

# Detecting Faces, Visual Medium Types, and Gender in Historical Advertisements, 1950-1995

Melvin Wevers<sup>1</sup>[0000-0001-8177-4582] and Thomas Smits<sup>2</sup>[0000-0001-8579-824X]

<sup>1</sup> DHLab, KNAW Humanities Cluster, Amsterdam, the Netherlands  
[melvin.wevers@dh.huc.knaw.nl](mailto:melvin.wevers@dh.huc.knaw.nl)

<sup>2</sup> Utrecht University, Utrecht, the Netherlands [t.p.smits@uu.nl](mailto:t.p.smits@uu.nl)

**Abstract.** Libraries, museums, and other heritage institutions are digitizing large parts of their archives. Computer vision techniques enable scholars to query, analyze, and enrich the visual sources in these archives. However, it remains unclear how well algorithms trained on modern photographs perform on historical material. This study evaluates and adapts existing algorithms. We show that we can detect faces, visual media types, and gender with high accuracy in historical advertisements. It remains difficult to detect gender when faces are either of low quality or relatively small or large. Further optimization of scaling might solve the latter issue, while the former might be ameliorated using upscaling. We show how computer vision can produce meta-data information, which can enrich historical collections. This information can be used for further analysis of the historical representation of gender.

**Keywords:** face detection, heritage, historical advertisements, medium detection, gender detection

## 1 Introduction

In 1925, psychologist Alfred Poffenberger encouraged ad makers to “short-circuit the consumer’s mind through vivid, pictorial appeals to fundamental emotions” [28]. By tapping into a representational system that produces meaning outside the realm of the advertised product, images could help to sell products [14]. As a result, advertisements have been an important source for historians studying the development of consumerism and consumer society. An important focus in consumer studies is the representation of gender [37,42,32]. How is gender constructed in advertisements? In the late 1970s, sociologist Erving Goffman undertook a systematic investigation of the semiotic content of advertisements printed in contemporary newspapers and glossy magazines. He was particularly interested in what he called Gender Displays [13]. This research interest resulted in his seminal book *Gender Advertisements* 1976].

In this study, Goffman proposed five categories to study the depictions of and the relations *between* men, women, and children in advertisements: relative size, feminine touch, function ranking, the ritualization of subordination, and licensed withdrawal. By manually examining advertisements, he reflected how

men and women could be placed in these categories, from which he then drew generalizations. Building on this approach, Jonathan Schroeder, contends that in advertisements, men are regularly presented as “the active subject, the business-like, self-assured decision maker, while the female occupies the passive object, the observed sexual/sensual body, eroticized and inactive [36].”

Studies like those by Goffman and Schroeder rely on a qualitative analysis, or close-reading, of a limited number of advertisements published in a relatively short period. Goffman only examined around 400 different advertisements, which he selected from “newspapers and magazines easy to hand - at least to my hand.” As Kang [18] notes, Goffman was often criticized for this method. Instead of relying on a random sample, he purposefully “selected images that mirrored gender differences.” To replicate Goffman’s findings and demonstrate the validity of his categories with more robust statistical analyses, some studies have used professional annotators to encode small sets of images [4]. They found that individual females were indeed positioned in the top half of adverts more frequently than individual men. At the same time, groups of males dominated the top position compared to groups of women. However, they could not confirm Goffman’s hypothesis that there exists a left-right bias sensitive to gender. Goffman’s study has not only been directly replicated [2,18] but also served as the primary theoretical framework for studies of a very similar nature [5,24,20].

In sum: Goffman’s work and subsequent replication and follow-up studies have yielded formalizations of visual semiotics, i.e., how the relationship between visual elements of images produces gender displays. These formalizations offer guidelines on what we should look for when studying images. These guidelines also function as a clear starting point for the operationalization of computer vision algorithms. We need to have a clear idea of what we are looking for in images before we can turn to methods that detect specific elements in images.

Moreover, computer vision algorithms and the growing amount of digitized visual material make it possible to study gender displays on larger samples over more extended periods. This approach can add more robustness or offer possible correctives to the claims made by Goffman and existing replication studies.

