

https://dspace.org/

Development of the Wits Face Database: an African database of high-resolution facial photographs and multimodal closedcircuit television (CCTV) recordings [version 1; peer review: 2 approved]

Bacci, Nicholas; Davimes, Joshua; Steyn, Maryna; Briers, Nanette

2021

https://hdl.handle.net/10539/32728

Chwistek M. Recent advances in understanding and managing cancer pain. F1000Res. 2017 Jun 20;6:945. doi: 10.12688/f1000research.10817.1.

Downloaded from a DSpace repository. https://dspace.org/

Check for updates

DATA NOTE

Development of the Wits Face Database: an African database of high-resolution facial photographs and multimodal closed-circuit television (CCTV) recordings [version 1; peer review: 2 approved]

Nicholas Bacci, Joshua Davimes, Maryna Steyn, Nanette Briers

Human Variation and Identification Research Unit (HVIRU), School of Anatomical Sciences, University of the Witwatersrand, Johannesburg, Gauteng, 2193, South Africa

v1 **First published:** 19 Feb 2021, **10**:131
https://doi.org/10.12688/f1000research.50887.1
Latest published: 19 Feb 2021, **10**:131
https://doi.org/10.12688/f1000research.50887.1

Abstract

Forensic facial comparison is a commonly used, yet under-evaluated method employed in medicolegal contexts across the world. Testing the accuracy and reliability of facial comparisons requires large scale controlled and matching facial image databases. Databases that contain images of individuals on closed-circuit television (CCTV), with matching formal and informal photographs are needed for this type of research. Although many databases are available, the majority if not all are developed in order to improve facial recognition and face detection algorithms through machine learning, with very limited if any measure of standardisation. This paper aims to review the available databases and describe the development of a high resolution, standardised facial photograph and CCTV recording database of male Africans. The database is composed of a total of 6220 standardised and uncontrolled suboptimal facial photographs of 622 matching individuals in five different views, as well as corresponding CCTV footage of 334 individuals recorded under different realistic conditions. A detailed description of the composition and acquisition process of the database as well as its subdivisions and possible uses are provided. The challenges and limitations of developing this database are also highlighted, particularly with regard to obtaining CCTV video recordings and ethics for a database of faces. The application process to access the database is also briefly described.

Keywords

face database, CCTV, facial photographs, facial identification, facial comparison, morphological analysis, facial recognition

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
version 1 19 Feb 2021	✓ report	✓ report

1. **Ching-Yiu Jessica Liu**, Liverpool John Moores University, Liverpool, UK

2. **Won-Joon Lee**, National Forensic Service Seoul Institute, Seoul, South Korea

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Nicholas Bacci (nicholasbacci@gmail.com)

Author roles: Bacci N: Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Davimes J**: Data Curation, Investigation, Methodology, Project Administration, Resources, Visualization, Writing – Review & Editing; **Steyn M**: Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing; **Briers N**: Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The research of N. Bacci (Grant No.: 11858) and N. Briers (Grant No.: CSUR160425163022 (UID:106031)) is sponsored by the National Research Foundation of South Africa. Any opinions, findings and conclusions or recommendations expressed in this study are those of the authors and therefore the NRF does not accept any liability in regard thereto. N. Bacci was also partially funded by the J.J.J. Smieszek Fellowship, School of Anatomical Sciences, University of the Witwatersrand.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Bacci N *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Bacci N, Davimes J, Steyn M and Briers N. **Development of the Wits Face Database: an African database of high-resolution facial photographs and multimodal closed-circuit television (CCTV) recordings [version 1; peer review: 2 approved]** F1000Research 2021, **10**:131 https://doi.org/10.12688/f1000research.50887.1

First published: 19 Feb 2021, **10**:131 https://doi.org/10.12688/f1000research.50887.1

Introduction

Facial comparison is utilised by law enforcement to associate two sets of images, captured on video or photographically, to one another. Although different approaches exist, facial comparison by morphological analysis is currently considered the most reliable method[1]. In this method, a target image, such as a snapshot from closed circuit television (CCTV) recordings obtained during criminal activity, and a standardised optimal image, such as a police mugshot, are compared to ascertain whether the two individuals are the same person. Facial comparison by morphological analysis has no directed standardised stepwise procedure as a validated methodology[2]. Although facial comparison still does employ the analysis, comparison, evaluation and verification (ACE-V) tenets of forensic examinations[1,3], its accuracy and reliability has not been tested extensively. As such, assessing its reliability should be considered a priority[4]. The lack of validation can potentially be attributed to the logistic complexity required for rigorous scientific testing, of which a considerable limitation is the lack of standardised and actualistic databases to use[5]. While several facial image databases exist (e.g. 5,6), the incongruity of their composition is a major limiting factor.

The composition of many face image databases tends to be specific to the original intended use with a focus on various controlled conditions which makes them difficult to use for general purposes[6]. For example, there is a variety of pre-landmarked databases available for use in the field of facial recognition (e.g. 7–11) with many variations and controlled conditions. Some of the most commonly controlled-for conditions are orientation of the head/pose, illumination/lighting conditions, facial expressions, and age-related variations (e.g. 6,7,12). Despite the large number of databases available (Table 1), there is a tendency towards either highly controlled data sets captured under very specific conditions with limited actualistic applications (e.g. 13,14) or highly randomised images collected under inconsistent conditions (e.g. 7–9,15). In addition, most of these databases include a limited number of subjects with many replications under niche conditions and no standard baseline control images. Some of the datasets with a limited number of unique individuals include under 100 subjects (e.g. 13,16,17), with a handful including only 10 to 15 subjects (e.g. 14,18–20). Lastly, many of these databases are, by today's standards, of subpar resolution with only two databases including images of resolutions greater than 640 x 480 pixels[21,22]. Between the highly specialised and varied conditions of capture, the lack of controlled images with matching realistic informal photographs of the same subjects, the low resolution, lack of methodological standardisation in image capture, and limited subject numbers, these facial image datasets provide very limited use in a forensic facial comparison context or more generalised facial studies.

Recently, databases consisting of video recordings of faces have become more common (e.g. 21,23–26) due to the increase in CCTV surveillance cameras. Generally, these databases have either been collated *via* recording of known participants in controlled environments (e.g. 21,23,26), or were based on pre-recorded videos obtained from various media on the internet or movies[24,25]. To the best of the authors' knowledge, only a single other database (SCface – Surveillance Cameras Face Database) published includes both facial photographs of high resolution and corresponding still images extracted from CCTV recordings of varying resolutions[21]. These databases are primarily used in testing and developing head/face detection and tracking or automated face annotation systems under a variety of video recording conditions.

