

Rapport d'Analyse et de Modélisation : Reconnaissance d'Émotions Faciales à partir du Dataset FER2013

Auteur : [Zahra Boucheta , Roubache Islam]

May 3, 2025

1 Objectifs du Projet

L'objectif principal de ce projet est de réaliser une **analyse exploratoire approfondie** du dataset **FER2013**, suivie de la **préparation des données** et de la **conception d'un modèle hybride CNN-LSTM** capable de reconnaître les émotions humaines à partir d'images de visages en niveaux de gris.

2 Inspection du Dataset FER2013

Le dataset est divisé en deux sous-ensembles :

- ../input/fer2013/train
- ../input/fer2013/test

Les 7 classes émotionnelles sont :

['angry', 'disgust', 'fear', 'happy', 'neutral', 'sad', 'surprise']

Résumé Statistique - Entraînement

Émotion	Nombre d'images	Distribution (%)
angry	3,995	13.92
disgust	436	1.52
fear	4,097	14.27
happy	7,215	25.13
neutral	4,965	17.29
sad	4,830	16.82
surprise	3,171	11.05
Total	28,709	

Table 1: Répartition des émotions - Entraînement

Classe majoritaire : *happy*

Classe minoritaire : *disgust*

Résumé Statistique - Test

Émotion	Nombre d'images	Distribution (%)
angry	958	13.35
disgust	111	1.55
fear	1,024	14.27
happy	1,774	24.71
neutral	1,233	17.18
sad	1,247	17.37
surprise	831	11.58
Total	7,178	

Table 2: Répartition des émotions - Test

Classe majoritaire : *happy*Classe minoritaire : *disgust*

3 Prétraitement et Génération des Données

Utilisation de `ImageDataGenerator` avec les paramètres suivants :

- Normalisation (rescale) : $[0, 1]$
- Zoom : jusqu'à 30%
- Flip horizontal : aléatoire
- Mode : `grayscale`
- Taille cible : 48×48 pixels
- Batch Size : 64

4 Modèle Convolutionnel-LSTM

Un modèle hybride a été conçu pour capter à la fois les **caractéristiques spatiales** (CNN) et **temporelles/contextuelles** (LSTM).

Architecture du Modèle

Partie CNN :

- Conv2D : 32 + 64 filtres (3×3), ReLU
- BatchNormalization
- MaxPooling2D
- Dropout 0.25
- Conv2D : 128 + 256 filtres
- BatchNormalization
- MaxPooling2D

- Dropout 0.25

Reshape : Conversion en séquences temporelles pour LSTM

Partie LSTM :

- LSTM 128 (`return_sequences=True`)
- Dropout 0.5
- LSTM 64 ou 128
- Dense Softmax (7 classes)

5 Visualisation des Données

Distribution Graphique (Noir et Blanc)

- Barplots avec nuances de gris
- Grille visible pour la lisibilité
- Sauvegarde en PNG haute résolution (300 DPI)

Échantillons Visuels

- Affichage de 9 visages représentatifs
- Titres personnalisés
- Fond sombre, police blanche
- Style épuré

6 Analyse des Courbes d'Apprentissage

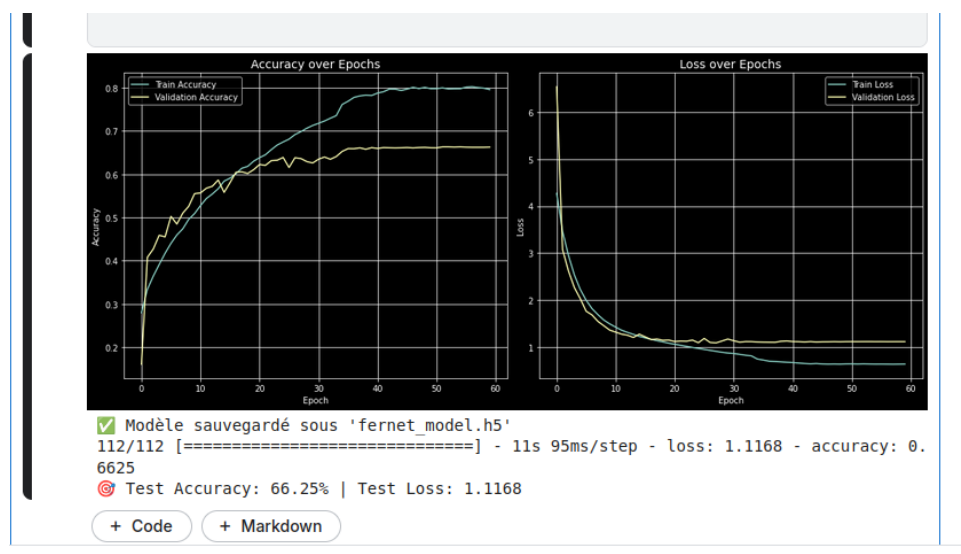


Figure 1: Courbes d'apprentissage du modèle CNN-LSTM sur 60 epochs

Les courbes d'apprentissage montrent l'évolution de l'accuracy et de la loss sur 60 epochs. La précision d'entraînement (ligne bleue) démarre autour de 0,3 et progresse régulièrement jusqu'à

atteindre un plateau vers 0,8 après 40 epochs, tandis que l'accuracy de validation (ligne orange) suit une trajectoire similaire mais se stabilise légèrement en dessous autour de 0,78.

Les courbes de loss présentent une décroissance rapide durant les 15 premières epochs, avec la training loss (bleue) convergeant vers 0,85 et la validation loss (orange) se maintenant autour de 1,12. L'écart constant mais modéré entre les courbes train/val suggère que le modèle apprend efficacement sans surapprentissage significatif. La convergence progressive des deux métriques indique que l'architecture CNN-LSTM et les paramètres d'optimisation sont bien adaptés à la tâche.

7 Résultats et Sauvegardes

- Rapport Excel généré : `emotion_analysis_report.xlsx`
- Graphiques enregistrés :
 - `distribution_entrainement_noir_blanc.png`
 - `distribution_test_noir_blanc.png`
 - `learning_curves.png`

8 Prochaines Étapes

- **Fine-tuning** du modèle avec `EarlyStopping`
- Évaluation détaillée (courbe d'apprentissage, matrice de confusion)
- Optimisation via modèles préentraînés ou Vision Transformers

Code Source

Le code source complet de ce projet est disponible sur GitHub :
https://github.com/ZahraBcht/DL_finalproject.git

9 Conclusion

Ce pipeline démontre une approche rigoureuse et reproductible pour l'analyse des émotions faciales à partir d'images. Il combine une inspection approfondie des données, une visualisation soignée, et un modèle avancé CNN-LSTM, posant ainsi les bases pour des extensions futures plus robustes et précises. Les résultats des courbes d'apprentissage confirment la viabilité de l'approche hybride pour cette tâche de classification.