Recognizing Injured Faces via SCIFI Loss

Puspita Majumdar, *Student Member, IEEE,* Saheb Chhabra, Richa Singh, *Fellow, IEEE,*
and Mayank Vatsa, *Senior Member, IEEE*

Abstract—Deaths and injuries are common in road accidents, violence, and natural disaster. In accidents and natural disasters scenarios, one of the tasks of responders is to retrieve the identity of the victims to reunite families or ensure proper identification of deceased persons. Apart from this, the identification of unidentified dead bodies due to violence and accidents is crucial for the police investigation. In the absence of identification cards, different forensic techniques such as DNA profiling and dental profiling may be used for identification. In this research, we present face recognition as a fast and viable approach for recognizing individuals with injuries. Face, which can be captured easily, is one of the most commonly used and widely accepted biometric modalities. However, face recognition is challenging in the presence of injuries as facial injuries change the appearance and geometric properties of the face due to swelling, bruises, blood clots, and accidental cuts. These changes introduce large intra-class variations among the same subject and small inter-class separability among different subjects. To address the challenge, we propose a novel Subclass Injured Face Identification (SCIFI) loss which is used in learning feature representation agnostic to injury variations. Additionally, an extended Injured Face (IF-V2) database of 150 subjects is presented which helps in evaluating the performance of face recognition models. Multiple experiments and comparisons are performed to showcase the efficacy of the proposed SCIFI loss based face recognition.

Index Terms—Injured Face Recognition, Biometrics, Face Identification, Injury.

✦

1 INTRODUCTION

Road accidents, violence, and natural disasters are common causes of deaths and injuries. A report of the World Health Organization (WHO) [1] states that every year, 1.25 million people are killed and 50 million are injured in road accidents worldwide. Several studies have reported that in such cases, face is one of the most affected regions of the human body [2], [3], [4]. A survey has shown that 50% to 70% of people surviving traffic accidents have a facial injury [5]. In the majority of such instances, victim identification becomes difficult as the facial regions are partially or completely damaged. The problem is exaggerated if the victim is in an unconscious state without any identity proof.

Apart from accidents, natural calamities and violence often result in unclaimed and unidentified dead bodies with facial injuries. According to the National Security Strategy [16], 80-90% of the recovered unidentified dead bodies never get identified. Retrieving the identity of the deceased person becomes difficult because the primary source of identity retrievals such as print or electronic media does not accept images with facial injuries as they may cause anxiety and discomfort to the viewers [17]. In such scenarios, techniques such as DNA profiling and dental profiling [18], [19], [20] are used by forensic experts and odontologists to recover the identity of injured victims and unidentified dead bodies. While these are popular primary modalities, they may be relatively slow to process or may not have the corresponding databases for identification purposes. For instance, not everyone has their dental or DNA records stored in some databases to be matched. Further, processing these modalities also require human expertise which may
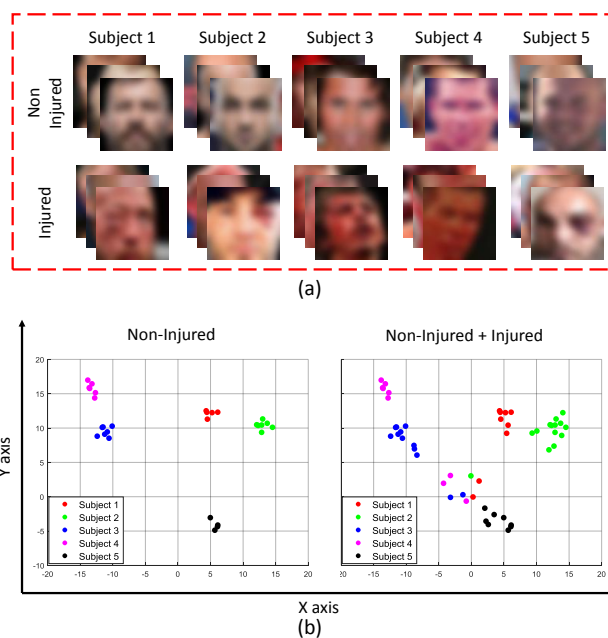

Fig. 1. (a) Sample images of five subjects from the Injured Face-Version2 (IF-V2) database to showcase the challenging nature of the problem statement. (b) t-SNE visualization of the corresponding subjects in the feature space obtained using pre-trained LCNN-29 [6] (Images are blurred for user privacy purposes).

not be available in all geographic locations. As an attractive supplement, we postulate that faces can be used to help the law enforcement agencies to reduce the search space and speed up the identification process. Face is one of the most commonly used and widely accepted biometric modality for recognition. A face can be captured easily from the injured victim or the dead body, which can be matched with the

• *P. Majumdar and S. Chhabra are with IIIT-Delhi, India, 110020 (email: pushpitam@iiitd.ac.in, sahebc@iiitd.ac.in).*

• *R. Singh and M. Vatsa are with IIT Jodhpur, India, 342037 (email: richa@iitj.ac.in, mvatsa@iitj.ac.in).*

TABLE 1
Literature review of the studies related to facial injuries and identification of dead bodies.

Authors	Year	Summary
Nassar et al. [7]	2008	Appearance-based features and string matching are used for the automatic construction of dental charts. It is observed that the appearance-based features takes less time for identification and outputs comparable accuracy compared to the methods in the literature.
Singh et al. [8]	2012	Analyzed 1,038 patients with maxillofacial injuries. It is concluded that road accidents are the main cause of maxillofacial injuries, and the age group of 21-30 years have more involvement in such cases.
Qiu et al. [9]	2017	Analyzed the Victoria Admitted Episodes Dataset that consists of facial trauma in Victoria between 2004 to 2014. It is found that falls from height are the major cause of facial trauma.
Black et al. [10]	2017	Studied different sports-related injuries and protective measures taken to reduce such injuries.
Sauerwein et al. [11]	2017	Evaluated human observers, including professionals and students, for the task of visual identification of a missing person given the antemortem facial photograph.
Majumdar et al. [12]	2018	Proposed an automatic framework to distinguish victims of domestic violence from others using facial images.
Canzi et al. [13]	2019	Provided a measuring scale for facial trauma by introducing the Comprehensive Facial Injury (CFI) score.
Trokielewiez et al. [14]	2019	Studied postmortem human iris recognition of 37 deceased subjects. It is shown that high performance is achieved when the images are taken shortly after death.
Bolme et al. [15]	2019	Evaluated face recognition models for decomposed human body recognition. Images are divided into six categories based on the decomposition level.

reference images in large (national level) databases.

Even though face recognition is a well research area, identifying victims using faces is still challenging because the injuries often affect the appearance and facial features. These changes induce large intra-class variations among the images of the same subject. Fig. 1 shows the t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization of the non-injured and non-injured + injured images of five subjects obtained using pre-trained LCNN-29 [6] model. It is observed that while non-injured images are well separated in the feature space, the overlap increases when both the injured and non-injured samples of a subject are given as input to the pre-trained model.

To address the variations due to injured faces, in this research, we have proposed a novel loss function, termed as **Subclass Injured Face Identification (SCIFI)** loss[1]. It considers injured and non-injured images as two different subclasses and operates in the **Subclass Space**, which is a 2-dimensional space of scores computed using non-injured and injured images. The proposed loss function optimizes the scores to achieve a two-fold objective in the feature space: (a) the distance between the samples of different subjects should be same while maintaining equidistance between the samples of same subject and (b) the distance between the samples of different subjects should be greater than the distance between the samples of the same subject. We have also extended the Injured Face (IF) database [21] to 150 (Injured Face-Version2 database) subjects and annotated the injured facial regions.

The proposed approach can be used as a stand alone identification framework followed by human-in-the-loop verification of the results. Further, it can be used as a supplement to a secondary modality such as DNA or dental records. Specifically, the proposed approach can be used to reduce the identification search space and then positive identification can be done using another modality.

2 RELATED WORK

Face is one of the widely accepted biometric modalities, and there has been a lot of progress in the field of face recog-

1. This paper is an extension of our work presented at BTAS2019 [21].

nition using deep models. However, recognizing faces with injuries during accidents or natural disasters still requires significant research. As shown in Table 1, a very few studies have been conducted related to this problem domain. To present an overview, we have divided the literature as follows: (i) injuries analysis and detection, (ii) face recognition using deep models, and (iii) recognition of injured faces.

Injury Analysis and Detection: Injuries and Violence Prevention Department of the World Health Organization has developed the International Classification of External Causes of Injury (ICECI) for studying injuries [22]. ICECI collects injury-related data such as the mechanism of injury, objects/substances producing injury, place of occurrence, activity when injured, the role of human intent, use of alcohol, and use of psychoactive drugs. This data is used by researchers to study injuries in different body parts.

