# 3D Face Reconstruction with Feature Enhancement using Bi-FPN for Forensic Analysis

[1]Sincy John and [2]Ajit Danti
[1,2] Department of Computer Science and Engineering, Christ University, Bangalore, India.
[1]sincyjohn2@gmail.com, [2]ajitdanti@yahoo.com

Correspondence should be addressed to Sincy John: sincyjohn2@gmail.com

**Article Info**
Journal of Machine and Computing (http://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202404037
Received 02 September 2023; Revised from 20 January 2024; Accepted 11 February 2024.
Available online 05 April 2024.


**Abstract –** The representation of facial features in three-dimensional space plays a pivotal role in various applications such as facial recognition, virtual reality, and digital entertainment. However, achieving high-fidelity reconstructions from two-dimensional facial images remains a challenging task, particularly in preserving fine texture details. This research addresses this problem by proposing a novel approach that leverages a combination of advanced techniques, including Resnet, Flame model, Bi-FPN, and a differential render architecture. The primary objective of this study is to enhance texture details in reconstructed 3D facial images. The integration of Bi-FPN (Bi-directional Feature Pyramid Network) enhances feature extraction and fusion across multiple scales, facilitating the preservation of texture details across different regions of the face. The objective is to accurately represent facial features from 2D images in three-dimensional space. By combining these methods, the proposed framework achieves significant improvements in preserving fine texture details and overall facial structure. Experimental results demonstrate the effectiveness of the approach, suggesting its potential for various applications such as virtual try-on and facial animation.

**Keywords -** Bi-FPN, 3D Face Reconstruction, Flame Model, Resnet.

## I. INTRODUCTION

In the realm of computer vision, the task of reconstructing three-dimensional (3D) facial representations from single 2D images poses a significant challenge, yet holds immense potential across diverse fields such as facial recognition, augmented reality, and virtual communication [1,2]. While traditional image processing predominantly relies on 2D images, the incorporation of depth information through 3D reconstructions has the capacity to enhance accuracy and realism. However, conventional methods for 3D face reconstruction often encounter limitations due to their reliance on handcrafted features and geometric constraints. These approaches may struggle to capture the intricate nuances and subtle variations inherent in facial features [16].

Traditional methods for 3D face reconstruction often rely on handcrafted features and geometric constraints. These methods can be effective, but they can also be limited in their ability to capture the complex and subtle variations in facial features [3]. In response to these challenges, deep learning techniques have emerged as a formidable solution, offering superior performance and generalization capabilities. By leveraging the vast amount of data available and the ability to learn complex patterns, deep learning models have demonstrated promising results in various computer vision tasks, including 3D face reconstruction [14,15]. Nonetheless, while deep learning models excel in capturing intricate details, they may still face limitations in certain scenarios. To address these challenges and capitalize on the strengths of both traditional methods and deep learning techniques, hybrid models have gained traction. These models combine multiple deep learning architectures, harnessing their complementary strengths and mitigating their individual weaknesses. By employing ensemble techniques, enhance robustness, generalization, and overall reconstruction accuracy, thereby pushing the boundaries of what is achievable in 3D face reconstruction. Ensemble models, which combine multiple deep learning models, have proven to be particularly effective in various computer vision tasks, including 3D face reconstruction. Ensemble models can effectively address the limitations of individual models by leveraging their strengths and mitigating their weaknesses. This approach leads to improved robustness, generalization, and overall reconstruction accuracy.

The main goal of this study is to enhance texture details in reconstructed 3D facial images. To achieve this, we employ a Resnet-based model to extract feature representations from input images, capturing both high-level semantic information and fine-grained texture details. Subsequently, the Flame model is utilized to model facial geometry and expressions, enabling a more accurate representation of facial structure. The integration of Bi-FPN (Bi-directional Feature Pyramid

Network) further enhances feature extraction and fusion across multiple scales, facilitating the preservation of texture details across different regions of the face. Then a differential Render, which accurately translates 2D facial images into their corresponding 3D representations. This architecture incorporates differential rendering techniques to handle complex facial geometries and ensure faithful reconstruction of texture details. By combining these components, our proposed framework aims to significantly improve the quality of reconstructed 3D facial images, enabling more realistic and detailed representations.

