Kin-Wolf: Kinship-established Wolfs in Indirect Synthetic Attack

Pallabi Ghosh
University of Florida
Gainesville, FL 32611
pallabighosh@ufl.edu

Sumaiya Shomaji
University of Kansas
Lawrence, KS 66045
shomaji@ku.edu

Mengdi Zhu
University of Florida
Gainesville, FL 32611
zhum@ufl.edu

Damon L. Woodard
University of Florida
Gainesville, FL 32611
dwoodard@ece.ufl.edu

Domenic Forte
University of Florida
Gainesville, FL 32611
dforte@ece.ufl.edu

Abstract

Two common attacks against biometric systems are direct (or physical) access and indirect (or logical) access. While most detection techniques focus on the former, often called presentation attacks, that occur at pre-sensor level, the attack surface for indirect access, that takes place post-sensor, is larger. In this paper, an indirect attack in the realm of faces is explored that utilizes a unique soft-biometric feature called 'Kinship Cues'. Unlike gender and ethnicity, kinship is less explored but powerful; we find that its knowledge can significantly increase the chances of an attacker getting access to a system. Due to lack of kin data in other domains, our attack is only performed against facial biometric systems. Nevertheless, the results underscore the impact of kinship cues and their need to be investigated in other domains such as fingerprint and iris. This kinship artifact boosts the convergence speed of state-of-the-art iterative adaptive Bayesian hill climbing attacks. Further, it is exploited to generate a dictionary of input images, commonly called wolf images, in a novel kinship-based non-iterative indirect attack that we call Kin-Wolf. A classical image fusion technique (morphing) and a deep learning based kinship framework utilizing pre-trained StyleGAN2 are investigated to generate the wolf images. The trade-off between kinship cues and randomization is also studied and a 6× average improvement in attack accuracy is achieved for Kin-Wolf over random probes.

1. Introduction

Over the years, eight different attack points are identified in a biometric system to falsely get authenticated as a system's registered user [1]. Out of these eight identified points, only one point, which is pre-sensor belongs to the direct (or physical) attack class. The most common type of attack performed at this point is the presentation attack. The remaining seven different attack points are places where an indirect (or logical) attack can take place.

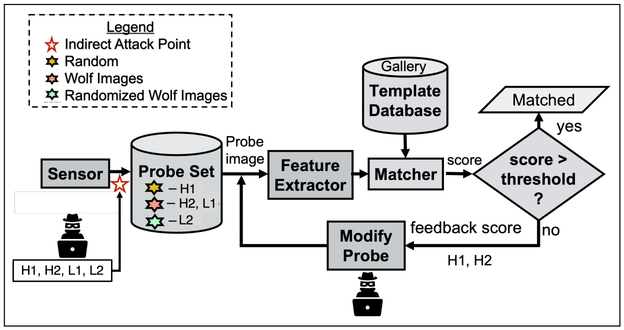


Figure 1. Block diagram of attack points and methods. Here, morphing refers to classical morphing as well as deep learning-based.

In this paper, we demonstrate how an attacker can use a less explored field of soft-biometric, i.e., *kinship properties or cues*, to bypass the requirement of knowledge about the facial biometric of the registered person to create a range of templates called 'wolfs' capable of attacking the system. Out of the seven post-sensor (indirect) points mentioned above, our proposed attack occurs just after the sensor in the communication channel between sensor and feature extractor, as shown in Figure 1. The main advantage of this attack is the attacker can perform it with high success rate compared to using a random image, without even using the biometric of the actual person. Our approach employs traditional morphing as well as a deep learning (DL) based child image generation approach from literature to generate the so-called wolf images. Also, while genetics play a crucial role in determining facial features, environmental factors can also have an impact. For example, nutrition and health can influence how certain facial features develop [2]. This environment factor is random and thus leads us towards also mixing the wolf images with different proportions of randomness.

Note that biometric modalities like iris or fingerprint may be more secure, as the registered person's biometric is not available online, unlike faces and is difficult for the attacker to gain access to that. However, since it is impossible to investigate our proposed attack against them due to scarcity of standardized kinship data for these modalities, our only option is to explore the wolf attack in the (*admittedly less realistic*) face domain where such datasets are available to researchers. The main objective of this paper is to provide initial exploration to show how kinship may be a useful soft-biometric property that increases the vulnerability of biometric systems and demands more attention. Our hope is that this research motivates the collection of kin datasets in the iris and fingerprint domains, thereby enabling researchers to explore if their associated systems are also vulnerable to kinship-based wolf attacks and, if so, how they can be protected. In addition, other than as a wolf attack, the approaches investigated in this paper for generation of synthetic images[1] may have other applications in forensics, finding lost children, etc.

To the best of our knowledge, use of kinship cues in attacking a face recognition system has never been studied before. Our major contributions are summarized as follows:

- Explore a post-sensor indirect attack with synthetic wolf images generated from kin instead of the registered person's biometric.
- Re-implement state-of-the-art Bayesian adaptation hill climbing attacks [3, 4] with and without kinship (proposed), and demonstrate how kinship aids in convergence. We also achieve a 4.3× and 1.4× average improvement in attack accuracy after 100 iterations for morphing and DL-based wolf image generation, respectively.
- Propose a novel kinship-based wolf image generation / attack framework that operates without using any feedback score, complicated training, or knowledge about the biometric of the authorized person. It is evaluated on the two largest facial kinship datasets, Family101 and Family-in-the-Wild.
- Investigate tradeoffs in attack accuracy as randomness in wolf images is varied. We find that there is an optimal point for the DL-generated wolves that improves attack accuracy by 2.2× on average.

