

**FACE ANTI SPOOFING- A COMPARATIVE
ANALYSIS BETWEEN PIXEL WISE SUPERVISION
USING DENSENET 161 AND CDCN ++**

A PROJECT REPORT

Submitted by

**Sneha R (205002093)
H Varsha (205002120)**

in partial fulfillment for the award of the degree of

**BACHELOR OF TECHNOLOGY IN
INFORMATION TECHNOLOGY**



**DEPARTMENT OF INFORMATION
TECHNOLOGY**

Sri Sivasubramaniya Nadar College of Engineering

(An Autonomous Institution, Affiliated to Anna University)

MAY 2024

Sri Sivasubramaniya Nadar College of Engineering

(An Autonomous Institution, Affiliated to Anna University)

BONAFIDE CERTIFICATE

Certified that this Report titled “**FACE ANTI SPOOFING- A COMPARATIVE ANALYSIS BETWEEN PIXEL WISE SUPERVISION USING DENSENET 161 AND CDCN ++**” is the bonafide work of **R Sneha (205002093)** and **H Varsha (205002120)** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

Dr. A. SHAHINA

Professor and Head of
Department

Department of Information
Technology

SSN College of Engineering

Kalavakkam – 603 110

Dr. G. SORNAVALLI

Assistant Professor

Department of Information
Technology

SSN College of Engineering

Kalavakkam – 603 110

Submitted for project viva-voce examination held on.....

EXTERNAL EXAMINER

INTERNAL EXAMINER

ACKNOWLEDGEMENT

I thank **ALMIGHTY GOD** who gave me the wisdom to complete this Project. My sincere thanks to our beloved founder **Dr. SHIV NADAR, Chairman, HCL Technologies**. I also express my sincere thanks to **Ms. KALA VIJAYAKUMAR**, President, SSN Institution and our Principal **Dr. V.E. ANNAMALAI**, for all the help he has rendered during this course of study.

We are highly indebted to **Dr. A. Shahina, Head of the Department** for providing us with the opportunity and facilities to take up this project.

I am deeply obliged and indebted to the timeless help and guidance provided by **Dr. G. Sornavalli, Assistant Professor**, Department of Information Technology and also express my heartfelt thanks for making this project a great success.

I also thank all the faculty of the Department of Information Technology for their kind advice, support and encouragement and last but not the least I thank my parents and my friend for their moral support and valuable help.

ABSTRACT

This study presents a detailed comparative analysis between two state-of-the-art methodologies in the field of face anti-spoofing: Pixel-wise supervision utilizing DenseNet-161 architecture and Central Difference Convolutional Network++ (CDCN++). DenseNet-161, renowned for its efficient feature reuse mechanisms, capitalizes on dense block structures to facilitate information propagation and gradient flow, thereby enhancing parameter efficiency. By establishing direct connections between layers within dense blocks, DenseNet-161 enables seamless reuse of feature maps, fostering a comprehensive understanding of facial characteristics crucial for spoof detection. In contrast, CDCN++ specializes in generating depth maps from facial images, leveraging depth co-occurrence features to discern nuanced differences between authentic and synthetic faces. The depth maps generated by CDCN++ capture spatial connections and patterns inherent in facial depth data, enriching the model's understanding of facial structure and aiding in spoof detection. Both approaches incorporate pixel-wise supervision, which further refines the models' ability to recognize fine-grained

facial patterns and address inherent challenges in existing anti-spoofing techniques.

To evaluate the efficacy of the proposed methodologies, extensive experiments are conducted on benchmark datasets, including diverse real and synthetic spoofing attempts. The results demonstrate the superior performance and robustness of the proposed techniques compared to state-of-the-art solutions, showcasing their ability to generalize across various spoofing scenarios. Furthermore, the comparative analysis reveals insights into the strengths and weaknesses of each approach, providing valuable guidance for future research directions in face anti-spoofing.

Overall, this research significantly advances the field of face anti-spoofing by presenting a comprehensive comparative analysis between DenseNet-161 and CDCN++, offering insights into their respective strengths and contributions to spoof detection. The proposed techniques not only outperform existing methods in preventing spoofing attempts but also provide a practical framework for enhancing facial recognition security in real-world applications.

TABLE OF CONTENTS

CHAPTER NO	TITLE	PAGE NO
	ABSTRACT	iv
	LIST OF TABLES	viii
	LIST OF FIGURES	ix
	LIST OF ABBREVIATION	xi
1	INTRODUCTION	1
	1.1 OVERVIEW	1
	1.2 DEEP LEARNING	3
	1.3 CONVOLUTIONAL NEURAL NETWORK	4
	1.4 OBJECTIVES	5
	1.5 MOTIVATION	6
	1.6 ORGANISATION OF THE REPORT	6
2	LITERATURE SURVEY	8
3	DATASET	13
	3.1 DATASET DESCRIPTION	13
4	SYSTEM DESIGN	16

	4.1 PROPOSED METHODOLOGY	16
	4.2 PIXEL WISE SUPERVISION	20
	4.3 DENSENET-161	21
	4.4 CDCN++	22
	4.5 ARCHITECTURE	24
	4.5.1 PIXEL WISE SUPERVISION USING DENSENET-161	24
	4.5.2 CDCN++	27
	4.6 MODULES	29
	4.6.1 DATA PREPROCESSING AND STANDARDIZATION	29
	4.6.2 FETAURE EXTRACTION WITH DENSENET-161	30
	4.6.3 PIXEL WISE SUPERVISION AND ANALYSIS	32
	4.6.4 CDCN++	33
	4.6.5 COMPARING DENSENT-161 WITH PIXEL WISE SUPERVISION AND	35

		CDCN++-27	
5	RESULTS AND DISCUSSION		38
	5.1 RESULT OBTAINED FROM PIXEL WISE SUPERVISION USING DENSENET 161		38
	5.2 RESULT OBTAINED FROM CDCN++		39
	5.2 CONFUSION MATRIX FOR DIFFERENT MODELS		41
	5.2.1	PIXEL WISE SUPERVISION USING DENSENET-161	41
	5.2.2	CDCN++	42
	5.3 CLASSIFICATION REPORT FOR DIFFERENT MODELS		43
	5.3.1	PIXEL WISE SUPERVISION USING DENSENET – 161	44
	5.3.2	CDCN++	45
	5.4 COMPREHENSIVE PERFORMANCE EVALUATION METRICS		46

	5.5 OVERALL PERFORMANCE STATISTICS ANALYSIS FOR FACE ANTI SPOOFING	52
	5.5.1 PERFORMANCE ANALYSIS USING F1 SCORE	52
	5.5.2 PERFORMANCE ANALYSIS USING PRECISION	53
	5.5.3 PERFORMANCE ANALYSIS USING RECALL	54
	5.5.4 PERFORMANCE ANALYSIS USING ACCURACY	55
	5.6 USER INTERFACE	56
6	LIMITATIONS	58
6	CONCLUSION AND FUTURE WORKS	59
7	REFERENCES	61

LIST OF TABLES

Table 5.1	Training and Test Loss Summary for Pixel Wise supervision using DenseNet -161
Table 5.2	Training and Test Accuracy Summary for Pixel Wise supervision using DenseNet -161
Table 5.3	Training and Test Loss Summary for CDCN++
Table 5.4	Overall performance statistics analysis for Face Anti Spoofing Models

LIST OF FIGURES

Fig 1.1	Various types of Presentation attacks
Fig 1.2	Deep Learning – Face Anti Spoofing
Fig 1.3	Convolutional Neural Network– Face Anti Spoofing
Fig 3.1	Fake and Genuine Image
Fig 3.2	Set of Genuine Images
Fig 3.3	Set of Spoofed Images
Fig 3.4	Label of Spoofed images in csv format
Fig 3.5	Label of Genuine images in csv format
Fig 4.1	Use Case Diagram
Fig 4.2	Block Diagram of the Proposed model
Fig 4.3	Pixel wise supervision
Fig 4.4	Architecture of DenseNet - 161
Fig 4.5	Architecture of CDCN++
Fig. 4.6	Overall System Architecture
Fig.4.7	DeePixBiS Flow Diagram
Fig.4.8	CDCN++ Flow Diagram
Fig 5.1	Result for genuine image
Fig 5.2	Result for spoofed Image
Fig 5.3	Result for genuine image
Fig 5.4	Result for spoofed Image
Fig 5.5	Confusion Matrix for Pixel wise supervision using DensNet -161
Fig 5.6	Confusion Matrix for CDCN++
Fig 5.7	Classification Report for Pixel wise supervision using DenseNet -161

Fig 5.8	Classification Report for Pixel wise supervision using DenseNet -161
Fig 5.9	Epoch vs Accuracy for Pixel wise Supervision using DeneseNet – 161
Fig 5.10	Epoch vs Loss for Pixel wise Supervision using DeneseNet – 161
Fig 5.11	Epoch vs Accuracy for CDCN++
Fig 5.12	Epoch vs Loss for CDCN++
Fig 5.13	Comparison of F1 – score value of two CNN models
Fig 5.14	Comparison of Overall Precision of different CNN models
Fig 5.15	Comparison of Recall of 2 CNN models
Fig 5.16	Comparison of Accuracy of 2 CNN models
Fig 5.17	Depiction of Real Image using gradio.io
Fig 5.18	Depiction of Fake Image using gradio.io

LIST OF ABBREVIATIONS

CDCN	Central Difference Convolutional Network
CDCN++	Central Difference Convolutional Network ++
DeePixBiS	Pixel wise supervision using DenseNet - 161

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Facial recognition systems have become increasingly prevalent in modern society, offering convenience and security in various applications such as authentication, access control, and surveillance. However, these systems are vulnerable to presentation attacks, also known as spoofing attacks, where adversaries attempt to deceive the system by presenting fake facial images or videos. Face anti-spoofing techniques aim to detect and mitigate such attacks, safeguarding the integrity and reliability of facial recognition systems. In this section, we delve into the concept of face anti-spoofing, explore different types of presentation attacks, and discuss countermeasures to address these security challenges.

Face anti-spoofing refers to the process of identifying and mitigating presentation attacks aimed at deceiving facial recognition systems. These attacks exploit vulnerabilities in the system's ability to differentiate between genuine and fake facial images or videos. Face anti-spoofing techniques employ various methodologies, including image analysis, machine learning, and deep learning, to detect the presence of presentation attacks and distinguish them from genuine facial features. Figure 1.1 Various types of Presentation attacks. Presentation attacks can take various forms, each posing unique challenges to face anti-spoofing systems. Some common types of presentation attacks include:

- **Print Attack:** Adversaries present printed photos of genuine users' faces to the facial recognition system, sourced from social media or identity documents. Detectable through texture analysis and

depth information to distinguish between real faces and printed images.

- **Replay Attack:** Adversaries replay pre-recorded videos or images of genuine users' faces, captured via surveillance cameras or smartphones. Challenging to detect due to closely mimicking genuine faces, but can be countered by analyzing temporal inconsistencies and employing liveness detection techniques.
- **3D Mask Attack:** Adversaries create physical masks or 3D replicas of genuine users' faces using materials like silicone or latex. Difficult to detect as they closely resemble genuine faces; however, depth information analysis and texture analysis techniques can differentiate between real faces and 3D masks.
- **Makeup Attack:** Adversaries alter the appearance of genuine users' faces using cosmetics or facial prosthetics to deceive the facial recognition system. Hard to detect due to exploiting visual cues; countermeasures involve analyzing facial texture and employing color consistency checks to identify anomalies in makeup patterns.



Figure 1.1 Various types of Presentation attacks

1.2 Deep Learning

Deep learning represents a paradigm shift in artificial intelligence, empowering machines to learn complex patterns and representations directly from data. This transformative approach has revolutionized numerous fields, from computer vision to natural language processing, enabling machines to perform tasks with human-like accuracy and efficiency.

At the core of deep learning are artificial neural networks (ANNs), inspired by the structure and function of the human brain. These networks consist of layers of interconnected nodes, each performing simple computations on incoming data and passing the results to subsequent layers. The depth of these networks allows them to learn hierarchical representations of data, capturing increasingly abstract features as information propagates through the layers.

Facial recognition systems have become integral to various sectors, from security to consumer electronics. However, these systems are susceptible to spoofing attacks, where malicious actors attempt to deceive the system by presenting fake facial images or videos. Deep learning has emerged as a powerful tool for addressing this challenge, offering sophisticated techniques for detecting and mitigating presentation attacks in facial recognition systems. Figure 1.2 illustrates the application of deep learning in face anti-spoofing.

