Face anti spoofing - A Comparative Analysis between Pixel wise supervision using DenseNet 161 and CDCN ++

Dr.G.Sornavalli
H Varsha
Sneha R
Information Technology Department,
Sri Sivasubramaniya Nadar College of Engineering
SH 49A, Kalavakkam, Tamil Nadu 603110

Abstract - This research compares the DenseNet 161 architecture with CDCN++ in terms of pixel-wise supervision, and then presents a strong face anti-spoofing approach that uses this technique. Feature reuse is promoted by DenseNet-161 by means of feed-forward connections between all of the layers inside a dense block. Feature maps are so well-connected that the layers above them may directly reuse them, which boosts the network's information flow and gradient propagation. Because DenseNet-161 often uses fewer parameters than competing topologies, it achieves better parameter efficiency. The CDCN++ module is a deep learning architecture that was built with face anti-spoofing applications in mind. Using the supplied face photos, it generates depth maps or depth images. With the use of depth information, we can identify more telltale signs that distinguish between actual and synthetic faces. The depth co-occurrence features are extracted from the input depth photos by the CDCN++ module so that spatial connections and patterns may be captured in the depth data. Along with enhancing the model's understanding of fine-grained facial patterns, the pixel-wise supervision approach aids the model in identifying and resolving any issues with the current anti-spoofing methods. Overall, the topic of face anti-spoofing is significantly advanced by this study, which offers a new and practical paradigm.

Keywords: face anti-spoofing, DenseNet 161, CDCN++, pixel-wise supervision, deep learning, convolutional neural network, spoof detection, benchmark datasets

I. INTRODUCTION

Face anti-spoofing, a crucial technology in biometrics and security, aims to fortify face recognition systems by differentiating between authentic face inputs and spoof attacks. A novel approach in this field that effectively counters spoofing attempts by using deep learning and convolutional neural networks is pixel-wise supervision with DenseNet 161 and CDCN++. DenseNet-161 is often at the top of its game when it comes to computer vision tasks like segmentation, object identification, and picture classification. The three primary factors contributing to its exceptional accuracy and generalization capacity are its ability to capture rich and varied input, promote optimum parameter use, and facilitate consistent training. It is readily adjustable to various activities and datasets. Its numerous connections and flexible architecture allow it to be easily modified to match the necessary network depth, breadth, and complexity. This means that it may be used in a variety of contexts and applications. The pixel-wise supervision of DenseNet-161 is contrasted with that of the CDCN++ module. Comparing projected depth co-occurrence features with ground truth depth information at the pixel level is known as "pixel-by-pixel monitoring." The model can therefore identify minute clues and characteristics that may point to a spoofing attempt. A loss function is used to punish differences between predicted depth co-occurrence features and ground truth depth data in order to improve the CDCN++ module during training. The loss function helps the model understand spatial dependence and contextual information, enabling accurate prediction of depth features. DenseNet-161 is very good at extracting appearance-based features from photos. In order to achieve good feature extraction, it uses a lot of connections, which also helps to promote feature reuse and information flow across layers. It is trained using standard supervised learning techniques on massive picture categorization datasets like ImageNet.

II. RELATED WORKS

[1] There is a spike in increasing reliance on facial recognition technologies for security applications. Solomon and Cios (2023) proposed the FASS system, which uses deep learning and image quality criteria, as a defense against spoof attacks.

[2] Wang et al.'s use of consistency regularization allowed deep face anti-spoofing techniques to be improved (2023). By ensuring that feature representations are consistent, this approach seeks to increase the model's resilience.

[3] Liu and colleagues (2023) presented a deep face anti-spoofing solution that is capable of handling several types of attacks without the need to identify which ones they are. This method improves the adaptability of the system in a variety of circumstances.

[4] Fang et al.'s research from 2023 on anti-spoofing techniques for surveillance faces highlights the need of having systems that can accurately detect spoof attempts in surveillance footage that has been captured.

[5] Verissimo et al. (2023) looked at transfer learning techniques for face anti-spoofing detection and highlighted the advantages of using pre-trained models and domain knowledge to improve system performance.

[6] Yu et al. (2023) created a flexible-modal face anti-spoofing benchmark and emphasized the importance of assessing the system's functionality over a range of modalities and scenarios.

[7] Lin et al. (2023) developed DEFAEK, a fast adaptive network for face anti-spoofing that prioritizes network flexibility and is successful and domain-effective.

[8] In their study on anti-spoofing surveillance face technologies, Wang et al. (2023) presented the dynamic feature queue methodology, which showed how effective progressive training methods can be in improving system performance.

[9] In a recent work, Solomon (2023) investigated how face anti-spoofing and deep learning for unsupervised image recognition systems relate to one another, emphasizing the possibility of combining the two technologies to provide more robust security solutions.

[10] Muhammad and Oussalah (2023) emphasized the value of data quality and sampling techniques in enhancing system performance and resilience in their research on face anti-spoofing.

[11] In addition to addressing open issues and highlighting future research directions, Yu, Zitong, Yunxiao Qin, Xiaobai Li, Chenxu Zhao, Zhen Lei, and Guoying Zhao (2022) emphasize the importance of deep learning in face anti-spoofing, particularly in incorporating pixel-wise supervision, domain generalization, and multi-modal sensor applications.

[12] PatchNet, a face anti-spoofing (FAS) framework, is proposed by Wang, Chien-Yi, Yu-Ding Lu, Shang-Ta Yang, and Shang-Hong Lai (2022). It reformulates FAS as a fine-grained patch-type recognition issue, taking into account local factors as capturing devices and presenting materials. PatchNet outperforms current methods on intra-dataset, cross-dataset, and domain generalization benchmarks by using auxiliary pixel-level supervision and novel loss functions to improve data variation and generalization ability. This opens the door to real-world applications and further research into spoof-related cues.

