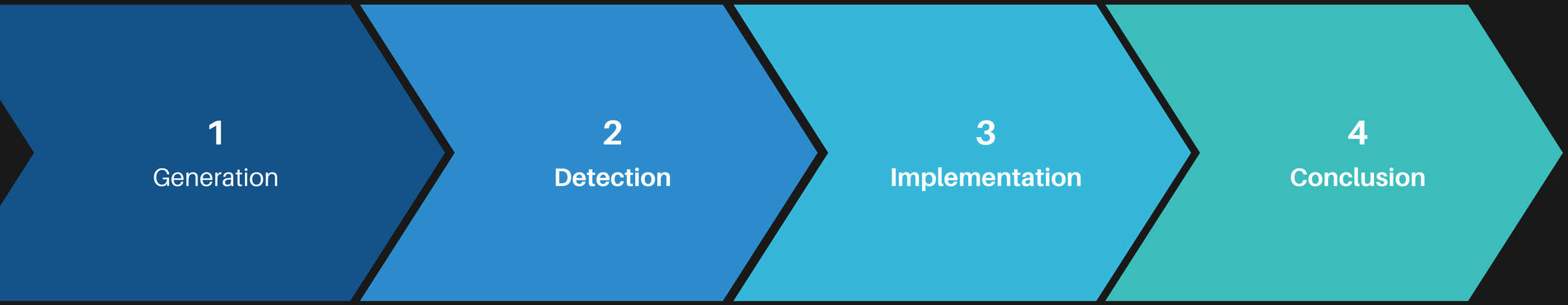
DEEPFAKES GENERATION AND DETECTION

BY TEAM 3

LAYOUT

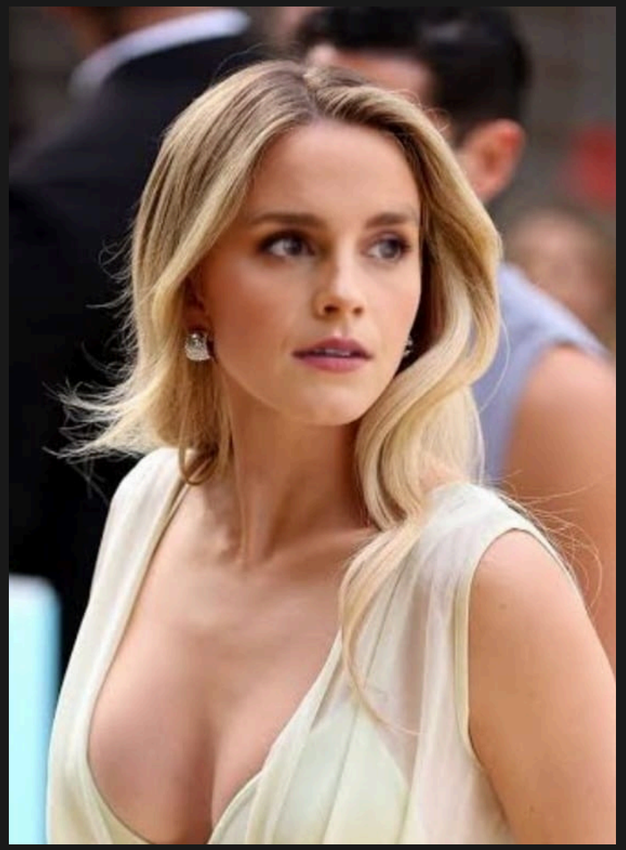


In this session, we will discuss the tools and metrics utilized for creating and evaluating deepfakes. We will also go through the implementation of deepfake detection project !

DEEPFAKE GENERATION

Few of our generated content






1

NODOWN

2

GLFF

3

DMIMAGEDETECTION

4

HIFI

5

CLIP-VIT

METRICS USED!

We’ve also curated a comprehensive list of tested tools used to generate Deepfakes:

Deepfake Generation

NAME	TOOL USED	NO. OF IMAGE	ACCURACY
Shree	Face ai tool	43	Low to medium
	FaceDancer	1065	Medium to High
	hugging face (stable diffusion)	56	Low to Medium
	faceswap	66	Low to Medium
Subham	Vidnoz	325	Medium to Hign
	Reface	567	Low to Medium
	Remaker	304	Medium to High
Omkar	Roop	700	Medium to High
	Faceswapper	8	Low to Medium
	Remaker ai	14	Low to Medium
Mohnish	FaceFusion	247	Medium to High
	SimSwap	13	Low to Medium
	SwapFace	16	Low to Medium

NAME	TOOL USED	NO. OF IMAGE	ACCURACY
Anirudh	SwapFace	50	Low to Medium
	HuggingFace Model- SwapFace	20	Medium to High
	ArtGuru	50	Low to Medium
	FaceSwapper	50	Low to Medium
	Akool	50	Medium to High
Aditi	Roop	40	Low to High
	HuggingFace	35	Medium to High
	SwapFace	25	Low to Medium
Ganesh Kumar	Roop Unleashed	180	Medium to High
	Akool	23	medium
	Face Swapper	18	Low to Medium
	Remaker	18	Low to medium
	Miocreate	31	Medium To
	Hugging Face	21	Medium to High

