Kinship identification using age transformation and Siamese network

Arshad Abbas [Corresp., 1] , **Muhammad Shoaib** [1]

[1] Department of Computer Science, University of Engineering and Technology, Lahore, Punjab, Pakistan

Corresponding Author: Arshad Abbas
Email address: 2016phdcs01@student.uet.edu.pk

Facial images are used for kinship verification. Traditional convolutional neural networks and transfer learning-based approaches are presently used for kinship identification. Transfer-learning approach is useful in many fields. However, it lacks for identification of humans' kinship accurately because transfer-learning models are trained on a different type of data that is significantly different as compared to human face image data, a technique which may be able for kinship identification by comparing images of parents and their children with transformed age instead of comparing their actual images is required. In this paper, a technique for kinship identification by using a Siamese neural network and age transformation algorithm is proposed. The results are satisfactory as an overall accuracy of 76.38% has been achieved. Further work can be carried out to improve the accuracy by improving the Life Span Age Transformation (LAT) algorithm for kinship identification using facial images.

1
2
Kinship Identification using Age Transformation and Siamese Network

3 Arshad Abbas and Muhammad Shoaib
4 Department of Computer Science, University of Engineering and Technology Lahore
5
6 Corresponding author: Arshad Abbas (2016phdcs01@student.uet.edu.pk)
7

8 **Abstract**

9   Facial images are used for kinship verification. Traditional convolutional neural networks and
10 transfer learning-based approaches are presently used for kinship identification. Transfer-
11 learning approach is useful in many fields. However, it lacks for identification of humans'
12 kinship accurately because transfer-learning models are trained on a different type of data that is
13 significantly different as compared to human face image data, a technique which may be able for
14 kinship identification by comparing images of parents and their children with transformed age
15 instead of comparing their actual images is required. In this paper, a technique for kinship
16 identification by using a Siamese neural network and age transformation algorithm is proposed.
17 The results are satisfactory as an overall accuracy of 76.38% has been achieved. Further work
18 can be carried out to improve the accuracy by improving the Life Span Age Transformation
19 (LAT) algorithm for kinship identification using facial images.

20 **Introduction**

21 The pictorial data generated by the business, social media, public industry, non-profit sectors,
22 and scientific research have increased tremendously [1]. This graphical data contains many
23 useful and worthwhile information that could be used for various purposes [2, 3].
24 In the last few years, researchers became interested in extracting kinship information from
25 pictorial data having human faces, which can be used for different purposes. As face image data
26 provides different unique features of humans and contains a wealth of information that can be
27 used for various purposes [4]. The purpose of extracting genetic relationships between human
28 images is to verify human kinship, which is useful information for medical sciences,
29 psychologists, security agencies and family album organizations. Furthermore, it can be utilized

30   in image annotation, searching for missing children, human trafficking, and solving immigration

31   and border patrol [4, 5, 14].

32   Face Recognition and verification have been an active area of research for the last two decades.

33   It has been studied enthusiastically to make computers capable as more intelligent applications

34   have been developed for HCI (Human-Computer Interface), security, robotics, entertainment and

35   games, etc. [6]. In parallel, after face recognition, now age and gender detection techniques have

36   also been proposed. For example, Levi and Hassner proposed a classification technique for age

37   and gender using convolutional neural networks [7]. Similarly, Dehghan A. et al. proposed the

38   genetic identification technique in which they used gated autoencoders and tried to determine the

39   resemblance between parent and child [8]. They found the resemblance by using father, mother

40   and children's facial features similar to those found in anthropological studies. However, it only

41   works if there is much resemblance of offspring with parents and therefore, it causes poor

42   performance in case children have little or no resemblance with their parents. On the other hand,

43   Hu J proposed another technique of kinship verification for videos using the video face dataset

44   KFVW, which was prepared in wild conditions to handle kinship verification for the video-based

45   study. This technique addresses some pose identification. However, the experimental data

46   indicates that matric-based learning is not an effective technique for kinship identification [9].

47   Amongst these challenges, researchers are adopting the pre-trained network Convolutional

48   Neural Network as a transfer learning with CNN layered architecture and training algorithm to

49   get better results for kinship identification & verification. However, the shortcoming of such

50   models is that they lack kinship identification as transfer learning models are trained on a

51   significantly different type of data compared to human face images. Moving onwards, to solve

52   the problems of limited datasets. Joseph P. R. et al introduced a database named Recognizing

53   Family In the Wild (RFIW). It is the first large-scale image database, especially for kinship

54   recognition and exploits the challenges of kinship recognition [10].

55   Meanwhile, methodologies proposed so far have several challenges, such as limited pair of

56   images for parents and children. Moreover, a classifier trained using transfer learning and a

57   limited scale dataset fails in recognizing real-world images. In this regard, this study aims to

58   propose and provide an effective technique for performing kinship identification through image

59   data.

60   Instead of comparing direct images of parents and children, this research work suggests an

61   approach of age transformation and converts images of parents and children into the same age of

62   15-19-year age and then compares them to get better accuracy of similarity. In this model, we

63   have used a pre-processing stage of age transformation before image comparison for kinship

64   identification and verification. At first, our model uses an age transformation algorithm to

65   transform facial images by increasing or decreasing the age of face images and making them at

66   the same age stage. After making images at the same stage of age, makes it easy and will make it

67   easy to compare images and finding similarities between them to exploit kinship identification

68   between them. Furthermore, the robustness of our technique is validated through extensive

69   experiments and analysis on a huge dataset. Figure 1 shows a high level overview of the

70   proposed methodology.

71   The contributions of this paper include 1) proposing improved pre-processing of dataset images

72   through employing the use of the Life Span Age Transformation (LAT) algorithm for

73   transforming the images onto the same scale of age, 2) using the Siamese network for performing

74   the feature extraction from the transformed images, 3) introduced technique is validated by using

75   the state-of-the-art benchmarked dataset namely RFIW (Recognizing Family in the Wild), 4)

76   finally, extensive experiments conducted on the dataset using the proposed technique to identify

77   the improved effectiveness. Moreover, the comparative analysis indicates that the proposed

78   technique outperformed the existing methods.

79   *Figure 1:* High Level Methodology of Proposed Study

80   **Related Work**

81   Researchers are interested in kinship verification (family or not family) in the computer vision

82   community by applying different face recognition and machine learning techniques. Fang et al.

83   introduced the problem and used simple features for kinship identification like eyes and skin

84   color and distances between facial parts for kinship verification [11]. Subsequently. Xia et al.