The majority of facial image databases are primarily or exclusively inclusive of males of either European or East Asian descent (e.g. 17,27,28). Only a few databases contain individuals from other ancestry groups, particularly African individuals (e.g. 6,7,13), and they have a limited number of individuals. This is evident when considering there is only a single other South African face database[29]. Although a great initiative for a large, landmarked face database specifically developed to increase the variety of lighting, ethnicity and age available, it consists primarily of low-resolution webcam-based images with often very distorted lighting conditions. In addition, a total of 276 subjects were used for this database with no control images or demographic specifications provided.

The majority of existing facial databases have been developed for the purposes of facial recognition and machine learning training and do not contain target and control images of the same individual. These databases tend to be sourced from public, internet images containing faces in a variety of inconsistent conditions (e.g. 9,11,12). To the best of the authors' knowledge, no databases exist that are intended specifically for use with facial comparison by morphological analysis. As such, we aimed at creating a system for developing a consistent face database with corresponding individuals across various photographic and video recording conditions, resulting in the Wits Face Database (WFD). This database is intended to be a functional, actualistic African database of facial images that can be utilised for facial comparison analyses and research in craniofacial identification. This database is intended to be a free resource strictly for non-commercial scientific research, provided access has been cleared by an ethics committee overseeing its use.

Materials and methods

The database was established by collecting photographs of willing participants on the University of the Witwatersrand campus, Johannesburg, South Africa, including corresponding CCTV recordings. The database is comprised of CCTV recordings and photographs gathered on university premises, between July 2018 and October 2019, *via* the pre-existing CCTV systems used by campus security. Facial photographs were standardised and in five different views. The CCTV recordings were captured in a variety of conditions, such as from different quality cameras, in different formats and heights of recording, and with disguises (sunglasses and caps).

Table 1. Overview of available face databases with available descriptives.

Database name and reference	No. of unique individuals	No. of images	Image resolution (pixels)	Database description and condition variations
AR Face Database[30]	126 (70 Male, 56 Female)	4000	576 x 768	Various facial expressions, lighting, glasses, scarf
CVL Database[31]	114	798	640 x 480	Various poses, varying facial expression
FERET Database[32]	1199	14051	256 x 384	Various slight facial expressions, poses
Labelled Faces in the Wild (LFW) Database[7]	5749	13233	250 x 250	Landmarked faces in various poses, expressions, lighting, ethnicity, age, clothing, hairstyles
Face Recognition Grand Challenge (FRGC) Database[33]	688	Undefined	"High resolution"	Various facial expression, lighting
CAS-PEAL Face Database[34]	1040 (595Male, 445 Female)	30900	360 x 480	Various poses, facial expressions, changes in lighting, glasses, caps
The MUCT Face Database[29]	276	3755	640 x 480	Variations in pose, lighting, and annotated faces
The Yale Face Database[18]	15	165	320 x 243	Facial expression variations, glasses
The Yale Face Database B[19]	10	5760	640 x 480	Various poses and changes in lighting
CMU Pose, Illumination, and Expression PIE Database[10]	68	41368	640 x 486	Various poses, facial expressions, changes in lighting, and glasses
Olivetti – Att – ORL[35]	40	400	92 x 112	None
Japanese Female Facial Expression (JAFFE) Database[20]	10	70	256 x 256	Various facial expressions
FIDENTIS 3D Face Database[36]	2476	2476 complete 3D scans	12 megapixels	3D scans, landmarked, ear to ear facial pose equivalent
Caltech Occluded Face in the Wild (COFW)[12]	Undisclosed	1852	Undisclosed	Various poses, expressions, lighting, occlusion focus, annotated
Ibug 300 Faces In-the-Wild (ibug 300W) Challenge database[9]	600	>4000	Undisclosed	Various poses, expressions, lighting, annotated
Labeled Face Parts in the Wild (LFPW) Dataset[11]	3000	3000	Undisclosed	Various poses, expressions, lighting
Quality labeled faces in the wild (QLFW) database[8]	5749	277809	250 x 250	Various poses, expressions, lighting, ethnicity, age, clothing, hairstyles, distortions
Helen dataset[37]	Undisclosed	2330	>500 width	Various poses, expressions, lighting, annotated
Facial expressions of emotion (KDEF) database[17]	70 (35 Male, 35 Female)	4900	Undisclosed	Various facial expressions
NimStim facial expression database[13]	43	672	Undisclosed	Various facial expressions

F1000Research 2021, 10:131 Last updated: 03 SEP 2021

Database name and reference	No. of unique individuals	No. of images	Image resolution (pixels)	Database description and condition variations
Annotated Facial Landmarks in the Wild (AFLW) database[15]	25993 (11437 Male, 14556 Female)	21997	Undisclosed	Various poses, expressions, lighting, ethnicity, age, clothing, and hairstyles
Pointing Head Pose Image Database[14]	15	2790	384 x 288	Various poses, glasses
BioID Database[38]	23	1521	382 x 288	Various facial expressions, poses, lighting, accessories
University of Olulu Physics-Based Face Database[39]	111	2112	428 x 569	Various poses (minor), lighting, glasses
Chicago Face Database[22]	158 (73 Male, 85 Female)	Undisclosed	2444 x 1718	Various poses, facial expressions, lighting
SCface – Surveillance Cameras Face Database[21]	130 (114 Male, 16 Female)	4160	Varied from 3072 x 2048 to 224 x 168	Various poses, normal and infrared
M2VTS multimodal face database[16]	37	Minimum of 185	286 x 350	Videos and video to still images, various poses, some with and without eyeglasses, various time intervals (weeks)

F1000Research 2021, 10:131 Last updated: 03 SEP 2021

Ethics and participant recruitment

Ethical approval was obtained from the Human Research Ethics Committee (Medical) of the University of the Witwatersrand (clearance certificate No.: M171026). Permits from the campus head of security and deputy registrar were obtained prior to data capture, in their capacity as site managers. Facial photographs were captured by an experienced photographer following recruitment of participants. CCTV recordings of matched participants were collected through the university's security systems. The original target of participants intended for the database were 600. However, due to data loss as a result of power failures and data transfer corruption a greater number of participants had to be recruited for redundancy purposes. Potential participants were identified and recruited near the data collection sites based on facial anthropological features resembling males of South African descent and being older than 18 years of age. Approached participants were then informed of the greater project orally and with an information sheet[40] and asked for voluntary participation in the study and database. Once agreed, they signed an informed consent form[40] prior to being photographed or recorded. If participants requested to be removed from the database, all their photographs and recordings were erased and their recruitment information shredded. None of the personal details and images given by the participants are to be freely distributed. As agreed with the ethics committee, and according to the consent forms signed by the participants, no identifiable images of any individuals are to be published without following up with participants for additional consent. As a result, for publication purposes to indicate examples of the face database, one of the author's images have been utilised.