In the literature, facial injuries have been studied in different fields, including medical science and forensic science for measuring the severity of facial trauma, identification of dead bodies, distinguishing injuries of domestic violence from accidents, and for studying the psychological effect of facial trauma. The common form of facial injuries include facial cuts, knocked out tooth, nosebleeds, eye injuries, and bleeding from the ears [23]. Qiu et al. [9] have analyzed 54,730 patients with facial trauma over ten years and found that interpersonal violence, transport accidents, and injuries from falls are the three common causes of facial trauma. According to Singh et al. [8], facial injuries are more frequent in males as compared to females. Researchers have also studied sports-related injuries and the requirement of medical treatment in such cases. Black et al. [10] have found that more than 41% of sports injuries require medical treatment. It is also found that helmets, eye protection, mouth guards, and face guard play a key role in reducing facial injuries. However, in most of the high-risk sports, these safety measures are not followed, which in turn results in facial injuries. Kampakis et al. [24] have studied football-related injuries and developed a framework using neural network and support vector machine to predict the recovery time of a football player automatically. A measuring scale of the facial trauma is proposed by Canzi et al. [13] that summarizes the surgical duration and care required for the

treatment. Facial injuries are also common in violence. A ten-year survey related to interpersonal violence (IPV) is provided by Coulibaly et al. [25]. It is found that facial injuries due to IPV occur mostly in young and male adults involved in brawls. An automatic framework to distinguish injuries of domestic violence from others is proposed by Majumdar et al. [12]. VGG-Face [26] model is fine-tuned to discriminate injuries of domestic violence. Authors have shown that features extracted from the injured regions are useful in classifying victims of domestic violence from others. However, they have not performed the recognition of injured faces to identify the victim.

Different techniques such as face, iris, fingerprint, DNA, and dental recognition have been proposed for identifying unidentified dead bodies. Caplova et al. [27] evaluated the performance of human observers for recognizing deceased persons with their corresponding facial photographs. It is found that the alteration of facial hair and age-separated images lead to lower performance. Researchers have also studied the effect of dead body decomposition on face recognition. Bolme et al. [15] collected a custom dataset of 42 subjects with different levels of decomposed face images. They evaluated the performance of commercial-off-the-shelf systems and existing pre-trained models for recognizing decomposed faces. It is observed that all the algorithms failed to recognize faces that were deformed or decomposed. Trokielewiez et al. [20] performed postmortem iris recognition and evaluated four independent iris recognition methods. They observed that iris recognition performs well up to 5-7 hours of death and are correctly recognizable with more than 90% accuracy. Sauerwein et al. [11] have shown that face and fingerprint are long-lasting biometric identifiers in postmortem cases while iris images lose their viability in a short span of time due to the loss of the moisture. Nassar et al. [7] have used appearance-based features and string matching for postmortem identification by constructing dental charts. Abaza et al. [28] proposed an algorithm on efficient retrieval of dental records to assist the forensic experts in rapid identification of deceased individuals. Balla [29] compared the practice of forensic dental identification in India with other Western countries. Kitayama et al. [30] evaluated the rapid DNA system for disaster victim identification (DVI). They observed that the success rate of fresh DVI samples ranges from 80% to 100%. Johnson and Riemen [31] discussed the role of digital fingerprint capture technology in accelerating the process of disaster victim identification. Levinson and Domb [32] discussed the role of police and medical examiner for the task of disaster victim identification using fingerprint matching.

Face Recognition using Deep Models: From the literature related to injuries, it is observed that existing methods for identifying the victims with facial injuries are time-consuming, non-scalable, and requires manual efforts. This demands the need for an automated system to speed up the process of identification. In recent years, a lot of progress has been made towards designing deep learning models for face recognition. Researchers have proposed several methods to enhance the discriminative power of the learned features for the task of face recognition. Hadsell et al. [33] have proposed Contrastive loss, where the loss function runs over pairs of samples. It minimizes the Euclidean distance between the pairs of similar samples while maximizes the distance between the pairs of dissimilar samples. The aim is to reduce the intra-class distance and increase the inter-class separability in the feature space. Wen et al. [34] have proposed center loss and combined it with softmax loss to enhance the supervision of the softmax loss to learn more discriminative features. Schroff et al. [35] proposed FaceNet, which uses triplet loss to unify the face embeddings in the Euclidean space. Later, Liu et al. [36] have shown that the Euclidean margin-based losses combined with softmax loss do not perform well to learn the discriminative features and proposed the A-softmax loss function to learn discriminative features angularly. This loss function modified the existing softmax loss to incorporate the angular information and introduced a margin parameter to control the angular margin. Wang et al. [37] have proposed a large margin cosine loss by modifying the softmax loss function as a cosine loss by L2 normalization of features and weight vectors. This helped in the removal of radial variations. Deng et al. [38] have proposed Additive Angular Margin Loss (ArcFace) to learn discriminative features for face recognition. This loss function computes the loss between the feature and the target weight and further adds an angular margin. It optimizes the geodesic distance margin on a hypersphere. Zheng et al. [39] have proposed a feature normalization technique, Ring loss, which constrains the norm by applying soft normalization and preserves the convexity. Zhang et al. [40] have shown the limitations of cosine based loss functions for face recognition, as they require tuning the scale parameter and angular margin. They have further proposed a loss function termed as AdaCos, which is hyperparameter free and uses an adaptive method for adjusting the scale parameter.

Recognition of Injured Faces: The literature on face recognition with injury variations is very limited. We have introduced the problem of injured face recognition in [21] where a database of 100 subjects (with and without injury) and Subclass Contrastive Loss (SCL) based deep learning model are proposed. The aim of SCL loss is to achieve an optimized feature space where the feature representations of the injured and non-injured subclasses of the same subject are close to each other while the feature representations of other subjects are far apart.

3 INJURED FACE RECOGNITION WITH PROPOSED SCIFI LOSS

We have observed that facial injuries increase the overlap among different subjects due to high intra-class variations and low inter-class separability in the feature space. Algorithms trained on non-injured faces generally do not perform well for injured face recognition. Therefore, we propose to model the injured and non-injured as two different subclasses. The concept of subclasses has been used in the literature for different classification tasks; for instance, Zhu and Martinez [41] have proposed subclass discriminant analysis (SDA). It approximates the underlying distribution of each class as a mixture of Gaussians and finds the optimal division of the classes into a set of subclasses. Mandal et al. [42] extended the concept by dividing each class into subclasses to capture intra-class variations and proposed a
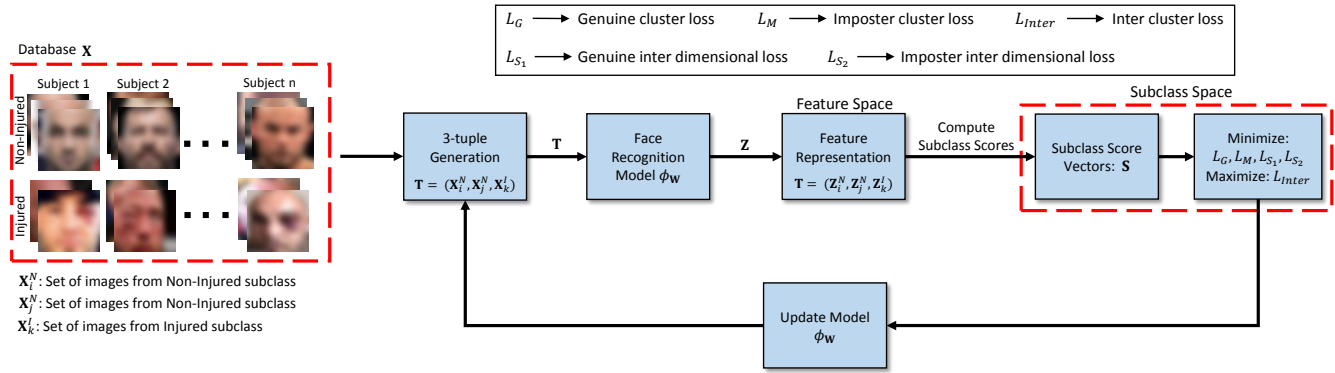
Fig. 2. Illustrating the steps involved in training a model for injured face recognition using the proposed SCIFI loss. In the first step, 3-tuples are generated by taking two samples from the non-injured subclass and one from the injured subclass. 3-tuples are then given as input to the model for obtaining the corresponding feature representations. In the next step, subclass score vectors are computed and projected to the 2-dimensional subclass space, where L_G, L_M, L_{S_1}, and L_{S_2} are minimized and L_{Inter} is maximized to train the recognition model (best viewed in color).

method of eigen-spectrum modeling. However, these methods estimated the number of subclasses in each class and require a large number of images per class.

In this research, we define non-injured and injured images as two subclasses, and propose a new space, termed as Subclass Space, to address the problem of injured face recognition. It is a 2-dimensional space of scores, where the scores are optimized to reduce the overlap between different subjects in the feature space. For this purpose, a novel *Subclass Injured Face Identification (SCIFI)* loss is proposed with two-fold objective in feature space: (a) to maintain equal inter-class distance between the samples of different subjects and equal intra-class distance between the samples of same subject and (b) the inter-class distance should be greater than the intra-class distance.

