

# IIIT-CFW: A Benchmark Database of Cartoon Faces in the Wild

Ashutosh Mishra<sup>1</sup>, Shyam Nandan Rai<sup>1</sup>, Anand Mishra<sup>2</sup>, C. V. Jawahar<sup>1,2</sup>

<sup>1</sup> IIIT Chittoor, Sri City, India

<sup>2</sup> CVIT, KCIS, IIIT Hyderabad, India

**Abstract.** In this paper, we introduce the cartoon faces in the wild (IIIT-CFW) database and associated problems. This database contains 8,928 annotated images of cartoon faces of 100 public figures. It will be useful in conducting research on spectrum of problems associated with cartoon understanding. Note that to our knowledge, such realistic and large databases of cartoon faces are not available in the literature.

**Keywords:** Cartoon faces, cartoon face synthesis, sketches, fine art

## 1 Introduction

We introduce a database of 8,928 annotated cartoon faces designed as an aid in studying the problem of unconstrained cartoon understanding. The database can be downloaded from our project website.<sup>1</sup>

Our database which we refer as IIIT cartoon faces in the wild (or IIIT-CFW in abbreviation) is designed to study spectrum of problems associated with cartoon understanding. Some of these problems are listed in Section 5. Oxford dictionary defines cartoon as follows: *a simple drawing showing the features of its subjects in a humorously exaggerated way, especially a satirical one in a newspaper or magazine*. However, the modern usage of cartoon has extended to any non-realistic, semi-realistic drawing or painting intended for satire, caricature, or humor, or to the artistic style of such works. Following the modern definition, our database of cartoon faces contains caricatures, paintings, cartoons and sketches of internationally well-known public figures.

Research in face recognition has been significantly advanced in last few years [1–6]. Thanks to multiple databases introduced (we have discussed few of them in Section 2). These databases have triggered the research and study in face recognition domain. On the other hand, there have been only very few attempts to address the problem of recognizing cartoon faces [7, 8], that too with experiments on small test sets. Cartoon understanding has many application areas. One of the fundamental application is: given an image containing cartoons of people, answering questions, such as who are in the image and what are they doing? (See Fig. 1). Such understanding, (i) can help visually impaired people

---

<sup>1</sup> <http://cvit.iiit.ac.in/projects/cartoonFaces/>



(a)



(b)

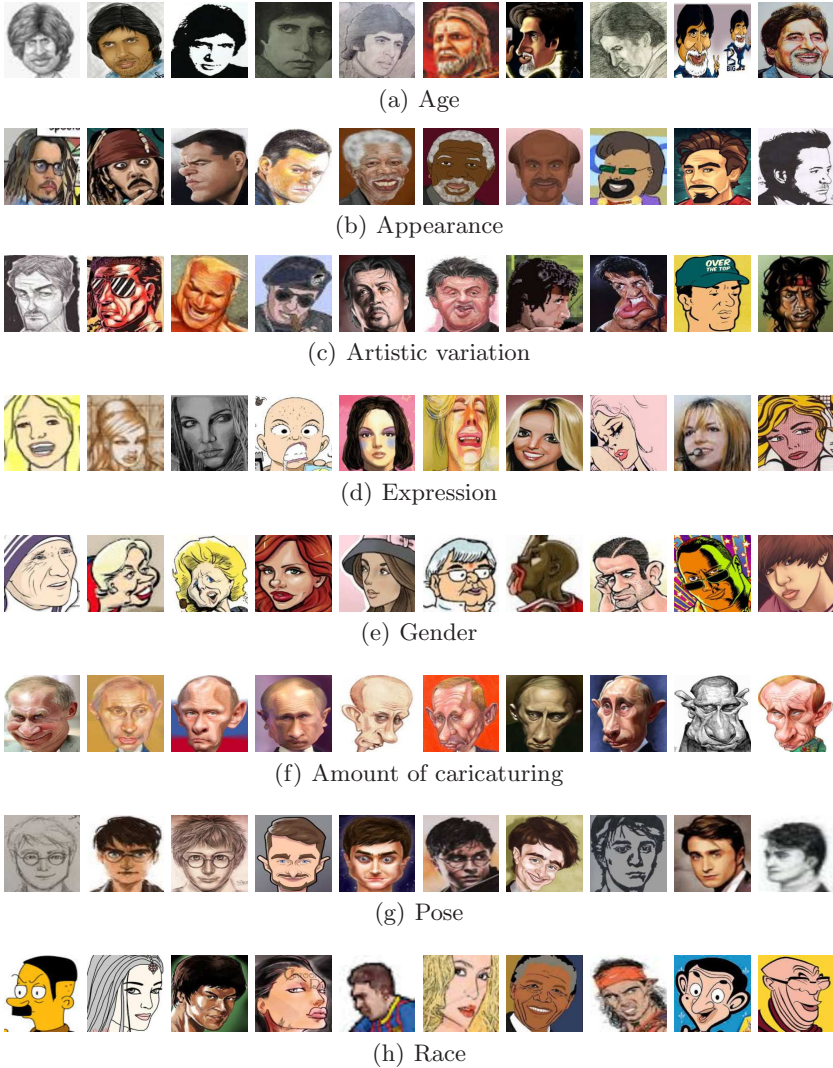
**Fig. 1.** Who are these and what are they doing? There has been lots of advancement in face recognition. However, the problem of recognizing cartoon characters has not been looked rigorously. Our database: the IIIT-CFW will be useful for comprehensive studies of understanding cartoon images.

