Does fusion of complementary spectral bands improves the cross-illumination on the performance of gender prediction?

Narayan Vetrekar† Marissa de Ataide† Krishna Patel† Raghavendra Ramachandra‡ R. S. Gad†
†*School of Physical and Applied Sciences, Goa University, Goa, India*
‡*Norwegian University of Science and Technology (NTNU), Gjøvik, Norway*
E-mail: {vetrekarnarayan;rsgad}@unigoa.ac.in, {raghavendra.ramachandra}@ntnu.no

***Abstract*—The automatic prediction of gender from the face has been studied extensively because of its potential relevance in numerous applications related to security. Although the problem of gender classification based on the face is substantial, it remains far from being solved under difficult environmental exposure, especially for different illuminations. In this work, we demonstrate the merits and demerits of classifying gender under cross-illumination variants. We present our approach by employing multi-spectral imaging in nine narrow-spectrum bands stemming from the visible to near-infrared range. The experimental evaluation results were obtained on 78300 sample face images of 145 subjects captured under six different illumination conditions. Further, we present quantitative and qualitative experimental evaluations to determine the average classification accuracy for setting the benchmark results. To demonstrate the goal of this work, we present the results based on three image fusion techniques independently processed using five feature extraction methods for cross-illumination scenarios. This work obtained the highest classification accuracy of 96.32% for cross-illumination conditions, demonstrating the reliability of employing an image fusion approach to combine complementary information from spectral bands in difficult environmental exposure.**

***Index Terms*—Face Biometrics, Multi-spectral Imaging, Cross-illumination, Fusion, Gender Classification**

I. Introduction

Soft biometric traits such as gender, age, height, and weight are based on human physiological and behavioral characteristics, which provide discriminative features that can be used to identify gender (i.e., male or female) [1]. Being a permanent and stable soft biometric trait, gender classification is useful in many real-life applications such as social media, security, advertisements, entertainment, social interactions, human-computer interaction communication, demographic research, and computer vision [2]; hence, a successful gender classification approach can provide a boost to all these applications. The problem of predicting gender is a two-class problem, ideally used in database management to partition a larger set of databases into two groups by representing them using soft labels [3]. In surveillance and security, it finds major applications to reduce the search space in security databases and improve the verification accuracy of biometric systems [4].

Different methods can be used to identify the gender of an individual based on their physiological and behavioral traits. Physiological traits include fingerprints, the retina, iris, face, and hand geometry. Behavioral traits are the features of an individual, including walking style, writing style, signature style, typing speed, and mouse usage. Physiological traits are more easily observed [5] than the behavioral traits of an individual, and hence find more applications for many security-related purposes. In particular, the human face [6] is the most common trait that can be easily acquired to provide distinct information about gender as well as other soft information such as expression, age, and ethnicity for security and surveillance purposes, while also providing seamless verification.

Apparently, the classification of gender based on facial biometric traits has been well studied towards the improved performance of face recognition systems by using gender as additional complementary information. Specifically, under surveillance conditions, a sample face image is partially occluded or acquired under non-standard poses [3]. Moreover, compared with other biometric traits, face acquisition is non-intrusive and does not involve human cooperation. However, it also renders the privilege to acquire face images in a varying standing position and in a covert manner [7].

Facial features are stable and reliable but challenging [8] owing to illumination variations, large dimensionality, uncontrolled environments, aging, motion blur, face alteration by plastic surgery, presentation attack, occlusion, etc. There are also instances in which individuals try to falsify their gender identity. Hence, gender detection before recognition is considered an important phenomenon for improved performance analysis. With conventional surveillance systems operating in the visible spectrum, the majority of gender classification works are in the visible spectrum, perform well under controlled environmental conditions, and suffer significantly when the sensor is exposed to difficult environmental conditions. For example, the visible spectrum can explicitly provide only reflectance information, and its robustness can be challenging under different illumination conditions. Therefore, to mitigate these limitations of the visible spectrum, the advancement of sensing technology has led to the use of a spectrum beyond the visible (VIS) spectrum, mainly in near-infrared (NIR), Thermal, and Multispectral imaging in nine narrow bands of the VIS-NIR range, to acquire complementary information for