Fig. 2 summarizes the training process of the proposed approach for injured face recognition. Let d_1 and d_2 be the distances between {non-injured, non-injured} and {non-injured, injured} samples of the same subject, respectively. Similarly, d_3 and d_4 represent the distances between {non-injured, non-injured} and {non-injured, injured} samples of different subjects, respectively. The problem statement can be defined as "*given a dataset* $\mathbf{X}$ *of non-injured and injured subclasses of each subject, train a model* $\phi_{\mathbf{W}}$ *such that, in the feature space, the intra-sample distance* d_1, d_2 *is equal while maintaining same inter-sample distance* d_3, d_4, *where* $d_1, d_2 < d_3, d_4$.". The details of the loss function are discussed below.

3.1 SCIFI Loss

Let $\mathbf{X}$ be the dataset with n subjects, represented as $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n\}$. Each subject $\mathbf{X}_i$ consists of non-injured and injured subclasses, denoted by:

$$\mathbf{X}_i = \{\mathbf{X}_i^N, \mathbf{X}_i^I\} \tag{1}$$

where, $\mathbf{X}_i^N$ and $\mathbf{X}_i^I$ represent the set of non-injured images and the set of injured images, respectively. In order to train the network for injured face recognition, 3-tuples are generated by taking two samples from the non-injured subclass and one from the injured subclass.

$$\mathbf{T}_a = \{\mathbf{X}_{i,p}^N, \mathbf{X}_{j,q}^N, \mathbf{X}_{k,r}^I\} \tag{2}$$

where, $\mathbf{T}_a$ is the ath 3-tuple. $\mathbf{X}_{i,p}^N$ and $\mathbf{X}_{j,q}^N$ are the pth and q^{th} samples corresponding to $\mathbf{X}_i^N$ and $\mathbf{X}_j^N$ subclasses, respectively. Similarly, $\mathbf{X}_{k,r}^I$ is the rth sample of $\mathbf{X}_k^I$ subclass. Let $\phi_{\mathbf{W}}$ be a Convolutional Neural Network (CNN) with weights $\mathbf{W}$, which outputs the feature $\mathbf{Z}_i$ for input $\mathbf{X}_i$. The feature of the 3-tuple $\mathbf{T}_a$ is represented as:

$$\mathbf{T}_a = \{\mathbf{Z}_{i,p}^N, \mathbf{Z}_{j,q}^N, \mathbf{Z}_{k,r}^I\} \tag{3}$$

Let Y_a be the label of the 3-tuple $\mathbf{T}_a$, where $Y_a = 0$ denotes the genuine 3-tuple and $Y_a = 1$ denotes the imposter 3-tuple. Mathematically, it is represented as:

$$Y_a = \begin{cases} 0 & \text{if} \quad i = j = k \\ 1 & \text{if} \quad i \neq j, i \neq k \end{cases} \tag{4}$$

The next task is to compute the subclass scores and project it to the 2-D subclass space. Subclass score represents the distance between the feature representations of the subclasses. The first dimension of the subclass space represents the distance between the features of the samples from the non-injured subclass. The second dimension represents the distance between the features of the samples from non-injured and injured subclasses. The subclass score corresponding to the first and second dimensions of subclass space for the 3-tuple $\mathbf{T}_a$ is computed as:

$$S_{a,NN} = D(\mathbf{Z}_{i,p}^N, \mathbf{Z}_{j,q}^N) \tag{5}$$

$$S_{a,NI} = D(\mathbf{Z}_{i,p}^N, \mathbf{Z}_{k,r}^I) \tag{6}$$

where, $D(.)$ is a distance metric. It is important to note that the subclass score corresponding to the first dimension will be d_1 for genuine 3-tuple and d_3 for imposter 3-tuple. Similarly, the subclass score corresponding to the second dimension will be d_2 for genuine 3-tuple and d_4 for imposter 3-tuple. Mathematically, it is represented as:

$$S_{a,NN} = \begin{cases} d_1 & \text{if} \quad Y_a = 0 \\ d_3 & \text{if} \quad Y_a = 1 \end{cases} \tag{7}$$

$$S_{a,NI} = \begin{cases} d_2 & \text{if} \quad Y_a = 0 \\ d_4 & \text{if} \quad Y_a = 1 \end{cases} \tag{8}$$

The subclass score $\mathbf{S}_a$ of the 3-tuple $\mathbf{T}_a$ is represented as:

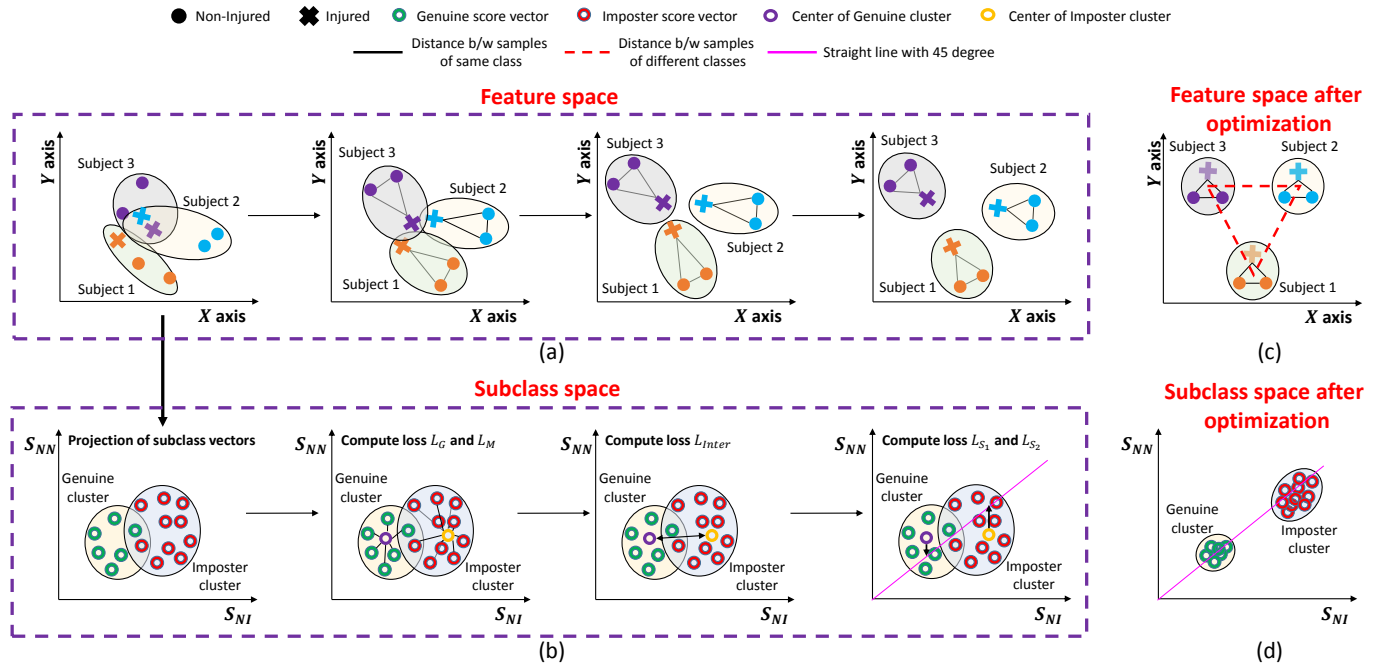$$\mathbf{S}_a = \begin{bmatrix} S_{a,NN} \\ S_{a,NI} \end{bmatrix} \tag{9}$$

Fig. 3. Illustration of the proposed SCIFI loss. (a and b) Represents the feature space and subclass space, where (b) demonstrates the optimization among genuine and imposter clusters for three subjects in subclass space and (a) shows the corresponding feature representations of the three subjects in feature space. (c and d) Represents the optimized feature space and subclass space (best viewed in color).

If m represents the number of 3-tuples generated for training the model, then there will be m subclass score vectors in 2-D subclass space. The subclass score vectors corresponding to genuine and imposter 3-tuples result in genuine and imposter clusters, respectively, in the subclass space. Let $\mathbf{C}_G$ and $\mathbf{C}_M$ represent the centers of the genuine and imposter clusters, respectively. The centers are computed as follows:

$$\mathbf{C}_G = \frac{\sum_{a=1}^{m}(1-Y_a)\mathbf{S}_a}{\sum_{a=1}^{m}(1-Y_a)} \qquad \mathbf{C}_M = \frac{\sum_{a=1}^{m}Y_a\mathbf{S}_a}{\sum_{a=1}^{m}Y_a} \qquad (10)$$

The aim is to maintain equal inter-class distance between the samples of different subjects and equal intra-class distance between the samples of same subject in the feature space, such that the inter-class distance is greater than the intra-class distance. For this purpose, the subclass score vectors corresponding to genuine and imposter clusters are pulled towards the center of their respective clusters. Mathematically, it is represented as:

$$L_G = \frac{\sum_{a=1}^{m}(1-Y_a)D(\mathbf{S}_a, \mathbf{C}_G)}{\sum_{a=1}^{m}(1-Y_a)} \qquad (11)$$

$$L_M = \frac{\sum_{a=1}^{m}Y_a D(\mathbf{S}_a, \mathbf{C}_M)}{\sum_{a=1}^{m}Y_a} \qquad (12)$$