[13] Fang et al. (2023) introduce the SuHiFiMask dataset for surveillance face anti-spoofing, addressing challenges in low-resolution and noisy environments. Their proposed Contrastive Quality-Invariance Learning (CQIL) network effectively handles image quality variations, demonstrating superior performance through extensive experiments.

[14] Echo-FAS, a unique acoustic-based face anti-spoofing solution for smartphones, is introduced by Kong, Chenqi, Kexin Zheng, Shiqi Wang, Anderson Rocha, and Haoliang Li in 2022. For face liveness recognition, Echo-FAS uses acoustic signals instead of costly sensors, providing robustness and cost-effectiveness. By merging local and global frequency hints, the suggested system achieves outstanding performance and offers fresh perspectives for FAS systems on mobile devices.

[15] In order to gather more liveness cues, Chen et al. (2022) suggest a two-stream network for face anti-spoofing that combines a convolutional network with a local difference network. Utilizing a straightforward binary face mask supervision scheme, their method outperforms others, exhibiting effectiveness with 2.79M parameters and real-time inference speeds of up to 118 frames per second.

[16] The novel acoustic-based face anti-spoofing system for smartphones, called Echo-FAS, is proposed by Kong et al. (2022). It uses a specially designed acoustic signal to identify the liveness of a face. By capturing 3D geometrical information and reducing overfitting in RGB-based FAS models, Echo-FAS provides robustness and cost-effectiveness while also developing FAS technology for mobile devices.

[17] In their 2021 study, Yu et al. reconsider pixel-wise supervision for face anti-spoofing (FAS) and suggest a novel pyramid supervision to improve the capacity of deep models to learn both global semantics and local features. By delivering fine-grained pixel/patch-level indications, their method enhances efficiency and interpretability and provides insights for future architecture and supervision design in FAS.

[18] A unified pixel- and patch-wise self-supervised learning system for domain adaptive semantic segmentation is presented by Chen et al. (2023) as PiPa. PiPa attains competitive accuracy on unsupervised domain adaptation benchmarks by explicitly promoting discriminative pixel-wise features and robust patch learning against various contexts, all the while maintaining compatibility with alternative techniques for additional performance enhancement.

[19] In order to lessen the cost of pixel-level annotations, Shen et al. (2023) explore label-efficient deep image segmentation algorithms, with a focus on weak supervision techniques. With implications for future research paths, their taxonomy classifies approaches based on segmentation difficulties and weak labels, providing insights into bridging the gap between dense prediction and weak supervision in picture segmentation.

[20] Fu et al. (2022) introduce Shape Alignment FacE (SAFE), a lightweight approach for cross-spectral face hallucination in Heterogeneous Face Recognition (HFR). By aligning image shapes using a 3D face model and incorporating probabilistic pixel-wise loss, SAFE achieves superior performance with a more lightweight generator and the ability to train on low-shot datasets, enhancing practical applications in heterogeneous face recognition.

III. EXISTING SYSTEM

Current face anti-spoofing technology is mostly based on early deep learning architectures and more conventional machine learning approaches. Typically, these systems use facial image processing software to extract features, and then use early Convolutional Neural Networks (CNNs) or Support Vector Machines (SVM) for classification. Hand-crafted traits are often used in the feature extraction process to distinguish between real and fake facial emotions. The next stage involves utilizing this data to train classifiers to differentiate between actual and phony faces.
Problem with Features Representation: Existing algorithms are based on handcrafted features, however they may not be able to distinguish between a real and fake face. This may result in poor detection performance, particularly when faced with complex spoofing strategies.

Because these systems depend on qualities that have been manually generated, they may find it difficult to adapt to various datasets and real-life settings. People may thus find it more difficult to deal with novel forms of spoofing.

Facial picture feature vectors are notoriously dimensional, which wastes resources and makes training and inference difficult. As a consequence, scaling becomes difficult and efficiency suffers.

In difficult cases when the differences are minimal, classifiers that are too simple and trained on features that were either hand-crafted or employed in early CNN architectures may find it difficult.

Adversarial attacks may exploit a weakness in early deep learning models and traditional machine learning: the potential for data misclassification brought on by minute changes in the input data. This flaw makes Face Anti-Spoofing systems less reliable and secure when used in production.

Adaptability issues: Because current systems lack mechanisms for ongoing learning and updating, they may find it difficult to handle new threats and spoofing tactics. These systems may eventually lose their effectiveness if they aren't updated often.

In conclusion, although some degree of spoofing detection is possible with current Face Anti-Spoofing systems, these systems suffer from a number of drawbacks due to their dependence on hand-crafted features and outdated deep learning architectures. The list includes problems with generalizability, computational inefficiencies, limited discriminative strength, vulnerability to attacks, and adaptability. To counteract spoofing efforts in face recognition applications, these flaws must be fixed and stronger and more dependable Face Anti-Spoofing solutions must be created.

IV. PROPOSED SYSTEM

The proposed work compares pixel-wise supervision utilizing DenseNet-161 and CDCN++ with the goal of strengthening face anti-spoofing. It is well recognized that small differences between actual and synthetic faces may be captured by Pixelwise Supervision with DenseNet-161, which takes use of the complex hierarchical features of deep networks. By concentrating on the finer elements of the picture, it can identify spoofing attempts with more accuracy. On the other hand, CDCN++ provides more indicators than only display-based capabilities. Through the analysis of depth data, CDCN++ may identify minute variations in the structure and form of the face, potentially enhancing its resistance against spoofing efforts that make use of 2D facial representations. Without these features, which are packed with background data, you can't tell a synthetic face from an actual one. A collection of dense blocks make up the Pixelwise Supervision utilizing DenseNet-161 architecture. Because every layer in a dense block has access to feature mappings from every layer below it, feature reuse is encouraged and information flow between levels is facilitated. Pixelwise Supervision with DenseNet-161 adds a set of learnable filters to the input feature maps via a convolutional layer in order to retrieve hierarchical features. Batch normalization layers correct the activations that come after convolutional layers in order to accelerate training and improve convergence. It does this by replacing DenseNet-161 instances with DenseNet-161-using Pixelwise Supervision. Using pixel-wise supervision and DenseNet-161's strong feature extraction capabilities, this method improves the model's capacity to recognize minuscule anomalies that may be signs of face spoofing.

