Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

Face Recognition: Innovative Advances and Future Impact

Neeraj Pratap[1], Sunil Kumar[2], Suraj Malik[3], Bhanu Pratap Singh[4], Himanshu[5]

[1,5]Department of Computer Science and Engineering, Meerut Institute of Technology, Meerut

[2,4]School of Information Technology, AURO University, Surat

[3]Department of Computer Science and Engineering, IIMT University, Meerut

(Received: 07 January 2024 Revised: 12 February 2024 Accepted: 06 March 2024)

KEYWORDS

Face recognition, Multispectral image, Hyperspectral images, Infrared images

ABSTRACT:

Images in which secondary channels are obtained in several levels or spectral extents of the Electromagnetic Scale (EMS) are referred to as multidimensional images (also known as hyperspectral or multispectral). Everyone has a claim to information that can be manipulated in face recognition (FR) applications. Apart from Visible Spectrum (VS) data, a facial investigation system can also utilize infrared (IR) imagery to identify presentation attacks, stance variations, and facial look dissimilarities. This research presents an evaluation of current multi-modal face recognition (MDFR) algorithms using imagery from VS, the Near Infrared (NIR), Short Wavelength Infrared (SWIR), and Long Wave Infrared (LWIR) sub-bands of EMS. An assessment about their requirement is conducted, and the public multispectral and hyperspectral datasets for facial analysis are acknowledged. The basic operating standards of the MDFR approaches are categorized, ranging from the traditional Fusion and Subspace schemes to the most recent Deep Neural Networks (DNN).

1. Introduction

There is now a promise to witness an increase in the number of apps utilizing FR systems, whether they are meant for individual use, like those in android phones, or for communal use, like those in industries. Using several spectral extents improves the results in FR. In FR, images can be obtained in two main modes. First, in a controlled environment where a person helps to take pictures, and second, in an uncontrolled environment, also known as "in the wild," where a person either doesn't help or doesn't know at some time during the picture taking process.

Systems that exclusively employ the VS provide a number of challenges, including as occlusions, position differences, personalized assistance, and, for the most part, problematic brightness variations. To lessen the difficulties that result from it, it must be supplemented with the use of other biometric detectors (such as those that recognize iris or dermatoglyphics) or EMS bands. As a supplement to the VS in FR systems, the infrared scale—that is, the NIR, SWIR, MWIR, and LWIR spectral bands—has been effectively employed. [1], [2]. Given that they employ many spectral bands, the most commonly utilized spectral groups that are helpful in FR are listed in Table 1 and are referred to as multispectral or hyperspectral.

Spectral Band Name	Wavelength(μm)
Visible Spectral(VS)	0.38 μm -0.75 μm
NIR Scale	0.75 μm -1.40 μm
SWIR Scale	1.40 μm -3.00 μm
MWIR Scale	3.00 μm -8.00 μm
LWIR Scale	8.00 μm -15.00 μm

Table 1: Details of Spectral ranges in EMS [3]

The infrared spectrum has several advantages over the visible spectrum as compared to the visible spectrum. However, it is not as perceptible to human eye and does not provide as much information about brightness variations. Using LEDs that secrete in IR scale to illuminate the scene—a technique sometimes utilized in film surveillance—nighttime surveillance can be carried out covertly. NIR and SWIR being relatively close to the VS allowed for an easy modification of the standard learning process skilled with images of the VS. The spectral bands MWIR and LWIR enable the use of FR systems at night, when brightness is very low or even nonexistent.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

Most multispectral or hyperspectral FR systems included the following processes, which are illustrated in the outline: picture acquisition, facade finding, facade position, characteristics mining, and categorization.

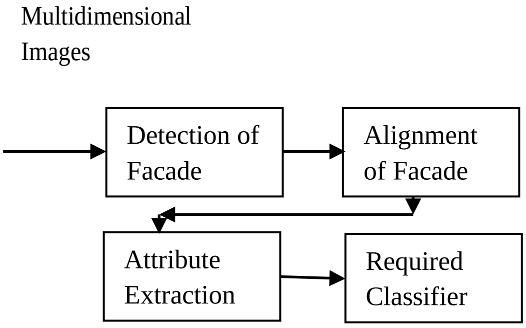

Fig. 1: Architecture of Multidimensional FR System

The attainment of multidimensional pictures is where the structures start. A facade finding involves using the illustration to locate a facade-bound container so that things that don't belong to the person can be recognized and separated.

The following phase involves using face familiar sight to perform a facial pose. Facial familiar sights are distinct facial structures like the nose, eyeball, jaw, and so forth. After removing the familiar face sight, a facade change can be completed from start to finish using the eyeball or the eyes combined with the maw. .

Facial landmarks and facial discoveries can be retrieved from the full set of EMS photos, or just the VS band illustration. Compared to additional EMS bands, visible images produce better results since facade detection and familiar sight extractor models are more adept at VS illustration. It is preferable to use the VS illustration for this reason. If the imagery is taken in a similar position at a parallel time, further spectral band descriptions can be provided to the bounding container and face landmarks.

The primary goal of the feature drawing out phase eliminates each person's fundamental identity. This phase is solely dependent on the plan employed. It is feasible to extract the personality of the person in the image by classifying the features, or embeddings, that indicate the individuality.

By using MDFR systems instead of merely visible FR systems, an extra security layer can be employed to precisely identify a person in order to guarantee access only to the authoritative community while entering a high security location. These places can be found in military forces installations, schools, hospitals, and laboratories. [1].

The creation of a better FR system can guarantee a more dependable and healthful access power, protecting material goods and increasing common security. The manuscript's goal is to represent a methodical, in-depth literary works examination of the MDFR's current development. Unlike previous assessments in the subject [4]–[6], this book compiled several noteworthy contributions, as highlighted below.

We first conduct an evaluation of the most recent techniques in MDFR by utilizing only international journals with impact factors. We offer broad overviews of condensed and summarized data from the examined databases, containing all the material in one package to assist researchers in selecting the most suitable catalog for their study. Regular use of the databases is done in connection with the real study, although Google Scholar citations were utilized for similar assignments in the other review submissions. [5]. we disclose how our study is conducted, allow other researchers to summarize the manuscript, and are particularly not hopeful to accomplish with more manuscripts.

Subdivision II sheds information on the meticulous analytical approach we used, which included a division by year research and study specialty. Subdivision III contains illustrations of the most often used databases and provides summarized statistics from the databases examined. Section V, the main body of the publication, explores the many techniques utilized in the investigated papers, which are gathered by schemes and publicizing era. Subdivision IV shows the performance assessment scheme used in MDFR. Section V also includes a final result and future developments in the MDFR field.

A. Motivation, Objectives, and Contributions

Due to its abundant information in every subject's face cube in multidimensional face datasets, MDFR offers improved categorization rates. Thin band information acquisition allows multidimensional images to show extensive information about any entity. Every pixel in multidimensional imagery represents a whole spectrum, and they are frequently used to address a wide range of problems in various fields like defense research, municipal investigation, harvest investigation,

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

geographical mapping, stone investigation, and armed forces supervision. Multidimensional imaging is therefore a highly popular technique for analyzing and researching many issues [86].

Recently, the decreasing cost of multidimensional sensors and the increasing speed of imaging have attracted computer visualization researchers to use multidimensional imaging to solve computer vision problems such as substance classification, crop cultivation, chemistry, and manuscript illustration research. A series of photos taken at different adjacent wavelengths is what makes up multidimensional images. There is a range in the number of wavelengths from a dozen to several hundred layers. More detailed information about the skin and facial tissues can be found in a multidimensional image of a facade. The image is shaped like a cube, with each image within the cube representing the spatial importance gathered at a specific wavelength. The more advantageous spectrum information can be used to address several major FR problems, such as pose and lighting. Several FR constraints are circumvented by multidimensional images, including lightning; pose changes, face appearance, and many others. As a result, researchers have been drawn to work in this specific field by MDFR.

Deep learning, or DNN, has demonstrated remarkable results recently in a variety of computer vision applications. Multiple uses of DNN for hyperspectral imaging were reported in a recent review. CNN has been projected for the MDFR's sparse band selection in the report. Even yet, more research on MDFR using DNN is desired [85]. The details of the following facts are being highlighted with regard to contributions.

- The paper that is being given examines multidimensional FR in great detail.
- The importance of multidimensional imaging is investigated in order to lessen the difficulties associated with FR, such as lightning circumstances, position fluctuations, and face emotions.
- A thorough architecture and methodology for multidimensional FR are shown.
- An extensive examination of the literature is used to examine the significance of deep neural networks (DNNs).

2. Methodical Review

This segment presents a thorough analysis of the papers in MDFR, broken down by years and fields of research. Only articles that achieve FR or facial detection with two or more EMS bands, such as VSNIR-LWIR, NIR-VS, or some other promising arrangements, have been selected for the current analysis. All articles published in academic journals between the years 2000 and 2022 with a good citation score and IF have been preferred.

Figure 2 reveals that the relevance of MDFR has fully increased. Over the last seven years, from 2016 to 2022, there has been a noticeable increase in manuscripts as compared to previous years. There are three plausible explanations for this phenomenon. Firstly, the expense of infrared cameras, particularly those that capture images in the near-infrared and mid-infrared (EMS) range, is prohibitive. (ii) It is necessary to minimize human intervention in access control so that human resources can be assigned to other duties. (iii) The application of deep learning in FR systems yields really encouraging outcomes.