The following are the contributions made by this paper:

- We propose a novel ensemble model for 3D face reconstruction from 2D images, combining three deep learning architectures:
- We develop an effective fusion strategy Bi-FPN to enhance the texture details, which result in efficient 3d face reconstruction.
- Experiment evaluations show that the proposed model have low Normalized mean error (NME), Mean and Standard deviation compared to existing state-of-the-art methods.

## II. RELATED WORKS

Advancements in computer vision and facial recognition technologies have spurred the development of innovative approaches for 3D face reconstruction. This survey provides an overview of various methods proposed in recent literature, each offering unique strategies and techniques to address the complexities inherent in reconstructing accurate and realistic 3D facial models.

The proposed method utilizes web-sourced images to reconstruct a 3D face mesh with 468 landmarks, extracting various distance metrics for comparison. This approach, coupled with hyper-parameter optimization of diverse machine learning classifiers, achieves promising accuracy (up to 78% with Extreme Gradient Boosting) in identifying faces. However, effectively distinguishing between highly similar faces remains a challenge, requiring further research. Future work will focus on building a multi-modal facial recognition system incorporating additional data modalities to tackle this challenge. Additionally, expanding the dataset and refining the classifier selection hold potential for further accuracy improvements [1].

Proposed an innovative "exemplar-coherent" approach that leverages an external face database. Facial Part Composition in which the face into individual components like eyes, nose, and mouth. These parts are then "composed" from the external database, ensuring they match the specific characteristics of the mugshot image. formulate an energy function incorporating three key cues: Shape from Shading, Multi-view Colour Consistency and Depth Smoothness Prior. Optimization and Segmentation is done by solving the energy function by first estimating shading parameters and then converting it to a multi-view image segmentation problem. This optimizes the reconstruction process while maintaining accuracy. Advantage of the work is that this method addresses challenging uncalibrated mugshots and by incorporating an external face database, the reconstructed features are realistic and consistent with the individual's appearance. Disadvantage of this methods are it requires more computational resources compared to simpler reconstruction techniques and external database reliance: the accuracy of the reconstruction depends on the quality and diversity of the external face database [2].

This paper introduces a regression-based approach for 3D face reconstruction that utilizes pre-defined facial landmarks and cascaded regressors. The method uses facial landmarks pre-defined on the input face image. These landmarks are crucial points on the face, and the approach aims to adjust the initial 3D face shape based on deviations between these landmarks and corresponding landmarks obtained from the reconstructed 3D faces. The cascaded regressors are pre-trained on a dataset of paired 3D faces and 2D images from various viewpoints, with corresponding landmark annotations. Advantages of this paper is high computational efficiency, Improved accuracy and Unified handling of arbitrary views. The disadvantages include dependence on Manually Annotated Landmarks and single view reconstruction [4].

This paper introduces a learning-based approach for 3D facial reconstruction using a CNN. To address the challenge of limited 3D face data for training, the proposed method suggests generating synthetic facial images with known geometric forms. The model demonstrates success in recovering facial shapes from real images, showcasing its robustness in handling diverse facial expressions and lighting conditions. Limitation includes Ethnic Bias in 3DMM Model and Issues with Unseen Facial Attributes [5].

This paper highlights the use of a UV position map for 3D facial structure representation, end-to-end learning with a CNN, integration of a weight mask for improved performance, independence from prior face models, and a lightweight network architecture with exceptional processing speed [6].

This paper highlights a significant advancement in 3D face reconstruction by addressing the data scarcity challenge and proposing a novel method that effectively combines hybrid loss function with multi-image information. Pros include hybrid-level image information (combining low-level and perception-level cues) for learning without requiring ground-truth 3D shapes. The paper mainly focuses on reconstruction accuracy and robustness. Other aspects like texture detail, fine-grained facial features, and handling extreme situations like heavy occlusions might require further improvement [7].
This paper focuses on face detection and landmark localization called the Deformable Hough-Transform Model (DHTM). Instead of using a separate face detector to initialize facial landmarks, the proposed method employs cascaded regression for both face detection and landmark localization. The PO-CR (Part-Ordinary Cascade Regression) algorithm is used to fit a facial deformable model to the image. The proposed method are efficient and robust face detection and landmark

localization, with a specific emphasis on improved landmark localization performance. Disadvantage includes Computational Complexity and Training Data Dependency [8].