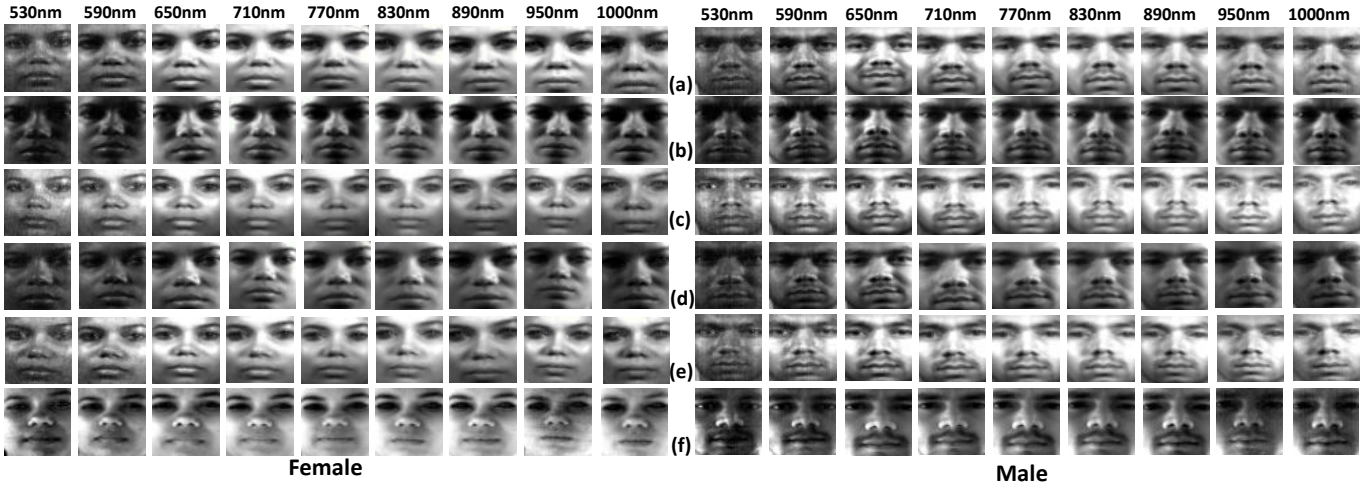


Fig. 1: Sample images from multi-spectral sensor collected in nine bands across visible and near-infra-red spectrum. Database collected in six different illuminations:(a) QTH-QTH, (b) Incad-Incad, (c) Xenon-Xenon, (d) QTH-Incad, (e) QTH-Xenon, and (f) Outdoor

a reliable gender classification approach.

A. Related Work

Most of the earlier works based on gender classification are limited to the visible spectrum, which is performed using handcrafted features, dictionary-based descriptors, and support vector machines (SVM) as a classifier. Furthermore, a deep learning-based approach was employed on a large-scale visible spectrum database to predict gender from facial images [9] [4], which are computationally complex. Performance in the visible spectrum degrades because of the difficulties associated with the environment, such as uncontrolled lighting conditions, which forces researchers to obtain data in a controlled environment [10]. The InfraRed spectrum promises better accuracy, as the images captured in these spectra are less impacted by ambiance and are invariant to illumination variations. Considering this, previous work in this direction has collected data in multispectral imaging using two to three broad spectra, visible, near-infrared, and thermal spectra for gender classification. The significance of these approaches was further analyzed and demonstrated using texture descriptor techniques [11] [12]. In addition, recent work has considered the use of multispectral imaging in nine narrow spectrum bands across the visible and near-infrared ranges to investigate the use of complementary information from the spatial and spectral domains for the performance analysis [13].

Previous studies to predict gender are more limited to the visible spectrum, in which system performance degrades when samples are presented with unknown or uncontrolled environmental conditions. However, there are few studies reported in the literature that use the spectrum beyond the visible spectrum, mainly multispectral imaging employing two to three broad spectra (i.e., VIS, NIR and Thermal spectra) or in nine narrow bands to leverage the complementary information for gender classification. Although substantial work is evident from earlier studies, automatic gender classification remains a challenging task under cross-illumination conditions, which certainly presents the generalizability of the classification algorithm. More importantly, the vulnerability of gender classification when probe samples are compared with enrollment samples belongs to different illumination conditions.

Therefore, in this work, we present extensive benchmark results employing cross-illumination protocols for gender classification owing to a generalized approach. We used a multispectral imaging sensor to acquire unique spatio-spectral imaging features across nine spectral bands for gender prediction, thereby exploring the photometric reflectance properties of the human skin [14] [15]. These photometric reflectance properties vary across genders owing to the presence of varying concentrations of chromophores present on the sub-surface of the human skin, and these discriminative features can be extracted in the form of complementary details (reflectance and/or emittance) across various bands of the electromagnetic spectrum.

Now, multi-spectral imaging can solve the problem pertaining to different illumination conditions, which results from the intrinsic difference in the spectral distribution of faces, making it invariant to illumination changes [16]. However, owing to the difference in reflectance properties of individuals across the spectral bands, comparison of bands is always challenging, and as a result, some bands may not provide the discriminative attributes necessary for gender prediction. Hence, in this study, we combined all spectral bands using three different fusion techniques: Image Matting Fusion (IMF) [17], guided filtering-based fusion (GFF) [18] and 2-Discrete Wavelet Transform Average Fusion (2-DWT) [19], independently for the performance analysis of our cross-illumination-based gender prediction.

In this work, we perform gender classification based on a multispectral face database of 78300 samples belonging to 145

subjects acquired under six different illumination conditions. We performed gender classification for the cross-illumination framework independently on five feature extraction methods followed by a Support Vector Machine Classifier. In this benchmark study, we present the experimental evaluation results using 10-fold cross-validation approach for the random selection of training and testing sets in a disjoint manner to compute the average classification accuracy. The major contributions of this work are summarized as follows.

- Present gender classification by exploring inherent properties of multi-spectral imaging to acquire the complimentary spatial and spectral face information stemming from Visible to Near InfraRed spectrum ($530nm$ to $1000nm$).
- Experimenting three different image fusion approach to combine complimentary spectral bands information for cross-illumination based gender prediction, an approach towards the generalisation.
- Extensive qualitative and quantitative benchmark results of gender prediction based on five different feature extraction methods using 10 fold cross-validation approach to present the merits and demerits of performing band fusion for cross-illumination evaluations.

The rest of the paper is organised as follows: section II provide the details of multispectral face database in six different illumination condition. This section also includes details of data collection and pre-processing approach employed. Section III consist of detailed methodology used for gender classification under cross-illumination condition. Section IV presents the experiments related to gender classification across three different fusion techniques and five feature extraction methods and Section VI summarizes final conclusive remarks.

TABLE I: Summarizes the number of samples acquired under six different illumination conditions

Database	Illumination Conditions					
	QTH-QTH	Incad-Incad	Xenon-Xenon	QTH-Incad	QTH-Xenon	Outdoor
Subject	145	145	145	145	145	145
Sessions	2	2	2	2	2	2
Samples	5	5	5	5	5	5
Bands	9	9	9	9	9	9
Total	**78300**					

II. Multi-spectral Face Database

This section explains the details of the multispectral face database employed in this study for gender classification. Facial data were collected using a custom-built imaging sensor [19] that can capture images in nine narrow spectral bands, which include $530nm$, $590nm$, $650nm$, $710nm$, $770nm$, $830nm$, $890nm$, $950nm$, and $1000nm$ belonging to the visible (VIS) and near-infrared (NIR) spectra. Data from 145 subjects were captured, comprising 87 males and 58 females.