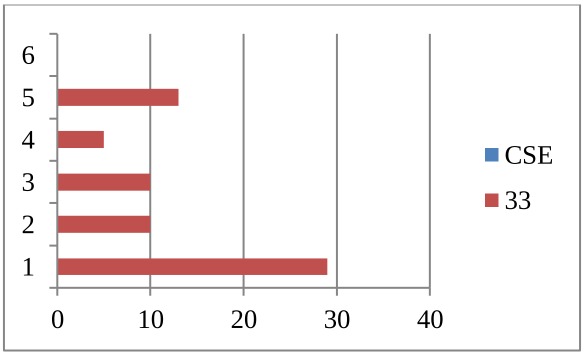

Fig. 2 Details of Manuscript Published (Year Wise)

Figure 3 shows the article split of the selected papers according to their research study field. This figure shows that the two primary fields of MDFR research are engineering (29%), and computer science (33%). The growing use of DNN, which is a modern and creative technology, is also being dispersed to the MDFR program. It seems sense that it is being taken into consideration because of this.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

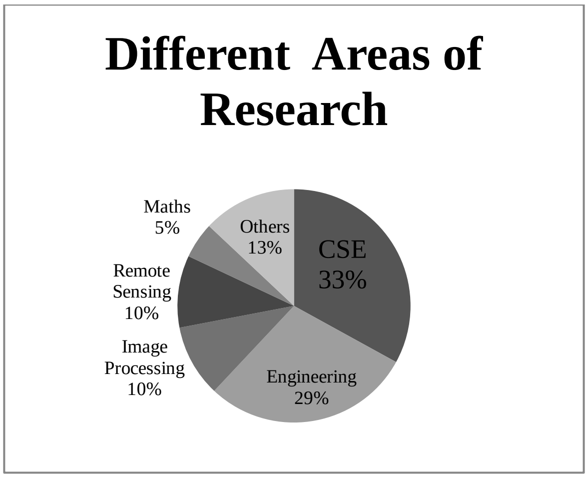

Fig. 3 Related Manuscripts Published in different domains

3. Multidimensional Databases

Most relevant manuscripts made use of the public databases this section illustrates. The first segment's research of database occurrence use is followed by a study of their attributes, which highlights the contrasts and similarities between them. While private databases are solely used by their authors, shared datasets (the dark blue bar in Figure 4) are often used and allow the evaluation of many schemes, making the investigator's task of selecting the best dataset for its purpose easier.

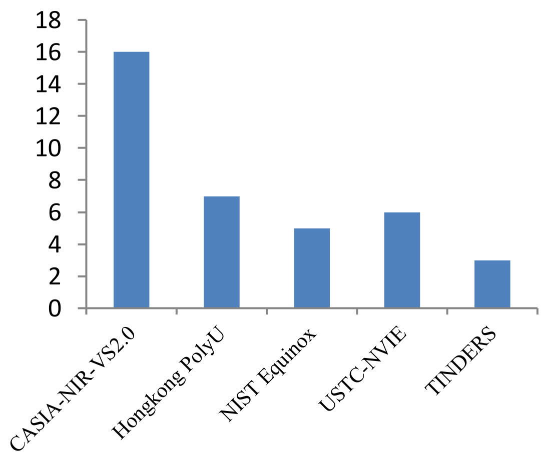

Fig. 4 Details of Public Databases available and number of times used

The CASIA NIR-VS 2.0 datasets are distinct from other datasets due to the processes from the dataset that have already been investigated, such as which imagery to use in the training and assessment stages. The database is comprised of two datasets: the original photos and equivalent imagery in resolution.

Table 2 lists publicly available databases for educational use. These databases are arranged by provided name, year or day of creation, EMS spectral bands chosen, population number, number of images per person, and number of images in dataset. Several observations are also given.

Table 2 shows that most datasets are mature enough to have been around for a normal duration of eight or nine years. One other pertinent detail is extracted using Table 2, which shows that the standard population number in the multidimensional database is 138, which is small when compared to the datasets containing photos from VS.

According to Masi et al. [62], a large number of photos of dissimilar people are helpful for training DNN because they cover the huge range of individual appearance. A database including a number of photos of the same person with different brightness changes and position conditions enables better learning through a DNN, as achieved by Masi [62]. Consequently, a larger image collection of all humans enables us to retrain DNNs that were previously trained on datasets containing imagery of various populations. Figure 5 shows the division of the spectral bands (NIR, SWIR, MWIR, and LWIR) by the shared multidimensional databases.

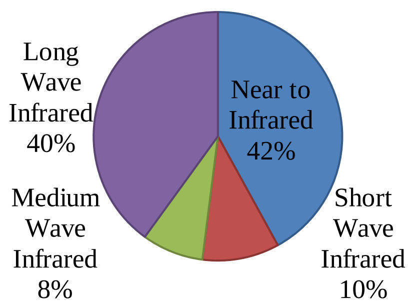

Fig. 5 Division of Spectral Bands used in EMS for FR

The fact that there are actually very few databases with SWIR and MWIR imagery—10% and 8%, respectively—is due to the high cost of SWIR and MWIR cameras as compared to NIR or LWIR cameras.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

Name of Database	Year	Spectral Range	Number of images	Best Rank-1	Features
ASUMS	2011	VS, LWIR	576	97.9	With 4 brightness variations
BUAA-VisNir	2012	VS, NIR	24,300	97.4	With 9 diverse facial expressions.
Carl Dataset	2013	VS, NIR, LWIR	7,380	75.6	With 3 luminosity variations and 5 different facial expressions.
CASIA NIR-VIS 2.0	2013	VS, NIR	17,580	99.4	Contains 725 subjects, imaged by VS and NIR cameras in four recording sessions.
CASIA-HFB	2009	VS, NIR	1,616	95.2	With changes in facial expressions
HK PolyU-HSFD	2010	VS, NIR	22,500	99.8	With 2 luminosity variations, 3 different poses and 2 different facial expressions
IRIS	-	VS, LWIR	4,228	96.0	With 5 luminosity variations and 3 different facial expressions
LDHF	2014	VS, NIR	800	78.0	With 2 luminosity variations and 4 different distances human-camera
NIST Equinox	2007	VS,SWIR,MWIR, LWIR	-	99.6	With 3 luminosity variations and 3 different facial expressions
NIVI.	2012	VS, NIR	24,605	94.5	
NVESD	2013	VS,MWIR, LWIR	-	82.3	
Oulu-CASIA NIR-VIS	2009	VS, NIR	2,880	99.9	With 3 luminosity variations and 6 different facial expressions
Polarimerrie Thermal	2019	VS, LWIR	-	98.0	With 2 luminosity variations
Pola Thermal	2016	VS, LWIR	Video	76.3	With several different facial expressions, 3 different distances human-camera
RGB-D-T	2016	VS, LWIR	45,900	86.9	-
TINDERS	2009	VS, NIR,SWIR	1,255	97.8	With 2 different facial expressions
UGC-JU	2015	VS, LWIR	6,552	99.2	With 22 different poses (one with glasses and) 2 different facial expressions
UND-XI	2004	VS, LWIR	4,584	99.1	-
USTC-NVIE	2010	VS, LWIR	34,830	97.4	With 3 luminosity variations, 9 different poses and 5 different facial expressions
UWA-HSFD	2013	VS, NIR	3,960	99.8	-
WVU NIR, Mid-Range	2015	VS, NIR,SWIR, LWIR	515 videos	56.0	With 3 luminosity variations

Table 2: Detailed description of public Multidimensional databases

4. Performance Assessment

Most often, FR is used to identify or validate identities by comparing a person's identification in a dataset with additional individuality, which shows a mapping of one-to-many assessment. When evaluating recognition, we differentiate between two types of problems: closed-set and open-set, depending on whether the human subject to be recognized exists in the given dataset or not.

When the employee's individuality is compared to the claim made about it, it is verified whether or not it is accurate. This is accomplished by having a human verify the individuality by inputting their identification number or validating their identification tag with the information that supports the individuality. Since this means that we are working with a one-to-one confirmation, the recognition process, although involuntary, does not require any human intervention. Nevertheless, it has the drawback that if the imaging dataset is really large, this process could take too long because it would have to go through the complete collection of images.

To evaluate the effectiveness of a certain method in comparison to others, conventional techniques like Computational instance, Rank-N, and False Acceptance Rate (FAR) [35] must be taken into consideration.

Matching performance is defined as the percentage of recognition attempts for which the facade imagery forecast is acquired in the highest ranking N outcome. While rank-10 refers to the fraction of face image predictions that appropriately correspond to their alike identities in the top ten highest gain outcomes, rank-1 is the fraction of forecasted

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

individuality that provide their similarity as accurate (forecasted appropriately individual individuality).

The total number of images that are properly recognized (Appropriately Recognized) is divided by the total number of identifications made (Recognition Attempts) to determine rank one:

$$\text{Rank} - \text{One}(\%) = \frac{AR}{RA} \times 100$$

Although rank-K is an addition of rank one, in this case, it is examined to see if the accurate image is among the top K imagery options in order to determine whether the most persuasive image is the correct one.

The False Positive Rate (FPR) approximates the likelihood that the system would incorrectly classify an identification system illustration fit in to the claimed individuality (impostor) when the input image actually matches to a dissimilar subject (right individual). One of the serious biometric safety issues caused by this depiction is the granting of unauthorized permission to users [2]. FPR is a measure used in admission control systems to quantify the probability that an unauthorized user will be granted access to an identifying system, such as the FR system.

$$FPR\ \% = \frac{False\ Positive}{False\ Positive + True\ Negative} \times 100$$

Where the numbers for inappropriate authorizations are represented by False Positive (FP) and True Negative (TN). As an example, an identification system with a 2% false positive rate (FPR) indicates that, out of a total of hundreds of attempts, 98 were genuinely accurate and 2 were wrong.