NAME	TOOL USED	NO. OF IMAGE	ACCURACY
Arnav Sud	FaceSwapper	30	Medium to High
	Insightface	20	Medium to High
	pixlr	50	Low to Medium
	HuggingFace	20	Medium

KEY FINDINGS

1. Dataset Creation: Successfully generated a large and diverse dataset of deepfakes using various tools.
2. Metrics Validation: Established a method to assess the validity and quality of the generated deepfakes.
3. Data Volume: Generated a substantial amount of data suitable for training and testing deepfake detection models.

PHASE 2
DEEPFAKE DETECTION

BENCHMARK DATASETS

1. DFFD: Diverse Fake Face dataset

2. FFHQ: Flickr Faces HQ 70K from StyleGAN

3. SRM: created by us

CHALLENGES AND LIMITATIONS

- Newer deepfake methods often find ways to bypass standard detection techniques.
- Detection systems that work well in controlled settings often struggle in real-world conditions
- The availability of GPUs and the difficulty of accommodating large datasets in the working space are significant concerns.

HYPERVERGE DEEPFAKE SYSTEM IN KYC FRAUD PREVENTION

1
Unparalleled Accuracy : 98.5%

2
Global Reach : 195+ countries covered

3
Lightning-Fast Detection in under 3 sec

KEY FINDINGS

1. Deepfake Evolution: Gained insights into the ongoing evolution of deepfake generation techniques.
2. Detection Challenge: Generalizability Gap Deepfake detection models often struggle to generalize to unseen deepfake techniques, requiring continuous adaptation.

PHASE 3
DEEPFAKE DETECTION IMPLEMENTATION
(SRM DETECTOR)

VGG16 MODEL

- MTCNN :
  - for face feature extraction
  - - face location and cropping
- Preprocessing layer :
  - SRM filter kernel
- Pretrained model with custom FC Layers for classification

- Input Shape : 128 x 128 x 3
- Optimizer : SGD
- Learning rate : 0.01
- Weight Decay : 1 x 10^-6
- Momentum : 0.9
- Loss Function : Categorical Crossentropy
- Epochs : 125

Model: "model_5"

Layer (type)	Output Shape	Param #
input_6 (InputLayer)	[(None, 128, 128, 3)]	0
block1_conv1 (Conv2D)	(None, 128, 128, 64)	1792
block1_conv2 (Conv2D)	(None, 128, 128, 64)	36928
block1_pool (MaxPooling2D)	(None, 64, 64, 64)	0
block2_conv1 (Conv2D)	(None, 64, 64, 128)	73856
block2_conv2 (Conv2D)	(None, 64, 64, 128)	147584
block2_pool (MaxPooling2D)	(None, 32, 32, 128)	0
block3_conv1 (Conv2D)	(None, 32, 32, 256)	295168
block3_conv2 (Conv2D)	(None, 32, 32, 256)	590080
block3_conv3 (Conv2D)	(None, 32, 32, 256)	590080
block3_pool (MaxPooling2D)	(None, 16, 16, 256)	0
block4_conv1 (Conv2D)	(None, 16, 16, 512)	1180160
block4_conv2 (Conv2D)	(None, 16, 16, 512)	2359808
block4_conv3 (Conv2D)	(None, 16, 16, 512)	2359808
block4_pool (MaxPooling2D)	(None, 8, 8, 512)	0
block5_conv1 (Conv2D)	(None, 8, 8, 512)	2359808
block5_conv2 (Conv2D)	(None, 8, 8, 512)	2359808
block5_conv3 (Conv2D)	(None, 8, 8, 512)	2359808
block5_pool (MaxPooling2D)	(None, 4, 4, 512)	0
global_average_pooling2d_2 (GlobalAveragePooling2D)	(None, 512)	0
dense_15 (Dense)	(None, 1024)	525312
dropout_6 (Dropout)	(None, 1024)	0
dense_16 (Dense)	(None, 1024)	1049600
dropout_7 (Dropout)	(None, 1024)	0
dense_17 (Dense)	(None, 512)	524800
dropout_8 (Dropout)	(None, 512)	0
dense_18 (Dense)	(None, 2)	1026

