

# Current Advances in Hyperspectral Face Recognition

<sup>1</sup>Anam Zahra, <sup>2</sup>Rizwan Qureshi, and <sup>3</sup>Muhammad Uzair

<sup>1</sup>University of Bordeaux, France, <sup>2</sup>City University of Hong Kong, Hong Kong, <sup>3</sup>Defence and Systems Institute, University of South Australia, Australia

Email: anam.zahra@etu.u-bordeaux.com, engr.rizwanqureshi786@gmail.com, muzair27@gmail.com

**Abstract**—Hyperspectral imaging systems are well established, for satellite, remote sensing and geosciences applications. Recently, the reduction in the cost of hyperspectral sensors and increase in the imaging speed has attracted computer vision scientists to apply hyperspectral imaging to ground based computer vision problems such as material classification, agriculture, chemistry and document image analysis. Hyperspectral imaging has also been explored for face recognition; to tackle the issues of pose and illumination variations by exploiting the richer spectral information of hyperspectral images. In this article, we present a detailed review on the potential of hyperspectral imaging for face recognition. We present hyperspectral image acquisition process and discuss key preprocessing challenges. We also discuss hyperspectral face recognition databases and techniques for feature extraction from the hyperspectral images. Potential future research directions are also highlighted.

**Index Terms**—Hyperspectral Face Recognition, Spectral Reflectance, Pose Variation, Illumination Variation, Deep Learning.

## I. INTRODUCTION

Machine vision systems for face recognition have achieved immense progress in the past decade. Unlike other biometrics (fingerprints, DNA, voice, iris, gait) face recognition neither involves direct physical interaction nor the user's consent. Compared to its counterpart, no physical interaction is required. Face recognition technology presents an easy solution for the users and economical to the vendors [1], [2]. Face recognition systems can be used both for security and commercial applications. Some of the key applications include surveillance, identification at airports, gaming, automatic tagging, unlocking mobile phones, automatic attendance marking, entertaining potential customers at banks and malls, classrooms evaluation, and many more [3].

Facial recognition is a very easy task for the human brain but it is a challenging problem for computer vision. Minor variations in pose or illumination can lead to inaccurate results. The human brain can easily identify the face, regardless of challenging appearance variations (makeup, hair style, glasses, masks, expressions) or from different poses; but for computer vision systems, a small change in appearance or background is not easy to distinguish.

Digital image based face recognition was initially treated as a 2D pattern recognition problem in which very basic features of symmetry and distance between important points of the face were used [4]–[6]. These approaches considered only a few points on the face, known as facial land mark points.

Digital video based face recognition was treated by modeling the spatio-temporal dynamics of face features [7]. After that, better approaches i.e. Eigenfaces [8] Fisherfaces [9] and graph matching [10] proved to be more efficient. These systems are known as holistic approaches; where instead of facial points, the whole face is used. Face recognition experienced tremendous advancements in the deep learning era, where human level performance has been achieved by using big datasets and deep neural network models [11]–[13]. Most current face recognition systems use color or RGB images for face recognition [14].

Hyperspectral imaging initially found its applications for aerial and remote sensing applications such as satellite imaging for earth observation [15] and astrophysical image application [16]. During the last few decades, due to the reduction in the hyperspectral camera cost and the increase in the imaging speed, hyperspectral imaging is finding new potential avenues in fields such as chemistry [17], agriculture [18], material classification [19], [20], forensic examination of document image analysis [21]–[23], and face recognition [24]–[27].

Hyperspectral images consists of a series of images captured at different contiguous wavelengths (Fig. 1). The number of wavelengths varies from a dozen to several hundreds of bands. A hyperspectral image of a face contains more fine details of facial tissues and skin. The image is in the form of a cube where each image in the cube, depicts the spatial values captured at a certain wavelength. The richer spectral information can be exploited to deal with some of the key challenges of face recognition such as illumination and pose [28], [29].

Recently, many attempts have been made to use hyperspectral images for face recognition [24], [26], [28]–[31]. In this paper, We present a comprehensive literature review of hyperspectral face recognition methods. We first present the methods and challenges involved in a typical ground system for the acquisition of hyperspectral face images (Section 2). Next, we present the details of the currently available hyperspectral face image databases (Section 3). Then, we highlight the key hyper-spectral pre-processing techniques and challenges in detail (Section 4). We then explain and analyze different techniques of hyperspectral face recognition (Section 5). Finally, the paper is concluded by commenting on the future potential of hyperspectral imaging for face recognition and presenting new insightful research directions.

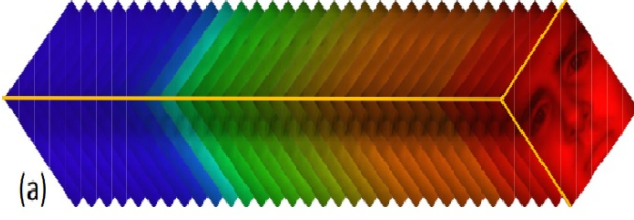


Fig. 1. A hyperspectral face image. Images acquired at different wavelengths are shown as a series of bands along the spectral dimension from blue to red. Figure taken from [26].

## II. HYPERSPECTRAL FACE IMAGING SET-UP

A typical ground based hyperspectral face imaging system is shown in Figure 2. The system consists of electronic tunable filters, optical devices and imaging sensors. The purpose of tunable filter is to perform wavelength selection. Modern filters are usually controlled electronically, mostly via software and their operating frequency can be tuned at a very high speed.

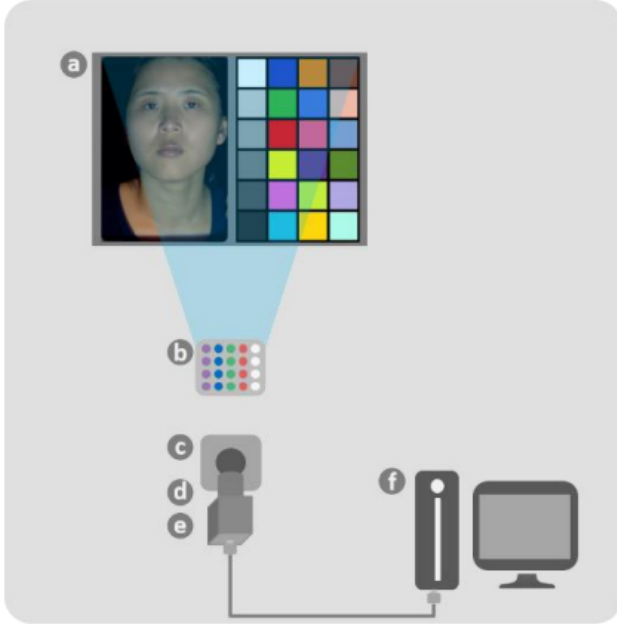


Fig. 2. A typical hyperspectral face imaging system: (a) Object and calibration pattern (b) Tunable filter such as Liquid Crystal Tuneable Filter (LCTF) (c) Optical device such as a lens (d) detector/sensor (e) Control software interface (f) Display computer. Figure taken from [32].