To compute the performance assessment, the algorithm's processing time is also calculated. When several algorithms reach similar rank values with fixed FPR values, the authors calculate the amount of time needed to identify a specific number of persons to demonstrate the superiority of their scheme.

5. Methods

This subgroup mostly represents acceptable texts in the MDFR area. The manuscripts have been categorized based on the techniques employed. Each method's study includes a description of the various methods each researcher took on, the dataset they used, the results they produced, and the conclusions they came to.

In the process of studying each paper, three distinct approaches to the accomplishment of FR methods have been noted: Multi-layer (ML) to Multi-layer (ML), ML to Single-layer (SL), and SL to SL. In these cases, a layer may be a spectral layer or alternatively referred to as a spectral range inside EMS.

Throughout the thorough investigation, a sizable number of relevant publications with five essential techniques—Feature Illustration (FI), Coupled Subspace Learning (CSL), Image Synthesis (IS), Fusion, and DNN—are collected together. The table shows that DNN is the most often utilized methodology since it is a contemporary method that has produced good results in MDFR. Table 3 shows the five strategies as well as the publications that have used them.

S.No	Methods Name	Methods Used (In Percentage)
1	Feature Representation	11%
2	Coupled Subspace Learning	23%
3	Image Synthesis	15%
4	Fusion	23%
5	Deep Neural Network	28%

Table 3: Brief description (In Percentage) of different methods used in FR

Figure 6 exposes most of the studies that employed the fusion technique for MDFR up until 2018. Since then, most papers have adopted DNN because it produces better results.

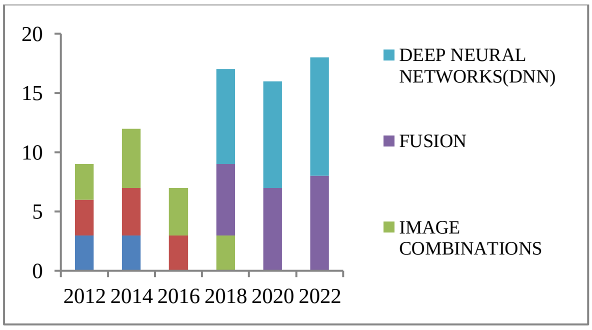

Fig. 6 Brief description of paper published using different methods for FR

A. Feature Illustration (FI)

Using techniques based on FI, the dominant features or traits that are constant across the EMS subbands used in each illustration are attempted to be retrieved. Eliminating features like corners, noses, eyes, maws, and so on reduces the initial image data and eliminates irrelevant information; by using this method, the classifier's calculations become easier to understand. It is possible to reduce the method space that links the various spectral levels in EMS in this way [21].

This method can be used independently, suggesting that it is the only method for performing FR, or it can be used as a starting point for new FR methods [21]. One limitation of feature extraction approaches like LBP is that they tend to ignore the facade spatial configuration, which makes a great presentation in various FR systems all the more important [22].

Nicolo and Schmid [54] processed the size and stages of the Gabor Filtered Imagery (GFI) using methods called Simplified Weber Local Descriptor (SWLD), LBP, and a generalized LBP. A varied FR system connecting an illustration of the SWIR subbands to an illustration of the VS subbands band was presented. The three histograms, each with 135 bins, combine to form a single feature vector. The researchers used symmetric Kullback-Leibler (KL) divergence distance to assess the feature vectors (FV) they extracted from two images. The TINDERS database [52] employed the aforementioned strategy and yielded a classification rate of 97.8%.

The TINDERS dataset [52] is used by Cao et al. [53] to project the use of Composite Multilobe Descriptor (CMD) to extract attributes, and the symmetric KL divergence distance was employed to assess features retrieved from two images. Ninety one and five four (91.54%) authentication rate at one percent FPR and seventy point one four percent rank-1 are acquired from a varied NIR-VS FR, whereas in the spectral sub-bands of SWIR-VS, a percentage of 99.46 authentication at one percent FPR was obtained. Researchers using various techniques, such as HOG and LBP, compared their results, and they found that the projected technique produced better results.

Shamia and Chandy [41], who had utilized a grouping of HOG and LBP to extract facial attributes or characteristics from NIR imagery to carry out FR at remoteness of one, sixty, one hundred, and one hundred fifty meters, evaluated the features extracted from the two images using the Euclidian Distance (ED). For the LDHF dataset [40], which contains VS and NIR images recorded at various remoteness of one meter, sixty meter, one hundred meter, and one hundred fifty meters accordingly, a rank-1 gain of 72%, 78%, and 32% is achieved for distances of 60, 100.

Peng et al. [22] created a graphic demonstration based HFR utilizing the CASIA NIR-VIS 2.0 [7] and USTC-NVIE [9] dataset. The HFR used a Markov Network (MN) [75] to represent different image patches discretely, considering the spatial compatibility among neighboring image patches. After evaluating the suggested approach, accuracy for the relevant datasets is 83.32% and 95.38%, respectively.

Peng et al. [22] created a graphical illustration-based HFR that used an MN [75] whose links represent equal probability dependencies to express different image patches independently, taking into account the spatial compatibility involving nearby image patches. When the CASIA NIR-VS 2.0 [7] and USTC-NVIE [9] databases were tested using the approach, the corresponding accuracy for the given databases was 83.32% and 95.38%.

Peng et al. [27] predicted the use of Sparse Graphical Representation based Discriminant Analysis (SGR-DA) to represent various facial imagery of various spectral layers. MN is used to generate adaptive sparse vectors, which are shown to be particularly useful for diverse FR. The application of the spatial partition approach improves the ability to distinguish between different face photos. After processing two datasets, CASIA NIR-VS 2.0 [7] and USTC-NVIE [9], the projected scheme obtained accuracy of 87.84% and 93.08%, respectively. Sparse Graphical Representation based Discriminant Analysis produced an increase of 4.52% for CASIA NIR-VS database and a drop of 2.30% for USTC-NVIE database, according to a comparison of the method with graphical illustration based HFR [22].

B. Coupled Subspace Learning(CSL)

The CSL technique is mostly useful for identifying the relevant information that represents the properties of various spectral layers inside a regular subspace. While the current method can lessen variation between images from different spectral bands, it also significantly reduces CSL's discriminative effectiveness in situations where modality space is vast [21]. This scheme's drawback is that, when mapping illustrations on regular subspaces, material is continuously overlooked that could lessen the FR organism's appearance. [22]. Huang et al. [69] have created a novel technique called Discriminative Spectral Regression (DSR), which maps the

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

S.No	Authors	Year	Brief Description
1	Yi Jin, Jiwen Lu, Qiuqi Ruan[21]	2015	Introduced an approach for diverse FR called coupled discriminative feature learning (CDFL).
2	Chunlei Peng, Xinbo Gao, Nannan Wang, Jie Li[22]	2017	A new technique based on graphical representation was applied, and MNs were used to individually represent heterogeneous image patches.
3	Robert B. Martin, Mikhail Sluch, Kristopher M. Kafka, Robert Ice, Brian E. Lemoff[52]	2013	An imaging device was created that could provide close-up facial images at night and at a distance of several hundred meters.
4	Cao Zhicheng, Schmid Natalia A., Bourlai Thirimachos[53]	2016	The development of a unique form and two alternative local operators for facial characteristic extraction is based on the cross-spectral similarity between multiple VS images and NIR, SWIR, MWIR, and LWIR.
5	D. Shamia, D. Abraham Chandy[41]	2018	Using the LDHF Dataset, which consists of VS and NIR imagery, a unique approach for nighttime FR has been proposed. The most crucial phases, including pre-processing, feature removal, and similarity, have been explored and proven.
6	Dongoh Kang, Hu Han, Anil K. Jain, Seong-Whan Lee[40]	2014	Created a system capable of collecting façade imagery at a great distance during both day and night, and offered a more advanced, varied FR method for cross-spectral and cross-distance facade similarity.
7	J. Pearl[75]	1998	The author searched for a useful visual representation of the interdependence ingrained in probabilistic modeling.
8	Chunlei Peng, Xinbo Gao, Nannan Wang, Jie Li[27]	2019	A concept for an adaptive sparse pictorial demonstration was created to represent mixed facade imagery. It was discovered that adaptive sparse vectors could be generated using MN's model, and this method was found to be highly effective for mixed FR.

Table 4: Brief descriptions of FR based on Feature Illustration Technique

facial imagery of VS and NIR in a regular discerning subspace. On the CASIA-HFB dataset [32], an accuracy of 95.33% is reached.

Using the CASIA NIR-VIS 2.0 [7] database, Jin et al. [21] introduced the Coupled Discriminative Feature Learning (CDFL) feature extraction method, which is applied for various FR. By using this method, the intraclass variations are decreased and the interclass variations are increased. Applying the suggested technique to the given database results in rank-1 authentication rates with FPR at 1% and authentication rates with FPR at 0.1%, which are, respectively, 71.5%, 55.1%, and 67.7%.

Applying the Mutual Component Analysis (MCA) established by Li et al. [11] to the CASIA NIR-VIS 2.0 database yields an accuracy of 92.71% when examining the characteristics common to the two forms of images in VS and LWIR.