The paper proposes a new solution to the challenges in face alignment through a novel alignment framework called 3D Dense Face Alignment (3DDFA). In 3DDFA, a dense 3D face model is fitted to the image using convolutional neural networks (CNNs). This approach helps address issues related to self-occlusion in modeling and high nonlinearity in fitting, suggesting a more robust and accurate representation of facial features. The authors have suggested that more complex network architecture, such as bigger input sizes and deeper network architecture can improve 3DDFA [9].

The proposed approach focuses on simultaneously extracting the 3D shape of the face and achieving semantically consistent 2D alignment using a 3D Spatial Transformer Network (3DSTN). The 3DSTN is used to model both the camera projection matrix and the warping parameters of a 3D model. This indicates a comprehensive approach that integrates both geometric and appearance aspects. Disadvantage is that limited data used in the generation of a 3DMM implies that the model's ability to represent unseen face shapes may be constrained. If the training data is not diverse enough, the 3DMM may struggle to accurately capture the full range of facial variations [10].

This paper focus on the shift from sparse to dense 3D alignment in face alignment, introducing a comprehensive approach that includes face contours and SIFT feature points. The method addresses challenges in training CNNs with multiple datasets. The limitation is highlighted in terms of the use of shape constraints, specifically 3DMM. The implication is that the employed shape constraints might not be as effective or advantageous when dealing with large-pose face alignment challenges [11].

This paper maximizes the use of 2D face images for 3D face model learning. The approach employs self-supervised learning, introduces multiple forms of self-supervision, incorporates self-critic learning for model improvement, and demonstrates excellent performance on tasks of 3D face reconstruction and dense face alignment. One of the limitations is that the method relies on sparse 2D facial landmarks as input for the CNN regressor. The accuracy and reliability of 3D face reconstruction and dense face alignment may be limited by the quality and precision of these sparse landmarks [12].
After going through the literature survey, the following research gap was found, reliance on manually annotated landmarks can be time-consuming and error prone. Limited accuracy for texture and fine details Limitations of single-view reconstruction and computational complexity and training requirements

## III. PROPOSED METHODOLOGY

The primary aim of this research is to enhance the texture details by deploying Bi-FPN and reconstructing the 3D facial images. For this, this research deploys a Resnet, Flame model, Bi-FPN and a differential render architecture which accurately represents the 3d representation of the face image from a 2d face. The stages involved in the proposed approach are shown in **Fig 1** and are discussed in the below subsections:

In **Fig 1** the input image is pre-processed and given as input to the machine learning model in this resnet 50 model is used for generating course and detail image. ResNet-50 [13] is a deep convolutional neural network architecture commonly used for image classification and feature extraction. Resnet is used to extract feature representations from input facial images. It is to capture both high-level semantic information (such as facial structures and features) and fine-grained texture details (such as wrinkles, pores, etc.).

The Flame model [17] is utilized to model facial geometry and expressions. It helps in creating a more accurate representation of facial structure by capturing details related to the shape of the face, including variations in facial expressions. Output from the resnet and flame model is used to generate the texture details, course shape and displacement map. The albedo map, which represents the surface color of the 3D face model, is enhanced using Bi-FPN (Feature Pyramid Network). Bi-FPN is a feature enhancement technique that likely helps improve the quality of the albedo map, capturing more detailed features. Enhanced albedo map along with course shape and displacement map is given to the differential render.

The displacement map allows us to generate images with mid-frequency surface details. To reconstruct the detailed geometry D′, we convert D and its surface normal $N$ to UV space, and combine them with M as

$$D'_{uv} = D_{uv} + M \odot N_{uv} \tag{1}$$

Where M is the is the detail decoder

$$M = F_d(\delta, \psi, \theta_{jaw}) \tag{2}$$

By calculating normal $N'$ from D′, we obtain the detail rendering $Q'_r$ by rendering D with the applied normal map, where D is the course shape, B is textured shape, and c is the camera parameter.