Image acquisition

Facial photographs and recordings were collected at three designated access-controlled locations with a large influx of potential participants on the Braamfontein Campus of the University of the Witwatersrand, Johannesburg, South Africa. The following three collection sites were set-up and utilised:

- Site A: outdoors CCTV camera (installation height: 3100 mm) with a view of a student card terminal near the food concourse and Student Union Building, the Matrix.
- Site B: outdoors CCTV camera positioned at eye-level height (1700 mm) in one of the more frequented pedestrian entrances in proximity to the Oppenheimer Life Sciences building.
- Site C: indoors analogue CCTV camera (installation height: 2500 mm) in the administration building concourse in view of the cashier's offices, Solomon Mahlangu House.

In the vicinity of the installed CCTV cameras at each site, a photography station was set up in a standard manner as demonstrated in Figure 1 and Figure 2. Within this set-up, facial photographs were captured in five different views using two different cameras with slightly different parameters and conditions. All cameras were arranged in a marked fixed location near the CCTV field of view on tripods at a fixed height of 1600 mm. This height was maintained to attempt centring the field of view of the photographs on the face, as the mean height for black South African males[41] is 1710 mm. Specifically, the eye level of each participant was composed on the top horizontal rule of thirds line for all photographs taken. All photographs were taken in portrait orientation.

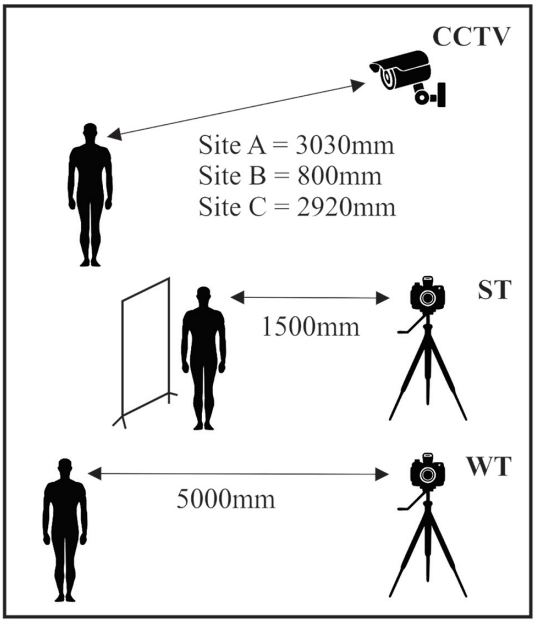


Figure 1. Schematic diagram of camera set-up for closed-circuit television (CCTV) and facial photograph capture. ST = standardized photograph; WT = wildtype photograph.

An example of the arrangements for database image capture and recording on pre-existing CCTV systems is shown in Figure 2. Distances were controlled for participants to be photographed and recorded at each site. Standardised (ST) photographs, with a solid black backdrop and participant clothing covered, were captured at a distance of 1500 mm. Wildtype (WT) photographs were captured at a 5000 mm distance. These photographs included a mixed background that was intentionally meant to simulate real-life photographic conditions. WT photographs were taken in a simulated scenario of suboptimal facial images with a comparable quality camera and facial poses. The background was purposefully not controlled for with a mixed environment visible and varied based on the location site of data collection. This background was intentionally out of focus to simulate suboptimal photographic conditions. Despite a minor level of variation, consistency was maintained across all photographs with regard to distance to subject, aperture, and composition.

The first set of photographs were captured under standardised conditions using a Canon 1300D 18MP DSLR camera (18–55 mm DC Canon lens) with the following settings: image



Figure 2. Actual photographic camera set-up in the process of database development. Arrangement of cameras and backdrop for standardised and wildtype photographs at site A (**a**), site B (**b**) and site C (**c**).

sensor sensitivity (ISO) of 800, aperture F/9, shutter speed between 1/125 and 1/40, focal length of 55 mm and daylight white balance. For these standard photographs, the objective to face distance was fixed at 1500 mm. A set of standardised photographs were taken, with a black backdrop and with the participants' clothing covered by a black velvet cloth – similar to the backdrop – in order to prevent matching participants based on clothing appearance. These standard photographs were captured in the following five views (Figure 3):

a. Anterior frontal view

b. Right 45-degree view

c. Right lateral view

d. Left 45-degree view

e. Left lateral view

The second set of WT photographs captured for each participant, corresponded to the same views as described above, using a Sony SLT A57 (18 – 250 mm Sony Zoom Lens) with the following settings: image sensor sensitivity (ISO) of 200, aperture F/9, shutter speed between 1/125 and 1/40, focal length of 250 mm and daylight white balance. Indoor photographs at site C were captured at a range of ISOs between 400 and 1600 depending on the varying light conditions.

All photographs were captured as both native .jpeg and RAW format. Images were then downloaded from the SD cards of each camera, stored, and sorted on a desktop computer. All images were then batch cropped at a 4x5 (8x10) aspect ratio, using Adobe Photoshop CS6, to only include the participants as centred in the frame of the photograph. The resulting image resolution for the ST photographs was 3456 x 4320 pixels at 300 dpi and sRGB colour representation. The WT



Figure 3. Example of the five views of standardised (left) and wildtype (right) facial photographs captured. The five views of facial photographs are demonstrated in this image, showing anterior (**a**), right 45-degree (**b**), right lateral (**c**), left 45-degree (**d**) and left lateral (**e**) views.

photographs' resolution was 3264 x 4080 at 350 dpi and with sRGB colour representation. Following batch cropping, the standardised images were imported into Adobe Lightroom (v. 5.3) for basic editing. The only adjustments made included exposure level, highlights, and shadow correction. Additionally, removal/spot healing of any exposed clothing or background features were done if the cloth or backdrop did not fully obscure the participant and surroundings. The wildtype images were left unaltered post cropping. The above discussed image processing (batch cropping) and adjustments (exposure level, highlights, shadows correction and spot healing) can alternatively be performed using open source software such as GNU Image Manipulation Program (GIMP), Photivo, and darktable.

CCTV recordings from internet protocol (IP) cameras were transferred live to a HikVision (model: DS-9664NI-I16) server for storage. Videos from the analogue camera were stored on a DS-ENC-V120B20121026-7054D2ABA6FB digital video

recorder (DVR) device. All recordings were then extracted in .mp4 format from the University of the Witwatersrand's CCTV systems through the university's protection services software (iVMS 4200 v. 2.71.9). Only footage recorded during data collection times was extracted from the CCTV systems. During recording, participants were asked to stand at a marked location and rotate through the same five views that the photographs were captured in for approximately 5 seconds. Following this, the participant acted out the process of utilising a student card terminal in view of the camera (sites A and C) or attempting to exit the campus (site B).

