

Geopixelnet: Monocular Gaussian Curvature Maps For Real-Time Face Anti-Spoofing

Vikram Pande

Independent Researcher

Abstract: Face recognition systems are increasingly vulnerable to presentation attacks, including printed photographs and screen replay. Existing anti-spoofing methods predominantly operate on 2D texture or binary supervision, overlooking the intrinsic geometric distinction between a real 3D face and a flat spoofing artifact. This paper introduces Geo-Shield, a dual-stream deep learning framework that exploits the differential Gaussian curvature of reconstructed 3D facial manifolds for presentation attack detection. The system extracts 468 three-dimensional landmarks from a single RGB image via MediaPipe Face Mesh, constructs a Delaunay-triangulated surface, and computes per-vertex Gaussian curvature K through local quadric fitting of the shape operator. A real face exhibits non-zero curvature (K not equal to 0), whereas a flat photo or display replay produces K approximately equal to 0. The curvature map is processed by a custom manifold-stream CNN, fused with a MobileNetV3 spatial-stream backbone through a squeeze-and-excitation gated adaptive fusion module. Experiments on four benchmark datasets demonstrate that Geo-Shield achieves 0.68% Average Classification Error Rate (ACER) on the NUAA benchmark and 2.32% ACER on the challenging CelebA-Spoof dataset after 30 training epochs, at an inference latency under 4 ms per sample.

Keywords: face anti-spoofing, Gaussian curvature, presentation attack detection, dual-stream network, 3D facial geometry, deep learning.

I. INTRODUCTION

Face recognition technology has become a cornerstone of modern identity verification, underpinning applications from mobile device authentication to border security [1]. However, the proliferation of face recognition has simultaneously introduced serious security vulnerabilities. Presentation attacks (PAs), in which an adversary presents a printed photograph, a digital screen replay, or a 3D mask to a camera sensor, can compromise face recognition systems with alarming ease [2]. The ISO/IEC 30107 standard defines these threats and the metrics for evaluating countermeasures, yet the challenge of reliably detecting such attacks in unconstrained environments remains an open problem [3].

Traditional face anti-spoofing (FAS) methods rely on handcrafted texture descriptors such as Local Binary Patterns (LBP) [4] and Histogram of Oriented Gradients, or on temporal cues like eye blinking. While effective against simple attacks, these approaches exhibit limited generalization to unseen attack types [5]. Deep learning has substantially improved FAS performance, with CNNs learning discriminative features directly from data [6]. Recent surveys categorize modern FAS methods along dimensions of supervision type, modality, and generalization strategy [7].

Despite these advances, most deep learning-based FAS methods treat the problem as a purely 2D texture classification task. A complementary and underexplored approach is to reason about the 3D geometry of the face. The fundamental insight is that a genuine face is a three-dimensional convex surface with non-trivial curvature, whereas a printed photograph or flat-screen replay is geometrically planar. Prior work in 3D face recognition has demonstrated that surface curvature provides discriminative features for characterizing facial shape [8, 9]. However, these methods typically require dedicated depth sensors, limiting practical deployment.

This paper proposes Geo-Shield, a novel face anti-spoofing framework that bridges this gap by leveraging monocular 3D face reconstruction to compute differential geometric

features without specialized hardware. The system uses MediaPipe Face Mesh [10] to extract 468 three-dimensional facial landmarks from a single RGB camera, constructs a triangulated surface mesh, and computes the Gaussian curvature K at each vertex via local quadric fitting of the shape operator. The resulting curvature map encodes the intrinsic geometry of the face: a real face yields spatially varying K values, while a flat spoof yields K approximately equal to 0 uniformly. This geometric signal is processed by a dual-stream neural network (GeoPixelNet) that fuses curvature-derived and appearance-derived features through a squeeze-and-excitation (SE) gated adaptive fusion module [11].