One such commonly used filter category include the Liquid Crystal Tuneable Filters (LCTF). The name finds its origin in the mechanism used by these filters (i.e. to filter the wavelength by controlling the liquid crystal). The LCTFs can be tuned at a broad spectral range. Generally, An LCTF is constructed by sandwiching the fixed filters and the liquid-crystal filters along with the linear polarizers. The wavelength for visible devices ranges from  $400nm$  to  $700nm$  [33], whereas for hyperspectral imaging ranges between  $400nm$  to  $2450nm$  [34]. So far, *Lyot filtering* is the basic principal applied, but some other designs are also available. The high resolution, better life

span, lower weight, and less power requirements makes LCTF a desired imaging filter. Furthermore, the degradation is caused by the mechanical or thermal shock and higher exposure to heat and humidity which can easily be avoided, thus making it a major image acquisition tool for remote as well as general hyperspectral imaging systems.

The Optical devices include lenses that are used to concentrate and focus the light for better image quality. The lenses are specially designed with ZnS and ZnSe coated glasses [35]. The imaging sensors are basically digital cameras. Usually commercial machine vision cameras are used for this purpose because of their easy customization to work in conjunction with the LCTF.

**A control software synchronizes the whole operation by first selection the appropriate wavelength and then allowing the light to reach the CCD through shutter control. There are two main classifications of the softwares i.e. perception studio & perception core [36]. The perception studio software is used to record the data for the given camera/setup, manipulate the bands and finally generates the model. Whereas, the perception core utilizes the GPU for the manipulation of the models recorded through perception studio.**

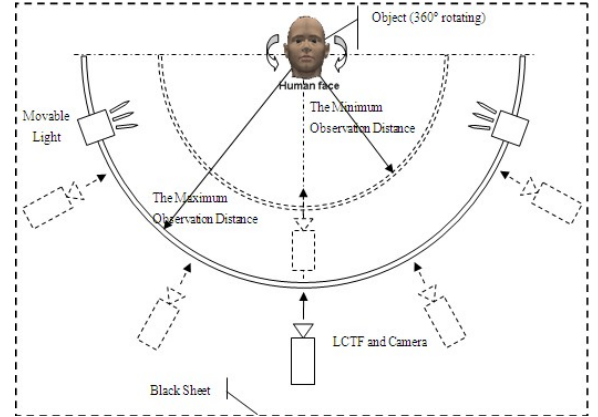


Fig. 3. Hyperspectral image acquisition setup of CMU dataset. Figure taken from [34].

## III. DATABASES OF HYPERSPECTRAL FACE IMAGES

There are several public databases available for hyperspectral face recognition research. In this section, we describe several popular hyperspectral face databases.

### A. CMU Hyperspectral Face Database

The CMU hyperspectral face database was prepared at the Robotic Institute Carnegie Mellon University by Denes et al. [14]. Starting from October 2001, multiple hyperspectral images of 54 subjects were collected. The spectral range is between  $450nm$  to  $1100nm$  with a step size of  $10nm$ .

1) *Acquisition Setup*: A 'Spectro Polarimetric Imaging Camera' was used for image acquisition. Hyperspectral images were acquired under controlled conditions (Fig 3). Subjects position were stabilized using head rest stop, whereas subjects

TABLE I  
FACE RECOGNITION DATABASES

| Database                         | Wavelength Range | No.of Bands | No.of Subjects |
|----------------------------------|------------------|-------------|----------------|
| CMU Hyperspectral [37]           | 450-1100nm       | 65          | 54             |
| IRIS-M Hyperspectral [38]        | 480-720nm        | 25          | 82             |
| PolyU-HFSD [39]                  | 400-720nm        | 33          | 25             |
| Stanford-1 Hyperspectral [40]    | 400-1000nm       | 160         | 45             |
| Stanford-2 Hyperspectral [41]    | 970-2500nm       | 256         | 45             |
| IRIS-HFD-2014 Hyperspectral [41] | 420-700nm        | 29          | 130            |
| UWA-HSFD Hyperspectral [26]      | 400-720nm        | 33          | 70             |

maintained still pose and height. Three 600-Watt halogen lamps were used at the angle of -45, 0, and 45 degrees for illumination. Cameras switching was automatic starting from all three and then turning off from left to right. A fast computer was used for both controlling, collection and storage of the data. Around 20MB data was collected in each session, comprised of 640x480 frontal views.

2) *Limitations of the Acquisition Setup*: The camera used for the acquisition of hyperspectral images was calibrated such that lightening and charge couple device (CCD) set to the highest response at 650nm with maximum light exposure at 1/60 seconds. The system suffers from noise at shorter and longer wavelengths.

#### B. IRIS-M Database

IRIS-M Hyperspectral Database prepared at University of Tennessee consists of different classes, one of them being the face. The face database comprises of 82 subjects. Hyperspectral images were taken in 10 different sessions, both in indoor and outdoor environments.

1) *Acquisition Hardware and Setup*: The images acquired in controlled conditions (indoor environment) were captured by using 2 halogen or fluorescent lamps for illumination; whereas outdoor images were collected under daylight. Frontal views of resolution 640x480 were captured in the wavelength range of 480nm to 720nm with a step of 10nm, comprising of a total 25 bands. Raytheon Palm-IR Pro Camera was used for capturing images in the thermal infrared wavelength range.

#### C. PolyU Hyperspectral Face Database

PolyU Hyperspectral Face Database was collected at Biometric Research Centre (UGC/CRC)-The Hong Kong Polytechnic University. 25 subjects comprising of 17 males and 8 females volunteers, and 300 samples were captured. Considering the changes occurring in the human face's appearance over time, multiple sessions were conducted with an average time difference of 5 months (minimum 3 months and maximum 10 months).

1) *Acquisition Setup*: For each subject 3 hyperspectral cubes i.e frontal, left, and, right, with neutral expressions were captured under indoor halogen illumination. Overall 33 spectral bands were covered from 400nm to 720nm with a difference of 10nm.

#### D. IRIS-Hyperspectral Face Database-2014

IRIS-HSFD was developed at the University of Tennessee in 2014, comprising of 86 males and 44 females subjects. Overall 490 hyperspectral face cubes were acquired over 420nm to 700nm with a difference of 10nm. This database covers more fine details due to the exposure time adaption. It does address several other problems, such as pose variation (frontal and side views) and structural features (glasses).

1) *Acquisition Setup*: For IRIS-HFD-2014 130 subjects' images were acquired without spectacles and with neutral face images whereas 51 individuals with spectacles were captured. RGB images for both scenarios were also collected. Frontal as well as poses from -45 and 45 degrees were collected by using *Lumia 5.1 reef version*. The light source comprised of 5 LED channels (neutral white, royal blue, hyper violet, deep red, and turquoise). X-rite Color Checker Classic was placed by the side of each subject while capturing hyperspectral face images for the purpose of calibration analysis and facial color. To adjust the exposure time for each wavelength, the spectral transmittance property of LCTF was used.