Hu et al. [47] developed the Difference of Gaussian (DoG) filter during the preprocessing stage to reduce brightness variations in VS imagery and local deviations in LWIR imagery. In addition, a sixteen-by-sixteen-pixel Histogram of Oriented Gradients (HoG) is used to extract features from images. The pre-processing and feature extraction stages are intended to reduce the modality space that lies between VS and LWIR imaging. Using the NVESD [46] database, diverse FR LWIR-VS at remoteness's of one, two, and four meters is carried out, yielding accuracy values of 82.3%, 70.8%, and 33.3%, respectively. Furthermore, varied FR MWIR-VS was conducted at 1 m, 2 m, and 4 m distances, yielding accuracy values of 92.7%, 81.3%, and 64.6%, in that order. Using the UND-X1 [58] dataset, an accuracy of 72.7% is likewise obtained for different FR in the LWIR-VS spectral region.

Gong et al. [67] created a novel feature descriptor termed the Common Encoding Feature Discriminant (CEFD) technique to perform various FR on the CASIA NIR-VIS 2.0 [7] dataset. This approach eliminated the significant modality gap between the NIR and VS pictures and achieved an accuracy of 85.6%.

Lei et al. [33] created the Coupled Discriminant Analysis Method (CDAM), which uses two LCKS practices—LCKS-CDA and LCKS-CSR—for varying FR between VS and NIR pictures. The technique is used on the CASIA-HFB database [32], where the accuracy of the LCKS-CDA method was

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

73.18%, while the accuracy of the LCKS-CSR method was 81.43%.

Prototype Random Subspace (PRS) is used by Klare and Jain [66] to perform various FR between VS-NIR images and VS-LWIR images on the CBSR database [76]. It is determined to be very promising to use diverse features lemmas to express the inquiry and gallery imagery. When this technique is evaluated against FaceVACS, LWIR images yield better results than NIR images, and a 95.8% accuracy level is attained.

The projected technique yields excellent results in varied FR and may be effective for identifying incorrect façade photos and facade appearance recognition. Bhowmik et al. [38] use log-ICA as an alternative to ICA, resulting in two topologies called logarithmic-ICA I and logarithmic-ICA II, with logarithmic-ICA II yielding better results. For Logarithmic-ICA I and Logarithmic-ICA II, respectively, in the first database, accuracy is eighty eight point one eight percent and ninety five percent; for dataset USTC-NVIE [9], accuracy is 95.9% and 97.4%, respectively, using Log-ICA I and Log-ICA II.

C. Image (Illustration) Combinations

The illustration combination practice, which converts an illustration from one layer of EMS to another layer of EMS, offers the benefit of being able to quickly compare two images and generate an illustration in the VS range with an illustration from another layer of EMS, such as LWIR. The main advantage of image combining is that when a LWIR illustration is combined as a VS image, FR methods intended for VS images can be used [66]. The main issue with this technology is that image synthesis is a challenging process, and the accuracy of the synthesized image is dependent on the efficacy of the FR system in most circumstances [22].

Osia and Bourlai [49] created equivalent images in the VS spectral region using LWIR images, and the benefits of the projected approach were demonstrated by applying FR to the synthesized images using LBP. The method was used to the NVESD dataset [46], which consists of VS, MWIR, and LWIR pictures. After investigation, FR MWIR-VS had a 75.3% accuracy rate, while various FR LWIR-VS had an 81.4% correctness rate after completion. Cao et al. [59] put up combined images into the learning progression using a data augmentation-based combined learning approach. The aggregated data increased the size of the intraclass set, perhaps improving diversified information. VS and LWIR pictures from the USTC-NVIE collection [9] were used, and an accuracy of 95.35% was obtained.

The author modified the FusionNet (FN) structure [77] and its training processes to decrease overfitting, add up dropout after bridge connections, ReLUs, and orthogonal regularization. Litvin et al.'s proposal [14] to combine LWIR and VS images was also modified. A technique was used to achieve an accuracy of 86.94%, 97.52%, and 99.19%, respectively, for each of the three illustration distinctions found in the RGB-D-T [13] dataset: posture appearances and intensity distinction.

He et al. [18] developed the Adversarial Cross-spectral Face Completion (ACFC), which synthesized VS images from NIR vision using a generative adversative algorithm. The employment of an inpainting element that combines and inpaints the VS illustration surface from the NIR illustration surface set this system apart from prior approaches. By using a warping technique, two components were combined into a continuous deep network. Additionally, a method was employed to create paired NIR and VS surfaces by transforming every stance in NIR imaging into a fore pose in VS imagery. By using a warping technique, two components were combined into a continuous deep network. Additionally, a method was employed to create paired NIR and VS surfaces by transforming every stance in NIR imaging into a fore pose in VS imagery. After that, LightCNN [78] was used to perform FR on the combined imagery. Three datasets—CASIA NIR-VS 2.0 [7], Oulu-CASIA NIR-VS [8], and BUAAVisNir [16]—were used to test the accuracy of ACFC, and the results were 98.6%, 99.9%, and 99.7%, respectively.

S.No	Authors	Year	Brief Description
1	Xiangsheng Huang, Zhen Lei, Mingyu Fan, Xiao Wang, Stan Z. Li[69]	2013	Strong categorization is achieved by mapping a variety of facade pictures into a regular discriminative subspace using a novel approach called DSR.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

2	Zhifeng Li, Dihong Gong, Qiang Li, Dacheng Tao, Xuelong Li[11]	2016	A novel system called MCA is investigated, wherein an abundant model is projected initially, and the EM method is then discussed to continuously learn model parameters.
3	Shuowen Hu, Jonghyun Choi, Alex L. Chan, and William Robson Schwartz[47]	2015	Following a discussion of preprocessing, feature extraction, and PLS model creation, it is stated that the proposed technique performs strongly for the tested state of affairs when addressing the cross-modal recognition problem utilizing PLS based on regression.
4	Dihong Gong, Zhifeng Li, Weilin Huang, Xuelong Li, Dacheng Tao[67]	2017	At the characteristic extraction stage, a novel characteristic lemma for diverse FR known as CEM may be greatly condensed. This lemma is capable of gathering regular diverse information such as enormous modality space.
5	Zhen Lei Shengcai Liao, Anil K. Jain, Stan Z. Li[33]	2012	In order to improve heterogeneous FR performance, a narrative coupled discriminant analysis method is created. It has the benefit that all samples from various modalities can be used to represent the coupled projections, allowing for the extraction of sufficient discriminative information.
6	Brendan F. Klare, Anil K. Jain[66]	2013	Images are represented using a generic framework that projects nonlinear similarities to a set of prototype face images. Projecting features onto a linear discriminant subspace improves the accuracy of this nonlinear prototype representation.
7	Mrinal Kanti Bhowmik, Priya Saha, Anu Singha, Debotosh Bhattacharjee, Paramartha Dutta[38]	2019	By applying a logarithmic transformation to the basic ICA, a new variant known as Log-ICA is created. This variant is capable of reducing the overlapping likelihood densities of the combined signal, which eliminates additional self-regulating components due to condensed gaussianity.

Table 5: Brief descriptions of FR based on coupled subspace learning

S.No	Authors	Year	Brief Description
1	Brendan F. Klare, Anil K. Jain[66]	2013	A novel feature descriptor for heterogeneous frequency range modeling, known as the common encoding model, may effectively capture common discriminant information, thereby significantly narrowing the enormous modality gap during the feature extraction stage.
2	Nnamdi Osia, Thirimachos Bourlai[49]	2017	Developed a method known as the MWIR (3–5 µm) and LWIR (8–14 µm) bands that bridges the VS and IR band of EMS. They also looked at the advantages and drawbacks of employing thermally synthesized visible facial pictures and vice versa.
3	Bing Cao, Nannan Wang, Jie Li, Xinbo Gao[59]	2019	Suggested a data augmentation-based joint learning (DA-JL) method that uses synthesized images to assist the learning process and mutually alter the cross-modality differences.
4	Andre Litvin, Kamal Nasrollahi, Sergio Escalera, Cagri Ozcinar, Thomas B. Moeslund, Gholamreza Anbarjafari[14]	2019	Suggested using fully convolutional network architecture to generalize RGB face images from a particular thermal input face image. The suggested technique, which uses dropout after bridge connections to boost robustness against overfitting, is based on the FusionNet architecture.
5	Ran He, Jie Cao, Lingxiao Song, Zhenan Sun[18]	2020	Created a high-illumination model for a variety of face combinations by combining two elements in parallel: position correction and texture inpainting. VS image textures are created by synthesizing and inpainting NIR image textures.

Table 6: Brief descriptions of FR based on image combinations

D. Fusion

After applying the proper illustration combination techniques to FR, a characteristic extractor was used to combine the characteristics of many illustration causes, resulting in the creation of a feature vector. To perform face segmentation or detection, a single characteristic vector that includes details about edges, corners, lines, and textures is computed and

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

combined [79]. The final characteristic vector's dimensionality is moderated by using characteristic combination [2].

The score fusion method improved classification effectiveness fundamentally by combining the output of multiple classifiers into a single, comprehensive classifier. The most used method for score combination is majority voting, which includes selecting the category in which classification occurs most frequently and sending that decision to the comprehensive classifier. Each classifier was given a dynamic weight, and those who showed poor presentation received a tiny weight. Consequently, one additional technique employed in the score combination—the modified weighted—assigned a reduced amount of significance in the global classification.

The use of picture combining in FR systems has several benefits, including a reduction in the rate of inaccuracy and decrease in effort when using multiple inexpensive cameras instead of one expensive camera [80].

In order to produce the merged image that is applied to achieve the gain, Singh et al. [43] performed an illustration combination of the VS and LWIR imagery using a Support Vector Machine (GSVM) to calculate both vigorously and nearby the weights. A 2-dimensional LPGT was applied to remove the overall facial characteristic, and LBP was used to remove the local facial features. After applying the score fusion and running tests, a verification rate of 99.91% and 99.54%, respectively, were obtained on the UND-X1 [58] and NIST Equinox [42] datasets.