Before we can proceed with such a replication study, we need to first establish to what extent existing computer vision algorithms and models *work* when they are applied to historical visual material. Therefore, this paper examines the use of computer vision techniques for studying gender in historical advertisements from 1950-1995. This study consists of three main tasks: 1) face detection, 2) detection of visual medium types, and 3) gender detection. We show how these tasks produce information that can be used to enrich historical collections, which can subsequently be used to answer long-standing questions about, for instance, the representation of gender. In simple terms, to be able to detect how males and females are represented in images, we first need to be able to detect them.

## 2 Related Work

In recent years, bias—systematic underrepresentation of certain groups or traits in data collection and/or statistical analysis—in computer vision models, algorithms, and data sets has become a primary concern for computer scientists. The presence of bias in a computer vision model could be due to certain design choices in an algorithm or the data set, or both. In relation to gender bias, [38] looked at the geo-diversity of the popular ImageNet and Open Images sets and concluded that they exhibited “an observable amerocentric (sic) and eurocentric (sic) representation bias”. [8] presented a method to audit the demographics (age and gender) of the ‘person’ category of ImageNet. In February 2020, Google announced that its popular Cloud Vision service would no longer label persons as either ‘male’ or ‘female’, noting that “a person’s gender cannot be inferred by appearance” [12]. Some of the founders of ImageNet have taken the almost exact opposite route by arguing that a large-scale annotation of gender, age, and skin color of the person category in their dataset will lead to a more representative dataset and, as a result, fairer algorithms [44]. Our paper builds on this kind of work by pointing to the historical dimension of gender bias in computer vision datasets and models.

While there is a considerable amount of work in the art-historical domain [3,27,17], cultural historians have only recently started to apply computer vision to their large collections of popular images. Advertisements are an excellent source for research that uses computer vision techniques to examine large-scale trends in modern visual culture. Computer vision algorithms have been applied to study how advertising ‘works’. Facial expressions and the representations of objects have been analyzed for the visual rhetorics and persuasiveness [40,46]. While these studies offer avenues to explore for historical analysis, the focus in the cited papers is on contemporary advertising, and their primary aim seems to be to better our understanding of the practice of advertising. In contrast, our approach is cultural-historical and our focus is on the historical dimension in advertising. Because gender displays, the way men and women are expected to look, change over time, the accuracy of gender recognition will fluctuate for tasks that have a historical dimension. By evaluating and adapting existing algorithms, this paper sheds light on the extent of this historical side of gender bias.

## 3 Data

Our data set is SIAMESET, which contains approximately one million advertisements that appeared in the Dutch newspaper *NRC Handelsblad* for the period 1945-1995 (Figure 1).<sup>3</sup> During the interwar period (1918-1940), advertising changed rapidly in the Netherlands. Technological advances made it cheaper to print images in newspapers. Influenced by developments in the American advertisement industry, Dutch agencies started to use visual material on a large scale

---

<sup>3</sup> Wevers, M., Lonij, J. (2017) SIAMESET. KB Lab: The Hague. <http://lab.kb.nl/dataset/siameset>

to increase sales and convey particular brand identities [35]. Throughout the German occupation of the Netherlands (1940-1945) and in the immediate post-war years, paper was scarce. However, starting in the early 1950s, the number of advertisements containing visual elements started to increase considerably. Figure 1 confirms the low number of adverts in the years right after the Second World War. Because of this, we focus on the post-war period.

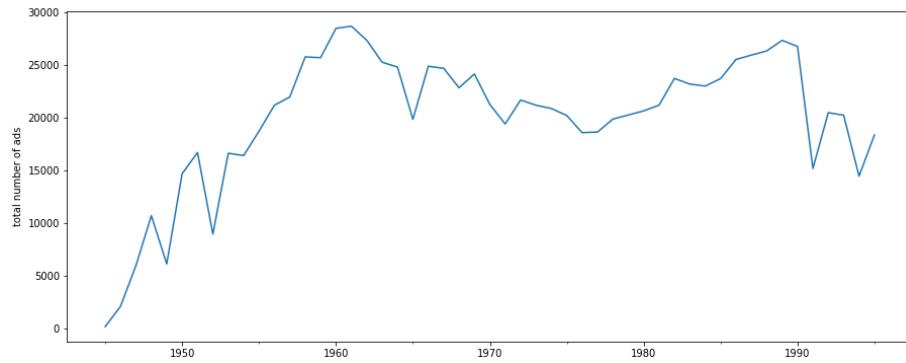


Fig. 1: Total number of advertisements per year in SIAMESET

From SIAMESET, we randomly sampled 1K images per year for the period 1950-1995, resulting in a data set of forty-five thousand advertisements. In these advertisements, a group of six annotators drew bounding boxes around faces and annotated their gender. We followed standard protocol by requiring every bounding box to be as small as possible while including all visible parts of the face [34,10]. We asked annotators to categorize the faces as either “male” or “female”. Although we are aware that binarization of gender is problematic and reductionist, advertisers, however, were quite explicit in their representation of gender. Therefore, there was little ambiguity found in establishing gender for adults. We also annotated children as “girl” and “boy”. Because of the ambiguity of these annotations, we decided to exclude these from the analysis.