85   claimed the similarity between parents and their children is quite large and proposed an approach

86   of kinship learning by removing the gap between two facial images of a parent, one image of

87   young age and another image of old age and children's images [12]. Lu et al. used a metric

88   learning approach for kinship verification and found effective features, which provided the most

89   discriminative results [13]. Levi G and Hassner T proposed a classification methodology using

90 age and gender by applying convolutional neural networks and got better results [7]. Dehghan A.
91 et al. proposed the genetic identification technique by determining resemblance between parent
92 and offspring via gated autoencoders. They used deep learning techniques to learn the most
93 discriminative features between parents and children to find out their resemblance. That
94 approach deals with resemblance by using the father and mother's facial shapes and extracting a
95 similar face with a combination of facial feathers of the father and mother [8]. Yan H, Hu J.
96 revealed that Euclidian similarity metric is not a powerful way to measure the similarity of facial
97 images, especially when captured in wild conditions. They clarified that the similarity metric can
98 handle the problem better to deal with face variations compared to Euclidian similarity. They
99 used a mid-level feature vector with discriminative metric learning and proposed a prototype-
100 based feature learning approach for kinship verification [13, 14]. Yan H, Hu J. proposed a
101 methodology of video-based kinship verification by using data set of video faces called Kinship
102 Face Videos in the Wild (KFVW). Dataset was built by capturing facial images from videos for
103 kinship verification. This methodology analyzes the human faces in the video by getting training
104 set from video poses and then applying distance metric learning approaches to get a positive
105 semi definite matrix (PSD) for face recognition and kinship identification [9]. Joseph P.R. et al.
106 introduced the first large-scale image database for kinship recognition called Families In the
107 Wild (FIW) and exploits the challenges in kinship recognition. The FIW database consists of
108 thousands of images of faces for kinship recognition [10]. Yong L. et al. presented a framework
109 in which knowledge of face recognition from large-scale data-driven transferred and then fine-
110 tuned metric space to get discrimination of kin related people. They also proposed an augmented
111 strategy to balance family members' images and used triplet and ResNet to extract face encoding
112 for kinship identification [15]. In early techniques, kinship verification uses handcrafted
113 descriptors from facial images to perform classification for learning. Fang et al. used facial
114 features like eye and skin colors and distance of eye-to-nose for kinship verification [11]. Zhou
115 et al. proposed an approach based on spatial pyramid features for kinship verification. This
116 approach used Gabor-based facial image gradient orientation features [16]. Liu et al. applied a
117 transferrable approach of fisher vectors derived from each facial image to extract similarity for
118 kinship verification [17]. Kohli et al. proposed an approach to achieving kinship similarity using
119 a self-similarity descriptor. They introduced that kinship verification is a two-factor classification
120 problem. They revealed that low-level features could not be used as an underlying source of

121   visual resemblance between people with kinship relations [18]. In Shallow metric-based
122   approaches, metric learning methods are used to learn discriminative features for kinship
123   verification. These approaches learn a Mahalanobis distance using handcrafted features
124   identification and get a better score of similarity between kinship-related pairs with non-kinship-
125   related pairs [5]. In the Deep learning-based approach, Kaiming He [19], Kohli et al [20]
126   motivated kinship identification and verification after getting impressive success by applying
127   deep learning approaches to classify different facial images. Many techniques have been adopted
128   for deep metric learning to get discriminant features for kinship verification. Dehghan et al.
129   introduced an approach of fusing the features using gated auto-encoders. They extracted optimal
130   features by reflecting parent-offspring resemblance [8]. Wang S et al. proposed the Kinship
131   Verification on Families in the Wild with Marginalized Designing Metric Learning (DML). That
132   technique used the largest kinship verification using Auto-encoder and Discriminative Low-rank
133   Metric Learning (DLML) algorithm for feature discrimination [21]. After using matric learning,
134   researchers found a better way to find similarity for kinship identification by using a
135   Convolutional neural network. Zhang et al. adopted an approach of kinship verification using a
136   convolution neural network (CNN) to train the algorithm with concentrated image pairs [22].
137   Kaming He et al. introduced a deep residual learning approach for image recognition. Their
138   approach used residual training with neural networks and multiple layers as learning residual
139   functions [19]. Duan Q, Zahng L., Zuo W proposed a deep kinship verification technique named
140   Coarse to Fine Transfer(CFT) using Convolutional Neural Network (CNN) from face
141   recognition to kinship recognition and used Deep Transfer Learning [23]. Yana H and Hu J
142   proposed a kinship verification technique, which works on videos. This technique uses distance
143   metric learning on Kinship Face Videos in the Wild (KFVW) dataset for kinship verification
144   [24]. Lu J. et al. developed a discriminative deep multi-metric learning (DDMML) methodology.
145   They used multiple neural networks jointly to maximize the association of different features of
146   each sample and reduce the distance of positive pair and increase the distance of negative pair
147   [25]. Yong Li et al. introduced the kinship verification technique using KinNet: Fine-to-Coarse
148   Deep Metric Learning and Pre-training the network and minimizing a soft triplet loss. They used
149   four CNN networks to boost the performance [15]. Liu W. et al. introduced SphereFace, a deep
150   hyper sphere embedding for face recognition. They addressed angular SoftMax loss and angular
151   margins problem. Their technique uses a 64-layer CNN neural network for training and

152 discriminative constraints on a hypersphere to get a better face recognition (FR) problem under

153 open-set protocol [26]. Savas and Akin introduced an approach of synthesizing child faces with a

154 pre-trained model by analyzing facial images of parents [27]. Habin Y. & Chaohui S worked on

155 Multi-scale Deep Relational Reasoning for Facial Kinship Verification and used two

156 convolutional neural networks, which shared network parameters and extracted different scales

157 of features for kinship identification [28]. After using a convolutional neural network,

158 researchers moved to find kinship using the Generative Adversarial Network (GAN), introduced

159 by Iain Good fellow [29]. Fady S. et al. proposed GANKIN: generating Kin faces using

160 disentangled GAN and image synthesis approach from parents to children, they also used

161 pertained FaceNet and GAN network [30]. Tuan H et al. proposed an approach of recognizing

162 families through images with pertained encoders. They used pre-trained networks FceNet,

163 Siamese and FGG network to get face image encoding and find kinship between facial images

164 [31]. Keeping in view the efficiency factor of GAN based approaches, we also used GAN based

165 age transformation algorithm and Siamese network to build and train our model.