The rest of this paper is organized as follows. Section 2 briefly discusses the related work while Section 3 provides background on kinship and different tools used in this paper. In Section 4, detailed mechanisms used to generate the wolf images are proposed and an overview of the algorithms used for demonstrating the attacks using the wolf images are explained. The experimental setup, results, and discussions are presented in Section 5. This paper is concluded with directions for future research in Section 6.

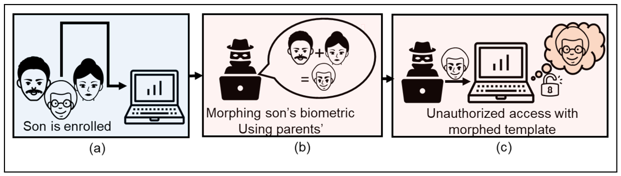


Figure 2. Use of kinship cues in attacking a face biometric system.

2. Related Work

Indirect synthetic attacks using kinship cues, without using knowledge about the person under attack and without using feedback score, as proposed in the paper, are a new domain of attack. A previous work in a similar domain is [5], discusses indirect synthetic attack on thermal face biometric system where the attack point is same as ours. Before this, the closest research performed in this domain includes iteratively attacking a system with random images and modifying the image based on feedback score, a method popularly known as hill climbing [3, 4]. This type of attack also needs to be performed post-sensor to bypass the liveliness detector. Yet another alternative involves creating morphed faces of random people to match multiple faces [6], a method popularly known as wolf attack, where the morphed images are called 'Wolfs'. Most of the work related to wolf attacks has been performed in the domains of signature and fingerprint. The existing hill climbing approaches on face recognition system can be found in [3, 4, 7–9]. The only existing iterative wolf attack method [10] uses feedback score to modify the input image.

3. Background

3.1. Kinship

Kinship cues are properties that convey resemblance between human faces among family members which help in recognizing the relatedness between different people. Hence, these traits, if identified accurately, can be used to *mimic or spoof a person's face with only the knowledge of their relative's faces.* The high-level scenario of our kinship attack is shown in Figure 2. These visual capacities to detect relatedness become weaker with lower degrees of relatedness. Hence only the parents' images are considered because they contain the maximum degree of relatedness [11]. The high-level scenario of our kinship attack is shown in Figure 2. A detailed investigation on kin relationship is provided in [11], which also states possible effects of age and gender on kinship cues. Another paper com-

[1] Throughout the paper, the terms 'synthetic image', 'fake child' and 'wolf image' are all used to represent the synthetically generated images.

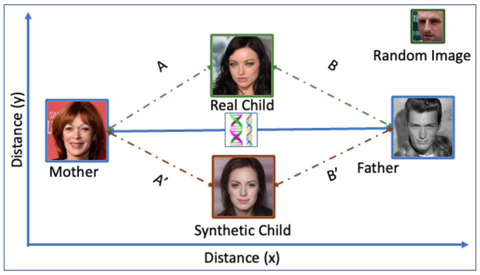


Figure 3. The synthetic child image (wolf) is closer to the parents but still away from the real child. It can be anywhere between parents. Random image can be at any random distance from victims.

pares the accuracy of kinship recognition from face images between humans and computers [12]. The results show that kinship recognition by computers is as precise as humans. Random environment factors also play a role in determining human facial cues. Thus, in one of our experiments, we investigate how synthetically generated images of authorized users (children) from parents remain closer to parents but still away from the real user (child) image, as illustrated in Figure 3 and later in Table 1. This is because without randomness the synthetic child image is generated exactly between the parent images, whereas the actual child image may remain at a location away from the straight-line distance between the parents in euclidean space due to inclusion of environmental randomness.

3.2. Face vs. Kin Recognition Systems

As shown in Figure 1, a biometric face recognition system has five main components. The sensor records a biometric trait, which is a face in this case. This recorded face image, also called a *probe*, is then forwarded to the feature extractor. The feature extractor uses some advanced algorithms to extract facial features and forms a template using these features. This template is then passed to the matcher. The system contains similar templates of registered users in the database. These templates are created using the same feature extractor during registration, also referred to as a *gallery*. The matcher matches the templates obtained from the database and authentication system's feature extractor and outputs a score. Based on the similarity threshold set for the system, an output decision is generated by the decision maker. A *genuine score* is the score between the probe and the gallery of the same class.

In case of our kin-wolf attack, the probe is the synthetic image, attack-genuine score is the euclidean distance between the probe and corresponding gallery child image, X. On the other hand, an *attack-imposter score* is the distance between the same probe and the gallery of different class. In our case, it is synthetic image and child Y, where X and Y are child images of different families.