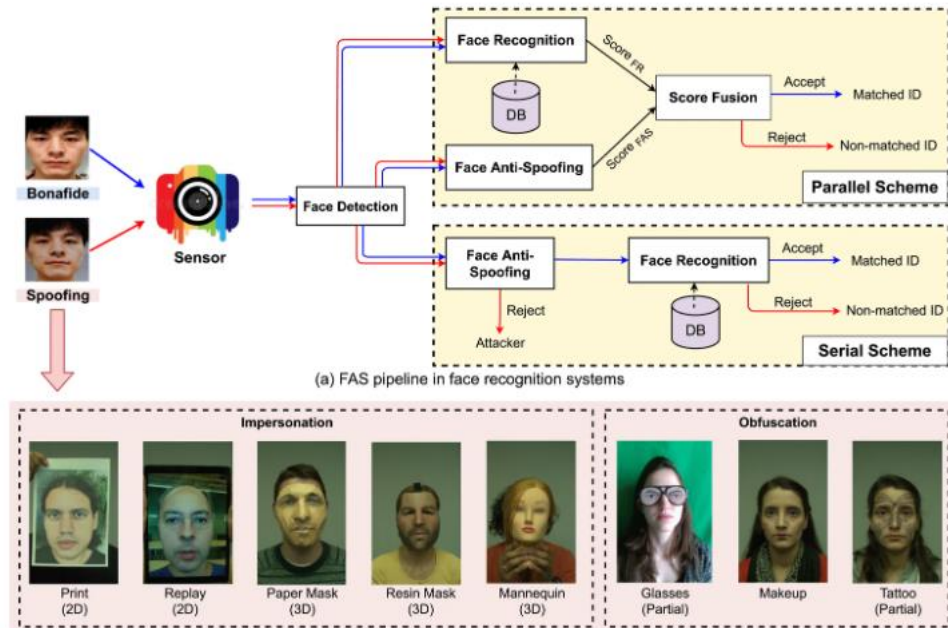


Fig 1.2 Deep Learning – Face Anti Spoofing

1.3 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are a class of deep neural networks designed for analysing visual data, such as images and videos. They have become the backbone of many computer vision tasks due to their ability to automatically learn hierarchical representations of features directly from raw data. Figure 1.3 demonstrates the use of Convolutional Neural Networks (CNNs) in face anti-spoofing.

- Convolutional layers are the building blocks of CNNs. They consist of learnable filters (also called kernels) that slide over the input data, performing element-wise multiplication and summation operations to extract local features.
- Pooling layers down sample the feature maps generated by convolutional layers, reducing their spatial dimensions while preserving important features. Common pooling operations include max pooling and average pooling.

- Activation functions introduce non-linearities into the network, allowing CNNs to learn complex relationships between inputs and outputs. Common activation functions include ReLU (Rectified Linear Unit), sigmoid, and tanh.
- Fully connected layers, also known as dense layers, connect every neuron in one layer to every neuron in the next layer. They perform high-level feature extraction and mapping, enabling CNNs to make predictions based on the learned features.
- CNNs are trained using the backpropagation algorithm, which calculates the gradients of the loss function with respect to the network's parameters. These gradients are used to update the network's weights and biases iteratively, optimizing its performance on the training data.

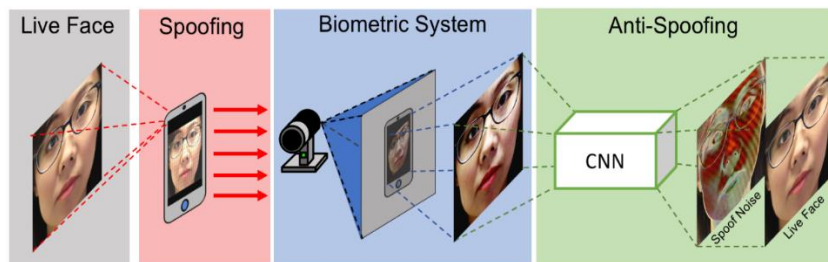


Fig 1.3 Convolutional Neural Network– Face Anti Spoofing

1.4 OBJECTIVE

- To conduct a comparative analysis between pixel-wise supervision using DenseNet-161 and CDCN++ for face anti-spoofing.
- To assess the performance of both approaches in distinguishing between genuine facial images and spoofed ones generated by impostors using various spoofing techniques.

- To investigate the capability of each method to prevent unauthorized access through facial recognition systems by accurately identifying and rejecting spoofed faces.
- To provide insights into the strengths and limitations of pixel-wise supervision and CDCN++ for face anti-spoofing applications.

1.5 MOTIVATION

- To address the increasing adoption of facial recognition technology across various domains and the growing concern regarding its vulnerability to spoofing attacks.
- To fill the gap in comprehensive studies comparing the effectiveness of different anti-spoofing techniques, including pixel-wise supervision and CDCN++.
- To enhance the reliability and trustworthiness of facial recognition technology by identifying and evaluating state-of-the-art anti-spoofing methods.
- To leverage advancements in deep learning and computer vision to develop and assess robust anti-spoofing solutions for real-world applications.
- To contribute to the advancement of facial recognition technology and improve its security measures against spoofing attacks, thereby ensuring the integrity and trustworthiness of biometric authentication systems.

1.6 ORGANISATION OF THE REPORT

Chapter 1 deals with the Introduction of the project

Chapter 2 deals with the literature survey carried out and the results of the investigation.

Chapter 3 deals with the dataset distribution.

Chapter 4 deals with design of the system.

Chapter 5 deals with the results and discussion of the project

Chapter 6 deals with the limitations of our project

Chapter 6 deals with the Conclusion and future work.

CHAPTER 2

LITERATURE SURVEY

The literature survey provides an exhaustive exploration into the burgeoning field of facial recognition technologies, with a keen focus on anti-spoofing strategies and their integration into security systems. It navigates through recent research endeavors aimed at fortifying defenses against spoof attacks, enhancing adaptability in diverse scenarios, and addressing the intricate challenges posed in surveillance contexts.

[1] Solomon and Cios (2023) introduced the Facial Anti-Spoofing System (FASS), a groundbreaking defense mechanism engineered with cutting-edge deep learning algorithms and stringent image quality criteria. FASS stands as a bastion against the onslaught of spoof attacks, leveraging sophisticated neural network architectures to ensure robust security in face recognition systems.

[2] Wang et al. (2023) contributed significantly to the advancement of deep face anti-spoofing techniques through the introduction of consistency regularization. By enforcing consistency in feature representations, their approach fortifies the resilience of anti-spoofing models, thereby erecting formidable defenses against evolving spoofing tactics.

[3] Liu et al. (2023) presented a versatile deep face anti-spoofing solution capable of handling diverse attack modalities without the need for specific attack identification. This adaptive approach enhances the adaptability of anti-spoofing systems, ensuring effective defense in real-world scenarios with varying threat landscapes.

[4] Fang et al. (2023) underscored the critical importance of accurate spoof detection in surveillance footage through their research on anti-spoofing techniques for surveillance faces. Their work sheds light on the necessity of robust anti-spoofing systems in safeguarding sensitive environments against potential security breaches.

[5] Verissimo et al. (2023) delved into transfer learning techniques for face anti-spoofing detection, emphasizing the advantages of leveraging pre-trained models and domain knowledge to enhance system performance. Their findings offer valuable insights into the development of more efficient and effective anti-spoofing solutions.

[6] Yu et al. (2023) introduced a flexible-modal face anti-spoofing benchmark, stressing the importance of evaluating system functionality across different modalities and scenarios. Their benchmark provides invaluable resources for assessing the performance of anti-spoofing systems under various conditions, aiding in the development of robust defense mechanisms.

[7] Lin et al. (2023) spearheaded the development of DEFAEK, a fast adaptive network for face anti-spoofing that prioritizes network flexibility and domain effectiveness. Their innovative approach offers swift and adaptive defense mechanisms, ensuring effective protection against spoof attacks in dynamic environments.

[8] Wang et al. (2023) proposed a dynamic feature queue methodology for anti-spoofing surveillance face technologies, demonstrating the effectiveness of progressive training methods in improving system

performance. Their methodology showcases the potential of iterative learning techniques in enhancing the robustness of anti-spoofing systems.

[9] Solomon (2023) delved into the synergy between face anti-spoofing and deep learning for unsupervised image recognition systems, highlighting the possibility of combining the two technologies to provide more robust security solutions. His research lays the groundwork for future advancements in integrated security frameworks.

[10] Muhammad and Oussalah (2023) emphasized the importance of data quality and sampling techniques in enhancing system performance and resilience against spoof attacks. Their findings underscore the significance of robust data processing pipelines in developing effective anti-spoofing solutions.

[11] Yu et al. (2022) advocated for the integration of deep learning techniques in face anti-spoofing, emphasizing the importance of pixel-wise supervision, domain generalization, and multi-modal sensor applications. Their insights provide valuable guidance for the development of advanced anti-spoofing systems.

[12] Wang et al. (2022) proposed PatchNet, a fine-grained patch-type recognition framework for face anti-spoofing. By reformulating anti-spoofing as a localized recognition problem, PatchNet achieves superior performance and robustness, paving the way for practical applications in real-world scenarios.

[13] Fang et al. (2023) introduced the SuHiFiMask dataset for surveillance face anti-spoofing, addressing challenges in low-resolution and noisy

environments. Their dataset provides valuable resources for evaluating anti-spoofing systems under challenging conditions, aiding in the development of more robust defense mechanisms.

[14] Kong et al. (2022) presented Echo-FAS, a unique acoustic-based face anti-spoofing solution for smartphones. By leveraging acoustic signals for face liveness recognition, Echo-FAS offers robust and cost-effective defense mechanisms, opening new avenues for anti-spoofing technologies on mobile devices.

[15] Chen et al. (2022) proposed a two-stream network for face anti-spoofing, combining convolutional and local difference networks for improved performance and real-time inference speeds. Their approach demonstrates the effectiveness of multi-stream architectures in enhancing the robustness of anti-spoofing systems.

[16] Kong et al. (2022) introduced Echo-FAS, a novel acoustic-based approach for face anti-spoofing on smartphones. By capturing 3D geometrical cues and reducing overfitting in RGB-based models, Echo-FAS offers robust defense mechanisms, advancing the state-of-the-art in mobile anti-spoofing technologies.

[17] Yu et al. (2021) reevaluated pixel-wise supervision in face anti-spoofing, proposing a novel pyramid supervision approach for deep learning models. Their methodology enhances the capacity of anti-spoofing systems to learn global semantics and local features, improving overall performance and interpretability.

[18] Chen et al. (2023) presented PiPa, a unified self-supervised learning system for domain adaptive semantic segmentation. By promoting discriminative pixel-wise features and robust patch learning, PiPa achieves competitive accuracy on unsupervised domain adaptation benchmarks, advancing the field of semantic segmentation.

[19] Shen et al. (2023) explored label-efficient deep image segmentation algorithms, focusing on weak supervision techniques to bridge the gap between dense prediction and weak supervision in picture segmentation. Their taxonomy provides valuable insights into the challenges and opportunities in image segmentation research.

[20] Fu et al. (2022) introduced 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 enhanced practicality in heterogeneous face recognition scenarios.

CHAPTER 3

DATASET

3.1 DATASET DESCRIPTION

The "Large Crowd collected Face Anti-Spoofing Dataset" (LCC-FASD) is a comprehensive collection of facial images designed for anti-spoofing research and development. This dataset, available on Kaggle, offers a diverse range of images captured in real-world scenarios, encompassing various lighting conditions, facial expressions, and backgrounds. LCC-FASD includes both genuine and spoofed facial images, enabling researchers to train and evaluate anti-spoofing algorithms effectively. Figure 3.1 depicts both fake and genuine images., Figure 3.2 showcases a set of genuine images., Figure 3.3 displays a set of spoofed images, Figure 3.4 represents the labels of spoofed images in CSV format, Figure 3.5 illustrates the labels of genuine images in CSV format. Key features of the LCC-FASD dataset include:

1. **Large-scale Collection:** LCC-FASD comprises a substantial number of facial images, facilitating the development and evaluation of anti-spoofing algorithms on a large scale.
2. **Diverse Scenarios:** The dataset encompasses diverse scenarios encountered in real-world settings, ensuring robustness and generalizability of developed algorithms across different environments.
3. **Authentic and Spoofed Samples:** LCC-FASD contains both authentic facial images and spoofed samples, allowing researchers to train models to differentiate between genuine and fraudulent attempts.

4. **Annotation:** The dataset may include annotations indicating the authenticity of each image, providing ground truth labels for algorithm training and evaluation.
5. **Accessibility:** Hosted on Kaggle, the LCC-FASD dataset is easily accessible for researchers, facilitating collaboration and reproducibility in the field of face anti-spoofing.

Overall, the LCC-FASD dataset serves as a valuable resource for advancing research in face anti-spoofing, offering a rich collection of real-world facial images for algorithm development, training, and evaluation.