B is the function of $A_{original}$ (albedo), l(Light), $N'$ is the surface normal.

$$Q'_r = R(D, B(A_{original}, l, N'), c) \tag{3}$$

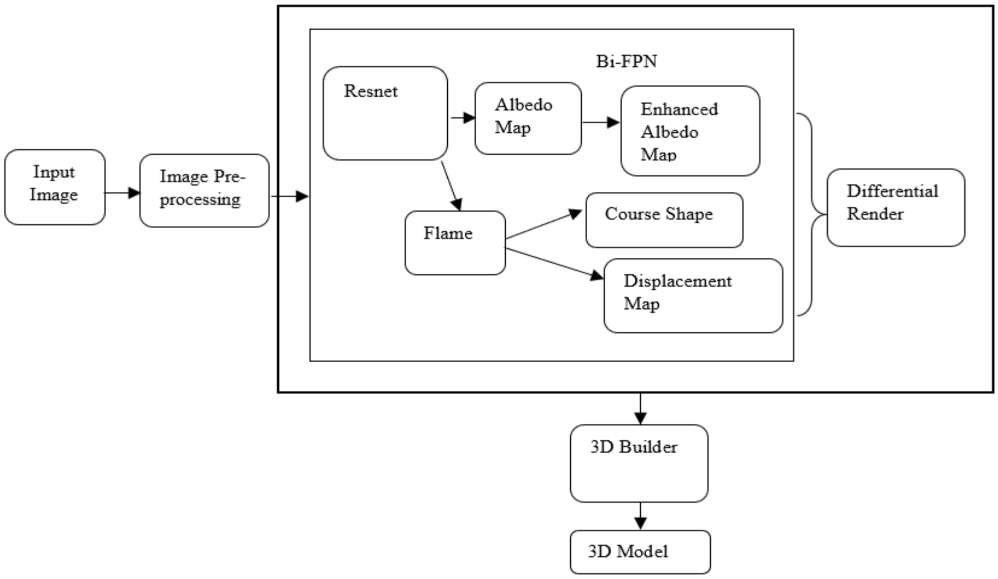


**Fig 1**. Proposed Methodology

***Algorithm 1.*** *Algorithm for 3D Reconstruction*

Input: Images from AFLW300 database
Step 1: Start
Step 2: Read the input image.
Step 3: Re-sizing the input image.
Step 4: Passing the processed image to Resnet 50 for generating course and detail image.
Step 5: Splitting the course code to c, α, l, β, θ, ψ.
Step 6: α is used to generate albedo map.
Step 7: β, θ, ψ used to generate course shape.
Step 8: δ used to generate displacement map.
Step 9: The generated albedo map, course shape and displacement map to the differential render.
Step 10: The output from the differential encoder is passed to the 3D builder to visualize the 3D model
Step 11: Stop

*Feature Enhancement*

Bi-FPN combines features from multiple levels of the pyramid, enabling the network to capture context and details at various scales. This multilevel feature fusion enhances the overall representation of the input. The bidirectional connections in Bi-FPN ensure that information from both high and low-resolution levels is effectively utilized. This helps in enhancing features with both semantic richness and fine-grained details.

***Algorithm 2.*** *Algorithm for Feature Enhancement*

Input: Albedo Code
Step 1: Start
Step 2: Read the Albedo code.
Step 3: Split the albedo code.
Step 4: Passing the split code to Bi-FPN.
Step 5: Multiscale feature fusion is done by bi-directional cross scale connection and weighted feature fusion
Step 6: Enhanced albedo map.
Step 7: Stop

Integrate Bi-FPN into the network architecture to enhance features. Let $F_{enhanced}$represent the feature map after applying Bi-FPN. Combine the enhanced features with the original albedo map to obtain the enhanced albedo map:

$$T_{enhanced} = Enhanced\ Albedo(A_{original}, F_{enhanced}) \tag{4}$$

The function Enhance Albedo incorporates the information from the enhanced feature map to improve the original albedo map. The specific details of this enhancement will depend on the chosen method and the characteristics of the albedo map.

Once the texture details have enhanced the detailed rendering equation can be written as:

$$Q'_r = R(D, B(T_{enhanced}, l, N'), c) \tag{5}$$

## IV. EXPERIMENTAL ANALYSIS &RESULTS

Experimental analysis was done on AFLW2000-3D [21] is an extension of the AFLW dataset, providing three-dimensional facial images with 68 points of articulation. It includes depth information, enhancing its utility for tasks like facial recognition and expression analysis. With annotations capturing key facial features, it offers a realistic representation of facial geometry and expression.