Out of the data set’s forty-five thousand advertisements, 8,522 contained at least one face. In total, the data set contained approximately 10.2k men and 8.5k women. Figure 2 shows the average number of faces in a single ad, containing a face, either a male or a female. The error bars show the 95% bootstrap confidence intervals. We see that the average number of male faces has more variance than that of female faces. In part, this variation is due to pictures with groups of people, or a sports team, which often-times only included males. In other words, in cases of pictures with large numbers of faces, these are more often male faces (see Figure 3). This finding in itself is already a remarkable result in terms of gender representations.

Also, Figure 2 shows that around 1975, the average number of female faces started to decrease and remained lower than that of males for the next twenty

years. This decreased presence of females in the newspaper might be related to the merger of the *Algemeen Handelsblad* into the *NRC Handelsblad*, a more liberal newspaper. After the merger, the newspaper also set out to target business-men as its main readership [15]. A comparative analysis with other newspapers would shed more light on these kinds of issues. During annotation, it became clear that the data set contained predominantly white males and females. For this reason, we could not determine how well our algorithms performed on people with an ethnicity other than white [8].

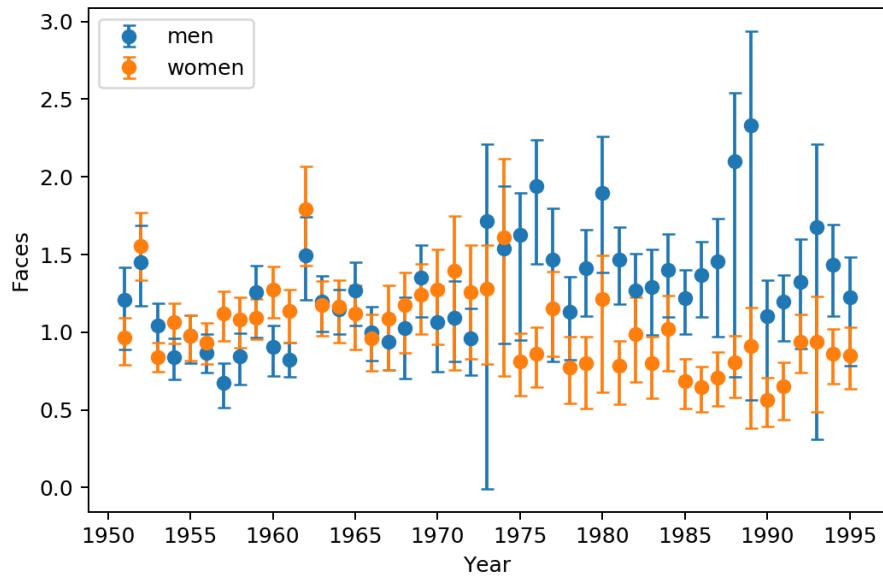


Fig. 2: Average number of male and female faces per advert in SIAMESET

## 4 Analysis

We divided our analysis of historical advertisements into three tasks. First, we benchmarked face detection algorithms on the annotated collection. Second, we created a classifier to detect the visual medium type in which a face was represented as either an illustration or a photograph. Third, we trained a convolutional neural network to classify the gender of the extracted faces. The following subsections describe these three tasks in more detail.<sup>4</sup>

<sup>4</sup> The code can be found here: <https://github.com/melvinwevers/detecting-faces-medium-types-gender>



Fig. 3: Group of males in a 1988 advert for mattresses

#### 4.1 Face Detection

Face detection is a common computer vision task [47,41,29]. It involves the estimation of a bounding box around a face in an image. Recent innovations in Deep Learning have resulted in increased accuracy of face detection techniques. Especially, the accuracy of the detection of tiny faces [16], occluded faces, and face categories with large variance [48] has improved greatly.

