**Dataset for Large-scale Multi-modal Face Anti-spoofing**

**Abstract:**

Face anti-spoofing is important to avoid face recognition systems from a security breach. In recent years most of the progresses have been made by the availability of face antispoofing benchmark datasets. To perform face anti-spoofing research, a large-scale multi-modal dataset was introduced called CASIASURF, both subjects and visual modalities dataset was largely available for anti-spoofing. The dataset of 1,000 subjects of 21,000 videos and of 3 modalities (i.e., RGB, Depth and IR) in each sample. For developing a new benchmark anti-spoofing, we are provided with a measurement set, evaluation protocol and training/validation/testing subsets. However, we present another multi-modular combination strategy as standard, which performs highlight re-weighting to choose the more useful channel highlights while stifling the less helpful ones for each modal.

**Introduction:**

Face anti-spoofing to parodying expects to decide if the caught face of a face acknowledgment framework is genuine or counterfeit. With the improvement of profound convolutional neural system (CNN), face acknowledgment has accomplished close to consummate acknowledgment execution and as of now has been applied in our everyday life, for example, telephone open, get to control, face instalment, and so on. Be that as it may, these face acknowledgment frameworks are inclined to be assaulted in different manners, including print assault, video replay assault and 2D/3D cover assault, which cause the acknowledgment result to get temperamental. Thusly, face introduction assault location (PAD) is an indispensable advance to guarantee that face acknowledgment frameworks are in a safe dependable condition.

Research questions:

1. How Face Presentation Attack Detection (PAD) achieved success using Face Anti-spoofing Datasets?
2. How Multi-modal Face Presentation Attack can be detected?

**Background:**

A large portion of existing face hostile to parodying datasets just contain the RGB modality. Replay-Attack and CASIAFASD are two broadly utilized PAD datasets. Indeed, even the as of late delivered SiW dataset, gathered with high goal picture quality just contains RGB information. With the far-reaching utilization of face acknowledgment in cell phones, there are additionally some RGB datasets recorded by replaying face video with cell phone, for example, MSU-MFSD, Replay-Mobile and OULU-NPU.

As attack techniques are constantly upgraded, some new types of presentation attacks (PAs) have emerged, e.g.,3D and silicone masks. These are more realistic than traditional 2D attacks. Therefore, the drawbacks of visible cameras are revealed when facing these realistic facemasks. Fortunately, some new sensors have been introduced to provide more possibilities for face PAD methods, such as depth cameras, multi-spectral cameras, and infrared light cameras.

**Methods:**

Before demonstrating some exploratory examination on the dataset, we previously assembled a solid standard strategy. We point at finding a clear engineering that gives great execution in our CASIA-SURF dataset. Hence, we characterize the face hostile to ridiculing issue as a paired order task (counterfeit vs. genuine) and lead the tests dependent on the ResNet-18 arrangement organize. ResNet-18 comprises of five convolutional squares (to be specific res1, res2, res3, res4, res5), a worldwide normal pooling layer and a SoftMax layer, which is a generally shallow system with high arrangement execution.

1. **Naive halfway fusion**

CASIA-SURF is described by multi-methodology (i.e., RGB, Depth, IR) and a key issue is how to combine the corresponding data between the three modalities. We utilize a multi-stream engineering with three subnetworks to study the dataset modalities, in which RGB, Depth and IR information are found out independently by each stream, and afterward shared layers are attached at a highlight learn joint portrayals furthermore, perform collaborated choices. The midway combination is one of the ordinarily utilized combination strategies, which consolidates the subnetworks of various modalities at a later stage, i.e., following the third convolutional square (res3) by means of the element map link. Along these lines, highlights from various modalities can be intertwined to perform characterization. In any case, direct linking these highlights cannot make full utilization of the qualities between various modalities.

1. **Squeeze and excitation fusion**

The three modalities provide with complementary information for different kind of attacks: RGB data have rich appearance details, Depth data are sensitive to the distance between the image plane and the corresponding face, and the IR data measure the amount of heat radiated from a face. Inspired by [19], we propose the squeeze and excitation fusion method that uses the “Squeeze-and-Excitation“ branch to enhance the representational ability of the different modalities’ feature by explicitly modelling the interdependencies between their convolutional channels.

**References:**

1. Samarth Bharadwaj, Tejas I Dhamecha, Mayank Vatsa, and Richa Singh. Computationally efficient face spoofing detection with motion magnification. In CVPR, 2013.
2. Sushil Bhattacharjee, Amir Mohammadi, and Sebastien ´ Marcel. Spoofing deep face recognition with custom silicone masks. In BTAS, 2018.
3. Zinelabidine Boulkenafet, Jukka Komulainen, and Abdenour Hadid. Face spoofing detection using colour texture analysis. TIFS, 2016.
4. Neslihan Kose and Jean-Luc Dugelay. Countermeasure for the protection of face recognition systems against mask attacks. In FG, 2013.
5. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In CVPR, 2016.
6. Artur Costa-Pazo, Sushil Bhattacharjee, Esteban VazquezFernandez, and Sebastien Marcel. The replay-mobile face presentation-attack database. In BIOSIG, 2016.
7. Jukka Maatta, Abdenour Hadid, and Matti Pietikainen. Face spoofing detection from single images using texture and local shape analysis. IET biometrics, 2012.
8. Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. TPAMI, 2002.
9. Sooyeon Kim, Sunjin Yu, Kwangtaek Kim, Yuseok Ban, and Sangyoun Lee. Face liveness detection using variable focusing. In ICB, 2013.
10. Maria De Marsico, Michele Nappi, Daniel Riccio, and JeanLuc Dugelay. Moving face spoofing detection via 3d projective invariants. In ICB, 2012.

.