The outdoor CCTV camera at site A was an IP camera (HikVision, model: DS-2CD2142FWD-I, 4-megapixel, 4 mm fixed lens, aperture F/2). This camera was installed at a height of 3100 mm and a floor distance of 2690 mm from the marked location where participants stood in the process of video capturing. The resulting distance between the participants' face and the IP camera objective was 3030 mm (Figure 4). This distance was calculated based on the mean height of the average black South African male of 1710 mm[41]. The angle of incidence from the camera objective to the face was approximately 27°.

The outdoor "eye-level" IP CCTV camera at site B had the same model and specifications of the camera at site A and was installed at a height of 1700 mm from the ground, marked location to floor distance of 800 mm. The resulting distance from target face to camera was 800 mm and the angle of incidence was virtually zero.



Figure 4. Demonstration of estimated objective to face distance at site A. Vertical distance indicates the mean height of a South African male (1710 mm) and the oblique distance indicates the calculated approxiamte camera to face distance at Site A (3030 mm).

The indoor analogue CCTV camera at site C (Securi-Prod 1/3" Sony Effio E 700TVL indoor dome, model: CC217, 2.8 – 12 mm vari-focal lens) was installed at a height of 2500 mm and at a floor distance from the marked area of 2810 mm. The resulting distance between camera and face was 2920 mm and the angle of incidence was approximately 22°.

A total of approximately 30 seconds of footage was recorded for each participant and out of the 30 seconds five still images were captured from the footage at each of the five views previously described (Figure 5, Figure 6, Figure 7 and Figure 8). The photographs and recordings of each individual were



Figure 5. Example of the five views of standard closed-circuit television (CCTV) stills captured at site A. The five views of facial photographs are demonstrated in this image, showing anterior (**a**), right 45-degree (**b**), right lateral (**c**), left 45-degree (**d**) and left lateral (**e**) views.

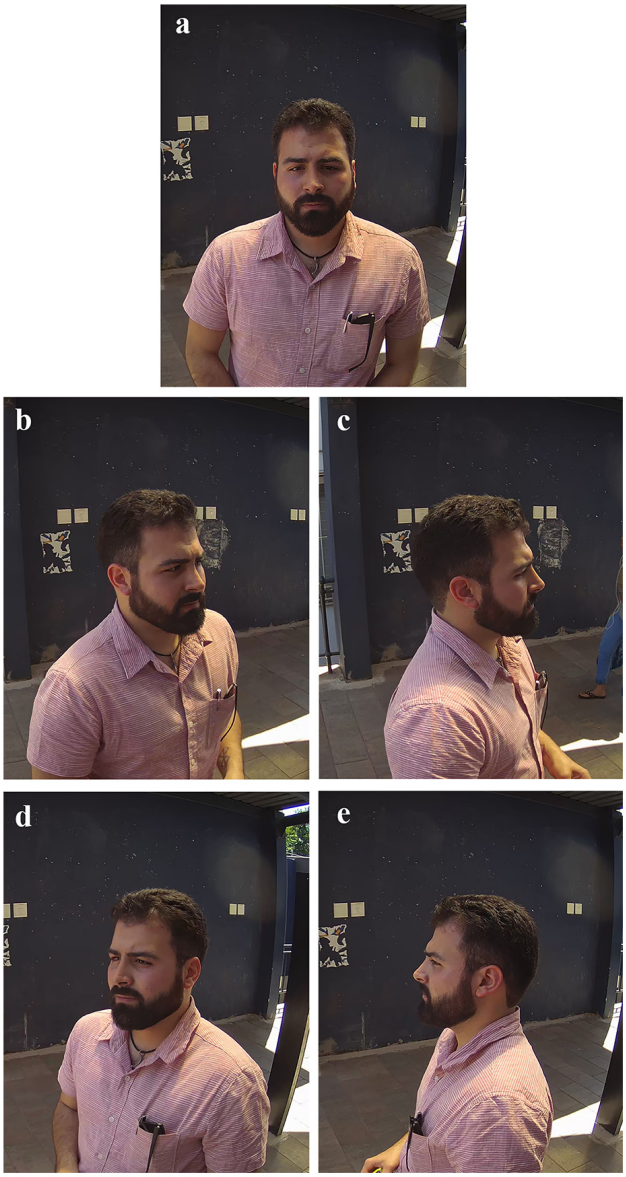


Figure 6. Example of the five views of eye level closed-circuit television (CCTV) stills captured at site B. The five views of facial photographs are demonstrated in this image, showcasing anterior (**a**), right 45-degree (**b**), right lateral (**c**), left 45-degree (**d**) and left lateral (**e**) views.



Figure 7. Example of the five views of analogue closed-circuit television (CCTV) stills captured at site C. The five views of facial photographs are demonstrated in this image, showing anterior (**a**), right 45-degree (**b**), right lateral (**c**), left 45-degree (**d**) and left lateral (**e**) views.

coded with a participant number to maintain anonymity of the participants and a separate record of the identity of the participants was retained. Complete anonymity of the participants included in the facial image database was maintained. Images and videos included in the database were stored on three separate password protected desktop computers and encrypted on a constantly monitored external hard disk drive. In addition, the dataset was also transferred to the University of the Witwatersrand's Library repository under restricted access to data management services. This repository retains the database in triplicate for cataloguing and in order to allocate searchable metadata to it to facilitate use.

Database composition

The database was composed strictly of male South Africans of African descent. Participants were over the age of 18, for consent purposes, and in the young adult age-ranges, between 18 and 35 years of age. The participants were subdivided

