Ayouba Anrezki

14408

Analyse des invariants de dissimulation

détection de la stéganographie dans les images JPG

Transition, transformation, conversion

Introduction

Stéganographie

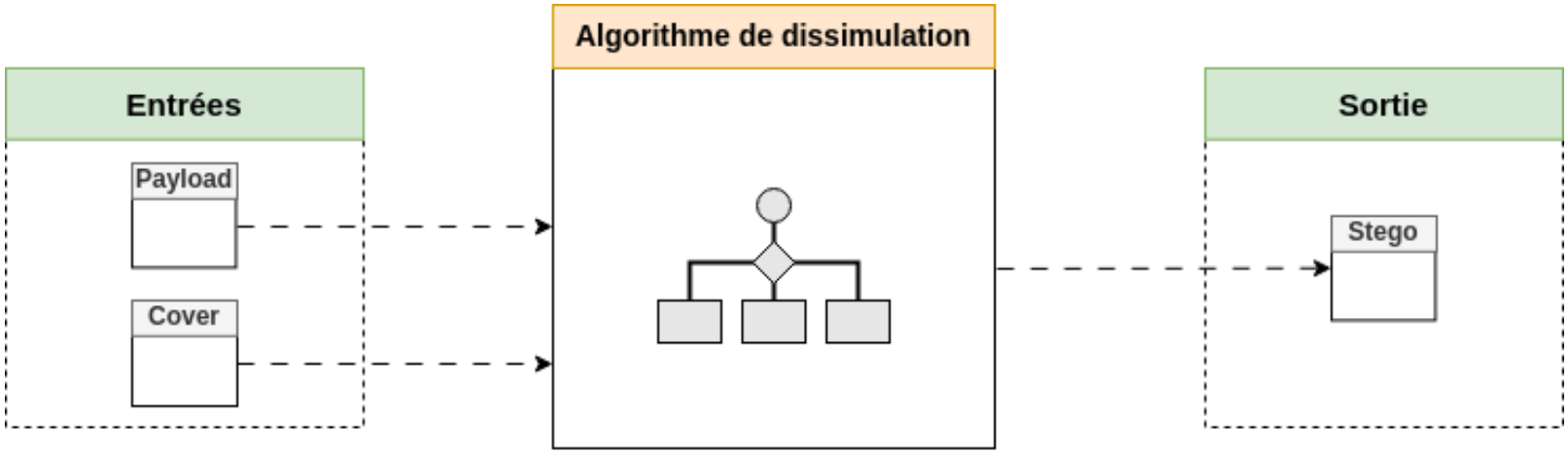


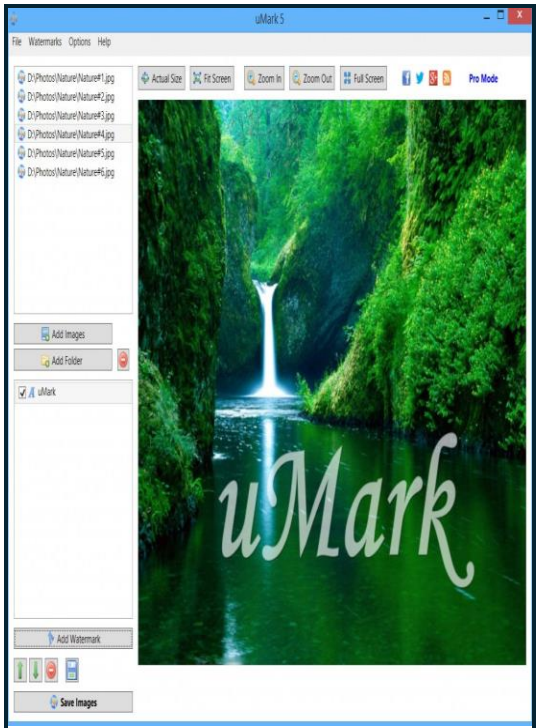
Figure 1 :
Illustration de la stéganographie

