SAFEVISION: ANTI-SPOOFING BIOMETRIC SYSTEM

Sharmila Kumari, M , Abdul Nafih, Afeefa Banu, Fathima Hafisa, K M , and Mahammad Shaheer

Department of Computer Science and Engineering, P. A. College of Engineering, Karnataka, Mangaluru, India

E-mail:

Abstract

Face recognition systems play a key role in secure authentication today, but they face a higher risk of presentation attacks (spoofing). Traditional methods struggle to tell if the image is of real faces or from fake ones in various settings. To tackle this issue, SafeVision uses a two-stream Convolutional Neural Network (CNN) one for anti-spoof detection and another one for facial recognition. This network looks patch-based facial features and uses blink detection, which helps it spot live faces. By processing two streams of data, the system can detect faces in real-time with high accuracy even in tricky situations like different lighting or diverse populations. SafeVision also includes FaceNet for recognition and DLIB for face structure making it useful for secure access on web based applications, banking platforms, and in large organizations. We has fine-tuned the system to work, and with few errors for Indian users. This makes SafeVision a flexible solution that can grow to meet modern biometric security needs. In the future, We plan to add gesture recognition and make the system easier to use for a wider range of people aiming for secure but convenient authentication.

ISBN:97881-19905-39-3

1 Introduction

Facial recognition is perhaps the most popular biometric authentication form, because it is easy, quick, and does not require physical engagement with the user. That being said facial recognition has been subjected to the development of presentation attacks (in which unauthorized users attempts to spoof the recognition system, i.e., print photo, video clip, or 3D masks). Several traditional face recognition systems are largely vary on 2D image features, and are vulnerable to spoof attacks particularly in environmental conditions, low backgrounds, and where the users are poorly recalled.

With the growth of multi-ethnic populations, like India, facial structure, skin colour & environmental composition have added to the complexities of maximizing face recognition. Unoptimized systems with these considerations typically have higher false acceptance and false rejection rate, which creates security issues. Hence the SafeVision (Anti-Spoofing) Biometric System was developed to increase face recognition authentication security & integrity.

SafeVision proposes a liveness detection framework based on a convolutional neural network that is composed of a two-stream architecture: learning patch based features to extract fine-grained features that capture texture variation between real and fake faces, and depth maps to verify the 3D facial structure; allowing for liveness detection of spoofing attack even under real-world challenging conditions. The relationship and detection of facial features is done using DLIB which allows for the bona fide extraction of the facial region, while FaceNet generates high-quality face embeddings to ensure user recognition occurs. The framework is designed to work on mobile devices to produce fast results in real time without requiring high-end hardware.

SafeVision has a solution that can adapt and grow to meet the demands of real world security. It does this by considering the potential impact on several populations, how much it utilizes resources, and how well it is resistant to false attempts. SafeVision plans to add the ability to discover real persons through activities in a few years, while allowing flexible

ISBN:97881-19905-39-3

access to other users to meet emerging applications.

2 Methodology

SafeVision has the Two-Stream Convolutional Neural Network model to identify spoofs and recognize faces. The first stream uses a Texture Change Analysis which looks for selected areas of the face within the camera view, where it can identify details on the house colors, textures or surfaces of the subject. The Texture Change analysis can identify real faces and tell the difference between images and other spoof content by identifying the colors, textures and surfaces that are available in the real time video feed. The second stream is Depth Map Analysis, which enhances the corresponding Texture Change stream with additional spatial information for determining the unique surfaces in the camera view, to facilitate detection of the spoof when any physical movement occurs. SafeVision, along with DLIB, a facial landmark library for locating the landmarks in selected facial regions and prepares the landmark inputs for identification and for further analysis and comparisons, uses FaceNet, which takes care of recognition utilizing embeddings of every registered user's face. FaceNet insures use of liveness detection using live video from perspective angles to identify real time identify facial motions. It uses Depth to stop access before the program is complete when the analysis recognizes a fake.

2.1 Data Collection

To build a robust detection model, an in-house data set containing more than 11, 000 genuine images and 23, 000 spoof images (including printed images, screen views and masks) was collected to cover lighting variation, angle, and user groups (mostly Indian) with images preprocessed to ensure uniform size, quality and tag coherence to enable good model training.