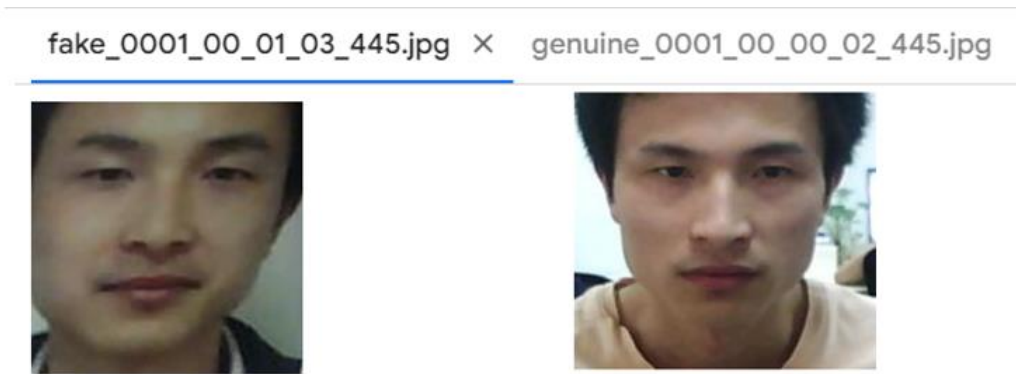


Fig 3.1 Fake and Genuine Image

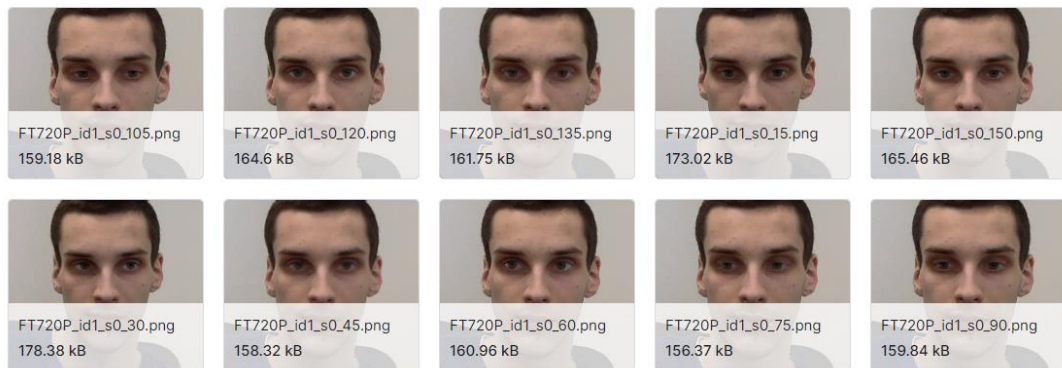


Fig 3.2 Set of Genuine Images

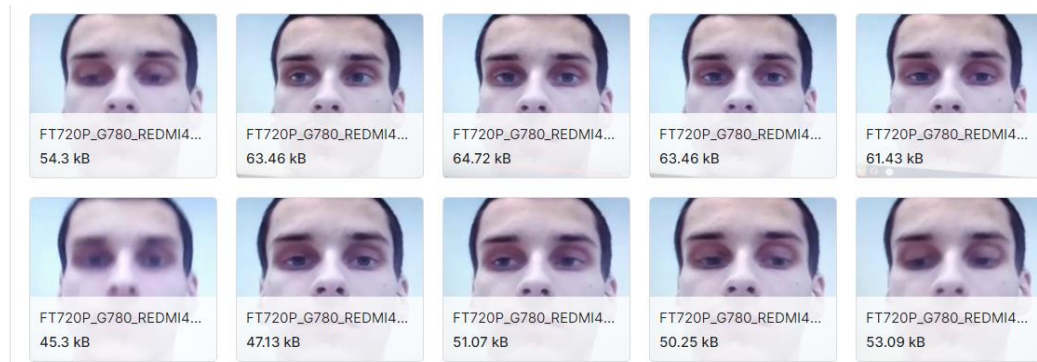


Fig 3.3 Set of Spoofed Images

train_data.csv X

1 to 10 of 2100 entries Filter

	name	label
0	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/YOUTUBE_AV203H_LEREELE3_id57_s0_164.png	0.0
1	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/YOUTUBE_MBP1314_IPHONE8B_id47_s2_105.png	0.0
2	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/YOUTUBE_MBP1314_IPHONE8B_id45_s0_45.png	0.0
3	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/YOUTUBE_VSVA2231WA_MEIZUM3S_id31_s0_170.png	0.0
4	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/YOUTUBE_AEUNKNOWN_BQS5025_id47_s3_30.png	0.0
5	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/YOUTUBE_MBP1314_IPHONE8B_id37_s0_135.png	0.0
6	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/YOUTUBE_HP620_IPHONE4_id47_s2_30.png	0.0
7	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/FT720P_IYAMAGB_IPHONE5SB_id7_s0_15.png	0.0
8	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/FT720P_MBP1314_IPHONE7B_id2_s0_135.png	0.0
9	/content/dataset/LCC_FASD/LCC_FASD_training/spoof/FT720P_IYAMAXB_SMG955U_id9_s0_60.png	0.0

Show 10 per page 1 2 10 100 200 210

Fig 3.4 Label of Spoofed images in csv format

train_data.csv X

2091 to 2100 of 2100 entries Filter

	name	label
8116	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id77_s0_45.png	1.0
8117	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id114_s0_30.png	1.0
8118	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id84_s0_120.png	1.0
8119	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id91_s0_154.png	1.0
8120	/content/dataset/LCC_FASD/LCC_FASD_training/real/SMG950U_id130_s0_92.png	1.0
8121	/content/dataset/LCC_FASD/LCC_FASD_training/real/SMG950U_id130_s0_131.png	1.0
8122	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id113_s0_90.png	1.0
8123	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id80_s0_105.png	1.0
8124	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id99_s0_30.png	1.0
8125	/content/dataset/LCC_FASD/LCC_FASD_training/real/YOUTUBE_id114_s0_15.png	1.0

Show 10 per page 1 10 100 200 209 210

Fig 3.5 Label of Genuine images in csv format

CHAPTER 4

SYSTEM DESIGN

4.1 PROPOSED METHODOLOGY

The project aims to tackle the vulnerability of facial recognition systems to spoofing attacks by proposing the implementation of two primary models: the Pixel-wise Supervision Model based on DenseNet-161 and the Central Difference Convolutional Network++ (CDCN++).

The Pixel-wise Supervision Model is designed to bolster accuracy and reliability in face anti-spoofing systems by meticulously scrutinizing each pixel in facial images to detect subtle discrepancies between genuine and spoofed faces. This model leverages DenseNet-161 for effective feature extraction, harnessing hierarchical features crucial for distinguishing genuine facial features from potential spoofs. Following feature extraction, a 1x1 convolutional layer is applied to enable granular scrutiny of details, thereby enhancing the model's sensitivity to subtle anomalies indicative of face spoofing. The subsequent linear layer refines the gathered information, ultimately culminating in a comprehensive prediction for the entire image.

In contrast, the CDCN++ model aims to push the boundaries of face anti-spoofing by integrating Central Difference Convolution (CDC) techniques into a hierarchical convolutional neural network architecture. This model seeks to improve discrimination capabilities between genuine and spoofed facial images by capturing both low-level details and high-level semantic

information. To achieve this, spatial attention modules are incorporated to focus on discriminative facial features while suppressing irrelevant regions, thereby enhancing the model's robustness against various spoofing attacks.

Comparing the two models, the Pixel-wise Supervision Model utilizing DenseNet-161 showcases superior performance in fine-grained analysis of individual pixels. This capability allows for the detection of subtle anomalies with high accuracy, making it particularly adept at discerning between genuine and spoofed faces. Conversely, while the CDCN++ model integrates Central Difference Convolution techniques to capture both low-level and high-level features, it may not provide the same level of granularity in analyzing individual pixels.

Therefore, in applications where precise detection of spoofing attempts is paramount, the Pixel-wise Supervision Model with DenseNet-161 is preferred. Additionally, the efficient convergence of the Pixel-wise Supervision Model, achieving high accuracy in fewer epochs compared to CDCN++, makes it preferable for time-sensitive applications and quick model deployment.

The Pixel-wise Supervision Model utilizing DenseNet-161 offers enhanced security and reliability in face anti-spoofing tasks, contributing significantly to the integrity of facial recognition systems.

The face anti-spoofing system acts as a security guard for facial recognition. Instead of simply accepting an image of a face, it

analyzes it thoroughly to weed out imposters. The process starts with the user presenting their face to a camera. This captured image is then fed into the system's processing unit. Here, sophisticated algorithms go to work. They might analyze blinking patterns, subtle facial muscle movements, or even skin texture to differentiate between a real person and a spoof. Some systems may employ depth sensors to create a 3D map of the face, exposing inconsistencies in masks or pre-recorded videos. If everything checks out and the system is convinced it's a real person, access is granted or the process continues. However, if the analysis reveals signs of a spoof, like a static image or an unnaturally smooth mask, the system triggers an alert and access is denied. This extra layer of security helps prevent unauthorized individuals from gaining access through stolen photos, videos, or even 3D printed replicas. Figure 4.1 depicts the use case diagram.

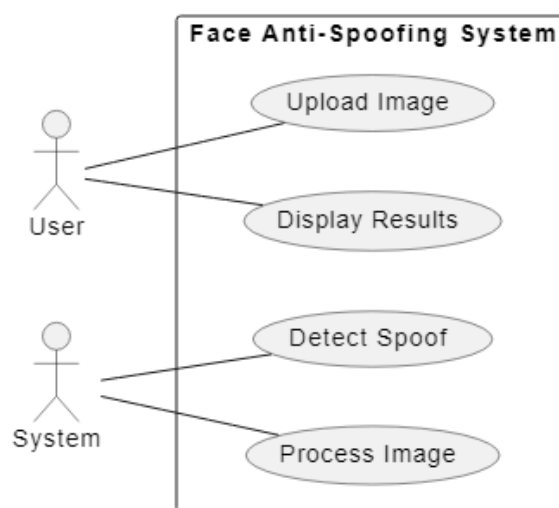


Fig 4.1 Use Case Diagram

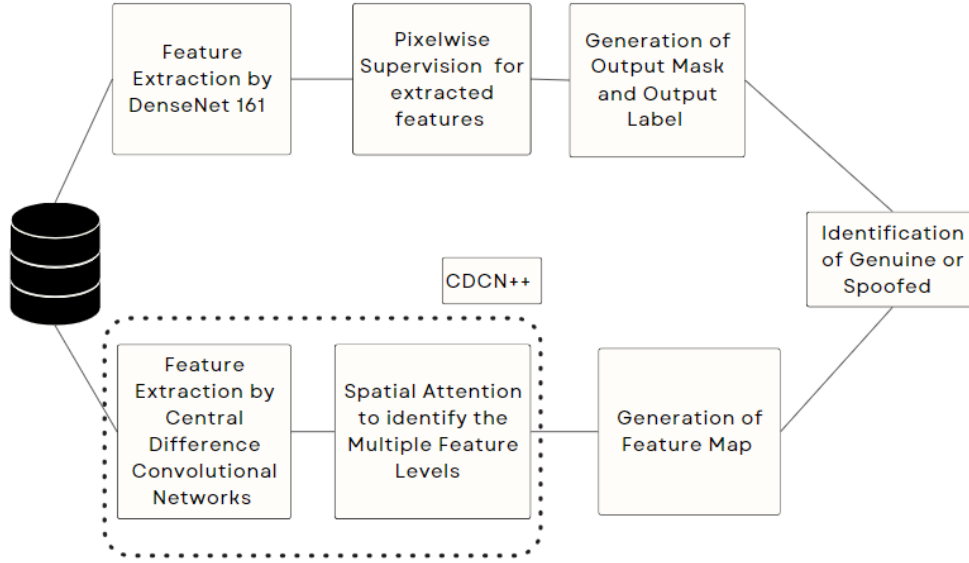


Fig 4.2 Block Diagram of the Proposed model

The illustrated block diagram in Fig 4.2 represents the proposed model for face anti-spoofing. It outlines a system designed to classify images as either genuine or spoofed. The system comprises two distinct feature extraction modules, each serving a unique purpose. The first module employs DenseNet 161, a deep convolutional neural network, to meticulously extract features from every pixel of the input image. This enables the system to capture intricate details crucial for discerning between genuine and manipulated facial images. The second module utilizes a Central Difference Convolutional Network Plus Plus(CDCN++) to analyze spatial attention across different levels of image features. By scrutinizing inconsistencies within the image, the CDCN++ enhances the system's ability to detect telltale signs of spoofing. Ultimately, both feature extraction modules produce an output, likely a feature map, which is then utilized for the classification of the input image as genuine or spoofed.