A. Data Collection

Using a multispectral imaging sensor, the face database was collected under six different illumination conditions in indoor and outdoor environments. For indoor conditions, five different illumination conditions were created using three light sources: Quartz Tungsten Halogen (QTH), incandescent (Incad), and xenon. These five different illumination conditions were created either by operating the light sources alone or by combining them. For simplicity, the five different illumination conditions created using these three light sources are given notations in this study: $QTH - QTH$, $Incad - Incad$, $Xenon - Xenon$, $QTH - Incad$, and $QTH - Xenon$. A brief description of these notations is presented in Table II. The outdoor data collection was based on natural day lighting conditions, and to maintain simplicity throughout the experiment, this set of data collection was given $Outdoor$ notation (Refer Table II). Two sessions were conducted for data acquisition with a time span of 3–4 weeks, and five sample face images were captured for each band under each illumination condition. Thus, we created a database consisting of 145 subjects $\times$ 2 sessions $\times$ 9 bands $\times$ 6 illumination conditions $\times$ 5 samples = 78300 sample face images. Table I summarizes the number of samples acquired under the six different illumination conditions.

TABLE II: Notations used to represents the six different illumination conditions of multi-spectral face database

Condition	Notations	Details
Indoor	**QTH-QTH**	Illumination from two Quartz Tungsten Halogen light source
	Incad-Incad	Illumination from two Incandecent light source
	Xenon-Xenon	Illuminations from two Xenon light source
	QTH-Incad	Illuminations from a Quartz Tungsten Halogen and a Incandecent light source
	QTH-Xenon	Illuminations from a Quartz Tungsten Halogen and Xenon light source
Outdoor	**Outdoor**	Illuminations from varying daylight

B. Pre-Processing

The acquired samples were high-dimensional images of 1024×1280 pixels captured from a custom-built sensor [19]. These images contain an unwanted background, apart from facial information. To remove unwanted background details, we used an eye-coordinate-based face detection algorithm to crop the facial region. This algorithm automatically detects the facial region based on left- and right-eye coordinates [20]. We then used histogram normalization to enhance the contrast of the cropped images for all the bands. We also resized the images to 120×120 spatial resolution to reduce the computational cost and the time required for gender classification. Figure 1 shows the sample spectral band face images of male and female subjects after preprocessing.

III. Methodology: Band fusion, Feature Extraction and Classification

This section presents a detailed description of the methodology used in this work. The study carried out gender classification on a multispectral face database by performing spectral band fusion for cross-illumination evaluation. Therefore, the scope of this study is to perform an evaluation in which the training and testing samples belong to different illumination levels. Specifically, to present the generalization of our approach, we train the fused spectral band images belonging to

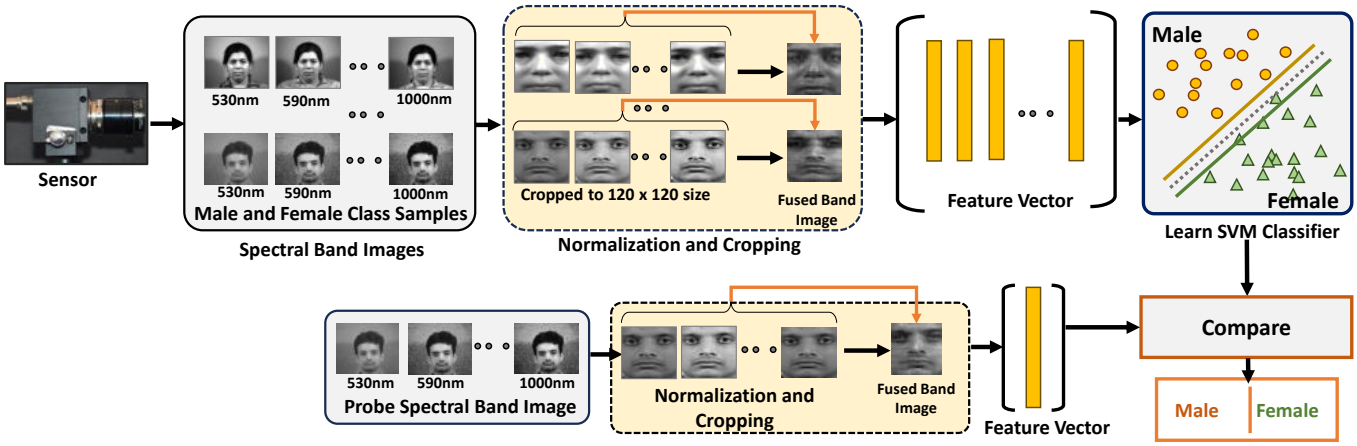


Fig. 2: Gender classification framework employed in this work for cross-illumination evaluation protocol

one type of illumination (say $QTH - QTH$) and perform the testing using the fused spectral band images belonging to different illuminations (say $Outdoor$). A block diagram of this study for evaluating the gender classification is shown in Figure 2.