to understand cartoon images or movies, and (ii) can be used to automatically censor communal or politically incorrect cartoons in the social media. Other applications of understanding cartoons are generating realistic cartoon faces, generating various realistic facial expression in cartoons, etc. These applications have enormous importance in fine art.

Cartoon face recognition is the first step towards larger understanding of cartoon images. The problem of cartoon face recognition is closely related to real face recognition, however poses many additional challenges. For example, (i) artistic variations, (ii) limited examples, (iii) magnitude less data than real faces, (iv) highly caricatured. The other challenges such as pose, expression, age, illumination variations, race (e.g., cartoon faces of Asian characters are very different in look from that of European characters) and gender variations remain similar as face databases. We show some examples of our database illustrating the above mentioned challenges in Fig. 2. Another closely related area to cartoon face recognition is *sketch recognition* [9, 10]. However, in sketch recognition, most of the attempts were made to recognize handwritten geometric objects. Our database of cartoon faces is much broader and challenging than simple geometric sketches or drawing, and sketch recognition techniques may not be directly applicable to these images. We believe that the IIIT-CFW will prove a first important step towards conducting research in broad area of cartoon understanding.

Before proceeding with the details, we briefly summarize the IIIT-CFW database here:

- This database contain cartoon images of 100 international celebrities (politician, actor, singer, sports person, etc) and 8,928 images in all.
- The images in this database are harvested from the web. These images contain cartoons drawn in totally unconstrained setting.
- The IIIT-CFW contains detailed annotations and it can be used for wide spectrum of problems including cartoon face recognition, cartoon face verification, cartoon image retrieval, relative attributes in cartoons, gender or age group estimation from the cartoon faces, and cartoon faces synthesis.
- We provide face bounding box along with some additional attributes such as age group, view, expression, pose etc.



**Fig. 2.** Sample faces of the IIIT-CFW. It has a lot of variations, such as (a) age (young and old), (b) appearance (with beard, with glass), (c) artistic (pure cartoon, sketch and caricature) (d) expression (happy, sad, seducing, angry, etc), (e) gender (male and female), (f) amount of caricaturing, (g) pose (frontal and non-frontal), and (h) race (Asian, American, African, etc).

- Most of the images of this database are in color with very few in gray scale.
- We provide standard train and test splits according to the problem.