V. BLOCK DIAGRAM

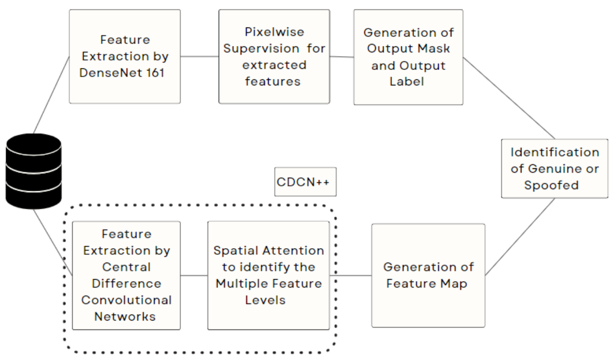

Fig 5.1: Block Diagram of Overall Architecture

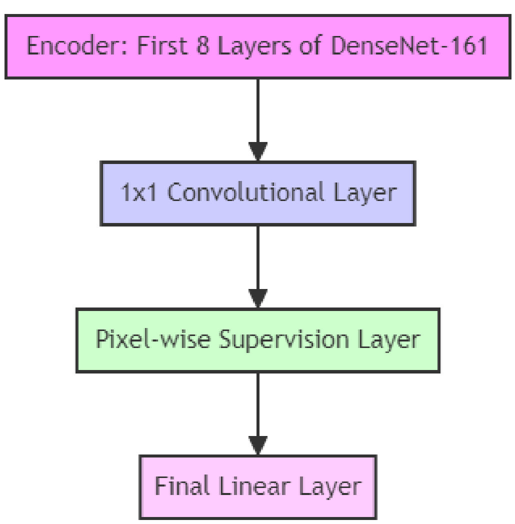

Fig 5.2: Pixel-wise Supervision Flow Diagram

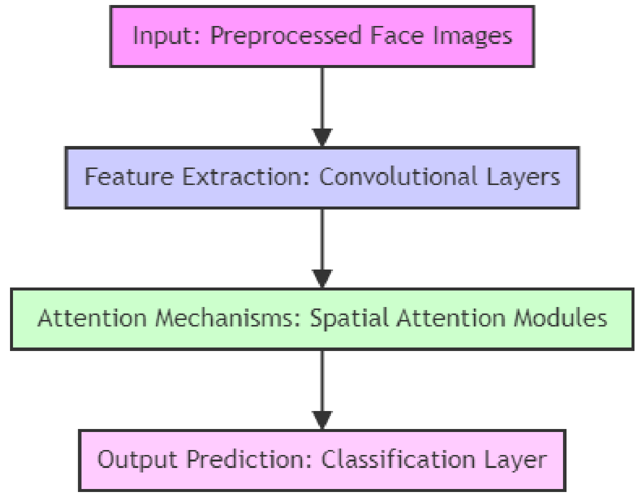

Fig 5.3: Central Difference Convolution Network++(CDCN++) Flow Diagram

VI. METHODOLOGY

Module 1: Preparing and Standardizing Data

Data preparation is a crucial initial step in any machine learning pipeline to guarantee successful model training. The input data is formatted and standardized by it. This module employs a number of preprocessing techniques to enhance the models' performance when it comes to face anti-spoofing. To guarantee uniformity, the input face images are first scaled and normalized. Resizing is used to ensure that every picture in the dataset has the same size in order to make processing the dataset consistent. In contrast, by scaling pixel values to a common range, often between 0 and 1, normalization techniques help to lessen the effect of illumination and contrast changes. To go one step further, by using further preprocessing techniques like alignment or cropping to the facial features, the accuracy of the model might be further enhanced. By standardizing the input data in this manner, the models are able to learn the discriminative features and patterns required to distinguish between genuine and fabricated face photographs.

Module 2: DenseNet-161 Feature Extraction

A key element of face anti-spoofing systems is feature extraction, which is in charge of extracting discriminative information from input facial photos. This module focuses on feature extraction using DenseNet-161 architecture, taking use of its capacity to extract complex hierarchical characteristics that are essential for differentiating real face features from possible spoofs. DenseNet-161 is a deep convolutional neural network (CNN) that excels in efficiently reusing features and propagating information between layers because to its densely connected blocks. DenseNet-161's initial eight layers function as the encoder, specifically removing hierarchical information from the input face pictures. After then, these characteristics are processed via other layers, which results in an all-encompassing representation of the input picture. In addition, a 1x1 convolutional layer is used to lower dimensionality and provide predictions at the pixel level, improving the model's sensitivity to minute irregularities suggestive of face spoofing. The model is able to learn and distinguish between real and fake face characteristics by using DenseNet-161 for feature extraction. This provides a strong basis for further research.

Module 3: Supervision and Analysis Pixel-by-Pixel

The Pixelwise Supervision Model, which is based on DenseNet-161, examines every pixel in the images to determine which faces are real and which are not. This section covers pixelwise supervision theory and analytical methods for spoofing detection. The model uses the hierarchical properties that DenseNet-161 acquired to find tiny abnormalities that might be caused by face spoofing. It explores data more thoroughly. Through breaking down the image into its pixel-level constituents, the model has enhanced capability to detect intricate efforts at spoofing that could otherwise remain undetected at higher abstraction levels. Once the final linear layer has processed all the input, a thorough picture prediction is generated. Through pixel-by-pixel analysis and monitoring, this technique improves the safety and dependability of face recognition systems by increasing their capacity to identify spoofing efforts.