Introduction

Utilisation de la stéganographie

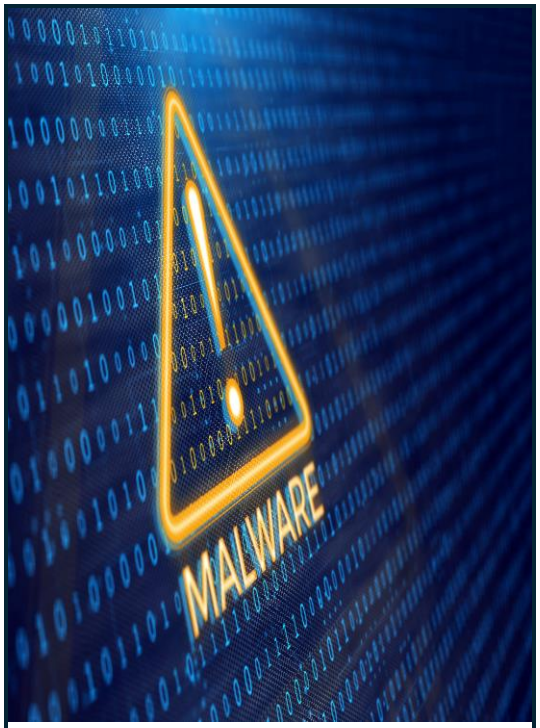
ThePhoto par PhotoAuthor est fournie sous licence CCYYSA.

Communication secrète



ThePhoto par PhotoAuthor est fournie sous licence CCYYSA.

Marquage de droits d'auteur



ThePhoto par PhotoAuthor est fournie sous licence CCYYSA.

Cacher des logiciels malveillants

Figure 2 :
Utilisation de la stéganographie

Introduction

Stéganalyse

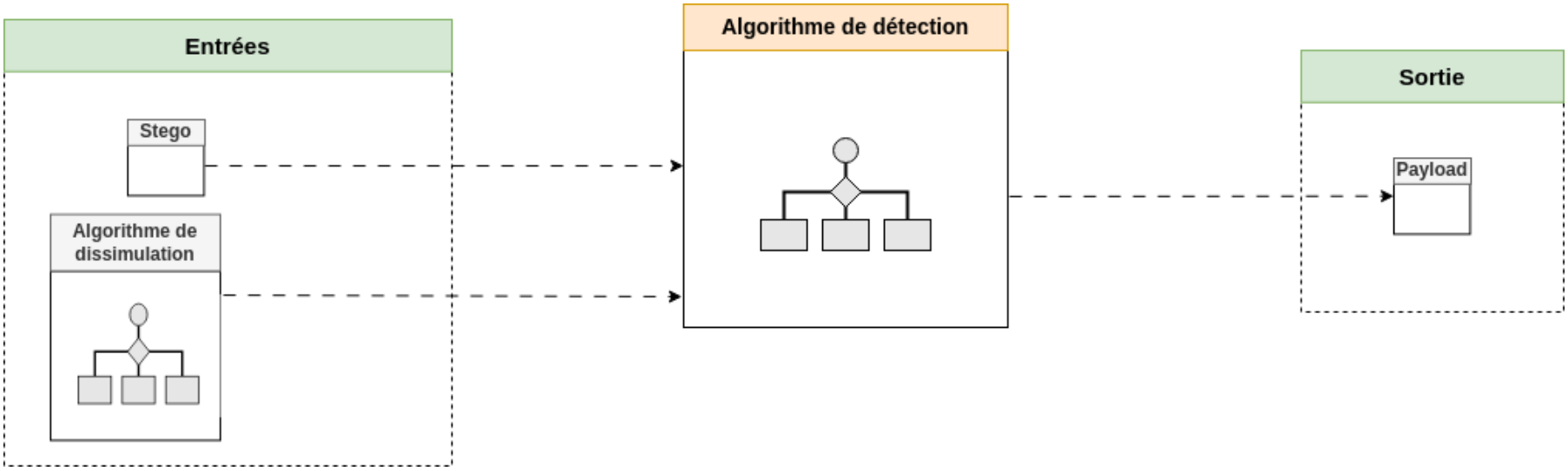


Figure 3 :
Illustration de la stéganalyse

Problématique

Est-il possible d'identifier un invariant de dissimulation dans des fichiers JPG, c'est-à-dire une caractéristique commune à toutes les données issues d'un processus de stéganographie, indépendamment de l'algorithme utilisé ou du type de données, permettant ainsi de détecter la présence d'information cachée dans ces fichiers ?