More details are provided in the remainder of this paper which is organized as follows. In Section 2, we discuss some of the related databases. Since cartoon

Databases	Variation					Size (# people, # images)	Target applications
	R	G	P	E	A		
ORL [11]	✓	✓	✓	✓	✓	(40, 400)	$P_1$
Yale [12]	✓	✓	×	✓	✓	(15, 165)	$P_1$
Indian Face [13]	×	✓	✓	✓	✓	(40, 440)	$P_1$
LFW [14]	✓	✓	✓	×	✓	(5749, 13233)	$P_1, P_2$
Texas 3D [15][16]	✓	✓	×	✓	✓	(105, 1149)	$P_1$
FEI [17]	×	✓	✓	✓	✓	(200, 2800)	$P_1$
PubFig83 [18]	✓	✓	✓	✓	✓	(100, 8300)	$P_2$
Celeb Faces [19]	✓	✓	✓	✓	✓	(5346, 87628)	$P_1, P_2$
IMFDB [20]	×	✓	✓	✓	✓	(100, 34512)	$P_1, P_2$
LWF [22]	✓	✓	✓	✓	✓	(1500, 8500)	$P_1, P_2$
FaceScrub [21]	✓	✓	✓	✓	✓	(530, 106863)	$P_1, P_3$
Facebook [5]	-	-	-	-	-	(200M, 8M)	*
Google [4]	-	-	-	-	-	(4.4M, 4030M)	*
VGG [6]	✓	✓	✓	✓	✓	(2363, 2.6M)	$P_1, P_2$
<b>The IIIT-CFW</b>	✓	✓	✓	✓	✓	(100, 8927)	$P_1, P_2, P_3, P_4, P_5, P_6, P_7$

**Table 1.** Variations, size and target applications of various face databases vs our cartoon database. Our database provides suitable annotations and train-test splits to study wide range of problems (R: race, G: gender, P: pose, E: expression, A: appearance,  $P_1$ : face recognition,  $P_2$ : face verification,  $P_3$ : face detection,  $P_4$ : photo2cartoon  $P_5$ : cartoon2photo  $P_6$ : attributes based search  $P_7$ : relative attributes in faces, \*: Not available publicly).

understanding is novel problem, however, closely related to face recognition, we provide brief discussion on popular face recognition databases, and compare them with ours. Section 3 describes image collection and annotation scheme. We give detail database statistics and analyze the complexity of the database in Section 4. We then discuss intended usage of this database where we discuss about spectrum of problems which can be studied on this database. We finally provide concluding remarks for this database paper in the Section 6.

## 2 Related Databases

The problem of cartoon face understanding lacks standard benchmark databases. On the other hand, the real face recognition community has a clear advan-

tage of having plethora of standard benchmarks. Starting from popular ORL database [11] in the year 1994 and Yale database [12] in the year 1997, numerous standard benchmarks have been released to address face recognition problem. Each of these databases are more challenging and realistic as compared to their past counterparts. In this section we present a brief survey of some of the popular face databases and compare them with our cartoon face database. We summarize this comparison in Table 1.

- **ORL face database** [11] This database consists of 400 face images of 40 distinct characters. The images for some characters were taken at different time, with varying lighting and facial conditions. The images are face centered and taken against a dark background.
- **Yale face database** [12] This database contains 165 face images of 15 characters. The faces in this database are frontal and have very small variations.
- **Labeled Wikipedia faces (LWF)** [22] The LWF database contains 8.5 K faces of approximately 1.5 K individuals. These images were collected from the Wikipedia living people category. The database, in addition to images, contains some meta data, e.g., the source images, image captions (if available) and person name detection results.
- **FaceScrub** [21] It contains 1,06,863 images of 530 celebrities, with about 200 faces per person. The images in the database were harvested from the web and are taken under real-world situations (i.e., uncontrolled settings). Images apart from the queried person were discarded in order to build the database.
- **FEI face database** [17] is a Brazilian face database that contains a set of 2800 face images, 14 images for each of 200 individuals. Equal number of male and female subjects are present in the database. Colored images taken against a white homogeneous background in an upright frontal position with profile rotation of up to about 180 degrees are present in the database.
- **Indian face database** [13] is a database consisting eleven different images, each of 40 distinct subjects. The images were acquired in an upright frontal position against a bright homogeneous background. The database also provide detail annotations for face orientation. Some emotions such as neutral, smile, laughter, sad or disgust are also included in annotation of this database.
- **Labeled faces in the wild** [14] The LFW database of face images is designed for studying the problem of unconstrained face recognition. The database contains more than 13,000 images of 1680 characters. These images were collected from the web. The faces in the database are result of Viola-Jones face detector [23] and hence, are roughly frontal.
- **CelebFaces** [19] This database contains 5436 characters and 87628 face images in all. It contains faces of those celebrity who do not exist in LFW. Images in this database were collected from web.
- **Youtube face** [24] The database contains 3,425 videos of 1,595 different people. All the videos were downloaded from YouTube.