Module 4: Comparing Pixelwise Supervision using DenseNet-161 with CDCN++

Pixelwise supervision with DenseNet-161 performs better in fine-grained analysis of individual pixels, enabling very accurate identification of small abnormalities. By using pixelwise supervision, our model carefully analyzes every pixel in face photos, improving its capability to distinguish real faces from fakes. Robustness in face anti-spoofing tasks is ensured by DenseNet-161, which captures complex hierarchical characteristics essential for differentiating real face features from possible spoofs. Conversely, CDCN++ improves the ability to distinguish between real and fake face pictures by incorporating Central Difference Convolution (CDC) methods into a hierarchical CNN design. Even though CDCN++ collects both high-level and low-level semantic information, it may not be as detailed in its analysis of individual pixels as DenseNet-161. This restriction could affect the model's capacity to identify minute differences that point to face spoofing, especially in situations when accurate detection is essential. Furthermore, the Pixelwise Supervision Model with DenseNet-161 is more suitable for time-sensitive applications and rapid model deployment because to its efficient convergence, which achieves high accuracy in fewer epochs than CDCN++. In summary, pixelwise supervision with DenseNet-161 provides better security and dependability than CDCN++ when dealing with anti-spoofing duties. Because of its capacity for conducting in-depth analysis of individual pixels as well as its effective convergence and deployment, it is the recommended option for applications where accurate spoofing attempt detection is crucial.

VII. RESULT AND DISCUSSION

For pixel-wise supervision, the face anti-spoofing system compares DenseNet 161 and CDCN++ models to determine which one is better. It was shown that DenseNet 161 outperformed CDCN++ in identifying and thwarting efforts at face spoofing, with an astounding accuracy rate of 0.904. On the other hand, CDCN++ achieved an accuracy of 0.70. It is clear from the results that DenseNet-161 performed better than CDCN++. The combination of various state-of-the-art technologies allows the system to achieve the highest standards in face anti-spoofing, which is responsible for its excellent reliability and durability in real-world scenarios.

7.1.Training Convergence

The DenseNet-161 model's training behavior under Pixelwise supervision is shown in Figures 7.1 and 7.2. Figure 7.1 shows a consistent rise in accuracy across each epoch, indicating effective learning and convergence to a stable state. On the other hand, Figure 7.2 shows that training loss decreases continuously throughout epochs, suggesting that the model is capable of efficiently optimizing its parameters. Comparable convergence graphs for the CDCN++ model are shown in Figures 7.3 and 7.4.

7.2.Model Comparison

Figure 7.5 presents a direct comparison of the accuracy obtained by the DenseNet-161 and CDCN++ models. The DenseNet-161 model clearly performs better than CDCN++ in terms of accuracy; it achieves a much higher value of 0.9 compared to CDCN++'s 0.77.This illustrates the degree to which DenseNet-161 can discriminate between authentic and fraudulent faces.

7.3. Performance Evaluation

We examined a number of indicators in order to comprehend the models' performance on a deeper level than merely accuracy. For both models on the test dataset, confusion matrices and classification reports are shown in Figures 7.6 and 7.7. For each model, the true positives, true negatives, false positives, and false negatives are broken down in depth in these figures. A high percentage of accurately categorized samples (both real and faked) and low misclassification rates are desirable characteristics of a well-performing model. DenseNet-161's supremacy is further supported by the categorization reports in Figures 7.6 and 7.7. DenseNet-161

shows a distinct advantage in detecting spoof faces, even though both models show comparable precision for the "genuine" class (around 0.85-0.92). For the "spoofed" class, it performs better than CDCN++ (0.67), with a precision of 0.88, suggesting a decreased chance of misclassifying a spoofed face as genuine.
Furthermore, when compared to CDCN++, DenseNet-161 exhibits a marginally higher recall for both real (0.88) and fake (0.92) faces. This demonstrates its ability to identify every actual and false face in the sample. The reports provide F1-scores that account for both recall and precision. The two models' F1-scores are similar, ranging from 0.76 to 0.90. However, DenseNet-161's consistently better F1-scores suggest a more equal performance across classes.

7.4. Output Analysis

As examples, Figures 7.11 through 7.14 present outputs produced by both models for real and fake photos. These figures show that the models may produce a classification result for a given facial image, even when the precise content of these outputs (classification label, confidence score, etc.) may not be stated directly. The entire system output screenshot for a real and a faked image is shown in Figures 7.15 and 7.16, respectively. These illustrations most likely show the user's final interface, which presents the genuine/spoof categorization result.

The results unequivocally demonstrate the superiority of the proposed face anti-spoofing system utilizing DenseNet-161 with Pixelwise supervision over the CDCN++ model. DenseNet-161's much improved accuracy, improved ability to identify spoof faces, and balanced performance across classes show how valuable it is in real-world anti-spoofing cases.

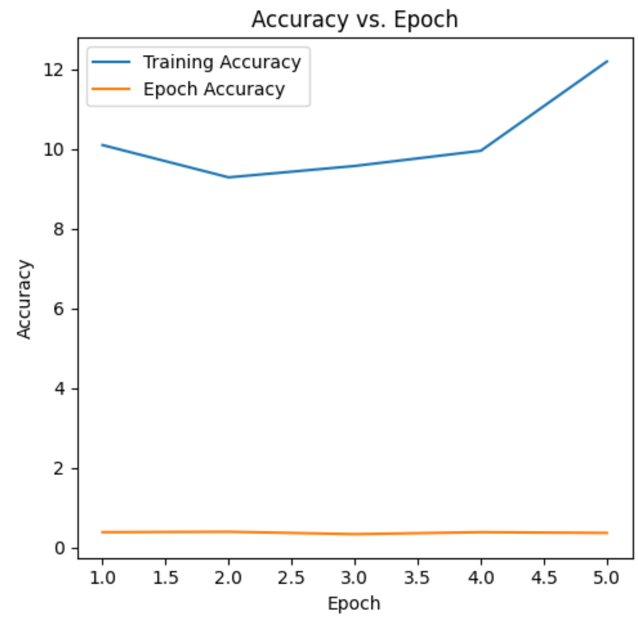

Fig 7.1: Graph Representing Epoch vs Accuracy of Pixelwise Supervision using DenseNet 161

