

Differential Morph Face Detection using Discriminative Wavelet Sub-bands

Baaria Chaudhary*, Poorya Aghdaie*, Sobhan Soleymani, Jeremy Dawson, Nasser M. Nasrabadi
West Virginia University

{bac0062, pa0002, sssoleyma}@mix.wvu.edu, {jeremy.dawson, nasser.nasrabadi}@mail.wvu.edu

Abstract

Face recognition systems are extremely vulnerable to morphing attacks, in which a morphed facial reference image can be successfully verified as two or more distinct identities. In this paper, we propose a morph attack detection algorithm that leverages an undecimated 2D Discrete Wavelet Transform (DWT) for identifying morphed face images. The core of our framework is that artifacts resulting from the morphing process that are not discernible in the image domain can be more easily identified in the spatial frequency domain. A discriminative wavelet sub-band can accentuate the disparity between a real and a morphed image. To this end, multi-level DWT is applied to all images, yielding 48 mid and high-frequency sub-bands each. The entropy distributions for each sub-band are calculated separately for both bona fide and morph images. For some of the sub-bands, there is a marked difference between the entropy of the sub-band in a bona fide image and the identical sub-band's entropy in a morphed image. Consequently, we employ Kullback-Liebler Divergence (KLD) to exploit these differences and isolate the sub-bands that are the most discriminative. We measure how discriminative a sub-band is by its KLD value and the 22 sub-bands with the highest KLD values are chosen for network training. Then, we train a deep Siamese neural network using these 22 selected sub-bands for differential morph attack detection. We examine the efficacy of discriminative wavelet sub-bands for morph attack detection and show that a deep neural network trained on these sub-bands can accurately identify morph imagery.

1. Introduction

Face recognition systems are increasingly replacing human inspectors in border control and other security applications. Face capture is non-invasive, can be performed at a distance, and benefits from a relatively high social acceptance. Furthermore, face recognition systems also have a

natural safeguard: if the algorithm triggers a false alarm, a human expert on site can easily perform the verification. For these reasons, the International Civil Aviation Organization (ICAO) has mandated the inclusion of a facial reference image in all passports worldwide [16]. Still, all the characteristics that make face recognition systems appealing also make them vulnerable. The mass adoption of automatic biometric systems for border control has exposed the inability of these systems to reliably detect a falsified image in a passport, particularly a morphed image, which has been identified as a serious threat. A morph attack is when a single morphed face image can be positively verified as two or more distinct identities [12]. This type of attack requires no complex forgery of passport technology, but rather a simple manipulation of the passport photo at time of application. Many morphing applications are easily and freely accessible and have no knowledge barrier [25]. It follows that a criminal attacker, who otherwise cannot travel freely, could obtain a legitimate travel document by morphing his face with that of an accomplice with similar features, resulting in existing face recognition systems verifying the morph image as either of the two individuals.

Morphed images are not visually perceptible to the human eye, which makes them especially difficult to detect. As characteristics of both subjects are taken into account when morphing an image, face recognition systems are easily deceived and, as such, the false acceptance rate is very high. In addition, these systems are designed to tolerate a large intra-class variance to account for the significant changes in facial appearance that occur in the 5 to 10-year life cycle of a passport. Many commercial off-the-shelf (COTS) systems have repeatedly failed to detect morphed images [34]. Similarly, studies also show human recognizers are unable to correctly differentiate between a morphed image and an authentic one [34] [30] [5]. Even after instruction on how to detect a morphed image, human recognizers still perform worse than face recognition algorithms [24] [18].

In this paper, we propose a differential morph attack detection algorithm using an undecimated 2D Discrete Wavelet Transform (DWT). By decomposing an image to

*Authors contributed equally.

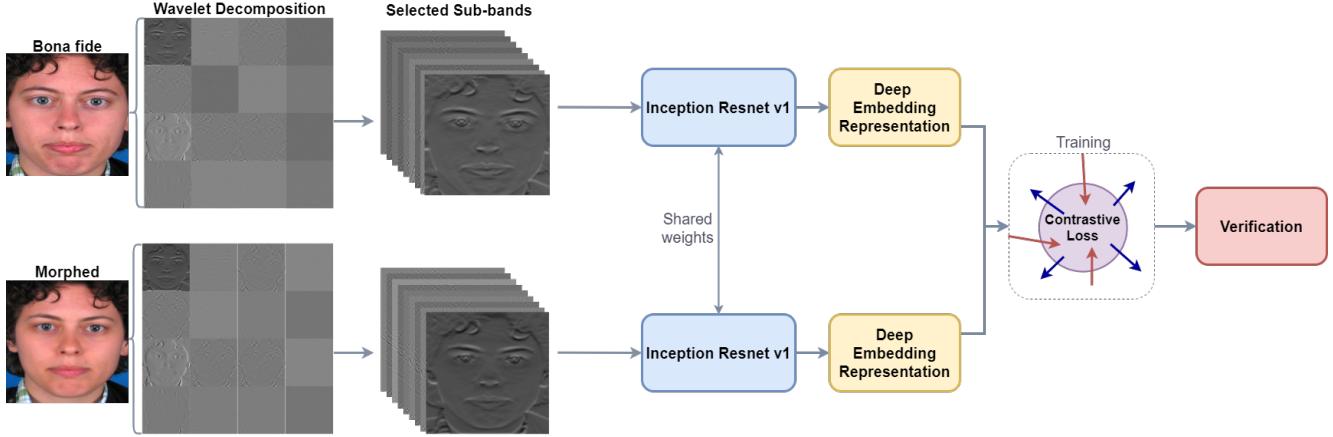


Figure 1. **Network architecture.** Image pairs are decomposed into wavelet sub-bands, only selecting the discriminative wavelet sub-bands for network training. The selected sub-bands are then sent to the Inception ResNET v1, where contrastive loss is applied during training. From the feature embedding representation, verification is performed by computing the L_2 distance between the feature vectors and a decision score is produced.