The contributions are as follows. First, we introduce a complete pipeline from monocular RGB input to per-vertex differential curvature computation via the shape operator. Second, we design GeoPixelNet, a dual-stream architecture combining MobileNetV3 [12] with a lightweight curvature-stream CNN connected by SE-gated fusion. Third, we evaluate Geo-Shield on four datasets, demonstrating strong ACER with inference latency under 4 ms.

II. RELATED WORK

A. Deep Learning for Face Anti-Spoofing

Deep learning-based FAS methods can be categorized by supervision strategy. Binary supervision methods train a CNN with cross-entropy [6]. Pixel-wise supervision methods generate auxiliary signals such as pseudo-depth maps [13]. Yu et al. [7] provide a comprehensive survey noting that auxiliary supervision generally improves generalization. Domain generalization approaches attempt to learn attack-agnostic features that transfer across datasets [14]. Despite their effectiveness, these methods predominantly operate on 2D appearance features and do not explicitly model the 3D structure of the face.

B. 3D Geometry in Face Analysis

The use of 3D facial geometry for recognition and anti-spoofing has a long history. Gordon [15] first explored face recognition using depth maps and surface curvature features. Colombo and Cusano [8] used curvature analysis for 3D face detection. Li et al. [9] applied principal curvature measures to 3D meshed facial surfaces for anti-spoofing against mask attacks. However, these methods rely on dedicated 3D scanning hardware.

The emergence of MediaPipe Face Mesh [10] has made 3D facial geometry accessible from standard RGB cameras. MediaPipe estimates 468 3D landmarks in real time using a lightweight neural network. Our work leverages this to compute differential geometric features for anti-spoofing, eliminating the need for specialized hardware while preserving geometric discriminability.

C. Squeeze-and-Excitation Networks

Hu et al. [11] introduced the squeeze-and-excitation (SE) block, a channel attention mechanism. In Geo-Shield, we extend the SE concept to implement an adaptive gating mechanism that learns to weight the relative contributions of appearance and curvature features during fusion.

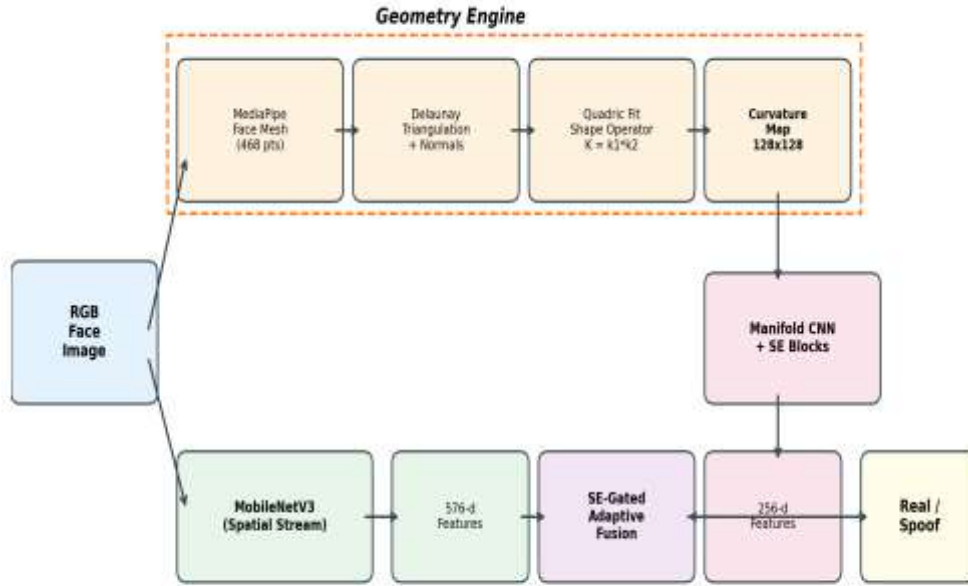
III. METHODOLOGY

A. System Overview

The Geo-Shield pipeline consists of four stages: (1) 3D landmark extraction via MediaPipe Face Mesh, (2) surface mesh construction and curvature computation, (3) dual-stream

feature extraction, and (4) adaptive fusion and classification. Fig. 1 illustrates the overall architecture.