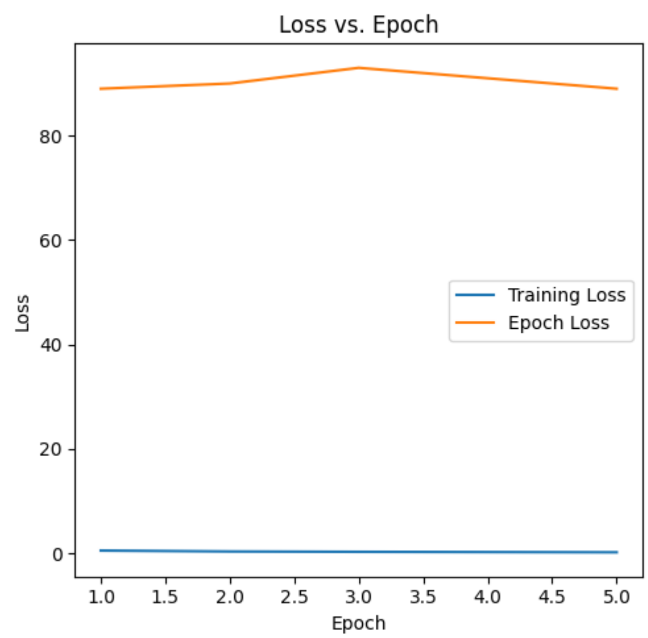

Fig 7.2: Graph Representing Epoch vs Loss of Pixelwise Supervision using DenseNet 161

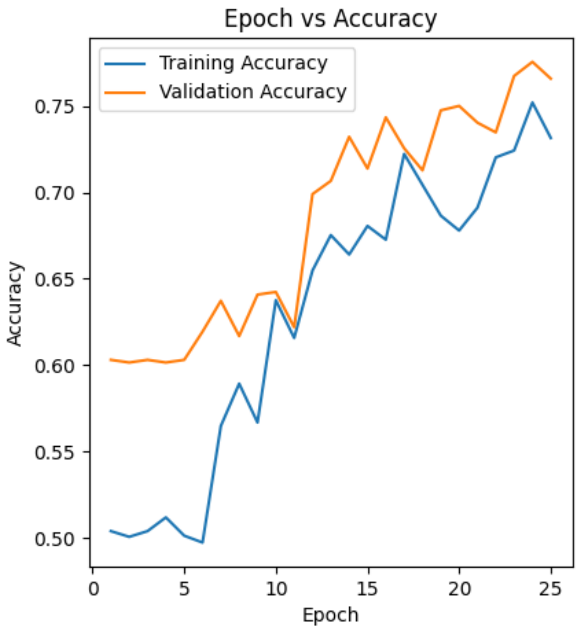

Fig 7.3: Graph Representing Epoch vs Accuracy of CDCN++

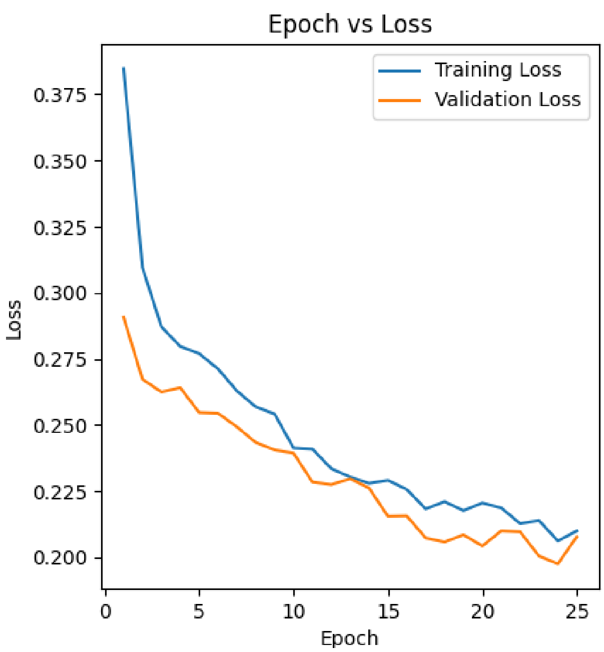

Fig 7.4: Graph Representing Epoch vs Loss of CDCN++

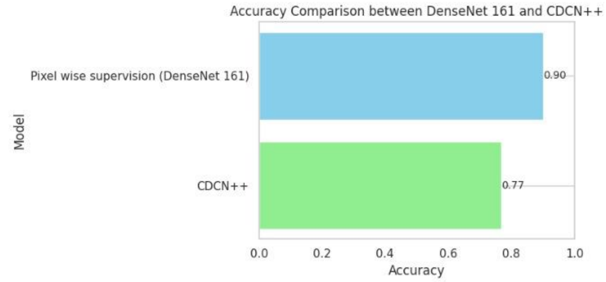

Fig 7.5: Graph Comparing Accuracy of Pixelwise Supervision using Densenet 161 and CDCN++

Accuracy: 0.9
Classification Report:

	precision	recall	f1-score	support
0.0	0.92	0.88	0.90	50
1.0	0.88	0.92	0.90	50
accuracy			0.90	100
macro avg	0.90	0.90	0.90	100
weighted avg	0.90	0.90	0.90	100