- **PubFig83** [18] The PubFig83 database contains 8300 cropped facial images of 83 unique public figures. This database aims to provide the solution for the problem of recognizing identities from near-frontal faces.
- **Indian Movie Face Database (IMFDB)** [20] is a large unconstrained face database consisting of 34512 images of 100 Indian actors from more than 100 videos. The images have been obtained manually by cropping them from video frames in turn leading to high degree of variability in terms of scale, pose, expression, illumination, age, resolution, occlusion, and makeup. The IMFDB provides a detailed annotation of every face in terms of age, pose, gender, expression and type of occlusion. Standard train-val-test division is missing in this database.
- **Industrial benchmarks** [4, 5]. These face databases introduced by Google and Facebook are the largest in size. However, they are not publicly available. The Google and Facebook database contains 200 Million and 4.4 Million faces of 4030 and 8 Million characters respectively.
- **VGG face database** [6]. This database is introduced by Visual geometry group of Oxford university. It contains 2.6 Million faces of 2363 characters harvested from the web.

Contrary to the real face databases, ours is a first attempt to create a realistic and large database of cartoon faces. In many ways our database is similar to pubFig and celebFace, e.g., both these databases and ours contain faces of public figures harvested from the web. However, our database has additional challenges due to the fact that it contains cartoon faces. Moreover, unlike pubFig which can only be used for face verification, our database can be used to address many different problems (see Section 5). In short, the major highlights of our database are as follows: (i) it is a first large database of cartoon faces, (ii) it contains detail annotations, and hence, it can be used to study spectrum of problems associated with cartoon understanding.

### 3 Data collection and Annotations

The CFW is harvested from Google image search. We have used queries like ‘Obama + cartoon’, ‘Emma Watson + cartoon’, and so on, to collect cartoon images of 100 public figures. These public figures are chosen from different categories such as sports, politics, art, science, etc and different countries such as USA, UK, Australia, India, etc. Once images are collected, we use manual filtering to remove irrelevant images. We then gave these images to our annotation team of three people who were asked to draw face bounding boxes and provide following attributes to each face:

- Type of cartoon: cartoon, cartoon sketch, caricature
- Pose: frontal, non frontal
- Expression: happy, sad, thoughtful, seductive, sorrow, angry, serious, frightened, crying, shocked
- Age group: young, old





Fig. 3. Example of annotations provided in the IIIT-CFW.

- Gender: male, female
- Beard: yes, no
- Glass: yes, no

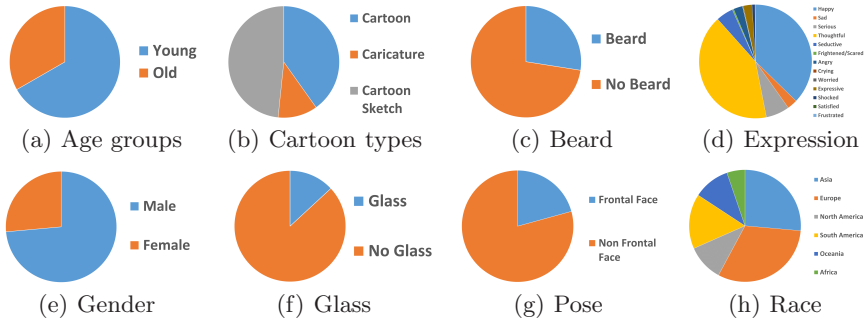
For the subjective attributes like age group and expression, we used voting scheme among the annotators. In addition to face annotation, attributes for each face, some metadata details such as the name of the personality and the identity number (id number) is also assigned.