wavelet sub-bands, we can identify the morph artifacts that are hidden in the image domain more clearly in the spatial frequency domain. Our analysis of wavelet sub-bands demonstrates that specific high-frequency components are more discriminative for morph attack detection. To this end, multi-level DWT is applied to the images, yielding 48 mid- and high-frequency sub-bands each and discarding the low-frequency bands altogether. To isolate the most informative sub-bands, we employ Kullback-Liebler Divergence (KLD) [19] on the sub-band entropy distributions. The higher the KLD value, the more discriminative the sub-band is for morph detection. We then use the selected informative sub-bands to train a deep Siamese network for the differential morph attack detection scenario. The Siamese network takes bona fide and morph pairs as input and yields a confidence score on the likelihood of the pairs being from the same person as shown in Figure 1. Siamese networks are ideal for this scheme as they are primarily employed in tasks that require finding similarities between two inputs. We examine the usefulness of wavelet sub-bands for differential morph attack detection and show a deep neural network trained on these sub-bands can accurately identify morph imagery. The experiments are conducted on three different morph image datasets: VISAPP [23], MorGAN [8] [9], and LMA [8]. Additionally, cross-dataset performance is evaluated on AMLS [27].

The paper is organized as follows: we briefly summarize the related works in Section 2, explain the methodology in the Section 3, and discuss our experiments and subsequent results in Section 4. Finally, conclusions are presented in Section 5.

2. Related Work

The vulnerability of face recognition systems to morph attacks was first introduced by [12]. Many morph detection algorithms have been proposed since then for both single (no reference) and differential morph attack detection scenarios. Single (no reference) morph attack detection algorithms rely only on the potential morphed image to make their classification. Conversely, differential morph attack detection algorithms compare the potential morphed image with an additional trusted image, typically a live capture at border security. As such, differential morph attack detection algorithms have more information at their disposal to make their classification and, therefore, generally perform better than single morph detection algorithms [37].

With respect to single morph attack detection, many classical hand-crafted feature extraction techniques have been explored. The most well-performing of these general image descriptors is Binarized Statistical Image Features (BSIF) [17], in which extracted BSIF features were classified using a Support Vector Machine (SVM) [31]. However, deep learning methods consistently perform better than these general feature extraction techniques [40] [48] [32]. In [32], complementary features from pre-trained VGG-19 and AlexNet models are concatenated and then used to train a Probabilistic Collaborative Representation-based Classifier (ProCRC). The authors of [36] employ a multi-algorithm fusion approach by extracting feature vectors using BSIF, LBP [20], SIFT [22], SURF [4] and HOG with additional deep feature embeddings from OpenFace DNN [3] used as the last feature vector. These feature vectors are then used to train separate SVMs, applying score-level fusion at the

end to obtain the final decision score. Photo Response Non-Uniformity (PRNU) spectral analysis has also been proposed to analyze the alterations caused by morphing features [35]. In [26], the authors design a face morphing detector by combining spatial and frequency feature descriptors from an image. Fuzzy LBP in color channels of HSV and YCbCR color spaces are investigated in [33]. Additionally, studying the residual noise computed on color channels using deep CNN-based denoising has also been presented for reliable face morphing detection [45] [46]. This paper aggregates several denoised instances of an image in the wavelet domain.

There are a few papers that also address differential morph attack detection. Face demorphing has provided some encouraging results [13] [14]. In simplistic terms, the demorphing algorithm subtracts the potential morph image from the trusted image. The difference image is then compared to the trusted image and a low similarity score signifies a morphed image. Face demorphing has also been approached using a GAN framework to restore an accomplice's image [29]. Classical feature extraction methods have also been applied to the differential scenario by taking the difference of the feature vectors of the images being compared. This difference vector along with the original feature vector for the potential morph is then used to train a difference SVM and a feature SVM, respectively. This method is explored in [37], using LBP, BSIF, SIFT, SURF, and HOG descriptors. Scherhag et al. [38] uses deep face representations from feature embeddings extracted from ArcFace [10] to detect a morph attack. The authors also emphasized the need for high variance and trained their network on a morph database constructed using multiple different morph generation techniques. Disentanglement of appearance and landmarks is another method recently proposed for differential morph detection [42]. The use of Siamese networks for differential morph image detection has also been explored [43], but only in the image domain.

Because of the lack of large, publicly available morph database, many morph detection algorithms train on small morph datasets, usually created in house. However, researchers can submit their morph detection algorithms to the NIST FRVT morph detection test [28] for objective evaluation. Most of the algorithms submitted to NIST exhibit less than ideal performance on almost all tested morph datasets, which vary in quality and method. The deep learning method in [38] outperforms the other models in the NIST test, most likely due to the training protocol employing cross-database training.

3. Method

Our morph attack detection framework centers around applying undecimated 2D wavelet decomposition and training a Siamese deep neural network to classify morphs based

