ENHANCED CONTEXTUAL BASED DEEP LEARNING MODEL FOR NIQAB FACE DETECTION

ABDULAZIZ ALI SALEH AL-ASHBI

UNIVERSITI TEKNOLOGI MALAYSIA

ENHANCED CONTEXTUAL BASED DEEP LEARNING MODEL TTFOR
NIQAB FACE DETECTION

ABDULAZIZ AL-ASHBI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Computer Science)

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

FEBRUARY 2022

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to Allah (SWT), who has shown me the right path and provided me with the strength and knowledge to complete this research. May Allah (SWT) accept my work and make it a source of goodness for me in this world and the hereafter, may Allah (SWT) forgive us any mistake and fault (Alhamdulillah firstly and lastly).

I would like to express my immense gratefulness and appreciation to my main thesis supervisor, Professor Dr. Mohd Shahrizal Sunar, for encouragement, guidance, critics, and friendship and support. He has a great personality, good morality, and long patience, without his continuous supervision and help, this thesis would not have been completed.

I am grateful to Universiti Teknologi Malaysia and all its staff for their support and help. Universiti Teknologi Malaysia (UTM), Librarians at UTM, deserve special thanks for their assistance in supplying the relevant literature.

My brother Dr. Fuad Ghaleb should also be recognized for his support and valuable remarks, fellow postgraduate students should also be recognized for their support. My sincere appreciation also extends to all my colleagues and others who have assisted on various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family members.

ABSTRACT

Human face detection is one of the most investigated areas in computer vision which plays a fundamental role as the first step for all face processing and facial analysis systems, such as face recognition, security monitoring, and facial emotion recognition. Despite the great impact of Deep Learning Convolutional neural network (DL-CNN) approaches on solving many unconstrained face detection problems in recent years, the low performance of current face detection models when detecting highly occluded faces remains a challenging problem and worth of investigation. This challenge tends to be higher when the occlusion covers most of the face which dramatically reduce the number of learned representative features that are used by Feature Extraction Network (FEN) to discriminate face parts from the background. The lack of occluded face dataset with sufficient images for heavily occluded faces is another challenge that degrades the performance. Therefore, this research addressed the issue of low performance and developed an enhanced occluded face detection model for detecting and localizing heavily occluded faces. First, a highly occluded faces dataset was developed to provide sufficient training examples incorporated with contextual-based annotation technique, to maximize the amount of facial salient features. Second, using the training half of the dataset, a deep learning-CNN Occluded Face Detection model (OFD) with an enhanced feature extraction and detection network was proposed and trained. Common deep learning techniques, namely transfer learning and data augmentation techniques were used to speed up the training process. The false-positive reduction based on max-in-out strategy was adopted to reduce the high false-positive rate. The proposed model was evaluated and benchmarked with five current face detection models on the dataset. The obtained results show that OFD achieved improved performance in terms of accuracy (average 37%), and average precision (16.6%) compared to current face detection models. The findings revealed that the proposed model outperformed current face detection models in improving the detection of highly occluded faces. Based on the findings, an improved contextual based labeling technique has been successfully developed to address the insufficient functionalities of current labeling technique.

ABSTRAK

Pengecaman wajah manusia merupakan salah satu bidang yang paling banyak dikaji dalam visi komputer yang memainkan peranan asas sebagai langkah pertama dalam semua sistem pemprosesan wajah dan analisis wajah seperti pengecaman wajah, pemantauan keselamatan dan pengecaman emosi wajah. Walaupun terdapat impak yang besar daripada pendekatan Rangkaian Neural Konvolusi-Pembelajaran Mendalam (DL-CNN) dalam menyelesaikan masalah pengecaman wajah tanpa batasan dalam tahun-tahun kebelakangan ini, prestasi yang rendah pada model semasa pengecaman wajah apabila pengecaman wajah yang terhalang masih lagi kekal menjadi masalah yang mencabar dan wajar dikaji. Cabaran ini semakin ketara apabila halangan pada wajah itu menutupi sebahagian besar wajah yang seterusnya secara dramatik mengurangkan bilangan ciri-ciri perwakilan dipelajari yang digunakan oleh Rangkaian Penyarian Sifat (FEN) untuk membezakan bahagian-bahagian wajah daripada latar belakang. Kekurangan set data wajah yang terhalang dengan imej yang mencukupi bagi wajah yang terhalang dengan teruk merupakan cabaran lain yang mengurangkan prestasi pengecaman. Oleh itu, kajian ini menangani masalah prestasi rendah dan membangunkan model pengecaman wajah terhalang yang dipertingkat untuk mengecam dan mengesan dengan tepat wajah-wajah yang terhalang dengan teruk. Pertama, set data wajah-wajah yang terhalang dengan teruk dibangunkan untuk menyediakan contoh latihan mencukupi digabungkan dengan teknik penganotasian berdasarkan kontekstual untuk memaksimumkan jumlah ciri-ciri wajah yang menonjol. Kedua, dengan menggunakan latihan separuh daripada set data, sebuah model Pengesanan Wajah Terhalang (OFD) Pembelajaran Mendalam-CNN dengan ciri pengekstrakan dan rangkaian pengecaman dipertingkat telah dicadangkan dan dilatih. Teknik pembelajaran mendalam yang lazim, iaitu teknik pembelajaran pemindahan data dan teknik peningkatan data diguna pakai untuk mempercepatkan proses latihan. Pengurangan positif-palsu berdasarkan strategi maksimum-masuk-keluar diguna pakai untuk mengurangkan kadar positif-palsu yang tinggi. Model yang dicadangkan dinilai dan ditanda araskan dengan lima model pengecaman wajah semasa di set data. Hasil yang diperoleh menunjukkan bahawa OFD mencapai prestasi yang lebih baik dari segi ketepatan (purata 37%) dan purata ketepatan (16.6%) berbanding dengan model pengecaman wajah semasa. Hasil kajian mendedahkan bahawa model yang dicadangkan mengatasi model pengesanan wajah semasa dengan ketara dalam menambah baik pengecaman wajah-wajah yang terhalang teruk. Berdasarkan dapatan, satu teknik pelabelan berdasarkan kontekstual yang dipertingkat berjaya dibangunkan untuk mengatasi ketidakcukupan fungsi-fungsi teknik pelabelan semasa.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Adaboost	-	Adaptive boosting
ACM	-	Association for Computing Machinery
ATM	-	Automated Teller Machine
CNN	-	Convolutional Neural Network
CPU	-	Central Processing Unit
DL-CNN	-	Deep learning Convolutional Neural Network
DN	-	Detection Network
DPM	-	Deformable Part Model
Fast-RCNN	-	Fast Region-based Convolutional Neural Network
Faster-RCNN	-	Faster Region-based Convolutional Neural Network
FDDB	-	Face detection database
FEN	-	Feature Extraction Network
FN	-	False Negative
GIoU	-	Generalized Intersection Over Union
GPU	-	Graphical Processing Unit
FP	-	False Positive
GT	-	Ground truth
HCI	-	Human-Computer Interaction
HOG	-	Histogram of Gradient
IOU	-	Intersection over Union
LBP	-	Local Binary Patterns
MAFA	-	Masked Faces
ML	-	Machine Learning
MTCNN	-	Multi-Task Cascaded Convolutional Neural Networks
NMS	-	Non-Maximum Suppression
OFD	-	Occluded Face Detection
PC	-	Personal Computer
PCA	-	Principal Component Analysis
RAM	-	Random Access Memory

R-CNN	-	Region-based Convolutional Neural Network
RGB	-	Red Green and Blue
RPN	-	Region Proposal Network
SSD	-	Single Shot Detection
YCbCr	-	Luminance; Chroma: Blue; Chroma: Red
YOLO	-	You Only Look Once

LIST OF APPENDECES

OPERATIONAL DEFINISIONS

Extended five face-regions: is a classification of the face regions into five equal areas, i.e., forehead ,two eyes, nose, mouth and chin, the purpose of this classification is to measure the degree of face occlusal based on the number of occluded regions.

Face regions: are all face related areas that include the forehead, the two eyes and elbows , nose , mouth and chin.