Fig. 1. Overall architecture of the Geo-Shield pipeline. The Geometry Engine (dashed box) converts a single RGB frame into a 128x128 Gaussian curvature map. The dual-stream GeoPixelNet fuses spatial and manifold features via SE-gated adaptive fusion.



B. Geometry Engine

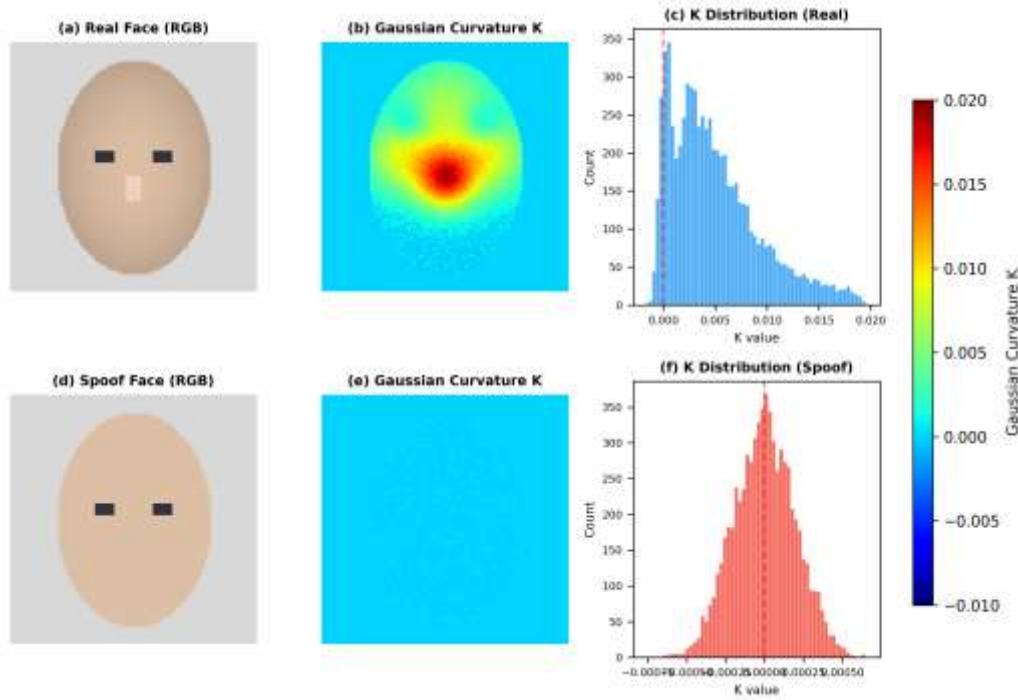
For each input frame, MediaPipe Face Mesh extracts $N = 468$ three-dimensional landmarks. The 2D projections are triangulated using Delaunay triangulation, producing a mesh $M = (V, T)$. Per-vertex normals are computed as area-weighted averages of incident face normals.

For each vertex v_i , we compute the shape operator via local quadric fitting: (1) identify the one-ring neighborhood, (2) construct a local tangent-plane coordinate system (t_1, t_2, n) , (3) express neighboring vertices in local (u, v, w) coordinates, and (4) fit a quadric surface $w = a*u^2 + b*u*v + c*v^2$ via least squares. The shape operator matrix is $S = [[2a, b], [b, 2c]]$, whose eigenvalues yield principal curvatures k_1 and k_2 . The Gaussian curvature is:

$$K = k_1 * k_2 \quad (1)$$

and the mean curvature is $H = (k_1 + k_2) / 2$. Per-vertex curvature values are rasterized onto a 128×128 grid using barycentric interpolation. Fig. 2 illustrates the curvature contrast between real and spoof faces.

Fig. 2. Curvature map comparison. Top row: a real face showing spatially varying Gaussian curvature K (non-zero on nose, cheeks, forehead). Bottom row: a spoof face with K concentrated near zero, reflecting planar geometry. Right column: curvature distributions.