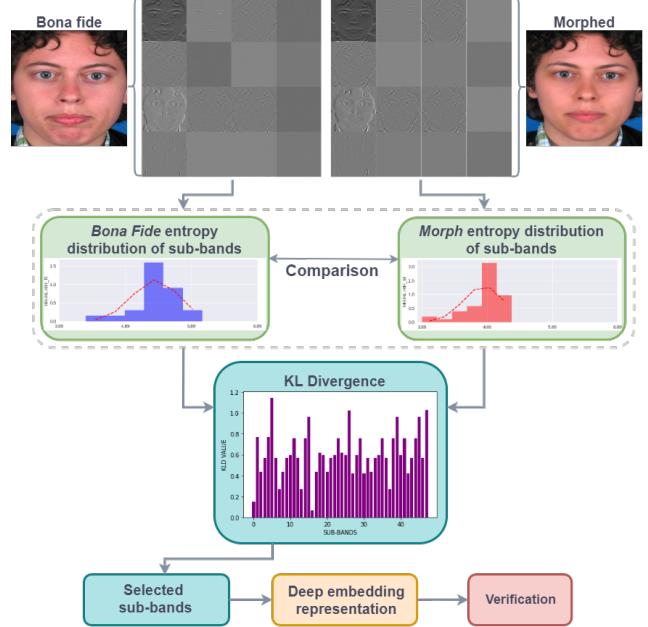


Figure 2. Discriminative wavelet sub-band selection algorithm. Bona fide (left) and morphed (right) images are decomposed into 48 wavelet sub-bands each. The entropy distributions and corresponding KL divergence values are found. For a given sub-band, the dissimilarity between the bona fide and morph entropy distribution represents how informative the sub-band is for morph detection. KL divergence is applied to isolate the discriminative sub-bands. A Siamese network is trained with the selected informative sub-bands.

on the most discriminative wavelet sub-bands. Because the morphing process can involuntarily introduce artifacts in the final morph image, the proposed method aims to isolate these artifacts in the wavelet domain and effectively utilize them for morph detection. A close study of the wavelet sub-bands shows that most morphing artifacts reside in the high frequency spectrum. As such, we do not consider the Low Low (LL) sub-band for decomposition and drop the LL sub-band completely after the first level of wavelet decomposition. Instead, we decompose only the Low High (LH), High Low (HL) and High High (HH) sub-bands down to the third level. After three levels of uniform decomposition, 48 sub-bands are obtained per image. We determine the most optimal sub-bands for network training using Shannon entropy [41] and KL divergence [19]. After instituting a threshold over the KLD values of the 48 sub-bands, we obtain 22 sub-bands with the highest KLD values. Figure 3 displays the sub-bands that are selected for network training and their location in the wavelet decomposition. A final set of 22 informative sub-bands is then used to train a Siamese deep neural network, consisting of the Inception ResNET

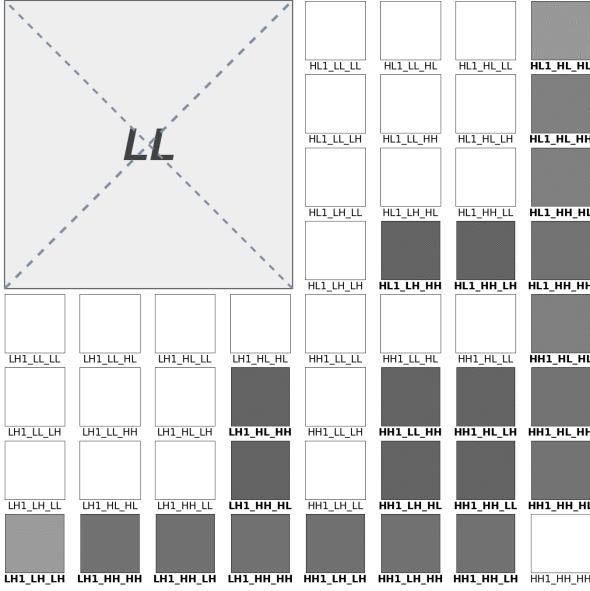


Figure 3. **Selected sub-bands.** The selected sub-bands are shown with regards to their location in wavelet decomposition. Most of the informative sub-bands chosen by KL divergence are those that have been filtered with the HH filter.

v1 architecture as the base network.

3.1. 2D Discrete Wavelet Transform

A 2D Wavelet transform decomposes an image in the frequency domain, essentially capturing different frequencies at different resolutions. This means that wavelet transform allows us to separately examine the approximation and detail data in an image. Particularly for morph detection, we can pinpoint the sub-bands where the morph artifacts appear and discard the sub-bands that are not informative for our problem.

Wavelet decomposition occurs by applying the low-pass and high-pass filters both vertically and horizontally simultaneously on a given image. After one level of decomposition, the LL, LH, HL, and HH sub-bands are obtained. We can continue decomposing the image further by filtering each sub-band separately. In our framework, we adopt undecimated wavelet decomposition, which maintains the resolution of the image with each decomposition, and decompose the LH, HL, and HH sub-bands specifically down to the third level. Our chosen naming convention for the sub-bands is such that each sub-band is labeled LH_HL_HH, where LH is the sub-band after the first decomposition, HL is the sub-band after the second decomposition, and HH is the sub-band after the third level of decomposition. As morphed images are, in essence, approximations of the original, our research indicates that the LL sub-band is unhelpful for

morph detection.

3.2. Sub-band Selection using KL Divergence

Training on all 48 sub-bands does not isolate which sub-bands truly contribute to the classification result. Consequently, we employ Shannon entropy and Kullback-Liebler divergence to identify the optimal sub-bands for morph attack detection. Shannon entropy, in particular, is used to measure the embedded information in each sub-band. Figure 2 illustrates the wavelet sub-band selection algorithm. For each of the three datasets, Shannon entropy and the entropy distributions are computed for all 48 sub-bands. Since we are interested in the comparison of bona fide and morphed images, we calculate the entropy distributions for all bona fide and morphed images in a dataset separately. Then, KL divergence (relative entropy) is calculated between the bona fide entropy distribution and the morph entropy distribution for each sub-band.

The method for finding the KLD values is as follows: after the entropy distributions of each sub-band are found, we find the histograms of entropy of all 48 sub-bands for both morphed and bona fide images. Accordingly, 96 normal distributions (48 bona fide and 48 morphed) are estimated using these histograms. \hat{f}_{b_i} represents the estimated bona fide normal distribution for the i^{th} sub-band, and similarly, \hat{f}_{m_i} represents the estimated morph normal distribution for the i^{th} sub-band. Dissimilarity of the two probability distribution functions, namely $(\hat{f}_{b_i}, \hat{f}_{m_i})$ is calculated for all 48 sub-bands and the KL divergence is computed for each relative entropy distribution.