*Data Preparation:* To make data appropriate for creating and improving machine learning models, preprocessing is done. The facial input images are in the RGB format, it is then converted to grey scale images. After this Haar cascade filter is utilized to detect the face, it helps in accurately locating faces within images using a pre-trained cascade classifier that efficiently identifies facial features. The images are smoothed using a mean filter, which reduces noise and smoothness features by averaging pixel values within a narrow region. This increases the model's resistance to fluctuations. Because of their differences in size and orientation, they are resized or scaled using a scaling factor. For face detection in this work, a scaling factor of 1.2 is employed.

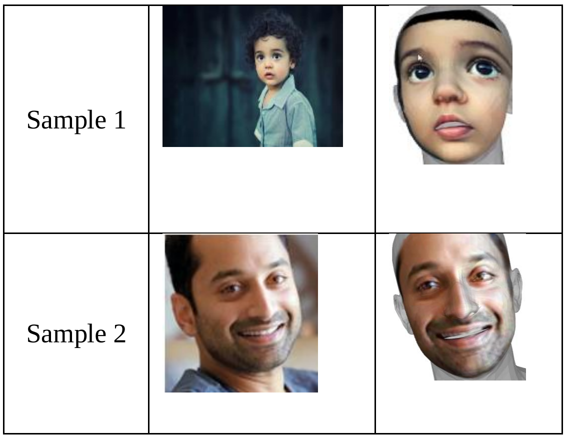


**Fig 2**. 3D Reconstructed image from 2D Image

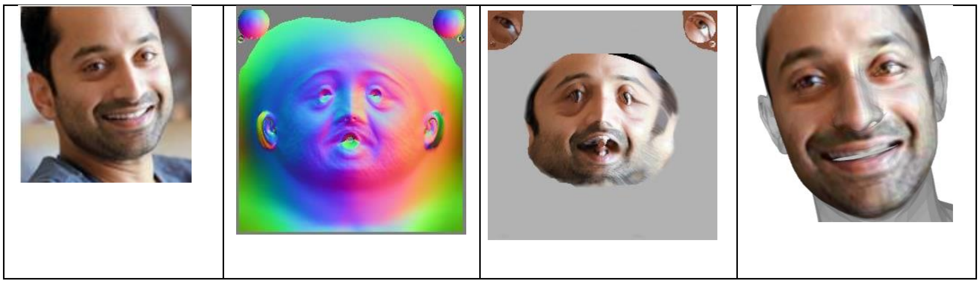

**Fig 3**. 3D Reconstruction with The Enhanced Albedo Map

**Fig 2** depicts 3D reconstruction of the respective 2D image. **Fig 3** portrays 2D image, displacement map, enhanced albedo map and 3D reconstruction of the same.

**Table 1.** Quantitative Analysis of 3D Face Reconstruction Methods using NME

| Methods | NME |
|---|---|
| **3DDFA [21]** | 6.56 |
| **DeFA [11]** | 6.04 |
| **PRNet [6]** | 4.41 |
| **3DDFA-V2 [18]** | 4.2 |
| **Propose Method** | 4.04 |

The **Table 1** compares the performance of different methods for 3D face reconstruction based on their Normalized Mean Error (NME) scores. The proposed method achieves the lowest NME score of 4.04, indicating the smallest average error in 3D face reconstruction. 3DDFA [21] method has the highest NME score of 6.56, indicating the largest average error in 3D face reconstruction. **Fig 4** represent the graphical representation for the same with x axis as methods and y axis as NME.