SIAMESET contains images with both a high variance in image size as well as variance in the size of the different faces depicted on one single advertisement. Moreover, advertisements contain faces represented in different types of visual media, using photography, illustration, and, in some cases, photography and illustration concurrently. Faces, thus, both appear in iconic ways, as in being the central object of the image, and also in non-iconic ways, for example, in adverts where a product is the central object. Since we are working with historical material, the printing techniques used to reproduce the images also varies considerably: from high-definition, full-color, full-page advertisements, to small, smudgy, black-and-white notices. Furthermore, significant differences in quality might arise from the digitization process [11]. Recently, non-iconic views of objects and scenes have been purposefully collected to improve the accuracy of computer vision techniques [23]. Following these efforts, we present a highly heterogeneous annotated set that can be used to test the performance of existing face and gender detection algorithms.<sup>5</sup>

For the evaluation of face detection algorithms, we contrasted the widely-used OpenCV’s DNN module with two state-of-the-art models, Dual Shot Face Detection (DSFD) and RetinaFace [22,7]. OpenCV DNN’s module uses a Single Shot Detector, which predicts the bounding boxes and the confidence in one single shot, with ResNet-10 as a backbone. Unfortunately, it is not specified on

---

<sup>5</sup> The annotated data set and models can be found on Zenodo using the following DOI: 10.5281/zenodo.4008991

what data and in what manner the model was trained. DSFD adds a step to a Single Shot Detector that improves its ability to detect tiny faces and occluded faces. Part of this additional step is the use of a Feature Enhance Module that improves discriminability and robustness of features [22]. RetinaFace, which is based on a single-stage design, employs a multi-task learning strategy that predicts a face score, a bounding box, five facial landmarks, and 3D position and correspondence of each facial pixel. The use of facial landmarks improves the algorithm’s ability to learn. RetinaFace runs on either a MobileNet or a ResNet-50 backbone. The former is considerably faster and less memory-intensive, while the latter has slightly better results [7].

When applied to the often-used WIDERFACE data set, RetinaFace outperforms DSFD, and is much faster than two-stage methods, such as DSFD (see Table 1). The WIDERFACE data set consists of a total of 32,203 images, containing 393,703 annotated faces with significant variations in scale, pose, and occlusion. Each image is further defined into three levels of difficulty: “Easy”, “Medium”, “Hard” based on the detection rate of a baseline detector [45].

Table 1: Average precision on WIDER FACE taken from [22,7]

Method	Easy	Medium	Hard
DSFD	96.6	95.7	90.4
RetinaFace	<b>96.9</b>	<b>96.2</b>	<b>91.9</b>

**Results** For OpenCV’s DNN, we used the default FaceDetector weights; for the other two algorithms, we relied on weights pre-trained on the WIDER FACE data set. We selected the RetinaFace with ResNet-50 as a backbone rather than MobileNet, the former is slower but has better accuracy on the WIDER FACE data set. For this study, the decreased speed is no issue since we value accuracy above speed. Using these weights, we applied the three algorithms to our annotated data set. We used Average Precision to compare the performance of the algorithms. Average Precision (AP) captures the area under the curve (AUC) of the Precision and Recall curve. We interpolated all data points, following the method of the PASCAL VOC challenge [6]. We employ an IOU (intersection over union) threshold of 0.5 when calculating the AP. This score is calculated by dividing the area of overlap between the predicted bounding box and the ground-truth box by the area of the predicted bounding box and ground-truth bounding box, i.e. the area of union.

DSFD (AP: 71%) and RetinaFace (AP: 68.13%) clearly outperform the Open CV DNN module (AP: 20.54%) (Figure 4). The slightly better performance by DSFD over RetinaFace can most definitely be explained by the former’s ability to detect faces with a large size variance, which is quite common in advertisements. Its default scaling algorithm is more versatile than the one that DSFD offers.