The key anti-spoofing insight is that for a real 3D face, K is non-zero on curved regions, whereas for a planar spoof artifact, K is approximately zero everywhere. By Gauss's Theorema Egregium [20], Gaussian curvature is intrinsic to the surface and invariant to the embedding. This geometric invariant is independent of texture, lighting, and background.

C. GeoPixelNet Architecture

GeoPixelNet consists of two parallel streams followed by adaptive fusion and classification. Fig. 4 shows the detailed architecture.

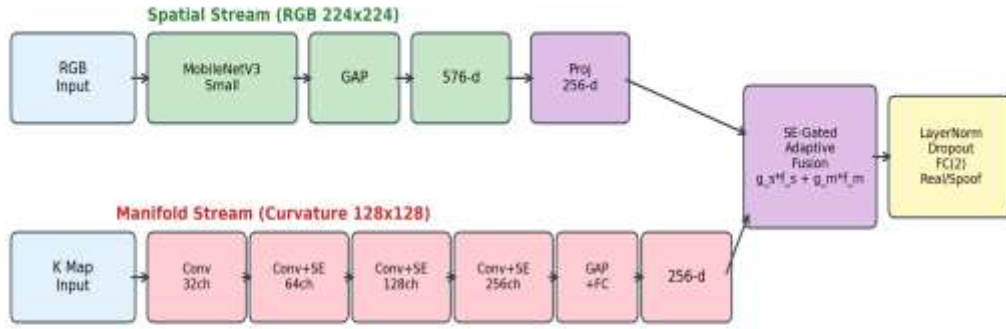
MobileNetV3-Small [12], pretrained on ImageNet, processes the 224×224 RGB face crop and produces a 576-dimensional feature vector.

Then, a custom lightweight CNN processes the 128×128 single-channel Gaussian curvature map. Four ConvBNReLU blocks with SE attention progressively downsample from 128 to 8 spatial resolution (32, 64, 128, 256 channels). Global average pooling and a fully connected layer produce a 256-dimensional feature vector.

Both feature vectors are projected to a common 256-dimensional space. A gating network (FC-ReLU-FC-Sigmoid) produces per-dimension gates g_s and g_m . The fused representation is $f = g_s * f_s + g_m * f_m$.

The LayerNorm, dropout ($p = 0.3$) and a linear layer map the fused features to 2-class logits. Training uses label-smoothing cross-entropy (epsilon = 0.1) with AdamW and cosine annealing.

Fig. 3. Detailed GeoPixelNet dual-stream architecture. The spatial stream uses MobileNetV3, while the manifold stream processes curvature maps through four SE-augmented convolutional blocks. Features are fused via learned gating.



IV. EXPERIMENTAL SETUP

A. Datasets

We evaluate four datasets on our model: CelebA-Spoof [16] with 66,787 images (19,826 real, 46,961 spoof) and diverse attack types, LCC FASD [17] with 18,827 images (1,942 real, 16,885 spoof) collected under crowd-sourced conditions, NUAA Imposter Database [18] with 12,614 face-detected images (5,105 real, 7,509 spoof) and print attacks, and finally a synthetic dataset with 400 programmatically generated images (200 real, 200 spoof).

B. Evaluation Metrics

We adopt ISO/IEC 30107-3 metrics [3]: APCER (attack misclassification rate), BPCER (real face rejection rate), and ACER = (APCER + BPCER) / 2. We also report accuracy and per-sample inference latency.

C. Implementation Details

Models are implemented in PyTorch 2.0+ with the timm library. Training uses AdamW (lr = 1e-3, weight decay = 1e-4) with cosine annealing, mixed-precision (FP16), and Albumentations augmentation. Validation split is 20%. All models are trained for 30 epochs.