#### E. UWA-HSFD

The University of Western Australia hyperspectral face database consists of facial hyperspectral images of 70 different subjects in the wavelength range of 400nm-720nm.

1) *Acquisition Hardware*: The face data base was gathered using CRI's Varispec LCTF, equipped with a photon focus camera helping in adjusting exposure time, luminance adaption and CCD sensitivity. Image cubes were captured over 33 bands from 400nm to 720nm with a difference of 10nm. Subjects didn't keep their head still so there are variations in this data set.

#### F. Stanford Hyperspectral Database

1) *Indoor*: Using Hypspec VNIR camera the images of 45 subjects with the resolution of 1600 spatial pixels were acquired. The studio tungsten light was used as the illumination source. Frontal views were captured over 148 bands from 415nm to 950nm. Whereas hyperspectral cubes from 970nm to 2500nm with the 6nm spectral resolution were also collected.

2) *Outdoor*: Another set of images of 70 subjects was captured in daylight using VNIR and SWIP. It took 20 seconds in face scanning where subjects were advised to keep heads still rested against wall and hold the breath. There are

TABLE II  
CAMERA SPECIFICATIONS

| AO Material    | TeO <sub>2</sub> |
|----------------|------------------|
| Spectral Range | 450nm-1100nm     |
| Resolution     | 10nm @ 600nm     |
| AO Efficiency  | ≥80%             |
| RF Range       | 25-70MHz         |
| Retarder Range | 400-1800nm       |
| IFOV           | -7 Degree        |

blinking artifacts quite visible. If it is scanned more accurately breathing artifacts are also visible.

#### IV. PREPROCESSING OF HYPERSPECTRAL IMAGES

A hyperspectral image consists of two spatial and one spectral dimension. The spectral dimension has series of channels across different wavelengths. During the acquisition process, several types of artifacts can add up. It is necessary to apply some preprocessing techniques to extract the relevant information from the hyperspectral cube and suppress the artifacts. Fortunately most of the hyperspectral devices includes state of the art preprocessing software packages that includes different image processing methodologies. In this section, we have highlighted the preprocessing challenges of HSI. We have listed the advantages and drawbacks of different algorithms for solving these challenges in Table 3.

##### A. Spikes

Spikes are a sudden rise followed by a sharp fall in the spectrum. Spikes are generated due to abnormal behavior of instrument, environment, or imperfection in electronic circuitry. They often mask the detail in an image and leads to an inaccurate analysis. One of the best method to detect spikes is by manual supervision. This requires human care and is time consuming. It becomes more difficult for hyperspectral cube because of large amount of data. Interpolation or removing spikes based on the neighboring pixel [42], [43] is a good choice for removal of spikes. Other solutions include median and median modified wiener filters for anomaly or spike detection [44].

##### B. Dead Pixels

Another challenge in HSI is the presence of dead pixels often replaced by zero or the maximum value. Dead pixels are caused by anomalies in the detector. The location and size of dead pixels varies between a specific pixel, group of pixels or a complete pixel line. For finding the dead pixels, thresholding can be a good choice. The heuristic algorithms, such as, genetic algorithms or other evolutionary techniques can produce better results. Once dead pixels are located, the best choice is to interpolate them with neighboring pixels

##### C. Spectral Preprocessing

Spectral preprocessing is required for avoiding the undesirable phenomena like scattering of light, the effect of particle size or morphological differences affecting the spectral

measurement. A variety of preprocessing techniques like denoising, Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) are widely used in spectral preprocessing.

##### D. Image Size and Compression

HSI contains thousands or even millions of pixels, e.g, an 8-bit HSI having 200 bands contains more than 13 million data points. This huge amount of information requires much memory and transmission time. This issue can be addressed by fast and computationally powerful systems. However in many applications, such as time series of images or super resolution images, storage is still an important issue [47]. Therefore compression of an image is often necessary to retain the desired information [48]. The most common approaches of compressing images are byte encoding and data binning. Heuristic algorithms like Genetic algorithm, Ant colony optimization can also be used to remove the redundant data [49]. Factor model approaches like principle component analysis and multivariate curve analysis significantly reduces the spectral dimensions of hyperspectral cube.

##### E. Background Removal

The non informative background or outliers (not consistent with rest of the dataset) further complicates the information extraction. These aberrant observations can be generated due to three reasons; such as the instrument, geometry of the sample and the radiation. To overcome these issues, classical image preprocessing techniques like the use of histogram to detect the sharp changes in an image or manual selection of a threshold value can be used to clean up and extract the relevant information from the hyperspectral cube.

#### V. HYPERSPECTRAL FACE RECOGNITION TECHNIQUES

In this section, we present major achievements in the area of hyperspectral face recognition techniques. Hyperspectral imaging technology provides overwhelmingly large amounts of information available for processing. It captures details over dozens of bands that is more finegrained compared to 3 bands RGB. It also captures data beyond the visible range. Keeping this in view, Pan *et al.* [28] studied the reflectance models of human skin tissues for face recognition through hyperspectral images. 200 subjects were captured with the spectral resolution of 468x494 over 31 bands ranging from 0.7 $\mu m$  to 1.0 $\mu m$ . It has been observed that near infrared spectral patterns of underlying skin tissues vary from person to person, thus this unique feature scan be employed for human recognition. Analyzing the underlying fine details of skin tissues offers a solution to the orientation problem often encountered in face recognition. The problem of luminance affecting face recognition can also be decreased. As reflectance over two different spectralon panels is measured i.e. (a panel with 99% reflectance is refereed to as white spectralon whereas 2% reflectance spectralon is dark spectralon). Reflectance is constant over both panels in the range of 0.7 $\mu m^*$  to 1.0 $\mu m$ .

