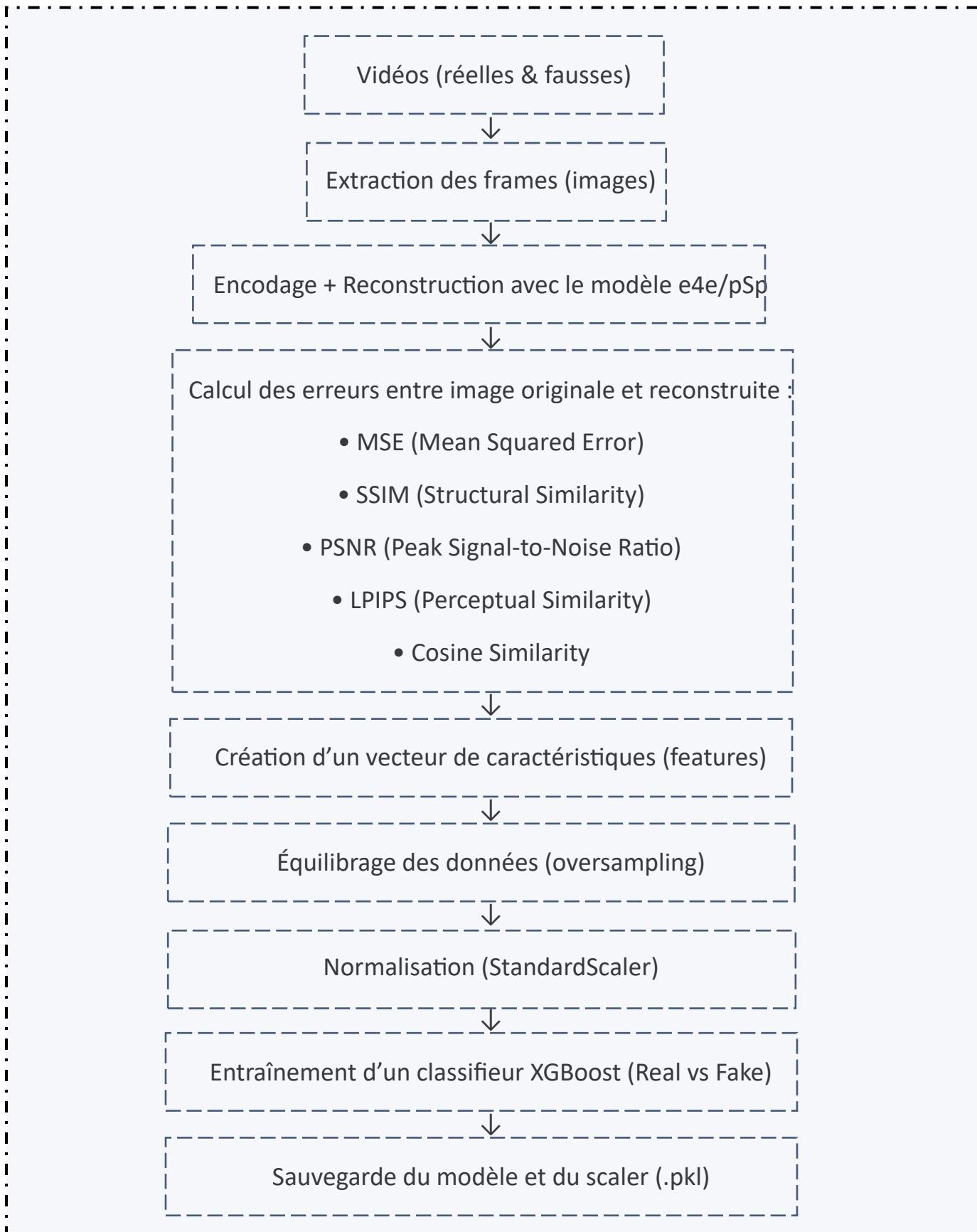


Détection de Deepfake via GAN Inversion(e4e/pSp)



Description du dataset :Celeb-DF v2

Nom du dataset : Celeb-DF (v2)

Source officielle : [GitHub – Celeb-DeepFakeForensics](#)

Type : Dataset vidéo pour la détection de deepfakes

Année de publication : 2019

• Construction d'un deepfake

Vidéo réelle cible : id0_0000 → contient le visage de la célébrité “cible”.

Vidéo donneur : id1_0000 → contient le visage de la personne source.

Résultat : la vidéo deepfake ressemble à la vidéo réelle id0_0000, mais avec le visage de id1_0000.

1.Extraction des frames

a. Chargement des images

Les frames obtenues à partir de vidéos authentiques et Deepfake sont tirées de dossiers dédiés.

Formats pris en charge : .jpg, .jpeg, .png.

2.Prétraitement

Redimensionnement des images à 256x256 pixels.

Transformation en tenseurs normalisés de PyTorch.

3.GAN Inversion (e4e / pSp)

Chaque image est traitée par le modèle e4e/pSp afin de réaliser une reconstruction.

Les reconstructions permettent de capturer les subtilités et anomalies propres aux Deepfakes.

4.Extraction de features

Pour chaque image, un vecteur de cinq caractéristiques est déterminé :

- ✓ MSE : erreur quadratique moyenne entre image originale et reconstruite.
- ✓ 1-SSIM : mesure structurelle de similarité inversée.

- ✓ PSNR : rapport signal/bruit basé sur l'erreur.
- ✓ LPIPS : distance perceptuelle basée sur un réseau pré-entraîné AlexNet.
- ✓ 1-Cosine Similarity : similarité cosinus inversée entre l'image originale et reconstruite.

5. Préparation du dataset

Les vecteurs de features sont combinés pour former la matrice X.

Les labels associés (0 pour réel, 1 pour fake) forment le vecteur y.

Un RandomOverSampler est utilisé pour équilibrer le dataset.

6. Entraînement du classifieur XGBoost

Hyperparamètres optimisés : 1200 arbres, profondeur maximale = 8, learning rate = 0.03.

StandardScaler appliqué pour normaliser les features.

Split train/test : 80/20, stratifié sur les classes.

7. Évaluation

Total samples: 4765 (Real: 790, Fake: 3975)

Accuracy train : 96,05 %

Accuracy test : 88,05 %

	precision	recall	f1-score	support
Real	0.85	0.93	0.89	795
Fake	0.92	0.83	0.87	795
accuracy			0.88	1590
macro avg	0.88	0.88	0.88	1590
weighted avg	0.88	0.88	0.88	1590

8. Confusion Matrix