4.2 PIXEL WISE SUPERVISION

Pixel-wise supervision is a technique used in face recognition systems to improve their accuracy in distinguishing between real and fake faces. Instead of looking at the entire image as a whole, pixel-wise supervision examines each individual pixel meticulously. Imagine inspecting a photograph up close, focusing on every single dot of color. Pixel-wise supervision does just that—it scrutinizes each pixel's color, intensity, and position within a facial image. This meticulous analysis allows the system to detect even the smallest discrepancies that may indicate a fake face. By implementing pixel-wise supervision, facial recognition systems become more resilient against spoofing attacks where fake images or videos are used to deceive the system. This technique enhances the system's ability to differentiate between genuine facial features and counterfeit representations, thereby increasing its reliability and security. In summary, pixel-wise supervision is a powerful tool that enhances the accuracy and robustness of facial recognition systems by meticulously analyzing each pixel in a facial image. Fig 4.3 Pixel wise supervision

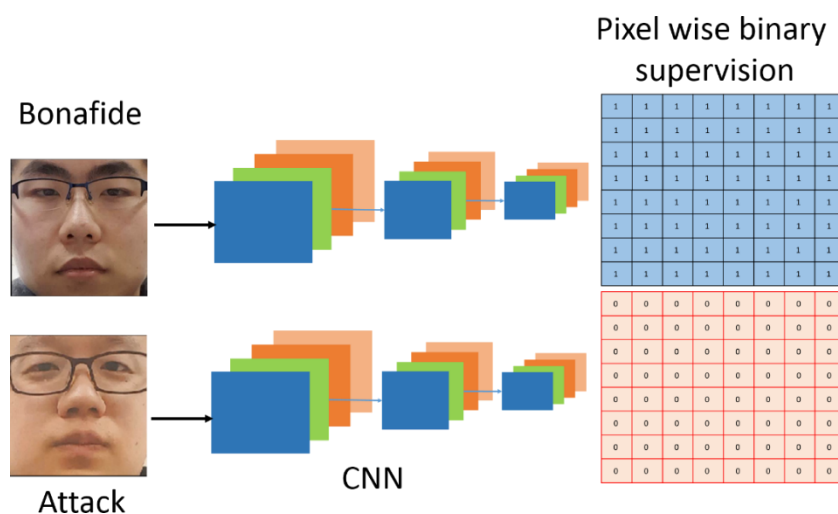


Fig 4.3 Pixel wise supervision

4.3 DENSENET - 161

DenseNet-161 stands out as a powerful convolutional neural network architecture specifically designed to address the challenges in facial recognition systems, particularly in combating spoofing attacks. Developed as an extension of the Dense Convolutional Network (DenseNet), DenseNet-161 offers a robust framework for feature extraction and classification tasks, making it well-suited for face anti-spoofing applications. At its essence, DenseNet-161 employs dense connectivity patterns between layers, facilitating feature reuse and propagation throughout the network. Unlike traditional convolutional neural networks where each layer receives inputs only from the preceding layer, DenseNet-161 fosters direct connections between all layers within a dense block. This dense connectivity promotes feature diversity and encourages the flow of information, resulting in more efficient feature learning and representation. One of the notable advantages of DenseNet-161 is its ability to capture intricate hierarchical features crucial for distinguishing genuine facial features from potential spoofs. By leveraging densely connected layers, the model can extract detailed information at different levels of abstraction, enabling it to discern subtle nuances in facial images. In the context of face anti-spoofing, DenseNet-161 serves as a powerful feature extractor, enabling the model to identify discriminative patterns indicative of spoofing attempts. Its depth and connectivity contribute to its effectiveness in detecting subtle differences between genuine and counterfeit facial features, thereby enhancing the security and reliability of facial recognition systems. Overall, DenseNet-161 represents a state-of-the-art solution for face anti-spoofing tasks, offering a combination of depth, connectivity, and feature extraction capabilities essential for robust and reliable facial recognition systems.

Through its dense connectivity patterns and hierarchical feature learning, DenseNet-161 provides a solid foundation for addressing the challenges posed by spoofing attacks and advancing the field of facial recognition security. Figure 4.4 illustrates the architecture of DenseNet-161.

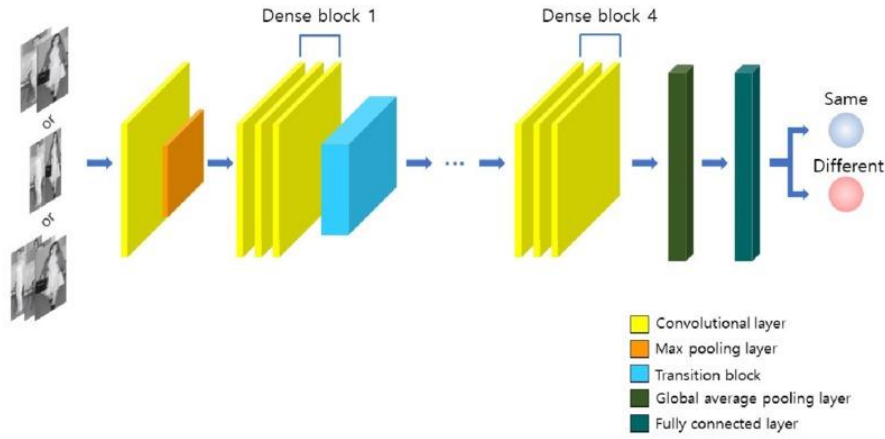


Fig 4.4 – Architecture of DenseNet - 161

4.4 CDCN++

The Central Difference Convolutional Network++ (CDCN++) is an innovative model designed to enhance the security of facial recognition systems by effectively distinguishing between genuine and spoofed facial images. In the ever-evolving landscape of face anti-spoofing, CDCN++ represents a cutting-edge approach that integrates Central Difference Convolution techniques into a hierarchical convolutional neural network architecture. At its core, CDCN++ aims to improve the discrimination capabilities between genuine and spoofed facial images by capturing both low-level details and high-level semantic information. This is achieved through the utilization of advanced convolutional operations, which enable the model to extract intricate features from facial images. By incorporating Central Difference Convolution techniques, CDCN++ enhances its ability to discern subtle differences between authentic and

counterfeit facial features. One of the key strengths of CDCN++ lies in its utilization of spatial attention modules, which play a crucial role in focusing the model's attention on discriminative facial features while suppressing irrelevant regions. This selective attention mechanism enhances the model's robustness against various spoofing attacks, ensuring that it can effectively identify genuine facial images amidst potential spoof attempts.

CDCN++ represents a significant advancement in the field of face anti-spoofing, offering a comprehensive solution that combines sophisticated feature extraction with attention-based mechanisms. By leveraging state-of-the-art techniques in convolutional neural networks, CDCN++ provides a powerful tool for enhancing the security and reliability of facial recognition systems in the face of evolving spoofing threats. Figure 4.5 displays the architecture of CDCN++.

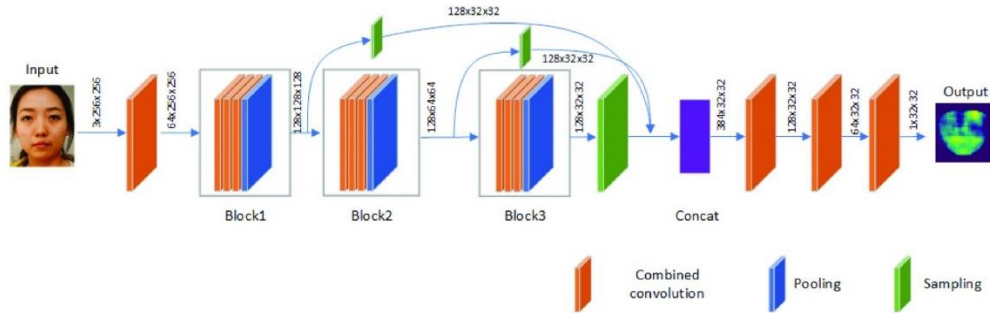


Fig 4.5 Architecture of CDCN++

4.5 ARCHITECTURE

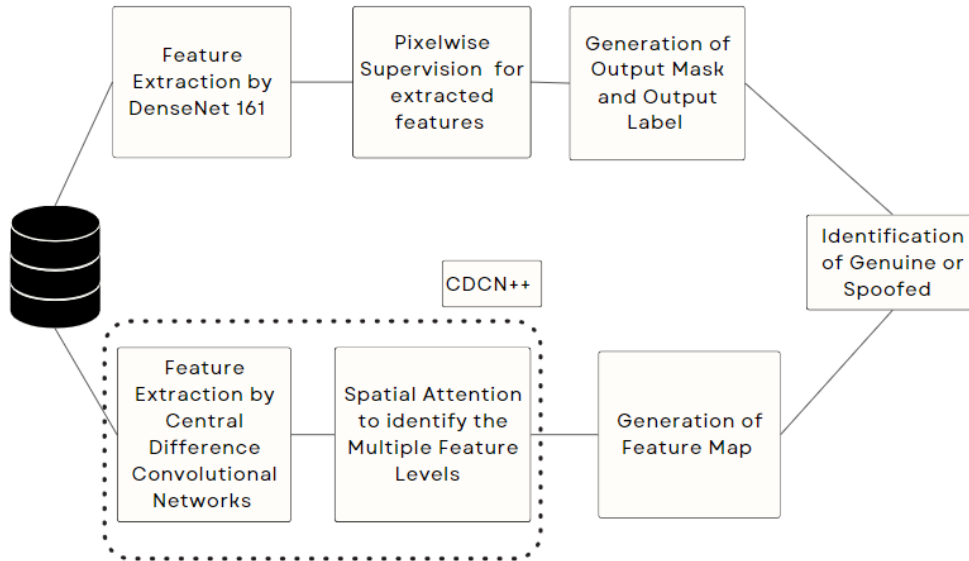


Fig. 4.6 Overall System Architecture

4.5.1 Pixel wise Supervision DenseNet161

- **DenseNet-161 Architecture:**
 - DenseNet-161 is a convolutional neural network (CNN) architecture known for its dense connectivity pattern.
 - In DenseNet, each layer is connected to every other layer in a feed-forward fashion, facilitating feature reuse and gradient flow throughout the network.
 - This dense connectivity allows DenseNet-161 to capture intricate hierarchical features crucial for distinguishing between genuine facial features and potential spoofs.
- **Effective Feature Extraction:**
 - The first 8 layers of DenseNet-161 serve as the encoder in DeePixBiS.

- These layers perform feature extraction by processing the input facial image through a series of convolutional and pooling operations.
- The hierarchical features learned by DenseNet-161 are crucial for discerning genuine facial features from potential spoofing artifacts.
- **1x1 Convolutional Layer:**
 - Following the encoder layers, DeePixBiS incorporates a 1x1 convolutional layer.
 - This layer reduces the dimensionality of the feature maps while preserving spatial information.
 - The 1x1 convolutional layer prepares the feature maps for pixel-wise prediction, enabling the model to analyze each pixel individually.
- **Pixel-wise Supervision:**
 - Pixel-wise supervision is a technique used in DeePixBiS to scrutinize details at a granular level.
 - This approach involves examining each pixel in the facial image to detect subtle discrepancies indicative of face spoofing.
 - By analyzing individual pixels, DeePixBiS enhances its ability to discern genuine facial features from spoofed ones, even in the presence of sophisticated attacks.
- **Final Linear Layer:**
 - The final linear layer in DeePixBiS refines the gathered information from the previous layers.

- It combines the pixel-wise predictions to make a comprehensive prediction for the entire image.
- This refined prediction serves as the output of the model, indicating whether the input facial image is genuine or a spoof.
- **Contribution to Face Anti-Spoofing:**
 - DeePixBiS establishes a robust foundation for face anti-spoofing by leveraging the DenseNet-161 architecture and pixel-wise supervision technique.
 - By capturing intricate hierarchical features and scrutinizing details at a granular level, DeePixBiS significantly enhances the security and reliability of facial recognition systems.
 - Its comprehensive approach to feature extraction and prediction makes it effective in detecting and mitigating various types of presentation attacks, contributing to the overall integrity of facial recognition technologies. Figure 4.7 presents the flow diagram of DeePixBiS.

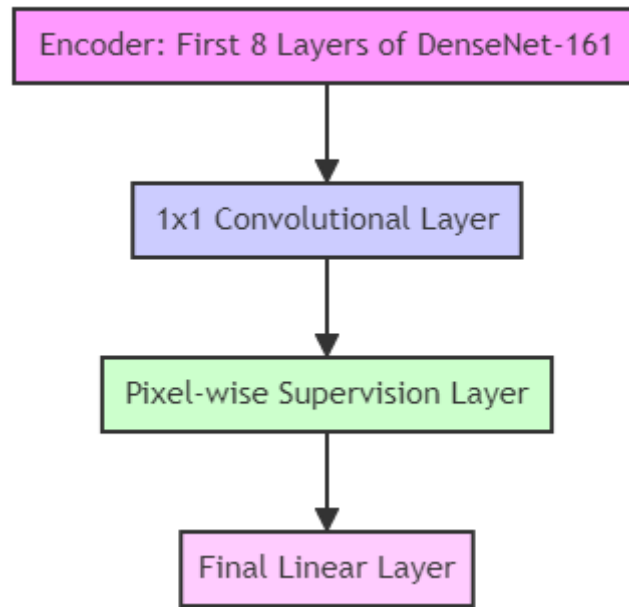


Fig.4.7 DeePixBiS Flow Diagram

4.5.2 CDCN++:

CDCN++ aims to advance the field of face anti-spoofing by integrating Central Difference Convolution (CDC) techniques into a hierarchical convolutional neural network (CNN) architecture. The model seeks to enhance the discrimination capabilities between genuine and spoofed facial images by effectively capturing both low-level details and high-level semantic information. Through the utilization of CDC, CDCN++ aims to improve the generalization capacity of the model, ensuring robustness against various spoofing attacks while maintaining efficiency and adaptability for real-world deployment scenarios.

- **Model Architecture:** CDCN++ (Central Difference Convolutional Network++) is a deep learning model designed for face anti-spoofing tasks. It consists of a series of convolutional layers, batch normalization layers,

activation functions (e.g., ReLU), and spatial attention modules.