Five facial regions i.e hair, forehead, right and left cheeks and lips are the regions of interest. For each region of

TABLE III  
PRE-PROCESSING CHALLENGES OF HSI

| Pre-Processing Challenges     | Algorithms  | Image Type   | Advantages  | Drawbacks  |
|-------------------------------|---|--|---|--|
| Spikes                        | Thresholding  | Images that contain more than one channel RGB, Multiband and Hyperspectral.                        | Robust and reliable   | Time consuming Specially in Hyperspectral Images. Difficult to find the best filter mask. The selection of wavelet for each type of image. |
|                               | Manual Inspection                                     |  | Robust and reliable   |  |
|                               | Averaging Filters.                                    |  | Robust and reliable   |  |
|                               | Wavelets  | All  | Robust and reliable   |  |
| Dead Pixels                   | Interpolation   | Multiband, Hyperspectral   | Easy to calculate   | Dependent on Signal to noise ratio, Risk of false positives.   |
|                               | Spectral median                                       |  | Easy to implement   |  |
|                               | Evolutionary Algorithms /Genetic algorithm            |  | Robust and reliable   |  |
|                               | Interpolation/Neighbouring                            | All  | Robust and reliable   | Large cluster of dead pixel. can result in losing resolution   |
| Spectral Pre-Processing       | De-noising<br>Scatter Correction<br>MSC,SNV [45]      | Suitable for Hyperspectral Images. Special care should be taken with multi-band images.            | Easy to Implement   | To find the best fit of filter parameters. specially the window size   |
| Compression                   | Spectral binning<br>Spatial binning<br>Bytes Encoding | All, but special applicability to multiband and hyperspectral images. All, depending upon software | Easy to Implement<br>Easy to Implement<br>Efficient Compression images, based on software | Loss of spectral information. Loss of spatial information For Specific type  |
|                               | Statistical models                                    |  | Information<br>Efficient compression  |  |
| Background Region of interest | Principle component analysis, Threshold [46]          | All  | Selection of ROI based on PC scores   | Selection of threshold is tedious  |

interest, reluctance vector is estimated. Overall, 90% accuracy is achieved using this technique for frontal images whereas around 75% accuracy is achieved with orientation variation. The strong aspect is the improved performance of the technique for the rotated faces. But on the other hand, significant change is observed with expression variation, except in the forehead.

Di *et al.* [29] proposed that different hyperspectral bands carry different information. The details from one set of bands vary from the information in the other set of bands. They proposed different strategies for the analysis of selected feature bands. Over certain sets of bands the feature information is very rich, once such sets called feature bands are selected, three different techniques are applied to analyze the efficiency. Their dataset comprised of 300 HSI cubes of 25 subjects with frontal, right and left view over 33 bands from 400nm to 720nm captured in 4 sessions at different times. After processing and feature band extraction, 540nm and 580nm band carries most significant feature information. First of all, Whole Band (WB) based  $(2D)^2$  PCA was implemented in which all the feature bands were input to the PCA. In the next technique, each band was processed separately and were fused at the decision level; this strategy called as single band based  $(2D)^2$  PCA. Whereas the final technique, included fused band sets passed as input to PCA, known as band subset fusion based  $(2D)^2$  PCA. Results are as accurate as 80% with the cumulative eigen value  $Cu=98.5\%$  for BS-WFD technique.

Wang *et al.* [50] proposed Gabor features for expression invariant face recognition. They acquired dataset of 400 images

from 200 subjects over 31 bands of Near Infra Red (NIR) from 0.7  $\mu m$  to 1.0  $\mu m$ . In the proposed algorithm, eye location was found as the first feature. Afterwards using eye locations skin regions were found. The skin was used for spectral feature extraction. Once all features were calculated, 3D Gabor filter is used for Gabor feature extraction.

Once Gabor features were extracted, then PCA was applied for testing and training. The maximum accuracy of 1 for training and 0.95 for testing was achieved by combining Gabor PCA and spectral features. They achieved high accuracy but there were a few limitations. Data set was very small to be generalized and the only expression used was *smile* that also could not lead to the general phenomenon for expression invariant face recognition.

A new technique for hyperspectral face recognition was proposed by rangaswamy *et al.* [51]. This method involves Dual Tree Complex Wavelet Transform (DTCWT) fused with Fast Fourier Transform (FFT). 5-level DTCWT was applied to 128 \* 512 face resized images. The features used for face recognition were magnitudes of both real and imaginary bands. FFT coefficients were sorted in descending order. Finally both vectors were arithmetically summed up. Algorithm was tested on ORL, JAFEE, L-SPACER & CMU-PIE databases. It achieves 90% recognition rate.

Raghavendra *et al.* [52] gathered hyperspectral face images from 40 subjects over a period of 90 days keeping neutral and smiling expression. The uniqueness was that, they collected images over 6 bands i.e. 425nm, 475nm, 525nm, 570nm, 625nm, and 680nm. Total of 1200 images were acquired. The

idea of fusing 2 images with highest entropy was used. The images were then decomposed through discrete Wavelet transform (DWT) and fused again after analyzing sub-band details. Face recognition was done through local binary patterns. They claimed 100% accuracy on their database.

To achieve better accuracy and recognition rate, some researchers have tried to study more fine details. Sharma and Gool [53] processed each band as separate image. This way, high level information could easily be exploited and more details can be explored. Hyperspectral face images in the range of 380nm to 700nm, 750nm to 1100nm, treating each slice of cube as a separate image. State of the art techniques for feature extraction were applied i.e. SIFT, HOG and LBP. Feature sets were passed to SVM for classification.

Sharma *et al.* [54] proposed hyperspectral CNN for image classification and band selection. Each band in the cube was treated as a separate image. Discriminative band selection was accomplished through AdaBoostSVM. The architecture contained 3 convolutional layers followed by 2 fully connected layers which was then connected to C-way softmax layer. Whereas a convolutional layer was formed as conv layer followed by s-norm followed by ReLU and Maxpooling, and then the next Convolutional layer. Through CNN 99.2% accuracy was achieved for PolyU-HSFD and 98.8% for CMU. On the other hand SVM achieved 99.3% accuracy for PolyU and 99.2% for CMU.

On the way to achieve better results Chen *et al.* came up with [55]. First of all, adaptive denoising of each spectral band was carried out. Afterwards, face is cropped from these denoised images using eye coordinates. And finally log polar Fourier features are extracted and voting classifiers are applied. In order to reduce distortion, log-polar transformation is carried out. Different translation, rotation and scale invariant features are extracted through 2D Fourier spectrum. The databases selected were PolyU and CMU, BM3D methodology for denoising was utilized. Finally CRC based method is used for the classification of original HSI cubes. The recognition rate achieved is  $92.7\% \pm 2.6$  using polyUHSFD.

In [56], Kim *et al.* proposed the solution to identifying synthetic and real human faces. They used CASIA NIR-VIS 2.0 database (comprised of 2055 visible (VIS) and NIR face images from 205 different subjects). Synthetic face image from the pair i.e. VIS & NIR preserves fine color and texture details. Joint bilateral filter is applied to the NIR image, the filtered image has three channels. This forms the base image. On the other hand, bilaterally filtered VIS image forms the detailed image. The synthesized image is the result of dot product of the base image and the detailed image. Color vector is extracted through LBPs. Accuracy of 97.59% is reported.

Vetrekar *et al.* extended the applications of face recognition to the different age groups and presented their findings in [57]. They established database of two age groups (i.e. less than 15 and greater than 20 years old) Images from 168 individuals were collected in two different sessions. The main purpose was to study the variations in results for two different age groups. Different state-of-the art algorithms were applied and better results for age group of greater than 20 years old. The findings for each technique are discussed in the Table IV.