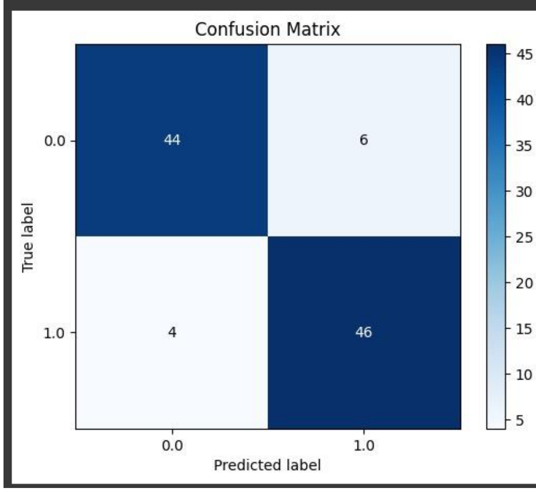

Fig 7.6: Confusion Matrix and Classification Report of Pixelwise Supervision using Densenet 161

```
Accuracy: 0.7679012345679013
Classification Report:
              precision    recall  f1-score   support

           0       0.85      0.76      0.80       250
           1       0.67      0.78      0.72       155

    accuracy                           0.77       405
   macro avg       0.76      0.77      0.76       405
weighted avg       0.78      0.77      0.77       405
```

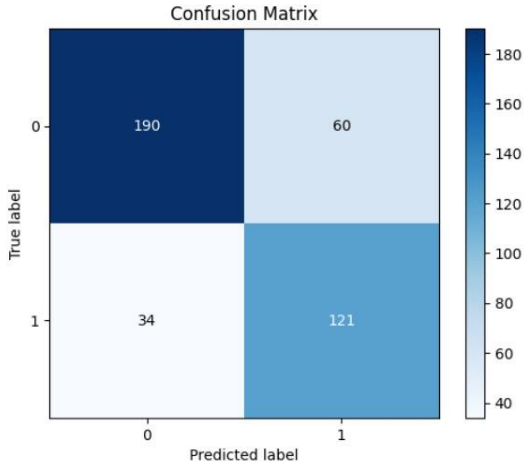

Fig 7.7: Confusion Matrix and Classification Report of CDCN++

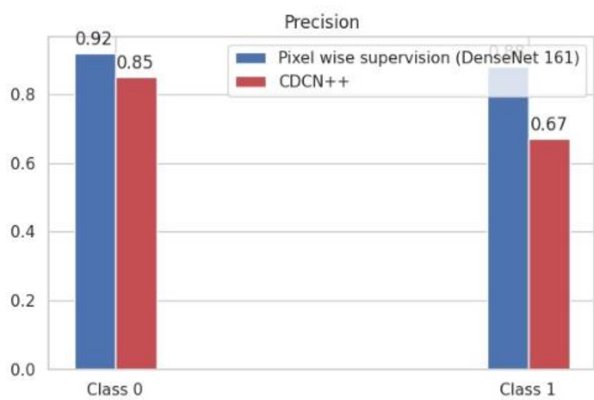

Fig 7.8: Graph Comparing Precision of Pixelwise Supervision using Densenet 161 and CDCN++

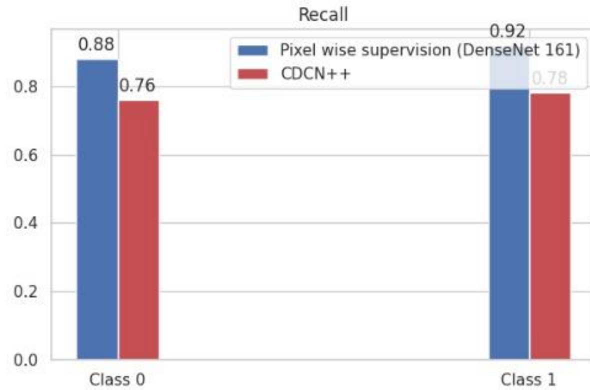

Fig 7.9: Graph Comparing Recall of Pixelwise Supervision using Densenet 161 and CDCN++

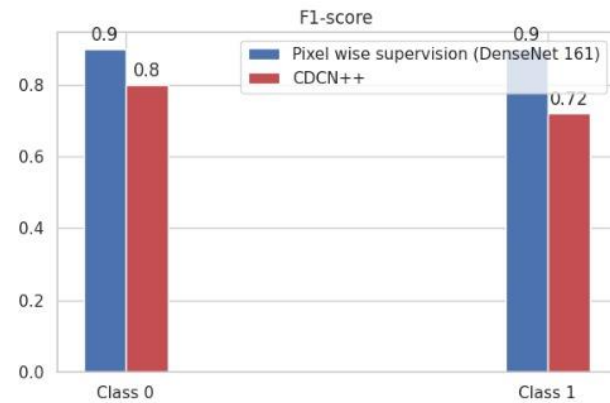

Fig 7.10: Graph Comparing F1-Score of Pixelwise Supervision using Densenet 161 and CDCN++