Seal et al. [56] proposed a VS and LWIR picture fusion technique that combines the constructive information found in VS and thermal images using illustrative entropy, translation, the invariant wavelet transform, and Random Forests (RF). After testing on the UGC-JU database, an accuracy of 99.07% was attained [55].

Bourlai et al. [61] used the MFSDF with characteristic extraction using LBP, GABOR, and HOG. Eleven states of affairs were completed with a mixture of scores: three individual characteristics (LBP, GABOR, and HOG), six groupings of two characteristics (e.g., LBP C GABOR), the summation of the three individual characteristics, and a weighted combination proposal where weights were assigned to each descriptor based on the presentation scores. In order to determine which state of things obtains the best rank-1 gain, an experimental evaluation was conducted on the WVU NIR Mid-Range database [61]. Following this, MFSDF was assessed using different FR techniques, such as PCA and LDA.

Seal et al. [57] employed a synthesis procedure that computed the weighted addition of LWIR and VS fore face information by two weighting features. For this purpose, two independent FR were conducted, the primary on the VS image and the subsequent on the LWIR illustration. All of this formed a gain that was equivalent to the possibility of accurate categorization for each illustrate Applying the merged illustration created by employing the projected combination process, where the weights were the gains previously estimated in the following phase, resulted in an accuracy value of 98.42% for FR on UGC-JU [55].

Simón et al. [13] combined a single feature vector on the RGB-D-T database using LBP, HOG, HAAR, and HOGOM to extract the characteristics of VS, LWIR, and depth imagery. This single feature vector was then used to train the W-kNN, which is used to give points that are closer to one another more weight and points that are farther away less weight. CNN processed each of the original photographs, and the three images were combined. Finally, by merging these two classifiers with different weights, the final classifier was obtained.

Kanmani and Narasimhan [39] proposed three optimization-based combination strategies that support varied FR dilemmas. In the first and second methods, input illustration was broken down into high and low frequency coefficients using twofold tree DWT. A population-based optimization process [81] was then used to find the most advantageous weights for combining VS and LWIR imagery. With the IRIS [37] dataset and STPS Optimization, which is used to delay the untimely union of the particle swarm, curvelet transform, which is used to carry out illustration breakdown care for edges as well as for the curves, and Brain storm optimization scheme, which is applied for optimization, a rank-1 attain of 94.1%, 94.5%, and 96.0% for the primary, next, and third schemes, respectively, was achieved.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

S.No	Authors	Year	Brief Description
1	Richa Singh, Mayank Vatsa, Afzel Noore[43]	2008	Presented a match gain combination of multispectral facial imagery and an integrated picture combination. -granular SVM, which applies several SVMs to be trained both the local and global features of the multispectral face data at different granularity levels and resolution, is used to combine VS and LWIR face imagery.
2	Ayan Seal, Debotosh Bhattacharjee, Mita Nasipuri[56]	2016	Presented a novel picture combination technique that utilizes the compensation of both types of data and is based on VS and thermal IR facial imaging. The proposed algorithm selected which VS and thermal IR face imagery to include in the generation of merged imagery using the translation-invariant à-trous wavelet transform and RF classifier.
3	Thirimachos Bourlai, Nikolaos Mavridis, Neeru Narang[61]	2016	Developed a creative internal solution using VS/NIR, variable distance, within/outside, and daylight/nighttime face data to solve intra/cross-distance NIR FR problems.
4	Ayan Seal, Debotosh Bhattacharjee, Mita Nasipuri, Consuelo Gonzalo-Martin, Ernestina Menasalvas[57]	2017	Based on the VS and thermal images for FR, a novel illustration combination algorithm is presented, yielding the benefits of both modalities' imagery and combination process: the weighted addition of VS and thermal face information using two weighting factors, α and β, respectively.
5	Marc Oliu Simón, Ciprian Corneanu, Kamal Nasrollahi, Olegs Nikisins, Sergio Escalera, Yunlian Sun, Haiqing Li, Zhenan Sun, Thomas B. Moeslund, Modris Greitans[13]	2016	Combined the most recent achievements in two methods by solving the multimodal RGB and thermal (RGB-D-T) based FR problem with DL CNN. Furthermore, a late synthesis of the CNN-based recognition block is started, and it establishes even better recognition presentation on a benchmark RGB-D-T dataset with a range of hand-crafted characteristics.
6	Madheswari Kanmani, Venkateswaran Narasimhan	2020	Suggested an innovative Eigen FR to improve the accuracy, utilizing thermal face imaging and VS synthesis. Three different synthesis approaches are chosen, in which the face data is combined using the most advantageous weights determined by a variety of optimization methods.

Table 7: Brief descriptions of FR based on Fusion

E. DEEP NEURAL NETWORKS (DNNs)

Artificial brain Networks (ANNs), which are based on human brain networks, have produced capable results beyond those demonstrated by previous methodologies. The application of ANN in FR is really straightforward. An artificial neural network (ANN) receives an image and extracts a set of features from it. When the ANN receives another image of the same person, it must generate a set of highly relevant features, and the opposite must happen when the input is an illustration of a human being that is not similar to the original. These days, DNNs are mostly utilized in the ANN industry due to their higher decision level than traditional ANNs.

The Graphic Processing Unit (GPU) has been highlighted as a weakness, and the training duration of the existing DNNs is highly dependent on its performance. Not just because ANNs can be extremely similar, but also because the score achieved, the calculating instance of the training and the categorization phases [35] [12] are periodically compared.

The DNN employed by Sarfraz and Stiefelhagen [48] limited the random association between two modalities, reducing the modality space between LWIR and VS images. A 10% improvement was obtained in the UND-X1 [58] dataset and a 15% to 30.00% improvement in the NVESD [46] dataset when the proposed strategy was assessed using PLS-based representations. In order to advance the distinctive learning using the two EMS layers—VS and NIR imagery—and to classify a vast amount of data, Jin et al. [23] established the concept of MTC-ELM. Using the CASIA HFB [32] and CASIA NIR-VIS 2.0 [7] datasets, the corresponding accuracy values were 95.2% and 89.1%.

Oh et al. [24] projected the use of SHLGBN to perform various FR using the CASIA NIR-VIS 2.0 [7] dataset, and an accuracy of 97.52% was attained. Guei and Akhlou [25] used the same dataset [7] with DCGAN to enlarge the image's dimension while preserving important facial details. The finished imagery's dimension was 6464, while the unique

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

imagery's dimension was 1616. MDNDC, which was created by Hu et al. [26], reduced intra-class dissimilarities and increased inter-class dissimilarities. It was possible to reduce modality space by scatter hammering, so the identifiable human features were retained. The MDN extracted features, and DC was used to adaptively control each DN's weight. Using MDNDC, the CASIA NIR-VIS 2.0 [7] and Oulu-CASIA NIR-VIS [8] datasets were examined, yielding verification rates of 88.1% and 99.6%, respectively.

As projected by Peng et al. [28] on the CASIA NIR-VIS 2.0 [7] dataset for various FR systems, the proposed method achieved an overall accuracy of 96.6% by incorporating a CNN for deep local descriptor removal and having the ability to learn discriminant along with dense confined information straightforwardly from facade imagery. It also had the ability to eliminate modality space on confined patch plane. Pereira et al. [10] applied the decreased characteristics from DCNN to DSU, which functioned as domain-specific lessened feature detectors. The same sets of dominant qualities from the source province were communal without requiring retraining, even when the reduced layers were altered. The presentation of DCNN was tested using three datasets: CASIA NIR-VIS 2.0 [7], NIVL [45], and PolaThermal [51]. The author got accuracy intensities of 96.3% and 90.1%, 94.5% and 92.2%, and 76.3% and 50.9% for Siamese and Triplet Neural Network, respectively.

He et al. [17] used a CNN known as Wasserstein CNN (WCNN) to reduce modality space involving VS and NIR pictures by implementing the Wasserstein Distance (WD). After being tested on WCNN, three databases—CASIA NIR-VIS 2.0 [7], Oulu-CASIA NIR-VIS, and BUAAVisNir [16]—achieved accuracy of 98.7%, 98.0%, and 97.4%, respectively. Hu and Hu [29] projected a novel varied FR practice called DDFLJM, and they used the CASIA NIR VIS 2.0 [7] and Oulu-CASIA NIR-VIS [8] datasets to do a relative comparison with WCNN [17]. For the initial and subsequent datasets, results were 99.3% and 98.8%, respectively. The author also discussed the CF using the CASIA NIR-VIS 2.0 [7] dataset, and it was found that SL produced better results (98.4%) than Softmax (84.4%).

Peng et al. [30] devised a very effective dimensional deep local illustration re-ranking approach for handling heterogeneous VS-NIR FR. The LLRe-Rank method was utilized to differentiate the standing outcomes. The CASIA NIR VIS 2.0 [7] and Oulu-CASIA NIR-VIS [8] datasets were used for the experiments, which produced gains of 98.71% and 98.91%, respectively. Wu et al. [35] created a deep CNN for multispectral FR dubbed IDICN, which examined intrascale diverse details and interscale associating information. Tests were run on the UWA [60] and HK PolyU [34] databases, yielding scores of 99.85% and 99.76%, respectively.