Problématique

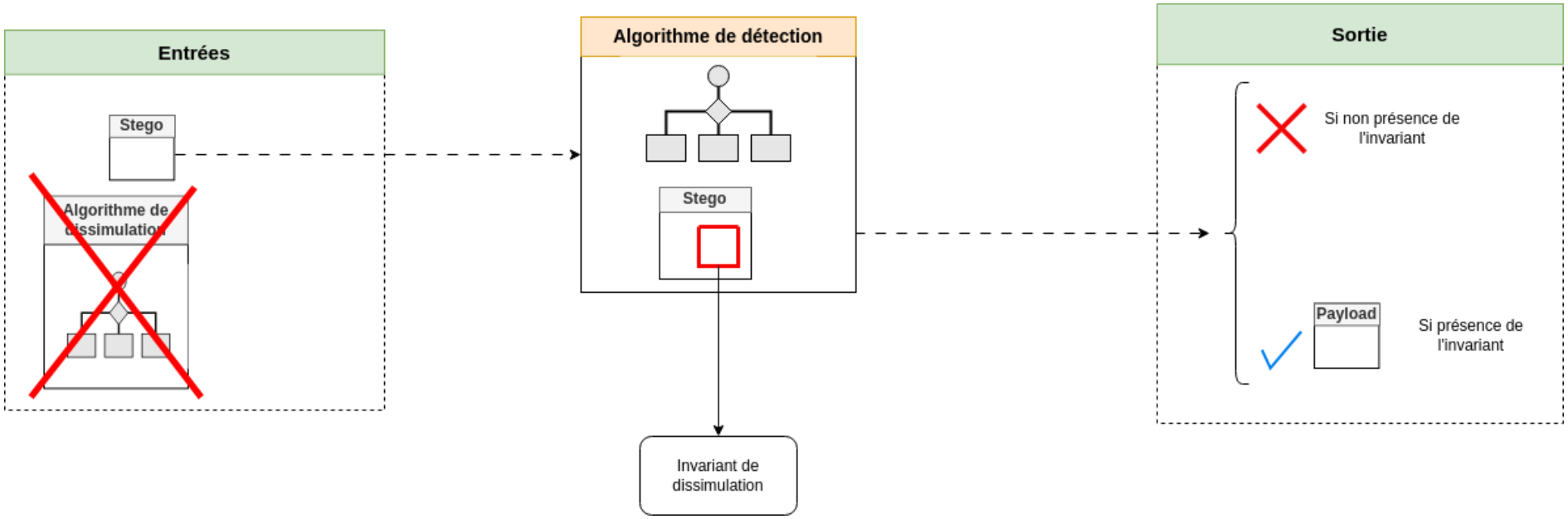


Figure 4 :
Illustration de la problématique

Objectifs du TIPE

1. Choisir trois méthodes de dissimulations en stéganographie sur les images JPG.
2. Créer une base de données d'image JPG (payload, cover).
3. Réalisé l'opération de dissimulation sur les données avec les trois méthodes et récupérer la sortie (stego).
4. Faire des mesure statistique sur le jeu de données obtenu.
5. Étudier le résultat de la mesure pour identifier un invariant de dissimulation.
6. Implémenter une méthode détection.
7. Évaluer la méthode.

Plan de la présentation

I. Mise en place
   A. Le format JPG
   B. Les méthodes de dissimulation (LSB,PVD, F5)
   C. Dissimulation
   D. Base de données.
II. Études statistique
   A. Série de mesures sur le jeu de données
   B. Mise en évidence de l'invariant
   C. Étude de l'invariant
III. Implémentation
   A. Algorithmes
   B. Evaluation
IV. Conclusion

Mise en place

Le format JPG : qu'est-ce que le format JPG ?