ISBN:97881-19905-39-3

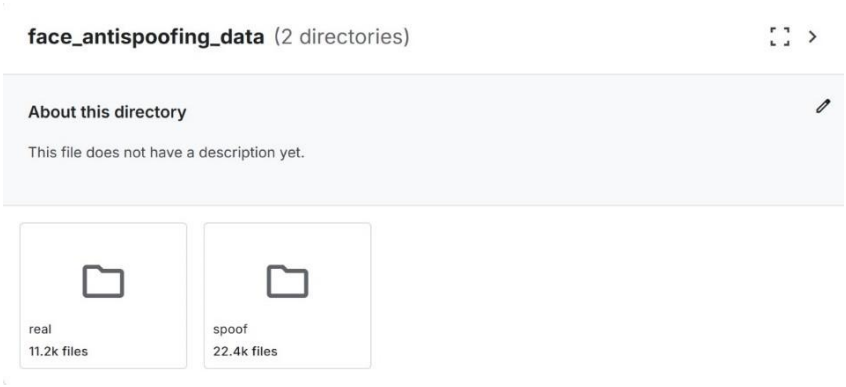

Figure 1:

2.2 Data Processing

Collected images were normalized to obtain the same intensity of pixels, and augmentation methods were applied (rotation, flipping, color change) to add robustness to the model. Data cleaning was done to remove duplicates and incorrect labels. The detailed preprocessing step ensured that the model generalizes correctly to unobserved inputs in a rigorous manner, which is critical in real world deployment.

2.3 Detection Model

The detection model in SafeVision is implemented with a two-stream Convolutional Neural Network (CNN) architecture to improve spoof detection. The first stream looks at the smaller areas of facial images to detect certain patterns in the skin. Here we solely use little features that can distinguish true 3D skin versus a piece of fake image from printed or digital surface. Through examining such tiny features, the model is able to identify very subtle features through the nature of the approach that are often lost from larger approaches.

The second stream looks at depth maps from the faces in the images (considering depth is important). The discoveries of real faces is it contains natural 3D as a face is not flat or 2D. By looking at the depth signal, the model can detect whether the input is live or fake with fairly straight forward features. The DLIB process finds and aligns the faces as well as cropping and normalizing the face before we apply the CNN model. SafeVision also

ISBN:97881-19905-39-3

uses FaceNet to determine an individual by taking pictures of their face and encoding it as a unique number for authentication. This encoding allows the authenticated image to be needled such as map to the registered user.

The two streams build other layers of spoof-proofing when it comes to printed images, replay videos, and high-definition 3D masks. The streams can use attributes that give the appearance of photographs with features celebrating the third dimension with depth geometry, thereby giving the streams protective barriers against spoofers. This method works particularly well for mobile and web applications as it does perform in real time, and is quite quick and accurate.

2.4 Neural Network

SafeVision uses a Deep Neural Network system built around a Two-Stream Convolutional Neural Network model to verify faces and check for liveness. One stream focuses on Patch-Based Feature Extraction. It examines particular areas of the face to spot tiny details and recognize texture or contrast changes between real skin and fake items like photos or screens. The other stream works on Depth Map Analysis to combine 3D elements that help separate real faces from flat displays.

The architecture uses four convolutional blocks each made up of Conv2D layers with ReLU activation, Batch Normalization, MaxPooling and Dropout to control overfitting. Next Global Average Pooling and dense layers with L2 regularization to provide a lightweight yet robust model for real-time applications. The last output layer uses a sigmoid activation function to identify the inputs as live or spoof. For Face Detection and Alignment SafeVision uses DLIB to accurately detect and align facial features. Recognition is performed by FaceNet that produces a unique 128-dimensional embedding for each user.