166 Although some encouraging results have been obtained from proposed methodologies for

167 kinship identification and verification in the last few years, automatic kinship verification is

168 being performed poorly in the real-world applications used in daily life. Due to the non-

169 availability of large-scale datasets, results are not too accurate to handle the kinship identification

170 problems. Existing datasets like Family101, UB KinFace, Cornell KinFace, KinFaceW-I, and

171 KinFaceW-II provide a few examples, but they fail to achieve accurate distributions of genetic or

172 kinship relationships. Moreover, they have a limited pair of images for parents and children;

173 Classifier trained on a limited scale dataset fails while recognizing real-world images.

174 To handle these issues, we proposed an approach to find the kinship relationship between

175 parents and children. Our methodology uses Age Transformation and converts images of

176 parents and children to the age of 15-20 because images of this age have maximum facial

177 features, which can be a good source for the discrimination of features between facial images.

178 After the process of age transformation and converting facial images to a young age for both

179 parents and children, these faces get closer to each other in facial look and expression and then

180 it makes it easy to find the similarity between them. With these images, there is much

181 probability of getting parent's faces and images close to each other. Ultimately, it will make it

182   easy for the face encoder to generate close face encoding. As a result, we get a low distance

183   value while finding cosine similarity. Figure 2 shows the effect of age transformation

184   **Proposed Work**

185   This section outlines the proposed methodology for performing the kinship identification. In the

186   proposed method, we presented a model of a deep relational network that uses a preprocessing

187   stage of age transformation of two facial images before comparing them to exploit kinship

188   relationships from facial images. This scheme first transforms facial images by increasing or

189   decreasing the age factor and making two images into the same age stage and then compares

190   them to find and verify kinship. After transforming facial images, we proposed the use of a

191   Siamese network with two convolutional neural networks by sharing parameters between them.

192   Afterward, it extracts different scales of features to find similarities between images by using

193   triplet loss. We also aimed to conduct experiments on a widely used facial kinship dataset,

194   namely RFIW. In this methodology, the proposed model uses age transformation and converts

195   facial images at the same stage of age, between 15 and 19 years. However, we considered this

196   age because, in this age period, a person's face looks strong and can provide clear facial features

197   and better encode facial images. Furthermore, after encoding transformed faces, we applied

198   triplet loss on three faces of parents and images and extracted the kinship relationship between

199   parents and images. In addition, we have employed parent's images as anchor and negative part

200   of the triplet while children's images as a positive part of the triplet. We fixed the father and

201   mother position of being positive or negative to each other while training in the Siamese

202   network. Likewise, we used an age transformation algorithm that provided close pair of facial

203   images of parents and children for processing to exploit kinship identification between them.

204   This age transformation algorithm will provide images for processing to consider for kinship

205   identification. More graphical representation and the working flow of our proposed methodology

206   is depicted in Figure 2.

207   *Figure 2:* Effect of Age Transformation

208   **Model Training**

209   The proposed model uses age transformation and feature encoding of face images with triplet

210   loss to extract facial similarity to identify kinship. The first stage converts facial images to

211   images of persons having approximate ages of 15-19 years. After doing conversion of two input

212 images with ages between 15-19 years, these converted images are processed with the Siamese
213 network to extract feature encoding for further processing. Working of proposed methodology is
214 described in figure 3.

215   Figure 3: Working of Proposed Methodology

216 It uses ResNet 50 with two fully connected layers and one Dense output Layer to extract
217 features. It extracts a feature vector of 128x128 for all input facial images and uses triplet loss to
218 discriminate features for kinship identification. It maximizes the distance of the anchor image
219 with a negative image and minimizes the distance with a positive image. The size of input
220 images is 224x224x3 and the feature vector returned by the Siamese network is 128. During the
221 training process, hard sample selection for positive or negative pairs are not equally important.
222 The pairs with higher loss might have more impact on the model training. The training set can be
223 defined as: Let X^a, X^p and X^n are finite set of images for Father, Children and Mother having 'm'
224 number of images for each set.

225 $X^a = \{x^a_1, x^a_2, \ldots.. x^a_m\}$ (1)

226 Where X^a is set of anchor images for father images.

227 $X^p = \{x^p_1, x^p_2, \ldots.. x^p_m\}$ (2)

228 Where X^p is set of positive images which are taken from children's images.

229 $X^n = \{x^n_1, x^n_2, \ldots.. x^n_m\}$ (3)

230 X^n is a set of negative images taken from the set of mother images.

231 Then input sample taken from these three sets will be a powerful set of three sets to make a set of
232 triplets let X is the power of X^a, X^p and X^n, set then we get set X as a set of the triplet.

233 $X = \{(x^a_1, x^p_1, x^n_1), (x^a_2, x^p_3, x^n_4) \ldots (x^a_n, x^p_n, x^n_n)\}$ (4)

234 X is a power set of images having three members as triplet of anchor as x^a, positive as x^p and
235 negative image as x^n respectively where sequence of triplet members are anchor, positive and
236 negative members with images of father, child, and mother respectively. After getting feature
237 extracted from pertained Siamese network, we get a set of features:

238 $F(X) = \{[f(x^a_1), f(x^p_1), f(x^n_1)], \{[f(x^a_2), f(x^p_2), f(x^n_2)] \ldots \{[f(x^a_n), f(x^p_n), f(x^n_n)]\}$ (5)

239   This sequence of the set is used for extracting similarities of children with father and mother to

240   get kinship relation of Father-Son (F-S), Father-Daughter (F-D), Mother-Son (M-S) and Mother-

241   Daughter. For sibling relationships, we changed some sequence of power set. We took one

242   sibling image as an anchor, one as positive, and one as negative if the third image of the sibling

243   did not exist. For negative position, we took any random image from the set of mother or father.

244   So, for negative position random set of images: $Xr = P \{Xa \mid Xn\}$. Then set of triplets for

245   sibling relationship Brother-Brother (B-B), Sister-Sister (S-S) and Brother-Sister (B-S) is as

246   follows:

247    $X^s = \{(x^a_1, x^p_2, x^r_1), (x^a_2, x^p_3, x^r_2) \dots (x^a_n, x^p_m, x^r_n)\}$ (6)

248   Where X^s is a power set of images having three members as triplet of anchor as x^a, positive as x^p

249   and negative image as x^r respectively.

