of face crop
Gender	Categorical: Female, Male, or Unknown
Adult	Boolean: Yes (1) or No (0)
Face Angle	Categorical: Profile or Forward
Image Color	Boolean: Color (1) or Monochrome (0)
Image Quality	Categorical: Good, Fair, Poor, or Error (not a face)
Image Type	Boolean: Photo (1) or Illustration (0)
Context	Categorical: Advertisement, Cover Page, or Feature Story
Multi-face	Boolean: Yes (1) or No (0)
Race	Categorical: American Indian, Asian, Black, Pacific Islander, White, or Unknown (where 'white' includes Middle Eastern and Latinx)
Smile	Boolean: Yes (1) or No (0)
Face-Extraction	Categorical: Human or Machine
Tag group	<Integer>.<Integer>.<Integer>: Unique worker ID

Table 4: Description of columns in Dataset 2 tables.

Dataset 3

In the interest of transparency, *Dataset 3* consists of the raw data collected to create *Dataset 2*, and consists of 2 tables. Before explaining these tables we first briefly describe our data collection and verification procedures, which have been fully described elsewhere.[16]

A custom AMT interface was used to enable human laborers to classify faces according the categories in Table 4. Each worker was given a randomly-selected batch of 25 pages, each with a clearly highlighted face to be categorized, of which three pages were verification pages with known features, which were used for quality control. Each face was labeled by two distinct human coders, determined at random so that the paring of coders varied with the image. A proficiency rating

was calculated for each coder by considering all images they annotated and computing the average number of labels that matched those identified by the image's other coder. *Dataset 2* was created by resolving inconsistencies between an image's two coders by selecting the labels from the coder with the highest proficiency rating. Prior to calculating the proficiency score, all faces that were tagged as having 'Poor' or 'Error' image quality by either of the two coders were eliminated. Due to technical bugs when the AMT interface was first implemented, a small number of images were only labeled once; these were also eliminated from *Datasets* 2 and 3.

In *Dataset 3*, we present the raw annotations by each coder that tagged each face, along with demographic data for each coder. *Dataset 3* consists of two tables: the raw data from each of the two sets of coders, and the demographic information for each of the coders. The columns in the raw data table are the same as in Dataset 2. The columns in the demographic table is detailed in Table 5 below.

Column Name	Possible value
Tag group	<Integer>.<Integer>.<Integer>: Unique worker ID
Gender	Categorical: Female, Male, or Other
Age	Categorical: 18-25, 26-35, 36-45, 46-55, 56-65, or 66+
Race	Categorical: American Indian, Asian, Black, Pacific Islander, White, or (None of the above, mixed, or unknown)
Location	Categorical: Asia (including Turkey and Indonesia, excluding Russia and Egypt), Europe (including Russia), Africa (including the Sinai Peninsula in Egypt), North America (including Central America and the Caribbean), South America, Antarctica, or Australia and Oceania
Proficiency score	Numerical value between 0 and 1

Table 5: Description of columns in demographics tables in Dataset 3.

Discussion

What can you do with this data?

Dataset 1 allows us to ask the following questions. How has gender representation changed over time? How has the importance of the image of the face changed over time? How does gender representation correlate with the magazine's text and with the historical context? As noted in the introduction, we were able to address these questions using this data in a concurrent publication.[17]

Dataset 2 is more detailed than *Dataset 1* and allows more detailed questions to be asked, such as: How has race representation changed over time? How has the representation of children changed over time? How does race and/or age correlate with the magazine's text and with the historical context? What types of faces are more likely to be smiling? In what context (ads or news) do certain types of faces tend to appear, and how does this change over time? What types of faces are more likely to be presented as individualized portraits?

Dataset 3 was included to provide full transparency on our methods of arriving at *Dataset 2*, but it also contains demographic data of the coders for the examination of the relationship between the coders' demographics and how their labels for each face, a subject we are currently investigating.

The meaning and use of this data

Our dataset of faces is representative of the people deemed news-worthy and/or advertisement worthy in American culture, from the specific perspective of *Time* magazine. Overwhelmingly, the faces in this data belong to white men. One major motivation for collecting this data was to quantify and track the representation of women and ethnic minorities in a popular media source, and we hope that this data helps expose the bias implicit in at least one source of American print media. We believe that the biases in this data could provide an interesting source of study that could reveal patterns in the American media landscape.

For this data to be culturally meaningful, it needs to be studied in the context of a historical framework and a close-reading analysis of the magazine. We point to our own recent work as an example of how to work with this data. We used this data to study how the representation of women in the magazine changed over time, and we found that the peaks and valleys in the proportion of faces that were women tracked closely with historical context and with a close reading of the magazine's content.[18]

Limitations

Because of the bias in race and gender, we believe this data is more suitable for cultural and media studies rather than for machine learning training.

The under-representation of women created challenges when training the algorithm to automatically classify the gender of all the faces. We needed to add a significant number of female faces to our training set in order to improve the accuracy on correctly classifying female faces.