TABLE IV  
RESULTS OF EXTENDED MULTI-SPECTRAL .... STUDY

| Age Group | HOG   | LBP   | GIST | Log Gabor Transform |
|-----------|-------|-------|------|---------------------|
| $\leq 15$ | 92.85 | 34.52 | 87.5 | 96.42               |
| $\geq 20$ | 98.80 | 73.21 | 100  | 100                 |

Another step in face recognition through CNN was taken by Peng *et al.*, as they modified GoogleNet CNN for NIR Face Recognition [58]. They have used CASIA NIR face database and suggested that medium sized network suffice the need rather than a full sized GoogleNet. For full sized GoogleNet the recognition rates achieved by softmax0, softmax1, and softmax2 are 99.02%, 98.8% and 98.74% respectively. It is modified such that there are 8 layers and 2 feature extraction modules. The overall identification rate of 98.28% was achieved through this strategy.

Chen *et al.* have analyzed different state of the art techniques. Images from PolyU and CMU were corrupted through Gaussian white noise with standard deviation (0, 10, 20 and 30). These images were then classified through 5 state-of-the-art techniques (i.e. spectral angles, spectral signature, eigenfaces, 3D gabor Wavelet and log polar FFT2). 3DLDP and PLS+Band Fusion were also tested. Best results of  $95.2 + -1.6$  for PolyU and  $99.1 + -0.6$  for CMU were achieved through band fusion and PLS.

Vetkar *et al.* prepared a database of 230 subjects over the range of 530nm to 1000nm. The data was gathered using BCi5-U-M-40-LP camera in two different sessions, using the white background. [59] studies four state of the art techniques on the acquired database. The methods used are HOG, LPQ, GIST, and Log Gabor Transform along with CRC based classification. These algorithms were tested on both single image as well as fused bands of hyperspectral image. Recognition rate of 99.55% was achieved for fused image with HOG, CRC, and GIST CRC. On the other hand their claim is to reach 100% recognition rate for Log Gabor CRC method.

Using UWA database, Cho *et al.* carried out inter band alignment analysis [60]. They have investigated inter band misalignment using 10 state of the art techniques and have proposed a new method as well. Their approach is to carry out qualitative performance analysis through principal curvature based map and cumulative probability of target colors in HSV domain. Principal Curvature map of each band estimated through maximum or minimum eigen values of 2x2 Hessian Matrix. Through this pixel wise similarity index is calculated. To evaluate color distortion, HSV color domain for each band is considered.

Patch-Based Low-Rank Tensor Decomposition for hyperspectral images was proposed by Du *et al.* [61]. They studied the tensor decomposition based HSI comparison methods for the removal of spatial and spectral redundancies. Spatial dimensions and neighborhood relations as well as global correlations of spectral domain are preserved through third order tensors for each local HSI patch. Fourth order tensor per cluster is formed through nonlocal similarity of spatial domain. In the following step each cluster is then decomposed to a coefficient tensor and three dictionary matrices. HSI can now

TABLE V  
IMPORTANT STUDIES ON HYPERSPECTRAL FACE RECOGNITION

| Author                         | Year | Dataset Size | Bands        | Spectrum Range                        | Accuracy |
|--------------------------------|------|--------------|--------------|---------------------------------------|----------|
| Z.pan [29]                     | 2010 | 9,651        | 33           | 0.7 $\mu\text{m}$ - 1.0 $\mu\text{m}$ | 90%      |
| Han                            | 2014 | 400          | 31           | 0.7 $\mu\text{m}$ - 1.0 $\mu\text{m}$ | 100%     |
| Wang [50]                      |      |              |              |                                       |          |
| W.Di [29]                      | 2010 | 300          | 33           | 400 - 720nm                           | 80%      |
| Raghavendra <i>et al.</i> [52] | 2015 | 1200         | 6            | 425nm - 680nm                         | 100%     |
| Sharma [53]                    | 2016 |              | 380 - 1100nm | 85%                                   |          |
| Vetrekar [57]                  | 2016 | 230 / self   |              | 530 - 1000nm                          | 100 %    |
| Uzair [26]                     | 2015 | UWA          | 33           | 400 - 720 nm                          | 98%      |
| Chen [64]                      | 2016 | PolyU        | 33           | 400 -720nm                            | 95%      |

be reconstructed by the multiplication of coefficient tensor and dictionary matrices. This leads to the removal of redundancy. The performance is evaluated on Stanford HSFD.

In [62] Liang *et al.* used 3D-High Order texture pattern on Local Derivative Pattern ( LDP ) descriptor for hyperspectral face recognition. Hyperspectral image is encoded with multi-directional derivatives and spectral-spatial space through binarization function. Then spatial-spectral descriptor is generated through a 3D histogram on the derivative patterns. The recognition rates achieved are  $95.33 \pm 1.63\%$  &  $94.83 \pm 2.62\%$  for PolyU and CMU databases respectively.

In the year 2015, Han *et al.* carried out a study [63]. Graphite details were captured using NIR bands and graphite sketches were obtained. From the graphite sketches, the original design can be restored by removing damages occurred over years. Through hyperspectral imaging extraction of contour line (i.e. graphite distribution areas), by selecting such parts as region of interest based on spectrum analysis or spectral features can be done. Accuracy was evaluated with the use of spectral angle classification, spectral feature fitting classification and the matching degree (M). The maximum accuracy is 90.30%.

Multispectral Low Rank Structured dictionary learning for face recognition [65] suggests real time applications, during acquisition of multi spectra images, high level of noise may corrupt the multi spectral image which can degrade the recognition rate. To improve the performance of multispectral images corrupted with noise, a novel approach called MSDL is used. In the proposed approach, a structured common dictionary and a spectrum specific dictionary are learned to exploit both the spectrum specific and common information among the images. The low rank matrix regularization technique is applied on sub dictionaries to learn pure dictionaries, that can reconstruct the denoised images. To reduce the computational complexity of the classification problem, they have used L2 - norm instead of L1 - norm based on collaborative representation. The experiments performed on PolyU, CMU and HK datasets shows the effectiveness of the proposed approach. In addition to this, experimental results also show the effectiveness of low rank structured regularization and low rank structural incoherence terms and the efficiency of the designed classification scheme.

Hyperspectral imaging offers the ease of pose variation and ensures better accuracy in such scenarios. Large amount of information is available in the form of multiple bands but in the era of deep learning the data set with diversity is required which is generally not available in most of the cases. Therefore it can easily be hypothesized that the above studied algorithms doesn't give generic results. On the other hand the aforementioned strategies does not perform well for analysing the emotion variation which is the current challenge. The selection of bands though is critical for the success of various algorithms using the deep learning approaches. So far we observed that different strategies focus on keeping the spatial relations between various channels and manipulating that information to achieve optimum results. These studies pave the path towards the manipulation of hyperspectral imaging through deep learning approaches.