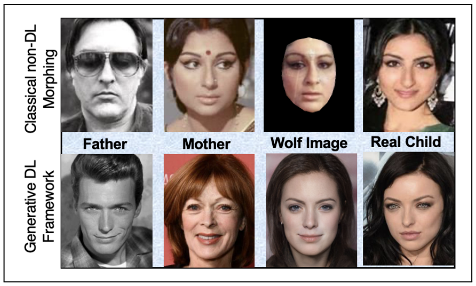


Figure 4. Example morphed wolf (synthetic child) images generated using non-deep learning based classical morphing (top row) and generative deep learning framework (bottom row).

3.3. Types of Attacks Explored

The vulnerability of face recognition systems at each of the components and the links between them has raised many concerns. Data can be manipulated at each of the eight sub-parts of the system to falsely get access to the system. Presentation attacks mainly occur in the sensor. Similar to the presentation attack, another type of attack is the indirect synthetic attack, which takes place just after the sensor, bypassing the liveness detector. In this paper, we have focused on such post-sensor indirect synthetic attacks. Effect of kinship is explored in two types of such attacks: hill climbing and wolf attacks. Background for each of these approaches and their purposes are explained in this section.

3.3.1 Hill-Climbing Attacks

This attack is used in this paper to demonstrate the power of kinship in an existing attack scenario and how it helps in converging the attack faster. In this attack, the score returned by the matcher can be considered as one major vulnerability of a biometric system. If an attacker gets hold of this score, then without having any knowledge about the gallery images an attacker can break into the system by manipulating a random image iteratively till expected score is achieved. There are several ways in which mutation of the random input value is executed. The state-of-the-art Bayesian adaptation hill climbing method [13], that we reimplemented in this paper, uses an input distribution for the same purpose. This method uses feedback score to modify the random input distribution. Here, a global distribution (G) is formed from all the input images. This distribution is matched with the latent distribution created from a subset of probe set, which is a set random images in the first iteration, whose distance from the attacked gallery is minimum. Depending on the matched score, G is modified into an adaptive distribution (A) at each iteration till a required match score threshold or maximum number of iterations is reached. For more details, see [13].

3.3.2 Wolf Attacks

Unlike hill climbing, a wolf attack does not modify the input template iteratively. Rather, it tries to attack a system using a dictionary of templates. Thus, it is more focused on wolf data creation, instead of manipulating a single template iteratively. The only existing wolf attack method in the realm of faces in the literature, discussed in [10], uses feedback score to generate the dictionary of random images. While this approach explores the StyleGAN space, our approach explores the kinship space, and instead of searching for images from the entire space, it uses kin features to locate the region of interest.

4. Proposed Kin-Wolf Methodology

In the kin-wolf attack, image generation plays a pivotal role. In this section, a detailed overview and algorithms used to generate such wolf images is given. These wolf images not only help convergence speed of the state-of-the-art hill climbing approaches but also help in indirect synthetic attacks without using any information or feedback of the registered person. No new models are trained for this wolf image creation technique. We have used a non-deep learning (non-DL) based classical morphing algorithm and a pre-trained state-of-the-art child image generation model as a deep learning (DL) based approach. Details of each generation technique are provided in Sections 4.1 and **??**.

4.1. nDL Method: Kin Morphing

The classical image processing based nDL technique that is often used to combine the cues from two images is morphing [14]. In our work we have used this technique to combine kinship cues from both the parents. In morphing, the facial features are blended in different portions to get an innumerable number of morphed faces. The morphing algorithm we use has four phases. In the first phase, the landmark points are detected in both the source and the destination faces. Then both the faces are aligned based on these points. In the third phase, given two images and their face points, one image is warped into the other and the algorithm triangulates face points. Then the affine transformation of each triangle is performed with bilinear interpolation. Finally, a series of morphed images are produced in which the transformation is performed at different percentages (determined by $0 \leq M, N \leq 100$). An example morphed wolf image with kinship properties is shown in Figure 4.

4.2. DL Method: Child Image Generator

A recent work [15] proposes a child image generation framework, ChildNet, to predict the child image from two parents by leveraging a state-of-the-art GAN for higher quality synthesis. It is based on a novel latent-space manipulation scheme designed around the StyleGAN2 model [16], inspired from the actual gene mixing and leads to convincing child synthesis results with photo-realistic, high-resolution, and artifact-free synthesized images. The model provides several synthesis control mechanisms that involve age and gender manipulation, image variability control, and determination of the dominant parental image. We have leveraged these mechanisms to generate synthetic images with different proportions of each parent dominance, age, and gender to increase the chances of the attack. More details can be found in [15].

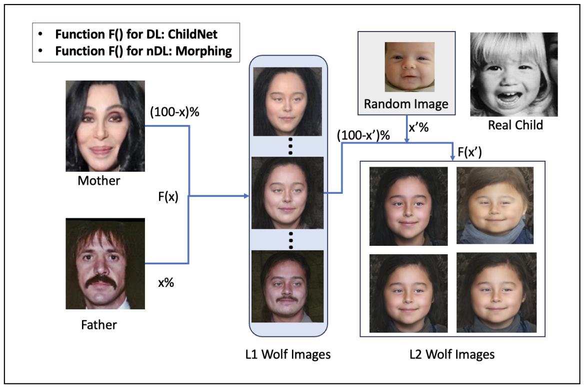


Figure 5. Probe set creation for Level 1 (L1) and Level 2 (L2) from parents' images to attack the registered child (i.e., victim).