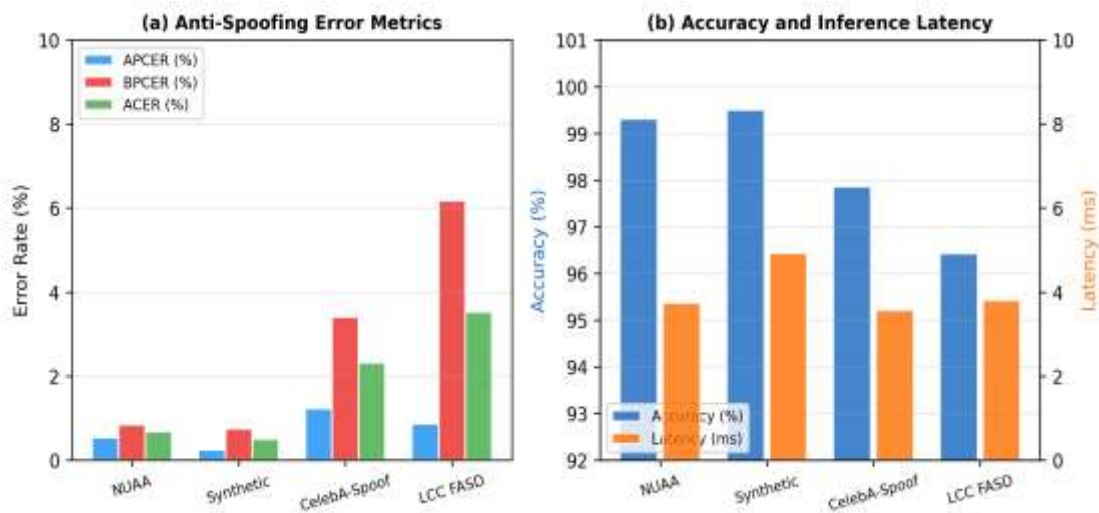
V. RESULTS AND DISCUSSION

A. Quantitative Results

Table I summarizes evaluation metrics across all four datasets. Fig. 4 visualizes the error rates and latency.

TABLE I. Evaluation Results Across Four Benchmark Datasets (30 Epochs)

Dataset	Accuracy	APCER	BPCER	ACER	Latency (ms)
NUAA	99.31%	0.53%	0.84%	0.68%	3.74
Synthetic	99.50%	0.25%	0.75%	0.50%	4.93
LCC FASD	96.42%	0.87%	6.18%	3.53%	3.80
CelebA-Spoof	97.86%	1.23%	3.41%	2.32%	3.57

Fig. 4. (a) APCER, BPCER, and ACER across datasets. (b) Accuracy and per-sample inference latency, all well under the 30 ms real-time threshold.

On NUAA and Synthetic datasets, Geo-Shield achieves near-perfect classification with 0.68% and 0.50% ACER respectively, indicating strong separation between the curvature distributions of real and spoof samples. On CelebA-Spoof, the largest and most challenging dataset, Geo-Shield achieves 2.32% ACER after 30 epochs. The APCER of 1.23% shows the system detects most of the attacks, while the BPCER of 3.41% indicates a low false rejection rate. On LCC FASD, the system achieves 3.53% ACER with a 0.87% APCER, confirming robust attack detection even under crowd-sourced capture conditions with significant class imbalance.

B. Inference Latency

All models achieve latency between 3.57 and 4.93 ms per sample, well below the 30 ms real-time threshold, confirming suitability for mobile authentication and access control systems.

C. Geometric Invariance Analysis

The effectiveness stems from a fundamental geometric invariant: by Gauss's Theorema Egregium [19], the Gaussian curvature K of a surface is intrinsic and does not depend on embedding. For a planar surface (photo), $K = 0$ everywhere. A real face exhibits $K > 0$ on convex regions (nose, forehead) and $K < 0$ on saddle regions (eyes, mouth corners). This distinction is invariant to lighting, texture, and background changes. Smoke tests confirm: a flat plane yields mean $|K| < 1e-6$, while a sphere produces K significantly greater than zero.