In order to reduce the overlap between the subjects in the feature space, the inter-cluster distance between the centers of the genuine and imposter clusters is increased. Following represents the inter-cluster loss:

$$L_{Inter} = max(0, \alpha - D(\mathbf{C}_G, \mathbf{C}_M)) \qquad (13)$$

where, α is the margin with a value greater than or equal to zero that controls the separation among genuine and imposter clusters. While increasing the separability among

the clusters, it is possible that the separability is increased along a particular dimension and not the other. This condition could lead to overlap among different subjects in the feature space. In order to overcome this problem, the centers of the genuine and imposter clusters are separated along the straight line with a 45^0 angle with respect to the first dimension. Since x and y coordinates of any point on a straight line with 45^0 angle are equal, the centers are shifted towards the line using the following loss functions:

$$L_{S_1} = D(C_{G,NN}, C_{G,NI}) \qquad (14)$$

$$L_{S_2} = D(C_{M,NN}, C_{M,NI}) \qquad (15)$$

where, $C_{G,NN}$ and $C_{G,NI}$ represents the first and second dimensions of center $\mathbf{C}_G$, respectively. Similarly, $C_{M,NN}$ and $C_{M,NI}$ represents the first and second dimensions of center $\mathbf{C}_M$, respectively. The overall loss function to be minimized is:

$$L = L_G + L_M + L_{Inter} + L_{S_1} + L_{S_2} \qquad (16)$$

where, L is the final loss that increases the inter-class separability while maintaining approximately equal distance between the samples of different subjects as well as equal distance between the samples of the same subject in the feature space.

Fig. 3 illustrates the proposed SCIFI loss with an example. Consider a dataset with three subjects having injured and non-injured images. The feature space representations of samples of all the subjects are shown in the first plot of Fig. 3(a). The corresponding projection of subclass vectors with genuine and imposter clusters in subclass space is shown in the first plot of Fig. 3(b). Next, genuine cluster loss L_G and imposter cluster loss L_M are computed, as shown

in the second plot of Fig. 3(b). Here, L_G enforces the non-injured samples of the same subject to maintain distance d_1 among them. Simultaneously, it enforces distance d_2 among non-injured and injured samples of the same subject. Similarly, L_M enforces distance d_3 and d_4 for samples of different subjects. The visualization in feature space is shown in the second plot of Fig. 3(a). The third plot of Fig. 3(b) computes the inter-cluster loss L_{Inter} among the genuine and imposter clusters in the subclass space. This will either increase the difference between d_1 and d_3, d_2 and d_4 or both to increase the inter-class separability. The corresponding feature representations are shown in the third plot of Fig. 3(a). In order to achieve the optimal inter-cluster separation in both the dimensions, the centers of genuine and imposter clusters are shifted along the straight line with a 45^0 angle w.r.t the first dimension is shown in the fourth plot of Fig. 3(b). The fourth plot of Fig. 3(a) shows the corresponding representation in the feature space. The plot shows that the separability among all the subjects is further enhanced. Fig. 3 (c and d) show the optimized feature space and subclass space, respectively. It is observed that the proposed SCIFI loss optimizes the genuine and imposter clusters in the subclass space, which in turn optimizes the feature representations of all the subjects as well. It is important to note that the aim of the proposed SCIFI loss is to maximize the relative difference among the intra-class and inter-class separation in the feature space.

3.2 SCIFI Loss for Face Recognition

For recognition of injured faces, faces are first detected and aligned from the input images using Dlib library [43]. Next, features corresponding to the aligned faces are extracted using the trained model with SCIFI loss. Extracted features are matched with the enrolled gallery using Euclidean distance for identification. In this research, LightCNN-29 network is used as the base network for fine-tuning it with the proposed SCIFI loss.

4 INJURED FACE-VERSION2 DATABASE

Injured Face (IF) database proposed in our previous work [21], was the first ever database for injured face recognition. The database was created by collecting images from online resources. All the images are captured in the unconstrained environment with pose (-20^0 to $+20^0$) and illumination variations. This paper extends the database to 150 subjects (120 males and 30 females). To promote research on this problem, we have released the database[2]. Details regarding the creation of the Injured Face-Version2 (IF-V2) database and annotation of injured regions are discussed below.

4.1 Creation of Injured Face-Version2 Database

Injured Face-Version2 (IF-V2) database comprises injured and non-injured images of different subjects collected from the Internet. The problem of injured face recognition required matching injured face images of a subject with non-injured images enrolled in the database. On the other hand,

2. The database (in the form of weblinks) and the trained models are available at: http://iab-rubric.org/resources/ifv2.html

TABLE 2
Number of images with injuries in different facial regions.

Region	Images	Region	Images
Forehead	179	Right Cheek	214
Left Eye	224	Nose	179
Right Eye	182	Mouth	148
Left Cheek	181	Multiple	367

retrieving information or whereabouts of a missing person sometimes requires matching the non-injured face image of the missing person with unidentified victims or dead bodies with facial injuries. Therefore, for each subject, two types of images are collected: i) faces with real injuries and ii) normal or non-injured face images. Injured face images include faces injured during accidents, violence, and sports. The database contains a total of 1363 images corresponding to 150 subjects, out of which 524 are injured and 839 are non-injured face images.

4.2 Annotation of Injured Facial Regions

Injuries on different facial regions may affect the recognition performance differently. To study this impact, the injured region in the faces of the IF-V2 database are annotated. In cases where the image contains injuries on multiple facial regions, all the injured regions are annotated. Table 2 summarizes the number of images with injuries in different facial regions.

5 EXPERIMENTS AND RESULTS

The performance of the proposed SCIFI loss and existing face recognition models and loss functions are evaluated for problems such as identification of injured victims and retrieving information of missing persons. The details of the experiments are given below.

- The first experiment involves establishing the baseline performance of existing pre-trained models namely, VGG-Face [26], OpenFace [44], LightCNN-9 [6], and LightCNN-29 [6].
- The second experiment evaluates the performance of the proposed SCIFI loss, and compares it with existing loss functions. This experiment is performed under two scenarios: i) probe with injured images and ii) probe with non-injured + injured images. The first scenario represents the case of matching injured images with non-injured gallery for identifying the victims of road accidents or unidentified dead bodies with facial injuries. The second scenario represents the case of matching both non-injured and injured images with non-injured gallery. This scenario is applicable where the recognition system is deployed for both non-injured and injured face recognition.
- The last experiment evaluates SCIFI loss for the task of injured image retrieval, given normal images.

Protocol: Experiments are performed by dividing the IF-V2 database into non-overlapping (unseen) training and testing sets with 70%-30% subject-wise partitioning. Training and testing partitions are created by three times repeated random sub-sampling. For all the experiments, the training

TABLE 3
Mean identification accuracy (%) with standard deviation of existing pre-trained models on the IF-V2 database.

	Rank 1	Rank 5	Rank 10
VGG-Face	4.28 ± 1.9	15.55 ± 2.7	29.34 ± 3.2
OpenFace	4.75 ± 1.5	19.55 ± 1.3	32.59 ± 2.5
LCNN-9	46.65 ± 6.4	68.18 ± 3.9	83.20 ± 0.5
LCNN-29	$\mathbf{63.93 \pm 3.6}$	$\mathbf{85.48 \pm 2.0}$	$\mathbf{90.47 \pm 1.5}$

and testing sets are divided into gallery and probe. During training, the gallery contains non-injured/normal images, and the probe contains injured images of each subject. During testing, multiple gallery and probe partitions are created for each experiment. In the first experiment, the gallery contains a single non-injured image per subject, and the probe contains multiple injured images per subject. The second experiment is performed with a gallery of single non-injured image per subject, and the probe of multiple injured images per subject and multiple non-injured + injured images per subject. In the third experiment, the gallery contains a single injured image per subject while the probe contains multiple non-injured images per subject.

Implementation Details: All the experiments are performed on aligned faces. Experiments are performed on NVIDIA Tesla P100 server with 96GB RAM and 16GB GPU memory, in PyTorch environment. Adam optimizer is used with a learning rate of 6×10^{-5}. For SCIFI loss function, the LightCNN-29 network is fine-tuned for 10 epochs with a batch size of 50, and α is set to 2.6. While for the existing algorithms, a batch size of 50 is used, and the LightCNN-29 network is fine-tuned for 30 epochs with Adam optimizer. To fine-tune the network with Subclass Contrastive Loss (SCL) [21], the two margins are set to 1.5 and 2.6, respectively, and a learning rate of 3×10^{-6} is used. A margin of 2 is used to fine-tune the siamese network with Contrastive Loss (CL) with a learning rate of 3×10^{-6}. During fine-tuning with Triplet Loss (TL), a learning rate of 8×10^{-6} is used, and the margin is set to 0.4.