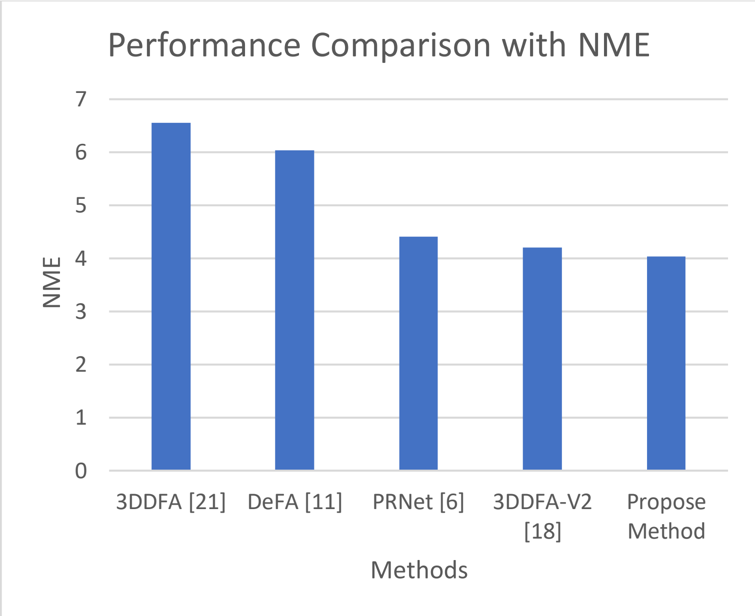


**Fig 4**. Performance Comparison With Existing Methods Using NME

The analysis compares the performance of various methods for 3D face reconstruction based on their mean values. The proposed method achieves the lowest mean of 3.02, indicating superior accuracy compared to other methods listed in the table. PRNnet and 2DASL also perform well, with mean of 3.62 and 3.53, respectively. In contrast, SDM exhibits higher error, with a mean value of 6.12. Overall, the analysis shows that, in comparison to other methods, the suggested method is effective in obtaining high accuracy in 3D face reconstruction. **Fig 5** depicts the graphical representation. **Table 2** Show in Quantitative Analysis of 3D Face Reconstruction Methods using Mean.

**Table 2.** Quantitative Analysis of 3D Face Reconstruction Methods using Mean

| Methods | Mean |
|---|---|
| **SDM [8]** | 6.12 |
| **3DDFA [9]** | 5.42 |
| **3DDFA + SDM [9]** | 4.94 |
| **3DSTN [10]** | 4.49 |
| **DeFA [11]** | 4.50 |
| **PRNnet [6]** | 3.62 |
| **2DASL[12]** | 3.53 |
| **Proposed Method** | 3.02 |

**Table 3.** Quantitative Analysis of 3D Face Reconstruction Methods using Standard Deviation

| Methods | Std |
|---|---|
| **3DDFA-V2 [18]** | 1.391 |
| **RingNet [19]** | 1.306 |
| **DAD-3DNet [20]** | 1.285 |
| **Proposed Method** | 1.087 |

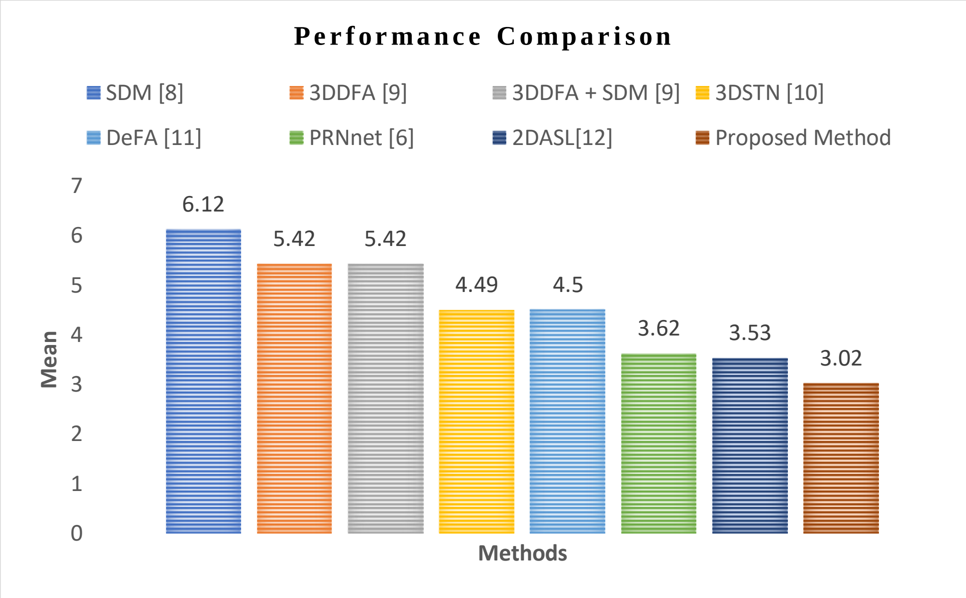


**Fig 5**. Performance comparison with existing methods using Mean

**Table 3** Show in Quantitative Analysis of 3D Face Reconstruction Methods using Standard Deviation. The proposed method stands out with a lower standard deviation of 1.087 compared to the other methods. A lower standard deviation suggests that the performance of this method is more consistent or less variable compared to the others. The standard deviation indicates the variability or spread of the performance scores around the mean. Lower standard deviation values indicate less variability and more consistent performance. **Fig 6** portrays the performance comparison of existing methods based on standard deviation.