Facial features: facial features can be defined as the distinguishing characteristics of the face that can be described in terms of the size and shape of the entire face and its component parts, which include the two eyes, two ears, nose, mouth, and chin.

Feature extraction network is a convolutional neural network (CNN) composed of set of network layers which process an input image and extract the distinctive features, the extracted features are then forwarded to another CNN network called detection network for classification and detection of targeted objects.

Heavily occluded face: is a face in a digital image with high degree of occlusion, in which four face regions of the face (or more) are blocked by any type of occlusion as in face with niqab for instance.

Image annotation and labelling: is the process of adding metadata in a form of bounding box coordinates of the faces in the images of the dataset to be used as ground truth for the training of deep learning convolutional neural network models.

Niqab: A niqab is a face covering veil which is worn by minority of Muslim women for a religious purpose to cover their faces while being outdoor or among none relatives. Niqab hides most of the face landmarks and sometimes leaves only the two eyes and some parts of the nose.

Niqab-face: is a woman face with niqab covering veil, normally all the face is covered except of the two eyes. Niqab face is considered highly occluded face (since four or more faces regions are hidden).

Occluded face: a face is considered occluded in the digital image if at least one face area is hidden or blocked by any object such as face mask, sun glasses, niqab. Or any other object which hides facial features to be detected.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Finding faces using the visual system is a trivial task that humans can do effortlessly in their daily life, but building machines or robots with vision ability is probably one of the most challenging problems, which humans try to solve. During the past three decades, the computer vision community started to pay attention to face processing as well as other researchers such as psychophysicists and neuroscientists, which have widely investigated recently, many commercial applications and research demonstrations are obtained and developed from these efforts (Voulodimos et al., 2018). The human face is among the most important and informative object that tells a person’s race, sex, identity, age emotion, and more in just a glimpse of a second, the demand of finding a face in digital images or videos has been incrementally increased recently due to wide practical applications in multimedia, biometric systems, surveillance, security applications and human-computer interaction (HCI) (Guo and Zhang, 2019; Tsao and Livingstone, 2008).

Face detection is a sub-task of object detection under the domain of computer vision (Zhao et al., 2019). It has been under hot investigation for more than two decades and still an active area of research in computer vision (Masi et al., 2018; Zhang and Zhang, 2010). It is the first step of all face-application-related including face recognition, face verification, face tracking, and facial expression detection (Zhao et al., 2019; Kortli et al., 2020). It is the building block for more sophisticated systems developed for consumer products like digital cameras, social networks, smartphone apps, etc. (Zafeiriou et al., 2015; Tikoo and Malik, 2017). Figure 1.1 represents the research area according to the Association for Computing Machinery Digital Library (ACM) classification system (Acm, 2012).

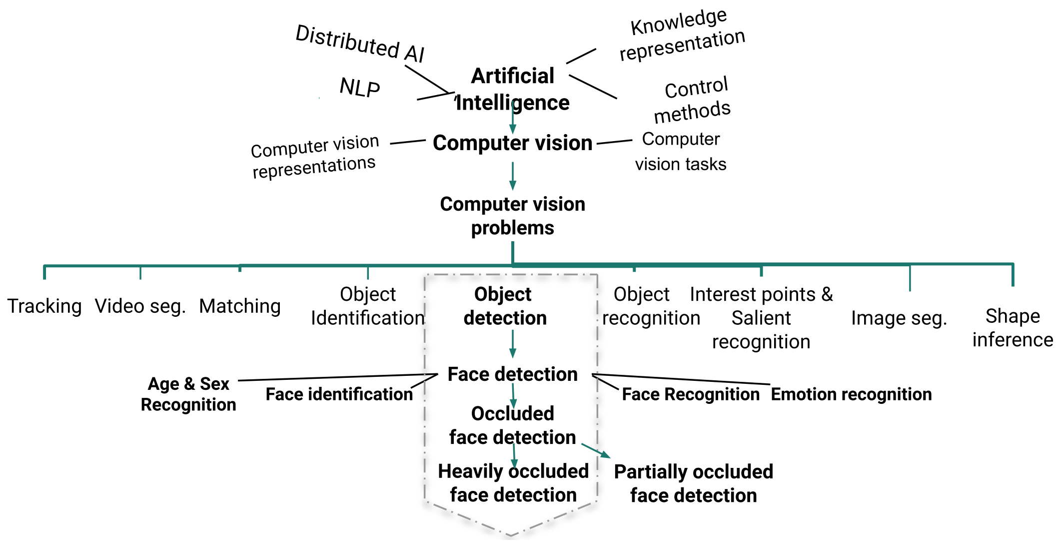


Figure 1.1 Research Focus Area

The enormous applications of face detection motivate researchers to find ways of improving its accuracy and performance. Hundreds of research papers had proposed several approaches for detection methods, the advancement before 2001 had been nicely surveyed and grouped by Hjelmås and Low (2001); (Zhang and Zhang, 2010). However early works within that period were inapplicable in real-world situations until the distinguished work of Viola and Jones (2001) which made face detection feasible and enabled the implementation of face detection in real-world applications, such as in digital cameras and photo software organization, due to the brilliant ideas which used a successful combination of machine learning with feature invariant techniques (Zafeiriou et al., 2015; Zhang and Zhang, 2010).

Since then, great progress has been made to boost the performance, for instance, deformable part models (DPM) were proposed to model face parts (Felzenszwalb et al., 2010). Histogram of Gradient (HOG) introduced by Dalal and Triggs (2005), Local Binary Patterns (LBP) by (Chehrehgosha and Emadi, 2016), and Integral Channels by Wang et al. (2009a) had a significant impact on face detection to achieve better accuracy on frontal faces in controlled environments, where no extreme variations of illumination and lighting conditions. All these aforementioned approaches were mainly dependent on manually designed and handcrafted features and classifiers (Li et al., 2019b).

The emergence of deep learning and convolutional neural network (DL-CNN) have recently shown remarkable successes and dominated various computer vision tasks such as image recognition, object detection, and face detection (Hao et al., 2017; Liu et al., 2020). Unlike hand-crafted features in machine learning as in Haar-features for instance DL-CNN is a hierarchical deep-learning approach that has been successfully applied as a powerful features extraction, that extracts and learns high-level representation features from a vast amount of training data (Alafif et al., 2017). Impressive progress has been made in detecting human faces from digital images where an average performance of 98% is achieved by Hu and Ramanan (2017) in the unconstrained face-detection benchmark.

Among the challenges in the design of face detection is the low performance of face detectors when detecting faces in certain scenarios due to illumination differences, facial expression changes, pose variations, and with the presence of occlusion which appears when covering faces either in partial or in total (Chen et al., 2018b). In fact, an accurate face detection model as a primary stage of any human face processing system has a major influence on the overall performance and usability of the entire practical applications.

1.2 Problem Background

The goal of face detection is to determine the existence of a face in an image, therefore, if a face exists, return its location in a form of bounding box coordinates for each face (Ranjan et al., 2018). It should be able to find faces among all non-face objects (Zhang and Zhang, 2010). Various face-related applications require face detection as a pre-processing step for finding face location, many of the techniques are proposed for these applications assume that the location of the face is pre-identified and available for the next step, therefore all these systems for sure will fail if face detection fails to accurately allocate the face (Zhao et al., 2019).

Detecting faces in constrained conditions where images are taken in controlled settings, with a fixed distance from the camera, and with specific pose and lighting conditions as in the frontal face, for instance, is no longer a challenging task (Masi et al., 2018) and is considered a solved problem (Guo and Zhang, 2019). It was made possible since 2001 after the aforementioned work of Viola and Jones (2001). They used haar-like features with adaptive boosting (Adaboost) designed in a cascaded manner and integral image. Modern face detection systems can straightforwardly detect faces in near-frontal and are embedded in e-albums and digital cameras (Yang et al., 2016a).

However, face detection models have to be able to detect not just the frontal faces, but they also have to deal with faces of various scales, poses, and appearances and with the presence of occlusion that can change the overall appearance of the face (Li et al., 2018; Yang et al., 2002).