Loss Function

251   The loss function for the triplet loss on the extracted feature, For three cases

252   1.   While comparing the father's image with the child's image, if D_f is distance of child's

253   image with father's image and D_m is distance of child's image with mother's image then we

254   define the loss function as:

255    $D_{f=} \|f(P) - f(A)\|^2, \ D_{m=} \ \|f(P) - f(N)\|^2$ (7)

256   and some margin 'm' as hyperparameter, whereas A, P, N are anchor, positive and negative

257   images, and f(A), f(P) and f(N) are features of father, child, and mother respectively. If father's

258   image is closer to child's image then we increase the distance of child's image with mother's

259   image and decrease the distance of child's image with father's image, so loss function to get

260   similarity between father and child will be:

261    $£_f (A, P, N) = max (D_f\text{-} D_m + m, 0) \ if \ D_f < D_m$ (8)

262   2.   While checking the similarity of children with mothers then, we revert the loss function.

263   To find the similarity of the child image with the mother image, we increase the distance of child

264   image with father image and decrease the distance of child image with mother image then loss

265   function will be:

266          $\pounds_m\,(A,\,P,\,N) = max\,(D_m - D_f + m,\,0)\,if\,\,D_f > D_m$ (9)

267   3.   While comparing images of siblings, we use distance measures of images of two siblings,

268   S1, S2. We find the distance between siblings and random images as:

269    $Ds_1 = ||f(S1) - f(S2)||^2, D_{s2} = ||f(S1) - f(N_r)||^2$ (10)

270   where $f(S1),\ f(S2)\ and\ f(Nr)$ features of siblings and a random image, respectively. After

271   calculating distance and using margin 'm' as hyper parameter, we can define the loss function as:

272          $\pounds s\,(A,P,N) = Max\,(\,D_{s1} - Ds_2 + m\,,0).$ (11)

273   D_{s1} is the distance between one sibling with other siblings. Similarly, D_{s2} is the distance of

274   siblings with a random image to find triplet loss and minimize the distance between the first and

275   second siblings.

276   **Network Structure**

277   To select information from different scales of features for input to the relational network, we use

278   the pre-trained Siamese network and get a feature map of size $R^{512\times1}$. Network contains three

279   dense layers to down sample the features map and get a features vector of size 128x1. Each

280   features vector of size 128x1 will provide information of the faces as face encoding. This face

281   encoding is then used to find the cosine similarity between face images respectively. After that

282   relational network analyzes these selected features with multi-layer perceptrons which consists

283   of some fully connected layers and relu activation functions.

284   Following are steps of model training:

285   1. Pictorial data is fetched from the data set and all images are converted to the same stage

286   of age by LATS. After age transformation, an intermediate dataset is prepared for

287   training from original images.

288   2. Transformed data is fetched into three vectors: father, mother and children, to prepare a

289   triplet for the Siamese network

290   3. One vector is used as positive, one for negative and one as anchor

291   4. The triplet is used by the Siamese network to extract face features

292   5. Define the Triplet loss function. It decreases the distance between positive and anchor

293   images and increases the distance between positive and negative images.

294   6. Setting up for training and evaluation

295   7. This multi-layer perceptron will extract the relation of features and output feature of size

296       $R^{128 \times 1}$. Then we compare these features of size $R^{128 \times 1}$ at the element level to represent the

297   distance between features of faces.

298   8. Lastly, we use another multi-layer perceptron to find the similarity of faces for kinship

299   identification from the relation of different face images. It also consists of some fully

300   connected layers and relu activation functions.

301   A flow of model traing is represented in figure 4.

302   *Figure 4:* Flow of *Model Training with Age Transformation and Face Encoding*

303   The CNN structure uses Saimese network; its input size is 3*256*256 and final output features

304   vector size is 128*1. This network has three dense layers subsequently with batch normalization

305   and relu activation function to minizise the size of feature vector.

306   The relational network has three convolutional layers, each layer uses 128 feature vector of 10

307   images with batch normalization and relu activation function. The input feature size of each layer

308   is $R^{10 \times 128 \times 128 \times 128}$ and last dense layer has output feature size of 1x128. It applies segmoid function

309   to establish kin relationship between images, detailed relational network with input parameters is

310   depicted in Table-1.

311   *Table 1: Parameters of Deep Relational Network*

312   To optimize the network, contrastive loss function is used with below specifications:

313       $L(d,Y) = \sum N_i \left[1/2 \, {}^*Y_i {}^* d_i^2 + (1 - y_i) \, {}^* 1/2 {}^* max(0, m - d_i) \right]$ (12)

314   where L denotes the loss, N represents the number of samples, y_i is the ground truth of ith

315   sample, and d_i is the distance between the output of the encoder, m is margin parameter.

316   Similar face images are pushed close and dissimilar images pushed away to get maximum

317   similarity between similar images.

318   **Data Set**

319   We used a dataset of Recognizing Family in the Wild (RFIW) and took images of 200 families

320   having good resolution images as a constraint of Life Span Age Transformation (LATS) that

321   requires images having good resolution. LATS generates ten age clusters and each age cluster

322 has ten images. We picked one age cluster between 15 and 19, so we used 10 images to train our
323 model. We used images of ages between 15 and 19 years because in this age period person's face
324 looks strong and can provide clear facial features and we can get the better encoding of facial
325 images. For model training, we used images of 200 families; each family has average four
326 members. For each member we used 10 transformed images and our model is trained on
327 approximate 200x4x10=8000, from this pool of data, we used 30% data for validation.

328 As we used the LATS model for preprocessing and Siamese Network for training our model,
329 which are CNN based network architecture, therefore we also adopted CNN network for feature
330 extraction and training model to get compatibility with the existing model and achieve efficient
331 results. Moreover, for high dimensional data CNN provides automatic feature extraction and
332 forwards extracted features to the classifier to get classification results.

333 **Results and Discussion**

334 The CNN-based deep relational network is utilized for extracting the features from the facial
335 images of the dataset. *Table 1.* outlines the details of the included parameters for the CNN-based
336 deep relational network. Unlike the previously existing models, it represents that our model
337 explicitly establishes relations between three feature maps rather than making relations within
338 one another. Additionally, it depicts that our model takes ten images of each member and finds
339 the triplet loss on 128 features maps of each ten images for one member. In total used, 30
340 features map for one comparison to find the similarity between them. The proposed model
341 delivers the optimal performance by utilizing this methodology.