## VI. DEEP LEARNING AND HYPERSPECTRAL IMAGING

Deep learning has recently shown remarkable results in many computer vision application [66], [67]. A recent review [68] reported several applications of deep learning for hyperspectral imagery. In the paper [69], a convolution neural network (CNN) has been proposed for the sparse band selection of hyperspectral face recognition. However more studies are needed of hyperspectral face recognition using deep learning methods.

CNNs have played a revolutionary role in the field of image and video recognition and classification. The technique served the purpose for hyperspectral imaging as well [70]–[79]. [80] is the initialising step towards the study of CNNs application for HSI. Overall there has been less progress made in the area of HSFR compared to HSI manipulation through deep learning methods. Recently, coupled generative adversarial network (GAN) is proposed for face recognition [81]. The major advantage of the HSI is the rich information availability. There is enough room to study the spectral content alongside the spatial material. One such experiment [82] proposes the deep architecture manipulating both. A two step technique first generates the spectral and spatial feature maps which are then used to train and then fine tune the DCCNs-LR classifier. This technique makes it easy to combine both spectral and spatial contents but has the disadvantage of feature selection by deep architecture.

The availability of a large number of bands and higher pixels offers room for the exploration of pixels pairing and its effects. One such study [73] pairs the training samples with similar and different class pixels and then fed it to the well designed CNN, which gives a final classification through voting. The strategy outperforms the previous with highly improved accuracy on a small number of samples with total processing time of 16.92 i.e 16 hours and 55.2 minutes.

In order to get better approximation and generalized neural network architecture, the requirement is to train with large amount of data. Many data sets have been prepared and are being used for RGB images but there is lack of availability when it comes to hyperspectral imaging. One solution is



the data augmentation [83]–[85]. The idea is to generate more and more data by changing the already available data. Data augmentation for hyperspectral images is bit tricky and common methods don't give promising results [86]. At first the general data augmentation i.e. rotations, translations etc have been tried but it doesn't give any satisfactory results. A new technique Pixel-Block Pair(PBP) is introduced that utilizes both spectral and spatial information is introduced. Pixel blocks are constructed and updated with augmented data and is then passed through the CNN and then labeled based on similarity of merged blocks. It leads to very high classification accuracy. Another data augmentation technique is online data augmentation through inference as well as the offline data augmentation based on PCA and generated promising results with accuracy of 89% [87]. This is still an emerging area and has many more to offer. The illumination problem for hyperspectral CNN is addressed and resolved using the data augmentation [88]. In the preprocessing stage of the hyperspectral images the raw pixels are given certain reflectance values which make use of the noisy or missing values. This process is efficiently done through radiometric normalization to enhance the spectral features and reduce the noise or environmental artifacts. A novel relighting data augmentation technique is proposed which estimates and improves the radiance of the scene. The classification results are then achieved through CNNs.

Is it possible to add a layer of spectral features and band reduction using deep neural networks?

Can we use pre-trained face recognition models on hyperspectral database??

Can we use data augmentation to generate large sample of hyperspectral faces. These are some of the unexplored questions which requires extensive research.

With the advancement in computational speed, and the availability of GPUs, TPUs, NPU's, and the rise of deep learning and enormous amount of data available today, we hope that, deep learning will lead the future research in hyperspectral imaging.

## VII. CONCLUSION AND FUTURE WORK

In this paper, we presented a comprehensive review of hyperspectral face recognition systems. We presented data acquisition systems, databases and major pre-processing challenge. A detailed overview of major achievements is also presented. We also discuss the rise of deep learning and highlight important future research directions. We hope this article may trigger the research in hyperspectral face recognition. Hyperspectral imaging has found its application in almost every area of image processing and computer vision in the past two decades. Its applications in face recognition are promising but can it offer something beyond recognition for instance emotion recognition or better say emotion understanding? We believe that the methodical engineering of the spectral information offered by HSI can lead to highly accurate emotion recognition and understanding systems. On the other hand the break through of deep learning has impact in almost every field, yet HSI research can have big break in relevance to the

data augmentation. Deep learning approaches are data hungry, training a deep neural network with a small number of samples is a hot research topic in machine learning, one approach is to use the matching networks [89]. Matching networks and other efficient deep learning architectures specifically tailored for hyperspectral face recognition systems are yet to be explored.

## REFERENCES

- [1] H. K. Ekenel and B. Sankur, "Multiresolution face recognition," *Image and Vision Computing*, vol. 23, no. 5, pp. 469–477, 2005.
- [2] L. J. Denes, P. Metes, and Y. Liu, *Hyperspectral face database*. Carnegie Mellon University, The Robotics Institute, 2002.
- [3] R. Poppe, "A survey on vision-based human action recognition," *Image and vision computing*, vol. 28, no. 6, pp. 976–990, 2010.
- [4] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM computing surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.
- [5] K. W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition," *Computer vision and image understanding*, vol. 101, no. 1, pp. 1–15, 2006.
- [6] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *International workshop on analysis and modeling of faces and gestures*, pp. 168–182, Springer, 2007.
- [7] E. J. Barbeau, M. J. Taylor, J. Regis, P. Marquis, P. Chauvel, and C. Liégeois-Chauvel, "Spatio temporal dynamics of face recognition," *Cerebral Cortex*, vol. 18, no. 5, pp. 997–1009, 2007.
- [8] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586–591, IEEE, 1991.
- [9] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 7, pp. 711–720, 1997.
- [10] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, and S. Lin, "Graph embedding and extensions: A general framework for dimensionality reduction," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 1, pp. 40–51, 2006.
- [11] O. M. Parkhi, A. Vedaldi, A. Zisserman, et al., "Deep face recognition," in *bmvc*, vol. 1, p. 6, 2015.
- [12] Y. Sun, D. Liang, X. Wang, and X. Tang, "Deepid3: Face recognition with very deep neural networks," *arXiv preprint arXiv:1502.00873*, 2015.
- [13] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in *Advances in neural information processing systems*, pp. 1988–1996, 2014.
- [14] Q. et al, "Face recognition applied to computer forensics," *International Journal of Forensic Computer Science*, vol. 1, pp. 19–27, 2006.
- [15] P. Arellano, K. Tansey, H. Balzter, and D. S. Boyd, "Detecting the effects of hydrocarbon pollution in the amazon forest using hyperspectral satellite images," *Environmental Pollution*, vol. 205, pp. 225–239, 2015.
- [16] S. Bourguignon, D. Mary, and E. Slezak, "Sparsity-based denoising of hyperspectral astrophysical data with colored noise: Application to the muse instrument," in *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pp. 1–4, IEEE, 2010.
- [17] J.-H. Cheng, D.-W. Sun, H. Pu, and Z. Zhu, "Development of hyperspectral imaging coupled with chemometric analysis to monitor k value for evaluation of chemical spoilage in fish fillets," *Food Chemistry*, vol. 185, pp. 245–253, 2015.
- [18] L. Ravikanth, D. S. Jayas, N. D. White, P. G. Fields, and D.-W. Sun, "Extraction of spectral information from hyperspectral data and application of hyperspectral imaging for food and agricultural products," *Food and bioprocess technology*, vol. 10, no. 1, pp. 1–33, 2017.
- [19] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Transactions on geoscience and remote sensing*, vol. 32, no. 4, pp. 779–785, 1994.
- [20] N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Transactions on Geoscience and remote sensing*, vol. 42, no. 7, pp. 1552–1565, 2004.