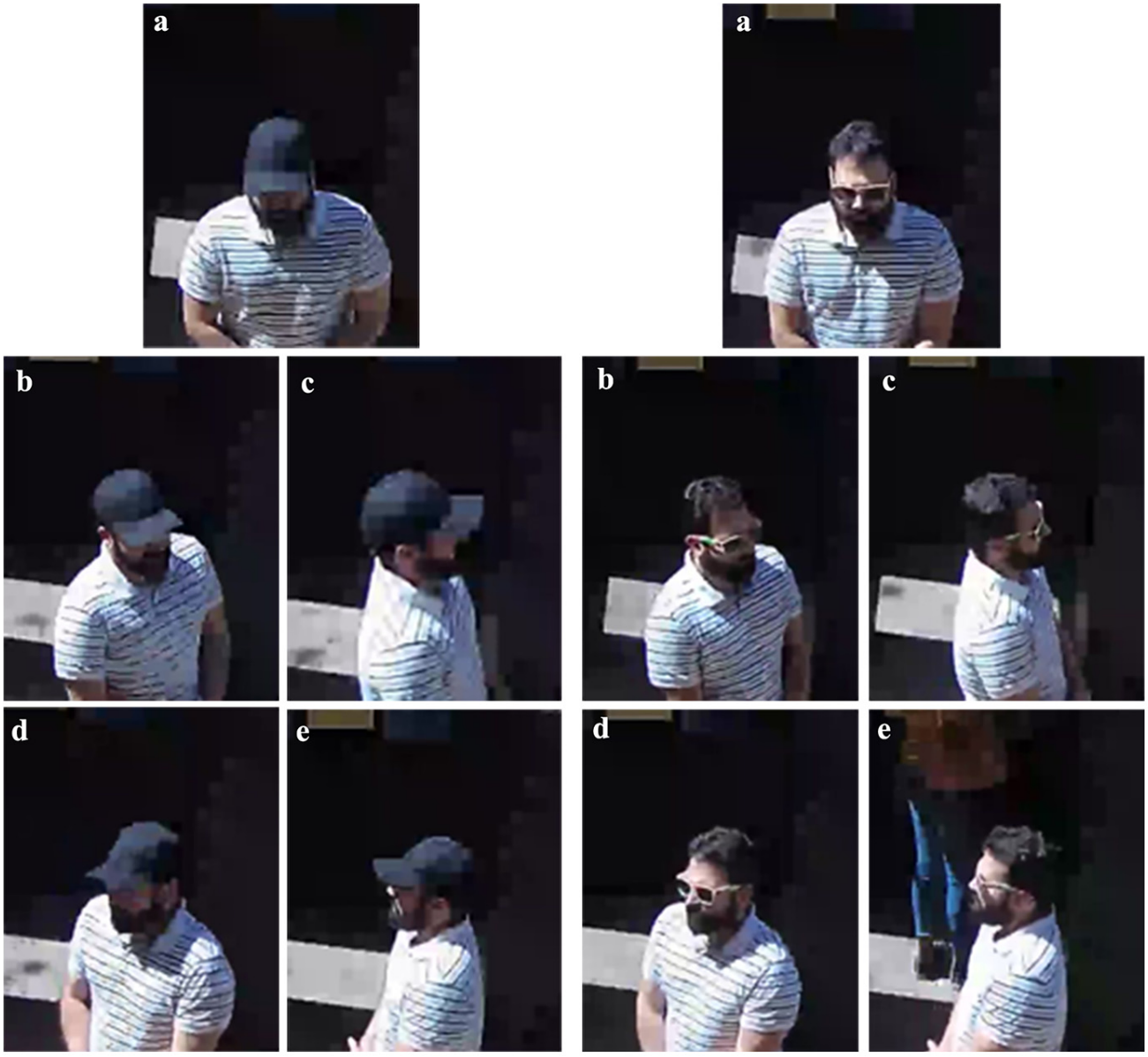


Figure 8. Example of the five views of closed-circuit television (CCTV) recordings with two obstruction types - brimmed cap (left) and sunglasses (right). The five views of facial photographs are demonstrated in this image, showing anterior (**a**), right 45-degree (**b**), right lateral (**c**), left 45-degree (**d**) and left lateral (**e**) views.

into a series of cohorts depending on the type of matched analysis possible, as outlined in Table 2. All participants were photographed in both the standardised and wildtype setting at the various sites. The first two groups were only photographed under natural outdoor lighting at site A (n=120) as well as with artificial indoor fluorescent lighting at site C (n=99). For each participant 10 photographs were captured, totalling 1200 photographs under outdoor conditions with natural lighting conditions and 990 photographs under indoor fluorescent lighting conditions. A second group, in addition to being photographed as above, was recorded on an outdoors setting CCTV camera at site A (n=98, n=86 with corresponding footage). A third group was recorded at site B with the eye-level camera (n=108, n=76 with corresponding footage). A fourth cohort was recorded at site C with the analogue CCTV camera and lastly a final group of participants were recorded with obstructive accessories, namely caps (n=45, n=34 with corresponding footage) or sunglasses (n=41, n=31 with corresponding footage),

Table 2. Detailed categorisation of Wits Face Database (WFD) composition by cohorts of potential analyses.

Cohort of matching individuals	No. of participants	No. of photographs	Participants with corresponding CCTV footage
Photo-Photo Outdoor (Site A)	120	1200	0
Photo-Photo Indoor (Site C)	99	990	0
Photo-Photo Totals	**219**	**2190**	**0**
Photo-CCTV Outdoor (Site A)	98	980	89
Photo-CCTV Outdoor Eye Level (Site B)	108	1080	76
Photo-CCTV Indoor Analogue (Site C)	111	1110	107
Photo-CCTV Outdoor with Cap (Site A)	45	520	34
Photo-CCTV Outdoor with Sunglasses (Site A)	41	490	31
Photo-CCTV Totals	**403**	**4030**	**334**
Grand Totals	**622**	**6220**	**334**

CCTV= closed-circuit television.

using the same IP camera at site A. Due to data corruption or data loss in the CCTV recording process, not all photographed individuals have corresponding footage associated to them. Overall, the database is inclusive of over 6200 facial photographs and 334 corresponding video recordings from various types of CCTV cameras within an African sample (Table 2).

Database utility and applications

A database of this scale can be utilised in a variety of training and research applications, particularly when considering this as the first database of facial images with such a large complement of facial photographs of African individuals. Among a variety of possible applications, the primary intended use is testing various methodologies and conditions for forensic facial comparison. Furthermore, this database can be utilised as a prime tool for training facial comparison experts to develop a high rate of competency and proficiency. Having an appropriate level of training is a crucial aspect of the judicial process[1,4]. In addition, the faces in this database could be used to generate stimuli for studies in psychology and marketing sciences relating to facial recognition. Likewise, the images in this database could be modified as required on an *ad hoc* basis and implemented to train and develop future machine learning and artificial intelligence systems for the purpose of facial recognition in a forensic context. Although a potentially laborious process, standardised landmarks could be added to allow for an estimated dimension calculation of the various facial proportions and features as required by facial recognition systems. However, the database is only available for *bona fide* researchers affiliated with academic research institutions and not for commercial use. An example use the dataset was designed for was validating the use of morphological analysis in forensic facial comparison across a photographic and CCTV sample[27]. This was achieved by sub-setting selected photographs and CCTV recording stills from the database into facial image pools that were independently analysed[27] following the Facial Identification Scientific Working Group morphological analysis feature list[42].

Challenges and limitations

Developing a database of facial photographs and surveillance footage is a logistically complex and tedious process. It can be difficult to recruit volunteers even in highly trafficked areas in a major university (the University of the Witwatersrand has over 39,500 students across five campuses), although people on a campus tend to feel more secure and therefore more willing to participate in this type of study. The recruitment and photography process require a large amount of manpower with a minimum of three to four volunteers needed per day as assistants for an efficient image acquisition process. The location management and site selection are also a highly demanding task as one needs to limit variations between recorded images as much as possible. This entire process needs to be carried out while still collecting images that are representative of a somewhat realistic scenario. The variations in conditions of recordings vary based on camera type, quality, and installation. The majority of CCTV recordings, for example, were collected in an outdoors area with sunny daylight conditions which is by design an actualistic sample, although inconsistent due to weather and lighting conditions changing throughout a given day.