342 In this section of the study, we have listed the experiments and achieved results by employing
343 the use of the proposed technique of utilizing a deep relational network along with the LAT age
344 transformation algorithm [32]. We have used the large dataset of Recognizing family in the wild
345 (RFIW) for the training and validation of our proposed technique. In the first phase, we
346 converted images of datasets RFIW to different life stages for age transformation. After the age
347 transformation of facial images, we converted images at the same stage of ages by adjusting the
348 age factor. In the first stage, we transformed facial images by increasing or decreasing the age
349 factor and making two images into the same stage of age. In the second phase, we trained our
350 algorithm by comparing two images and evaluating metrics and parameter settings to extract
351 kinship relation accuracy.

352   **Age Transformation**

353   For Age Transformation, we employed the Lifespan Age Transformation Synthesis algorithm,

354   proposed by Roy Or-El et al. [32]. Using this algorithm, we prepared our data set images for

355   comparison that converts images at different stages of life. Afterward, we picked the images of

356   ages between 15-19 and used them for feature extraction. Table 2. outlines the training and

357   validation accuracies observed in different relationships by utilizing the proposed model.

358   *Table 2: Achieved Model Performance for different Relationships*

359   Similarly, Table 3. represents the observed results on the baseline dataset. While comparing

360   accuracy with a model trained on dataset RFIW, the results from Table 3. indicate that our

361   proposed model has delivered better performance than the existing state-of-the-art models by

362   improving the overall accuracy.

363   *Table 3: Comparison of Results on Baseline Dataset*

364   Meanwhile, the previous models have failed to deliver improved performance for up to 73.21%

365   accuracy. On the other hand, the proposed model has outperformed existing state-of-the-art

366   models by delivering an accuracy of 76.38%. Furthermore, we plan to improve the model and

367   accuracy in the future by improving the underlying relational network and applying it to

368   transformed images with the same stage of age.

369   The major contribution of our research is to introduce a robust way of kinship identification by

370   comparing images of parents and their children with transformed ages instead of comparing their

371   actual images. Improved accuracy of methodology proved that we could get better results for

372   kinship identification if we compare images after age transformation instead of comparing direct

373   actual images. From the results obtained after training indicates that similarity between the same

374   genders is greater than opposite gender because the similarity score between father-son and

375   mother-daughter is greater than father-daughter and mother-son, respectively. The obtained

376   results show that due to the same gender factor, daughter looks more similar to the mother

377   compared to the father. Similarly, the son seems more similar to father rather than the mother.

378   **Conclusion**

379   Kinship identification is used for kinship verification by using facial images. Meanwhile, the

380   previous studies have explored this area by employing transfer learning-based solutions. This

381   study, however, presents a different approach to perform kinship verification.

382   In this study, we have introduced a technique that uses a pre-trained LAT model along with a

383   Siamese network for performing kinship identification. Additionally, we have employed the age

384   transformation approach to find similarities between parents with children. The extensive

385   experimental results were used to validate the performance of our proposed model. Furthermore,

386   the comparative analysis with previously carried out studies reflects that our model outperformed

387   the existing state-of-the-art models using a similar approach, thereby delivering an overall

388   accuracy of 76.38%. In the future, we aim to improve the model performance by improving the

389   underlying relational network and applying it on transformed images with the same age stage.

390   **References**

391   [1] Jin, Xiaolong, et al. "Significance and challenges of big data research." Big Data

392   Research 2.2 (2015): 59-64

393   [2] Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile networks and

394   applications, 19(2), 171-209.

395   [3] Emani, C. K., Cullot, N., & Nicolle, C. (2015). Understandable big data: a survey. Computer

396   science review, 17, 70-81.

397   [4] Zhou, X., Hu, J., Lu, J., Shang, Y., & Guan, Y. (2011, November). Kinship verification from

398   facial images under uncontrolled conditions. In Proceedings of the 19th ACM international

399   conference on Multimedia (pp. 953-956). ACM

400   [5] Lu, J., Zhou, X., Tan, Y. P., Shang, Y., & Zhou, J. (2013). Neighborhood repulsed metric

401   learning for kinship verification. IEEE transactions on pattern analysis and machine

402   intelligence, 36(2), 331-345.

403   [6] Bowyer, K. W., Chang, K., & Flynn, P. (2006). A survey of approaches and challenges in 3D
404   and multi-modal 3D+ 2D face recognition. Computer vision and image understanding, 101(1), 1-
405   15.

406   [7] Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural
407   networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition
408   Workshops (pp. 34-42).

409   [8] Dehghan, A., Ortiz, E. G., Villegas, R., & Shah, M. (2014). Who do I look like? Determining
410   parent-offspring resemblance via gated autoencoders. In Proceedings of the IEEE Conference on
411   Computer Vision and Pattern Recognition (pp. 1757-1764).

412   [9] Yan, H., & Hu, J. (2018). Video-based kinship verification using distance metric
413   learning. Pattern Recognition, 75, 15-24.

414   [10] Robinson, J. P., Shao, M., Zhao, H., Wu, Y., Gillis, T., & Fu, Y. (2017, October).
415   Recognizing families in the wild (RFIW) data challenge workshop in conjunction with acm mm
416   2017. In Proceedings of the 2017 Workshop on Recognizing Families in the Wild (pp. 5-12).

417   [11] Fang, R., Tang, K. D., Snavely, N., & Chen, T. (2010, September). Towards computational
418   models of kinship verification. In Image Processing (ICIP), 2010 17th IEEE International
419   Conference on (pp. 1577-1580). IEEE.

420   [12] Xia, S., Shao, M., Luo, J., & Fu, Y. (2012). Understanding kin relationships in a
421   photo. IEEE Transactions on Multimedia, 14(4), 1046-1056.

422   [13] Lu, J., Zhou, X., Tan, Y. P., Shang, Y., & Zhou, J. (2014). Neighborhood repulsed metric
423   learning for kinship verification. IEEE transactions on pattern analysis and machine
424   intelligence, 36(2), 331-345.

425   [14] Yan, H., Lu, J., Deng, W., & Zhou, X. (2014). Discriminative multimetric learning for
426   kinship verification. IEEE Transactions on Information forensics and security, 9(7), 1169-1178.

427   [15] Li, Y., Zeng, J., Zhang, J., Dai, A., Kan, M., Shan, S., & Chen, X. (2017, October). Kinnet:

428   Fine-to-coarse deep metric learning for kinship verification. In Proceedings of the 2017

429   workshop on recognizing families in the wild (pp. 13-20).