Annotation is stored and described in the form of XML. The XML contains the information about the unique id number assigned to each personality in addition to the name of the personality and the above seven attributes. Moreover, the coordinates  $(x_1, y_1, width, height)$  of the face bounding box are also included in the XML description. Fig. 3 shows an example of annotations we provide in the CFW. For annotation, image markup tool which was released by Mozilla under public license version-1.1 was used.<sup>2</sup>

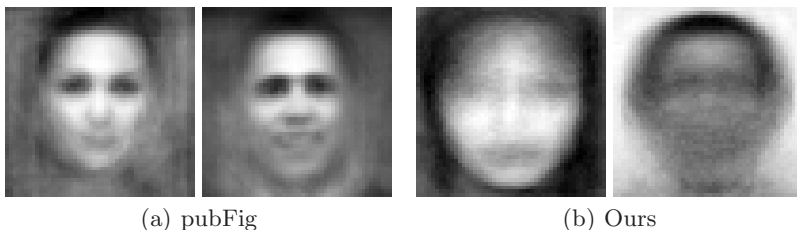
#### 4 Statistical details and Complexity analysis of database

We now provide statistical details of the IIIT-CFW which is collected in totally unconstrained manner. The IIIT-CFW has large variations in attributes. These variations make the database realistic and challenging. This database contains 8928 cartoon images of 100 public figures. The variations in this database across

<sup>2</sup> [http://tapor.uvic.ca/~mholmes/image\\_markup/](http://tapor.uvic.ca/~mholmes/image_markup/)



**Fig. 4.** Wide variations in attributes in our database. We show distribution of face images across eight attributes and observe that our database has lots of diversity across various attributes.

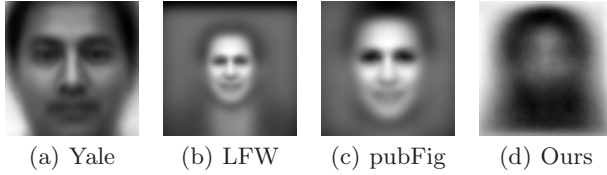


**Fig. 5.** Mean faces of Angelina Jolie (1st and 3rd column) and Barack Obama (2nd and 4th column) in pubFig [18] and ours. We observe that in pubFig [18] celebrities are clearly recognizable but not in ours.

the attributes are shown through pie-charts in Fig. 4. In this figure, we illustrate variations in data across seven attributes (discussed in Section 3) and variations in race. We observe that the introduced database possesses large diversity.

We also analyze the complexity of the database by analyzing the mean face and eigenfaces of the database and compare it with some of the popular face database, namely, Yale [12], LFW [14] and pubFig [18]. These comparisons are shown in Fig 6 and Fig. 7 respectively. Here, we show mean and eigenfaces of Yale database [12] and relatively harder databases such as LFW [14] and pubFig [18]. We observe that as compared to ours, the mean and the top eigenfaces of these databases clearly look like faces. This implies that these popular database have lesser variations in pose and appearance of faces as compared to ours. We also compare celebrity wise mean faces of our database with that of pubFig [18] which also contain faces of public figures. This comparison is shown in Fig. 5. We choose faces of Angelina Jolie and Barack Obama for this study. As can be seen, one can easily recognize the mean faces of these two celebrities in case of pubFig [18]. However, the mean face of these celebrities in our cartoon database is hard to recognize. These observations justify our claim that our database contains cartoon images collected in a totally unconstrained setting and it has large variations.





**Fig. 6.** Mean faces of our cartoon database in (d) as compared to (a) Yale [12] (b) LFW [14] and (c) pubFig [18]. We observe that mean face of Yale database clearly looks like a face as compared to the relatively harder databases such as LFW and pubFig. On the other hand, our database: the IIIT-CFW contains large variations, and hence the mean face does not clearly look like a face.