5.1 Injured Face Recognition using Pre-trained Models

In this experiment, the performance of existing pre-trained models, namely, VGG-Face, OpenFace, LightCNN-9, and LightCNN-29, are evaluated for the task of injured face recognition. For this purpose, features are extracted from the gallery and probe of the testing set using the pre-trained models. Extracted features of the probe images are matched with the gallery images using Euclidean distance. Three times random subsampling based cross-validation is performed, and the mean identification accuracy with standard deviation at rank 1, rank 5, and rank 10 are shown in Table 3. Fig. 4 shows the Cumulative Match Characteristic (CMC) curves. The results indicate that existing pre-trained models do not perform well for the task of injured face recognition. Injuries change the facial features and deteriorate the performance of existing models by incorporating large intra-class variations among the images of the same subject and reduce the inter-class separability. From Table 3, it is observed that LightCNN-29 yields the best performance with an average identification accuracy of 63.93% at rank 1.
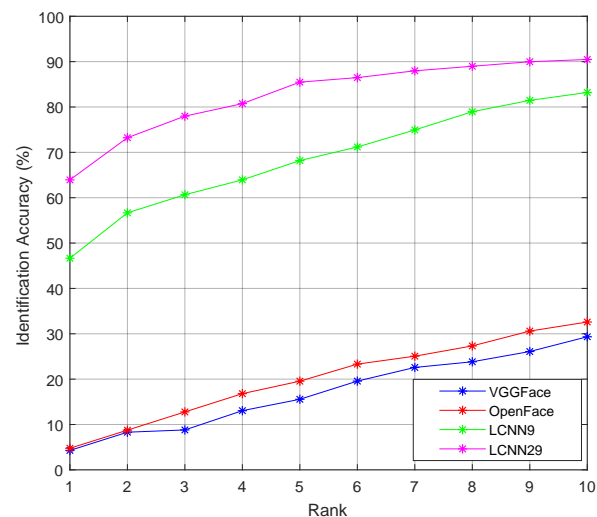

Fig. 4. CMC curves of the four existing pre-trained face recognition models (best viewed in color).
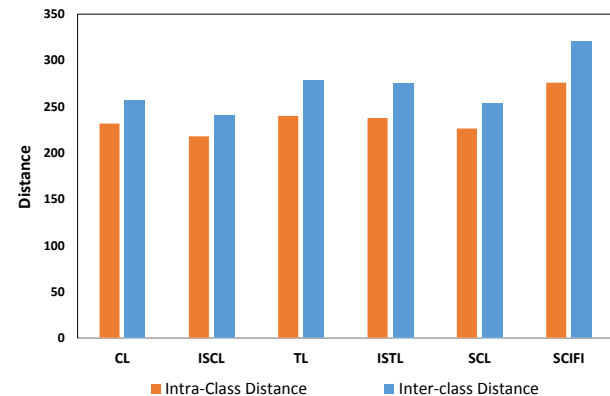

Fig. 5. Comparison of mean intra-class and inter-class distance of gallery image of each subject with probe images of same and other subjects using existing (CL, ISCL, TL, ISTL, and SCL) and proposed algorithms in the testing set (best viewed in color).

5.2 Injured Face Recognition using SCIFI Loss

In order to evaluate the performance of the proposed SCIFI loss, experiments are performed on the IF-V2 database and the results are compared with existing loss functions including Contrastive Loss (CL) [33], Triplet Loss (TL) [35], Center Loss [34], CoCo Loss [45], Ring Loss [39], SphereFace [36], CosFace [37], ArcFace [38], AdaCos [40], Subclass Contrastive Loss (SCL) [21], and variants of Contrastive Loss and Triplet Loss, termed as, Injury Specific Contrastive Loss (ISCL) and Injury Specific Triplet Loss (ISTL), respectively, by fine-tuning LCNN-29 model. To incorporate the injury specific information in CL and TL, we have modified the sampling technique for pairs and triplets generation. In ISCL, pairs are generated using one sample from the injured subclass and other from the non-injured subclass. Similarly, for ISTL, triplets are generated by taking anchor from the non-injured subclass while positive and negative from the injured subclass. However, for CL and TL, subclasses are not considered during pairs or triplet generation. LightCNN-29

TABLE 4
Identification accuracy (%) on the IF-V2 database using the proposed and existing loss functions. The left block shows the results for the probe with injured images, while the right block shows the results for the probe with injured + non-injured images.

Loss Function	Probe: Injured			Probe: Injured + Non-injured		
	Rank 1	Rank 5	Rank 10	Rank 1	Rank 5	Rank 10
Contrastive Loss [33]	64.68 ± 3.6	85.23 ± 2.9	91.46 ± 1.0	82.23 ± 2.6	92.58 ± 2.3	95.72 ± 0.4
Injury Specific CL	66.70 ± 4.4	84.98 ± 4.1	91.73 ± 0.7	82.68 ± 2.8	92.25 ± 2.5	95.72 ± 0.7
Triplet Loss [35]	67.68 ± 2.2	86.99 ± 3.7	94.01 ± 2.6	83.01 ± 2.1	93.03 ± 2.1	96.74 ± 1.7
Injury Specific TL	68.71 ± 4.4	87.48 ± 2.2	94.76 ± 2.9	83.57 ± 2.4	92.69 ± 1.8	97.08 ± 1.5
Center Loss [34]	63.97 ± 6.0	82.74 ± 4.1	88.73 ± 2.1	81.34 ± 4.0	91.13 ± 2.9	94.48 ± 1.5
CoCo Loss [45]	65.94 ± 2.7	85.23 ± 2.3	91.22 ± 0.7	82.67 ± 2.4	92.58 ± 1.9	95.71 ± 0.6
Ring Loss [39]	64.93 ± 3.5	84.22 ± 1.9	90.23 ± 1.1	82.11 ± 2.7	92.35 ± 1.6	95.04 ± 0.7
SphereFace [36]	67.45 ± 5.4	84.47 ± 2.4	90.22 ± 0.5	83.47 ± 3.5	92.24 ± 1.9	95.15 ± 0.4
CosFace [37]	66.96 ± 5.2	85.98 ± 1.6	91.74 ± 1.5	83.25 ± 3.7	93.02 ± 1.6	95.72 ± 1.3
ArcFace [38]	68.92 ± 4.1	84.47 ± 3.3	91.72 ± 1.0	84.48 ± 2.7	92.58 ± 1.7	95.83 ± 0.6
AdaCos [40]	69.21 ± 3.9	87.48 ± 2.4	91.73 ± 0.9	83.81 ± 2.7	93.70 ± 1.5	95.82 ± 0.5
SCL [21] (BTAS2019 version)	69.97 ± 5.0	$\mathbf{88.98 \pm 1.7}$	92.99 ± 1.0	84.26 ± 3.3	$\mathbf{94.16 \pm 1.9}$	96.28 ± 0.6
Proposed SCIFI Loss	$\mathbf{72.22 \pm 4.8}$	88.75 ± 3.1	$\mathbf{96.27 \pm 3.9}$	$\mathbf{84.60 \pm 3.6}$	93.82 ± 2.2	$\mathbf{97.65 \pm 2.0}$

TABLE 5
Comparing the value of metric R.

	CL	ISCL	TL	ISTL	SCL	SCIFI
Ratio	0.0992	0.0963	0.1377	0.1363	0.1081	**0.1389**

performs the best among the baseline models. Therefore, experiments are performed using LightCNN-29 as the base network. The base network is fine-tuned using the proposed and existing loss functions for the task of injured face recognition. The features of 256-dimension from the last fully connected (fc) layer of the LCNN-29 model are used for performing experiments.

Table 4 shows the average identification accuracy with standard deviation at ranks 1, 5, and 10 with injured probe and injured + non-injured probe. It is observed that the proposed SCIFI loss outperforms existing algorithms for both cases. For instance, the classification accuracy at rank-1 is 72.22% and 69.46%, corresponding to SCIFI loss and SCL, respectively, when the probe are injured images. Similarly, when the probe are injured + non-injured images, the proposed SCIFI loss outperforms at rank-1 and rank-10 with a comparable performance at rank-5. This showcases that the proposed SCIFI loss works well for both the cases. It is also observed that incorporating the injury information (ISCL and ISTL) improves the performance of CL and TL. It is important to note that the proposed SCIFI loss outperforms recent state-of-the-art face recognition algorithms [36], [37], [38], [40]. To further validate the improvement in the accuracies, McNemar's test is performed to statistically compare ArcFace and AdaCos with the proposed SCIFI loss. It is observed that the results of SCIFI loss are significantly different at 5% significance level. The p-value corresponding to the comparison of ArcFace and SCIFI loss is 0.03, while the p-value for comparison of AdaCos and SCIFI loss is 0.04.