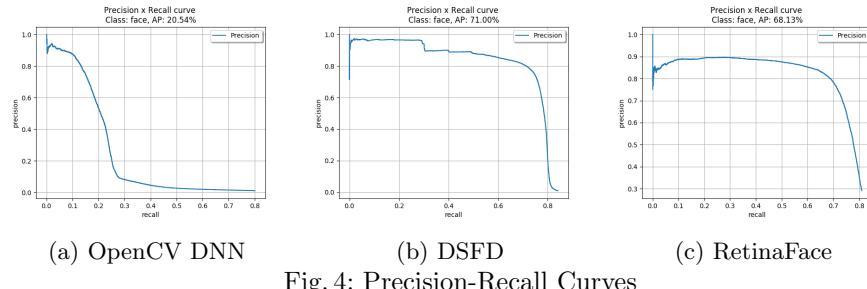


Fig. 4: Precision-Recall Curves

A review of false negatives suggests that relatively large faces are often not detected by the algorithms. In Figure 5, for example, the large face is not detected. The scaling algorithm probably prohibits drawing a bounding box around objects occupying the majority of the image's area. Further optimization of the scaling algorithms might solve this issue.



Fig. 5: Advertisement with a large face

## 4.2 Visual Medium Detection

Persons in advertisements can be depicted using different visual media types, ranging from photograph to illustration, or sometimes both in the same advertisement or even the same face. Furthermore, faces can be illustrated using a wide range of different styles, from highly realistic to highly abstract (see Figure 6). Previous work shows that convolutional neural networks can be used with high accuracy to distinguish between different media types [44,43]. Because of

the co-existence of different visual media and visual styles within a single advertisement, we built a classifier on the level of the face rather than on the level of the full advertisements. Instead of constructing an object detector that detects faces in their specific styles, we combined the face detection with an additional classifier for the visual medium. This two-step approach allowed us to benefit from the ability of existing algorithms, such as RetinaFace, to detect tiny faces and large variance between faces.



Fig. 6: Advertisement containing illustrations and a photograph (1959)

The visual medium classifier was trained on a set of 242 illustrated and 259 photographed faces drawn from the set of forty-five thousand advertisements. We extracted the faces from the advertisements and manually classified them as either illustration or photograph. Next, we applied data augmentation on this image set (resizing, center cropping, horizontal flipping, and random rotation). This augmentation generated new training data from the original set by adding perturbations to the images. In every batch of data, the network is presented with new, slightly different versions of the input data, which forces the network to learn more robust features, making the resulting less prone to overfitting.

Finally, we fed these augmented images into a pre-trained VGG16, a comparatively small network, and extracted the encoded features as a flattened vector from the penultimate Max Pooling layer.<sup>6</sup> We trained a Linear SVC with default settings, using Scikit-learn, on these encoded features.

**Results** With ten-fold cross-validation, our classifier reaches a mean accuracy of .91 with a standard deviation of 0.04. In other words, using a relatively small set of training data, our classifier can distinguish between photographs and illustrations with high accuracy. Closer inspection of the misclassified predictions shows

<sup>6</sup> A parameter search found that this layer provided the best results for our classification task.

that these are mostly low-quality images or images of which the visual medium is difficult to determine, even for a human annotator. A possible strategy is augmenting the WIDERFACE data set with illustrated faces. Future work will explore whether Deep Adaptation Networks might be leveraged to learn faces across different modalities [25].

### 4.3 Gender Detection

The detection of gender in images using some form of computer vision has been explored extensively [30,19,9]. Convolutional neural networks can pick up the necessary features to make highly accurate predictions of gender [21,1,33]. The conception of gender used in this kind of research is binary, static, and highly normative. Partly based on Goffman’s work, gender studies scholars like Judith Butler have argued that gender is performed. Notions of masculinity and femininity—the way that men and women are expected to look and behave—are not derived from biological sex but from the constant performance of gender in society, for example, in advertisements [39].

These insights are problematic for gender detection techniques because they reveal that annotating faces into binary gender categories is not a clear-cut analytic procedure but a gender performance in itself. The fact that the performance of gender changes over time further complicates the matter. Are we using computer vision to study historical gender displays, or are we superimposing our notion of gender on the historical archive? We will address this question in a forthcoming study.

For the classification of gender in advertised persons, we constructed two data sets. The first set consists of two classes: “male” and “female”, excluding images with a width or height smaller than 100 pixels. We separated the data set into a training, validation, and test set according to an 80, 10, 10 division. Table 2 shows the number of images in each set.

Table 2: Training set gender classifier

	train	validation	test	total
male	5,068	649	644	6,361
female	4,489	555	1,204	5,617
total	9,557	1,204	1,848	11,888

For the second set, we used the visual medium classifier described in section 4.2 to divide the data set into four classes: male photo, female photo, male illustration, and female illustration. This set was also divided according to an 80, 10, 10 division, resulting in the set described in Table 3.