Total params: 16815426 (64.15 MB)
Trainable params: 2100738 (8.01 MB)
Non-trainable params: 14714688 (56.13 MB)

loss: 0.3284

accuracy: 0.8558

val_loss: 0.7721

val accuracy: 0.7259

EfficientNet-B7 and Vit

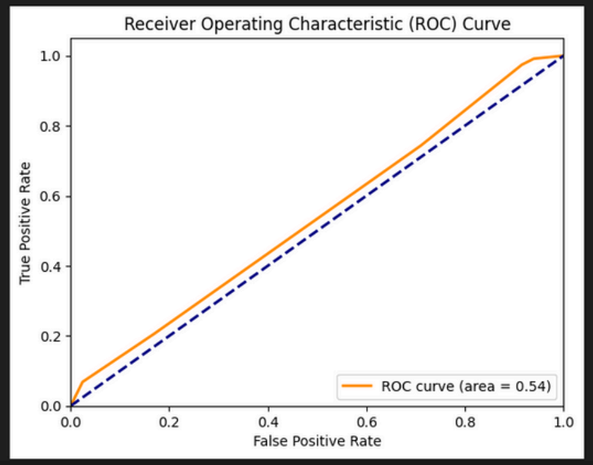
```
Epoch 38/40
54/54 [==============================] - 55s 1s/step - loss: 0.4286 - accuracy: 0.8469 - val_los
s: 0.4030 - val_accuracy: 0.8625
Epoch 39/40
54/54 [==============================] - 55s 1s/step - loss: 0.4285 - accuracy: 0.8469 - val_los
s: 0.4029 - val_accuracy: 0.8625
Epoch 40/40
54/54 [==============================] - 55s 1s/step - loss: 0.4285 - accuracy: 0.8469 - val_los
s: 0.4028 - val_accuracy: 0.8625
7/7 [==============================] - 13s 1s/step - loss: 0.4151 - accuracy: 0.8550
Test accuracy: 0.8550000190734863
```

```
7/7 [==============================] - 4s 77ms/step
Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        83
           1       0.58      1.00      0.74       117

    accuracy                           0.58       200
   macro avg       0.29      0.50      0.37       200
weighted avg       0.34      0.58      0.43       200

Sensitivity: 1.0
Specificity: 0.0
/opt/conda/lib/python3.10/site-packages/sklearn/metrics/_classification.py:1344: U
  _warn_prf(average, modifier, msg_start, len(result))
/opt/conda/lib/python3.10/site-packages/sklearn/metrics/_classification.py:1344: U
  _warn_prf(average, modifier, msg_start, len(result))
/opt/conda/lib/python3.10/site-packages/sklearn/metrics/_classification.py:1344: U
  _warn_prf(average, modifier, msg_start, len(result))
```



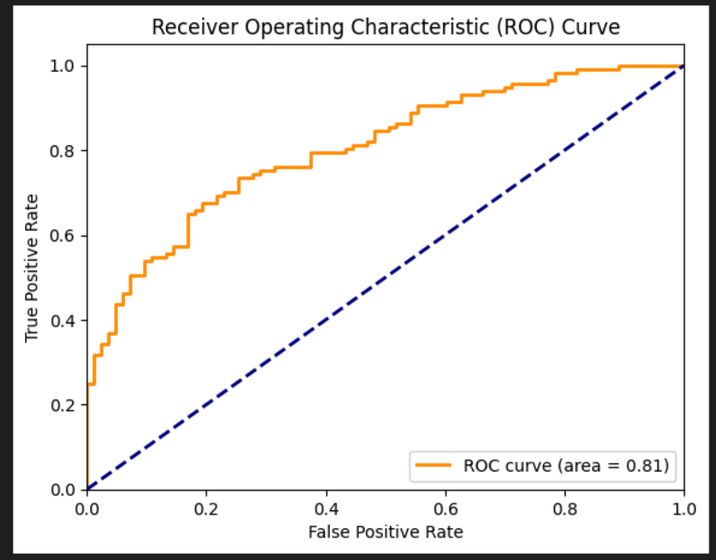
Vit-Vision transformer

```
7/7 [==============================] - 7s 194ms/step
Classification Report:
              precision    recall  f1-score   support

           0       0.58      0.86      0.69        83
           1       0.85      0.56      0.68       117

    accuracy                           0.69       200
   macro avg       0.71      0.71      0.68       200
weighted avg       0.74      0.69      0.68       200

Sensitivity: 0.5641025641025641
Specificity: 0.8554216867469879
```



EfficientNet-B7

COMPARITIVE STUDY (SRM DETECTOR VS HYPERVERGE)

Feature/Aspect	Hyperverge Nexus	Research Paper Model
Architecture	Modular, microservices-based, cloud-native	Hybrid, centralized framework, scalable
Models Used	ML models, CNNs, NLP	Ensemble learning, supervised & unsupervised models
Preprocessing	Data cleaning, feature extraction, normalization	Data augmentation, dimensionality reduction, synthetic data generation

KEY FINDINGS

- Project Success: Successfully implementing a deepfake detection project demonstrates the feasibility of such systems.
- GPU Limitations: The encountered GPU bottleneck suggests the need for resource optimization or exploring alternative hardware solutions for future projects.

THANK YOU