Let the set of spectral band images $H_\lambda(s,r)$ acquired using a custom-built multispectral imaging sensor be expressed by Equation 1.

$$H_\lambda(s,r) = \{H_1(s,r), H_2(s,r), \ldots, H_9(s,r)\} \quad (1)$$

where $H_\lambda \in \mathbb{R}^{s\times r}$ consists of nine narrow spectral band images, that is, $\lambda = \{1,2,3,\ldots,9\}$ and each image has $s \times r$ spatial dimension that corresponds to 120×120 size. Considering the differential reflectance properties of individual face images across different illumination conditions, we perform the spectral band fusion approach to form a composite image from the complementary spectral band images. The composite image formed after band fusion not only reduces the dimensions, but also preserves the discriminate information required for cross-illumination-based gender classification. We employ three image fusion techniques: image matching fusion (IMF) [17], guided filtering-based fusion (GFF) [18] and 2-Discrete Wavelet Transform Average Fusion (2-DWT) [19]. Therefore, spectral band images (Equation 1) were preprocessed, and image fusion algorithms were employed to obtain the composite image represented as $U_{fus}(s,r)$ in Equation 2.

$$U_{fus}(s,r) = \{\omega_1 * H_1 + \omega_2 * H_2 + \ldots + \omega_9 * H_9\} \quad (2)$$

where, $\omega_1, \omega_2, \ldots, \omega_9$ are the weight vectors. Furthermore, we extracted dominant features from the fused spectral band images using texture descriptor methods. In this work, we employed five texture descriptor methods that include Local Phase Quantization (LPQ), Histogram of Oriented Gradients (HOG), GIST, logGabor (LG), and binarized statistical image features (BSIF) independently before learning the classifier model. Further, to learn the classifier model, the histogram feature vectors obtained from the texture descriptor method were then processed to train the two-class model. Let K denote the number of training samples, and let the training feature matrix employed to learn the classifier model be represented using Equation 3 as follows:

$$T_d = \left\{\left(J^i_\lambda, \phi^i\right)\right\}_{i=1}^{K} \quad (3)$$

where T_d represents the training feature matrix after processing with the feature extraction method and the variable d denotes the illumination variation belonging to six different illumination conditions, that is, $d = 1,2,3,\ldots,6$. Further, J represents the individual training vectors belonging to the spectral bands, and ϕ denotes the class information that can be represented as follows:

$$\phi(v) = \begin{cases} v = 0 & \in \text{ Male} \\ v = 1 & \in \text{ Female} \end{cases} \quad (4)$$

where 0 denotes the male class label, and 1 denotes the female class label. In this study, we employed a linear Support Vector Machine (SVM) classifier to learn the features and perform gender classification on fused spectral band images [13]. The SVM divides the data points of the two classes in the feature space and finds the optimal hyperplane that divides these two classes into male and female. The margin, which is the distance between the closest data points of both classes, determines the optimum hyperplane, and can be represented using Equation 5 as follows:

$$\phi(v) = \psi^{'} \cdot \sigma(v) + b \quad (5)$$

where $\phi(v)$ represent the gender class labels, ψ denotes the learning weights, b indicates the bias and $\sigma(v)$ presents the kernel function. Now for the given probe fused spectral band images belongs to different illumination other than used in training set, processed to compare with each class and based on threshold values, the probe sample is classified as male or female class label.

IV. Experiments

This section of the paper is based on an experimental evaluation protocol and related experimental results. In this work, we perform gender classification on fused spectral band images of a multispectral face database (refer to Section II) in which training and testing samples are considered from different illumination conditions. Thus, the experimental classification results are based on analyzing the influence of employing cross-illumination frameworks on fused spectral band images for the performance analysis of gender classification. We present an extensive and quantitative experimental benchmark evaluation of 78300 sample images from a multispectral face database that belongs to 145 subjects and nine bands. To demonstrate the objective of this work, we present the results independently using three different fusion methods and five different feature extraction algorithms, along with a linear Support Vector Machine (SVM) classifier. Furthermore, to obtain the average classification accuracy along with the standard deviation, the experiments were repeated using 10-fold cross-validation approach, thereby randomly selecting the training and testing samples for performance analysis with statistical significance.

A. Experimental Evaluation Protocol

Considering the scope of this work and to present the performance analysis of the gender classification algorithm, we present an experimental evaluation protocol that partitions face multispectral data into training and testing sets. Based on the database samples, the training set consisted of 50% samples belonging to the male class and 50% samples from the female class. Hence, the training set comprised 40 male $\times$ 2 sessions $\times$ 5 samples + 40 female $\times$ 2 sessions $\times$ 5 samples = 800 images. The testing set consisted of the remaining samples belonging to the male and female classes, and comprised 47 male $\times$ 2 sessions $\times$ 5 samples + 18 female $\times$ 2 sessions $\times$ 5 samples = 650 images.

V. Results and Discussion

In this section, we present the performance of our approach in performing cross-illumination-based gender classification. Figure 3 illustrates the mean-variance plot illustrating the performance analysis of cross-illumination for gender classification. Figure 4 and Figure 5 present the average classification accuracy across the three different fusion methods and across five different feature extraction methods. Furthermore, quantitative experimental results in the form of the average classification accuracy obtained after 10-fold cross validations are summarized in Tables III, IV, V, VI, VII and VIII.