The KLD values vary by dataset as each dataset is created using a different morphing technique. Therefore, we focus on selecting the sub-bands that are discriminative across different morphing techniques. As such, the KL divergence values of each dataset are normalized by removing the mean. Then, the normalized values are averaged over the KLD values of each sub-band for each of the three datasets. Figure 4 presents the distribution of the normalized KLD divergence values for the three morphed datasets and their average values. The higher the normalized KLD value for a single sub-band, the more informative the sub-band is for morph classification. After sorting the normalized KLD values from highest to lowest, we institute a threshold for selecting the sub-bands for training. According to the method described in [2], the optimal number of sub-bands is found to be 22. Thus, our final network is trained with 22 input channels, consisting of the top-22 most discriminative sub-bands for morph detection.

3.3. Siamese Network

A Siamese neural network [6] is the architecture used to train with the wavelet sub-bands. A Siamese network consists of two identical sub-networks which share weights.

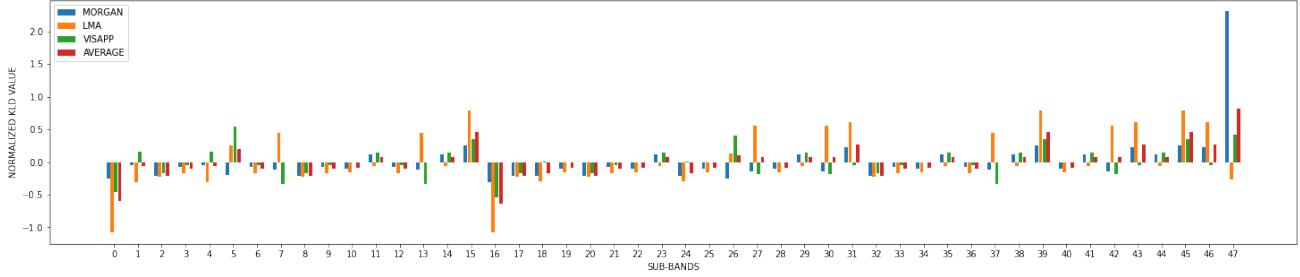


Figure 4. Normalized KL divergence values in the sub-bands 0 to 48 for all three morphing techniques: LMA (orange), MorGAN (blue), and VISAPP (green). The averaged KL divergence value is represented in red.

Siamese networks are ideal for morph attack detection as they are primarily designed to find similarities between two inputs. Contrastive loss [15] is the loss function utilized for training the Siamese network. Contrastive loss is a distance-based loss function which attempts to bring similar images closer together in a common latent space. At the same time, the loss function distances the dissimilar ones even more. Essentially, contrastive loss seeks to emphasize the similarity between samples of the same class and exaggerate the differences between images of different classes. The distance is found from the feature embeddings of the input pair produced by the Siamese network. The margin is the distance threshold that regulates the extent to which pairs are separated. The equation for calculating contrastive loss is as follows:

$$L_c = (1 - y_g)D(I_1, I_2)^2 + y_g \max(0, m - D(I_1, I_2))^2, \quad (1)$$

where I_1 and I_2 are the input face images, m is the margin or distance threshold to control the separation and y_g is the ground truth label for a given pair of training images and $D(I_1, I_2)$ is the L_2 distance between the feature vectors:

$$D(I_1, I_2) = \|\phi(I_1) - \phi(I_2)\|_2. \quad (2)$$

Here, $\phi(\cdot)$ represents a non-linear deep network mapping image into a vector representation in the embedding space. According to the loss function defined above, y_g is 0 for genuine image pairs and y_g is 1 for imposter (morph) pairs.

To streamline training, a Siamese Inception ResNET v1 architecture [44] is adopted, using weights pre-trained on the VGGFace2 dataset [7]. The network is then re-trained with the morphing datasets for the differential Siamese implementation. The model is optimized by enforcing contrastive loss on the embedding space representations of the genuine and imposter morph samples. The pre-trained Siamese network is then additionally fine-tuned using the training portion of each morph database. The feature embeddings are taken from the last fully connected layer and

the L_2 distance between the two embeddings is calculated for verification.

4. Experiments

4.1. Datasets

We train our network on three different morph datasets that apply three different morphing techniques: splicing, GAN generation, and landmark manipulation to investigate how our model generalizes. The two morph image databases used in this experiment are VISAPP [23] and MorGAN [8] [9]. VISAPP is a collection of complete and splicing morphs generated using the Utrecht FCVP database [1]. The images are 900×1200 in size. This dataset is generated by warping and alpha-blending two face images together [49] and then splicing the resulting face into one of the faces of the original contributing images. This preserves the background and hairline of one of the contributing faces, which helps avoid blurry artifacts and ghosting that typically occurs in these regions and makes morphs easier to spot [23]. For our network, we only use a subset of 183 high quality splicing morphs that is constructed by selecting the morph images that have no recognizable artifacts (VISAPP-Splicing-Selected dataset) along with 131 genuine neutral and smiling images for a total of 314 images.

The MorGAN database is generated from a selection of full frontal face images manually chosen from the CelebA dataset [21]. It consists of a custom morph image generation pipeline (MorGAN) [8], created by the authors that uses a GAN, inspired by the learned inference model [11], to generate morphs. The encoder in the GAN transforms the images into a latent space and when two latent spaces related to two different subjects are combined, a morphed image is synthesized. The database consists of 1,500 bona fide reference images, 1,500 bona fide probe images and 1,000 MorGAN morphs of size 64×64 pixels. To compare their GAN morphs, the authors also generate 1,000 LMA (landmark manipulation) morphs [25]. The VISAPP

Table 1. Performance of the proposed framework and baselines. With the exception of RGB-66 testing on MorGAN, BW-22 exhibits superior performance.

