

Compte Rendu 2

Projet Face-Swap M2
T.Bonetti A.Conde-Salazar

1 Objectif Global

Construire une représentation compacte (*embedding*) des visages permettant :

- Séparation des identités.
- Classification supervisée ultérieure.
- Reconstruction/génération indicative pour analyse qualitative.
- Tester nos méthodes sur les visages des élèves de la promo (dataset personnalisé)

2 Choix Structurants

- Taille d'image : 128x128
- Embedding : dimension 128..
- Séparation pipeline : apprentissage métrique puis classification.

3 Apprentissage Metric (Triplet)

Réseau d'Embedding

Deux blocs convolutionnels simples + BatchNorm + ReLU + MaxPool puis couche linéaire.
Simplicité voulue pour tester rapidement :

- Convolutions larges (kernel 5) pour capter traits faciaux globaux.
- BatchNorm pour stabilité et éviter collapse.
- Normalisation finale.

Optimisation

Adam (lr = 1e-4) sans scheduler pour analyse initiale. Possibles améliorations :

- Scheduler (Cosine, ReduceLROnPlateau).
- Data augmentation (flip léger, color jitter).

4 Analyse Embeddings

- Centroïdes calculés (moyenne par classe) pour évaluation *nearest centroid*.
- Visualisation 2D : PCA (alternative t-SNE / UMAP pour plus tard).

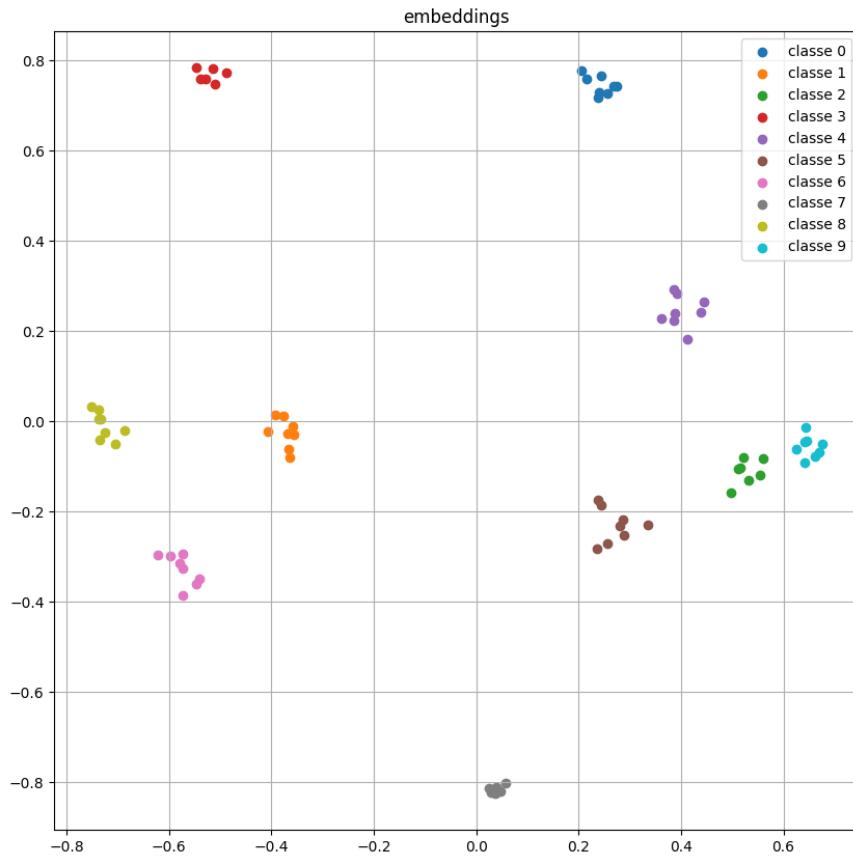


FIGURE 1 – Projection PCA des embeddings.

5 Classification Supervisée

- Classifieur linéaire sur embeddings
- Perte : CrossEntropyLoss (standard multi-classe).
- Adam pour convergence rapide.