Despite our best efforts, we note that in *Dataset 1*, the machine classification of gender isn't perfect: the error rate was about 6.5% for women and about 4.2% for men.[19] Furthermore, the algorithm cannot differentiate between someone who identifies as a woman (or man) and someone who presents as female (or male). *Dataset 2* mitigates this problem by using human coders who see the context of the page, but that set of data is much smaller.

Having encountered issues with training an algorithm to classify gender, we expect that the lack of representation of races other than white will make it difficult to use *Dataset 2* to train an algorithm to classify faces by race. There are other factors that make race classification difficult. For one, race categories are somewhat arbitrary: we found that census categories have changed significantly over the past century, and some researchers may not agree with these categories. Secondly, the concept of race is highly context-dependent. We found that the race of a face is often not recognized unless it is embedded within a stereotyped setting. Finally, we found that when the face was not white, human coders tended to disagree on race more than with other categories.

Summary and Future Work

In summary, we categorized faces extracted from a *Time* magazine archive that contains 3,389 issues ranging from 1923 to 2012. The data we are publishing consists of 1.) the gender labels and image characteristics for each of the 327,322 faces that were automatically-extracted from the entire *Time* archive, 2.) a subset of 8,789 faces from a sample of 100 issues that were labeled by AMT workers according to ten dimensions, and 3.) the raw data collected from the AMT workers before being processed to produce *Dataset 2*.

We believe that *Time* magazine is interesting for humanities researchers because it is so culturally ubiquitous in the United States. Perhaps more than any comparable publication, it has both reflected and influenced American culture, as well as popular attitudes to domestic and global politics. Tracking the view of

the world through Time magazine can provide a historical perspective through one lens of American culture.

In future work, we plan to create a dataset of the faces that appear in the advertisements. In preliminary work, we found that as the representation of women in the overall magazine increases, the proportion of women in advertisements decreases. We intend to further investigate this and other trends in advertising.

Acknowledgements

Notes

[1] Curtis Prendergast and Geoffrey Colvin, *The World of Time Inc: The Intimate History of a Changing Enterprise 1960 – 1980* (New York: Atheneum, 1986).

[2] John Berger, *Ways of Seeing* (London: British Broadcasting Corporation and Penguin), 129; Sontag, *Photography*, 21–22.

[3] Catherine J. Mondloch, Terri L. Lewis, D. Robert Budreau, Daphne Maurer, James L. Dannemiller, Benjamin R. Stephens, and Kathleen A. Kleiner-Gathercoal, "Face Perception During Early Infancy," *Psychological Science* 10, no. 5 (1999): 419–22.

[4] Anthony Giddens, *The Consequences of Modernity*. Stanford: Stanford University Press, 1990, pp. 80-83.

[5] Ernst H. Gombrich, *The Story of Art*, 13th ed. (New York: Phaidon, 1978), 20.

[6] Tessa Morris-Suzuki, *The Past Within Us: Media, Memory, History* (New York and London: Verso): 118.

[7] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., and Matthews, I. "The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression", Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101 (2010). (http://www.pitt.edu/~emotion/ck-spread.htm)

[8] Gary B. Huang and Erik Learned-Miller, "Labeled Faces in the Wild: Updates and New Reporting Procedures." University of Massachusetts, Amherst, Technical Report UM-CS-2014-003 (2014). (http://vis-www.cs.umass.edu/lfw/)

[9] http://selfiecity.net/

[10] Taylor Arnold and Lauren Tilton, "Distant viewing: analyzing large visual corpora," *Digital Scholarship in the Humanities* (2019). For more, also visit the Distant Viewing Lab at the University of Richmond at distantviewing.org.

[11] Taylor Arnold, Lauren Tilton, and Annie Berke, "Visual Style in Two Network Era Sitcoms," *Journal of Cultural Analytics*. July 19, 2019. DOI: 10.22148/16.043.

[12] Ana Jofre, Josh Cole, Carl Bennett, Michael Reale, and Vincent Berardi, "What's in a Face? Gender representation of faces in Time, 1940s-1990s," *Journal of Cultural Analytics*. March 16, 2020.

[13] https://time.com/vault/

[14] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar, "Focal Loss for Dense Object Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2018): 1–1, https://doi.org/10.1109/TPAMI.2018.2858826.

[15] P Omkar M. Parkhi, Andrea Vedaldi, and Andrew Zisserman, "Deep Face Recognition," in *Proceedings of the British Machine Vision Conference 2015*, 41.1–41.12 (Swansea: British Machine Vision Association, 2015), https://doi.org/10.5244/C.29.41

[16] Ana Jofre, Vincent Berardi, Carl Bennett, Aisha Cornejo, John Harlan, and Kathleen P.J. Brennan. "Crowdsourcing Image Extraction and Annotation: Software Development and Case Study" submitted to *Digital Humanities Quarterly*. Under review.

[17] Jofre et al. "What's in a Face?" *Journal of Cultural Analytics.*

[18] Jofre et al. "What's in a Face?" *Journal of Cultural Analytics.*

[19] Jofre et al. "What's in a Face?" *Journal of Cultural Analytics*.