The high variations of face orientation, facial expression, and occlusion in an unconstrained environment where images are taken in different illumination variations and occlusion is a challenging problem for face detection systems and may degrade the performance of the detection, it may also be responsible for the increase of the false detection rate (Chen et al., 2018b; Zafeiriou et al., 2015). The following factors are considered the main challenges associated with face detection performance:

(a) **Pose**: The position of the face in images may differ due to variation in-plane rotation (frontal, half profile, profile, upside down), and some facial features such as eye or nose may become partially or fully occluded (Moallem et al., 2015).

(b) **The existence of certain facial components:** some facial components like glasses, caps, mustaches, and beards, for instance, can exist or not on the face, these components vary greatly in their form, color, and size (Sharifara et al., 2014).

(c) **Facial expression:** the appearance of faces such as happiness, angriness, or sadness can directly affect a person’s facial expression (Dagar et al., 2016).

(d) **Image orientation and Illumination:** when the camera's optical axis varies, image orientation will exist and be formed, other aspects like illumination and lightning variant intensity, and camera calibration (sensor-response, type, and size of lenses) (Zou et al., 2007).

(e) **Scale variations and small face: s**mall faces in crowed in which an image could have tens or hundreds of faces makes the detection task difficult and lower the performance of face detection models (Bai et al., 2018).

(f) **Occlusion:** when some or more parts of the face is unavailable, either blocked by other object or covered by face-veil partially or in total (Peng et al., 2020; Chen et al., 2018b).

The existence of one or more of the aforementioned challenges may degrade the performance of face detectors. A lot of progress has been done to improve the performance of face detectors under many unconstrained scenarios. Many previous researches have addressed these challenges of unconstrained scenarios (Chen et al., 2014; Li et al., 2015; Mathias et al., 2014). For example, pose, scale, and lighting variations were addressed by feature invariant approaches which focused on finding face features robust to changes in pose and lighting (Zafeiriou et al., 2015).

The recent improvement of face detection in the unconstrained scenario could be attributed to two factors, i) the emergence of deep learning approaches that have a direct impact on the extraction of facial representative features and analysis tasks, which enabled the existence of sophisticated face detections models, and ii) the availability of large-scale face detection datasets with varieties of training images (Nada et al., 2018).

Deep learning has been recently behind the great advancement of object detection and face detection (Liu et al., 2020; Zhao et al., 2019; Liang et al., 2020). The application of DL and CNN layers have provided excellent feature extraction and learning methods for accurate face detection models, while multi-stage region proposal network and single-stage detection such as Faster-RCNN proposed by Ren et al. (2015) and YOLO introduced by Redmon and Farhadi (2017) have significantly enhanced the performance of face detection in an unconstrained environment. The availability of public face detection datasets such as Annotated Faces in The Wild (AFW) (Zhu and Ramanan, 2012), FDDB (Jain and Learned-Miller, 2010), and Widerface (Yang et al., 2016a) have contributed to the advancement in face detection research.

Despite the progress that has been achieved in face detection, detecting faces in certain scenarios as in face with occlusion has not reached saturation yet, there have been open questions that remains unsolved and has to be explored more when dealing with occlusion (Yang et al., 2018; Mathias et al., 2014). Face detection under patrial or heavy occlusion remains a challenge to face detection algorithms and worth investigation (Chen et al., 2018b; Alafif et al., 2017).

Although the detection of heavily occluded faces is critical for several applications, it is highly demanded for security monitoring and people-counting applications. Very few researches have been done in that direction and directly addressed the detection under occlusion. For example, Hotta (2007) used local features with Support Vector Machine (SVM) to detect faces under partial occlusion, Chen et al. (2017) proposed an occlusion aware framework based on convolutional neural network model to address the occlusion problem in face detection, Alafif et al. (2017) trained a single CNN model on large partial occluded faces images to detect unconstrained multi-view partially occluded and non-partially occluded faces.

Faces under partial occlusion have been addressed in general as in the aforementioned works, however, the challenge of highly occluded faces was not considered. In heavily occluded faces most of the face features are hidden and blocked due to the occlusion.

Occluded face or face under occlusion usually appears when the face is covered or blocked either in partial or in total, which is due to work requirement such as medical masks as in hospitals or due to pandemic awareness of the current COVID-19, it could be also due to religious concern as in some Muslim societies where Muslim ladies wear niqab, a face covering veil, which is a practice of veiling their faces while being outdoor or in the presence of non-relatives (Khan, 2016; Zempi, 2016).

The word ‘Niqab’ is used to refer to head and face covering worn by Muslim women, thus they can often be distinguished by the way they dress and cover their faces (Chowdhury et al., 2017a). Figure 1.2 shows examples of heavily occluded faces for some Muslim women are wearing the niqab veil in different cultural styles referred as niqab, burqa or khimar (Chowdhury et al., 2017b; Khan, 2016). The whole faces are hidden and almost covered and blocked by niqab; therefore, faces are heavily occluded.

Figure 1.2 Examples of heavily occluded faces with different niqab styles

There are some researchers who tried to define occluded faces according to the degree of occlusion. For example, in Yang et al. (2016a) they classified face under occlusion into three categories: face with no occlusion, partially occluded faces, and heavily occluded faces. Partially occlusion is defined as a face where 1% to 30% of its area is occluded, whereas heavily occlusion is when more than 30% of the face area is covered or blocked.

In Ge et al. (2017) they divided the face into four main regions as shown in Figure 1.2 section (a), these major regions are chin, mouth, nose and the eyes. they defined the degree of occlusion based on the number of occluded regions; therefore, weak occlusion is where one to two regions are occluded, medium occlusion is with three regions and heavy occlusion is when there are four occluded regions.

However, it seems a loose classification when defining heavily occluded faces as faces occluded or covered with over 30%. There is a concern with describing highly occluded faces, defining heavily occluded faces when four regions are occluded still not cover all the heavily and fully occluded faces. Therefore, based in Ge et al. (2017) instead of four regions we extend the classification of the face regions into five equal areas which include the forehead, the two eyes, the nose, the mouth and the chin, so that heavily occluded and fully occluded can be distinguished by the number of occluded areas.

The degree of heavily occlusion is shown in Figure 1.3 and 1.4. below. Figure 1.3 section (a) shows the four face regions of Ge et al. (2017). The extended five-regions are shown in section (b). Section (c) shows the overlay of the extended five-regions with the four regions of Ge et al. (2017) in order to emphasis and point out the distinction between the two definitions. Section (c) shows an example of a face heavily occluded with four occluded regions according to Ge et al. (2017), however the degree of occlusion is approximately 50% when overlaid by the extended five-regions. Figure 1.4 shows faces in high degree of occlusion with four occluded regions. The faces are arranged according to the number of occluded areas, from left to right. Although all faces are considered heavily occluded according to the four occluded area definition of Ge et al. (2017), however there is still a distinction in the degree of occlusion between the faces from left to right which ranges from 70% till 100% of full occlusion, the degree of occlusion can be measured more clearly when using the extended five face regions.

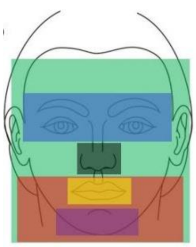
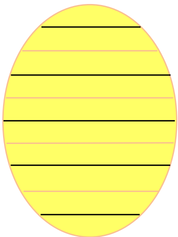
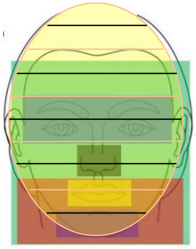
				
(a) The four face regions (Ge et al., 2017)	(b) The Face oval divided into five regions	(c) The five-face-regions is overlied on foue-face regions	(d) Heabvily occluded face with four occluded regions according to Ge et al. (2017)	(e) 50% of occluded areas

Figure 1.3 The four-regions of the occluded face of Ge et al.(2017) alongside with the five-regions

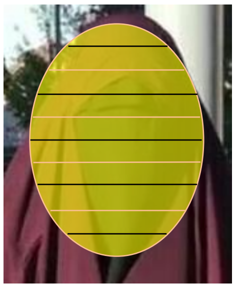
Variations of heavily occluded faces				
Five occluded regions overlaid on heavily occluded faces				
	occluded area $\geq$ 70%	occluded area $\geq$ 80%	occluded area $\geq$ 90%	occluded areas $\approx$ 100%

Figure 1.4 High occluded faces covered with niqab in different degree of occlusion

The existence of recently proposed large-scale datasets has attributed to the high performance of face detection models which almost recent state-of-the-art deep learning-CNN face detection models have leveraged a large amount of training dataset and the powerful features extractors network of deep-CNN (Nada et al., 2018). The amount of training images used during the training determines the performance of the detector (Shepley, 2019; Kim et al., 2018).