430   [16] Zhou, X., Lu, J., Hu, J., & Shang, Y. (2012, October). Gabor-based gradient orientation

431   pyramid for kinship verification under uncontrolled environments. In Proceedings of the 20th

432   ACM international conference on Multimedia (pp. 725-728). ACM.)

433   [17] Robinson, J. P., Shao, M., Wu, Y., Liu, H., Gillis, T., & Fu, Y. (2018). Visual kinship

434   recognition of families in the wild. IEEE Transactions on pattern analysis and machine

435   intelligence, 40(11), 2624-2637.

436   [18] N Kohli, R Singh, and M Vatsa. 2012. Self-similarity representation of Weber faces for

437   kinship classification. In IEEE Fifth International Conference on Biometrics: Theory,

438   Applications and Systems. 245–250.

439   [19] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition.

440   In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-

441   778).

442   [20] Kohli, N., Vatsa, M., Singh, R., Noore, A., & Majumdar, A. (2017). Hierarchical

443   representation learning for kinship verification. IEEE Transactions on Image Processing, 26(1),

444   289-302

445   [21] Wang, S., Robinson, J. P., & Fu, Y. (2017, May). Kinship verification on families in the

446   wild with marginalized denoising metric learning. In 2017 12th IEEE International Conference

447   on Automatic Face & Gesture Recognition (FG 2017) (pp. 216-221). IEEE.

448   [22] Zhang, K., Huang, Y., Song, C., Wu, H., Wang, L., & Intelligence, S. M. (2015). Kinship

449   verification with deep convolutional neural networks.

450   [23] Duan, Q., Zhang, L., & Zuo, W. (2017). From face recognition to kinship verification: An

451   adaptation approach. In Proceedings of the IEEE International Conference on Computer Vision

452   Workshops (pp. 1590-1598).

453   [24] Yan, H., & Hu, J. (2018). Video-based kinship verification using distance metric

454   learning. Pattern Recognition, 75, 15-24.

455   [25] Lu, J., Hu, J., & Tan, Y. P. (2017). Discriminative deep metric learning for face and kinship

456   verification. IEEE Transactions on Image Processing, 26(9), 4269-4282.

457   [26] Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017). Sphereface: Deep hypersphere

458   embedding for face recognition. In Proceedings of the IEEE conference on computer vision and

459   pattern recognition (pp. 212-220).

460   [27] Ozkan, S., & Ozkan, A. (2018, October). Kinshipgan: Synthesizing of kinship faces from

461   family photos by regularizing a deep face network. In 2018 25th IEEE International Conference

462   on Image Processing (ICIP) (pp. 2142-2146). IEEE.

463   [28] Yan, H., & Song, C. (2021). Multi-scale deep relational reasoning for facial kinship

464   verification. Pattern Recognition, 110, 107541.

465   [29] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... &

466   Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139-

467   144.

468   [30] Ghatas, F. S., & Hemayed, E. E. (2020). Gankin: generating kin faces using disentangled

469   gan. SN Applied Sciences, 2(2), 1-10.

470   [31] Nguyen, T. D. H., Nguyen, H. N. H., & Dao, H. (2020, November). Recognizing Families

471   through Images with Pretrained Encoder. In 2020 15th IEEE International Conference on

472   Automatic Face and Gesture Recognition (FG 2020) (pp. 887-891). IEEE.

473   [32] Or-El, R., Sengupta, S., Fried, O., Shechtman, E., & Kemelmacher-Shlizerman, I. (2020,

474   August). Lifespan Age Transformation Synthesis. In European Conference on Computer

475   Vision (pp. 739-755). Springer, Cham.

Figure 1

High Level Methodology of Proposed Study
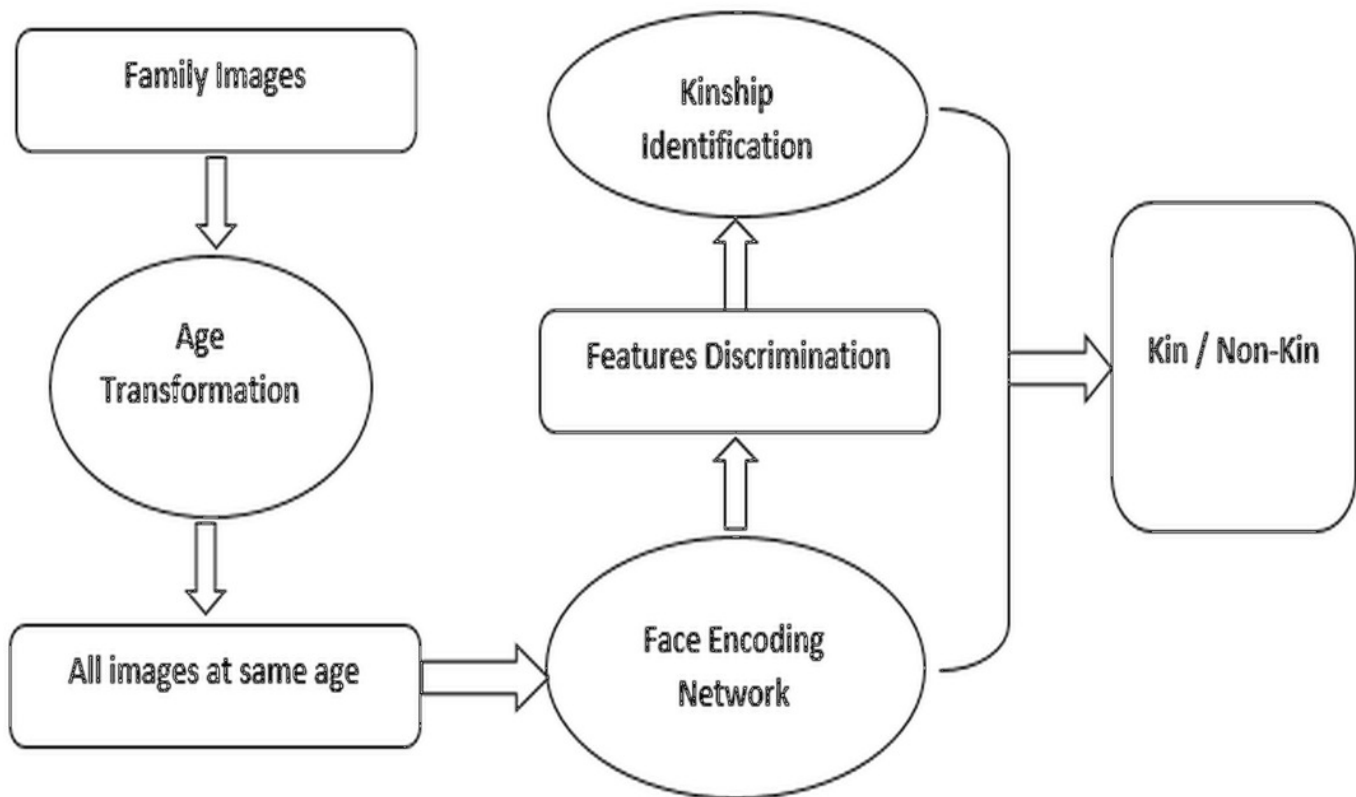
Diagram showing working of methodology