- [21] A. Berciano, D. Diaz-Pernil, W. Kropatsch, H. Molina-Abril, and P. Real, *Computer Analysis of Images and Patterns: 14th International Conference, CAIP 2011, Seville, Spain, August 29-31, 2011, Proceedings*, vol. 6854. Springer, 2011.
- [22] Z. Khan, F. Shafait, and A. Mian, "Hyperspectral imaging for ink mismatch detection," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, pp. 877–881, IEEE, 2013.
- [23] R. Qureshi, M. Uzair, K. Khurshid, and H. Yan, "Hyperspectral document image processing: Applications, challenges and future prospects," *Pattern Recognition*, vol. 90, pp. 12–22, 2019.
- [24] M. Uzair, A. Mahmood, and A. S. Mian, "Hyperspectral face recognition using 3d-dct and partial least squares," in *BMVC*, 2013.
- [25] W. Cho, J. Jang, A. Koschan, M. A. Abidi, and J. Paik, "Hyperspectral face recognition using improved inter-channel alignment based on qualitative prediction models," *Optics express*, vol. 24, no. 24, pp. 27637–27662, 2016.
- [26] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatio-spectral information fusion and pls regression," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 1127–1137, 2015.
- [27] M. Uzair, A. Mahmood, A. Mian, and C. McDonald, "Periocular region-based person identification in the visible, infrared and hyperspectral imagery," *Neurocomputing*, vol. 149, pp. 854–867, 2015.
- [28] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1552–1560, 2003.
- [29] W. Di, L. Zhang, D. Zhang, and Q. Pan, "Studies on hyperspectral face recognition in visible spectrum with feature band selection," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 40, no. 6, pp. 1354–1361, 2010.
- [30] S. A. Robila, "Toward hyperspectral face recognition," in *Image Processing: Algorithms and Systems VI*, vol. 6812, p. 68120X, International Society for Optics and Photonics, 2008.
- [31] M. Uzair, A. Mahmood, F. Shafait, C. Nansen, and A. Mian, "Is spectral reflectance of the face a reliable biometric?," *Optics express*, vol. 23, no. 12, pp. 15160–15173, 2015.
- [32] W. Cho, "Hyperspectral data acquisition and its application for face recognition," 2015.
- [33] Wikipedia, "Liquid crystal tuneable filters," 2019. [https://en.wikipedia.org/wiki/Liquid\\_crystal\\_tunable\\_filter](https://en.wikipedia.org/wiki/Liquid_crystal_tunable_filter).
- [34] M. R. Centre, "Liquid crystal tuneable filters," 2012. <http://olympus.magnet.fsu.edu/primer/java/filters/lctf/index.html>.
- [35] C. A. Klein, J. Pappis, et al., "Zns, znse, and zns/znse windows: their impact on flir system performance," *Optical Engineering*, vol. 25, no. 4, p. 254519, 1986.
- [36] "Perception tutorials." <https://wiki.perception-park.com/display/MAN/Hyperspectral+imaging+step+by+step>. Accessed: 2020-02-29.
- [37] R. Goh, L. Liu, X. Liu, and T. Chen, "The cmu face in action (fia) database," in *AMFG*, vol. 5, pp. 255–263, Springer, 2005.
- [38] W. Cho, A. Koschan, and M. A. Abidi, "Hyperspectral face databases for facial recognition research," in *Face Recognition Across the Imaging Spectrum*, pp. 47–68, Springer, 2016.
- [39] H. Bouchech, *Selection of optimal narrowband multispectral images for face recognition*. PhD thesis, Dijon, 2015.
- [40] W. Cho, A. Koschan, and M. A. Abidi, "Hyperspectral face databases for facial recognition research," in *Face Recognition Across the Imaging Spectrum*, pp. 47–68, Springer, 2016.
- [41] A. S. Georgiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 6, pp. 643–660, 2001.
- [42] C. J. Behrend, C. P. Tarnowski, and M. D. Morris, "Identification of outliers in hyperspectral raman image data by nearest neighbor comparison," *Applied spectroscopy*, vol. 56, no. 11, pp. 1458–1461, 2002.
- [43] L. Zhang and M. J. Henson, "A practical algorithm to remove cosmic spikes in raman imaging data for pharmaceutical applications," *Applied spectroscopy*, vol. 61, no. 9, pp. 1015–1020, 2007.
- [44] C. V. Cannistraci, F. M. Montevecchi, and M. Alessio, "Median-modified wiener filter provides efficient denoising, preserving spot edge and morphology in 2-de image processing," *Proteomics*, vol. 9, no. 21, pp. 4908–4919, 2009.
- [45] Å. Rinnan, F. van den Berg, and S. B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra," *TrAC Trends in Analytical Chemistry*, vol. 28, no. 10, pp. 1201–1222, 2009.
- [46] J. Burger and P. Geladi, "Hyperspectral nir image regression part i: calibration and correction," *Journal of chemometrics*, vol. 19, no. 5-7, pp. 355–363, 2005.
- [47] L. Wang and C. Zhao, "Anomaly detection technique of hsi," in *Hyperspectral Image Processing*, pp. 217–256, Springer, 2016.
- [48] J. M. Amigo, "Practical issues of hyperspectral imaging analysis of solid dosage forms," *Analytical and bioanalytical chemistry*, vol. 398, no. 1, pp. 93–109, 2010.
- [49] R. Leardi, "Genetic algorithms in chemometrics and chemistry: a review," *Journal of chemometrics*, vol. 15, no. 7, pp. 559–569, 2001.
- [50] H. Wang, T. C. Bau, and G. Healey, "Expression-invariant face recognition in hyperspectral images," *Optical Engineering*, vol. 53, no. 10, pp. 103102–103102, 2014.
- [51] Y. Rangaswamy, K. Raja, and K. Venugopal, "Frdf: Face recognition using fusion of dtcwt and fft features," *Procedia Computer Science*, vol. 54, pp. 809–817, 2015.
- [52] R. Raghavendra, S. Venkatesh, K. B. Raja, F. A. Cheikh, and C. Busch, "Mutual information based multispectral image fusion for improved face recognition," in *Signal-Image Technology & Internet-Based Systems (SITIS), 2016 12th International Conference on*, pp. 62–68, IEEE, 2016.
- [53] V. Sharma and L. Van Gool, "Image-level classification in hyperspectral images using feature descriptors, with application to face recognition," *arXiv preprint arXiv:1605.03428*, 2016.
- [54] V. Sharma, A. Diba, T. Tuytelaars, and L. Van Gool, "Hyperspectral cnn for image classification & band selection, with application to face recognition," 2016.
- [55] G. Chen, C. Li, and W. Sun, "Hyperspectral face recognition via feature extraction and crc-based classifier," *IET Image Processing*, vol. 11, no. 4, pp. 266–272, 2016.
- [56] H.-I. Kim, S. H. Lee, and Y. M. Ro, "Multispectral texture features from visible and near-infrared synthetic face images for face recognition," in *Multimedia (ISM), 2015 IEEE International Symposium on*, pp. 593–596, IEEE, 2015.
- [57] N. Vetrek, R. Raghavendra, 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*, p. 78, ACM, 2016.
- [58] M. Peng, C. Wang, T. Chen, and G. Liu, "Nirfacenet: A convolutional neural network for near-infrared face identification," *Information*, vol. 7, no. 4, p. 61, 2016.
- [59] N. Vetrek, 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)*, pp. 324–329, IEEE, 2016.
- [60] W. Cho, J. Jang, A. Koschan, M. A. Abidi, and J. Paik, "Hyperspectral face recognition using improved inter-channel alignment based on qualitative prediction models," *Optics express*, vol. 24, no. 24, pp. 27637–27662, 2016.
- [61] B. Du, M. Zhang, L. Zhang, R. Hu, and D. Tao, "Pltd: Patch-based low-rank tensor decomposition for hyperspectral images," *IEEE Transactions on Multimedia*, vol. 19, no. 1, pp. 67–79, 2017.
- [62] J. Liang, J. Zhou, and Y. Gao, "3d local derivative pattern for hyperspectral face recognition," in *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, vol. 1, pp. 1–6, IEEE, 2015.
- [63] X. Han, M. Hou, G. Zhu, Y. Wu, and X. Ding, "Extracting graphite sketch of the mural using hyper-spectral imaging method/dobivanje grafite skice murala primjenom metode hyper-spectral imaging," *Tehnicki Vjesnik-Technical Gazette*, vol. 22, no. 6, pp. 1567–1578, 2015.
- [64] G. Chen, C. Li, and W. Sun, "Hyperspectral face recognition via feature extraction and crc-based classifier," *IET Image Processing*, vol. 11, no. 4, pp. 266–272, 2016.
- [65] X.-Y. Jing, F. Wu, X. Zhu, X. Dong, F. Ma, and Z. Li, "Multi-spectral low-rank structured dictionary learning for face recognition," *Pattern Recognition*, vol. 59, pp. 14–25, 2016.
- [66] L. Deng, D. Yu, et al., "Deep learning: methods and applications," *Foundations and Trends® in Signal Processing*, vol. 7, no. 3–4, pp. 197–387, 2014.
- [67] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational intelligence and neuroscience*, vol. 2018, 2018.
- [68] A. Signoroni, M. Savardi, A. Baronio, and S. Benini, "Deep learning meets hyperspectral image analysis: A multidisciplinary review," *Journal of Imaging*, vol. 5, no. 5, p. 52, 2019.