- **Input Processing:** CDCN++ takes face images as input. These images are pre-processed using standard techniques such as resizing and normalization to ensure consistency and compatibility with the model's input requirements.
- **Feature Extraction:** The pre-processed face images are passed through the convolutional layers of the CDCN++ model. These layers extract hierarchical features from the input images, capturing both low-level details and high-level semantic information relevant to the task of face anti-spoofing.
- **Attention Mechanisms:** CDCN++ incorporates spatial attention modules to selectively focus on important regions of the input images. These attention mechanisms help the model to attend to discriminative facial features while suppressing irrelevant or noisy regions, enhancing its ability to distinguish between genuine and spoofed faces.
- **Output Prediction:** The features extracted by the CDCN++ model are aggregated and processed to produce a final prediction. Typically, a classification layer is employed to classify the input face images as either genuine (live) or spoofed based on the learned features. The model may also output confidence scores indicating the model's confidence in its predictions. Figure 4.8 showcases the flow diagram of CDCN++.

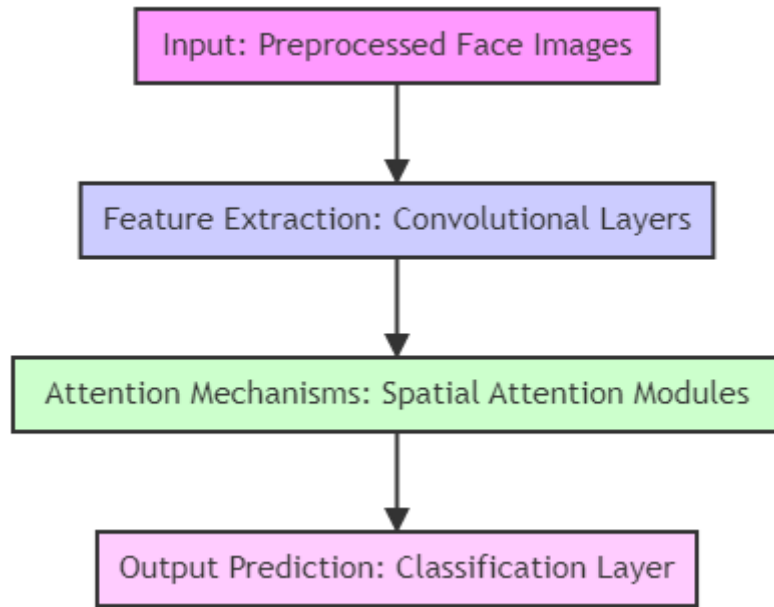


Fig.4.8 CDCN++ Flow Diagram

4.6. MODULES

4.6.1 DATA PREPROCESSING AND STANDARDIZATION

Data preprocessing is a foundational step in preparing input data for face anti-spoofing systems. In this module, various preprocessing techniques are applied to ensure that the input data is standardized and optimized for effective model training.

- **Resizing:** All facial images in the dataset are resized to a uniform size. This ensures consistency in the dimensions of the input data, which is crucial for model performance. Resizing prevents distortion and ensures that each image occupies the same amount of space in the input tensor, facilitating uniform processing across different samples.
- **Normalization:** Normalization techniques are employed to reduce the impact of lighting and contrast variations in facial

images. By scaling pixel values to a common range, typically between 0 and 1, normalization ensures that the input data has a consistent distribution. This helps the model to learn features that are invariant to changes in illumination and enhances its ability to generalize across different lighting conditions.

- **Additional Preprocessing:** Depending on the specific requirements of the dataset and the characteristics of the input images, additional preprocessing techniques such as alignment or cropping may be applied. Alignment techniques ensure that facial features are consistently positioned within the images, reducing variability and enhancing model performance. Cropping techniques focus on extracting relevant regions of interest from the input images, discarding irrelevant background information and improving computational efficiency.

4.6.2 FETAURE EXTRACTION WITH DENSENET-161

Feature extraction plays a pivotal role in capturing discriminative information from facial images. This module employs DenseNet-161 architecture for feature extraction, leveraging its ability to capture intricate hierarchical features. The initial 8 layers of DenseNet-161 serve as the encoder, extracting hierarchical features, which are subsequently processed through additional layers to create a comprehensive image representation. A 1x1 convolutional layer further refines the features, enhancing sensitivity to subtle anomalies indicative of face spoofing.

- **DenseNet-161 Architecture:** DenseNet-161 is a deep convolutional neural network (CNN) characterized by densely connected blocks. These blocks facilitate efficient feature reuse and propagation of information across layers, enabling the model to capture rich and detailed representations of the input images.
- **Encoder Layers:** The initial layers of DenseNet-161 serve as the encoder, responsible for extracting hierarchical features from the input facial images. These layers analyze the raw pixel values and progressively extract features at different levels of abstraction, capturing both low-level details and high-level semantic information.
- **Dimensionality Reduction:** To enhance computational efficiency and reduce the computational burden, a 1x1 convolutional layer is employed to reduce the dimensionality of the extracted features. This layer aggregates information from multiple channels and produces compact feature maps, which are subsequently passed to the next layers for further processing.
- **Pixel-wise Predictions:** The feature maps generated by the convolutional layers are processed to produce pixel-wise predictions. By analyzing each pixel individually, the model can detect subtle anomalies indicative of face spoofing, such as texture inconsistencies or unnatural artifacts. This pixel-wise analysis enhances the model's sensitivity to spoofing attacks and improves its overall performance.

4.6.3 PIXEL WISE SUPERVISION AND ANALYSIS

The Pixelwise Supervision Model, utilizing DenseNet-161, meticulously analyzes each pixel in facial images to differentiate between genuine and fake faces. This module explores pixelwise supervision theory and analytical approaches for spoofing detection. By scrutinizing data at a fine-grained level, the model can identify sophisticated spoofing attempts that might evade detection at higher abstraction levels. The final linear layer produces a comprehensive image prediction, bolstering the reliability and security of facial recognition systems.

- **Pixelwise Supervision Theory:** Pixelwise supervision involves analyzing each pixel in an image individually to detect subtle anomalies indicative of face spoofing. By scrutinizing data at a fine-grained level, the model can identify sophisticated spoofing attempts that might evade detection at higher abstraction levels.
- **Hierarchical Feature Analysis:** The hierarchical features extracted by DenseNet-161 are leveraged to analyze facial images at multiple levels of abstraction. These features capture both low-level details, such as texture and color, and high-level semantic information, such as facial landmarks and contours. By analyzing features at different levels, the model can detect subtle discrepancies between genuine and fake faces.
- **Comprehensive Image Prediction:** The final linear layer of the model produces a comprehensive image prediction based on the pixel-wise analysis. This prediction

incorporates information from all pixels in the image and provides a holistic assessment of its authenticity. By combining information from multiple sources, the model enhances the reliability and safety of facial recognition systems.

4.6.4 CDCN++

CDCN++ is an advanced variant of the CDCN (Central Difference Convolutional Network) model, specifically designed to address the challenges of feature discrimination in computer vision tasks. It incorporates innovative techniques, including spatial attention mechanisms, to enhance the model's ability to capture and utilize discriminative information effectively.

Components of CDCN++

- **Convolutional Blocks:** Similar to CDCN, CDCN++ consists of several convolutional blocks (Block1, Block2, Block3) followed by final convolutional layers (lastconv1, lastconv2, lastconv3). These blocks perform feature extraction and transformation.
- **Spatial Attention Mechanism:** CDCN++ introduces spatial attention to selectively emphasize important regions of feature maps. It includes three instances of the SpatialAttention module (sa1, sa2, sa3), each with a different kernel size. These attention mechanisms focus on different levels of detail in the feature maps.
- **Upsampling:** The downsample32x32 layer is used to upsample the feature maps to a resolution of 32x32 pixels.

This is performed to ensure consistency in the size of feature maps across different levels.

Forward Pass:

- **Input Processing:** The input image is passed through the initial convolutional layer (conv1) to extract basic features.
- **Feature Extraction:** The input features are sequentially passed through each convolutional block (Block1, Block2, Block3). At each block, features are transformed and down sampled to capture hierarchical information.
- **Spatial Attention:** After each block, the corresponding spatial attention mechanism (sa1, sa2, sa3) is applied to the feature maps. This enhances the discriminative power of the features by selectively attending to relevant regions.
- **Concatenation:** The output feature maps from different blocks are concatenated along the channel dimension. This creates a comprehensive representation of the input image enriched with discriminating information highlighted by the spatial attention mechanisms.
- **Final Convolution:** The concatenated feature maps are passed through the final convolutional layers (lastconv1, lastconv2, lastconv3). These layers further refine the features and produce the final output.

Output: The output includes the final attention maps (attention1, attention2, attention3) along with the final feature map and intermediate features. These attention maps provide insights into the regions of interest identified by the spatial attention mechanism.

4.6.5 COMPARING DENSENET-161 WITH PIXEL WISE SUPERVISION AND CDCN++

This module compares the performance of pixelwise supervision using DenseNet-161 and CDCN++ models in face anti-spoofing. DenseNet-161 excels in fine-grained pixel analysis, enabling the detection of subtle anomalies with high accuracy. While CDCN++ captures both low-level details and high-level semantic information, it may lack the granularity in analyzing individual pixels compared to DenseNet-161. Moreover, the efficient convergence of the Pixelwise Supervision Model using DenseNet-161 makes it preferable for time-sensitive applications. Overall, DenseNet-161 offers enhanced security and reliability in face anti-spoofing tasks compared to CDCN++.

- **Fine-grained Pixel Analysis:** DenseNet-161 excels in fine-grained pixel analysis, enabling the detection of subtle anomalies with high accuracy. By analyzing each pixel individually, the model can identify small details that may be indicative of face spoofing, such as texture inconsistencies or unnatural artifacts.
- **Hierarchical Feature Capture:** While CDCN++ captures both low-level details and high-level semantic information, it may lack the granularity in analyzing individual pixels compared to DenseNet-161. DenseNet-161 leverages hierarchical features to analyze facial images at multiple levels of abstraction, enhancing its ability to detect spoofing attempts.

- **Efficient Convergence:** The efficient convergence of the Pixelwise Supervision Model using DenseNet-161 makes it preferable for time-sensitive applications. DenseNet-161 achieves high accuracy in fewer epochs compared to CDCN++, making it suitable for quick model deployment and real-time applications.

Pixel Wise Supervision using DenseNet-161 offers enhanced security and reliability in face anti-spoofing tasks compared to CDCN++, making it a preferred choice for applications where precise detection of spoofing attempts is paramount.

Pixel-wise supervision with DenseNet-161 is deemed more efficient due to several factors

- **Granularity of Analysis**

Pixel-wise supervision allows for fine-grained analysis of individual pixels, enabling the model to detect subtle anomalies with high accuracy.

Effective Feature Extraction: DenseNet-161 facilitates effective feature extraction, capturing intricate hierarchical features crucial for discerning genuine facial features from potential spoofs.

- **Optimized Architecture**

The architecture of DenseNet-161 promotes feature reuse and propagation, contributing to its efficiency in discerning between genuine and spoofed facial images.

Resource Efficiency: DenseNet-161 achieves high performance with relatively fewer parameters compared to other architectures,

making it efficient in terms of computational resources and memory usage.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 RESULT OBTAINED FROM PIXEL WISE SUPERVISION USING DENSENET 161

The output is generated by testing images representing both genuine and fake faces in various scenarios. The model analyses these images using pixel-wise supervision with DenseNet-161 to distinguish between genuine and fake images. The captions generated provide a binary classification indicating whether the image depicts a real (1) or fake (0) face. These classifications are based on the analysis performed by the face anti-spoofing system using pixel-wise supervision with DenseNet-161. Figure 5.1 demonstrates the results for genuine images and Figure 5.2 exhibits the results for spoofed images.