Similarly, the photography collection was affected by uncontrollable varying lighting conditions for both the indoor and outdoor settings. These include dim lighting indoors leading

to lower quality, higher noise images, and dappled light and sun position outdoors leading to highlights, shadow, and contrast artefacts. The objective to face distances were controlled as much as possible, without compromising the actualistic nature of the data, in order to minimise perspective distortion due to the nature of relaying a three-dimensional scene/object into a two-dimensional medium of photography[43]. Despite the intent, at greater distances the focal plane of a camera lens requires adjusting and can result in lower clarity images[43], as is evident in the varied focal lengths of the CCTV cameras.

Furthermore, the nature of CCTV cameras depends on the entirety of the surveillance system installation. A variety of complications, including inconsistent IP camera network connectivity and power outages caused occasional complete data loss or footage corruption in the recording process, resulting in the overall reduced numbers of corresponding recordings. This data loss was particularly evident in the IP cameras, as in the particular set-up used by the university, they do not locally store footage but transfer it immediately to a central server. In the process, any interruption or fluctuation in local area network speeds or connectivity would result in data loss or corruption. Even though analogue cameras by default record at lower resolutions, the immediate local storage on a DVR device resulted in reduced data loss and no data corruption. In fact, a total of 710 participants were originally recruited and photographed and/or recorded. Following data loss, participants that requested to be excluded from the database and data corrupted during transfer from the CCTV cameras to the servers, only 622 of the 710 could be included in the database.

Male individuals were selected specifically as males are more commonly involved in criminal activity, both as victims and perpetrators, in both developed and developing nations[44,45]. However, other demographic factors such as age and ancestry were not strictly controlled, and sample composition resulted varied across all cohorts. This lack of strict control was primarily due to the sensitivity of labelling groups of individuals as belonging to specific descent groups, making requesting this information from participants one of the ethical limitations considered originally.

The intention of the authors is for this database to be further expanded to include female individuals, additional disguises in the form of make-up and face masks, as well as more variations of CCTV camera recordings, such as infrared night vision CCTV, to provide a better and more varied selection for its applications.

Ethics of face databases

Overall, face databases are quite common with 27 photographic image databases available for use in the fields of head and face detection, tracking and recognition. These databases are usually created with specific criteria, such as particular facial expressions and lighting condition variations for testing a specific aspect of facial recognition. Amongst the 27 databases outlined in Table 1, seven have been collected from various media of personal photographs available on the internet[7,9,11,12,15,24,37], which is a legal yet perhaps ethically questionable practice, as the individuals included in these face databases have not provided consent for that inclusion. This is particularly common in the most recently developed databases (e.g. 7,8,15), most likely due to the broad availability of facial images and the highly efficient and accurate search engines. This practice can result in privacy-intrusive practices of freely available and commercialised application of facial recognition[28], which is an ever-increasing concern for the public with specific regard to privacy of data and images[46]. The database compiled here is an example of a complex yet more ethical approach that limits commercialised application and potential research.

Data availability

Underlying data

The WFD is stored on the Wits Institutional Repository environment on DSpace (WIReDSpace) and published under the following unique identifier: http://doi.org/10.17605/OSF.IO/WMA4C (this registration also contains the PhD study protocol that led to the development of the WFD and an addendum to the protocol registration highlighting the major changes to the methodological approach of the original protocol). A sample of the dataset is freely accessible at https://hdl.handle.net/10539/29924.

The database is an open access resource for use in strictly non-commercial research. In order to access the WFD, prospective users will have to apply for access to the Institutional Review Board overseeing ethical and scientific use of the database in order to safeguard the privacy and decency of the database's participants. Once approved a researcher may use the database free of charge. Database access is restricted and limited to following the above-mentioned procedure, due to the nature of the data including potentially identifying information (facial physiognomy) of participants. In addition, strict limitations were imposed by the Human Research Ethics Committee (Medical) as well as the consent permissions agreed upon with the participants, which assign responsibility to the School of Anatomical Sciences to review access applications in ethical and scientific merit in order to exclusively conduct research. The access procedures and limitations are governed by a legally binding Conditions of Use document available on https://hdl.handle.net/10539/29924[40] in conjunction to the freely accessible sample. Data will be made available to successful applicants under a temporary restricted licence guided by the aforementioned conditions of use document.

Extended data

Open Science Framework: Wits Face Database: Description. https://doi.org/10.17605/OSF.IO/Q8V2R[40]

This project contains the following extended data:

- Participant information sheet
- Participant consent form
- Conditions of use

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

Authors' contributions

Conceptualization: Nicholas Bacci, Maryna Steyn, Nanette Briers; Data curation: Nicholas Bacci, Joshua Davimes; Methodology: Nicholas Bacci, Joshua Davimes, Maryna Steyn, Nanette Briers; Investigation: Nicholas Bacci, Joshua Davimes; Project Administration: Nicholas Bacci, Joshua Davimes; Resources: Nicholas Bacci, Joshua Davimes; Visualisation: Nicholas Bacci, Joshua Davimes; Writing - original draft preparation: Nicholas Bacci; Writing - review and editing: Nicholas Bacci, Joshua Davimes, Maryna Steyn, Nanette Briers; Funding acquisition: Nicholas Bacci, Maryna Steyn, Nanette Briers; Supervision: Maryna Steyn and Nanette Briers.

Acknowledgements

Thanks are due to all the participants who agreed to be photographed for the development of this database. Particular recognition is due to all the volunteers who assisted in participant recruitment for this study: Jesse Fredericks, Kiveshen Pillay, Rethabile Masiu, Sameerah Sallie, Daniel Munesamy, Laurette Joubert, Jordan Swiegers, Betty Mkabela, Johannes P. Meyer, Amy Spies, Natasha Loubser, Nicole Virgili, Dan-Joel Lukumbi, Tamara Lottering, Mathabatha Ntjie, Claudia Landsman, Raheema Dalika, Merete Goosen, Stephanie Souris, Rabelani Negota, Mahlatse Mahasha, Jessica Manavhela. Special thanks are due to Gideon LeRoux for his assistance in the capturing and extracting of the CCTV recordings from the University security systems. A great deal of thanks are also due to Nina Lewin who aided in policy development and establishing the database repository. All named persons above have granted permission to be included within the acknowledgements.