Testing	Method	APCER@BPCER		BPCER@APCER		D-EER %
		5%	10%	5%	10%	
MorGAN	BW images	7.88	6.17	13.1	3.1	5.57
	RGB images	4.5	3.3	3.22	1.74	4.17
	LL-removed BW images	5.5	3.14	4.5	3.28	5.53
	LL-removed RGB images	3.66	2.98	1.58	0.79	3.55
	BW-22 wavelets	3.71	1.85	3.06	0.26	3.89
LMA	RGB-66 wavelets	0.86	0.0	0.37	0.37	1.62
	BW images	22.7	14.3	36.5	15.1	11.6
	RGB images	11.1	6.68	12.2	5.62	8.8
	LL-removed BW images	25.9	14.4	19.0	11.5	11.5
	LL-removed RGB images	15.75	7.4	12	6.48	8.06
	BW-22 wavelets	4.95	2.67	4.38	1.46	4.52
VISAPP	RGB-66 wavelets	10.53	5.39	9.44	4.72	7.36
	BW images	5.97	0.0	0.0	0.0	3.17
	RGB images	1.32	0.08	0.0	0.0	0.0
	LL-removed BW images	1.57	0.08	5.63	4.22	0.0
	LL-removed RGB images	2.98	0.8	0.0	0.0	3.25
	BW-22 wavelets	0.0	0.0	0.0	0.0	0.0
UNIVERSAL	RGB-66 wavelets	0.0	0.0	0.0	0.0	0.0
	BW images	15.0	8.95	14.4	7.5	8.53
	RGB images	6.65	4.01	5.22	2.5	5.63
	LL-removed BW images	19.1	6.74	10.872	7.78	8.45
	LL-removed RGB images	10.9	3.53	5.52	4.56	5.52
	BW-22 wavelets	3.25	1.69	3.01	0.65	3.93

and MorGAN images differ significantly in terms of quality and resolution. However, varying quality and resolution during training can result in a network that is more robust and performs better on different morphing techniques, particularly when the bona fide images are of equal resolution to the morphed images. An additional publicly available dataset, the Advanced Multimedia Security Lab (AMSL) Face Morph Image dataset [27], is used as an “unseen” dataset to measure cross database performance. This dataset consists of approximately 102 bona fide images and 2,175 morph images, created using the Combined Morph tool, as described in [27]. This dataset is used for testing purposes only.

All images are preprocessed according to the FaceNet architecture [39]: face detection and alignment is performed via MTCNN [50]. All images are resized to 160×160 pixels before uniform wavelet decomposition is applied. According to the KLD sub-band selection algorithm, the top 22 wavelet sub-bands are selected for each image to prepare for network training. As the Siamese network expects pairs of input, the morph wavelet bands are paired off into genuine face pairs and imposter face pairs, where a genuine pair consists of two trusted 22 selected wavelet sub-bands and an imposter pair consists of a trusted image’s wavelet sub-bands and a corresponding morph image’s wavelet sub-bands. 50% of the subjects are considered for training while the other 50% are used to evaluate the performance of the network. In addition, 15% of the test set is selected during model optimization as the validation set. The training

data is further augmented with horizontal flips to increase the training set and improve generalization. By design, the train-test split is disjoint, with no overlapping morphs or contributing bona fides to morphs. This enables us to attain an accurate representation of performance. Batch size of 64 pairs of 22 selected wavelet sub-bands of size $22 \times 160 \times 160$ is used for training the model. The batch generator also compensates for class imbalance, ensuring that the network sees an equal number of morph pairs and genuine pairs every iteration.

4.2. Network Setup and Metrics

We fine-tune an Inception ResNET v1, already pre-trained on VGGFace2, on the 22 selected sub-bands. We train the network using the training portions of all three datasets, calling it the “universal” dataset. The margin m of contrastive loss is set to 1. Adam is the chosen optimizer and the initial learning rate is 0.0001. The performance is monitored by the validation loss and whenever the validation loss achieves a new low, the best weights are saved. Every time the validation loss plateaus, the best weights are re-loaded, the learning rate is divided by 10, and training continues from there down to 1e-07. After that, early stopping is implemented if the loss still does not improve after 35 epochs. The network is implemented using PyTorch and training is accelerated with the use of three 12 GB Titan X (Pascal) GPUs. We train the network on the universal dataset for 150 epochs. All experimental results in the paper are reported for the final iteration.

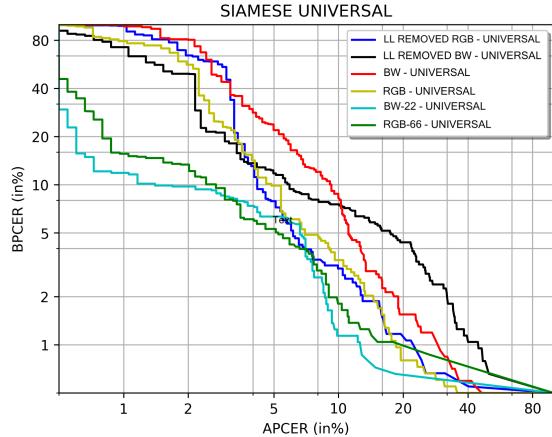


Figure 5. DET curves for all protocols, tested on the universal test set. The universal test set consists of the testing portions of all three datasets and serves as an indicator for overall network performance.