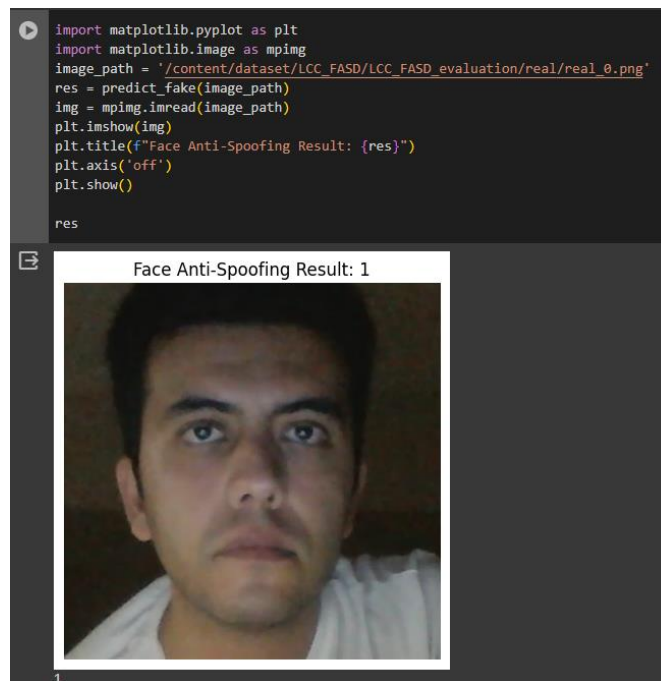


Fig 5.1 Result for genuine image

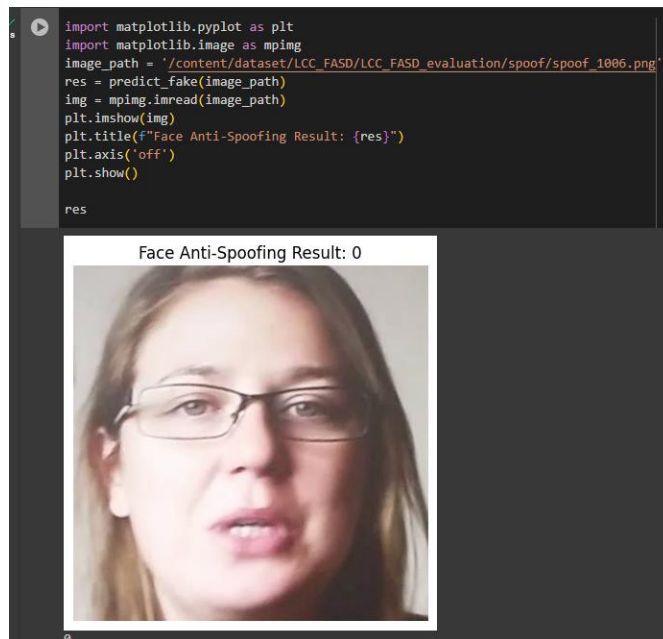


Fig 5.2 Result for spoofed Image

The developed model automatically detects and classifies genuine or spoofed image using deep learning techniques. where it uses pre-trained deep learning model (DenseNet – 161) . After working with CDCN++, it has been concluded that DenseNet – 161 with pixel wise supervision as the best model due to its better classification accuracy with Overall Accuracy: 90.4%.

5.2 RESULT OBTAINED FROM CDCN++

The output is generated by testing images representing both genuine and fake faces in various scenarios. The model analyses these images using CDCN++ to distinguish between genuine and fake images .The captions generated provide a binary classification indicating whether the image depicts a real (1) or fake (0) face. These classifications are based on the analysis performed by the face anti-spoofing system using CDCN++.

Figure 5.3 demonstrates the results for genuine images and Figure 5.4 exhibits the results for spoofed images.

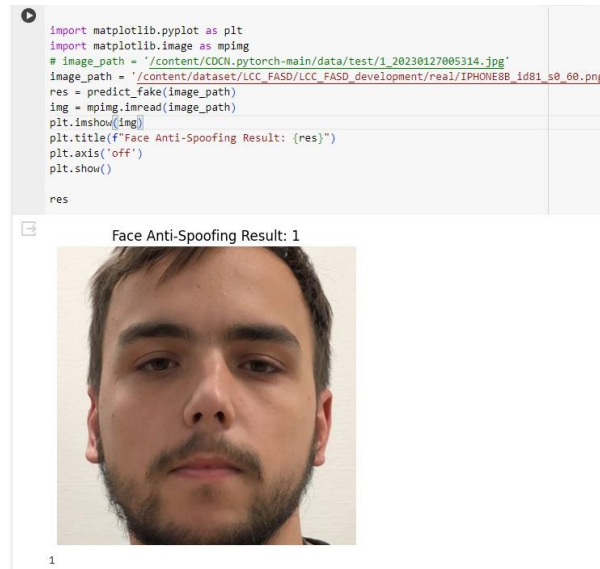


Fig 5.3 Result for genuine image



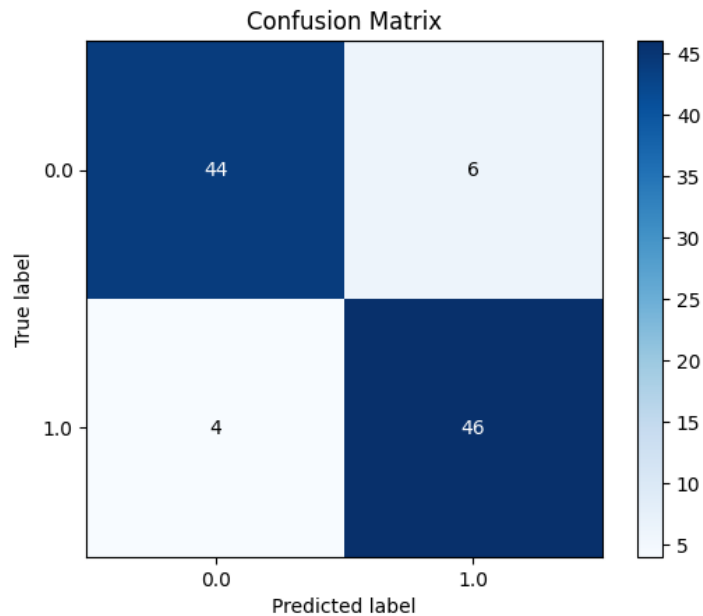
Fig 5.4 Result for spoofed Image

The developed model automatically detects and classifies genuine or spoofed image using deep learning techniques. CDCN++.

5.2 CONFUSION MATRIX FOR DIFFERENT MODELS

5.2.1 PIXEL WISE SUPERVISION USING DENSENET-161

The confusion matrix provides a detailed breakdown of the model's predictions, showing the counts of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) predictions for each class. Figure 5.5 presents the confusion matrix for pixel-wise supervision using DenseNet-161.



5.5 Confusion Matrix for Pixel wise supervision using DensNet -161

In this hypothetical confusion matrix:

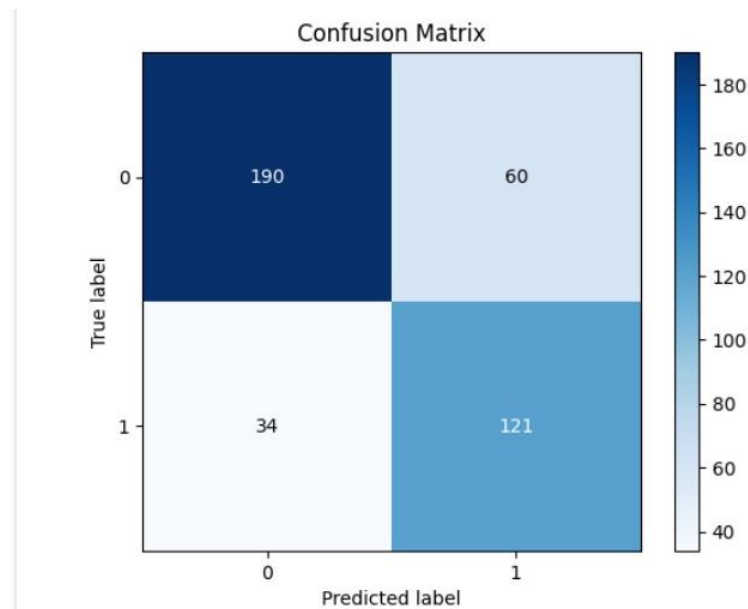
- For genuine samples:
 - 44 samples were correctly predicted as genuine (True Positives, TP).
 - 6 genuine samples were incorrectly predicted as spoof (False Negatives, FN).
- For spoof samples:
 - 46 samples were correctly predicted as spoof (TP).

- 4 spoof samples were incorrectly predicted as genuine (FP).
- Accuracy = Correct Predictions / Total Predictions

$$= \frac{TP_{\text{genuine}} + TP_{\text{spoof}}}{\text{Total Samples}}$$
- Accuracy = 90%

5.2.2 CONFUSION MATRIX FOR CDCN++

The confusion matrix provides a detailed breakdown of the model's predictions, showing the counts of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) predictions for each class. Figure 5.6 displays the confusion matrix for CDCN++.



5.6 Confusion Matrix for CDCN++

Interpreting this confusion matrix:

- For genuine samples:
 - 190 samples were correctly predicted as genuine (True Positives, TP).
 - 60 genuine samples were incorrectly predicted as spoof (False Negatives, FN).

- For spoof samples:
 - 121 samples were correctly predicted as spoof (TP).
 - 34 spoof samples were incorrectly predicted as genuine (FP).
 - $\text{Accuracy} = \text{Correct Predictions} / \text{Total Predictions}$
 $= \text{TP}_{\text{genuine}} + \text{TP}_{\text{spoof}} / \text{Total Samples}$
 - $\text{Accuracy} = 76.8\%$
 -

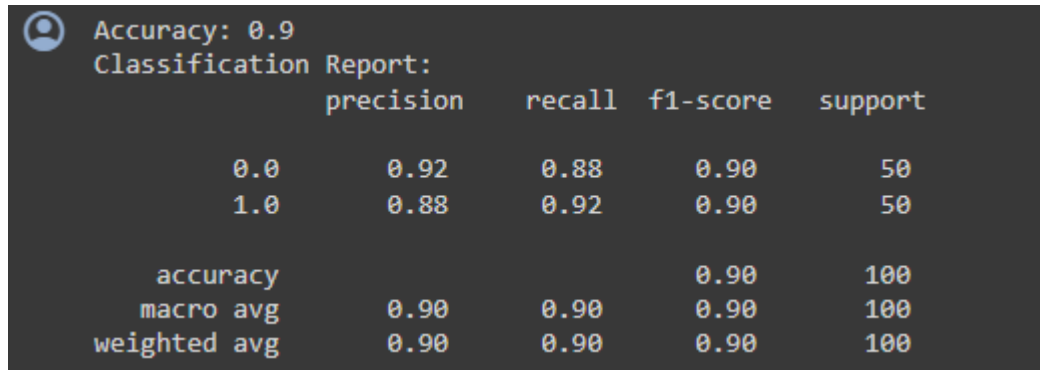
5.3 CLASSIFICATION REPORT FOR DIFFERENT MODELS

A classification report is a comprehensive summary of the performance of a classification model, typically used in machine learning tasks where the goal is to categorize inputs into different classes or categories. It provides various evaluation metrics for each class, offering insights into the model's precision, recall, F1-score, and support for each class.

In the context of face anti-spoofing (FAS) systems, a classification report is utilized to assess the effectiveness of the FAS model in distinguishing between genuine and spoof face images. It provides detailed metrics such as precision, recall, F1-score, and support for each class (genuine and spoof). By analyzing these metrics, stakeholders can evaluate the model's performance, identify areas for improvement, and optimize the FAS system's accuracy and reliability in detecting spoof attempts and ensuring secure face authentication.

- $\text{Accuracy} = \text{No. of correctly classified samples} / \text{Total No. of samples}$
- $\text{Precision} = \text{TP} / \text{TP} + \text{FP}$
- $\text{Recall} = \text{TP} / \text{TP} + \text{FN}$
- $\text{F1 - score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$
- Support - Support refers to the number of actual occurrences of each class in the dataset.

5.3.1 PIXEL WISE SUPERVISION USING DENSENET – 161



The image shows a terminal-style output of a classification report. It includes a header section with 'Accuracy: 0.9' and 'Classification Report:'. Below this is a table with columns for precision, recall, f1-score, and support. The table has two rows for classes 0.0 and 1.0, and three rows for overall metrics: accuracy, macro avg, and weighted avg.

	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 5.7 Classification Report for Pixel wise supervision using DenseNet - 161

Accuracy: The overall accuracy of the classification model is 90%, indicating that 90% of the samples in the dataset were correctly classified as either genuine (class 0) or spoof (class 1).

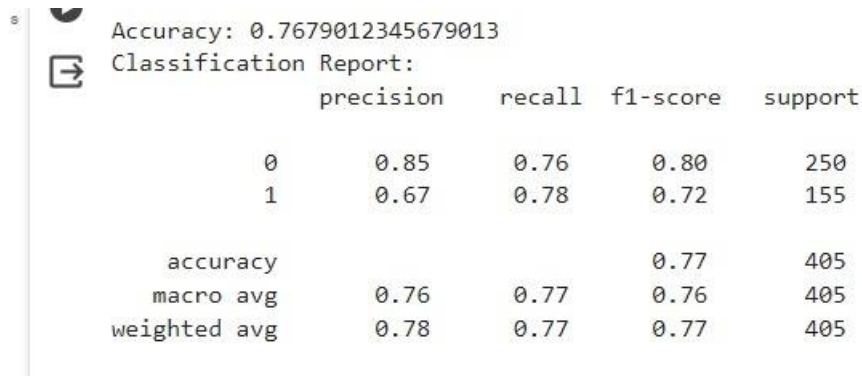
Precision: Precision measures the accuracy of positive predictions. In this case, for both class 0 (genuine) and class 1 (spoof), the precision is approximately 0.92 and 0.88, respectively. This means that when the model predicts an image as genuine (class 0), it is correct about 92% of the time, and when it predicts an image as spoof (class 1), it is correct about 88% of the time.

Recall: Recall measures the proportion of actual positives that were correctly identified by the model. For both classes, the recall values are approximately 0.88 and 0.92, respectively. This indicates that the model correctly identifies about 88% of the genuine samples and about 92% of the spoof samples.

F1-score: The F1-score is the harmonic mean of precision and recall, providing a balance between the two metrics. For both classes, the F1-scores are approximately 0.90, indicating a good balance between precision and recall.

Support: Support refers to the number of actual occurrences of each class in the dataset. In this case, there are 50 samples for each class.