In order to understand the variations in intra-class and inter-class distances with different loss functions, we have introduced a metric R that computes the ratio of the difference between mean inter-class and intra-class distance to the mean inter-class distance.
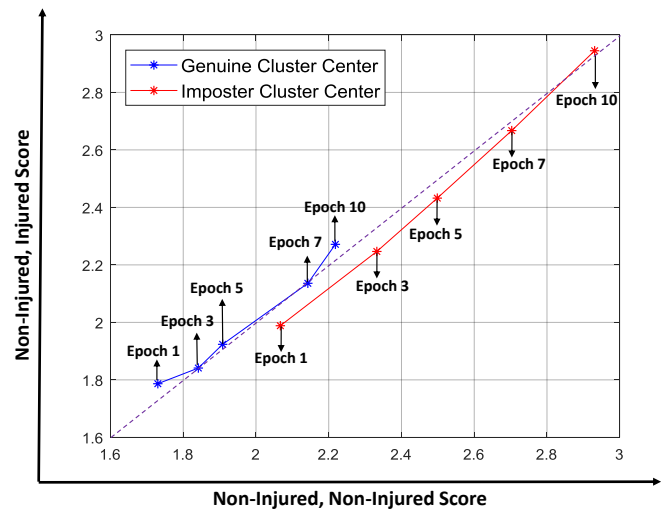
$$R = \frac{Inter - Intra}{Inter} \tag{17}$$

Fig. 6. Illustration of the shift of genuine and imposter cluster centers with different epochs during training (best viewed in color).

Here, it is assumed that $Inter > Intra$ and the range of R is 0 to 1. The value of metric R becomes 1 when the $Intra$ tends to zero or $Inter >> Intra$. Therefore, the higher the value of R, the better the performance of the algorithm. Fig. 5 compares the mean intra-class and inter-class distances of all the samples in the testing set. From Fig. 5, it is observed that the proposed SCIFI loss has higher intra-class as well as inter-class distance. However, we cannot directly compare the performance by observing intra-class and inter-class distance individually. Therefore, the metric R is computed corresponding to the proposed loss and existing loss functions. The intra-class and inter-class distances are computed using the 256-dimensional feature vector from the last fc layer of the trained LCNN-29 model and the results are summarized in Table 5. It is observed that the proposed SCIFI loss has the highest R value. Next, Fig. 6 shows the shift of the genuine cluster center and imposter cluster center with different epochs during training. It is observed that both the cluster centers move along the straight line, which represents the increase in intra-class and inter-class distances. However, the relative difference among the genuine and imposter cluster centers
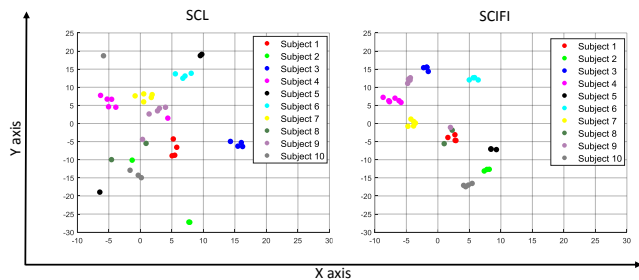
Fig. 7. t-SNE visualization of the SCL and SCIFI loss of 10 randomly selected subjects from the testing set (best viewed in color).
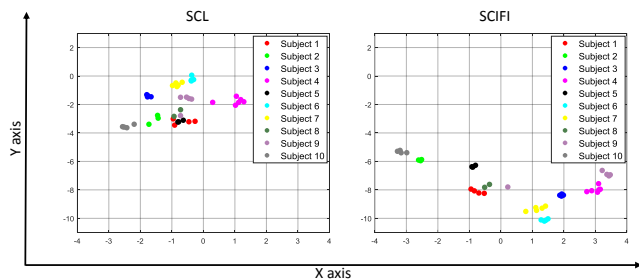

Fig. 8. UMAP visualization of the SCL and SCIFI loss of 10 randomly selected subjects from the testing set (best viewed in color).

increases with epochs, which in turn reduces the overlap among different subjects in the feature space.

The t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization of the feature representation of 10 randomly selected subjects from the testing set using SCL and SCIFI loss is shown in Fig. 7. It is observed that the subjects are better separated in feature space using the proposed SCIFI loss compared to SCL. For better interpretation, we have also shown Uniform Manifold Approximation and Projection (UMAP) visualization of 10 randomly selected subjects from the testing set in Fig. 8. t-SNE preserves local structure in the data while UMAP preserves both local and global structure in the data. Therefore, UMAP helps to interpret both within-cluster and between-cluster distances. Similar to the t-SNE visualization, UMAP also shows that the subjects are better separated in the feature space using the proposed SCIFI.

We have also evaluated the performance in terms of the verification accuracy. For this purpose, 100 genuine and 100 imposter pairs are randomly generated from the database and Fig. 9 shows the Genuine Accept Rate (GAR) at 0.1 False Accept Rate (FAR). It is observed that SCIFI loss outperforms other loss functions. For instance, the GAR corresponding to SCIFI loss and SCL is 0.92 and 0.83, respectively. Fig. 10 shows the genuine and imposter score distributions using SCL and the proposed SCIFI loss, respectively. It is observed that the proposed SCIFI loss reduces the overlap between the two distributions, which in turn results in improved performance. These results indicate that the proposed SCIFI loss performs well for both identification and verification of injured faces.
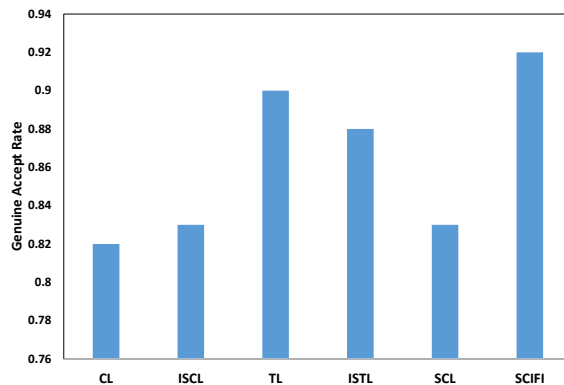

Fig. 9. GAR at 0.1 FAR on the IF-V2 database (best viewed in color).
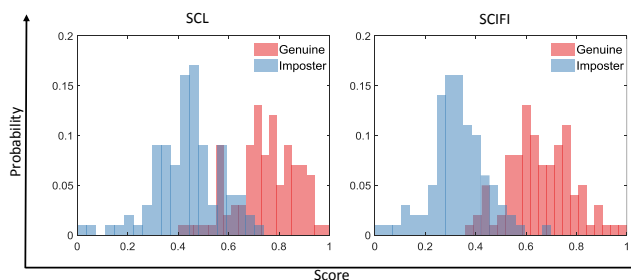

Fig. 10. Genuine and imposter score distributions of the SCIFI loss and SCL (best viewed in color).

TABLE 6
Identification accuracy (%) on the IF-V2 database by ablating different loss terms in the proposed SCIFI loss.

Loss Terms	Rank 1	Rank 5	Rank 10
L_G and L_{S_1}	67.42 ± 1.4	87.21 ± 1.1	92.48 ± 0.5
L_M and L_{S_2}	69.20 ± 3.0	89.21 ± 0.3	94.22 ± 1.5
L_{Inter}	58.45 ± 4.8	79.73 ± 3.6	87.47 ± 0.7
L_{S_1} and L_{S_2}	69.46 ± 4.0	87.72 ± 2.0	94.24 ± 1.8

5.3 Injured Image Retrieval

Injured image retrieval has several real-world applications. For instance, retrieving the information of a missing person sometimes requires matching the non-injured face images of the missing person with some unidentified victims or dead bodies with facial injuries. Therefore, the proposed SCIFI loss is also evaluated for the task of injured image retrieval. For evaluation, the model is trained on the training set of the IF-V2 database. For evaluation, the testing set of the IF-V2 database is divided into gallery with injured images and probe with non-injured images. In this experiment, non-injured images are matched with injured images, and it is observed that the proposed SCIFI loss yields the mean identification accuracy of **73.23%** at rank-1, which is **2.88%** higher than SCL.

5.4 Ablation Study and Choice of Hyper-parameters

To investigate the effect of individual terms in the proposed SCIFI loss function, we have performed four different experiments by ablating individual loss terms. The identification accuracy at ranks 1, 5, and 10 are shown in Table 6. The results indicate that each term is important and ablating

TABLE 7
Analyzing the effect of varying the margin α at learning rate $= 6 \times 10^{-5}$ and batch size = 50.

Alpha	Rank 1	Rank 5	Rank 10
1.6	69.45 ± 3.0	87.47 ± 2.0	94.72 ± 1.6
2.1	69.70 ± 3.2	88.97 ± 0.2	94.98 ± 0.6
2.6	72.22 ± 4.8	88.75 ± 3.1	96.27 ± 3.9
3.1	70.93 ± 2.2	88.97 ± 0.5	93.97 ± 1.2

TABLE 8
Analyzing the effect of varying the batch size at $\alpha = 2.6$ and learning rate $= 6 \times 10^{-5}$.

