

Problématique

La reconnaissance automatique des expressions faciales (FER) est un domaine clé pour les interactions homme-machine [1].

Cependant, elle reste un défi en raison :

- ✓ de la variabilité des expressions faciales,
- ✓ des variations d'éclairage et de pose
- ✓ du manque de données annotées.

Objectifs

L'objectif principal est de concevoir un système efficace de reconnaissance des émotions faciales basé sur des caractéristiques profondes.

Objectifs spécifiques :

- ✓ Exploiter **ConvNeXt pré-entraîné** avec fine-tuning partiel pour extraire des caractéristiques robustes.
- ✓ Réduire la dimension des caractéristiques via **PCA optimisé par base de données**.
- ✓ Utiliser un **SVM linéaire** pour une classification efficace en faible quantité de données.
- ✓ Garantir des résultats **robustes et reproductibles** sur plusieurs bases.

Base de données

Notre système utilise les bases suivantes :

- **RaFD** : 536 images, 67 sujets, 8 émotions
- **JAFFE** : 207 images, 10 sujets, 6 émotions + neutre
- **KDEF** : 980 images, 70 sujets, 6 émotions + neutre
- **WSEFEP** : 210 images, 7 émotions

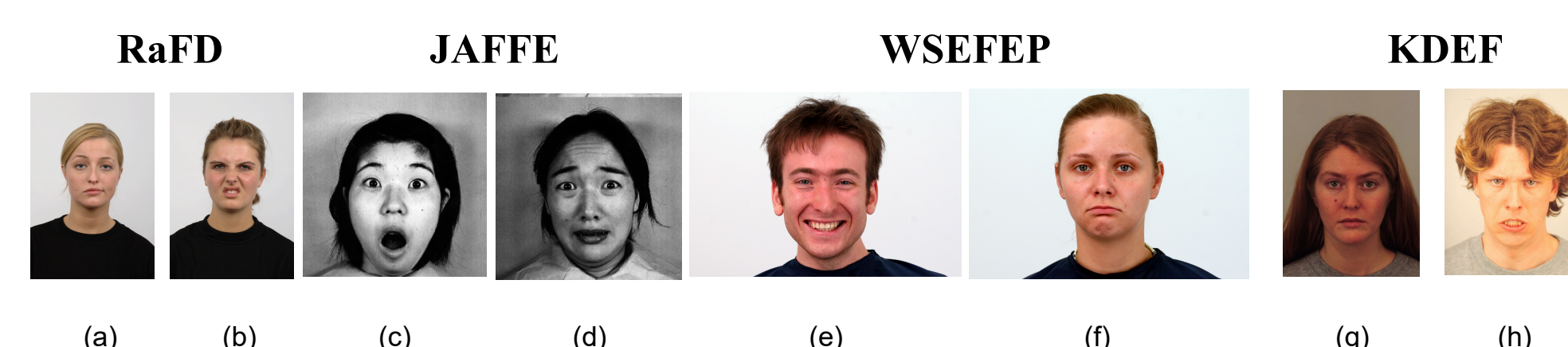


FIGURE 1 – Images représentatives des 8 classes d'émotions (a) Mépris, (b) Dégout, (c) Surprise, (d) Peur, (e) Heureux, (f) Triste, (g) Neutre, (h) Colère.

Méthodologie suivie

La méthode proposée repose sur un pipeline basé sur l'extraction de caractéristiques profondes, suivie d'une réduction de dimensionnalité et d'une classification supervisée, comme illustré à la Figure 2.