Bae et al. [31] developed two modules to advance diversified FR, with VS layer serving as the final example. A pre-processing chain was employed to ensure that the intensity variation between the translated and target illustrations was the same, and a CycleGAN was used to find the mapping between an input NIR image and an output VS image using a training set of allied image pair off. Using imagery from the training dataset and its transformed imagery, a ResNet-101 [82] trained with the Celeb-1M dataset [83] was employed in a later module to optimize the backbone model that had already been built. The database CASIA NIR-VIS 2.0 [7] was used throughout the experimentation phase. Without the pre-processing part, a rank-1 gain of 99.07% was obtained; with it, the outcome was improved by 99.40%.

Peng Lu et al. devised a method in 2021 [84] to address the issue of FR on minute unique dataset by combining CNN with augmented dataset. Through multiple transformations of face photography, the unique minute dataset is expanded into an enormous dataset. Based on the larger façade illustration dataset, the clever CNN effectively extracted the faces' characteristics and achieved higher FR correctness. Many tests were conducted to prove the effectiveness and superiority of the proposed scheme, and it was evaluated against other widely-used FR approaches.

FR has been used widely for person tracking and identification. However, because different facial images of the same sample seem differently due to differences in emotions, ages, individual locations, and lighting conditions, facial recognition (FR) has gotten more complex. It has been recognized that DL is a suitable option for both computer vision and FR. In a study conducted in 2022 by Thair A. Kadhim et al. [85], facial appearance and individuality were extracted from 14,126 images in a massive dataset known as FERET. Of these images, 80% were meant for training data and 20% were meant for testing data using CNN. Three unique deep learning representations were used: DenseNet-161, Resnet18, and AlexNet. The obtained results showed that DenseNet-161 has the highest accuracy, at 98.6%, while

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

Resnet18 and AlexNet had accuracy rates of 96.31% and 93.31%, respectively.

6. Conclusion and Future Scope

After a well-organized investigation, it was determined to be viable to conclude that the primary FR approaches and those that produced the best results were centered around ANN. In fact, according to the MDFR method, 36% of the majority of appropriate research manuscripts used ANN. It should also be noted that since 2019, image fusion or combinations methods have become more popular due to the use of ANN, primarily Generative Adversarial Networks (GAN), to perform this function.

Experience revealed that CASIA NIR-VIS 2.0 is the community dataset that is most frequently utilized [7]. The main drawbacks of the current MDFR systems were noted to be the accessibility of multidimensional databases, condensed numbers of imagery, the fact that there isn't a shared dataset with facial imagery of a similar human being at different spectral wavelength bands, and the lack of pose, brightness, and space differences amongst imagery in the similar dataset. Additionally, this work suggested that these constraints could cause an ANN to overfit during the training phase.

Compared to FR scheme that solely employed imagery from VS, MDFR approaches achieved better performance. Multidimensional pictures in FR can be used to conquer several feature spaces in this EMS. Several writers have noted in their investigations [39], [43], and [49] that the LWIR spectral band, which is unaffected by brightness differences, is capable of complementing the VS picture.

The limited amount of imagery (and population) in today's multidimensional datasets has limited the use of DNNs as an MDFR technique. However, since DNNs may produce incredibly powerful results, they are the primary approaches used for MDFR. When the CASIA NIR-VIS 2.0 [7] dataset was operational, the best results reached 99.41%. Safety and monitoring remain the primary goals of the MDFR plan, especially in hazardous areas like airfields or army-classified areas where identity verification is required. MDFR still has room to develop and improve.

It has been determined through a series of studies that the near future FR instruments will result in significant evolution from one end to the other. Large companies can promote in a variety of methods if they have the freedom to focus their marketing efforts on their potential customers. By the end of 2024, 1.28 billion devices should have FR solutions accessible. Companies like MasterCard and iProov are already using FR software in Smartphones, which is powered by AI, to validate credit transactions and carry out other sophisticated verification activities. .

S.No	Authors	Year	Brief Description
1	M. Saquib Sarfraz, Rainer Stiefelhagen[48]	2017	The primary challenging face matching problem in thermal-to-VS FR arises from an extraordinarily large modality space. provided a plan to capture the very non-linear relationship between the two modalities using DNN, thereby viaducting this modality space by a significant margin.
2	Y.Jin, Jie Li, Congyan Lang, Qiuqi Ruan[23]	2017	We present a novel multi-task clustering ELM designed for cross-modal feature learning. A linked cross-modal characteristic learning based face descriptor is predicted to reduce cross-modal dissimilarities, in contrast to conventional FR approaches.
3	B.Seok Oh, K.Oh, Andrew Beng Jin Teoh, Zhiping Lin, Kar-Ann Toh[24]	2017	For mixed FR, a single hidden-layer Gabor-based network was projected. The most recent computation units, which multiply geometrically localized input illustration sub-blocks to hidden nodes, are stored in the projected input layer.
4	Axel-Christian Guei, Moulay Akhloufi[25]	2018	Provided a deep learning framework based on the use of DCGAN for super-resolution infrared facade illustrations, which allowed the images to be upscaled by a factor of 4×4 after being first measured at a resolution of 16×16 to produce a 64×64 face picture.
5	Weipeng Hu, Haifeng Hu, Xinlong Lu[26]	2019	As a target purpose that may traverse the modality space while preserving the individuality information, SL is used to reduce the intra-class disparities while increasing the inter-class differences.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

6	Chunlei Peng, Nannan Wang, Jie Li, Xinbo Gao[28]	2019	For cross-modality FR, a deep local descriptor learning framework was projected with the goal of directly learning compressed and distinct local information from raw facial areas.
7	Tiago de Freitas Pereira, André Anjos, Sébastien Marcel[10]	2019	By utilizing Deep CNN's low-level characteristic in what are known as Domain Specific Units (DSU), a general framework for mixed FR is envisioned. The DSU adaption made it possible to learn shallow feature detectors unique to each new picture province.
8	Ran He, Xiang Wu, Zhenan Sun, Tieniu Tan[17]	2019	WCNN technique projected for learning invariant characteristic with VS and NIR imagery. The VS range of facial images is widely available for use in the low-level bands of WCNN, while the elevated layer is divided into three fractions: NIR, VS, and the NIR-VS shared layer.
9	Weipeng Hu, Haifeng Hu[29]	2019	A novel loss function called Scatter Loss (SL), which embeds both intra- and inter-class information for efficient deep model training, is projected in order to increase the discriminative ability of the deeply learnt attributes.
10	C. Peng, Nannan Wang, Jie Li, Xinbo Gao[30]	2019	With the goal of creating valuable insights from the preliminary standing consequences, a high-dimensional deep local representation for matching NIR and VS, or NIR-VIS FR, is envisioned. Using CNN, deep properties on local face areas are extracted and combined to first build a high-dimensional deep local representation.
11	F.Wu, Xiao-Yuan Jing, Xiwei Dong, Ruimin Hu[35]	2020	An IDICN technique was devised in order to enhance the multispectral FR presentation. Many scales are divided into multiple scale-sets, each containing a set of spectra in a small spectral series.
12	Han byeol Bae, Taejae Jeon[31]	2020	Suggested a two-step narrative structure made up of a characteristic learning unit and an illustration conversion unit to get a better cross-modality identification organism for a variety of datasets.
13	P.Lu, Lin Xu, Baoye Song[84]	2021	A novel approach linking CNN with larger dataset is developed to address the issue of human FR on minute unique dataset.
14	T. Kadhim, Nadia Smaoui Zghal, Walid Hariri, Dalenda Ben Aissa[85]	2022	Features were extracted from images of the massive FERET dataset, which consisted of 14,126 images that were divided into 80% for training and 20% for experimental data using CNN. CNN is first pre-trained using additional data and then trained with the target dataset to reveal more unidentified facial distinctiveness.

Table 8: Brief descriptions of FR based on DNN

REFERENCES

1. W. Zhang, X. Zhao, J.-M. Morvan, and L. Chen, ``Improving shadow suppression for illumination robust face recognition,'' IEEE Transactions on Pattern Analysis and Machine Intelligence., Volume. 41, no. 3, pp. 611624, Mar. 2019.
2. A. K. Jain, A. A. Ross, and K. Nandakumar, "Introduction to Biometrics", New York, NY, USA: Springer-Verlag, 2011, doi: 10.1007/978-0-387-77326-1.
3. A. D'Amico, C. Natale, F. Castro, S. Iarossi, A. Catini, and E. Martinelli, ``Volatile compounds detection by IR acousto-optic detectors,'' in Unex-ploded Ordnance Detection and Mitigation (NATO Science for Peace and Security Series B: Physics and Biophysics), J. Byrnes, Ed. Dordrecht, The Netherlands: Springer, 2009.
4. Krishna Rao Kakkirala, Srinivasa Rao Chalamala, "Thermal infrared face recognition: A review", in Proceeding of International Conference on Computer Modeling and Simulation, UKSIM, pp. 55-60, April 2017.
5. Rumaisah Munir, Rizwan Ahmed Khan, "An extensive review on spectral imaging in biometric systems: Challenges & advancements", Journal of Visual Communication and Image Representation, Volume 65, 102660, December 2019.
6. M. Kristo and M. Ivasic-Kos, "An overview of thermal face recognition methods", in Proceeding of 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), May 2018, pp. 1098-1103.
7. Stan Z. Li, Dong Yi, Zhen Lei, Shengcai Liao, "The CASIA NIR-VIS 2.0 face database", in Proceedings of

IEEE Conference on Computer Vision and Pattern Recognition Workshops, June 2013, pp. 348-353.