Batch Size	Rank 1	Rank 5	Rank 10
50	72.22 ± 4.8	88.75 ± 3.1	96.27 ± 3.9
75	71.44 ± 1.4	87.72 ± 1.1	95.23 ± 1.2
100	70.44 ± 2.1	89.45 ± 1.8	93.47 ± 0.9
125	70.19 ± 1.9	87.96 ± 1.0	95.23 ± 1.2

TABLE 9
Analyzing the effect of varying the learning rate (LR) at $\alpha = 2.6$ and batch size = 50.

| LR	Rank 1	Rank 5	Rank 10
3×10^{-5}	66.93 ± 2.3	87.46 ± 1.2	94.72 ± 0.6
6×10^{-5}	72.22 ± 4.8	88.75 ± 3.1	96.27 ± 3.9
9×10^{-5}	71.69 ± 1.9	89.48 ± 1.5	96.24 ± 1.1
1×10^{-4}	70.94 ± 2.8	87.22 ± 0.9	94.97 ± 1.9

Fig. 11. Rank 1 identification accuracy (%) with existing and proposed loss functions and ResNet-18 model. Results are shown for the probe as (a) injured images and (b) injured + non-injured images.

any of the loss terms degrades the overall performance. Among different loss terms, ablating L_{S_1} and L_{S_2} has the least effect on the performance and the accuracy reduces by 2.76%. However, L_{Inter}, which controls the separation among the genuine and imposter cluster centers, has the highest impact on the SCIFI loss. It is observed that on ablating L_{Inter}, the rank 1 accuracy of the SCIFI loss reduces by 13.77%. Removing this term increases the overlap among genuine and imposter clusters, which in turn degrades the performance of the model. From Equation 13, it is clear that L_{Inter} becomes zero when the distance between the genuine and imposter clusters i.e., $D(\mathbf{C}_G, \mathbf{C}_M) \geq \alpha$. In this scenario, the L_{Inter} loss does not have a role in the minimization of the SCIFI loss function and the L_{Inter} loss can be disregarded. However, in practical real world problems like injured face recognition, this smaller distance between these clusters is one of the most crucial factors, and should not be disregarded.

To compare the choice of hyper-parameters for model training, experiments are performed with different values of α, batch size, and learning rate. In the literature, different methods such as grid search, Bayesian approach, and gradient based approach have been proposed to estimate or optimize hyper-parameters [46], [47], [48], [49], [50], [51], [52]. In this research, hyper-parameters are chosen using grid search approach. Three times repeated random sub-sampling is performed for training and testing partitioning and hyper-parameter tuning. For each hyper-parameter, a set of values is used for performing grid search. For each hyper-parameter, the value that yields the best results across all the partitions (by averaging) is selected. Tables 7, 8, 9 show the impact of varying α, batch size, and learning rate, respectively, on the final result. From Table 7, it is observed that on varying the margin α, which is a parameter of L_{Inter}
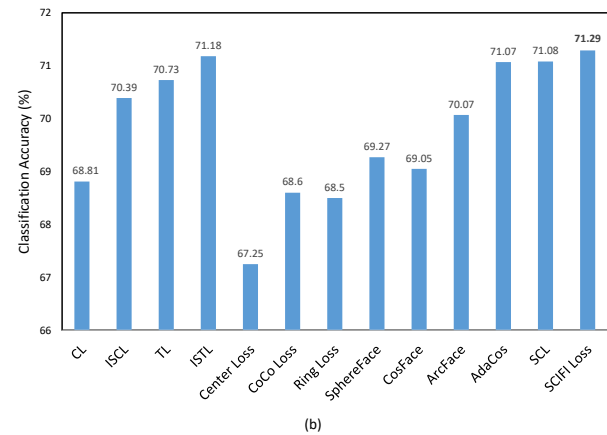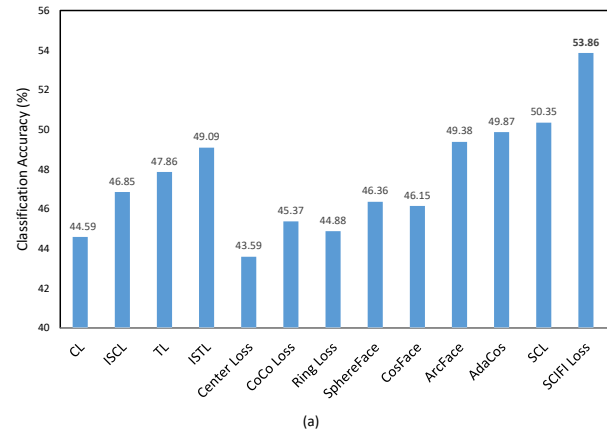
loss term, the rank-1 accuracy varies from 69.45% to 72.22%. Table 9 shows that a small learning rate results in lower rank-1 accuracy. These results indicate that carefully crafted hyper-parameters play an important role in improving the generalizability on unseen data. Further, among the three parameters, learning rate is a very important parameter to optimize followed by the margin α.

Apart from grid search based approach, we have performed experiment using the Tree-structured Parzen estimators (TPE) approach which is a sequential model-based optimization technique. In this experiment, we have obtained the learning rate as 4.15×10^{-5} and alpha as 2.58. The mean classification accuracy with standard deviation obtained at rank 1, 5, and 10 is 71.03 ± 4.7, 88.99 ± 2.9, and 96.41 ± 0.7, respectively. It is observed that the accuracy obtained using hyperparameters estimated via TPE approach is close to the grid search approach.

5.5 Results of ResNet18 Experiment

In order to evaluate the generalizability of the SCIFI loss with other deep learning networks, we have also performed experiments using ResNet-18 network [38] as the base model. Experiments are performed using the protocol mentioned in Section 5 under the following two scenarios: probe with injured images and probe with injured + non-injured images. Fig. 11(a) and 11(b) show the results of SCIFI

and existing loss functions for probe with injured images and injured + non-injured images, respectively. Compared to existing loss functions, the proposed SCIFI loss yields 3.51% to 10.27% higher rank-1 accuracy for the probes with injured images. Similarly, 0.11% to 4.04% improvements are observed for the probes with injured + non-injured images. This experiment showcases that the proposed SCIFI loss can be used in conjunction with other deep learning models and compared to existing loss functions, it has more expressive power to improve the performance for injured face recognition. The performance of LightCNN-29 and ResNet18 is summarized in Table 4 and Fig. 11, and it is observed that LightCNN-29 yields better performance compared to ResNet18.

6 CONCLUSION

Facial injuries are common in road accidents, violence, and natural disasters. Several real-world cases require timely identification of unconscious victims or unidentified dead bodies with facial injuries. However, injuries change the appearance, geometric properties, and physical structure of face due to swelling, blood clots, bruises, and accidental cuts. In such scenarios, identification becomes difficult as these variations increase the overlap among different subjects in the feature space. In this research, a novel Subclass Injured face Identification (SCIFI) loss is proposed to address the problem of injured face recognition. The proposed loss function performs optimization in the subclass space, which in turn optimizes the feature space. Additionally, the Injured Face database is extended to 150 subjects. Multiple experiments are performed on the IF-V2 database to showcase the efficacy of the proposed SCIFI loss along with analysis for the improved performance. In future, we plan to extend this work by detecting injuries in different facial regions and analyzing the impact of different types of injuries on the face recognition system.