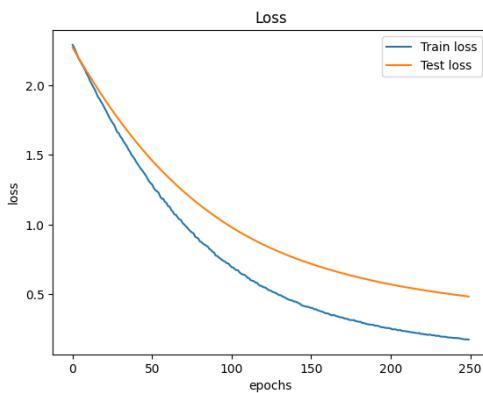


FIGURE 2 – Courbes Loss train / test .

6 Évaluation

- Accuracy globale.
- Matrice de confusion pour classes difficiles.
- Inspection visuelle échantillons : prédiction vs vérité (couleur code).

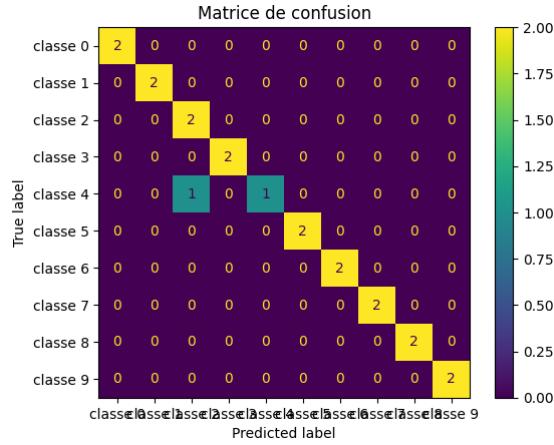


FIGURE 3 – Matrice de confusion (placeholder).



FIGURE 4 – Exemples d’images test avec prédictions avec notre dataset personnalisé.

7 Reconstruction / Décodeur

- Décodeur convolutionnel transposé (architecture ascendante progressive).
- Objectif : MSE entre image d'entrée et reconstruction (pas de terme KL : ce n'est pas un vrai VAE).
- Usage : vérifier quelles informations l'embedding conserve (texture, forme visage).
- Génération à partir des centroïdes pour images moyennes par identité.
- A ce stade cette partie est grandement améliorables comme vous pouvez le voir :

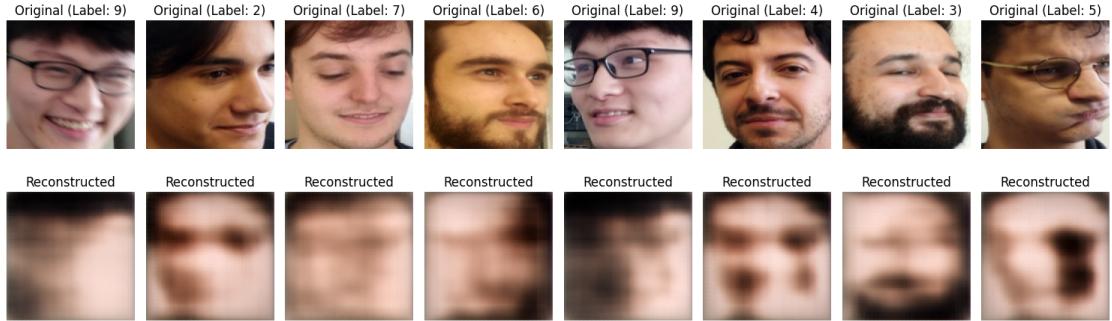


FIGURE 5 – Originaux vs reconstructions.

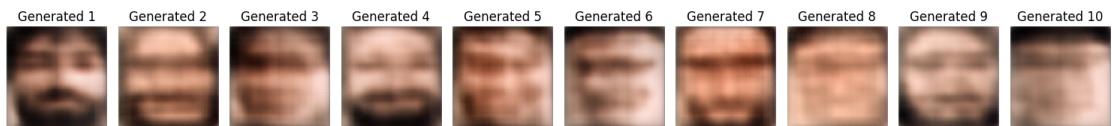


FIGURE 6 – Images générées depuis centroïdes.

8 La suite ?

- Décodeur : ajouter régularisation (Dropout / Weight decay).
- Embedding : essayer pré-entraînement (FaceNet / ResNet) pour base plus riche.
- Reconstruction : passer à VAE réel (terme KL) pour interpolation contrôlée.
- Évaluation : ajouter F1 par classe et distances intra/inter-classe.