Figure 2

Effect of Age Transformation

(A) Diagram to eloborate Effect of Age Transformation on Facial Images. (B)

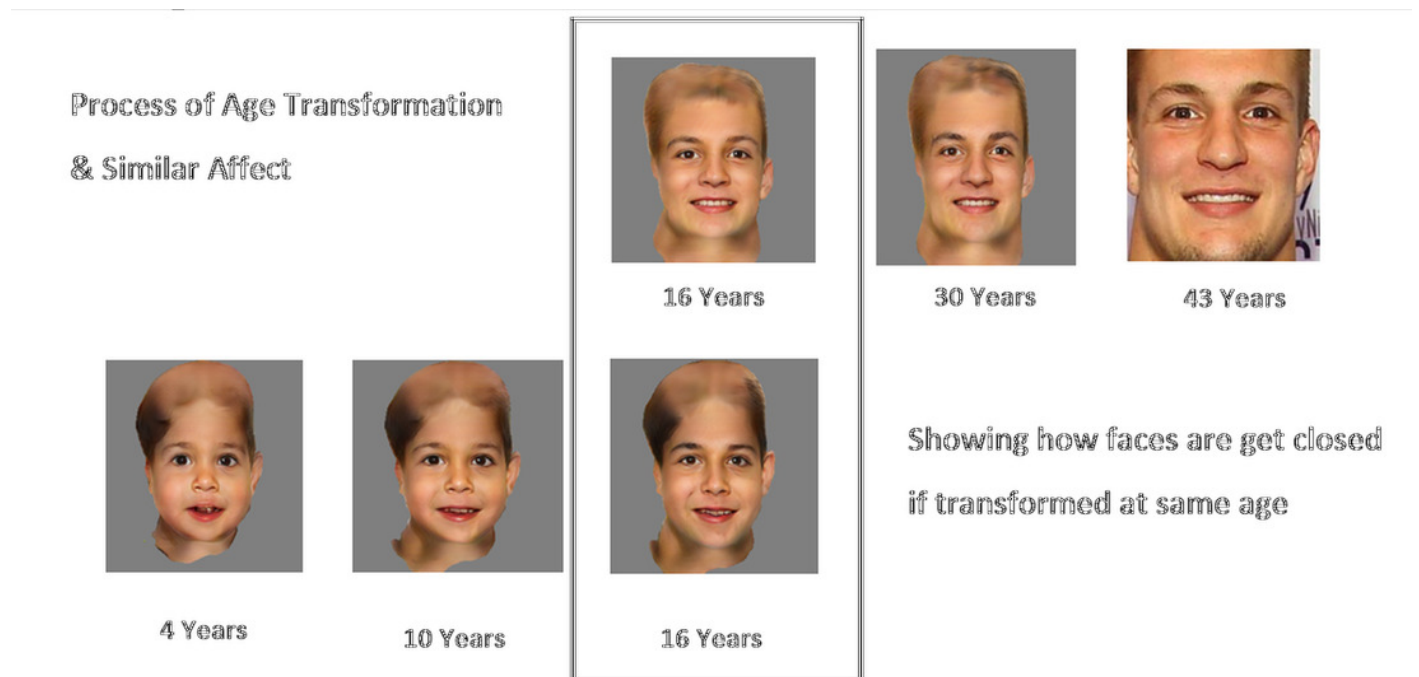https://web.northeastern.edu/smilelab/fiw/

Figure 3

Working of Proposed Methodology

(A) This Diagram shows Working of Proposed Methodology, how methodology works to achieve kinship identification

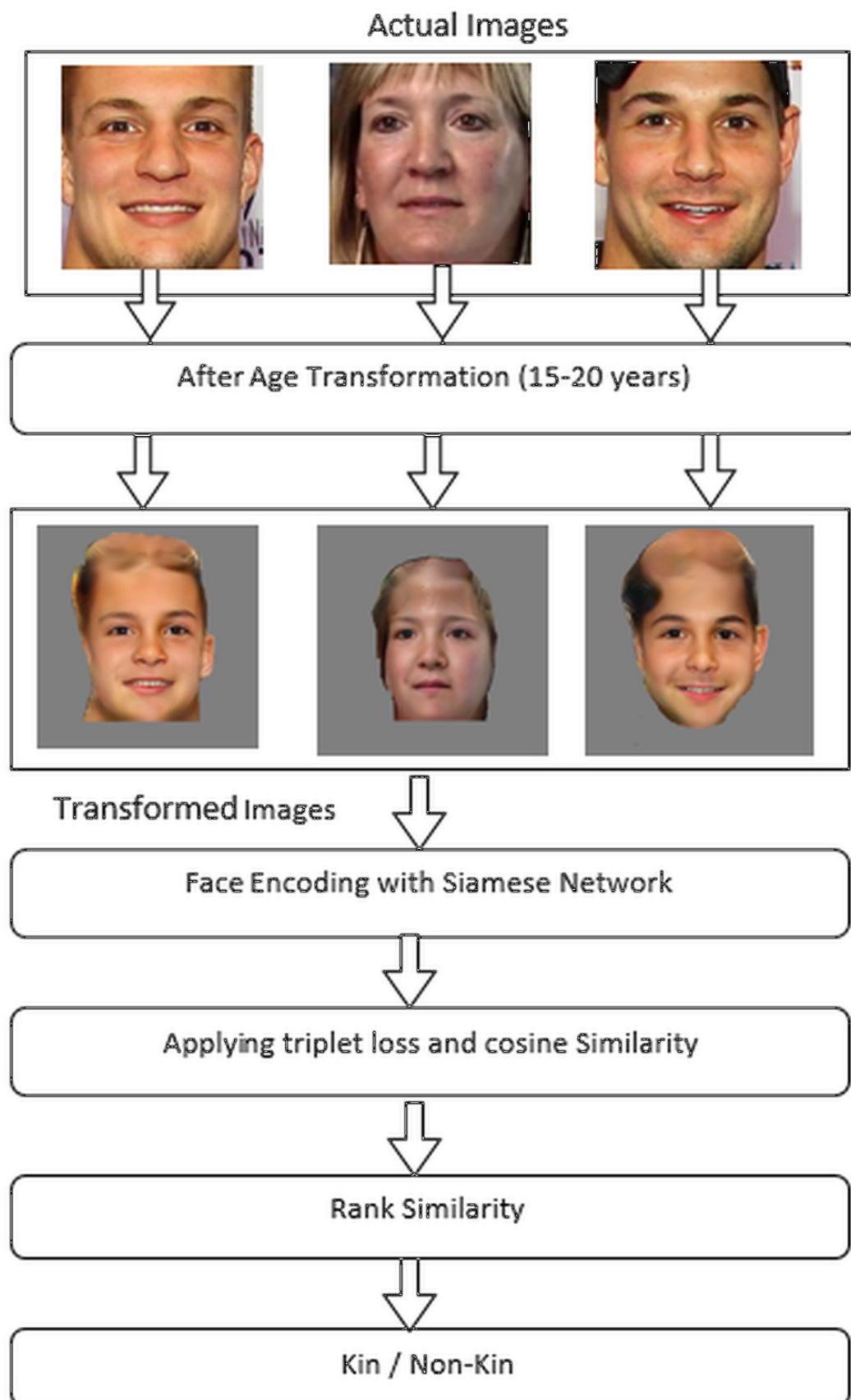(B) https://web.northeastern.edu/smilelab/fiw/