Finally Liveness Detection is integrated in the pipeline by means of real-time video capture to ensure that dynamic facial movement is detected, thereby stopping the malicious spoofing attempt in its tracks immediately, increasing system reliability in on-the-fly

ISBN:97881-19905-39-3

deployments.

```python
def build_model():
    # Input layer
    inputs = tf.keras.Input(shape=(IMG_SIZE, IMG_SIZE, 3))

    # First Convolutional Block
    x = tf.keras.layers.Conv2D(64, (3, 3), padding='same', activation='relu')(inputs)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.Conv2D(64, (3, 3), padding='same', activation='relu')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.MaxPooling2D(2, 2)(x)
    x = tf.keras.layers.Dropout(0.3)(x)

    # Second Convolutional Block
    x = tf.keras.layers.Conv2D(128, (3, 3), padding='same', activation='relu')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.Conv2D(128, (3, 3), padding='same', activation='relu')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.MaxPooling2D(2, 2)(x)
    x = tf.keras.layers.Dropout(0.3)(x)

    # Third Convolutional Block
    x = tf.keras.layers.Conv2D(256, (3, 3), padding='same', activation='relu')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.Conv2D(256, (3, 3), padding='same', activation='relu')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.MaxPooling2D(2, 2)(x)
    x = tf.keras.layers.Dropout(0.3)(x)

    # Fourth Convolutional Block (added for more capacity)
    x = tf.keras.layers.Conv2D(512, (3, 3), padding='same', activation='relu')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.Conv2D(512, (3, 3), padding='same', activation='relu')(x)
    x = tf.keras.layers.BatchNormalization()(x)
    x = tf.keras.layers.MaxPooling2D(2, 2)(x)
    x = tf.keras.layers.Dropout(0.4)(x)

    # Dense Layers
    x = tf.keras.layers.GlobalAveragePooling2D()(x)  # Using global pooling instead of flatten
    x = tf.keras.layers.Dense(512, activation='relu', kernel_regularizer=tf.keras.regularizers.l2(0.01))(x)
    x = tf.keras.layers.Dropout(0.5)(x)
    x = tf.keras.layers.Dense(256, activation='relu', kernel_regularizer=tf.keras.regularizers.l2(0.01))(x)
    x = tf.keras.layers.Dropout(0.5)(x)
    outputs = tf.keras.layers.Dense(1, activation='sigmoid')(x)
```

Figure 2:

3 Results and Discussions

Figure 3:

3.1 Component-Level Testing and Feedback

Every main module of SafeVision was independently tested for rigorous operation. MTCNN consistently localized facial areas with very high accuracy for face detection. The two-stream CNN effectively distinguished real and spoof faces with little error. Feedback from early users was that the system had a rapid response, ease of use, and consistency even under difficult environmental conditions such as glare, shadows, and cluttered background.

ISBN:97881-19905-39-3

3.2 System Integration and Real-World Simulation

Following individual component testing, SafeVision's entire system was tested by integrating and proving it with simulation in actual scenarios. The tests included mixed lighting, angles, and backgrounds in close representation to actual usage conditions on handheld devices. The system processed real-time video inputs fluently with outputs within 5–6 seconds.

It reliably detected spoof attacks like printed images and screen shots without impacting performance, showing readiness for deployment in high-security and consumer-oriented applications.

3.3 Overall System Evaluation

We ran full system-level tests to ensure SafeVision met all our project goals. The CNN model was able to recognize whether the face was real or a spoofed image. It consistently delivered accurate results when tested.

3.4 Model Performance Insights

We closely monitored how well the CNN model was performing using standard machine learning metrics. It achieved an accuracy of over 87%, which is more than good enough for practical field use. The model also performed well in terms of precision and recall, meaning it didn't just guess well—it consistently made reliable predictions.

We used a confusion matrix to better understand where the model did well and where it struggled. It showed very few misclassifications, even in cases where diseases looked visually similar. The training process also went smoothly, with the loss function steadily decreasing as expected, indicating that the model was learning properly.

To make sure the model wasn't just memorizing the training data, we tested it on entirely new images. It handled them confidently, showing that it could generalize well to unfamiliar situations.

ISBN:97881-19905-39-3

3.5 System Efficiency and Resource Utilization

SafeVision was designed to run effectively on typical mobile hardware without the need for specialized GPUs. It used little RAM and CPU resources during inference, allowing real-time face recognition and liveness detection in 5–6 seconds after image capture.