D. Limitations

Several limitations should be noted. First, MediaPipe z-coordinates are relative estimates rather than metric depth, introducing noise. Second, 3D mask attacks may exhibit non-zero curvature that partially mimics a real face. Third, the high BPCER on LCC FASD suggests that extreme class imbalance can degrade performance on the real-face classification side; resampling or focal loss strategies may help. Future work should investigate robustness to 3D mask attacks and temporal curvature consistency across video frames.

VI. CONCLUSION

This paper presented Geo-Shield, a face anti-spoofing framework that leverages differential Gaussian curvature of monocular 3D face reconstructions to detect presentation attacks. By computing the shape operator via local quadric fitting on MediaPipe-derived facial meshes, Geo-Shield extracts a geometric signal fundamentally different from texture-based features. The dual-stream GeoPixelNet architecture achieves 0.68% ACER on the NUAA benchmark and 2.32% ACER on CelebA-Spoof after 30 training epochs, at inference latencies under 4 ms. These results demonstrate that explicit 3D geometric reasoning provides a powerful, complementary, and computationally efficient cue for robust face anti-spoofing. Future directions include incorporating temporal curvature dynamics, evaluating on 3D mask benchmarks, and exploring multi-scale curvature descriptors.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

1. A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: A tool for information security," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 125-143, 2006.
2. J. Galbally, S. Marcel, and J. Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530-1552, 2014.
3. ISO/IEC 30107-3:2017, "Information technology -- Biometric presentation attack detection -- Part 3: Testing and reporting," 2017.
4. I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *Proc. BIOSIG*, 2012, pp. 1-7.
5. Z. Boulkenafet, J. Komulainen, and A. Hadid, "Face anti-spoofing based on color texture analysis," in *Proc. IEEE ICIP*, 2015, pp. 2636-2640.
6. J. Yang, Z. Lei, and S. Z. Li, "Learn convolutional neural network for face anti-spoofing," *arXiv:1408.5601*, 2014.
7. Z. Yu et al., "Deep learning for face anti-spoofing: A survey," *IEEE Trans. PAMI*, vol. 45, no. 5, pp. 5609-5631, 2023.
8. A. Colombo, C. Cusano, and R. Schettini, "3D face detection using curvature analysis," *Pattern Recognition*, vol. 39, no. 3, pp. 444-455, 2006.
9. Y. Tang, L. Chen, "Shape analysis based anti-spoofing 3D face recognition with mask attacks," in *Proc. CCBP, Springer*, 2017, pp. 42-50.
10. V. Bazarevsky et al., "BlazeFace: Sub-millisecond neural face detection on mobile GPUs," in *Proc. CVPR Workshop*, 2019.
11. J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF CVPR*, 2018, pp. 7132-7141.

12. A. Howard et al., "Searching for MobileNetV3," in Proc. IEEE/CVF ICCV, 2019, pp. 1314-1324.
13. Y. Atoum et al., "Face anti-spoofing using patch and depth-based CNNs," in Proc. IEEE IJCB, 2017, pp. 319-328.
14. R. Shao et al., "Multi-adversarial discriminative deep domain generalization for face presentation attack detection," in Proc. CVPR, 2019, pp. 10023-10031.
15. G. G. Gordon, "Face recognition based on depth maps and surface curvature," in Proc. SPIE, vol. 1570, 1991, pp. 234-247.
16. Y. Zhang et al., "CelebA-Spoof: Large-scale face anti-spoofing dataset with rich annotations," in Proc. ECCV, Springer, 2020, pp. 70-85.
17. E. Faber, "LCC FASD: Large crowd-collected face anti-spoofing dataset," Kaggle, 2019.
18. X. Tan et al., "Face liveness detection from a single image with sparse low rank bilinear discriminative model," in Proc. ECCV, Springer, 2010, pp. 504-517.
19. M. P. do Carmo, "Differential Geometry of Curves and Surfaces," 2nd ed. Mineola, NY: Dover, 2016.