ACKNOWLEDGMENTS

REFERENCES

[1] "WHO statistics," http://tinyurl.com/y6dz83t4.
[2] H. Granhed, E. Altgarde, L. Akyurek, and P. David, "Injuries sustained by falls-a review," *Trauma Acute Care*, vol. 2, p. 38, 2017.
[3] "Toy related deaths and injuries," http://tiny.cc/ykhhjz, 2016.
[4] "Impact injury," https://www.britannica.com/science/impact-injury, 2019.
[5] J. Jordan and K. Calhoun, "Management of soft tissue trauma and auricular trauma," *Bailey BJ, Johnson JT, Newlands SD. Head & Neck Surgery: Otolaryngology. Hagerstwon, MD: Lippincott Williams & Wilkins*, pp. 935–36, 2006.
[6] X. Wu, R. He, Z. Sun, and T. Tan, "A light cnn for deep face representation with noisy labels," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2884–2896, 2018.
[7] D. E. Nassar, A. Abaza, X. Li, and H. Ammar, "Automatic construction of dental charts for postmortem identification," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, pp. 234–246, 2008.
[8] V. Singh, L. Malkunje, S. Mohammad, N. Singh, S. Dhasmana, and S. K. Das, "The maxillofacial injuries: A study," *National Journal of Maxillofacial Surgery*, vol. 3, no. 2, p. 166, 2012.
[9] M. Qiu, A. Barberi, and K. Lee, "Facial trauma: epidemiology, demographics and fracture patterns," *International Journal of Oral and Maxillofacial Surgery*, vol. 46, pp. 234–235, 2017.
[10] A. M. Black, D. A. Patton, P. H. Eliason, and C. A. Emery, "Prevention of sport-related facial injuries," *Clinics in Sports Medicine*, vol. 36, no. 2, pp. 257–278, 2017.
[11] K. Sauerwein, T. B. Saul, D. W. Steadman, and C. B. Boehnen, "The effect of decomposition on the efficacy of biometrics for positive identification," *Journal of Forensic Sciences*, vol. 62, no. 6, pp. 1599–1602, 2017.
[12] P. Majumdar, S. Chhabra, R. Singh, and M. Vatsa, "On detecting domestic abuse via faces," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018.
[13] G. Canzi, E. De Ponti, G. Novelli, F. Mazzoleni, O. Chiara, A. Bozzetti, and D. Sozzi, "The CFI score: Validation of a new comprehensive severity scoring system for facial injuries," *Journal of Cranio-Maxillofacial Surgery*, vol. 47, no. 3, pp. 377–382, 2019.
[14] M. Trokielewicz, A. Czajka, and P. Maciejewicz, "Iris recognition after death," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, pp. 1501–1514, 2019.
[15] D. Bolme, D. Cornett, D. Steadman, K. Sauerwein, and T. Saul, "Effects of postmortem decomposition on face recognition," in *IEEE International Conference on Biometrics Theory, Applications and Systems*, 2019.
[16] "National security strategy," https://tinyurl.com/yyb3l7m4.
[17] P. Bodkha and B. Yadav, "A role of digital imaging in identification of unidentified bodies." *Journal of Indian Academy of Forensic Medicine*, vol. 34, no. 4, pp. 0971–0973, 2012.
[18] T. Clayton, J. Whitaker, and C. Maguire, "Identification of bodies from the scene of a mass disaster using DNA amplification of short tandem repeat (str) loci," *Forensic Science International*, vol. 76, no. 1, pp. 7–15, 1995.
[19] K. Quinet, S. Nunn, and A. Ballew, "Who are the unclaimed dead?" *Journal of Forensic Sciences*, vol. 61, pp. S131–S139, 2016.
[20] M. Trokielewicz, A. Czajka, and P. Maciejewicz, "Post-mortem human iris recognition," in *IEEE International Conference on Biometrics*, 2016.
[21] P. Majumdar, S. Chhabra, R. Singh, and M. Vatsa, "Subclass contrastive loss for injured face recognition," in *IEEE International Conference on Biometrics Theory, Applications and Systems*, 2019.
[22] "International classification of external causes of injury (iceci)," https://www.who.int/classifications/icd/adaptations/iceci/en/.
[23] "Common facial injuries and complications," https://www.verywellhealth.com/facial-injuries-1298721.
[24] S. Kampakis, "Comparison of machine learning methods for predicting the recovery time of professional football players after an undiagnosed injury." in *MLSA@ PKDD/ECML*, 2013, pp. 58–68.
[25] T. A. Coulibaly, R. Béogo, I. Traoré, H. M. Kohoun, and B. V. Ili, "Inter personal violence-related facial injuries: a 10-year survey," *Journal of Oral Medicine and Oral Surgery*, vol. 24, no. 1, pp. 2–5, 2018.
[26] O. M. Parkhi, A. Vedaldi, A. Zisserman *et al.*, "Deep face recognition." in *British Machine Vision Conference*, 2015.
[27] Z. Caplova, Z. Obertova, D. M. Gibelli, D. Mazzarelli, T. Fracasso, P. Vanezis, C. Sforza, and C. Cattaneo, "The reliability of facial recognition of deceased persons on photographs," *Journal of Forensic Sciences*, vol. 62, no. 5, pp. 1286–1291, 2017.
[28] A. Abaza, A. Ross, and H. Ammar, "Retrieving dental radiographs for post-mortem identification," in *IEEE International Conference on Image Processing*, 2009.
[29] S. B. Balla *et al.*, "Forensic dental identification: Practice in indian context compared to western countries," *Journal of Forensic Science and Medicine*, vol. 2, no. 1, p. 44, 2016.
[30] T. Kitayama, T. Fukagawa, H. Watahiki, Y. Mita, K. Fujii, K. Unuma, K. Sakurada, K. Uemura, K. Sekiguchi, and N. Mizuno, "Evaluation of rapid dna system for buccal swab and disaster victim identification samples," *Legal Medicine*, p. 101713, 2020.
[31] B. T. Johnson and J. A. Riemen, "Digital capture of fingerprints in a disaster victim identification setting: a review and case study," *Forensic sciences research*, vol. 4, no. 4, pp. 293–302, 2019.
[32] J. Levinson and A. Domb, "Applying new police technologies to disaster victim identification," *Forensic Research & Criminology International Journal*, no. 5, 2016.
[33] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality reduction by learning an invariant mapping," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2006.

[34] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *European conference on computer vision*. Springer, 2016, pp. 499–515.

[35] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[36] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "Sphereface: Deep hypersphere embedding for face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[37] H. Wang, Y. Wang, Z. Zhou, X. Ji, D. Gong, J. Zhou, Z. Li, and W. Liu, "Cosface: Large margin cosine loss for deep face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[38] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

[39] Y. Zheng, D. K. Pal, and M. Savvides, "Ring loss: Convex feature normalization for face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[40] X. Zhang, R. Zhao, Y. Qiao, X. Wang, and H. Li, "Adacos: Adaptively scaling cosine logits for effectively learning deep face representations," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

[41] M. Zhu and A. M. Martinez, "Subclass discriminant analysis," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1274–1286, 2006.

[42] B. Mandal, L. Li, V. Chandrasekhar, and J. H. Lim, "Whole space subclass discriminant analysis for face recognition," in *IEEE International Conference on Image Processing*, 2015, pp. 329–333.

[43] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, no. Jul, pp. 1755–1758, 2009.

[44] B. Amos, B. Ludwiczuk, J. Harkes, P. Pillai, K. Elgazzar, and M. Satyanarayanan, "Openface: Face recognition with deep neural networks," in *IEEE Winter Conference on Applications of Computer Vision*, 2016.

[45] Y. Liu, H. Li, and X. Wang, "Rethinking feature discrimination and polymerization for large-scale recognition," in *Advances in Neural Information Processing Systems*, 2017.

[46] H. C. Law, P. Zhao, L. S. Chan, J. Huang, and D. Sejdinovic, "Hyperparameter learning via distributional transfer," in *Advances in Neural Information Processing Systems*, 2019, pp. 6801–6812.

[47] I. Ilievski, T. Akhtar, J. Feng, and C. A. Shoemaker, "Efficient hyperparameter optimization for deep learning algorithms using deterministic rbf surrogates," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[48] X. Dong, J. Shen, W. Wang, L. Shao, H. Ling, and F. Porikli, "Dynamical hyperparameter optimization via deep reinforcement learning in tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[49] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 2623–2631.

[50] Y. Yoo, "Hyperparameter optimization of deep neural network using univariate dynamic encoding algorithm for searches," *Knowledge-Based Systems*, vol. 178, pp. 74–83, 2019.

[51] M. Parsa, J. P. Mitchell, C. D. Schuman, R. M. Patton, T. E. Potok, and K. Roy, "Bayesian-based hyperparameter optimization for spiking neuromorphic systems," in *IEEE International Conference on Big Data*, 2019, pp. 4472–4478.

[52] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, "An empirical evaluation of deep architectures on problems with many factors of variation," in *International Conference on Machine Learning*, 2007, pp. 473–480.

Puspita Majumdar received the Master of Technology degree in Computer Science from the National Institute of Technology - Delhi, India, in 2017 where she was awarded the President Gold Medal for her academic performance. She is currently pursuing her Ph.D. with the Image Analysis and Biometrics (IAB) Lab at IIIT-Delhi, India. She is a recipient of the DST INSPIRE Fellowship for pursuing her Doctoral Program. Her research interests are machine learning and deep learning with applications in face recognition. She received the Best Student Paper Award at IEEE BTAS 2019.

Saheb Chhabra received the Bachelor of Technology degree in Electronics & Communication from the Lovely Professional University - Punjab, India, in 2013. Currently, he is pursuing his Ph.D. degree in Computer Science from IIIT-Delhi, India. His research interests include machine learning and deep learning with applications in privacy preservation. He has received the Best Student Paper Award at IEEE BTAS 2019. He has also served as the Program Committee member of AAAI 2021.

Richa Singh is currently a Professor at IIT Jodhpur, India. She is a Fellow of IAPR and a Senior Member of ACM. She was a recipient of the Kusum and Mohandas Pai Faculty Research Fellowship at the IIIT-Delhi, the FAST Award by the Department of Science and Technology, India, and several best paper and best poster awards in international conferences. She has also served as the Program Co-Chair of IJCB 2020, FG2019, and BTAS 2016, and a General Co-Chair of ISBA 2017. She is currently serving as the General Co-Chair of FG2021. She is also the Vice President (Publications) of the IEEE Biometrics Council. She is an Associate Editor-in-Chief of Pattern Recognition, and Area/Associate Editor of several journals.

Mayank Vatsa is a Professor at IIT Jodhpur, India, and an Adjunct Professor with IIIT-Delhi and West Virginia University, USA. His areas of interest are biometrics, machine learning, computer vision, and information fusion. He is the recipient of the prestigious Swarnajayanti Fellowship award from Government of India, A. R. Krishnaswamy Faculty Research Fellowship at the IIIT-Delhi, the FAST Award Project by DST, India, and several Best Paper and Best Poster Awards at international conferences. He is an Area/Associate Editor of Information Fusion and Pattern Recognition Journals, General Co-Chair of IJCB 2020, and the PC Co-Chair of the ICB 2013 and IJCB 2014. He has served as the Vice President (Publications) of the IEEE Biometrics Council where he started the IEEE Trans. on BIOM.