Among the reasons that may explain why the performance of recent face detectors decrease and perform poorly when the occlusion is very high as in face covered with niqab for instance in which the occlusion blocks most of the face area is the lack of a large-scale dataset with abundant training images of particular scenarios as in images with faces covered with niqab. Several face detection datasets have been recently released. For example, Widerface proposed by Yang et al. (2016a) and FDDB proposed by Jain and Learned-Miller (2010) and were used for training and testing of current face detection models (Chen et al., 2017). For instance, MTCNN and TinyFace models which were proposed by Zhang et al. (2016) and Hu and Ramanan (2017) respectively were both trained on Widerface dataset.

Though these datasets may contain thousands of images with different pose rotation, different lighting, and some degree of occlusion. However, the high degree of occlusion scenario particularly, as in faces covered with a niqab has not existed among them. This led to the inability of current face detectors to successfully detect heavily covered faces as in niqab-face for instance, which is regarded as a consequent result of the absence of sufficient representative features extracted and learned by CNN-based face detection models during the training. Many state-of-the-art face detection models were initially trained on Widerface dataset which contains 18,839 images with faces in different poses and occlusion but with a limited amount of images of highly occluded faces, except for a few images with faces covered with medical masks and not sufficient for the training (Hu and Ramanan, 2017; Zhang et al., 2016) (Li et al., 2020; Llinzai, 2019).

Another assumption that may influence the performance of recent face detection models and contribute to their poor performance when detecting faces in a high degree of occlusion, is associated with the occlusion problem itself (Li et al., 2019a). Excessively, occluded faces have diminished discrimination and very few features due to occlusion, unlike typical faces with salient features (Shepley, 2019). Therefore, the existence of occlusion limits the distinctive features of faces and restricts the number of learned representative features during the training (Chen et al., 2018b).

1.3 Problem Statement

Face detection as the first and the fundamental step of any automated face processing and facial analysis systems motivates researchers to find ways of improving the accuracy and performance of the system. An accurate face detection system has a direct impact on the overall performance and accuracy of all face-related applications such as face recognition, face identification, security monitoring, and facial emotion detection. A robust face detection model should be effective under arbitrary variations in pose and occlusion; however, it is still an unresolved problem, one of the challenges in the design of face detection is the low performance of face detectors when detecting occluded faces.

A substantial gap exists between the accuracy of existing face detectors and the expected performance in the case of a high degree of occlusion. Several approaches were proposed to address the problem of partially occluded faces; however, the challenge of highly occluded faces was not considered, the problem of a high degree of occlusion where the majority of the face is covered still remains a challenge. State-of-the-art face detectors still have problems in dealing with faces in a high degree of occlusion. This challenge has not yet been entirely solved. The degree of occlusion has a direct effect on the performance of face detection, the detection rate decreases as occlusion level increases.

The lack of available labeled dataset of highly occluded faces with large and sufficient numbers of images with faces in varieties of a high degree of occlusion to be used for training of occluded face detectors broadens the existing gap. The scarcity of salient representative features in highly occluded faces complicates the task of the feature extraction network and restricts its ability to learn adequate discriminative features from the training examples during the training.

The general research question is how to enable face detection models to be able to detect and localize heavily occluded faces on digital images? Therefore, this thesis aims to design and develop an improved deep learning occluded face detection model for detecting heavily occluded faces as in faces covered with niqab for example.

1.4 Research Goal

The goal of this research is to propose and design an enhanced heavily occluded face detection model which is capable of detecting faces in a high degree of occlusion as in faces covered with niqab. The degree of occlusion in faces with niqab ranges from 50% to 90% (as illustrated previously in Figure 1.4) where most of the face features are hidden.

1.5 Research Objectives

To achieve the research goal, the following objectives must be accomplished:

1) To propose an occluded face detection dataset that takes into account the properties of heavily covered faces suitable to be used for deep-learning face detection training and evaluation.

2) To design a contextual-based deep learning scheme for enhancing the representative features with an improved feature extraction network of the heavily covered faces to improve the detection performance.

3) To propose a deep-learning-CNN face detection model capable of detecting faces in a high degree of occlusion where most of the face is veiled in unconstrained environment.

1.6 Research Scope and Limitation

The proposed occluded face detection (OFD) model is introduced to address occluded faces. Although, it potentially can deal with un-occluded faces such as frontal and unconstrained faces, however, the focus is more on heavily covered faces as in faces in niqab for instance.

This research is concerned with occluded face detection, mainly focused on heavily occluded faces which results due to wearing of the face-covering veil, with the aim to improving face detection performance under a high degree of occlusion as in heavily covered faces with niqab which is referred to as niqab-face. The focus on heavily covered faces is different from other types of occlusions which result due to the crowds for instance, so that, some faces are occluded by other faces or objects and may suffer from low-resolution quality due to crowd which makes it difficult to be detected.

Heavily covered faces may have no resolution issue, an image may contain one face only as shown previously in Figure 1.2 for instance, however, it may not be detected successfully, due to the poor performance of current face detection models, because they were not exposed to sufficient training examples of images that contain faces with similar type of occlusion, and also due to the limited features on occluded faces.

Since the unconstrained environment has extreme variations in where faces could appear, this research is constrained on occluded faces with a resolution not less than 80x80. Occluded faces in low resolution images are very hard to be detected this is because face in occlusion is already have limited features, therefore low-quality images and low resolutions make it very difficult to find the representative face features and harden the task of feature extraction to find distinguished features on low resolution images. On the other hand, images with occluded faces in extreme pose and poor illuminations are also another burden that makes the detection worse.

The following scenarios are considered out of the research scope.

- Small faces with high occlusion as in crowded faces or with extreme lighting and in low-quality images.
- Faces in extreme pose with high degree of occlusion are not addressed in this research.

1.7 Significance of the Study

It is expected that the proposed occluded face detection model will improve the performance of the detection of highly occluded faces since all face-related applications rely on face detection as an essential preprocessing step, therefore their overall performance and accuracy will be improved accordingly.

It is also expected that the proposed occluded face dataset as training and benchmark publicly available dataset will contribute to the improvement of current face detection researches and models. The idea of utilization of contextual information of occluded faces and the improvement of the feature extraction network is expected to attract other researchers to be used for similar situations.

1.8 Thesis Organization

This thesis is organized as follows.

Chapter 2 provides a comprehensive literature review of the related area highlighting problem background and existing solutions in the context of deep learning and occluded face detection. It started with the importance of face detection for all face-related applications, followed by a briefing of face detection challenges. More focus was shed on the remaining challenges related to occlusion. Two popular approaches for addressing face detection challenges were briefly discussed, more concern was given to the deep learning convolutional neural network approach as it gained a positive record in solving most of the computer vision and object detection tasks. Recently available solutions have made use of contextual information for addressing occlusion and small face issues in face detection with more attention on occluded face detection. The available solutions are discussed to highlight the novelty of this research. A review of face detection datasets was provided along with a comprehensive discussion. Research direction was highlighted with illustration on current challenges on dataset and availability of feature representation. Finally, a summary was provided.

Chapter 3 describes the research methodology to guide as a roadmap to achieve and verify research objectives, all required phases for designing and developing the proposed model were briefly highlighted.

Chapter 4 presents the design and implementation of the proposed occluded face detection model. All phases related to the design are described and illustrated in detail. The first phase of dataset construction, then Occluded face detection model design, the last phase related to the implementation involves training and evaluation of the proposed model.

Chapter 5 provides details of the experimental configurations and evaluation results of the proposed OFD model and the current benchmarked face detection models. Details analysis of the obtained result was interpreted, discussed, and compared with current models.