Figure 4

Flow of Model Training with Age Transformation and Face Encoding

(A) This Diagram shows Flow of Model Training with Age Transformation and Face Encoding
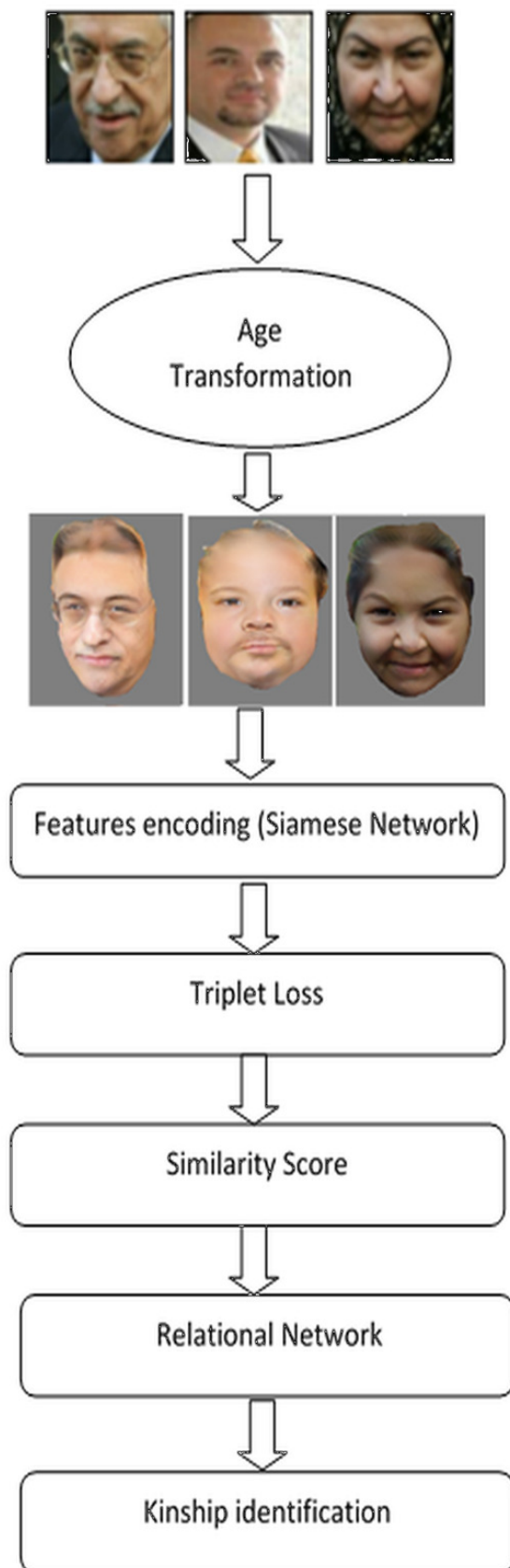
(B) https://web.northeastern.edu/smilelab/fiw/

Table 1(on next page)

Parameters of Deep Relational Network

Table showing parameters of training network

1   *Table 1:* Parameters of Deep Relational Network

Setting	Input Size	Output Size	Kernel	Stride	Padding
SIAMESE NETWORK	3*256*256	512*1	3	1	**0**
DENSE-1+ BN + RELU	512*1	256*1	2	2	**0**
DENSE-2 + BN + RELU	256*1	256*1	2	1	**0**
DENSE-3	**256*1**	**128*1**	**2**	**2**	**0**
Relational Network					
Conv-1+ BN + RELU	10x128x128x128		3	1	2
Conv-2+ BN + RELU	10x128x128x128		3	1	2
Conv-3+ BN + RELU	10x128x128x128		3	1	2
Dense (Flatten)+SIGMOID	1x128	**Kin / Not Kin**			

2

Table 2(on next page)

Achieved Model Performance for different Relationships

Table shows achieved Model Performance for different Relationships after training

1   *Table 2:* Achieved Model Performance for different Relationships

Training Accuracy		Validation Accuracy		
Relation	**Accuracy (%)**	**Relation**	**Accuracy (%)**	**Mean (%)**
Father-Son	80.12	Father-Son	76.16	78.14
Father-Daughter	77.15	Father-Daughter	73.00	75.07
Mother-Son	75.75	Mother-Son	72.75	74.25
Mother-Daughter	79.17	Mother-Daughter	77.0	78.08
Overall Mean				**76.38**

2

Table 3(on next page)

Comparison of Results on Baseline Dataset

Table showing Comparison of Results on Baseline Dataset with other models

1   *Table 3:* Comparison of Results on Baseline Dataset

Sr. No.	Methodology	Classifier	Accuracy
1	Joseph P. Robinson (2018). Visual Kinship Recognition of Families in the Wild	Cosine similarity SVM	69.18%
2	Ghatas, F. S., & Hemayed, E. E. (2020). GANKIN: generating Kin faces using disentangled GAN	GAN	71.16%
3	Tuan H et al. (2020). Recognizing Families through Images with Pretrained Encoder	Pre-trained CNN	73.21%
4	**Proposed Methodology**	**Pre-trained LATS + Siamese**	**76.38%**

2