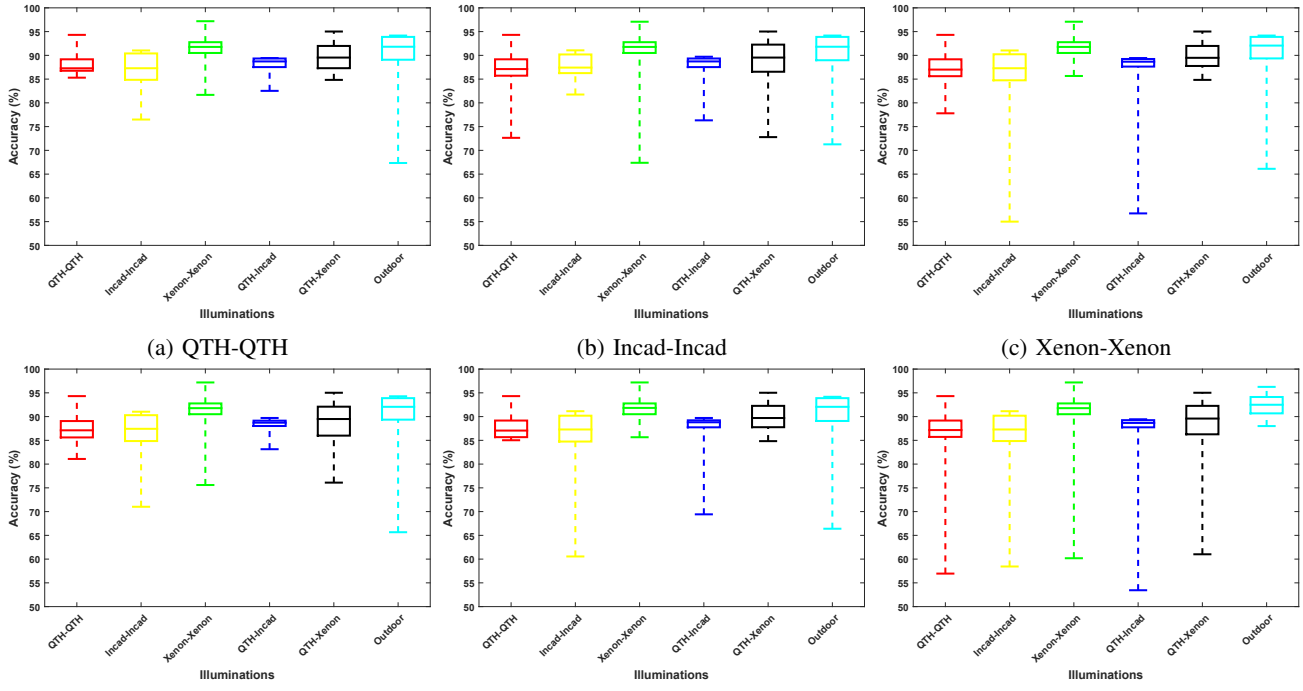
The purpose of conducting a cross-illumination-based gender classification experiment was to demonstrate a real-life surveillance scenario in which the samples were trained under controlled lighting conditions and tested against different illumination conditions for gender prediction. Therefore, in this section, we present an evaluation of the training and testing samples under different illumination conditions. The multi-spectral face database was acquired under six different illumination conditions; therefore, considering the scope of this work for cross-illumination performance analysis, we have presented the results independently across all the illumination conditions when trained and tested against different illuminations in cross-illumination-based gender classification. For instance, fused spectral band faces image samples is trained with $QTH - QTH$ illumination and then fused spectral band face image samples tested independently from different illuminations to present the benchmark results. The procedure was repeated for the remaining illumination conditions when trained and tested under different illumination conditions. Based on our quantitative results, the spectral band fusion approach for cross-illumination-based gender classification provides reasonable performance accuracy. A summary of the major observations related to the experimental results is as follows.

TABLE III: Average classification accuracy (%) when trained with $QTH - QTH$ illuminations

Testing	Training: Samples from QTH-QTH Illuminations					
	Fusion	LPQ-SVM	HOG-SVM	GIST-SVM	LG-SVM	BSIF-SVM
QTH-QTH	IFM	69.19	79.70	82.07	88.51	83.96
	GFF	63.35	77.37	77.47	87.96	80.51
	Wavelet	81.38	96.02	93.27	96.91	93.43
Incad-Incad	IFM	67.63	74.17	83.36	86.55	81.49
	GFF	68.51	76.13	78.87	86.46	77.53
	Wavelet	79.23	88.76	93.18	93.56	91.53
Xenon-Xenon	IFM	69.38	75.53	86.19	89.81	81.58
	GFF	69.88	77.99	87.44	90.96	81.70
	Wavelet	78.46	93.47	91.22	93.44	91.64
QTH-Incad	IFM	63.20	74.92	82.95	85.96	79.73
	GFF	63.00	77.32	81.14	87.63	80.53
	Wavelet	82.79	91.27	92.52	93.61	91.35
QTH-Xenon	IFM	66.03	73.30	87.17	88.52	83.45
	GFF	69.81	78.77	85.94	89.75	82.54
	Wavelet	81.95	95.28	92.91	95.83	94.16
Outdoor	IFM	78.71	85.53	89.66	89.41	86.49
	GFF	77.38	83.58	86.16	89.45	84.00
	Wavelet	84.20	89.23	90.66	92.47	89.59

TABLE IV: Average classification accuracy (%) when trained with $Incad - Incad$ illuminations

Testing	Training: Samples from Incad-Incad Illuminations					
	Fusion	LPQ-SVM	HOG-SVM	GIST-SVM	LG-SVM	BSIF-SVM
QTH-QTH	IFM	67.38	77.50	79.58	87.07	81.24
	GFF	61.66	76.02	75.71	86.50	77.52
	Wavelet	77.74	91.89	92.19	95.61	90.00
Incad-Incad	IFM	67.98	74.79	83.81	88.65	82.34
	GFF	69.34	76.71	79.38	87.58	78.80
	Wavelet	82.30	90.99	94.78	96.53	93.40
Xenon-Xenon	IFM	68.94	75.51	84.41	89.15	80.22
	GFF	68.92	77.19	86.67	89.52	80.73
	Wavelet	77.76	91.72	90.80	92.77	89.43
QTH-Incad	IFM	63.07	73.98	82.16	86.17	78.85
	GFF	62.42	76.43	80.46	87.04	79.19
	Wavelet	82.17	88.49	89.59	95.42	90.43
QTH-Xenon	IFM	64.67	72.47	85.49	87.20	81.54
	GFF	68.25	77.69	85.02	88.34	80.31
	Wavelet	80.39	92.42	93.29	94.97	91.16
Outdoor	IFM	78.52	85.15	88.14	89.78	85.86
	GFF	77.84	83.31	86.26	89.78	83.27
	Wavelet	85.30	88.94	90.05	89.78	90.16



(a) QTH-QTH (b) Incad-Incad (c) Xenon-Xenon

(d) QTH-Incad (e) QTH-Xenon (f) Outdoor

Fig. 3: Mean and variance plot represents the cross-illumination performance for gender classification: (a) illustrates the results when the samples from $QTH - QTH$ is in training set and tested independently with $QTH - QTH$, $Incad - Incad$, $Xenon - Xenon$, $QTH - Incad$, $QTH - Xenon$ and $Outdoor$. Similar illustrations can be drown for (b,c,d,e,f). For simplicity, the performance of GFF fusion and LG-SVM method is shown (Results are best viewed in colors).