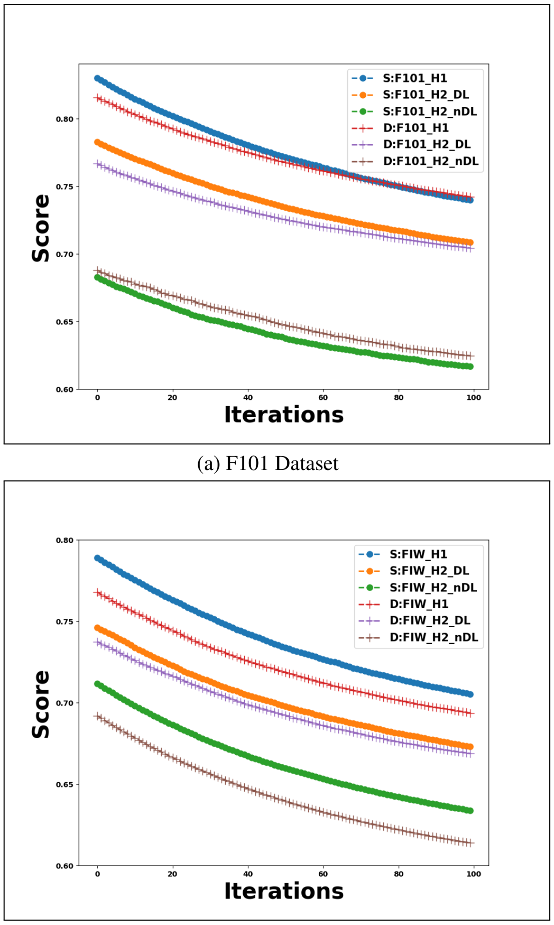
4.3. Kin-Wolf Probe Set Generation

Using the algorithms provided in Sections 4.1 and **??**, two sets of wolf image datasets are generated and referred to as 'L1 (Morph)' or 'L1 (nDL)' and 'L1 (ChildNet)' or 'L1 (DL)'. These sets are generated by combining the parents' images in different proportions using the kin morphing and ChildNet frameworks, respectively. For L1 Morph, the only variable parameter is the proportion of parents. The DL method using ChildNet provides more parameters soft biometric features to explore, like age and gender. The synthetic wolf or child image created for each pair of parents are generated for both male and female gender as well as for five age ranges. The five age ranges are {7–9}, {15–19}, {30–39}, and {50–69}.

Each of these L1 images are further combined, using the same framework, with different proportions of random image to generate 'L2 (nDL)' and 'L2 (DL)' datasets. This integration is done with a set of 100 random images for each L1 image. These random images are taken from a different disjoint dataset than the family dataset. For L2 images no added parameters are provided for the age and gender. Figure 5 shows the different levels of morphing.

5. Experimental Results and Discussion

In this paper, the kin-wolf images are generated using two datasets, namely, Family101 (F101) [17] and Family-



(a) F101 Dataset

(b) FIW Dataset

Figure 6. Change in distance score between input probe and the destination image (D: Daughter, S: Son) at every iteration for H1 and H2 (generated using DL and nDL methods).

Table 1. EERs for each best pair of probe and gallery for FIW dataset. Ideally, EER with L1 images as probes should be close to RC vs. RC and always less than RC vs. parents. The best pair is chosen across age and gender.

Probe	**Gallery**		
	Father	**Mother**	**Real Child (RC)**
Real Child (RC)	0.34	0.35	0.13
nDL Level 1 (L1) Wolf	0.26	0.24	0.20
Dl Level 1 (L1) Wolf	0.30	0.31	0.25

in-the-Wild (FIW) [18]. From these datasets, the parents and child images (having both parents) are separated and used in our experiments. For example, the separated daughter set contains images of daughters along with the respective parents' images. 27 such triplet families in total are extracted from Family101 dataset and 1,468 triplet families from FIW dataset. The son set contains images of sons along with their respective parents. 32 such triplet families in total are extracted from the Family101 dataset and triplet 1,609 families from FIW dataset.

Two sets of wolf images are generated by combing the parents' images of each dataset, as discussed in Section 4.3. A total of four sets of L1 wolf image datasets are generated, two for each family datasets, F101 and FIW. They are referred as 'F101_L1_DL', 'F101_L1_nDL', 'FIW_L1_DL', and 'FIW_L1_nDL', where DL and nDL symbolizes the frameworks used for generation and F101 and FIW symbolizes the dataset from which it is generated. Each of these sets has daughter and son subsets. The pre-trained Childnet model used in this paper is pre-trained with FIW dataset.

The four sets of L1 images are further combined with another set of 100 random images as described in Section 4.3. The random set of 100 selected images, used for level 2 (L2) morphed probe image generation, are taken from Kaggle Facial Age Dataset [19]. Like L1, four sets of L2 images are created which is referred as 'F101_L2_DL', 'F101_L2_nDL', 'FIW_L2_DL', and 'FIW_L2_nDL' based on the framework and kin dataset used in generation. In this paper, we utilize dlib pretrained model used in face recognition [20]. This model has an accuracy of 99.38% on the Labeled Faces in the Wild benchmark[21].It learns a mapping from face images to Euclidean space where distances directly correspond to a measure of face similarity. Another feature extractor, kinfacenet [22], is used to study the kinship content in the generated wolf images.