The standard quantitative measures for morph attacks are used to measure performance: APCER and BPCER. Attack Presentation Classification Error Rate (APCER) is the percentage of morphed samples incorrectly classified as bona fide. Conversely, the Bona-fide Presentation Classification Error Rate (BPCER) is the percentage of bona fide images classified as morphs. D-EER stands for Detection Equal Error Rate at which APCER equals BPCER. The APCER5 is the APCER rate when $BPCER = 5\%$ and similarly APCER10 is the APCER rate when $BPCER = 10\%$. In real world applications, the BPCER rate is the measure by which individuals are inconvenienced with a false alarm. We plot these rates in a Detection Error Tradeoff (DET) graph.

4.3. Results

We assess the performance of our morph detector using the test data of each individual dataset as well as the universal test set (which consists of all three individual test sets). The universal test set performs essentially as an “average” of how the universally trained network performs on each of the individual databases. We feel it is important to capture how the network learns each morphing technique separately and generally. As is standard for wavelet transform, the images are converted to grayscale before wavelet decomposition is applied. The 22 most discriminative sub-bands are selected for training according to the procedure described in Section 3.2. These grayscale 22 selected sub-bands, called BW-22, are then used to train a Siamese network for morph detection.

To determine how wavelets perform for RGB images, we also apply wavelet decomposition separately to each channel of an RGB image following the methodology described

in Section 3, yielding 144 wavelet sub-bands (48 sub-bands per channel). 22 sub-bands are chosen for each channel, totaling 66 sub-bands each (RGB-66). The 22 sub-bands selected from each channel are the same as BW-22 to facilitate comparison. The training protocol is identical for both BW-22 and RGB-66, except for the batch size. Because RGB-66 consists of 66 channels of 160×160 , the batch size is halved to 32 in our configuration to conserve memory.

Accordingly, we compare the performance of our wavelet Siamese networks, BW-22 and RGB-66, with other frameworks to validate the efficacy of our method. We apply each baseline to both color and grayscale images to gauge how color information plays a role in morph detection. The first baseline is RGB images (referred to as RGB) to compare how the original images’ performance varies from that in the wavelet domain. We also train a Siamese network for the grayscale (referred to as BW) images to act as a baseline for BW-22 before wavelet decomposition is applied. Additionally, as the LL sub-band has shown to be unhelpful for morph detection, we have only decomposed the mid- and high-frequency information for our wavelet Siamese network. To mirror the removal of the approximation data in the image domain, we decompose the images into wavelet sub-bands. Then, we remove the LL-band in the wavelet domain, and reconstruct the image using Inverse Wavelet Transform. We do this for both RGB images and grayscale images in our dataset, designating these baselines as LL-removed RGB and LL-removed BW respectively. Theoretically, the LL-removed RGB and LL-removed BW baselines should be approximately equivalent to the original 144 sub-bands of RGB-66 and the 48 sub-bands of BW-22 before sub-band selection. We train a separate Siamese network using the same training protocol for each of the above scenarios: RGB, BW, LL-removed RGB, LL-removed BW, RGB-66, and BW-22.

All networks are trained using the combined portions of the three datasets: VISAPP, LMA, and MorGAN to obtain a robust network that generalizes to many different morphing techniques. The performance of each network can be observed in Table 1. From Table 1, it is clear VISAPP performs consistently, regardless of framework. This is likely due to the small size of the VISAPP dataset that the models easily learn VISAPP’s morphing technique. Generally speaking, the networks trained on RGB information perform better. This implies that there is in fact information in the color channels that is useful for morph detection. Notably, grayscale images exhibit poor performance all around. Table 1 also shows how difficult LMA morphs are to detect in comparison to MorGAN and VISAPP. This is in line with research that GAN-generated morphs are easier to detect than landmark manipulation morphs [47]. BW-22 performs significantly better on LMA morphs than the

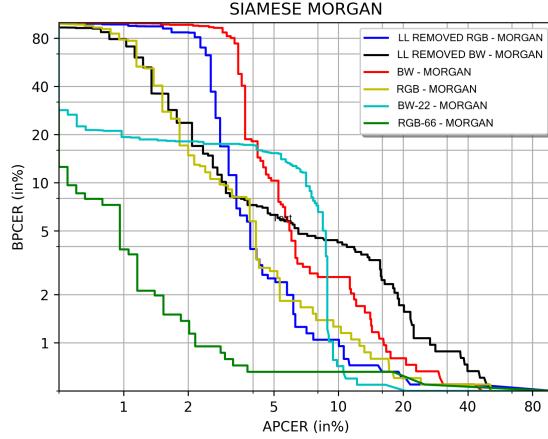


Figure 6. DET curves for all protocols, tested on the MorGAN test set. RGB-66 performs the best for this morphing technique.

other models. From Figure 6, it should be noted RGB-66 performs extremely well on the MorGAN dataset. This suggests that the MorGAN morphing technique contains more color information than LMA and VISAPP, which were creating using landmark-based techniques. MorGAN, on the other hand, was created using a GAN architecture, meaning it is essentially a synthesized image created using the two contributing images. Still, while RGB-66 performs unusually well for the MorGAN test set, BW-22 performs better on the LMA dataset (see Figure 7) as well as overall performance as can be seen in Figure 5, which shows the performance of the networks on the 'universal' dataset. From the results, we can derive that color information plays a smaller role in the classification of morphs in the wavelet domain. Figure 5 shows the DET curves for all tested networks. It is clear from Figure 5 that the wavelet Siamese networks, particularly BW-22, exhibit superior performance.

Additionally, in Table 2, we compare the performance of our wavelet-based morph detector with other classical feature extraction techniques that have been used for morph detection, namely BSIF [17], LBP [20], SIFT [22], and SURF [4], each combined with an SVM. Each method is trained using the universal dataset and evaluated on the test sets of the individual datasets. We also measure the baseline performance of FaceNet [39] on all the morph test sets. To evaluate cross-database performance, AMLS dataset is used for testing only.