8. Guoying Zhao, Xiaohua Huang, Matti Taini, Stan Z. Li, Matti Pietikäinen, “Facial expression recognition from near-infrared videos," Image and Vision Computing, Volume 29, Issue 9, August 2011, Pages 607-619.
9. Shangfei Wang, Zhilei Liu, Siliang Lv, Yanpeng Lv, Guobing Wu, “A natural visible and infrared facial expression database for expression recognition and emotion inference”, IEEE Transactions on Multimedia, Volume 12, Issue 7, November 2010.
10. Tiago de Freitas Pereira, André Anjos, Sébastien Marcel, “Heterogeneous face recognition using domain specific units”, IEEE Transactions on Information Forensics and Security, Volume 14, Issue 7, July 2019, pp. 1803 – 1816.
11. Zhifeng Li, Dihong Gong, Qiang Li, Dacheng Tao, Xuelong Li, “Mutual component analysis for heterogeneous face recognition”, ACM Transactions on Intelligent Systems and Technology, Volume 7, No. 3, pp. 123, April 2016.
12. Weipeng Hu, Haifeng Hu, “Fine tuning dual streams deep network with multi-scale pyramid decision for heterogeneous face recognition”, Neural Processing Letters, Volume 50, No. 2, pp. 1465-1483, October 2019.
13. Marc Oliu Simón, Ciprian Corneanu, Kamal Nasrollahi, Olegs Nikisins, Sergio Escalera, Yunlian Sun, Haiqing Li, Zhenan Sun, Thomas B. Moeslund, Modris Greitans, “Improved RGB-DT based face recognition," IET Biometrics, Volume 5, No. 4, pp. 297-303, 2016, https://doi.org/10.1049/iet-bmt.2015.0057.
14. Andre Litvin, Kamal Nasrollahi, Sergio Escalera, Cagri Ozcinar, Thomas B. Moeslund, Gholamreza Anbarjafari, “A novel deep network architecture for reconstructing RGB facial images from thermal for face recognition”, Multimedia Tools and Applications., Volume 78, No. 18, pp. 25259-25271, Sep. 2019.
15. Yufeng Zheng, “Orientation-based face recognition using multispectral imagery and score fusion”, Optical Engineering, Volume 50, No. 11, November 2011.
16. D. Huang, J. Sun, and Y.Wang, “The BUAA-VisNir face database instructions”, Beihang University, Beijing, China, Tech. Rep. IRIP-TR-12-FR-001, 2012.
17. Ran He, Xiang Wu, Zhenan Sun, Tieniu Tan, “Wasserstein CNN: Learning invariant features for NIR-VIS face recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 41, Issue 7, 2019.
18. Ran He, Jie Cao, Lingxiao Song, Zhenan Sun, “Adversarial cross-spectral face completion for NIR-VIS face recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 42, No. 5, pp. 1025-1037, May 2020.
19. Virginia Espinosa-Duró, Marcos Faundez-Zanuy, Jiří Mekyska, “A new face database simultaneously acquired in visible, near-infrared and thermal scales”, Cognitive Computation, Volume 5, No. 1, pp. 119-135, 2013.
20. Cunjian Chen, Arun Ross, “Matching thermal to visible face images using hidden factor analysis in a cascaded subspace learning framework”, Pattern Recognition Letters, Volume 72, 1 March 2016, pp. 25-32
21. Yi Jin, Jiwen Lu, Qiuqi Ruan, “Coupled discriminative feature learning for heterogeneous face recognition”, IEEE Transactions on Information Forensics and Security, Volume: 10, Issue: 3, March 2015.
22. Chunlei Peng, Xinbo Gao, Nannan Wang, Jie Li, “Graphical representation for heterogeneous face recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 39, Issue 2, 01 February 2017.
23. Yi Jin, Jie Li, Congyan Lang, Qiuqi Ruan, “Multi-task clustering ELM for VIS-NIR cross-modal feature learning”, Multidimensional Systems and Signal Processing, Volume 28, No. 3, pp. 905-920, Jul. 2017.
24. Beom-Seok Oh, Kangrok Oh, Andrew Beng Jin Teoh, Zhiping Lin, Kar-Ann Toh, “A Gabor-based network for heterogeneous face recognition”, Neurocomputing, Volume 261, pp. 253-265, Oct. 2017.
25. Axel-Christian Guei, Moulay Akhloufi, “Deep learning enhancement of infrared face images using generative adversarial networks”, Applied Optics, Volume 57, Issue 18, pp. D98-D107 (2018), https://doi.org/10.1364/AO.57.000D98
26. Weipeng Hu, Haifeng Hu, Xinlong Lu, “Heterogeneous face recognition based on multiple deep networks with scatter loss and diversity combination”, IEEE Access, Volume 7 pp. 75305-75317, 2019.
27. Chunlei Peng, Xinbo Gao, Nannan Wang, Jie Li, “Sparse graphical representation based discriminant analysis for heterogeneous face recognition”, Signal Processing, Volume 156, March 2019, pp. 46-61.

28. Chunlei Peng, Nannan Wang, Jie Li, Xinbo Gao, “DLFace: Deep local descriptor for cross-modality face recognition”, Pattern Recognition, Volume 90, pp. 161-171, June 2019.
29. Weipeng Hu, Haifeng Hu, “Discriminant deep feature learning based on joint supervision loss and multi-layer feature fusion for heterogeneous face recognition”, Computer Vision and Image Understanding, Volume 184, pp. 9-21, July 2019.
30. Chunlei Peng, Nannan Wang, Jie Li, Xinbo Gao, “Re-ranking high-dimensional deep local representation for NIR-VIS face recognition”, IEEE Transactions on Image Processing, Volume 28, Issue 9, September 2019.
31. Han byeol Bae, Taejae Jeon, Yongju Lee, Sungjun Jang, “Non-visual to visual translation for cross-domain face recognition”, IEEE Access, Vol. 8, pp. 50452-50464, 2020.
32. Stan Z Li, Zhen Lei, Meng Ao, “The HFB face database for heterogeneous face biometrics research”, in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 18, Jun. 2009.
33. Zhen Lei Shengcai Liao, Anil K. Jain, Stan Z. Li, “Coupled discriminant analysis for heterogeneous face recognition”, IEEE Transactions on Information Forensics and Security, Volume 7, Issue 6, pp. 1707 – 1716, December 2012.
34. Wei Di, Lei Zhang, David Zhang, Quan Pan, “Studies on hyperspectral face recognition in visible scale with feature band selection”, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, Volume 40, Issue 6, pp. 1354 – 1361, November 2010.
35. Fei Wu, Xiao-Yuan Jing, Xiwei Dong, Ruimin Hu, “Intrascale discrimination and interscale correlation analysis deep network for multispectral face recognition”, IEEE Transactions on Cybernetics, Volume 50, Issue 3, pp. 1009 – 1022, March 2020.
36. Haitao Zhao, Shaoyuan Sun, “Sparse tensor embedding based multispectral face recognition”, Neurocomputing, Volume 133, pp. 427-436, 10 June 2014.
37. B. Abidi, S. Huq, M. Abidi, “Fusion of visual, thermal, and range as a solution to illumination and pose restrictions in face recognition”, in Proceeding of 38th Annual International Carnahan Conference on Security Technology, pp. 325-330, Oct. 2004.
38. Mrinal Kanti Bhowmik, Priya Saha, Anu Singha, Debotosh Bhattacharjee, Paramartha Dutta, “Enhancement of robustness of face recognition system through reduced Gaussianity in log-ICA”, Expert Systems with Applications, Volume 116, pp. 96-107, February 2019.
39. Madheswari Kanmani, Venkateswaran Narasimhan, “Optimal fusion aided face recognition from visible and thermal face images”, Multimedia Tools and Applications, Volume 79, pp. 17859-17883, Feb. 2020.
40. Dongoh Kang, Hu Han, Anil K. Jain, Seong-Whan Lee, “Nighttime face recognition at large standoff: Cross-distance and cross-spectral matching”, Pattern Recognition, Volume 47, Issue 12, pp. 3750-3766, 2014.
41. D. Shamia, D. Abraham Chandy, “Intelligent system for cross-spectral and cross-distance face matching”, Computers & Electrical Engineering, Volume 71, pp. 915-924, October 2018.
42. George Bebis, Aglika Gyaourova, Saurabh Singh, Ioannis Pavlidis, “Face recognition by fusing thermal infrared and visible imagery”, Image and Vision Computing, Volume 24, Issue 7, pp. 727-742, 1 July 2006.
43. Richa Singh, Mayank Vatsa, Afzel Noore, “Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition”, Pattern Recognition, Volume 41, Issue 3, pp. 880-893, March 2008.
44. Shaoyuan Sun, Haitao Zhao, Bo Jin, “Robust tensor preserving projection for multispectral face recognition”, Mathematical Problems in Engineering, Vol. 2014, Article No. 597245, August 2014.
45. John Bernhard, Jeremiah Barr, Kevin W. Bowyer, Patrick Flynn, “Near-IR to visible light face matching: Effectiveness of pre-processing options for commercial matchers”, IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2015.
46. K. A. Byrd, “Preview of the newly acquired NVESD-ARL multimodal face database”, Proceedings of SPIE, Volume 8734, Art. no. 018734, March 2013.
47. Shuowen Hu, Jonghyun Choi, Alex L. Chan, and William Robson Schwartz, “Thermal-to-visible face recognition using partial least squares”, Journal of the Optical Society of America ,Volume 32, Issue 3, pp.