5.1. Kin-Wolf Probe Set in Hill Climbing Attack

As described in Section 3.3.1, the hill climbing method re-implemented in this paper uses Bayesian adaptation [13]. We have used two types of images as starting points. The first is explained in the original paper, a set of random images, and we refer to it as H1. The second attack is made with our proposed kinship cues wolf (L1) images and is referred as H2. The iteration vs. mean score is noted. In each of these techniques, a gallery set of 20 real child images are attacked using either random images, as in H1, or L1 images of the respective family, as in H2, as starting seed. Use of L1 image, created by combining parents of the attacked child, is the proposed modification. Figure 6 shows how the mean distance score across all the families changes with an increase in the number of iterations. For each and every dataset and its subsets, Daughter (D) and Son (S), a common trend is observed. The fake encoding is approaching towards the genuine encoding. Yet another observation is that the start image has notable impact in achieving the threshold distance faster. If the probe contains kinship properties (H2), then the starting score itself is better than starting with a random image. For the daughter (D) and son (S) subsets of both F101 and FIW, if a system has a threshold at 0.62 and the start seed is taken from nDL sets of kin generated L1 dataset, referred as F101_H2_nDL and FIW_H2_nDL, the system has considerable number of successful attacks after 100 iterations, as the average score falls below the threshold. Although hill-climbing attacks provide high accuracy, most state-of-the-art face recognition infrastructure does not provide feedback scores for safety and security of the biometric system. These experiments provide the initial justification for why kinship can be used in wolf attacks, and demonstrates our second contribution as listed in Section 1.

5.2. Kin-Wolf Distance Evaluation

Now that the impact of kinship has been established for hill climbing, the L1 kin images, or the sets used in H2, are further explored in a dictionary or wolf attack framework. The kin-wolf attack framework proposes a method to generate a dictionary of wolf synthetic child images which has potential to attack the real authorized child. It is independent of any modification in the template, or the face recognition system used. As described in Section 4.3, two types of sets are generated. One set, L1, is generated with only kin information from the parent images. This L1 set is further expanded to L2 set by injecting different proportion of randomness. This is done with the hypothesis that human facial features are created not only from the inherited genes, but environmental factors and randomness plays a pivotal role. The purpose of these experiments is also to show how the genuine scores are improved when kinship features are injected in random images. When no biometric information about the authorized person is available, the attacker can utilize the kinship cues to make the attacking probes better than just using random images. Once these two sets, L1 and L2, are obtained, experiments on efficiency of each of these dictionaries of wolf images, are performed and the genuine and imposter score distributions, attack accuracies using these images as probe and real child images as gallery are studied. The Equal Error Rate (EER) of the system is noted for different probe sets, as given in Table 1. The minimum EER is achieved when the face recognition system is queried using the image of the real child. The distance threshold obtained at this optimal point is **0.629** (shown as red dotted line in Figures 7–9b). As the face recognition system is not modified, the attack accuracy is computed based on this threshold later. If a fake probe achieves score less than or equal to this, it successfully attacks the system. Also, this table provides insight about the distributions of the genuine and imposter scores for each of the probe set. It clearly shows that the distributions are more separated for L1 and real child than real child images and the parents.

5.2.1 Experimental Results for L1 Wolf Images

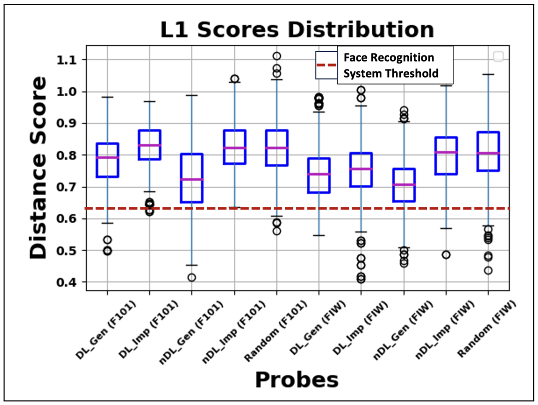


Figure 7. All L1 scores distribution. All the genuine scores are evaluated without using the authorized gallery entity as the probe. Kinship can be used as a soft biometric as it can generate synthetic probes which are closer to the actual probes. ’Gen’ represents genuine and ’Imp’ represents imposter.