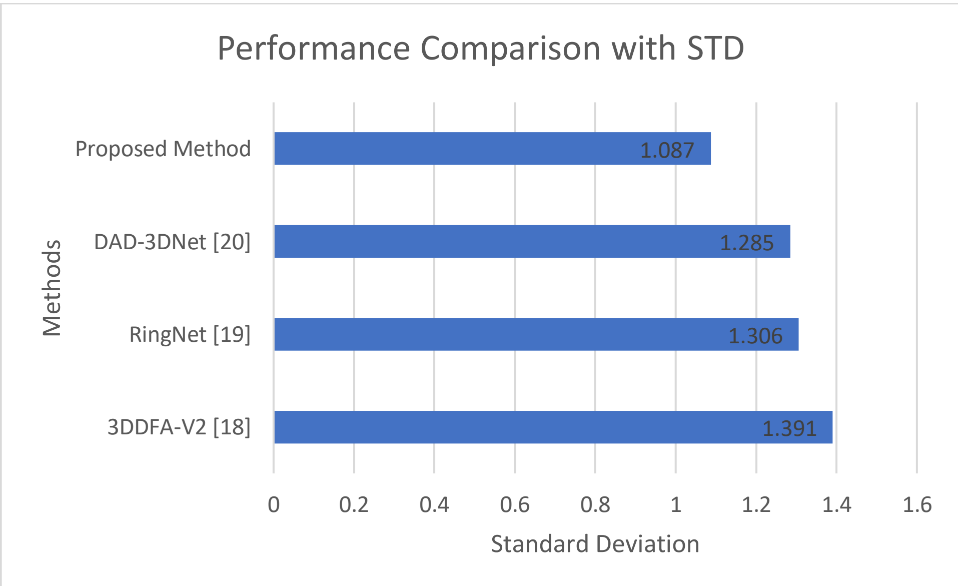


**Fig 6**. Performance Comparison With Existing Methods Using Standard Deviation

## V. CONCLUSION

In conclusion, this research makes notable contributions to the field of 3D face reconstruction by leveraging advanced techniques and addressing key challenges in preserving texture details. The proposed framework offers promising avenues for future research and application development in areas such as facial recognition, augmented reality, and virtual communication. This research proposes a novel approach that integrates advanced techniques such as Resnet, Flame model, Bi-FPN, and a differential render architecture to address this challenge. According to experimental data, the suggested method performs better than previous methods and achieves the lowest NME of 4.04, mean of 3.02, and standard deviation of 1.087 in 3D face reconstruction. This underscores the efficacy of the proposed framework in achieving high accuracy and realism in reconstructed 3D facial images compared to traditional methods and other state-of-the-art approaches.