TABLE V: Average classification accuracy (%) when trained with $Xenon - Xenon$ illuminations

Testing	Training: Samples from Xenon-Xenon Illuminations					
	Fusion	LPQ-SVM	HOG-SVM	GIST-SVM	LG-SVM	BSIF-SVM
QTH-QTH	IFM	67.96	79.25	81.59	88.42	82.74
	GFF	62.60	77.11	77.37	86.98	79.25
	Wavelet	79.81	95.98	91.33	96.32	91.83
Incad-Incad	IFM	66.70	73.72	82.87	86.10	79.99
	GfF	67.93	75.56	78.79	84.28	76.25
	Wavelet	79.03	87.18	91.88	94.70	89.24
Xenon-Xenon	IFM	69.87	77.41	86.63	90.93	83.52
	GFF	70.75	79.28	88.42	91.73	83.93
	Wavelet	80.92	96.33	93.46	95.80	93.62
QTH-Incad	IFM	63.12	74.36	82.51	84.93	77.65
	GFF	62.99	85.05	81.38	85.05	78.63
	Wavelet	82.34	88.41	90.32	91.76	89.67
QTH-Xenon	IFM	65.72	74.30	87.13	88.99	84.33
	GFF	69.73	76.70	86.36	89.83	82.98
	Wavelet	83.52	97.18	94.09	96.67	94.78
Outdoor	IFM	78.37	86.68	89.44	89.75	85.50
	GFF	76.88	84.88	85.98	89.34	83.13
	Wavelet	83.99	90.37	89.36	93.41	89.20

TABLE VI: Average classification accuracy (%) when trained with $QTH - Incad$ illuminations

Testing	Training: Samples from QTH-Incad Illuminations					
	Fusion	LPQ-SVM	HOG-SVM	GIST-SVM	LG-SVM	BSIF-SVM
QTH-QTH	IFM	68.16	79.22	81.84	87.33	83.16
	GFF	62.40	77.32	76.95	87.33	79.52
	Wavelet	80.64	94.64	92.81	97.15	83.67
Incad-Incad	IFM	66.42	74.26	83.52	87.20	81.09
	GFF	67.51	76.29	78.59	85.93	77.83
	Wavelet	81.06	90.75	94.51	95.70	92.64
Xenon-Xenon	IFM	69.74	76.16	85.77	89.38	81.06
	GFF	70.04	77.93	87.81	90.35	81.14
	Wavelet	78.32	92.14	92.04	95.41	90.64
QTH-Incad	IFM	63.98	75.03	82.75	86.93	79.88
	GFF	63.42	77.88	81.64	88.08	80.72
	Wavelet	85.30	90.94	92.76	95.69	92.79
QTH-Xenon	IFM	65.38	74.00	87.16	87.45	82.63
	GFF	69.29	79.14	85.79	88.59	81.40
	Wavelet	81.58	93.19	93.26	96.92	92.88
Outdoor	IFM	78.34	85.53	87.84	88.61	86.23
	GFF	76.99	83.87	85.27	89.32	83.67
	Wavelet	84.49	88.97	90.43	92.60	89.69

- From the tabular representation, it is evident that the gender classification accuracy based on cross-illumination is slightly poor compared to the performance when the same illumination conditions sample data in the training and testing sets. However, this decrease in performance is marginal (Figure 3), which mainly results from employing the spectral band fusion approach to combine the information in the form of a composite image from the complementary information across the spectral bands spanning the VIS to NIR wavelength range. For instance, when the training and testing datasets belong to the same illumination condition, the highest classification accuracy

TABLE VII: Average classification accuracy (%) when trained with $QTH - Xenon$ illuminations

Testing	Training: Samples from QTH-Xenon Illuminations					
	Fusion	LPQ-SVM	HOG-SVM	GIST-SVM	LG-SVM	BSIF-SVM
QTH-QTH	IFM	68.06	79.46	82.05	88.56	76.24
	GFF	63.12	77.29	77.37	87.70	79.64
	Wavelet	79.89	96.13	92.77	97.11	92.64
Incad-Incad	IFM	66.96	73.54	82.88	86.51	80.31
	GFF	67.74	75.92	77.94	84.88	76.24
	Wavelet	78.78	87.31	93.20	93.92	89.40
Xenon-Xenon	IFM	70.31	76.60	86.94	91.01	83.12
	GFF	70.80	78.79	88.19	91.55	83.27
	Wavelet	79.80	95.64	92.83	95.73	93.62
QTH-Incad	IFM	63.21	74.19	82.98	85.88	78.58
	GFF	63.15	76.91	81.27	86.37	79.30
	Wavelet	83.41	88.95	91.92	95.07	90.69
QTH-Xenon	IFM	66.05	73.60	87.65	89.30	84.46
	GFF	70.09	79.40	86.39	89.92	83.25
	Wavelet	83.72	97.05	94.67	97.03	95.19
Outdoor	IFM	78.26	86.77	88.94	89.70	85.43
	GFF	76.76	84.86	86.68	89.34	83.06
	Wavelet	84.11	89.10	89.59	92.25	89.27