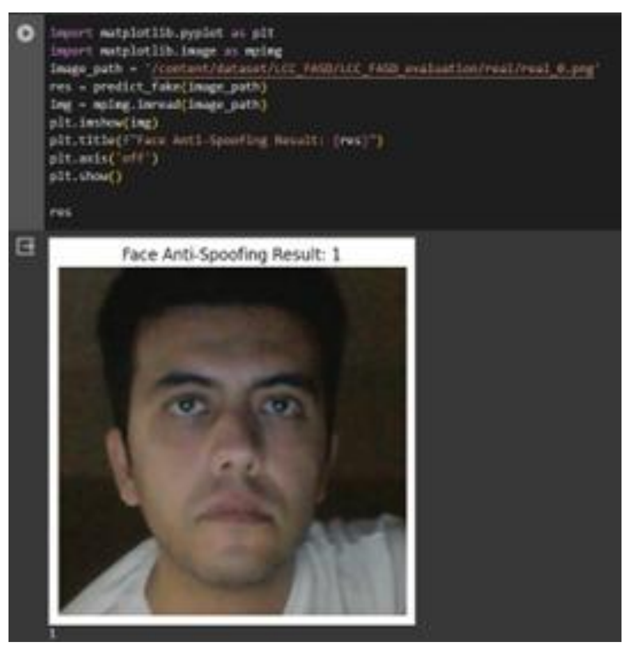

Fig 7.11: Output for Genuine Image of Pixelwise Supervision using DenseNet 161

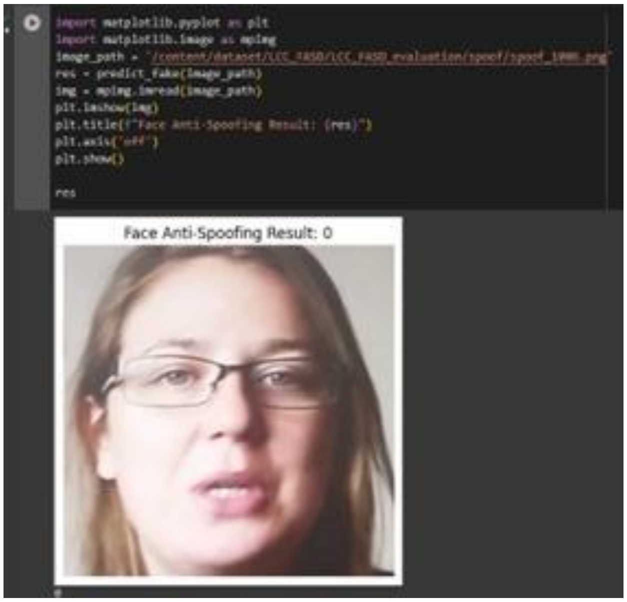

Fig 7.12: Output for Spoofed Image of Pixelwise Supervision using DenseNet 161

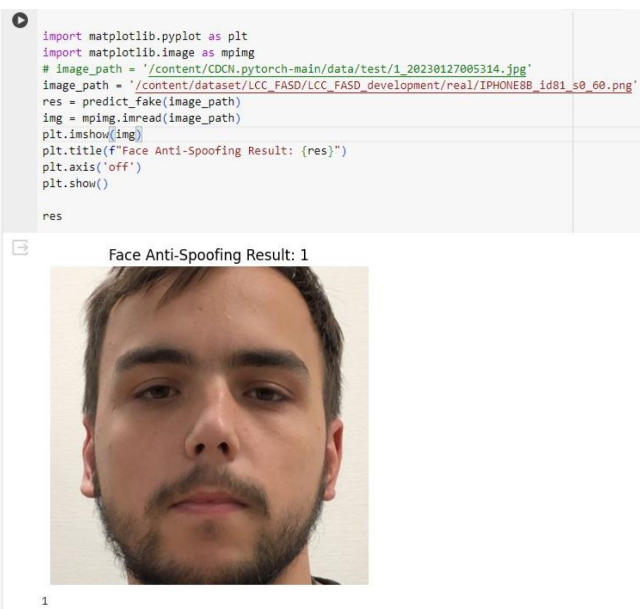

Fig 7.13: Output for Genuine Image of CDCN++

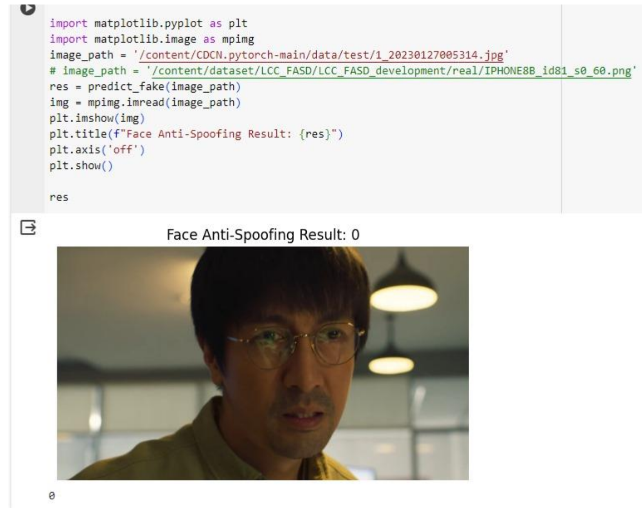

Fig 7.14: Output for Spoofed Image of CDCN++

Fig 7.15: Overall Output Screenshot of Genuine Image

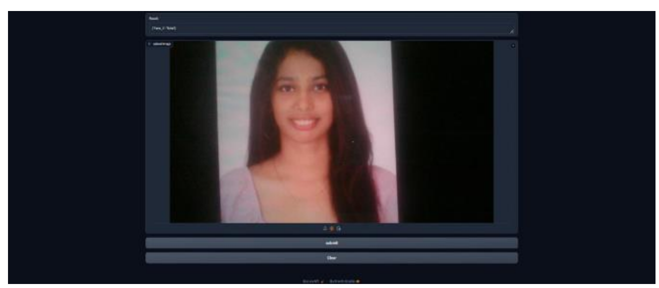

Fig 7.16: Overall Output Screenshot of Spoofed Image

VIII. CONCLUSION

All things considered, this is a significant advancement in face anti-spoofing technology. We compared Pixelwise Supervision with DenseNet-161 and CDCN++ in an effort to improve face recognition systems' accuracy and dependability, especially when it comes to spotting and thwarting spoofing assaults.DenseNet-161 is the best option for Pixelwise Supervision, which reliably and accurately detects minor spoofing attempts due to its extensive feature extraction capabilities and pixel-level inspection. Due to its efficiency and security, this approach is the greatest option for enhancing the dependability of face recognition systems. Simultaneously, CDCN++ demonstrated its ability to interpret depth data to detect minute changes in facial form, so strengthening its defenses against spoofing attempts. After thorough analysis and examination, it was determined that Pixelwise Supervision utilizing DenseNet-161 was the optimum model for face anti-spoofing tasks. Because of its fine-grained analysis and DenseNet-161's superior feature extraction capabilities, which enabled it to painstakingly study each pixel in face pictures, it was able to spot minute anomalies with astonishing precision. Pixelwise Supervision using DenseNet-161 showed effective convergence compared to CDCN++, achieving high accuracy in less training epochs. This efficacy not only streamlines training but also improves the model's suitability for real-world scenarios where deployment is critical for applications that need rapid responses.