Chapter 6 provides a conclusion about this study by emphasizing contributions and recommended future direction and possible further enhancement of this research.

REFERENCES

Acm, D. 2012. Acm Computing Classification System. The ACM Digital Library website.

Adhikari, B., Peltomaki, J., Puura, J. and Huttunen, H. Faster Bounding Box Annotation for Object Detection in Indoor Scenes. 2018 7th European Workshop on Visual Information Processing (EUVIP), 2018. IEEE, 1-6.

Alafif, T., Hailat, Z., Aslan, M. and Chen, X. On Detecting Partially Occluded Faces with Pose Variations. 2017 14th International Symposium on Pervasive Systems, Algorithms and Networks & 2017 11th International Conference on Frontier of Computer Science and Technology & 2017 Third International Symposium of Creative Computing (ISPAN-FCST-ISCC), 2017. IEEE, 28-37.

Angelova, A., Abu-Mostafam, Y. and Perona, P. Pruning Training Sets for Learning of Object Categories. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005. IEEE, 494-501.

Ansari, A.-N. and Abdel-Mottaleb, M. 3d Face Modeling Using Two Views and a Generic Face Model with Application to 3d Face Recognition. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003., 2003. IEEE, 37-44.

Bai, Y., Zhang, Y., Ding, M. and Ghanem, B. Finding Tiny Faces in the Wild with Generative Adversarial Network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018. 21-30.

Bargal, S. A., Barsoum, E., Ferrer, C. C. and Zhang, C. Emotion Recognition in the Wild from Videos Using Images. Proceedings of the 18th ACM International Conference on Multimodal Interaction, 2016. 433-436.

Bojarski, M., Choromanska, A., Choromanski, K., Firner, B., Jackel, L., Muller, U. and Zieba, K. 2018. Visualbackprop: Efficient Visualization of Cnns.

Bolme, D. S. 2003. Elastic Bunch Graph Matching. Colorado State University.

Branco, P., Torgo, L. and Ribeiro, R. P. 2015. A Survey of Predictive Modelling under Imbalanced Distributions. arXiv preprint arXiv:1505.01658.

Chan, H. and Bledsoe, W. 1965. A Man-Machine Facial Recognition System: Some Preliminary Results. Panoramic Research Inc., Palo Alto, CA, USA1965.

Chehrehgosha, A. and Emadi, M. 2016. Face Detection Using Fusion of Lbp and Adaboost. Journal of Soft Computing and Applications, 2016**,** 1-10.

Chen, D., Ren, S., Wei, Y., Cao, X. and Sun, J. Joint Cascade Face Detection and Alignment. European conference on computer vision, 2014. Springer, 109-122.

Chen, S., Liu, Y., Gao, X. and Han, Z. Mobilefacenets: Efficient Cnns for Accurate Real-Time Face Verification on Mobile Devices. Chinese Conference on Biometric Recognition, 2018a. Springer, 428-438.

Chen, W., Huang, H., Peng, S., Zhou, C. and Zhang, C. 2020. Yolo-Face: A Real-Time Face Detector. The Visual Computer**,** 1-9.

Chen, Y., Song, L. and He, R. 2017. Masquer Hunter: Adversarial Occlusion-Aware Face Detection. arXiv preprint arXiv:1709.05188.

Chen, Y., Song, L. and He, R. 2018b. Adversarial Occlusion-Aware Face Detection. arXiv preprint arXiv:1709.05188V6.

Chowdhury, N. A., Bakar, H. S. A. and Elmetwally, A. A. 2017a. Probing Niqab Wearing as an Islamic Identity, Cultural Piety and Women's Empowerment: A Phenomenological Approach. International Journal of Ethics in Social Sciences, 5**,** 57-76.

Chowdhury, N. A., Elmetwally, A. A. and Bakar, H. S. A. 2017b. Misconception of Islamic Apparel, Niqab: A Phenomenological Approach. Jurnal Komunikasi: Malaysian Journal of Communication, 33.

Dagar, D., Hudait, A., Tripathy, H. and Das, M. Automatic Emotion Detection Model from Facial Expression. 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 2016. IEEE, 77-85.

Dalal, N. and Triggs, B. Histograms of Oriented Gradients for Human Detection. 2005.

De, A. and Saha, A. A Comparative Study on Different Approaches of Real Time Human Emotion Recognition Based on Facial Expression Detection. 2015 International Conference on Advances in Computer Engineering and Applications, 2015. IEEE, 483-487.

Deng, J., Guo, J., Zhou, Y., Yu, J., Kotsia, I. and Zafeiriou, S. 2020. Retinaface: Single-Shot Multi-Level Face Localisation in the Wild. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**,** 5203-5212.

Deng, W., Fang, Y., Xu, Z. and Hu, J. 2018. Facial Landmark Localization by Enhanced Convolutional Neural Network. Neurocomputing, 273, 222-229.

Dong, X., Yan, Y., Ouyang, W. and Yang, Y. Style Aggregated Network for Facial Landmark Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018. 379-388.

Everingham, M., Eslami, S. A., Van Gool, L., Williams, C. K., Winn, J. and Zisserman, A. 2015. The Pascal Visual Object Classes Challenge: A Retrospective. International journal of computer vision, 111, 98-136.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J. and Zisserman, A. 2010. The Pascal Visual Object Classes (Voc) Challenge. International journal of computer vision, 88, 303-338.

Everingham, M., Zisserman, A., Williams, C. K., Van Gool, L., Allan, M., Bishop, C. M., Chapelle, O., Dalal, N., Deselaers, T. and Dorkó, G. 2012. Pascal Voc Challenge Performance Evaluation website.

Felzenszwalb, P. F., Girshick, R. B., Mcallester, D. and Ramanan, D. 2010. Object Detection with Discriminatively Trained Part-Based Models. IEEE transactions on pattern analysis and machine intelligence, 32, 1627-1645.

Feng, Z.-H., Kittler, J., Awais, M., Huber, P. and Wu, X.-J. Wing Loss for Robust Facial Landmark Localisation with Convolutional Neural Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018. 2235-2245.

Ffmpeg, M. F. 2000. Ffmpeg , an Open Source Gnu Multimedia Cross-Platform Framework.

Foundation, P. S. 2020. Python, a Programming Language.

Freund, Y., Schapire, R. and Abe, N. 1999. A Short Introduction to Boosting. Journal-Japanese Society For Artificial Intelligence, 14, 1612.

Ge, S., Li, J., Ye, Q. and Luo, Z. Detecting Masked Faces in the Wild with Lle-Cnns. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. 2682-2690.

Ghiasi, G. and Fowlkes, C. C. 2015. Occlusion Coherence: Detecting and Localizing Occluded Faces. arXiv preprint arXiv:1506.08347.

Girshick, R. Fast R-Cnn. Proceedings of the IEEE international conference on computer vision, 2015. 1440-1448.

Girshick, R. 2019. The Generalized R-Cnn Framework for Object Detection. CVPR 2019 Tutorial.

Goldstein, A. J., Harmon, L. D. and Lesk, A. B. 1971. Identification of Human Faces. Proceedings of the IEEE, 59**,** 748-760.

Grother, P. and Ngan, M. 2017. The Ijb-a Face Identification Challenge Performance Report. Technical Report. National Institute of Standards and Technology.

Günther, M., Hu, P., Herrmann, C., Chan, C.-H., Jiang, M., Yang, S., Dhamija, A. R., Ramanan, D., Beyerer, J. and Kittler, J. Unconstrained Face Detection and Open-Set Face Recognition Challenge. 2017 IEEE International Joint Conference on Biometrics (IJCB), 2017. IEEE, 697-706.

Guo, G. and Zhang, N. 2019. A Survey on Deep Learning Based Face Recognition. Computer Vision and Image Understanding, 189**,** 102805.

Guo, Z., Zhou, W., Xiao, L., Hu, X., Zhang, Z. and Hong, Z. Occlusion Face Detection Technology Based on Facial Physiology. 2018 14th International Conference on Computational Intelligence and Security (CIS), 2018. IEEE, 106-109.