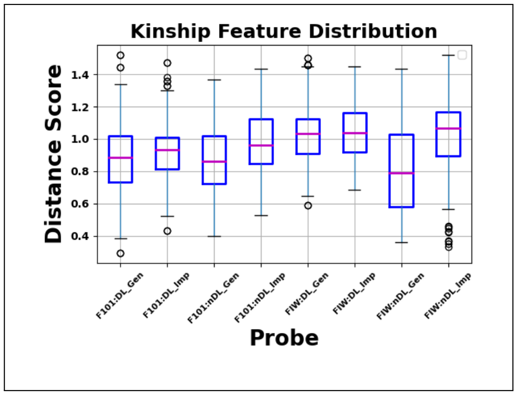
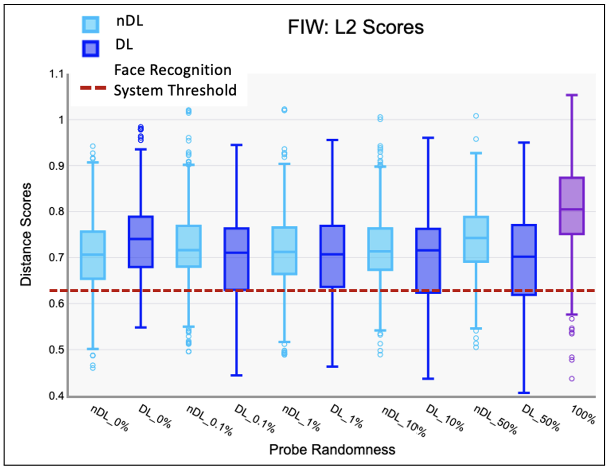
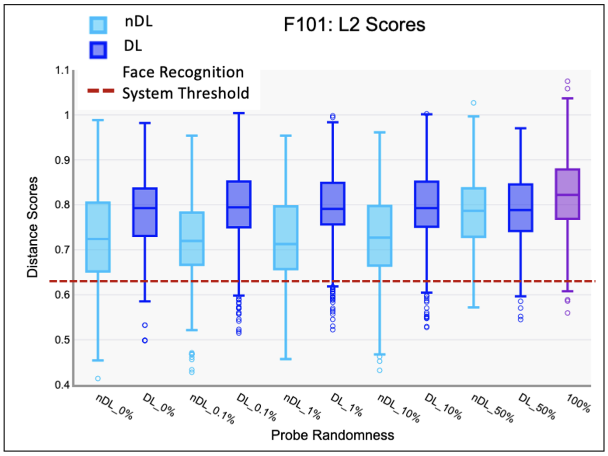


Figure 8. Kinfacenet kinship feature [22] distance score comparison between real child and synthetic probe.

The genuine and imposter distribution for each of the four L1 wolf sets are analyzed. All the score distributions are given in Figure 7 where the random FIW and random F101 column scores are created by taking random images as probe and real child images from respective datasets as gallery. The other columns ’F101:DL_Gen’, ’F101:DL_Imp’ belong to the attack genuine (gen) and attack imposter (imp) scores of F101 for DL generated wolfs. Scores of FIW dataset for DL generated wolfs are represented by columns ’FIW:DL_Gen’, ’FIW:DL_Imp’. In a similar naming sequence the columns of nDL are also represented. This analysis shows that while all the imposter score distributions are near the random score distribution, the genuine scores are much better. This clearly shows that kinship based wolf images are better than random images as wolf images. This set of experiments demonstrate our second contribution listed in Section 1. Another interesting observation is that the classical kin morphing method creates better wolf images than the deep learning based approach. One reason for that is, the wolf images generated using the classical morphing contains more information from the parents than the deep learning framework using ChildNet. We have tested the kinship content in the images for DL and nDL using the KinFaceNet framework from [22]. The score distribution is given in Figure 8. It clearly shows, the kinship



(a) FIW



(b) F101

Figure 9. L2 distance score for DL and nDL generated wolfs for different randomness.

distance scores between real child and DL generated child features for both FIW and F101 are overlapping. Although ChildNet produces high quality images of synthetic child, there is still plenty of scope to learn the kin features better. It clearly demonstrates, while using the pre-trained StyleGAN network embedded in the framework, it is including a lot of randomness along with the kin features. That is why, although the images are better than random images, there is still scope to improve. Nevertheless, the main advantage of using ChildNet is that it provides the flexibility to change the age and gender of the synthetic image.

5.2.2 Experimental Results for L2 Wolf Images

As discussed in Section 3, studies have shown that, while genetics play a crucial role in determining facial features, environmental factors can also have an impact. This environment factor is the randomness. In one of our experiments, it is demonstrated how the synthetic child generated is closer to the parents but still away from the real child image, as shown in Figure 3 and Table 1. That proves randomness plays an important role and it is introduced in L2 images by mixing random images to the L1 set.

Table 2. False Acceptance Rate (FMR) at Th=0.629. Genuine distribution represents score distribution, where score represents distance between two real child images. False Reject Rate (FNMR) is 0.13, which represents rejection rate at the specific threshold.

Imposter Distribution	FMR_Threshold
nDL Method vs Real Child	0.19
DL Method vs Real Child	0.10
Single Parent vs Real Child	0.11
Random Image vs Real Child	0.08

The amount of randomness chosen are $\{0.1, 1, 10, 50$ and $99.9\}\%$ of each of the 100 random images. For each L1 images a total of 500 L2 images are generated, where each L1 image is mixed with a random image in five proportions and the total number of random images used is 100. As a result, for N families, and M L1 probe images, the total number of images in the L2 dataset is $M \times N \times 500$. Due to its' large size, sampling of images is done to generate genuine and imposter scores. In all the experiments, the attack genuine score is referred as the distance between a selected probe set and real image of the same family and attack imposter score is represented as the distance between the same probes and gallery image of different family.