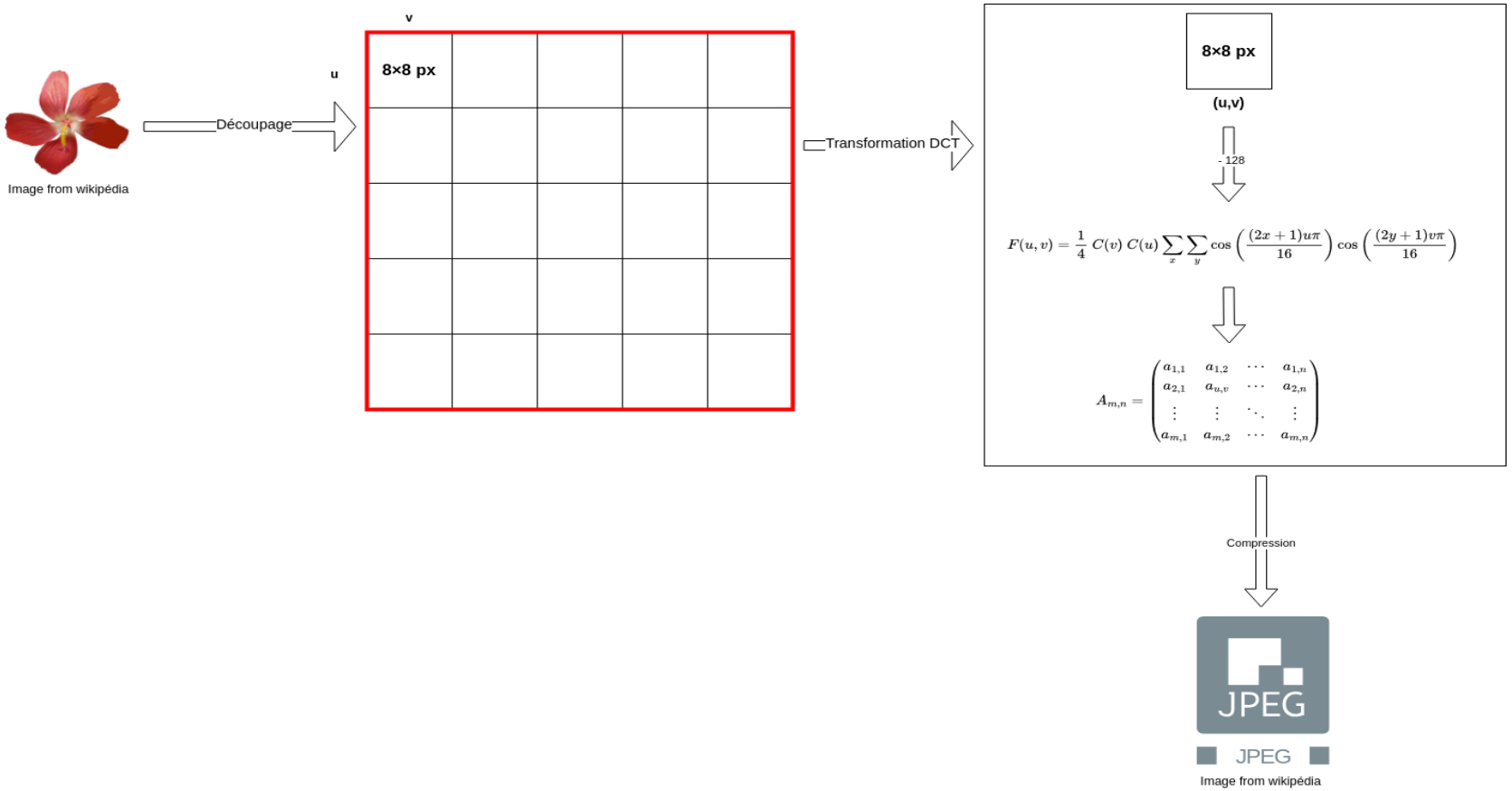


Figure 5 :
Format JPG et compression JPEG

Mise en place

Le format JPG : Avantages pour la stéganographie

Critère	JPG	PNG / BMP
Compression	Avec perte (DCT + quantification)	Sans perte (pixel par pixel)
Facilité d'intégration	Modifications dans les coefficients DCT	Modifications visibles dans les pixels
Taille du fichier	Petite	Grande
Invisibility (imperceptibilité)	Haute (modifs en fréquence)	Moins bonne (modifs visibles)
Robustesse	Moyenne à élevée (résiste à recompression)	Faible (modifications fragiles)
Diffusion / Usage	Très élevée	Moins courante