### Data Availability
No data was used to support this study.

### Conflicts of Interests
The author(s) declare(s) that they have no conflicts of interest.

### Funding
No funding agency is associated with this research.

### Competing Interests
There are no competing interests.

### References
[1] G. Sanil, K. Prakash, S. Prabhu, V. C. Nayak, and S. Sengupta, “2D-3D Facial Image Analysis for Identification of Facial Features Using Machine Learning Algorithms With Hyper-Parameter Optimization for Forensics Applications,” IEEE Access, vol. 11, pp. 82521–82538, 2023, doi: 10.1109/access.2023.3298443.
[2] D. Zeng, Q. Zhao, S. Long, and J. Li, “Examplar coherent 3D face reconstruction from forensic mugshot database,” Image and Vision Computing, vol. 58, pp. 193–203, Feb. 2017, doi: 10.1016/j.imavis.2016.03.001.
[3] W. N. Widanagamaachchi and A. T. Dharmaratne, “3D Face Reconstruction from 2D Images,” 2008 Digital Image Computing: Techniques and Applications, 2008, doi: 10.1109/dicta.2008.83.
[4] Liu, F., Zeng, D., Li, J., Zhao, Q.: Cascaded regressor based 3D face reconstruction from a single arbitrary view image. arXiv preprint (2015), arXiv:1509.06161.
[5] E. Richardson, M. Sela, and R. Kimmel, “3D Face Reconstruction by Learning from Synthetic Data,” 2016 Fourth International Conference on 3D Vision (3DV), Oct. 2016, doi: 10.1109/3dv.2016.56.
[6] Y. Feng, F. Wu, X. Shao, Y. Wang, and X. Zhou, “Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network,” Lecture Notes in Computer Science, pp. 557–574, 2018, doi: 10.1007/978-3-030-01264-9_33.
[7] Y. Deng, J. Yang, S. Xu, D. Chen, Y. Jia, and X. Tong, “Accurate 3D Face Reconstruction With Weakly-Supervised Learning: From Single Image to Image Set,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun. 2019, doi: 10.1109/cvprw.2019.00038.
[8] J. McDonagh and G. Tzimiropoulos, “Joint Face Detection and Alignment with a Deformable Hough Transform Model,” Computer Vision – ECCV 2016 Workshops, pp. 569–580, 2016, doi: 10.1007/978-3-319-48881-3_39.
[9] X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li, “Face Alignment Across Large Poses: A 3D Solution,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, doi: 10.1109/cvpr.2016.23.
[10] C. Bhagavatula, C. Zhu, K. Luu, and M. Savvides, “Faster than Real-Time Facial Alignment: A 3D Spatial Transformer Network Approach in Unconstrained Poses,” 2017 IEEE International Conference on Computer Vision (ICCV), Oct. 2017, doi: 10.1109/iccv.2017.429.
[11] Y. Liu, A. Jourabloo, W. Ren, and X. Liu, “Dense Face Alignment,” 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Oct. 2017, doi: 10.1109/iccvw.2017.190.
[12] X. Tu et al., “3D Face Reconstruction From A Single Image Assisted by 2D Face Images in the Wild,” IEEE Transactions on Multimedia, vol. 23, pp. 1160–1172, 2021, doi: 10.1109/tmm.2020.2993962.
[13] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, doi: 10.1109/cvpr.2016.90.
[14] P. Dou, S. K. Shah, and I. A. Kakadiaris, “End-to-End 3D Face Reconstruction with Deep Neural Networks,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017, doi: 10.1109/cvpr.2017.164.
[15] E. Richardson, M. Sela, R. Or-El, and R. Kimmel, “Learning Detailed Face Reconstruction from a Single Image,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017, doi: 10.1109/cvpr.2017.589.
[16] E. Richardson, M. Sela, and R. Kimmel, “3D Face Reconstruction by Learning from Synthetic Data,” 2016 Fourth International Conference on 3D Vision (3DV), Oct. 2016, doi: 10.1109/3dv.2016.56.
[17] T. Li, T. Bolkart, M. J. Black, H. Li, and J. Romero, “Learning a model of facial shape and expression from 4D scans,” ACM Transactions on Graphics, vol. 36, no. 6, pp. 1–17, Nov. 2017, doi: 10.1145/3130800.3130813.
[18] J. Guo, X. Zhu, Y. Yang, F. Yang, Z. Lei, and S. Z. Li, “Towards Fast, Accurate and Stable 3D Dense Face Alignment,” Lecture Notes in Computer Science, pp. 152–168, 2020, doi: 10.1007/978-3-030-58529-7_10.
[19] S. Sanyal, T. Bolkart, H. Feng, and M. J. Black, “Learning to Regress 3D Face Shape and Expression From an Image Without 3D Supervision,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2019, doi: 10.1109/cvpr.2019.00795.
[20] T. Martyniuk, O. Kupyn, Y. Kurlyak, I. Krashenyi, J. Matas, and V. Sharmanska, “DAD-3DHeads: A Large-scale Dense, Accurate and Diverse Dataset for 3D Head Alignment from a Single Image,” 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2022, doi: 10.1109/cvpr52688.2022.02027.
[21] X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li, “Face Alignment Across Large Poses: A 3D Solution,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, doi: 10.1109/cvpr.2016.23.