For gender detection, we relied on a VGGFACE, a VGG16 network with weights pre-trained on two large face data sets: Labeled Faces in the Wild and

Table 3: Training set gender classifier divided by style

	train	validation	test	total
male photo	2,865	373	374	3,612
male illustration	2,205	271	272	2,748
female photo	2,532	293	294	3,119
female illustration	1,959	280	281	2,520
total	9,561	1,217	1,221	11,999

the YouTube Faces dataset [31].<sup>7</sup> We compared three optimizers: SGD, ADAM, and ADAbound, and achieved the best performance using the latter. The latter is described to train as fast as ADAM with the performance of SGD [26]. We held out ten percent of the training data as a validation set. Next, we augmented the training data by rotating the images, shifting the width and height, changing the shearing, flipping them horizontally, and applying a random eraser. The latter randomly erases parts of the images and replaces them with random pixels, to make the detection of features more difficult, leading to more robust features [49].

We used a two-step approach to fine-tune the network. As part of the first step, we fine-tuned the head weights. For this, we loaded the VGGFace model with pre-trained weights and left the top layers off. We flattened the last layer and added two fully-connected dense layers with 512 hidden dimensions. Before adding the classification layer with a Sigmoid activation, we added Dropout to reduce overfitting. As part of our training strategy, we used early stopping (patience=10) and reduced the initial learning rate (0.001) when it did not change for three epochs. For the second step, we unfroze the fully-connected head layers and the final convolution block, allowing them to be fine-tuned as well. Next, we trained the network for a maximum of 20 epochs, or when the learning rate plateaued—this additional step slightly improved accuracy.

**Results** We achieved an  $F_1$ -score of .91 for females and .92 for males (Table 4). However, when we tried to detect both medium and gender—a more difficult task—the accuracy drops to an average of .84 (Table 5). If the task requires a focus on a particular medium, it might be beneficial to classify the medium and then apply gender detection to those particular media. However, as this result shows, we can estimate gender with a high degree of accuracy, regardless of the medium.

## 5 Conclusion

This paper has presented: 1) a heterogenous, non-iconic annotated data set of images—a subset of SIAMESET—which can be used to test the performance of

---

<sup>7</sup> We also tried a Resnet-101 architecture, but this approach decreased accuracy.

Table 4: Classification report gender detection

	prediction	recall	f1-score	support
female	0.91	0.92	0.91	573
male	0.93	0.91	0.92	644
weight avg	0.92	0.92	0.92	1,217

Table 5: Classification report gender and style detection

	prediction	recall	f1-score	support
female drawing	0.90	0.83	0.86	280
female photo	0.83	0.86	0.84	293
male drawing	0.81	0.83	0.82	271
male photo	0.84	0.84	0.84	373
weight avg	0.84	0.84	0.84	1,217

existing face and gender detection algorithms on historical material, 2) our efforts to evaluate and improve the performance of three face detection algorithms (OpenCV, DSFD, RetinaFace) on historical material; 3) the training of a visual medium classifier to separate photographs from illustrations, and 4) the training of a gender classifier on the extracted faces.

We have shown that we can detect faces, visual media, and gender with a high degree of accuracy in historical advertisements. It remains challenging to detect gender when faces are of low quality, relatively small, or relatively large. Further optimization of scaling and upscaling of the images to improve their resolution might solve these issues. Notwithstanding these concerns, we are confident that we can apply these algorithms and fine-tuned models to enrich other large visual historical collections and use them for a future study, in which we replicate Goffman’s formalizations of gender displays using larger data set.