431-442, 2015. https://doi.org/10.1364/JOSAA.32.000431.
48. M. Saquib Sarfraz, Rainer Stiefelhagen, "Deep perceptual mapping for crossmodal face recognition", International Journal of Computer Vision, Volume 122, No. 3, pp. 426-438, May 2017.
49. Nnamdi Osia, Thirimachos Bourlai, "Bridging the spectral gap using image synthesis: A study on matching visible to passive infrared face images", Machine Vision and Applications, Volume 28, No. 56, pp. 649-663, August 2017.
50. He Zhang, Benjamin S. Riggan, Shuowen Hu, Nathaniel J. Short, Vishal M.Patel, "Synthesis of high-quality visible faces from polarimetric thermal faces using generative adversarial networks", International Journal of. Computer Vision, Volume 127, No. 67, pp. 845-862, June 2019.
51. Shuowen Hu, Nathaniel J. Short, Benjamin S. Riggan, Christopher Gordon, Kristan P. Gurton, Matthew Thielke,Prudhvi Gurram, Alex L. Chan, "A polarimetric thermal database for face recognition research", In Proceeding of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 187-194, June 2016.
52. Robert B. Martin, Mikhail Sluch, Kristopher M. Kafka, Robert Ice, Brian E. Lemoff, "Active-SWIR signatures for long-range night/day human detection and identification", Proceeding of SPIE, Volume 8734, Art. No. 87340J, May 2013.
53. Cao Zhicheng, Schmid Natalia A., Bourlai Thirimachos, "Composite Multilobe descriptors for cross-spectral recognition of full and partial face", Optical Engineering, Volume 55, id. 083107, 2016.
54. Francesco Nicolo, Natalia Schmid, "Long range cross-spectral face recognition: Matching SWIR against visible light images", IEEE Transactions on Information Forensics and Security, Volume 7, No.6, pp. 1717-1726, December 2012.
55. Ayan Seal, Debotosh Bhattacharjee, Mita Nasipuri, Dipak Kr. Basu, "UGC-JU face database and its benchmarking using linear regression classifier", Multimedia Tools and Applications, Volume 74, No. 9, pp. 2913-2937, May 2015.
56. Ayan Seal, Debotosh Bhattacharjee, Mita Nasipuri, "Human face recognition using random forest based fusion of à-trous wavelet transform coefficients from thermal and visible images", AEU - International Journal of Electronics and Communications, Volume 70, Issue 8, Pages 1041-1049, August 2016.
57. Ayan Seal, Debotosh Bhattacharjee, Mita Nasipuri, Consuelo Gonzalo-Martin, Ernestina Menasalvas, "Fusion of visible and thermal images using a directed search method for face recognition", International Journal of Pattern Recognition and Artificial Intelligence, Volume 31, No. 4, Art. No. 1756005, Apr. 2017.
58. X. Chen, P. J. Flynn, and K. W. Bowye, "Visible-light and Infrared Face Recognition", in Proceeding of ACM Workshop Multimodal User Authentication, 2003, pp. 48-55.
59. Bing Cao, Nannan Wang, Jie Li, Xinbo Gao, "Data augmentation-based joint learning for heterogeneous face recognition", IEEE Transactions on Neural Networks and Learning Systems, Volume 30, Issue 6, June 2019.
60. Muhammad Uzair, Arif Mahmood, Ajmal Saeed Mian, "Hyperspectral face recognition using 3D-DCT and partial least squares", In Proceedings British Machine Vision Conference, pp. 110, 2013.
61. Thirimachos Bourlai, Nikolaos Mavridis, Neeru Narang, "On designing practical long range near infrared-based face recognition systems", Image and Vision Computing, Volume 52, pp. 25-41, August 2016.
62. Iacopo Masi, Yue Wu, Tal Hassner, Prem Natarajan, "Deep face recognition: A survey", in Proceeding of 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Oct. 2018, pp. 471-478.
63. Jian-Wu Lin, Hao Li, "HPILN: A feature learning framework for cross-modality person re-identification", Computer Vision and Pattern Recognition, Volume 13, No. 14, pp. 2897-2904, Dec. 2019.
64. Meiping Song, Xiaodi Shang, Chein-I Chang, "3-D receiver operating characteristic analysis for hyperspectral image classification", IEEE Transactions on Geoscience and Remote Sensing, Volume 58, Issue 11, pp. 8093-8115, November 2020.
65. N. Short, S.W. Hu, P. Gurram, K. Gurton, and A. Chan, "Improving crossmodal face recognition using polarimetric imaging," Opt. Lett., Volume 40, No. 6, pp. 882-885, 2015.
66. Brendan F. Klare, Anil K. Jain, "Heterogeneous face recognition using kernel prototype similarities", IEEE Transactions on Pattern Analysis and Machine

Intelligence, Volume 35, Issue 6, pp. 1410-1422, June 2013.
67. Dihong Gong, Zhifeng Li, Weilin Huang, Xuelong Li, Dacheng Tao, “Heterogeneous face recognition: A common encoding feature discriminant approach”, IEEE Trans. Image Process., Volume 26, No. 5, pp. 2079-2089, May 2017.
68. Shanmukhappa A. Angadi, Sanjeevakumar M. Hatture, “ Face recognition through symbolic modeling of face graphs and texture”, International Journal of Pattern Recognition and Artificial Intelligence, Volume 33, No. 12, Art. No. 1956008, November 2019.
69. Xiangsheng Huang, Zhen Lei, Mingyu Fan, Xiao Wang, Stan Z. Li, “Regularized discriminative spectral regression method for heterogeneous face matching”, IEEE Transactions on Image Processing, Volume 22, Issue 1, pp. 353-362, January 2013.
70. Kristan P Gurton, Alex J Yuffa, Gorden W Videen, “Enhanced facial recognition for thermal imagery using polarimetric imaging”, Opt. Lett., Volume 39, No. 13, pp. 3857-3859, 2014.
71. Z. Xie, S. Zhang, X. Yu, G. Liu, “Infrared and visible face fusion recognition based on extended sparse representation classification and local binary patterns for the single sample problem”, Journal of Optical Technology, Volume 86, Issue 7, pp. 408-413, 2019, https://doi.org/10.1364/JOT.86.000408.
72. Jie Gui, Zhenan Sun, Jun Cheng, Shuiwang Ji, Xindong Wu, “How to estimate the regularization parameter for spectral regression discriminant analysis and its kernel version?", IEEE Transactions on Circuits and Systems for Video Technology, Volume 24, Issue 2, , pp. 211-223, February 2014.
73. K. Siwek, S. Osowski, “Deep neural networks and classical approach to face recognition comparative analysis”, Przegl¡d Elektrotechniczny, Volume 94, No. 4, pp. 1-4, 2018.
74. J. Chmielinska, J. Jakubowski, “Face recognition based on deep learning techniques and image fusion”, Przegl¡d Elektrotechniczny, Volume 95, No. 11, pp. 150-154, 2019.
75. J. Pearl, “Markov and Bayesian networks: Two graphical representations of probabilistic knowledge”, in Probabilistic Reasoning in Intelligent Systems. San Mateo, CA, USA: Morgan Kaufmann, Chapter 3, 1998.
76. Stan Z. Li, Rufeng Chu, Shengcai Liao, Lun Zhang, “Illumination invariant face recognition using near-infrared images”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 29, No. 4, pp. 627-639, Apr. 2007.
77. Tran Minh Quan, David G. C. Hildebrand, Won-Ki Jeong, “FusionNet: A deep fully residual convolutional neural network for image segmentation in connectomics", Computer Vision and Pattern Recognition, arXiv: 1612.05360 [Online], Available: http://arxiv.org/abs/1612.05360, 2016.
78. Xiang Wu, Ran He, Zhenan Sun, Tieniu Tan, “A light CNN for deep face representation with noisy labels”, IEEE Transactions on Information Forensics and Security, Volume 13, No. 11, pp. 2884-2896, Nov. 2018.
79. B. Na, S. Js, “Detection of camouaged people”, International Journal of Sensor Networks and Data Communication, Volume 5, No. 3, pp. 143-148, 2016.
80. M. Liggins II, D. Hall, J. Llinas, “Handbook of Multisensor Data Fusion: Theory and Practice”, Boca Raton, FL, USA: CRC Press, 2017.
81. M. Omran, A. P. Engelbrecht, A. Salman, “Particle swarm optimization method for image clustering”, International Journal of Pattern Recognition and Artificial Intelligence, Volume 19, No. 03, pp. 297-321, May 2005.
82. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, “Deep residual learning for image recognition”, In Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, June 2016.
83. Yan Xu, Yu Cheng, Jian Zhao, Zhecan Wang, Lin Xiong, Karlekar Jayashree, Hajime Tamura, Tomoyuki Kagaya, Sugiri Pranata, Shengmei Shen, Jiashi Feng, Junliang Xing, “High performance large scale face recognition with multi-cognition softmax and feature retrieval”, In Proceeding of IEEE International Conference of Computer Vision and Workshops (ICCVW), pp. 1898-1906, Oct. 2017.
84. Peng Lu, Lin Xu, Baoye Song, “Human face recognition based on convolutional neural network and augmented dataset”, Systems Science & Control Engineering, Volume 9, Issue sup2, pp. 29-37 2021.
85. Thair A. Kadhim, Nadia Smaoui Zghal, Walid Hariri, Dalenda Ben Aissa, “Face Recognition in Multiple Variations Using Deep Learning and Convolutional Neural Networks”, in proceeding of IEEE 9th International Conference on Sciences of Electronics,

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 1024-1043 | ISSN:2251-6727

Technologies of Information and Telecommunications (SETIT), 2022.

86. Qureshi, Rizwan, Uzair, Muhammad, Zahra, Anam, "Current Advances in Hyperspectral Face Recognition". TechRxiv. Preprint, 2020.