IX. FUTURE WORK

Subsequent development of the Face anti-spoofing system combining Pixel-wise supervision with DenseNet 161 and CDCN++ may concentrate on a number of important areas. First, beyond DenseNet 161, investigating other deep learning architectures or combinations of them to improve the model's detection and prevention of spoofing assaults. Furthermore, adding other modalities like depth sensors or infrared imaging can strengthen the system's resistance to different kinds of spoofing efforts. In order to improve accuracy and speed, more study may be done on how to best optimize the model's hyperparameters. Furthermore, in order to guarantee the system's practical usability, it would be essential to examine its generalization capabilities across various datasets and real-world settings. Finally, investigating the system's implementation on edge devices or in real-time applications may be a fruitful avenue for further study.

REFERENCES

[1] Solomon, E., & Cios, K. J. (2023). FASS: Face anti-spoofing system using image quality features and deep learning. Electronics, 12(10), 2199.

[2] Wang, Z., Yu, Z., Wang, X., Qin, Y., Li, J., Zhao, C., ... & Lei, Z. (2023). Consistency regularization for deep face anti-spoofing. IEEE Transactions on Information Forensics and Security, 18, 1127-1140.

[3] Liu, A., Tan, Z., Liang, Y., & Wan, J. (2023). Attack-Agnostic Deep Face Anti-Spoofing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6335-6344).

[4] Fang, H., Liu, A., Wan, J., Escalera, S., Zhao, C., Zhang, X., ... & Lei, Z. (2023). Surveillance face anti-spoofing. IEEE Transactions on Information Forensics and Security

[5] Verissimo, S., Gadelha, G., Batista, L., Janduy, J., & Falcão, F. (2023). Transfer learning for face anti-spoofing detection. IEEE Latin America Transactions, 21(4), 530-536.

[6] Yu, Z., Liu, A., Zhao, C., Cheng, K. H., Cheng, X., & Zhao, G. (2023). Flexible-modal face anti-spoofing: A benchmark. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6345-6350).

[7] Lin, J. D., Han, Y. H., Huang, P. H., Tan, J., Chen, J. C., Tanveer, M., & Hua, K. L. (2023). DEFAEK: Domain Effective Fast Adaptive Network for Face Anti-Spoofing. Neural Networks, 161, 83-91.

[8] Wang, K., Huang, M., Zhang, G., Yue, H., Zhang, G., & Qiao, Y. (2023). Dynamic Feature Queue for Surveillance Face Anti-spoofing via Progressive Training. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6371-6378).

[9] Solomon, E. (2023). Face Anti-Spoofing and Deep Learning Based Unsupervised Image Recognition Systems.

[10] Muhammad, U., & Oussalah, M. (2023). Face anti-spoofing from the perspective of data sampling. Electronics Letters, 59(1), e12692.

[11Yu, Zitong, Yunxiao Qin, Xiaobai Li, Chenxu Zhao, Zhen Lei, and Guoying Zhao. "Deep learning for face anti-spoofing: A survey." *IEEE transactions on pattern analysis and machine intelligence* 45, no. 5 (2022): 5609-5631.

[12] Wang, Chien-Yi, Yu-Ding Lu, Shang-Ta Yang, and Shang-Hong Lai. "Patchnet: A simple face anti-spoofing framework via fine-grained patch recognition." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20281-20290. 2022.

[13] Fang, Hao, Ajian Liu, Jun Wan, Sergio Escalera, Chenxu Zhao, Xu Zhang, Stan Z. Li, and Zhen Lei. "Surveillance face anti-spoofing." *IEEE Transactions on Information Forensics and Security* (2023).

[14] Kong, Chenqi, Kexin Zheng, Shiqi Wang, Anderson Rocha, and Haoliang Li. "Beyond the pixel world: A novel acoustic-based face anti-spoofing system for smartphones." *IEEE Transactions on Information Forensics and Security* 17 (2022): 3238-3253.

[15] Chen, Suyang, Xiaoning Song, Zhenhua Feng,

Tianyang Xu, Xiaojun Wu, and Josef Kittler. "Face anti-spoofing with local difference network and binary facial mask supervision." *Journal of electronic imaging* 31, no. 1 (2022): 013007-013007.

[16] Kong, Chenqi, Kexin Zheng, Shiqi Wang, Anderson Rocha, and Haoliang Li. "Beyond the pixel world: A novel acoustic-based face anti-spoofing system for smartphones." *IEEE Transactions on Information Forensics and Security* 17 (2022): 3238-3253.

[17] Yu, Zitong, Xiaobai Li, Jingang Shi, Zhaoqiang Xia, and Guoying Zhao. "Revisiting pixel-wise supervision for face anti-spoofing." *IEEE Transactions on Biometrics, Behavior, and Identity Science* 3, no. 3 (2021): 285-295.

[18] Chen, Mu, Zhedong Zheng, Yi Yang, and Tat-Seng Chua. "Pipa: Pixel-and patch-wise self-supervised learning for domain adaptative semantic segmentation." In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 1905-1914. 2023.

[19] Shen, Wei, Zelin Peng, Xuehui Wang, Huayu Wang, Jiazhong Cen, Dongsheng Jiang, Lingxi Xie, Xiaokang Yang, and Qi Tian. "A survey on label-efficient deep image segmentation: Bridging the gap between weak supervision and dense prediction." *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).

[20] Fu, Chaoyou, Xiaoqiang Zhou, Weizan He, and Ran He. "Towards lightweight pixel-wise hallucination for heterogeneous face recognition." *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022).