## References

1. Antipov, G., Berrani, S.A., Dugelay, J.L.: Minimalistic CNN-based ensemble model for gender prediction from face images. *Pattern Recognition Letters* **70**, 59 – 65 (2016)
2. Belknap, P., Leonard, W.M.: A Conceptual Replication and Extension of Erving Goffman's Study of Gender Advertisements. *Sex Roles* **25**(3), 103–118 (1991)
3. Bell, P., Schlecht, J., Ommer, B.: Nonverbal Communication in Medieval Illustrations Revisited by Computer Vision and Art History. *Visual Resources* **29**(1-2), 26–37 (Jun 2013). <https://doi.org/10.1080/01973762.2013.761111>
4. Bell, P., Milic, M.: Goffman's Gender Advertisements revisited: Combining content analysis with semiotic analysis. *Visual Communication* **1**(2), 203–222 (Jun 2002). <https://doi.org/10.1177/14703572000100205>
5. Conley, T.D., Ramsey, L.R.: Killing Us Softly? Investigating Portrayals of Women and Men in Contemporary Magazine Advertisements. *Psychology of Women Quarterly* **35**(3), 469–478 (Sep 2011)
6. Davis, J., Goadrich, M.: The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning - ICML '06. pp. 233–240. ACM Press, Pittsburgh, Pennsylvania (2006)
7. Deng, J., Guo, J., Zhou, Y., Yu, J., Kotsia, I., Zafeiriou, S.: RetinaFace: Single-stage Dense Face Localisation in the Wild. arXiv:1905.00641 [cs] (May 2019)
8. Dulhanty, C., Wong, A.: Auditing ImageNet: Towards a Model-driven Framework for Annotating Demographic Attributes of Large-Scale Image Datasets. arXiv:1905.01347 [cs] (2019), arXiv: 1905.01347
9. Eidinger, E., Enbar, R., Hassner, T.: Age and Gender Estimation of Unfiltered Faces. *IEEE Transactions on Information Forensics and Security* **9**(12), 2170–2179 (Dec 2014), conference Name: IEEE Transactions on Information Forensics and Security
10. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision* **88**(2), 303–338 (2010)
11. Fyfe, P.: An Archaeology of Victorian Newspapers. *Victorian Periodicals Review* **49**(4), 546–577 (Dec 2016)
12. Ghosh, S.: Google AI will no longer use gender labels like ‘woman’ or ‘man’ on images of people to avoid bias (Feb 2020), <https://www.businessinsider.com/google-cloud-vision-api-wont-tag-images-by-gender-2020-2>
13. Goffman, E.: *Gender Advertisements*. Harper and Row, New York (1976)
14. Goldman, R.: *Reading Ads Socially*. Routledge, New York (1992)
15. van der Hoeven, P.: *Het succes van een kwaliteitskrant*. Prometheus, Soest (2012)
16. Hu, P., Ramanan, D.: Finding Tiny Faces. arXiv:1612.04402 [cs] (Apr 2017)
17. Impett, L.: *Painting by Numbers. Computational Methods and the History of Art*. Ph.D. thesis, École Polytechnique Fédérale de Lausanne, Lausanne (2020)
18. Kang, M.E.: The portrayal of women’s images in magazine advertisements: Goffman’s gender analysis revisited. *Sex Roles* **37**(11), 979 (Dec 1997)
19. Klare, B.F., Klein, B., Taborsky, E., Blanton, A., Cheney, J., Allen, K., Grother, P., Mah, A., Burge, M., Jain, A.K.: Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1931–1939 (Jun 2015), iSSN: 1063-6919