Hao, Z., Liu, Y., Qin, H., Yan, J., Li, X. and Hu, X. Scale-Aware Face Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. 6186-6195.

He, K., Gkioxari, G., Dollár, P. and Girshick, R. Mask R-Cnn. Proceedings of the IEEE international conference on computer vision, 2017. 2961-2969.

He, K., Zhang, X., Ren, S. and Sun, J. Deep Residual Learning for Image Recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 770-778.

Hjelmås, E. and Low, B. K. 2001. Face Detection: A Survey. Computer vision and image understanding, 83**,** 236-274.

Hotta, K. 2007. Robust Face Detection under Partial Occlusion. Systems and Computers in Japan, 38**,** 39-48.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H. 2017. Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861.

Hu, P. and Ramanan, D. Finding Tiny Faces. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. IEEE, 1522-1530.

Ioffe, S. and Szegedy, C. 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv preprint arXiv:1502.03167.

Jaccard, P. 1901. Étude Comparative De La Distribution Florale Dans Une Portion Des Alpes Et Des Jura. Bull Soc Vaudoise Sci Nat, 37**,** 547-579.

Jain, D. K., Shamsolmoali, P. and Sehdev, P. 2019. Extended Deep Neural Network for Facial Emotion Recognition. Pattern Recognition Letters, 120**,** 69-74.

Jain, V. and Learned-Miller, E. G. 2010. Fddb: A Benchmark for Face Detection in Unconstrained Settings. UMass Amherst Technical Report.

Jiang, H. and Learned-Miller, E. Face Detection with the Faster R-Cnn. 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 2016. IEEE, 650-657.

Jogin, M., Madhulika, M., Divya, G., Meghana, R. and Apoorva, S. Feature Extraction Using Convolution Neural Networks (Cnn) and Deep Learning. 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2018. IEEE, 2319-2323.

Joseph, A. and Geetha, P. 2020. Facial Emotion Detection Using Modified Eyemap–Mouthmap Algorithm on an Enhanced Image and Classification with Tensorflow. The Visual Computer, 36**,** 529-539.

Kang, M., Ji, K., Leng, X. and Lin, Z. 2017. Contextual Region-Based Convolutional Neural Network with Multilayer Fusion for Sar Ship Detection. Remote Sensing, 9**,** 860.

Kepenekci, B., Tek, F. B. and Akar, G. B. Occluded Face Recognition Based on Gabor Wavelets. Proceedings. International Conference on Image Processing, 2002. IEEE, I-I.

Khan, M. 2016. Niqab: An Approach Based on the Proofs. The Institute for the Revival of the Traditional Islamic Sciences. https://bukhari2013. files. wordpress. com/2016/12/niqab-an-approach-based-on-the-proofs. pdf. Accessed on, 30.

Kim, B., Ban, S.-W. and Lee, M. Multiple Occluded Face Detection Based on Binocular Saliency Map. International Conference on Neural Information Processing, 2009. Springer, 693-700.

Kim, J., Sung, Y., Yoon, S. M. and Park, B. G. A New Video Surveillance System Employing Occluded Face Detection. International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, 2005. Springer, 65-68.

Kim, K., Yang, Z., Masi, I., Nevatia, R. and Medioni, G. Face and Body Association for Video-Based Face Recognition. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018. IEEE, 39-48.

Klare, B. F., Klein, B., Taborsky, E., Blanton, A., Cheney, J., Allen, K., Grother, P., Mah, A. and Jain, A. K. Pushing the Frontiers of Unconstrained Face Detection and Recognition: Iarpa Janus Benchmark A. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 1931-1939.

Koestinger, M., Wohlhart, P., Roth, P. M. and Bischof, H. 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), 2011. IEEE, 2144-2151.

Kortli, Y., Jridi, M., Al Falou, A. and Atri, M. 2020. Face Recognition Systems: A Survey. Sensors, 20**,** 342.

Köstinger, M., Wohlhart, P., Roth, P. M. and Bischof, H. Annotated Facial Landmarks in the Wild: A Large-Scale, Real-World Database for Facial Landmark Localization. Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, 2011. IEEE, 2144-2151.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. Imagenet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems, 2012. 1097-1105.

Lapuschkin, S., Binder, A., Muller, K.-R. and Samek, W. Understanding and Comparing Deep Neural Networks for Age and Gender Classification. Proceedings of the IEEE International Conference on Computer Vision Workshops, 2017. 1629-1638.

Lecun, Y., Bengio, Y. and Hinton, G. 2015. Deep Learning. nature, 521**,** 436.

Levi, G. and Hassner, T. Age and Gender Classification Using Convolutional Neural Networks. Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2015. 34-42.

Li, C., Wang, R., Li, J. and Fei, L. 2020. Face Detection Based on Yolov3. Recent Trends in Intelligent Computing, Communication and Devices. Springer.

Li, H., Lin, Z., Shen, X., Brandt, J. and Hua, G. A Convolutional Neural Network Cascade for Face Detection. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 5325-5334.

Li, J., Wang, Y., Wang, C., Tai, Y., Qian, J., Yang, J., Wang, C., Li, J. and Huang, F. 2018. Dsfd: Dual Shot Face Detector. arXiv preprint arXiv:1810.10220.

Li, Y., Zeng, J., Shan, S. and Chen, X. 2019a. Occlusion Aware Facial Expression Recognition Using Cnn with Attention Mechanism. IEEE Transactions on Image Processing, 28**,** 2439-2450.

Li, Z., Tang, X., Han, J., Liu, J. and He, R. 2019b. Pyramidbox++: High Performance Detector for Finding Tiny Face. arXiv preprint arXiv:1904.00386.

Liang, X., Zhao, X., Zhao, C., Jiang, N., Tang, M. and Wang, J. Task Decoupled Knowledge Distillation for Lightweight Face Detectors. Proceedings of the 28th ACM International Conference on Multimedia, 2020. 2184-2192.

Liao, S., Jain, A. K. and Li, S. Z. 2016. A Fast and Accurate Unconstrained Face Detector. IEEE transactions on pattern analysis and machine intelligence, 38**,** 211-223.

Lienhart, R. and Maydt, J. An Extended Set of Haar-Like Features for Rapid Object Detection. Image Processing. 2002. Proceedings. 2002 International Conference on, 2002. IEEE, I-I.

Lin, T.-Y., Goyal, P., Girshick, R., He, K. and Dollár, P. Focal Loss for Dense Object Detection. Proceedings of the IEEE international conference on computer vision, 2017. 2980-2988.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C. L. Microsoft Coco: Common Objects in Context. European conference on computer vision, 2014. Springer, 740-755.

Lin, Y.-Y., Liu, T.-L. and Fuh, C.-S. 2007. Face Detection with Occlusions. Images & Recognition, 13**,** 4-21.

Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X. and Pietikäinen, M. 2020. Deep Learning for Generic Object Detection: A Survey. International journal of computer vision, 128**,** 261-318.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. and Berg, A. C. Ssd: Single Shot Multibox Detector. European conference on computer vision, 2016. Springer, 21-37.

Llinzai, L. 2019. Lightweight Facedetection Model Designed for Edge Computing Devices.

Ma, H., Liu, Y., Ren, Y. and Yu, J. 2020. Detection of Collapsed Buildings in Post-Earthquake Remote Sensing Images Based on the Improved Yolov3. Remote Sensing, 12**,** 44.

Masi, I., Trần, A. T., Hassner, T., Sahin, G. and Medioni, G. 2019. Face-Specific Data Augmentation for Unconstrained Face Recognition. International Journal of Computer Vision, 127**,** 642-667.

Masi, I., Wu, Y., Hassner, T. and Natarajan, P. Deep Face Recognition: A Survey. 2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI), 2018. IEEE, 471-478.

Mathias, M., Benenson, R., Pedersoli, M. and Van Gool, L. Face Detection without Bells and Whistles. European Conference on Computer Vision, 2014. Springer, 720-735.

Moallem, P., Mousavi, B. and Naghibzadeh, S. S. 2015. Fuzzy Inference System Optimized by Genetic Algorithm for Robust Face and Pose Detection. Int. J. Artif. Intell, 13**,** 73-88.