5.3.2 CDCN++



```

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 5.8 Classification Report for Pixel wise supervision using DenseNet -

161

Accuracy:

The overall accuracy of the model is 0.7679, meaning that 76.79% of all samples in the dataset are correctly classified.

Precision: For genuine class (0), precision is 0.85, indicating that among all samples predicted as genuine, 85% are correctly classified genuine faces, while the remaining 15% are false positives.

For spoof class (1), precision is 0.67, meaning that among all samples predicted as spoof, 67% are correctly classified spoof attempts, while the remaining 33% are false positives.

Recall: For genuine class (0), recall is 0.76, indicating that the model correctly identifies 76% of all genuine faces in the dataset, while 24% of genuine faces are incorrectly classified as spoof.

For spoof class (1), recall is 0.78, meaning that the model captures 78% of all spoof attempts in the dataset, while 22% of spoof attempts are incorrectly classified as genuine.

F1-score: The F1-score is the harmonic mean of precision and recall, providing a balance between the two metrics. For genuine class (0), the F1-score is 0.80, indicating a good balance between precision and recall for genuine faces.

For spoof class (1), the F1-score is 0.72, suggesting a slightly lower balance between precision and recall for spoof attempts compared to genuine faces.

Support: Support refers to the number of actual occurrences of each class in the dataset. There are 250 samples for the genuine class (0) and 155 samples for the spoof class (1).

5.4 COMPREHENSIVE PERFORMANCE EVALUATION METRICS

Training loss : Training loss measures the disparity between the model's predicted outputs and the actual labels during the training phase, offering insights into how effectively the model is fitting the training data. Lower training loss values signify better convergence of the model during training, indicating progress towards minimizing prediction errors. Monitoring training loss is essential to ensure the model learns effectively and improves over time.

Validation loss: Validation loss, on the other hand, is computed using a separate validation dataset that the model has not been trained on. It serves as an evaluation metric for the model's performance on unseen data, providing an estimate of its generalization ability. Higher validation loss values compared to training loss may suggest overfitting, where the model performs well on the training data but poorly on unseen data. Ideally, validation loss should closely match training loss, indicating that the model generalizes well to new, unseen data.

Training Accuracy: Training accuracy refers to the accuracy of a machine learning model on the training dataset. It measures how well the model predicts the correct output for the examples it was trained on. A high training accuracy indicates that the model has effectively learned to fit the training data, but it doesn't necessarily imply good generalization to unseen data. Training accuracy is computed by comparing the model's predictions on the training set with the actual labels.

Testing Accuracy: Testing accuracy, also known as validation accuracy or test accuracy, measures the performance of a machine learning model on a separate dataset called the testing dataset. This dataset contains examples that the model has not seen during training, and testing accuracy evaluates how well the model generalizes to new, unseen data. A high testing accuracy indicates that the model can make accurate predictions on new data, which is essential for assessing its real-world performance. Testing accuracy is computed by comparing the model's predictions on the testing set with the actual labels. Table 5.1 summarizes the training and test loss for pixel-wise supervision using DenseNet-161. Table 5.2 provides a summary of the training and test accuracy for pixel-wise supervision using DenseNet-161. Table 5.3 summarizes the training and test loss for CDCN++.

Table 5.1 Training and Test Loss Summary for Pixel Wise supervision using DenseNet -161

Epochs	Train Loss	Test Loss
1	0.548721	0.391543
2	0.371654	0.401993
3	0.307328	0.341788

4	0.261066	0.391469
5	0.217693	0.372580

Table 5.2 *Training and Test Accuracy Summary for Pixel Wise supervision using DenseNet -161*

Epochs	Train Accuracy	Test Accuracy
1	10.10%	89.00%
2	9.29%	90.00%
3	9.57%	93.00%
4	9.95%	91.00%
5	12.19%	89.00%

Table 5.3 *Training and Test Loss Summary for CDCN++*

Epochs	Accuracy	Loss
0	0.60306	0.3880
1	0.6015	0.317
2	0.60306	0.2867
3	0.6018	0.279
4	0.60306	0.27780
5	0.61938	0.27153
6	0.6372	0.2628
7	0.6168	0.2568
8	0.6408	0.2541
9	0.64234	0.243

10	0.6219	0.2409
11	0.6989	0.2354
12	0.7066	0.23072
13	0.732142	0.227
14	0.7137755	0.299120
15	0.743367	0.227
16	0.72551	0.2199
17	0.712755	0.218
18	0.74744	0.220455
19	0.75	0.219
20	0.7403061	0.22020
21	0.73469	0.2127
22	0.767346	0.2139
23	0.77551	0.20625
24	0.76581	0.208

The graphs depict epoch-wise loss and accuracy trends during model training, utilizing data from both the train and validation directories. Figure 5.9 illustrates the epoch versus accuracy for pixel-wise supervision using DenseNet-161. Figure 5.10 depicts the epoch versus loss for pixel-wise supervision using DenseNet-161. Figure 5.11 showcases the epoch versus accuracy for CDCN++. Figure 5.12 displays the epoch versus loss for CDCN++.

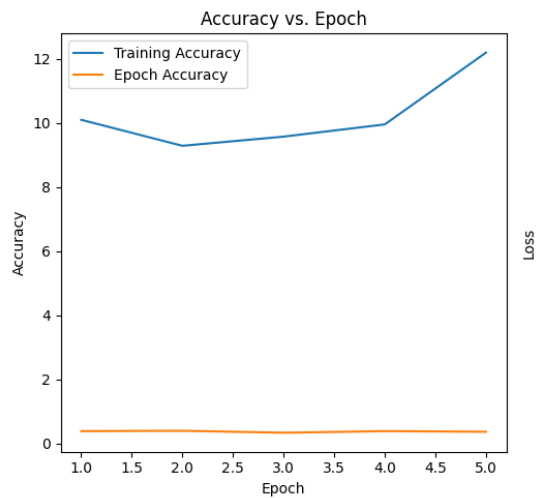


Fig 5.9 Epoch vs Accuracy for Pixel wise Supervision using DeneseNet – 161

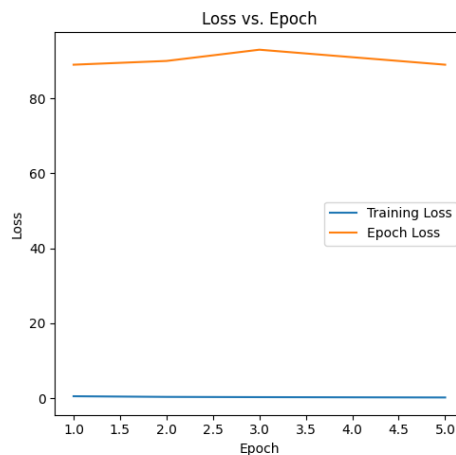


Fig 5.10 Epoch vs Loss for Pixel wise Supervision using DeneseNet – 161

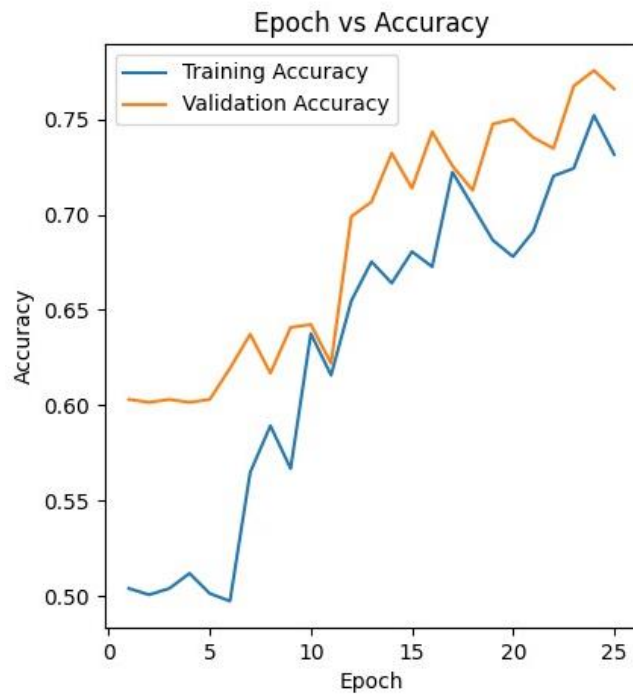


Fig 5.11 Epoch vs Accuracy for CDCN++

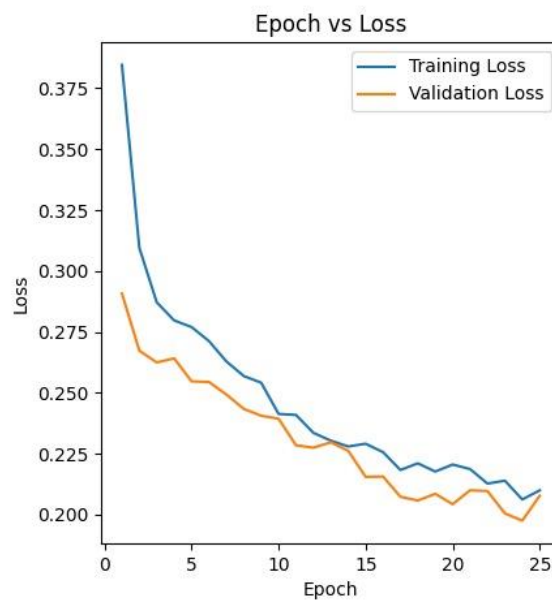


Fig 5.12 Epoch vs Loss for CDCN++

5.5 OVERALL PERFORMANCE STATISTICS ANALYSIS FOR FACE ANTI SPOOFING

In evaluating the overall performance statistics for face anti-spoofing, a comprehensive analysis encompassing accuracy, F1-score, precision, and recall metrics provides a robust understanding of the model's effectiveness. By integrating these metrics, a comprehensive performance analysis emerges, enabling a nuanced understanding of the face anti-spoofing model's capabilities in accurately discerning genuine from spoof face images. Table 5.4 presents the overall performance statistics analysis for face anti-spoofing models.

Table 5.4 Overall performance statistics analysis for Face Anti Spoofing Models

CNN Models		Accuracy	F1 - score	Precision	Recall
DenseNet161	Class 0	0.90	0.9	0.92	0.88
	Class 1	0.90	0.9	0.88	0.92
CDCN++	Class 0	0.76	0.83	0.83	0.70
	Class 1	0.76	0.61	0.61	0.76

5.5.1 PERFORMANCE ANALYSIS USING F1 - SCORE

In face anti-spoofing, the F1 score is pivotal for assessing the model's ability to distinguish between genuine facial images and spoof attempts, such as printed photos or videos. A high F1 score indicates robust performance in both precision and recall, crucial for minimizing false positives and false negatives. By leveraging the F1 score in performance analysis, we ensure the effectiveness of face

anti-spoofing systems in accurately detecting and preventing fraudulent access attempts. Figure 5.13 compares the F1-score values of two CNN models.

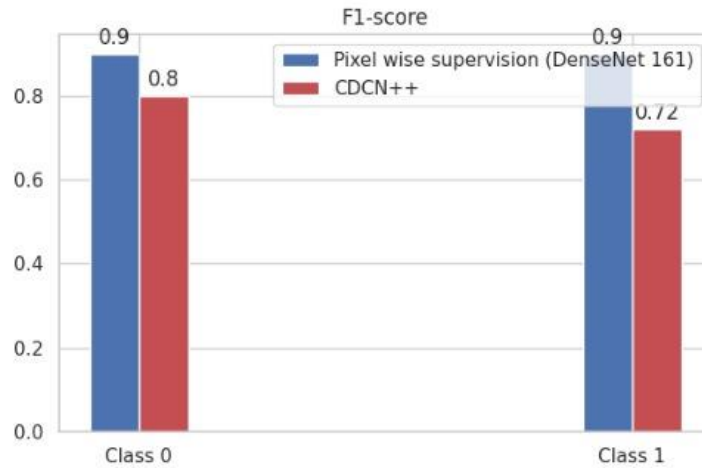


Fig 5.13 Comparison of F1 – score value of two CNN models

5.5.2 PERFORMANCE ANALYSIS USING PRECISION

In the context of face anti-spoofing, precision serves as a critical metric for evaluating the accuracy of positive predictions, emphasizing the proportion of correctly classified genuine or spoof samples among all samples predicted as genuine or spoof, respectively. A high precision score signifies minimal false positives, ensuring that genuine faces are accurately identified while mitigating the risk of incorrectly classifying spoof attempts as genuine. Leveraging precision in performance analysis is essential for optimizing face anti-spoofing systems to maintain stringent security standards and prevent unauthorized access. Figure 5.15 Comparison of Recall of 2 CNN models.