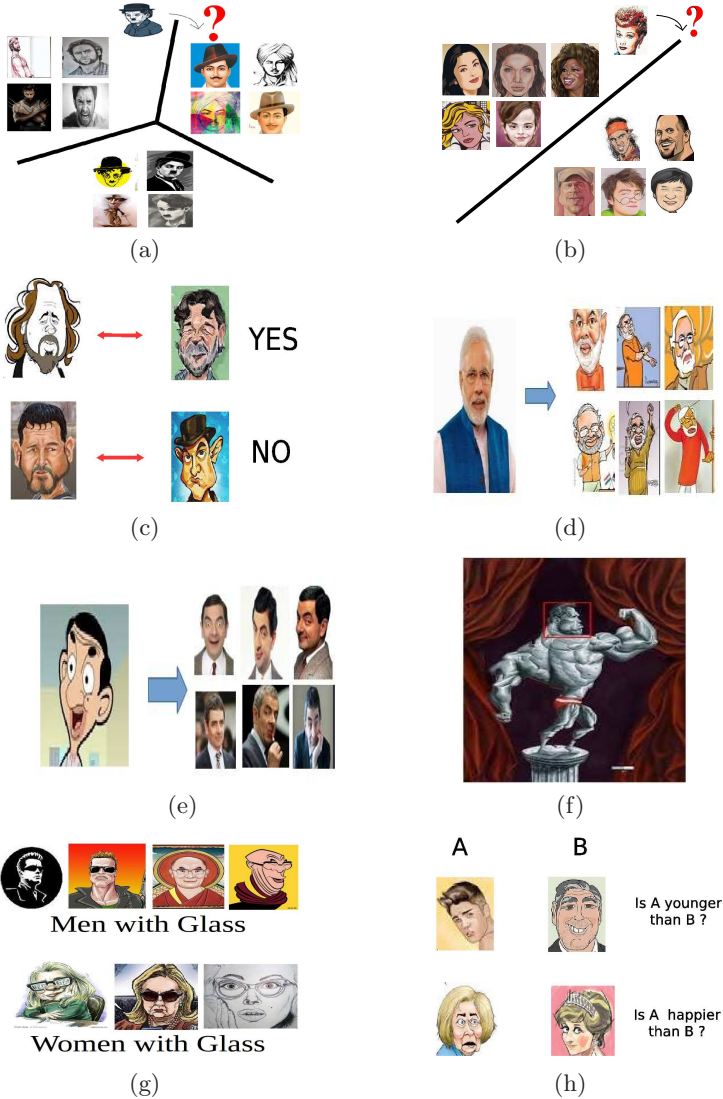
**Fig. 7.** Top-5 eigenfaces of (a) Yale [12] (ii) LFW [14], (c) pubFig [18] and (d) Ours

Problem	Train set	Validation set	Test set	Query set
Face recognition	4464 CF 100 classes	2232 CF 100 classes	2232 CF 100 classes	NA
Gender identification	6007 CF 2 classes	1857 CF 2 classes	1064 2 classes	NA
Face verification	1,80,39,021 CFP 2 classes	17,23,296 CFP 2 classes	5,65,516 CFP 2 classes	NA
Attribute based search	NA	NA	NA	Query: 10 strings Databse: 8928 CF
Photo2cartoon	NA	NA	NA	Query: 1000 RF Databse: 8928 CF
cartoon2photo	NA	NA	NA	Query: 8928 CF Databse: 1000 RF

**Table 2.** Training, testing and validation splits in our database (the CFG) for various problems (CF: cartoon faces, RF: real faces, CFP: cartoon face pairs).