The genuine score distribution of each of the five proportions of randomized probes for FIW and F101 datasets are provided in Figures 9a and 9b. In each figure, the last score represents the distance score distribution of random images as probe, without mixing any kinship information. This clearly shows that the match score distribution of the random images is improving after integrating the kinship information to it. The trend is more visible for the FIW dataset for DL method. One reason for this can be the pre-trained Childnet model used is pre-trained with FIW dataset. This also shows, in future, this attack can be improved by a dataset independent DL child image generation method. The random image set used in creating the L2 images is same as the random dataset used to get the random score distribution. This set of experiments, address our third contribution stated in Section 1.

5.3. Kin-Wolf Attack Evaluation

In Table 2, FMR and FNMR is reported for different attack-genuines at the threshold point of the system. It clearly shows that chances of falsely acceptance of morphed images as genuine is maximum and almost 2.5X times better than using random images, as is used in hill climbing type of attacks. Although the deep learning based images provide better scope to change other soft-biometrics like age and gender there is lot of scope for improvement. Another

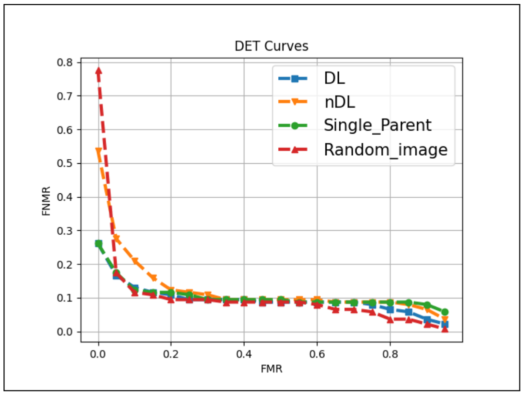


Figure 10. DET curves for score distribution with different attack probes (referred as imposter) and real vs real child images as genuine score distribution. The goal is to make imposters get falsely accepted as genuine in the system.

Table 3. Verification Accuracy (VA) from ROC and Attack Accuracy (AA) for different probe sets.

Accuracies	**Iterations/ Sample Size**	**nDL**		**DL**	
		F101	**FIW**	**F101**	**FIW**
VA L1 Wolf Images	-	77.16	77.09	66.03	56.87
VA L2 Wolf Images	-	56.34	50	50	49.9
AA H1(start:random seed)	100/100	8.86	22.46	8.86	22.46
AA (L1)	0/100	18.02	18.71	4.91	10.53
AA H2(start:L1)	100/100	58.65	50.47	11.4	34.73
Best AA (L2) (randomness=r)	0/100	14.90 (r=5%)	15.30 (r=30%)	8.10 (r=5%)	28.80 (r=30%)
AA (Random Image)	0/100	1.71	3.87	1.71	3.87

interesting observation is that, even without any modification, there is good chance of acceptance of a related person without even undergoing any change. These results are also shown using a DET curve in Figure 10, where a similar pattern is observed at the EER point. DET curves also show that nDL images are of the best type with higher FMR. For Table 2, the genuine score selected is the actual genuine score of the system, i.e. distance between the same person, and the distribution under inspection is the synthetic child distance with the real child(considered as imposter score of the system here).

In Table 3, two types of accuracies are reported, verification accuracy (VA) and attack accuracy (AA). VA is the 1:1 comparison where a fake probe generated using parents' images, is compared with the gallery image of the same family, and the score is referred as genuine score. If the distance is less than the threshold, then the probe is verified. This threshold is usually the distance score at EER in the genuine-imposter distribution, where the false acceptance rate is equal to the false rejection rate. However, for a face recognition system, this threshold is computed from the genuine imposter score distribution of the real probe and gallery images, where the genuine score is the distance between the real probe image of the gallery image, and not the kin generated probe. AA is computed based on this threshold of the face recognition system. For our case, the EER is computed at 0.629 so an attack is successful if the distance between the fake probe and gallery is less than this threshold. AA is the percentage of successful attacks when a dictionary of 100 random probes is used to attack the corresponding real child gallery images. These 100 random probes, L1 or L2, are independent of any information of the gallery images. The main observations from Table 3 are as follows:

- High verification accuracy proves the genuine and imposter scores are well separated, if the threshold is chosen at the value where the EER is achieved.
- All the attack methods, i.e., Bayesian hill climbing (H1), proposed kinship modified hill climbing (H2) and kin-wolf methods are better than trying to attack the system with random images.
- From AA H1, AA H2, and AA L1, it is observed that our proposed modifications to existing hill climbing attack (H2) and kin-wolf (L1) outperform the existing hill climbing approaches starting from random image as seed (H1).
- AA L2 shows that at certain percentage of randomness injected L1 images for each dataset, i.e., 30% random image injected with L1 FIW and 5% random image injected with L1 F101 dataset, improves the attack accuracy, when compared to the AA Random Image. For DL, AA (L1) is also outperformed.
- Although the results are better for modified hill climbing, it is difficult to extract the feedback score from modern face recognition systems. Kin-wolf achieves comparable attack accuracy without feedback.
- Results are better for nDL methods than DL methods. This is due to lack of enough kin features in DL methods as shown in Figure 8. Nevertheless, it is still promising, and in future we plan to improve the DL method since it more easily integrates other soft biometrics (e.g., age, gender, etc.).