References

1. Facial Identification Scientific Working Group: **Facial Comparison Overview and Methodology Guidelines.** 2019.
Reference Source
2. Dror IE, Charlton D, Péron AE: **Contextual information renders experts vulnerable to making erroneous identifications.** *Forensic Sci Int.* 2006; **156**(1): 74–8.
PubMed Abstract | Publisher Full Text
3. Speckeis C: **Can ACE-V be validated?** *J Forensic Identif.* 2011; **61**(3): 201–9.
Reference Source
4. Steyn M, Pretorius M, Briers N, *et al.*: **Forensic facial comparison in South Africa: State of the science.** *Forensic Sci Int.* 2018; **287**: 190–4.
PubMed Abstract | Publisher Full Text
5. Valentine T, Davis JP: **Forensic Facial Identification.** *Forensic Facial Identification.* 2015; 1–347.
6. Gross R: **Face Databases.** *Handbook of Face Recognition*. New York: Springer-Verlag; 2005; 301–27.
Publisher Full Text
7. Huang GB, Ramesh M, Berg T, *et al.*: **Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments.** October, University of Massachusetts, Amherst Technical Report. 2007.
Reference Source
8. Karam LJ, Zhu T: **Quality labeled faces in the wild (QLFW): a database for studying face recognition in real-world environments.** *Hum Vis Electron Imaging XX.* 2015; **9394**: 93940B.
Publisher Full Text
9. Sagonas C, Antonakos E, Tzimiropoulos G, *et al.*: **300 Faces In-The-Wild Challenge: database and results.** *Image Vis Comput.* 2015; **47**: 3–18.
Publisher Full Text
10. Sim T, Baker S, Bsat M: **The CMU Pose, Illumination, and Expression (PIE) database.** In: *IEEE Trans Pattern Anal Mach Intell.* 2003; 1615–8.
Publisher Full Text
11. Belhumeur PN, Jacobs DW, Kriegman DJ, *et al.*: **Localizing parts of faces using a consensus of exemplars.** *IEEE Trans Pattern Anal Mach Intell.* 2013; **35**(12): 2930–40.
Publisher Full Text
12. Burgos-Artizzu XP, Perona P, Dollar P: **Robust face landmark estimation under occlusion.** *Proc IEEE Int Conf Comput Vis.* 2013; 1513–20.
Publisher Full Text
13. Tottenham N, Tanaka JW, Leon AC, *et al.*: **The NimStim set of facial expressions: Judgments from untrained research participants.** *Psychiatry Res.* 2009; **168**(3): 242–9.
PubMed Abstract | Publisher Full Text | Free Full Text
14. Gourier N, Hall D, Crowley JL: **Estimating face orientation from robust detection of salient facial structures.** *International Workshop on Visual Observation of Deicitic Gestures.* 2004; 17–25.
Reference Source
15. Kostinger M, Wohlhart P, Roth PM, *et al.*: **Annotated Facial Landmarks in the Wild: a large-scale, real-world database for facial landmark localization.** *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops).* IEEE; 2011; 2144–51.
Publisher Full Text
16. Pigeon S, Vandendorpe L: **The M2VTS Multimodal Face Database.** *First Int Conf Audio-and Video-based Biometric Pers Authentication.* 1997; 403–9.
17. Calvo MG, Lundqvist D: **Facial expressions of emotion (KDEF): Identification under different display-duration conditions.** *Behav Res Methods.* 2008; **40**(1): 109–15.
PubMed Abstract | Publisher Full Text
18. Belhumeur PN, Kriegman DJ: **Eigenfaces vs. Fisherfaces : Recognition Using Class Specific Linear Projection.** *IEEE Trans Pattern Anal Mach Intell.* 1997; 711–20.
Publisher Full Text
19. Georghiades AS, Belhumeur PN, Kriegman DJ: **From few to many: Illumination cone models for face recognition under variable lighting and pose.** *IEEE Trans Pattern Anal Mach Intell.* 2001; **23**(6): 643–60.
Publisher Full Text
20. Lyons M, Akamatsu S, Kamachi M, *et al.*: **Coding facial expressions with Gabor wavelets.** *Proc - 3rd IEEE Int Conf Autom Face Gesture Recognition, FG 1998.* 1998; 200–5.
Publisher Full Text
21. Grgic M, Delac K, Grgic S: **SCface - Surveillance cameras face database.** *Multimed Tools Appl.* 2011; **51**(3): 863–79.
Publisher Full Text
22. Ma DS, Correll J, Wittenbrink B: **The Chicago face database: A free stimulus set of faces and norming data.** *Behav Res Methods.* 2015; **47**(4): 1122–35.
PubMed Abstract | Publisher Full Text
23. La Cascia M, Sclaroff S, Athitsos V: **Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3D models.** In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on.* 2002; **22**(4): 322–36.
Publisher Full Text
24. Dhall A, Dhall A, Member S, *et al.*: **Collecting Large, Richly Annotated Facial-Expression Databases from Movies.** *J LaTeX Cl Files.* 2007; **6**(1): 1–14.
Reference Source
25. Shen J, Zafeiriou S, Chrysos GG, *et al.*: **The First Facial Landmark Tracking in-The-Wild Challenge: Benchmark and Results.** In: *IEEE International Conference on Computer Vision.* 2016; 1003–11.
Publisher Full Text
26. Ariz M, Bengoechea JJ, Villanueva A, *et al.*: **A novel 2D/3D database with automatic face annotation for head tracking and pose estimation.** *Comput*

Vis Image Underst. 2016; **148**: 201–10.
Publisher Full Text

27. Bacci N, Houlton TMR, Briers N, *et al.*: **Validation of forensic facial comparison by morphological analysis in photographic and CCTV samples.** *Int J Legal Med.* 2021. (In Press).

28. Senior AW, Pankanti S: **Privacy Protection and Face Recognition**. In: Huang T, Xiong Z, Zhang Z. editors. *Handbook of Face Recognition.* 2nd ed. London: Springer; 2011; 671–91.
Publisher Full Text

29. Milborrow S, Morkel J, Nicolls F: **The MUCT Landmarked Face Database**. Pattern Recognit Assoc South Africa. 2008.
Reference Source

30. Martinez AM, Benavente R: **The AR Face Database CVC Tech. Report #24**. 1998; 24.
Reference Source

31. Solina F, Peer P, Batagelj B, *et al.*: **Color-based face detection in the" 15 seconds of fame" art installation**. *Proc Mirage, Conf Comput Vision/Computer Graph Collab Model Imaging, Render Image Anal Graph Spec Eff Rocquencourt, Fr*. 2003; 38–47.
Reference Source

32. Phillips JP, Moon H, Rizvi SA, *et al.*: **The FERET evaluation methodology for face-recognition algorithms.** In: *IEEE Trans Pattern Anal Mach Intell.* 2000; **22**(10): 1090–1104.
Publisher Full Text