Figure 6 :
Avantages du JPEG pour la stéganographie

Mise en place

Méthodes de dissimulations

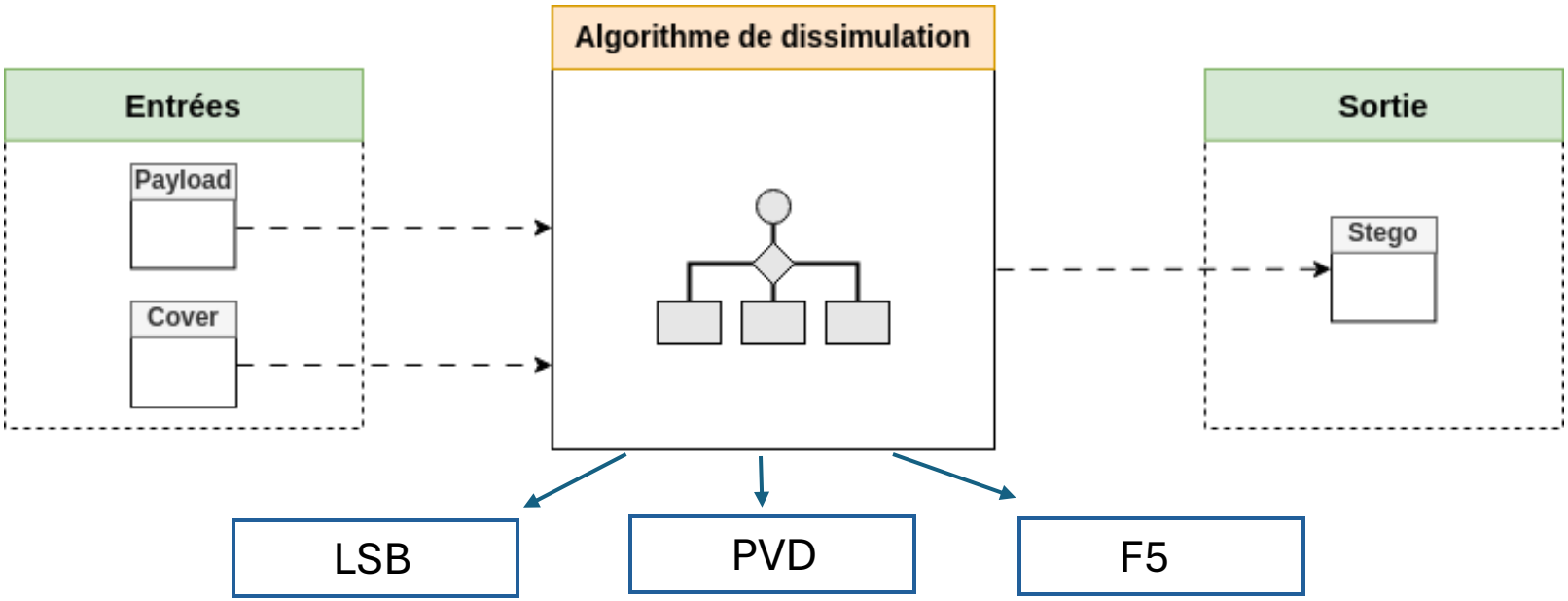


Figure 7 :
Illustration de la stéganographie

Mise en place

Méthodes de dissimulations : LSB

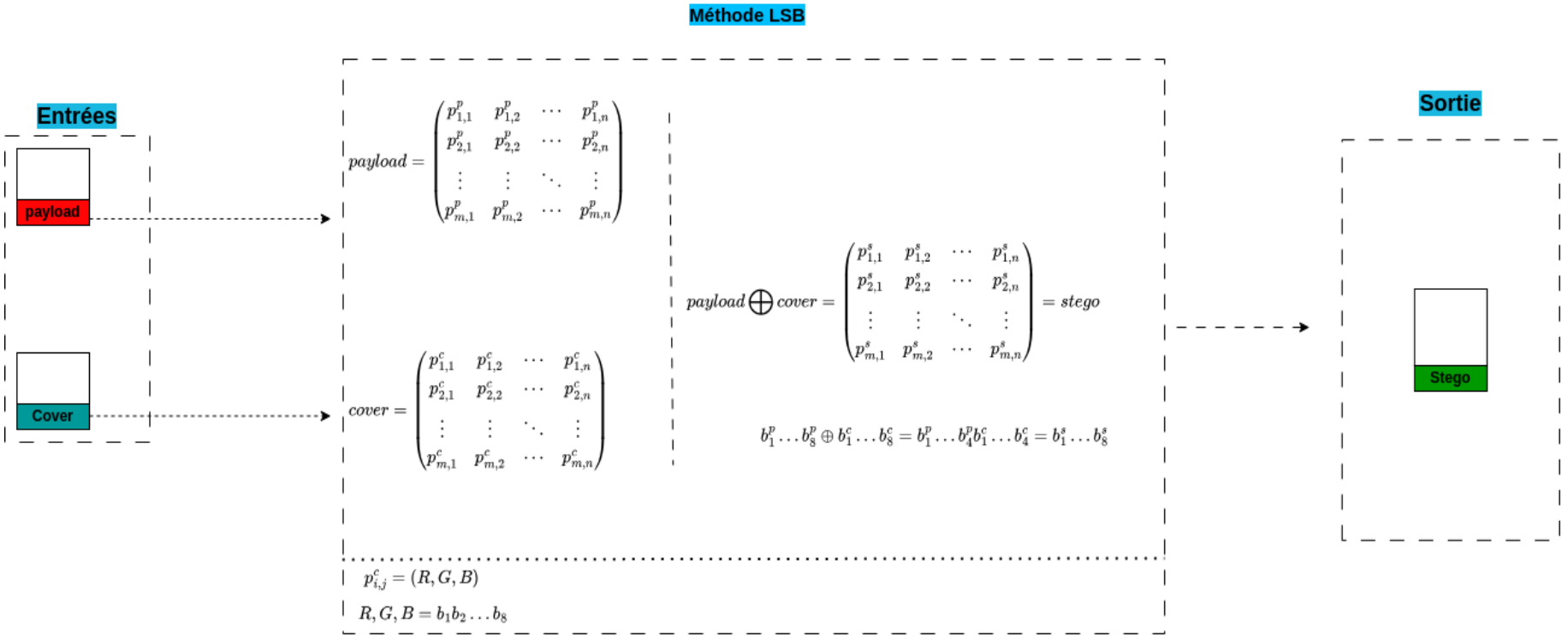


