

# Presentation PFE

## Conception et Entraînement d' un Modèle de Détection des Deepfakes Basé sur les Réseaux de Neurones Convolutifs (CNNs)

présenté par: Ramla Babaha El Moctar Salem

Université de Nouakchott  
Faculté des Sciences et Techniques

Encadrée par : Dr. Elveth Sidi (Académique), Dr. Rifaa Sadegh (Professionnel)



# Contents

- 1 Introduction
- 2 Problématique
- 3 Outils Utilisés
- 4 Données
- 5 Modèle et entraînement
- 6 Évaluation
- 7 Application Web
- 8 Conclusion

# Introduction

# Introduction

- Falsification multimédia
- Deepfakes (images/vidéos truquées)
- Progrès de l'IA (CNN, GAN)
- Menaces : désinformation, vie privée, confiance numérique
- Besoin de détection automatisée

# Problématique

# Problématique et Objectifs

- Sophistication croissante des deepfakes
- Défi de fiabilité et généralisation
- Objectif 1 : Modèle CNN robuste
- Objectif 2 : Pipeline complet (prétraitement à prédiction)
- Objectif 3 : Intégration dans une application web

## Outils Utilisés

# Environnement et outils techniques

- Python, TensorFlow, Keras, OpenCV, Flask
- Colab, ImageDataGenerator, Checkpoints
- MTCNN, VSCode, Pandas, numpy



# Données

# Jeux de données utilisés



- Datasets : DFDC, Celeb-DF, FaceForensics++, etc.
- 100 000 images, 20 techniques de synthèse
- Données équilibrées (réel/faux)

# Prétraitement des images

- Détection de visages : MTCNN
- Redimensionnement : 224x224
- Normalisation + split train/val/test

# Modèle et entraînement

# Architecture du modèle CNN

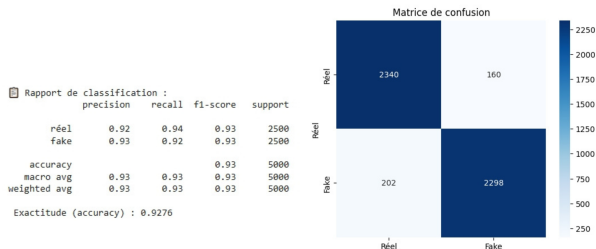
- EfficientNetB0 + couche dense
- Optimiseur Adam, loss binary crossentropy
- Régularisation : Dropout

# Pipeline d'entraînement

- Google Colab avec GPU
- Augmentation des données
- Checkpoints + EarlyStopping
- Enregistrement des résultats en JSON

# Évaluation

# Méthodologie et métriques



- Accuracy, Précision, Rappel, F1-score
- Matrice de confusion

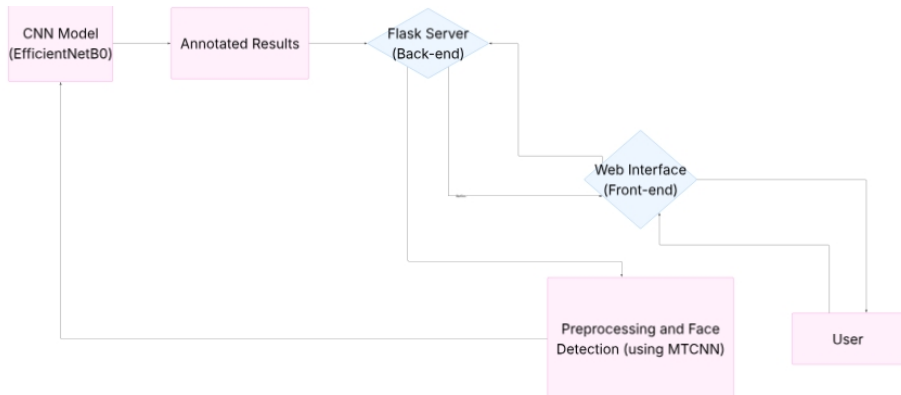


# Résultats et interprétation

- Accuracy obtenue : 93%
- Faibles taux de faux positifs/négatifs
- Limites : images floues, expressions extrêmes

# Application Web

# Architecture de l' application web



- Flask backend + CNN
- Interface HTML/CSS pour upload + prédiction
- Affichage des visages avec couleur selon verdict

## Conclusion

# Conclusion générale

- Pipeline complet : collecte → entraînement → interface
- Compétences acquises en IA, vision, déploiement web
- Limites rencontrées : variabilité, contraintes techniques
- Perspectives : détection audio/vidéo, GAN adversarial training

# Demonstration !

# Merci pour votre attention !

Des questions ?