TABLE VIII: Average classification accuracy (%) when trained with $Outdoor$ illuminations

Testing	Training: Samples from Outdoor Illuminations					
	Fusion	LPQ-SVM	HOG-SVM	GIST-SVM	LG-SVM	BSIF-SVM
QTH-QTH	IFM	67.76	77.71	81.07	85.85	80.66
	GFF	62.10	75.67	76.10	84.94	77.31
	Wavelet	77.88	92.27	91.97	93.46	89.06
Incad-Incad	IFM	67.05	73.57	83.46	85.48	80.92
	GFF	67.59	75.91	78.93	84.68	76.96
	Wavelet	78.91	89.37	91.96	93.46	90.06
Xenon-Xenon	IFM	69.01	75.80	85.32	87.68	80.52
	GFF	69.26	77.64	86.89	88.81	80.82
	Wavelet	77.80	93.94	91.49	94.02	89.50
QTH-Incad	IFM	63.27	73.19	82.31	83.57	77.80
	GFF	62.86	76.00	81.69	84.75	78.20
	Wavelet	82.20	89.41	89.74	92.55	90.18
QTH-Xenon	IFM	65.05	72.98	86.39	86.12	81.39
	GFF	68.47	77.79	85.03	87.15	80.58
	Wavelet	80.44	94.22	92.82	94.59	90.72
Outdoor	IFM	81.35	88.38	90.86	92.12	89.44
	GFF	79.89	86.12	88.40	92.34	87.26
	Wavelet	87.86	93.24	92.01	95.54	93.93

obtained was 97.05%. However, when the training and testing datasets belong to different illumination conditions, the highest average classification accuracy obtained is 96.32% which signifies the potential of employing spectral band fusion for gender classification in cross-illumination evaluation.

- Specifically, the performance under the five indoor artificial illumination conditions was better than that under outdoor illumination conditions. The possible reason for the better performance is mainly the controlled data acquisition condition adapted to all indoor illumination conditions, even though the illumination sources operated are different. It was further inferred that the spectral reflectance properties of the three illumination sources employed in this work have better reflectance properties across the visible and near-infrared spectra.
- Among the illumination conditions employed in this work, when fused spectral bands from $QTH - Xenon$ illumination were tested against $Xenon - Xenon$ training set, a better average classification accuracy of 97.18% was obtained. This could be due to the presence of one $Xenon$ illumination in the testing set, which is also a common illumination in the training dataset, as compared to totally different illumination conditions in training and testing. Therefore, for the same training illumination condition, when the test set is completely different, especially from $Outdoor$, the classification accuracy is reduced to 90.37%, which could be because varying daylight conditions result in nonuniform spectral reflectance properties across the visible and near-infrared

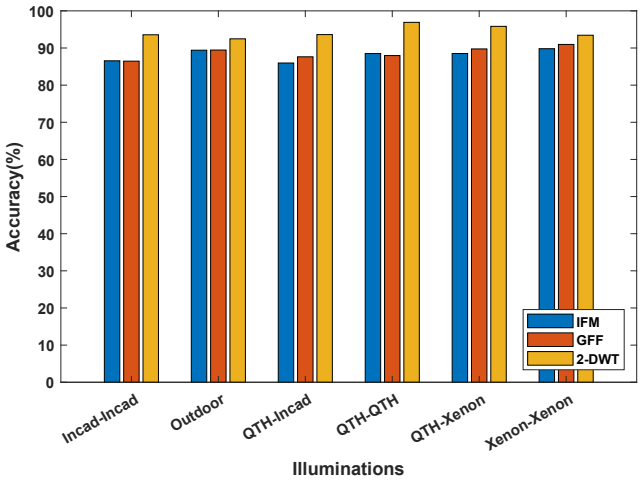


Fig. 4: Average classification accuracy across three fusion methods when samples are trained with $QTH - QTH$ and tested independently across other illumination data. For simplicity results are shown with of LG-SVM method

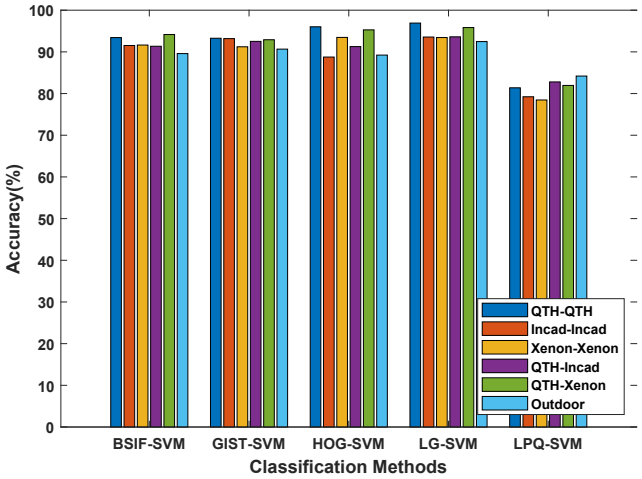


Fig. 5: Average classification accuracy independently across six different feature extraction methods when samples are trained with $QTH - QTH$ and tested independently across other illumination data. For simplicity results are shown with 2-DWT fusion method

spectra.

- Among the three image fusion methods employed, spectral band fusion based on 2-DWT outperforms the IFM and GFF methods, as illustrated graphically in Figure 4, which demonstrates the potential of employing the wavelet fusion method to obtain discriminate features that require robust gender classification.
- We also note that, of the five feature extraction method employed in this work, HOG, LG, GIST and BSIF performs better compared to LPQ method as can be viewed from Figure 5.