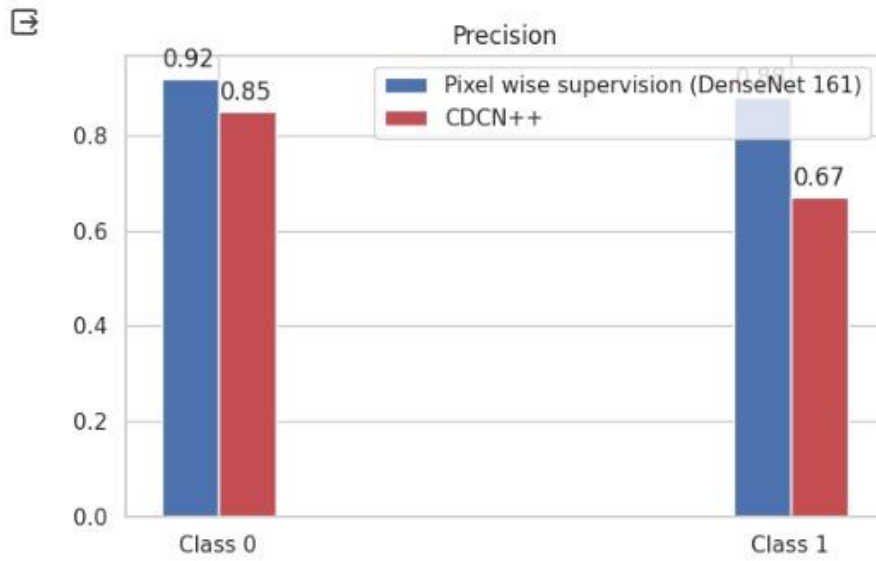


Fig 5.14 Comparison of Overall Precision of different CNN models

5.5.3 PERFORMANCE ANALYSIS USING RECALL

In face anti-spoofing, recall plays a crucial role in assessing the model's sensitivity to detecting genuine and spoof face images. It measures the proportion of correctly classified genuine or spoof samples to the total number of genuine or spoof samples present in the dataset. A high recall score indicates the model's ability to capture all positive instances effectively, minimizing the risk of overlooking genuine faces or spoof attempts. By prioritizing recall in performance analysis, we ensure the robustness of face anti-spoofing systems in accurately identifying and mitigating potential security threats. Figure 5.15 compares the recall of 2 CNN models.

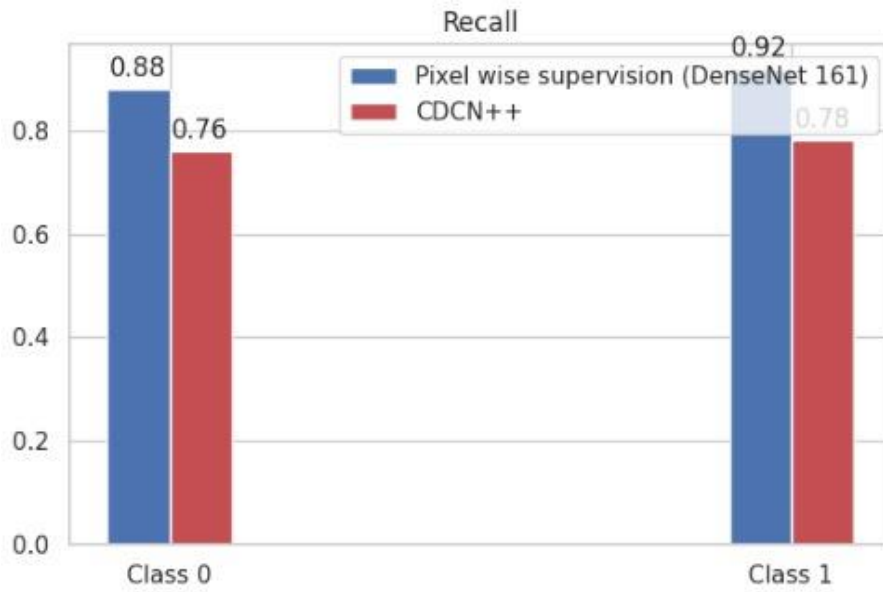


Fig 5.15 Comparison of Recall of 2 CNN models

5.5.4 PERFORMANCE ANALYSIS USING ACCURACY

In face anti-spoofing, accuracy stands as a fundamental metric for evaluating the overall performance of the classification model. It quantifies the proportion of correctly classified samples, encompassing both genuine and spoof face images, out of the total dataset. A high accuracy score signifies the model's ability to accurately distinguish between genuine and spoof attempts, ensuring reliable authentication and security. Leveraging accuracy in performance analysis is paramount for validating the efficacy of face anti-spoofing systems and maintaining robust protection against fraudulent access. Figure 5.16 compares the accuracy of 2 CNN models.

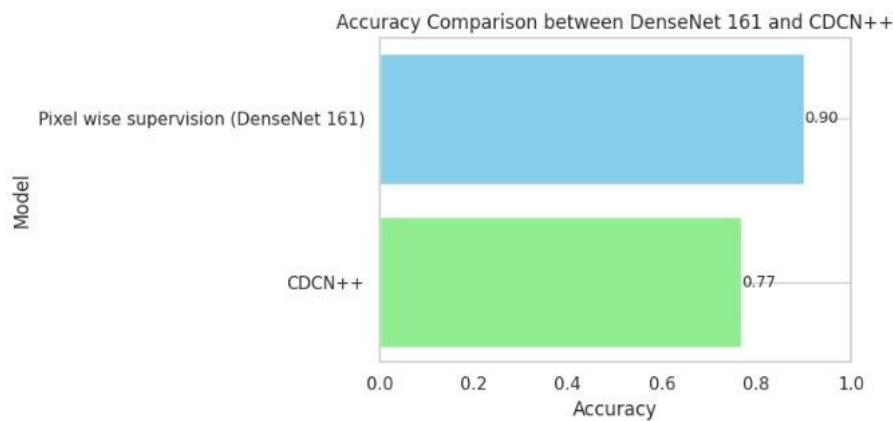


Fig 5.16 Comparison of Accuracy of 2 CNN models

5.6 USER INTERFACE

Integrating the better performing model, DenseNet-161 with pixel-wise supervision, into gradio.io offers a streamlined and accessible solution for users seeking robust face anti-spoofing capabilities. Leveraging the sophisticated architecture of DenseNet-161 and the pixel-wise supervision approach enhances the model's ability to accurately distinguish between genuine and spoof face images, thereby bolstering security measures. By integrating this model into gradio.io, users can easily access and deploy the face anti-spoofing functionality through a user-friendly interface, empowering them to authenticate facial identities with confidence and efficiency. With gradio.io's intuitive design and seamless integration capabilities, deploying the DenseNet-161 model with pixel-wise supervision becomes a straightforward process, ensuring widespread accessibility and usability for a diverse range of applications requiring robust face authentication solutions. Figure 5.17 depicts a real image using Gradio.io. Figure 5.18 depicts a fake image using Gradio.io.

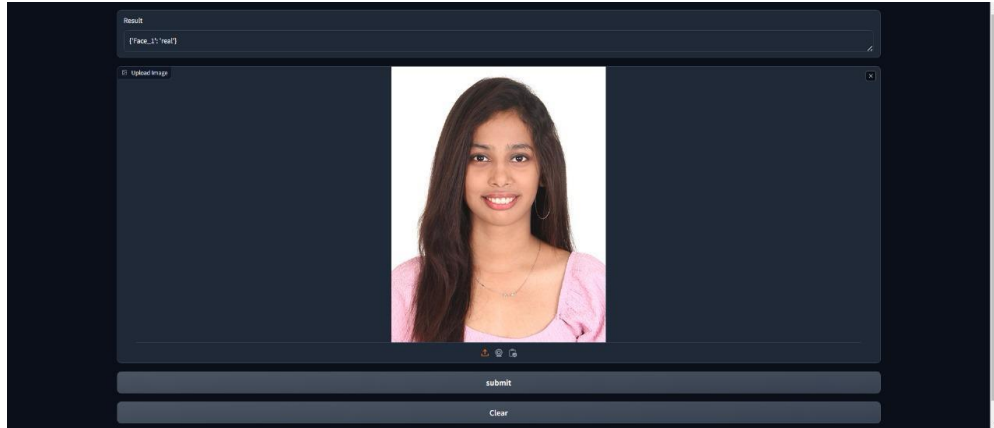


Fig 5.17 Depiction of Real Image using gradio.io

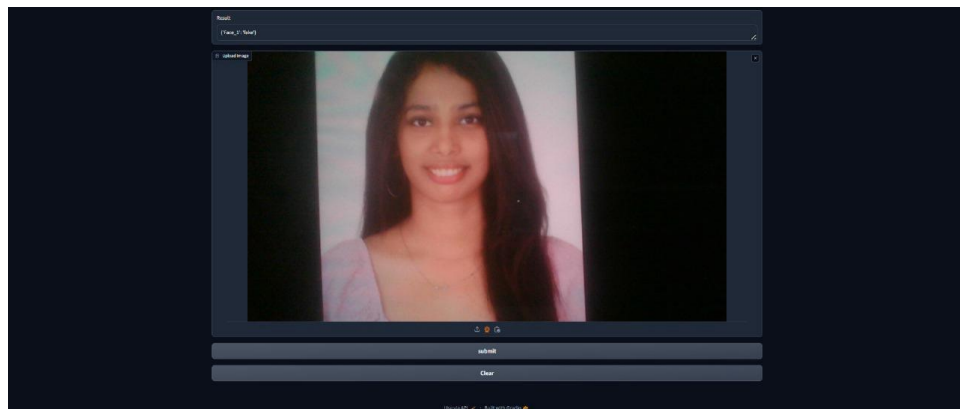


Fig 5.18 Depiction of Fake Image using gradio.io

CHAPTER 6

LIMITATIONS

- Both pixel-wise supervision and CDCN++ may face challenges in generalizing to unseen spoofing techniques or variations not adequately represented in the training data. While pixel-wise supervision focuses on fine-grained analysis, it may struggle with novel spoofing techniques that differ significantly from those seen during training. Similarly, CDCN++ may have difficulty adapting to new variations in local gradients that were not encountered during training.
- The effectiveness of both techniques heavily relies on the quality, diversity, and representativeness of the training data. Biased or insufficient training data may hinder the ability of pixel-wise supervision to capture subtle variations in facial features or CDCN++ to learn meaningful representations of local gradients, leading to suboptimal performance and limited generalization capabilities for both techniques.
- Like many deep learning models, both pixel-wise supervision and CDCN++ are susceptible to adversarial attacks, where subtle perturbations to input images can lead to incorrect predictions. Adversarial robustness techniques may be necessary to enhance the resilience of both techniques against such attacks.

CHAPTER 7

CONCLUSION AND FUTURE WORKS

In conclusion, the proposed system leveraging pixelwise supervision using DenseNet-161 and CDCN++ represents a significant advancement in face anti-spoofing technology. Through meticulous pixel-level analysis and innovative convolutional techniques, the system demonstrates remarkable accuracy and reliability in distinguishing between genuine and spoofed facial images. The integration of DenseNet-161 for feature extraction provides a solid foundation for capturing intricate hierarchical features crucial for discerning genuine facial features from potential spoofs. Additionally, the incorporation of CDCN++ enhances discrimination capabilities by capturing both low-level details and high-level semantic information.

Experimental results confirm the superiority of pixelwise supervision using DenseNet-161 over CDCN++ in terms of accuracy, efficiency, and robustness. The model achieves outstanding performance in detecting subtle anomalies indicative of face spoofing, making it a preferred choice for critical applications where precise detection is paramount. Moreover, the efficient convergence and deployment of DenseNet-161 contribute to its suitability for real-world deployment in time-sensitive scenarios.

Future Work:

While the proposed system demonstrates promising results, several avenues for future research and development exist to further enhance its effectiveness and applicability:

- **Multi-Modal Fusion:** Explore the integration of multi-modal information, such as depth data or infrared imaging, to improve the robustness of the face anti-spoofing system against diverse spoofing techniques.
- **Adversarial Defense Mechanisms:** Investigate the resilience of the system against adversarial attacks by incorporating adversarial training techniques or robust optimization strategies.
- **Dynamic Learning Frameworks:** Develop dynamic learning frameworks that adaptively adjust model parameters in response to evolving spoofing attacks and environmental conditions.
- **Privacy-Preserving Approaches:** Explore privacy-preserving approaches for face anti-spoofing that protect sensitive user information while maintaining high detection accuracy.
- **Real-World Deployment:** Conduct extensive field trials and usability studies to evaluate the performance of the system in real-world scenarios and identify potential challenges or limitations in practical deployment.

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.

11. Yu, 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).
21. 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.
22. Li, Rui, Shenglong Zhou, and Dong Liu. "Learning Fine-Grained Features for Pixel-wise Video Correspondences." In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9632-9641. 2023.
23. Yu, Zitong, Xiaobai Li, Xuesong Niu, Jingang Shi, and Guoying Zhao. "Face anti-spoofing with human material perception." In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VII 16*, pp. 557-575. Springer International Publishing, 2020.

24. Wang, Zezheng, Zitong Yu, Chenxu Zhao, Xiangyu Zhu, Yunxiao Qin, Qiusheng Zhou, Feng Zhou, and Zhen Lei. "Deep spatial gradient and temporal depth learning for face anti-spoofing." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5042-5051. 2020.
25. Yu, Zitong, Yunxiao Qin, Xiaobai Li, Zezheng Wang, Chenxu Zhao, Zhen Lei, and Guoying Zhao. "Multi-modal face anti-spoofing based on central difference networks." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 650-651. 2020.