- [69] F. Taherkhani, J. Dawson, and N. M. Nasrabadi, "Deep sparse band selection for hyperspectral face recognition," *arXiv preprint arXiv:1908.09630*, 2019.
- [70] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational intelligence and neuroscience*, vol. 2018, 2018.
- [71] X. Ma, J. Geng, and H. Wang, "Hyperspectral image classification via contextual deep learning," *EURASIP Journal on Image and Video Processing*, vol. 2015, no. 1, p. 20, 2015.
- [72] W. Zhao, Z. Guo, J. Yue, X. Zhang, and L. Luo, "On combining multiscale deep learning features for the classification of hyperspectral remote sensing imagery," *International Journal of Remote Sensing*, vol. 36, no. 13, pp. 3368–3379, 2015.
- [73] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 844–853, 2016.
- [74] A. Ioannidou, E. Chatzilari, S. Nikolopoulos, and I. Kompatsiaris, "Deep learning advances in computer vision with 3d data: A survey," *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, p. 20, 2017.
- [75] A. Santara, K. Mani, P. Hatwar, A. Singh, A. Garg, K. Padia, and P. Mitra, "Bass net: Band-adaptive spectral-spatial feature learning neural network for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 9, pp. 5293–5301, 2017.
- [76] B. Pan, Z. Shi, and X. Xu, "Mugnet: Deep learning for hyperspectral image classification using limited samples," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 108–119, 2018.
- [77] W. Song, S. Li, L. Fang, and T. Lu, "Hyperspectral image classification with deep feature fusion network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3173–3184, 2018.
- [78] J. M. Haut, M. E. Paoletti, J. Plaza, J. Li, and A. Plaza, "Active learning with convolutional neural networks for hyperspectral image classification using a new bayesian approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6440–6461, 2018.
- [79] N. He, M. E. Paoletti, J. M. Haut, L. Fang, S. Li, A. Plaza, and J. Plaza, "Feature extraction with multiscale covariance maps for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 755–769, 2018.
- [80] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *Journal of Sensors*, vol. 2015, 2015.
- [81] S. M. Iranmanesh, B. Riggan, S. Hu, and N. M. Nasrabadi, "Coupled generative adversarial network for heterogeneous face recognition," *Image and Vision Computing*, vol. 94, p. 103861, 2020.
- [82] J. Yue, W. Zhao, S. Mao, and H. Liu, "Spectral-spatial classification of hyperspectral images using deep convolutional neural networks," *Remote Sensing Letters*, vol. 6, no. 6, pp. 468–477, 2015.
- [83] P. Barré, B. C. Stöver, K. F. Müller, and V. Steinhage, "Leafnet: A computer vision system for automatic plant species identification," *Ecological Informatics*, vol. 40, pp. 50–56, 2017.
- [84] Y.-D. Zhang, Z. Dong, X. Chen, W. Jia, S. Du, K. Muhammad, and S.-H. Wang, "Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 3613–3632, 2019.
- [85] J.-J. Lv, X.-H. Shao, J.-S. Huang, X.-D. Zhou, and X. Zhou, "Data augmentation for face recognition," *Neurocomputing*, vol. 230, pp. 184–196, 2017.
- [86] W. Li, C. Chen, M. Zhang, H. Li, and Q. Du, "Data augmentation for hyperspectral image classification with deep cnn," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 4, pp. 593–597, 2018.
- [87] J. Nalepa, M. Myller, and M. Kawulok, "Hyperspectral data augmentation," *arXiv preprint arXiv:1903.05580*, 2019.
- [88] L. Windrim, R. Ramakrishnan, A. Melkumyan, and R. Murphy, "Hyperspectral cnn classification with limited training samples," *arXiv preprint arXiv:1611.09007*, 2016.
- [89] O. Vinyals, C. Blundell, T. Lillicrap, D. Wierstra, *et al.*, "Matching networks for one shot learning," in *Advances in neural information processing systems*, pp. 3630–3638, 2016.