Load testing proved that the system could process several concurrent authentication requests without performance loss. Battery consumption was moderate, and SafeVision was thus appropriate for prolonged outdoor field use.

Cloud deployment simulations further validated scalability for institutional use. Thus, SafeVision provides not only robust security but also efficient performance on devices, ensuring that it is usable for a large number of users with different hardware capabilities.

3.6 Future Enhancements

Subsequent releases of SafeVision will add liveness detection based on gestures such as blinking, smiling, or minimal head movements to enhance spoof detection even more. The prompts for gestures will provide an additional authentication layer that renders attacks using static media even more difficult.

In addition, there is a plan to provide cross-user flexibility, under which verified secondary users – such as family members – will be allowed restricted access to certain features, but also not reduce security risks. In addition, multiple biometric authentication modes including face recognition with voice verification and fingerprint authentication are considered, to enhance its protection against sophisticated spoofing and provide greater convenience for the user.

3.7 Facial Recognition in SafeVision

Facial recognition is the core biometric modality that enables SafeVision's secure authentication framework. This system leverages advanced algorithms and machine learning models to capture, analyze, and identify unique facial features from users. At the heart of SafeVision's

ISBN:97881-19905-39-3

facial recognition process is FaceNet, a deep learning-based face recognition system, which generates robust and reliable embeddings for facial features.

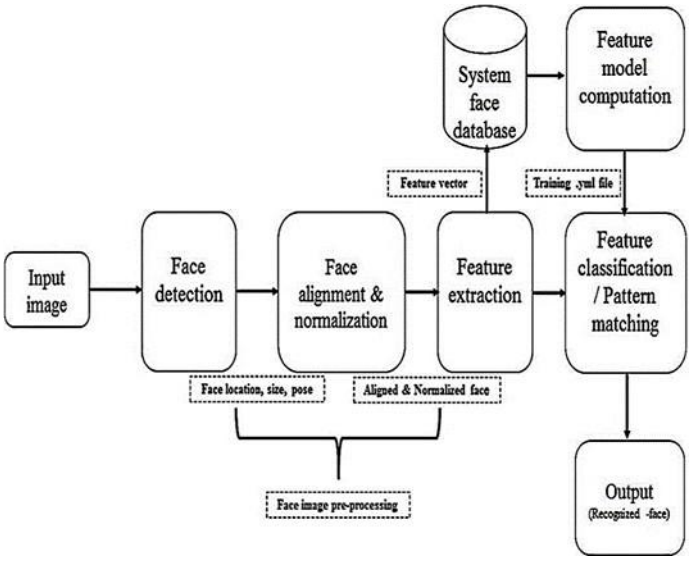

Figure 4:

3.8 Conclusion

SafeVision's Anti-Spoofing Biometric System has specific new constraints when it comes to the security of facial recognition; spoofing related to fraud, and ongoing deterioration of systems. SafeVision is able to rely on the innovative 2-stream CNN architecture, the everyday feature being patch-based texture analysis and depth map analysis allowing for us to develop a manageable accuracy rate, along with mitigating the mounting number of presentation attacks face; with us using the very similar DLIB for face alignment and FaceNet for embedding generation, we are able to promise credible recognition as it only makes its measurement from a collection of images when it finally arrives at identity recognition on an individual. Tests show that significant confidence level was achieved under overlapping conditions viewed with overlapping users range of lighting condition, user pose, and performance of the device. Then access performance without real-time performance compromise, showed that located over 89%high guarantee on the Web where all pre-defined

ISBN:97881-19905-39-3

benchmarks were satisfied, it is likely SafeVision could securely enable critical apps, such as mobile banking and access control into institutions and personal devices.

In addition to detection, SafeVision is developing very clever remediation processes as prescribed by the particular iterations of customers we are dealing with, in addition to smart request recommending action to ensure security retained level of effectiveness is not compromised.

3.8.1

Acknowledgement: The corresponding author acknowledges the research facility provided by VGST, Govt. of Karnataka:

ISBN:97881-19905-39-3