In summary, cross-illumination gender classification is a challenging task. However, the use of multi-spectral imaging demonstrates its potential for gender classification. A reasonable performance accuracy can be observed when the training and testing samples belong to different illuminations, which could be due to the invariant properties of the multi-spectral imaging and the use of spectral band image fusion method. However, cross-illumination gender classification remains a challenging task in outdoor illumination condition and requires further analysis and robust methods.

VI. Conclusion

Gender classification is important in many critical applications such as security, surveillance, and biometrics. The majority of studies employed in this direction for gender prediction are based on the visible spectrum and limited work exploring NIR, thermal, and multi-spectral imaging. With advancements in imaging technology across the spectrum, there are challenges in performing cross-illumination performance analysis under varying illumination conditions, which have not received significant attention. Thus, to present the generalization in the performance analysis, we performed gender classification by evaluating the cross-illumination approaches on a multispectral face database. The experimental results are presented on 78300 sample images captured from 145 subjects under six different illumination conditions. We employed three different image-fusion methods: Image Matting Fusion (IMF), Guided Filtering Based Fusion (GFF), and 2-Discrete Wavelet Transform Average Fusion (2-DWT) were used to combine the spectral band face images corresponding to nine narrow spectrum bands. Further, the results were obtained across five different feature extraction methods: Local Phase Quantization (LPQ), Histogram of Oriented Gradients (HOG), GIST, log Gabor (LG), and binarized statistical image features (BSIF), independently with a Support Vector Machine (SVM) classifier. Experimental evaluations were presented based on a cross-illumination evaluation, in which samples from training and testing were obtained under different illumination conditions. The use of multi-spectral imaging has shown reasonable classification across different illumination conditions, presenting the invariant properties of multi-spectral imaging sensors. The results were presented in the form of average classification accuracy by repeating the experiment 10 times for the random selection of training and testing samples. Based on the obtained results, the highest classification accuracy of 96.32% was obtained for the cross-illumination condition.

References

[1] C. Trixie and N. a. Clive, "Question of gender: Gender classification in international research," *International journal of Market Research*, vol. 65, pp. 575–593, 2022.

[2] A. Dhomne, R. Kumar, and V. Bhan, "Gender recognition through face using deep learning," *Procedia Computer Science*, vol. 132, pp. 2–10, 2018, international Conference on Computational Intelligence and Data Science.

[3] N. Narang and T. Bourlai, "Gender and ethnicity classification using deep learning in heterogeneous face recognition," in *2016 International Conference on Biometrics (ICB)*, June 2016, pp. 1–8.

[4] H. Moeini and S. Mozaffari, "Gender dictionary learning for gender classification," *Journal of Visual Communication and Image Representation*, vol. 42, pp. 1 – 13, 2017.

[5] R. P. Yuda, C. Aroef, Z. Rustam, and H. Alatas, "Gender classification based on face recognition using convolutional neural networks (cnns)," *Journal of Physics: Conference Series*, vol. 1490, no. 1, p. 012042, mar 2020.

[6] A. Swaminathan, M. Chaba, D. K. Sharma, and Y. Chaba, "Gender classification using facial embeddings: A novel approach," *Procedia Computer Science*, vol. 167, pp. 2634–2642, 2020, international Conference on Computational Intelligence and Data Science.

[7] N. T. Vetrekar, R. Raghavendra, A. A. Gaonkar, G. M. Naik, and R. S. Gad, "Extended multi-spectral face recognition across two different age groups: An empirical study," in *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*, ser. ICVGIP '16. New York, NY, USA: ACM, 2016, pp. 78:1–78:8.

[8] N. J. Siddiqui, "Novel approach for face recognition using cross-spectral environment," *International Journal of Research in Engineering and Science(IJRES)*, vol. 09, no. 10, pp. 70–76, 2021.

[9] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) workshops*, June 2015.

[10] P. Rai and P. Khanna, "Gender classification techniques: A review," in *Advances in Computer Science, Engineering & Applications*, D. C. Wyld, J. Zizka, and D. Nagamalai, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 51–59.

[11] N. Narang and T. Bourlai, "Face recognition in the {SWIR} band when using single sensor multi-wavelength imaging systems," *Image and Vision Computing*, vol. 33, pp. 26 – 43, 2015.

[12] C. Chen and A. Ross, "Evaluation of gender classification methods on thermal and near-infrared face images," in *2011 International Joint Conference on Biometrics (IJCB)*, Oct 2011, pp. 1–8.

[13] R. Raghavendra, N. Vetrekar, K. B. Raja, R. S. Gad, and C. Busch, "Robust gender classification using extended multi-spectral imaging by exploring the spectral angle mapper," in *2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, 2018, pp. 1–8.

[14] P. E. Pochi, J. S. Strauss, and D. T. Downing, "Age-related changes in sebaceous gland activity," *Journal of Investigative Dermatology*, vol. 73, no. 1, pp. 108 – 111, 1979.

[15] J. A. Iglesias-Guitian, C. Aliaga, A. Jarabo, and D. Gutierrez, "A Biophysically-Based Model of the Optical Properties of Skin Aging," *Computer Graphics Forum*, 2015.

[16] F. Wu, X.-Y. Jing, Y. Feng, Y. mu Ji, and R. Wang, "Spectrum-aware discriminative deep feature learning for multi-spectral face recognition," *Elsevier, pattern recognition*, 2021.

[17] S. Li, X. Kang, J. Hu, and B. Yang, "Image matting for fusion of multi-focus images in dynamic scenes," *Information Fusion*, vol. 14, no. 2, pp. 147 – 162, 2013.

[18] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2864–2875, 2013.

[19] N. T. Vetrekar, R. Raghavendra, and R. S. Gad, "Low-cost multi-spectral face imaging for robust face recognition," in *2016 IEEE International Conference on Imaging Systems and Techniques (IST)*, 2016, pp. 324–329.

[20] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2012, pp. 2879–2886.