20. Kuipers, G., van der Laan, E., Arfini, E.A.G.: Gender models: changing representations and intersecting roles in Dutch and Italian fashion magazines, 1982–2011. *Journal of Gender Studies* **26**(6), 632–648 (Nov 2017)
21. Levi, G., Hassner, T.: Age and gender classification using convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 34–42 (2015)
22. Li, J., Wang, Y., Wang, C., Tai, Y., Qian, J., Yang, J., Wang, C., Li, J., Huang, F.: DSFD: Dual Shot Face Detector. arXiv:1810.10220 [cs] (Apr 2019)
23. Lin, T.Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L., Dollár, P.: Microsoft COCO: Common Objects in Context. arXiv:1405.0312 [cs] (2015), arXiv: 1405.0312
24. Lindner, K.: Images of Women in General Interest and Fashion Magazine Advertisements from 1955 to 2002. *Sex Roles* **51**(7), 409–421 (2004)
25. Long, M., Cao, Y., Wang, J., Jordan, M.I.: Learning Transferable Features with Deep Adaptation Networks. arXiv:1502.02791 [cs] (May 2015)
26. Luo, L., Xiong, Y., Liu, Y.: Adaptive gradient methods with dynamic bound of learning rate. In: International Conference on Learning Representations (2019), <https://openreview.net/forum?id=Bkg3g2R9FX>
27. Madhu, P., Kosti, R., Mühlberg, L., Bell, P., Maier, A., Christlein, V.: Recognizing Characters in Art History Using Deep Learning. In: Proceedings of the 1st Workshop on Structuring and Understanding of Multimedia heritAge Contents. pp. 15–22. SUMAC ’19, Association for Computing Machinery, Nice, France (2019)
28. Marchand, R.: Advertising the American dream: making way for modernity, 1920–1940. University of California Press, Berkeley (1985)
29. Merler, M., Ratha, N., Feris, R.S., Smith, J.R.: Diversity in Faces. arXiv:1901.10436 [cs] (Apr 2019), arXiv: 1901.10436
30. Ng, C.B., Tay, Y.H., Goi, B.M.: Recognizing Human Gender in Computer Vision: A Survey. In: Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M.Y., Weikum, G., Anthony, P., Ishizuka, M., Lukose, D. (eds.) PRICAI 2012: Trends in Artificial Intelligence, vol. 7458, pp. 335–346. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
31. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep Face Recognition. In: Proceedings of the British Machine Vision Conference 2015. pp. 41.1–41.12. British Machine Vision Association, Swansea (2015)
32. Parkin, K.: Food Is Love: Advertising and Gender Roles in Modern America. University of Pennsylvania Press, Philadelphia (2007)
33. Ranjan, R., Patel, V.M., Chellappa, R.: HyperFace: A Deep Multi-task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition. arXiv:1603.01249 [cs] (Dec 2017)
34. Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T.: LabelMe: A Database and Web-Based Tool for Image Annotation. *International Journal of Computer Vision* **77**(1), 157–173 (May 2008)
35. Schreurs, W.: Geschiedenis van de Reclame in Nederland. Het Spectrum, Utrecht (2001)
36. Schroeder, J.E., Zwick, D.: Mirrors of Masculinity: Representation and Identity in Advertising Images. *Consumption Markets & Culture* **7**(1), 21–52 (2004)
37. Scranton, P.: Beauty and Business: Commerce, Gender, and Culture in Modern America. Routledge, London (2014)

38. Shankar, S., Halpern, Y., Breck, E., Atwood, J., Wilson, J., Sculley, D.: No Classification without Representation: Assessing Geodiversity Issues in Open Data Sets for the Developing World. arXiv:1711.08536 (2017), arXiv: 1711.08536
39. Smith, G.: Reconsidering Gender Advertisements. Performativity, Framing and Display. In: Jacobsen, M. (ed.) *The Contemporary Goffman*, pp. 165–184. Routledge, New York (2010)
40. Thomas, C., Kovashka, A.: Persuasive faces: Generating faces in advertisements. arXiv preprint arXiv:1807.09882 (2018)
41. Viola, P., Jones, M.J.: Robust real-time face detection. *International journal of computer vision* **57**(2), 137–154 (2004)
42. Weinbaum, A.E., Thomas, L., Ramamurthy, P., Poiger, U., Yue Dong, M., Barlow, T. (eds.): *The Modern Girl around the World: Consumption, Modernity, and Globalization*. Duke University Press, Durham (2008)
43. Wevers, M., Smits, T.: The visual digital turn: Using neural networks to study historical images. *Digital Scholarship in the Humanities* **35**(1), 194–207 (Apr 2020)
44. Yang, K., Qinami, K., Fei-Fei, L., Deng, J., Russakovsky, O.: Towards fairer datasets: filtering and balancing the distribution of the people subtree in the ImageNet hierarchy. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. pp. 547–558. Association for Computing Machinery, Barcelona, Spain (2020)
45. Yang, S., Luo, P., Loy, C.C., Tang, X.: Wider face: A face detection benchmark. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
46. Ye, K., Honarvar Nazari, N., Hahn, J., Hussain, Z., Zhang, M., Kovashka, A.: Interpreting the Rhetoric of Visual Advertisements. *IEEE Transactions on Pattern Analysis and Machine Intelligence* pp. 1–1 (2019)
47. Zafeiriou, S., Zhang, C., Zhang, Z.: A survey on face detection in the wild: past, present and future. *Computer Vision and Image Understanding* **138**, 1–24 (2015)
48. Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X., Li, S.Z.: Sfd: Single Shot Scale-invariant Face Detector. arXiv:1708.05237 [cs] (Nov 2017)
49. Zhong, Z., Zheng, L., Kang, G., Li, S., Yang, Y.: Random erasing data augmentation (2017)