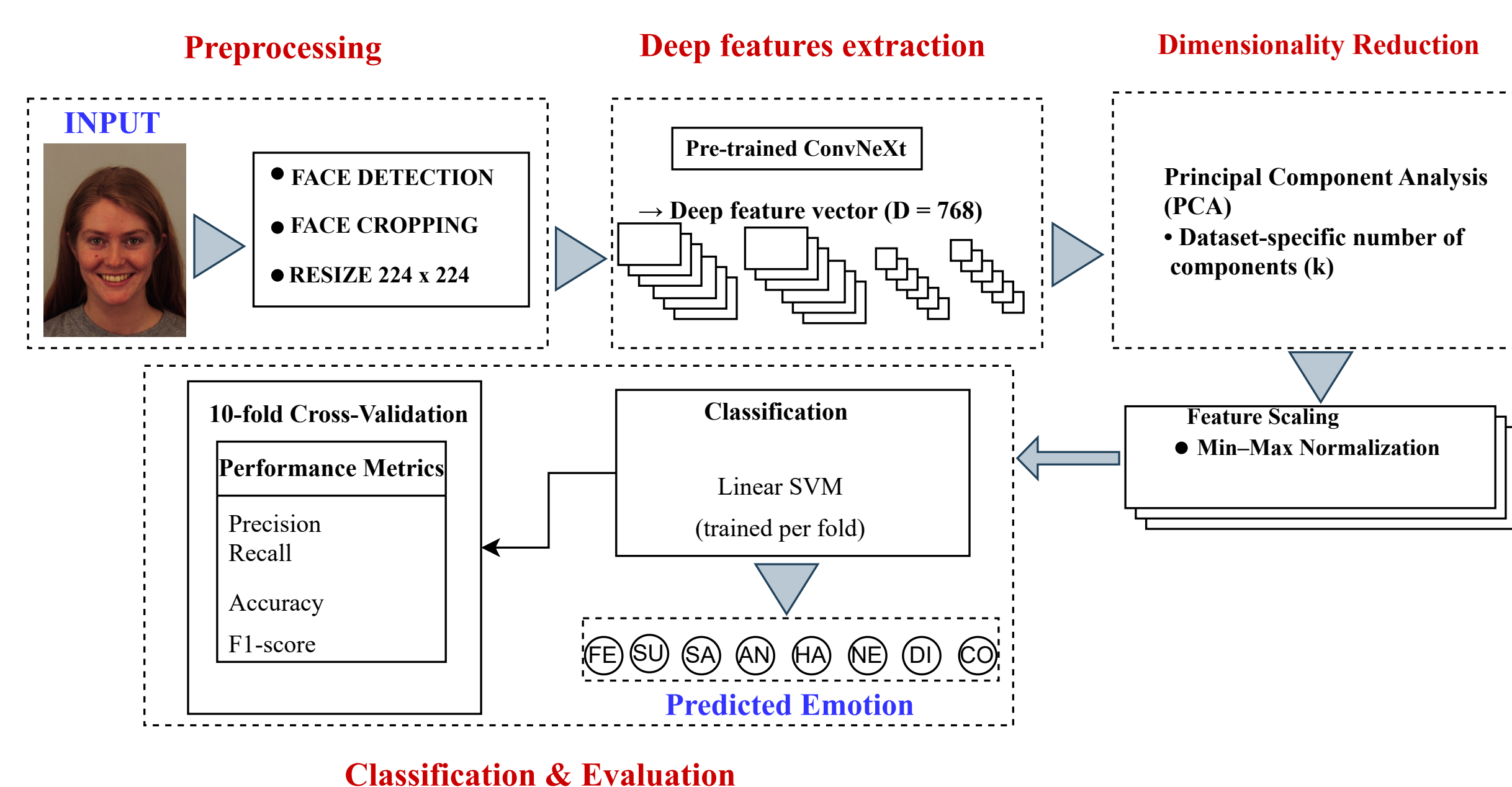


FIGURE 2 – Pipeline ConvNeXt + PCA + SVM.

1. Prétraitement

- ✓ Détection avec **MediaPipe**
- ✓ Redimensionnement 224×224
- ✓ Normalisation **ImageNet**

2. Extraction (ConvNeXt)

- ✓ **ConvNeXt-Tiny pré-entraîné**
- ✓ Fine-tuning partiel (100 époques, stage 2+)
- ✓ Extraction des caractéristiques puis moyenne globale
- ✓ Obtention d'un vecteur profond ($D = 768$)

3. Réduction (PCA)

- ✓ PCA
- ✓ Composantes déterminées empiriquement par base de données
- ✓ Évaluation via validation croisée à 10 plis

4. Classification

- ✓ Normalisation **Min-Max**
- ✓ **SVM linéaire** pour la classification
- ✓ Prise en compte du déséquilibre des classes

Évaluation

Une validation croisée stratifiée à 10 plis est utilisée. **Métriques utilisées** : Exactitude, Précision, Rappel, F1-score (macro). Les résultats correspondent à la moyenne sur les 10 plis.

Résultats

Les résultats montrent des performances élevées et stables malgré des bases de taille limitée.

Résultats moyens (%) sur les bases de données.

Dataset	Acc (%)	Prec	Rec	F1	PCA
KDEF	93.06	0.93	0.93	0.93	256
JAFFE	93.14	0.93	0.93	0.93	205
RaFD	96.82	0.96	0.96	0.96	483
WSEFEP	89.05	0.89	0.89	0.89	205

Matrice de confusion (%) (vrai vs prédit).

Base	AN	DI	FE	HA	NE	SA	SU
JAFFE	93.3	89.3	90.0	93.1	93.3	96.7	96.7
KDEF	93.6	92.1	85.0	100	87.9	95.0	97.9
WSEFEP	90.0	83.3	90.0	100	76.7	93.3	86.7
RaFD	95.5	98.5	97.0	100	92.5	98.5	97.0

- ✓ Meilleur résultat : **RaFD (96.82%)**
- ✓ Pipeline stable sur toutes les bases
- ✓ PCA adapté améliore la généralisation

Conclusion

- ✓ ConvNeXt + PCA + SVM offre de très bonnes performances
- ✓ Bonne généralisation malgré peu de données
- ✓ PCA améliore la robustesse
- ✓ Perspectives : intégration de caractéristiques géométriques

Références

- [1] Sanoar Hossain, Saiyed Umer, Vijayan Asari, and Ranjeet Kumar Rout. A unified framework of deep learning-based facial expression recognition system for diversified applications. *Applied Sciences*, 11(19):9174, 2021.
- [2] Yacine Yaddaden. Efficient dynamic emotion recognition from facial expressions using statistical spatio-temporal geometric features. *Bia Data and Coanitive Computina*. 9(8):213, 2025.