Morder, I. 2020. Facial Recognition Market | Growth, Trends, and Forecasts (2020—2025).

Mtcnn, P. 2019. Multi-Task Cascaded Convolutional Neural Networks for Face Detection, Based on Tensorflow.

Nada, H., Sindagi, V. A., Zhang, H. and Patel, V. M. Pushing the Limits of Unconstrained Face Detection: A Challenge Dataset and Baseline Results. 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2018. IEEE, 1-10.

Najibi, M., Samangouei, P., Chellappa, R. and Davis, L. S. Ssh: Single Stage Headless Face Detector. Proceedings of the IEEE International Conference on Computer Vision, 2017. 4875-4884.

Najibi, M., Singh, B. and Davis, L. S. Fa-Rpn: Floating Region Proposals for Face Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019. 7723-7732.

Ngan, M., Grother, P. and Hanaoka, K. 2020. Ongoing Face Recognition Vendor Test (Frvt).

Nguyen, N.-V., Rigaud, C. and Burie, J.-C. 2018. Digital Comics Image Indexing Based on Deep Learning. Journal of Imaging, 4**,** 89.

Olafenwa, M. O. J. 2020. Imageai Python Library.

Padilla, R., Netto, S. L. and Da Silva, E. A. A Survey on Performance Metrics for Object-Detection Algorithms. 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), 2020. IEEE, 237-242.

Peng, X.-Y., Cao, J. and Zhang, F.-Y. Masked Face Detection Based on Locally Nonlinear Feature Fusion. Proceedings of the 2020 9th International Conference on Software and Computer Applications, 2020. 114-118.

Pinterest, S. M. S. a. I. S. 2020. Pinterest, Mage Sharing and Social Media Service Designed to Enable Saving and Discovery of Information on the World Wide Web Using Images.

Qezavati, H., Majidi, B. and Manzuri, M. T. Partially Covered Face Detection in Presence of Headscarf for Surveillance Applications. 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA), 2019. IEEE, 195-199.

Ranjan, R., Patel, V. M. and Chellappa, R. 2017. Hyperface: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Ranjan, R., Sankaranarayanan, S., Bansal, A., Bodla, N., Chen, J.-C., Patel, V. M., Castillo, C. D. and Chellappa, R. 2018. Deep Learning for Understanding Faces: Machines May Be Just as Good, or Better, Than Humans. IEEE Signal Processing Magazine, 35**,** 66-83.

Rawat, W. and Wang, Z. 2017. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. Neural computation, 29**,** 2352-2449.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 779-788.

Redmon, J. and Farhadi, A. Yolo9000: Better, Faster, Stronger. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. 7263-7271.

Redmon, J. and Farhadi, A. 2018. Yolov3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.

Ren, S., He, K., Girshick, R. and Sun, J. Faster R-Cnn: Towards Real-Time Object Detection with Region Proposal Networks. Advances in neural information processing systems, 2015. 91-99.

Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I. and Savarese, S. Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019. 658-666.

Rodríguez, P., Cucurull, G., Gonfaus, J. M., Roca, F. X. and Gonzàlez, J. 2017. Age and Gender Recognition in the Wild with Deep Attention. Pattern Recognition, 72**,** 563-571.

Rosebrock, A. 2020. Detect and Remove Duplicate Images from a Dataset for Deep Learning. Pyimagesearch

Russell, B. C., Torralba, A., Murphy, K. P. and Freeman, W. T. 2008. Labelme: A Database and Web-Based Tool for Image Annotation. International journal of computer vision, 77**,** 157-173.

Sakai, T., Nagao, M. and Kanade, T. 1972. Computer Analysis and Classification of Photographs of Human Faces, Kyoto University.

Sapkota, A. 2013. Towards a Solution of Unconstrained Face Recognition. Ph.D. Thesis, University of Colorado Colorado Springs.

Schroff, F., Kalenichenko, D. and Philbin, J. Facenet: A Unified Embedding for Face Recognition and Clustering. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 815-823.

Scrapy, _Scrapy_Open_Source 2020. An Open Source and Collaborative Framework for Extracting the Data You Need from Websites.

Segui, S., Drozdzal, M., Radeva, P. and Vitria, J. 2012. An Integrated Approach to Contextual Face Detection.

Seo, J. and Park, H. Detection, Alignment, and Recognition of Partially Occluded Human Face Based on Statistical Learning of Local Features. Proceedings of the 9th International Conference on Machine Learning and Computing, 2017. 289-293.

Shafin, M., Hansda, R., Pallavi, E., Kumar, D., Bhattacharyya, S. and Kumar, S. Partial Face Recognition: A Survey. Proceedings of the Third International Conference on Advanced Informatics for Computing Research, 2019. 1-6.

Sharifara, A., Rahim, M. S. M. and Anisi, Y. A General Review of Human Face Detection Including a Study of Neural Networks and Haar Feature-Based Cascade Classifier in Face Detection. 2014 International Symposium on Biometrics and Security Technologies (ISBAST), 2014. IEEE, 73-78.

Shepley, A. J. 2019. Deep Learning for Face Recognition: A Critical Analysis. arXiv preprint arXiv:1907.12739.

Shorten, C. and Khoshgoftaar, T. M. 2019. A Survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6**,** 60.

Shuang, Y. 2018. Face Detection with Mobilenet-Ssd [Online]. GitHub. Available: https://github.com/bruceyang2012/Face-detection-with-mobilenet-ssd [Accessed December 10 2019].

Sikandar, T., Ghazali, K. H., Mohd, I. I. and Rabbi, M. Skin Color Pixel Classification for Face Detection with Hijab and Niqab. Proceedings of the International Conference on Imaging, Signal Processing and Communication, 2017. ACM, 1-4.

Sikandar, T., Samsudin, W. N. a. W., Rabbi, M. F. and Ghazali, K. H. 2020. An Efficient Method for Detecting Covered Face Scenarios in Atm Surveillance Camera.

Simonyan, K. and Zisserman, A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

Snapchat, S. I. 2021. Snap Inc. Is a Camera Company. https://www.snapchat.com.

Su, H., Deng, J. and Fei-Fei, L. Crowdsourcing Annotations for Visual Object Detection. Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence, 2012.

Sun, C., Shrivastava, A., Singh, S. and Gupta, A. Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. Proceedings of the IEEE international conference on computer vision, 2017. 843-852.

Systweak, S. I. 2020. Duplicate Photos Fixer Pro for Mac Os.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. Going Deeper with Convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 1-9.

Taigman, Y., Yang, M., Ranzato, M. A. and Wolf, L. Deepface: Closing the Gap to Human-Level Performance in Face Verification. Proceedings of the IEEE conference on computer vision and pattern recognition, 2014. 1701-1708.

Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J. N., Wu, Z. and Ding, X. 2020. Embracing Imperfect Datasets: A Review of Deep Learning Solutions for Medical Image Segmentation. Medical Image Analysis**,** 101693.

Tan, H., Yang, B. and Ma, Z. 2013. Face Recognition Based on the Fusion of Global and Local Hog Features of Face Images. IET computer vision, 8**,** 224-234.

Tang, X., Du, D. K., He, Z. and Liu, J. Pyramidbox: A Context-Assisted Single Shot Face Detector. Proceedings of the European Conference on Computer Vision (ECCV), 2018. 797-813.

Team_Inc., O. 2020. Duplicate Photos Finder for Mac Os.

Tikoo, S. and Malik, N. 2017. Detection of Face Using Viola Jones and Recognition Using Back Propagation Neural Network. arXiv preprint arXiv:1701.08257.

Tsang, R., Magalhaes, J. P. and Cavalcanti, G. D. Combined Adaboost and Gradientfaces for Face Detection under Illumination Problems. 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2012. IEEE, 2354-2358.

Tsao, D. Y. and Livingstone, M. S. 2008. Mechanisms of Face Perception. Annu. Rev. Neurosci., 31**,** 411-437.

Viola, P. and Jones, M. Rapid Object Detection Using a Boosted Cascade of Simple Features. Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001. IEEE, I-I.

Voulodimos, A., Doulamis, N., Doulamis, A. and Protopapadakis, E. 2018. Deep Learning for Computer Vision: A Brief Review. Computational intelligence and neuroscience, 2018.