33. Phillips PJ, Flynn PJ, Scruggs T, *et al.*: **Preliminary face recognition grand challenge results.** In: *International Conference on Automatic Face and Gesture Recognition*. 2006; 15–21.
Publisher Full Text

34. Gao W, Cao B, Shan S, *et al.*: **The CAS-PEAL large-scale chinese face database and baseline evaluations.** *IEEE Trans Syst Man Cybern Syst.* 2008; **38**(1): 149–61.
Reference Source

35. Samaria FS, Harter AC: **Parameterisation of a stochastic model for human face identification.** *IEEE Work Appl Comput Vis - Proc.* 1994; 138–42.
Publisher Full Text

36. Urbanová P, Ferková Z, Jandová M, *et al.*: **Introducing the FIDENTIS 3D Face Database.** *Anthropol Rev.* 2018; **81**(2): 202–23.
Publisher Full Text

37. Le V, Brandt J, Lin Z, *et al.*: **Interactive Facial Feature Localization.** In: *European Conference on Computer Vision*. 2012; 679–92.
Publisher Full Text

38. Jesorsky O, Kirchberg KJ, Frischholz RW: **Robust face detection using the Hausdorff distance.** In: *International Conference on Audio- and Video-based Biometric Person Authentication*. 2001; 90–5.
Publisher Full Text

39. Marszalec E, Martinkauppi B, Soriano M, *et al.*: **Physics-based face database for color research.** *J Electron Imaging.* 2000; **9**(1): 32–28.
Publisher Full Text

40. Lewin NS, Bacci N, Davimes J, *et al.*: **Wits Face Database: Description**. 2021.
http://www.doi.org/10.17605/OSF.IO/Q8V2R

41. Steyn M, Smith JR: **Interpretation of ante-mortem stature estimates in South Africans.** *Forensic Sci Int.* 2007; **171**(2–3): 97–102.
PubMed Abstract | Publisher Full Text

42. Facial Identification Scientific Working Group: **Facial Image Comparison Feature List for Morphological Analysis**. 2018.
Reference Source

43. Stephan CN: **Perspective distortion in craniofacial superimposition: Logarithmic decay curves mapped mathematically and by practical experiment.** *Forensic Sci Int.* 2015; **257**: 520.e1–520.e8.
PubMed Abstract | Publisher Full Text

44. Cooper A, Smith E: **Homicide Trends in the United States,1980-2008: Annual Rates for 2009 and 2010**. U.S. Department of Justice, Office of Justice Programs, Beuraeu of Justice Statistics. 2011.
Reference Source

45. Maluleke R: **Crime Statistics Series Volume V: Crime against Women in South Africa**. 2018.
Reference Source

46. Finn RL, Wright D, Friedewald M: **Seven Types of Privacy**. *European Data Protection: Coming of Age.* 1st ed. Dordrecht: Springer Science+Business Media; 2013; 3–32.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 22 March 2021

https://doi.org/10.5256/f1000research.53979.r80712

© **2021 Lee W.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Won-Joon Lee**
Department of Forensic Medicine, National Forensic Service Seoul Institute, Seoul, South Korea

This manuscript stated the procedure and results of developing a facial image database from South African males. Also, the authors discussed the applicability and limitations of their research.

This study is regarded to produce the best results from given conditions which they stated objectively in the section of challenges and limitations. It is considered that the database will benefit the research and researchers into facial recognition using CCTV images, especially South African males which available database has not been enough in this field.

Consequently, I would like to recommend this manuscript to be indexed.

Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Craniofacial Identification

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 05 March 2021

https://doi.org/10.5256/f1000research.53979.r79950

© **2021 Liu C.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ching-Yiu Jessica Liu
Liverpool John Moores University, Liverpool, UK

This article described the development of the 'Wits Face Database'. The authors collected CCTV footage from three different environments, along with standardised and uncontrolled facial photographs detailed in Table 2. A total of 622 adult male individuals with facial anthropological features resembling South African descent were recruited.

It will be beneficial to document the framerates of the CCTV recordings.

The mean height for black South African males was used for the photography station for ST images. Did you have to make any adjustments for different heights or were all participants within the capture range?

"individuals with facial anthropological features resembling South African descent were recruited." I am not sure I fully understand what this includes and excludes. You mentioned that age and ancestry were not strictly controlled along with the sensitivity of labelling groups. Does this mean anyone who was an adult male was recruited?

Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Facial Anthropology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 17 Mar 2021

Nicholas Bacci, University of the Witwatersrand, Johannesburg, South Africa

Dear Dr Liu,

Thank you for your prompt and positive review. We appreciate your insights for the missing information and points requiring clarification. We will be including the following information in a new version of the manuscript to provide the information to all readers:

1. Comment: It will be beneficial to document the framerates of the CCTV recordings. Response: The framerates and codec information for the IP camera CCTV recordings will be included. However, the analogue camera did not store framerate or codec information in its metadata or the manufacturer website. Amendment: The following statements will be included in the revised manuscript version for clarity: “The recordings from these two cameras were encoded in MPEG-H Part2/HEVC (H.265) codec and had a frame rate of 20 frames per second.”. “No codec and framerate information was available for this camera both on the manufacturer website and the recording metadata.”
2. Comment: The mean height for black South African males was used for the photography station for ST images. Did you have to make any adjustments for different heights or were all participants within the capture range? Response: The height of the tripod for the ST and WT photographs was not adjusted, it was maintained fixed at the average height of black South African males throughout the data collection process. Only a handful of participants were significantly above or below the mean height in order to affect the composition of the photographs. In those cases the angle of the camera was adjusted slightly in the vertical axis to retain the same composition. Amendment: The following clarification will be included in the manuscript: “In all cases, for both ST and WT photographs, the tripod height was not adjusted and left at the fixed height of 1600mm. The majority of participants were within the capture range; however, for a few of the individuals who were either vastly below or above average South African male height, the camera was tilted in the vertical axis of the tripod in order to align their eyes along the uppermost 1/3rd of the portrait composition. No other changes were made in terms of focal length or post crop editing.”
3. Comment: "individuals with facial anthropological features resembling South African descent were recruited." I am not sure I fully understand what this includes and excludes. You mentioned that age and ancestry were not strictly controlled along with the sensitivity of labelling groups. Does this mean anyone who was an adult male was recruited? Response: Clarification on the participants included in the study will be included. Male South Africans of facial appearance resembling African descent were specifically recruited for the database. Amendment: The statement speaking to participant descent will be amended for clarity as follows: “Potential participants were identified and recruited near the data collection sites based on facial anthropological features resembling South African males of African descent and being older than 18 years of age.”.

Best Regards,
Nicholas Bacci

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research