## 5 Spectrum of problems

The IIIT-CFW has been introduced to study a spectrum of problems. These problems are summarized in Fig. 8. Although many of these problems are fun-



**Fig. 8.** Spectrum of problems which can be studied on our database: (a) cartoon face recognition, (b) cartoon gender classification (c) cartoon face verification, (d) photo to cartoon search (e) cartoon to photo search, (f) cartoon face detection (e) attribute based search in cartoons, and (f) relative attributes in cartoons. We provide **suitable annotations** and **problem specific train-val-test split** to conduct study on various problems associated with cartoon faces.

damental in real face domain, but due to the unavailability of detail annotations and problem specific train-test splits existing face databases facilitate research only on selected problems. On the other hand, we not only provide problem specific train-val-test splits, but also detail annotations to study wide range of problems associated with cartoon faces. In this section, we briefly describe the problems which can be studied on our database, and explain the train-val-test split strategy. Table 2 summarizes the training, validation and test sets in our database.

1. **Cartoon face recognition.** The problem of cartoon face recognition is to recognize the given cartoon face as one of the  $C$  classes.
2. **Cartoon face verification.** Given two cartoon faces the problem of face verification is to answer whether the pair is of same person or not. For this problem we have labeled face pairs as same or different based on if they are of same public figure or not. The train-test division is done such that if a pair of cartoon characters  $c$  belong to train set, then it does not belong to validation and test set.
3. **Cartoon gender identification.** It is binary classification problem where given a cartoon face the system has to answer whether it is a male or a female. We provide a standard train and test split by making sure that if a cartoon character is present in train set, examples of this cartoon character will not be present in test set.
4. **photo2cartoon and cartoon2photo.** The problem here is: given real face of a public figure, can we retrieve all the cartoon faces of that public figure from a large database of cartoons, and vice-versa. This database provide support for studying such problems. Along with cartoon faces we provide 1000 real faces (ten real face of each character). These real faces are harvested from Google image search and can be used as query for photo2cartoon retrieval and database for cartoon2photo retrieval. Further for cartoon2photo retrieval all the cartoon in the database should be used as query.
5. **Face detection, pose estimation and landmark detection.** This has been a well studied problem of real face images [25]. On other side, cartoon faces have lots of artistic variations and detecting face, landmark points such as eyes, nose, mouth etc can prove challenging for the state of the art methods. For example: results of face detection using the famous Viola-Jones method [23] is shown in Fig. 9. We see that face detection in cartoon is not a trivial task and needs special attention. Our database can also facilitate research for this problem.
6. **Relative attributes in Cartoon.** Relative attributes in faces has been a well studied problem. Given a pair of cartoon faces ( $A, B$ ) the problem is to answer some questions such as: Is  $A$  older than  $B$ ? Is  $A$  happier than  $B$ ? Are  $A$  and  $B$  of similar age group? We also provide annotation for this problem and a standard train-val-test split similar to face verification problem.
7. **Attribute based cartoon search.** The IIIT-CFW database also provides annotation for attribute based cartoon search, such as search all cartoon faces with glass, search all female cartoon faces who are happy, and so on. We provide 10 such string queries along with associated relevant images.



**Fig. 9.** Results of Viola-Jones [23] face detection on our database: first four images show some successful face detection. The last four images are few examples where Viola-Jones fails to detect faces. In general, we observe that seminal method Viola-Jones is not very successful in detecting cartoon faces.

**Other exciting problems.** Generating photo-realistic images of human faces has been a fascinating yet a hard problem in computer vision and computer graphics. This problem has gained attention of researchers in the last decade, and some interesting works have been published in this space, such as, face hallucination [26], facial expression generation [27]. There have also been some works on face sketches, which are the subset of our database, e.g., face sketch synthesis and recognition [28], sketch inversion [29].

These problems in real faces or sketch faces domain can be also studied for cartoons, and our database can be used for conducting such study. Generating realistic cartoon faces with various facial expressions can be an exciting application in the fine art perspective.

## 6 Conclusions

We have introduced a large database of cartoon faces and discussed the spectrum of problems which can be studied on this database. We believe the database will trigger research in cartoon face understanding and can prove a turning point in the way of how computers see the cartoons.