Wang, J., Yuan, Y. and Yu, G. 2017. Face Attention Network: An Effective Face Detector for the Occluded Faces. arXiv preprint arXiv:1711.07246.

Wang, M. and Deng, W. 2018. Deep Face Recognition: A Survey. arXiv preprint arXiv:1804.06655.

Wang, N., Gao, X., Tao, D., Yang, H. and Li, X. 2018. Facial Feature Point Detection: A Comprehensive Survey. Neurocomputing, 275**,** 50-65.

Wang, X., Han, T. X. and Yan, S. An Hog-Lbp Human Detector with Partial Occlusion Handling. 2009 IEEE 12th international conference on computer vision, 2009a. IEEE, 32-39.

Wang, X., Han, T. X. and Yan, S. An Hog-Lbp Human Detector with Partial Occlusion Handling. Computer Vision, 2009 IEEE 12th International Conference on, 2009b. IEEE, 32-39.

Wang, X., Wang, K. and Lian, S. 2019. A Survey on Face Data Augmentation. arXiv preprint arXiv:1904.11685.

Weiss, K., Khoshgoftaar, T. M. and Wang, D. 2016. A Survey of Transfer Learning. Journal of Big data, 3**,** 9.

Weng, L. 2017. Object Detection for Dummies Part 3: R-Cnn Family. ilianweng.github.io/lil-log.

Weng, L. 2018. Object Detection Part 4: Fast Detection Models. lilianweng.github.io/lil-log.

Wu, Y. and Ji, Q. 2019. Facial Landmark Detection: A Literature Survey. International Journal of Computer Vision, 127**,** 115-142.

Wu, Y., Shah, S. K. and Kakadiaris, I. A. 2018. Godp: Globally Optimized Dual Pathway Deep Network Architecture for Facial Landmark Localization in-the-Wild. Image and Vision Computing, 73**,** 1-16.

Xia, Y., Zhang, B. and Coenen, F. 2016. Face Occlusion Detection Using Deep Convolutional Neural Networks. International Journal of Pattern Recognition and Artificial Intelligence, 30**,** 1660010.

Yang, B., Yan, J., Lei, Z. and Li, S. Z. Fine-Grained Evaluation on Face Detection in the Wild. 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2015. IEEE, 1-7.

Yang, M.-H., Kriegman, D. J. and Ahuja, N. 2002. Detecting Faces in Images: A Survey. IEEE Transactions on pattern analysis and machine intelligence, 24**,** 34-58.

Yang, S., Luo, P., Loy, C.-C. and Tang, X. Wider Face: A Face Detection Benchmark. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016a. 5525-5533.

Yang, S., Luo, P., Loy, C. and Tang, X. 2018. Faceness-Net: Face Detection through Deep Facial Part Responses. IEEE transactions on pattern analysis and machine intelligence, 40**,** 1845.

Yang, S., Wiliem, A. and Lovell, B. C. To Face or Not to Face: Towards Reducing False Positive of Face Detection. Image and Vision Computing New Zealand (IVCNZ), 2016 International Conference on, 2016b. IEEE, 1-6.

Youtube, O. V.-S. P. F. G. 2005. Youtube, Online Video-Sharing Platform.

Yu, F., Seff, A., Zhang, Y., Song, S., Funkhouser, T. and Xiao, J. 2015. Lsun: Construction of a Large-Scale Image Dataset Using Deep Learning with Humans in the Loop. arXiv preprint arXiv:1506.03365.

Yuan Yuan, M. S. 2019. Smaller Is Better: Lightweight Face Detection for Smartphones. China's leading media & information provider for #AI & #MachineLearning.

Zafeiriou, S., Zhang, C. and Zhang, Z. 2015. A Survey on Face Detection in the Wild: Past, Present and Future. Computer Vision and Image Understanding, 138**,** 1-24.

Zaid, G., Bossuet, L., Habrard, A. and Venelli, A. 2020. Methodology for Efficient Cnn Architectures in Profiling Attacks. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2020**,** 1-36.

Zempi, I. 2016. ‘It's a Part of Me, I Feel Naked without It': Choice, Agency and Identity for Muslim Women Who Wear the Niqab. Ethnic and Racial Studies, 39**,** 1738-1754.

Zhang, C. and Zhang, Z. 2010. A Survey of Recent Advances in Face Detection. Tech. rep., Microsoft Research.

Zhang, K., Zhang, Z., Li, Z. and Qiao, Y. 2016. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters, 23**,** 1499-1503.

Zhang, S., Wen, L., Shi, H., Lei, Z., Lyu, S. and Li, S. Z. 2019a. Single-Shot Scale-Aware Network for Real-Time Face Detection. International Journal of Computer Vision**,** 1-23.

Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X. and Li, S. Z. S3fd: Single Shot Scale-Invariant Face Detector. Proceedings of the IEEE International Conference on Computer Vision, 2017. 192-201.

Zhang, Y., Xu, X. and Liu, X. 2019b. Robust and High Performance Face Detector. arXiv preprint arXiv:1901.02350.

Zhang, Z., Luo, P., Loy, C. C. and Tang, X. Facial Landmark Detection by Deep Multi-Task Learning. European conference on computer vision, 2014. Springer, 94-108.

Zhao, Z.-Q., Zheng, P., Xu, S.-T. and Wu, X. 2019. Object Detection with Deep Learning: A Review. IEEE transactions on neural networks and learning systems, 30**,** 3212-3232.

Zhu, C., Zheng, Y., Luu, K. and Savvides, M. 2017. Cms-Rcnn: Contextual Multi-Scale Region-Based Cnn for Unconstrained Face Detection. Deep Learning for Biometrics. Springer.

Zhu, M., Shi, D., Zheng, M. and Sadiq, M. Robust Facial Landmark Detection Via Occlusion-Adaptive Deep Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019. 3486-3496.

Zhu, X. and Ramanan, D. Face Detection, Pose Estimation, and Landmark Localization in the Wild. Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012. IEEE, 2879-2886.

Zou, X., Kittler, J. and Messer, K. Illumination Invariant Face Recognition: A Survey. 2007 first IEEE international conference on biometrics: theory, applications, and systems, 2007. IEEE, 1-8.

LIST OF PUBLICATIONS

Indexed Journal

1. **Alashbi, A. A. S**., Sunar, M. S., and Alqahtani, Z. (2022). Deep-Learning-CNN for Detecting Covered Faces with Niqab. *In International Journal of Information Technology Management JITM)* ,Special Issue, 114-123 **(Indexed by SCOPUS)**.

Indexed Conference Proceedings

1. **Alashbi, A. A. S.,** Sunar, M. S., and Alqahtani, Z. (2020, December). Context-Aware Face Detection for Occluded Faces. *In 2020 6th International Conference on Interactive Digital Media (ICIDM)* (pp. 1-4). IEEE. https://doi.org/DOI: 10.1109/ICIDM51048.2020.9339647. **(Indexed by SCOPUS)**

2. **Alashbi, A. A. S**., and Sunar, M. S. (2019). Occluded Face Detection, Face in Niqab Dataset. *In International Conference of Reliable Information and Communication Technology*, (pp.209-215). Springer, Cham. https://doi.org/DOI/10.1007/978-3-030-33582-3_20. **(Indexed by SCOPUS)**

3. **Alashbi, A. A. S.**, Sunar, M. S. B., and AL-Nuzaili, Q. A. (2018). Two stages haar-cascade face detection with reduced false positive. *In International Conference of Reliable Information and Communication Technology (pp. 690-695). Springer, Cham.* https://doi.org/10.1007/978-3-319-99007-1_64. **(Indexed by SCOPUS)**

4. Zieb Alqahtani, **Alashbi, A. A. S**., and Sunar, M. S. (2020). Landmark Localization in Occluded, *In International Conference of Reliable Information and Communication Technology (IRICT, 2020)* (in press).

Non-indexed conference

1. **Alashbi, A. A. S.**, and Sunar, M. S. (2019). Face detection for covered faces, detecting a face in niqab. On universal wellbeing (issue 2019), 92.