5. Conclusion

In this paper, we introduced a framework to detect morphed face images using undecimated DWT. The core of our method was the ability to identify morph artifacts in the wavelet domain and to leverage the most informative sub-bands for differential morph detection. To select the optimal

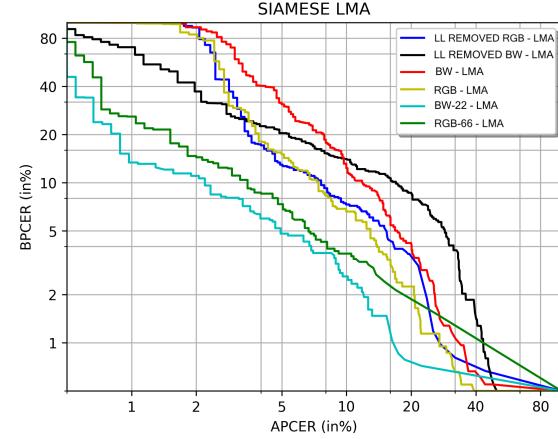


Figure 7. DET curves for all protocols, tested on the LMA test set.

Table 2. Performance Comparison of Proposed Framework. All algorithms trained with the Universal dataset.

Testing	Method	APCER@BPCER		D-EER
		5%	10%	
MorGAN	SURF	86.8	70.11	46.1
	SIFT	57.6	47.7	27.3
	LBP	90.13	82.2	41.6
	BSIF	86.8	71.6	31.7
	FaceNet	36.80	31.15	22.25
	BW-22 wavelets	0.86	0.0	1.62
LMA	SURF	81.1	63.69	51.1
	SIFT	63.2	55.8	36.7
	LBP	91.1	83.4	40.5
	BSIF	86.5	75.0	36.4
	FaceNet	43.70	40.90	30.35
	BW-22 wavelets	4.95	2.67	4.52
VISAPP	SURF	94.1	90.3	47.8
	SIFT	91.1	84.7	52.2
	LBP	31.1	19.5	16.0
	BSIF	30.6	22.73	16.4
	FaceNet	25.0	15.8	15.5
	BW-22 wavelets	0.0	0.0	0.0
AMSL	SURF	96.7	91.3	53.0
	SIFT	94.65	84.9	38.0
	LBP	91.0	72.9	43.0
	BSIF	91.0	82.0	41.3
	FaceNet	38.6	31.35	19.86
	BW-22 wavelets	33.78	23.61	16.4

sub-bands, a data driven approach based on KL divergence is employed. The 22 selected sub-bands were then used to train a deep Siamese network successfully. Our framework achieves an EER of 3.93% on the universal test set, significantly better than the other baselines. Furthermore, the framework performs well on an unseen morph dataset, AMLS, that uses a different morphing technique than our training set, achieving an EER of 16.4%. This shows how wavelet decomposition with selective sub-bands is useful in the morph problem domain.

References

- [1] Utrecht ecvp face database.
- [2] Poorya Aghdaie, Baaria Chaudhary, Sobhan Soleymani, Jeremy Dawson, and Nasser Nasrabadi. Detection of morphed face images using discriminative wavelet sub-bands. In *2021 International Workshop on Biometrics and Forensics (IWBF)*, 2021.
- [3] Brandon Amos, Bartosz Ludwiczuk, Mahadev Satyanarayanan, et al. Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science*, 6:2, 2016.
- [4] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *European conference on computer vision*, pages 404–417. Springer, 2006.
- [5] Thirimachos Bourlai. *Face recognition across the imaging spectrum*. Springer, 2016.
- [6] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature verification using a “siamese” time delay neural network. In *Advances in neural information processing systems*, pages 737–744, 1994.
- [7] Qiong Cao, Li Shen, Weidi Xie, Omkar M Parkhi, and Andrew Zisserman. Vggface2: A dataset for recognising faces across pose and age. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 67–74. IEEE, 2018.
- [8] Naser Damer, Alexandra M. Saladié, Andreas Braun, and Arjan Kuijper. Morgan: Recognition vulnerability and attack detectability of face morphing attacks created by generative adversarial network. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–10, 2018.
- [9] Luca Debiasi, Naser Damer, Alexandra Moseguí Saladié, Christian Rathgeb, Ulrich Scherhag, Christoph Busch, Florian Kirchbuchner, and Andreas Uhl. On the detection of gan-based face morphs using established morph detectors. In *International Conference on Image Analysis and Processing*, pages 345–356. Springer, 2019.
- [10] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4690–4699, 2019.
- [11] Vincent Dumoulin, Ishmael Belghazi, Ben Poole, Olivier Mastropietro, Alex Lamb, Martin Arjovsky, and Aaron Courville. Adversarially learned inference. *arXiv preprint arXiv:1606.00704*, 2016.
- [12] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. The magic passport. In *IEEE International Joint Conference on Biometrics*, pages 1–7. IEEE, 2014.
- [13] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face demorphing. *IEEE Transactions on Information Forensics and Security*, 13(4):1008–1017, 2018.
- [14] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face demorphing in the presence of facial appearance variations. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 2365–2369. IEEE, 2018.
- [15] Raia Hadsell, Sumit Chopra, and Yann LeCun. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, volume 2, pages 1735–1742. IEEE, 2006.
- [16] Doc ICAO. 9303-machine readable travel documents-part 9: Deployment of biometric identification and electronic storage of data in emrtds. *International Civil Aviation Organization (ICAO)*, 2015.
- [17] Juho Kannala and Esa Rahtu. Bsif: Binarized statistical image features. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, pages 1363–1366. IEEE, 2012.
- [18] Robin Kramer, Michael Mireku, Tessa Flack, and Kay Ritchie. Face morphing attacks: Investigating detection with humans and computers. *Cognitive Research: Principles and Implications*, 4, 12 2019.
- [19] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- [20] Shengcai Liao, Xiangxin Zhu, Zhen Lei, Lun Zhang, and Stan Z Li. Learning multi-scale block local binary patterns for face recognition. In *International Conference on Biometrics*, pages 828–837, 2007.
- [21] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaou Tang. Deep learning face attributes in the wild. In *Proceedings of International Conference on Computer Vision (ICCV)*, December 2015.
- [22] David G. Lowe. Sift-the scale invariant feature transform. *International Journal of Computer Vision*, 2:91–110, 2004.
- [23] Andrey Makrushin, Tom Neubert, and Jana Dittmann. Automatic generation and detection of visually faultless facial morphs. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 6 VISAPP: VISAPP, (VISIGRAPP 2017)*, pages 39–50. INSTICC, SciTePress, 2017.
- [24] Andrey Makrushin, Tom Neubert, and Jana Dittmann. Humans vs. algorithms: Assessment of security risks posed by facial morphing to identity verification at border control. In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 4: VISAPP*, pages 513–520. INSTICC, SciTePress, 2019.
- [25] Satya Mallick. Face morph using opencv-c++/python, 2016.
- [26] Tom Neubert, Christian Kraetzer, and Jana Dittmann. A face morphing detection concept with a frequency and a spatial domain feature space for images on emrtd. In *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*, pages 95–100, 2019.
- [27] Tom Neubert, Andrey Makrushin, Mario Hildebrandt, Christian Kraetzer, and Jana Dittmann. Extended stirtrace benchmarking of biometric and forensic qualities of morphed face images. *IET Biometrics*, 7(4):325–332, 2018.
- [28] Mei Ngan, Patrick Grother, Kayee Hanaoka, and Jason Kuo. Face recognition vendor test (frvt) part 4: Morph performance of automated face morph detection. *National Institute of Technology (NIST), Tech. Rep. NISTIR*, 8292, 2020.