## References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. *ACM computing surveys (CSUR)* **35**(4) (2003) 399–458
2. Tolba, A., El-Baz, A., El-Harby, A.: Face recognition: A literature review. *International Journal of Signal Processing* **2**(2) (2006) 88–103
3. Jafri, R., Arabnia, H.R.: A survey of face recognition techniques. *JIPS* **5**(2) (2009) 41–68
4. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In: *CVPR*. (2015)
5. Taigman, Y., Yang, M., Ranzato, M., Wolf, L.: Deepface: Closing the gap to human-level performance in face verification. In: *CVPR*. (2014)
6. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep face recognition. In: *BMVC*. (2015)
7. Glasberg, R., Samour, A., Elazouzi, K., Sikora, T.: Cartoon-recognition using video & audio descriptors. In: *European Signal Processing Conference*. (2005)
8. Takayama, K., Johan, H., Nishita, T.: Face detection and face recognition of cartoon characters using feature extraction. In: *Image, Electronics and Visual Computing Workshop*. (2012) 48

9. Field, M., Valentine, S., Linsey, J., Hammond, T.: Sketch recognition algorithms for comparing complex and unpredictable shapes. In: IJCAI. (2011)
10. Hammond, T.A., Logsdon, D., Paulson, B., Johnston, J., Peschel, J.M., Wolin, A., Taelle, P.: A sketch recognition system for recognizing free-hand course of action diagrams. In: IAAI. (2010)
11. Samaria, F.S., Harter, A.C.: Parameterisation of a stochastic model for human face identification. In: WACV. (1994)
12. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **19**(7) (1997) 711–720
13. Jain, V., Mukherjee, A.: The indian face database (2002). URL <http://vis-www.cs.umass.edu/vidit/IndianFaceDatabase>
14. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst (October 2007)
15. Gupta, S., Markey, M.K., Bovik, A.C.: Anthropometric 3d face recognition. *International journal of computer vision* **90**(3) (2010) 331–349
16. Gupta, S., Castleman, K.R., Markey, M.K., Bovik, A.C.: Texas 3d face recognition database. In: *Image Analysis & Interpretation (SSIAI), IEEE* (2010) 97–100
17. Thomaz, C.E., Giraldi, G.A.: A new ranking method for principal components analysis and its application to face image analysis. *Image and Vision Computing* **28**(6) (2010) 902–913
18. Pinto, N., Stone, Z., Zickler, T., Cox, D.: Scaling up biologically-inspired computer vision: A case study in unconstrained face recognition on facebook. In: CVPRW. (2011)
19. Sun, Y., Wang, X., Tang, X.: Hybrid deep learning for face verification. In: ICCV. (2013)
20. Setty, S., et al.: Indian Movie Face Database: A Benchmark for Face Recognition Under Wide Variations. In: NCVPRIPG. (2013)
21. Ng, H.W., Winkler, S.: A data-driven approach to cleaning large face datasets. In: ICIP. (2014)
22. Hasan, M.K., Pal, C.J.: Experiments on visual information extraction with the faces of wikipedia. In: AAAI. (2014)
23. Viola, P.A., Jones, M.J.: Robust real-time face detection. *International Journal of Computer Vision* **57**(2) (2004) 137–154
24. Wolf, L., Hassner, T., Maoz, I.: Face recognition in unconstrained videos with matched background similarity. In: CVPR. (2011)
25. Zhu, X., Ramanan, D.: Face detection, pose estimation, and landmark localization in the wild. In: CVPR. (2012)
26. Liu, C., Shum, H., Freeman, W.T.: Face hallucination: Theory and practice. *International Journal of Computer Vision* **75**(1) (2007) 115–134
27. Susskind, J.M., Anderson, A.K., Hinton, G.E., Movellan, J.R.: Generating facial expressions with deep belief nets. INTECH Open Access Publisher (2008)
28. Wang, X., Tang, X.: Face photo-sketch synthesis and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(11) (2009) 1955–1967
29. Güçlütürk, Y., Güçlü, U., van Lier, R., van Gerven, M.A.: Convolutional sketch inversion. arXiv preprint arXiv:1606.03073 (2016)