6. Conclusion and Future Work

In this paper, we demonstrated how kinship cues have the potential to create synthetic images which are closer to authorized system user images, thus increasing the chances of a successful indirect synthetic attack, without using feedback score or knowledge about gallery images. Although we have not proposed any synthetic child image generation method, we have re-implemented two state-of-the-art techniques and observed that although deep learning has the potential to integrate other soft biometrics like age and gender, kinship cues are transferred better by classical morphing. In future, we aim to create a better child image generation framework for a better attack as well as useful for other domains of applications like image forensics.

References

[1] André Anjos, Jukka Komulainen, Sébastien Marcel, Abdenour Hadid, and Matti Pietikäinen. Face anti-spoofing: Visual approach. In *Handbook of biometric anti-spoofing*, pages 65–82. Springer, 2014.

[2] Yong Zhi Foo, Leigh W Simmons, and Gillian Rhodes. Predictors of facial attractiveness and health in humans. *Scientific reports*, 7(1):39731, 2017.

[3] Javier Galbally, Julian Fierrez, Javier Ortega-Garcia, Chris McCool, and Sebastien Marcel. Hill-climbing attack to an eigenface-based face verification system. In *2009 First IEEE International Conference on Biometrics, Identity and Security (BIdS)*, pages 1–6. IEEE, 2009.

[4] Javier Galbally, Chris McCool, Julian Fierrez, Sebastien Marcel, and Javier Ortega-Garcia. On the vulnerability of face verification systems to hill-climbing attacks. *Pattern Recognition*, 43(3):1027–1038, 2010.

[5] Khawla Mallat and Jean-Luc Dugelay. Indirect synthetic attack on thermal face biometric systems via visible-to-thermal spectrum conversion. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1435–1443, 2021.

[6] Ulrich Scherhag, Christian Rathgeb, Johannes Merkle, Ralph Breithaupt, and Christoph Busch. Face recognition systems under morphing attacks: A survey. *IEEE Access*, 7:23012–23026, 2019.

[7] Andy Adler. Vulnerabilities in biometric encryption systems. In *International Conference on Audio-and Video-Based Biometric Person Authentication*, pages 1100–1109. Springer, 2005.

[8] GB Marshalko and LO Nikiforova. Spoofing attack on eigenfaces-based biometric identification system. *Automatic Control and Computer Sciences*, 53(8):980–986, 2019.

[9] M. Gomez-Barrero, J. Galbally, J. Fierrez, and J. Ortega-Garcia. Face verification put to test: A hill-climbing attack based on the uphill-simplex algorithm. In *2012 5th IAPR International Conference on Biometrics (ICB)*, pages 40–45, 2012.

[10] Huy H. Nguyen, Junichi Yamagishi, Isao Echizen, and Sébastien Marcel. Generating master faces for use in performing wolf attacks on face recognition systems, 2020.

[11] Gwenaël Kaminski, Slimane Dridi, Christian Graff, and Edouard Gentaz. Human ability to detect kinship in strangers' faces: effects of the degree of relatedness. *Proceedings of the Royal Society B: Biological Sciences*, 276(1670):3193–3200, 2009.

[12] Miguel Bordallo Lopez, Abdenour Hadid, Elhocine Boutellaa, Jorge Goncalves, Vassilis Kostakos, and Simo Hosio. Kinship verification from facial images and videos: human versus machine. *Machine Vision and Applications*, 29(5):873–890, 2018.

[13] Javier Galbally, Julian Fierrez, and Javier Ortega-Garcia. Bayesian hill-climbing attack and its application to signature verification. In *International Conference on Biometrics*, pages 386–395. Springer, 2007.

[14] Mark Steyvers. Morphing techniques for manipulating face images. *Behavior Research Methods, Instruments, & Computers*, 31(2):359–369, 1999.

[15] Martin Pernuš, Mansi Bhatnagar, Badr Samad, Divyanshu Singh, Peter Peer, Vitomir štruc, and Simon Dobrišek. Childnet: Structural kinship face synthesis model with appearance control mechanisms. *IEEE Access*, 11:49971–49991, 2023.

[16] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8110–8119, 2020.

[17] Ruogu Fang, Andrew C Gallagher, Tsuhan Chen, and Alexander Loui. Kinship classification by modeling facial feature heredity. In *2013 IEEE International Conference on Image Processing*, pages 2983–2987. IEEE, 2013.

[18] Joseph P Robinson, Ming Shao, Yue Wu, Hongfu Liu, Timothy Gillis, and Yun Fu. Visual kinship recognition of families in the wild. 2018.

[19] Kaggle. Facial age dataset.

[20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[21] Erik Learned-Miller, Gary B Huang, Aruni RoyChowdhury, Haoxiang Li, and Gang Hua. Labeled faces in the wild: A survey. *Advances in face detection and facial image analysis*, pages 189–248, 2016.

[22] Pallabi Ghosh, Sumaiya Shomaji, Damon L. Woodard, and Domenic Forte. Kinfacenet: A new deep transfer learning based kinship feature extraction framework. In *2023 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10, 2023.