[29] Fei Peng, Le-Bing Zhang, and Min Long. FD-GAN: face-demorphing generative adversarial network for restoring accomplice's facial image. *CoRR*, abs/1811.07665, 2018.

[30] R Raghavendra, KiranB Raja, Sushma Venkatesh, and Christoph Busch. Face morphing versus face averaging: Vulnerability and detection. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 555–563. IEEE, 2017.

[31] Ramachandra Raghavendra, Kiran B. Raja, and Christoph Busch. Detecting morphed face images. In *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–7, 2016.

[32] Ramachandra Raghavendra, Kiran B Raja, Sushma Venkatesh, and Christoph Busch. Transferable deep-cnn features for detecting digital and print-scanned morphed face images. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1822–1830. IEEE, 2017.

[33] Raghavendra Ramachandra, Sushma Venkatesh, Kiran Raja, and Christoph Busch. Towards making morphing attack detection robust using hybrid scale-space colour texture features. In *2019 IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, pages 1–8, 2019.

[34] David J Robertson, Robin SS Kramer, and A Mike Burton. Fraudulent id using face morphs: Experiments on human and automatic recognition. *PLoS One*, 12(3):e0173319, 2017.

[35] Ulrich Scherhag, Luca Debiasi, Christian Rathgeb, Christoph Busch, and Andreas Uhl. Detection of face morphing attacks based on prnu analysis. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 1(4):302–317, 2019.

[36] Ulrich Scherhag, Christian Rathgeb, and Christoph Busch. Morph detection from single face image: A multi-algorithm fusion approach. In *Proceedings of the 2018 2nd International Conference on Biometric Engineering and Applications*, ICBEA '18, page 6–12, New York, NY, USA, 2018. Association for Computing Machinery.

[37] Ulrich Scherhag, Christian Rathgeb, and Christoph Busch. Towards detection of morphed face images in electronic travel documents. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pages 187–192, 2018.

[38] Ulrich Scherhag, Christian Rathgeb, Johannes Merkle, and Christoph Busch. Deep face representations for differential morphing attack detection. *arXiv preprint arXiv:2001.01202*, 2020.

[39] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.

[40] Clemens Seibold, Wojciech Samek, Anna Hilsmann, and Peter Eisert. Detection of face morphing attacks by deep learning. pages 107–120, 07 2017.

[41] Claude E Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[42] Sobhan Soleymani, Baaria Chaudhary, Ali Dabouei, Jeremy Dawson, and Nasser M. Nasrabadi. Differential morphed face detection using deep siamese networks, 2020.

[43] Sobhan Soleymani, Ali Dabouei, Fariborz Taherkhani, Jeremy Dawson, and Nasser M. Nasrabadi. Mutual information maximization on disentangled representations for differential morph detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1731–1741, January 2021.

[44] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*, 2017.

[45] Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Luuk Spreeuwers, Raymond Veldhuis, and Christoph Busch. Morphed face detection based on deep color residual noise. In *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6, 2019.

[46] Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Luuk Spreeuwers, Raymond Veldhuis, and Christoph Busch. Detecting morphed face attacks using residual noise from deep multi-scale context aggregation network. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 269–278, 2020.

[47] Sushma Venkatesh, Haoyu Zhang, Raghavendra Ramachandra, Kiran Raja, Naser Damer, and Christoph Busch. Can gan generated morphs threaten face recognition systems equally as landmark based morphs? - vulnerability and detection. In *2020 8th International Workshop on Biometrics and Forensics (IWBF)*, pages 1–6, 2020.

[48] Lukasz Wandzik, Raul Vicente Garcia, Gerald Kaeding, and Xi Chen. Cnns under attack: On the vulnerability of deep neural networks based face recognition to image morphing. In *International Workshop on Digital Watermarking*, pages 121–135. Springer, 2017.

[49] George Wolberg. Image morphing: a survey. *The visual computer*, 14(8-9):360–372, 1998.

[50] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, 2016.