Figure 8 :
Illustration de la méthode LSB (least significant bit)

Mise en place

Méthodes de dissimulations : PVD

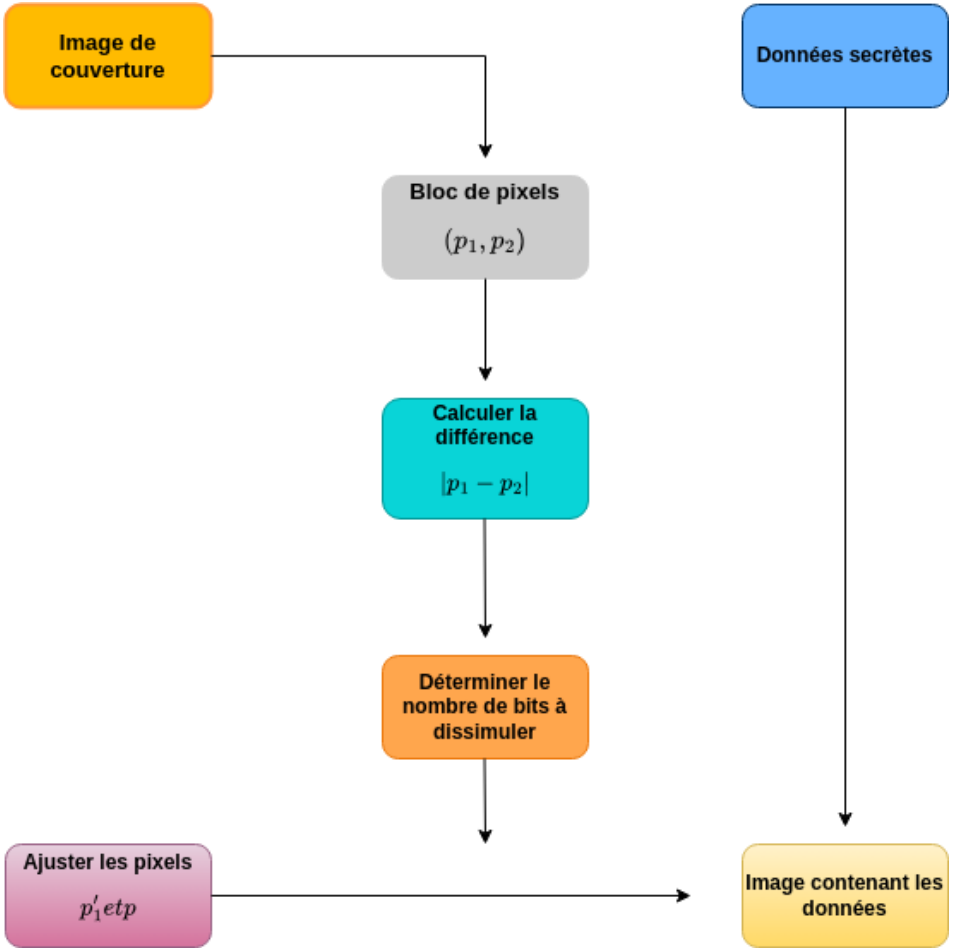


Figure 9 :
Illustration de la méthode PVD

Mise en place

Méthodes de dissimulations : F5

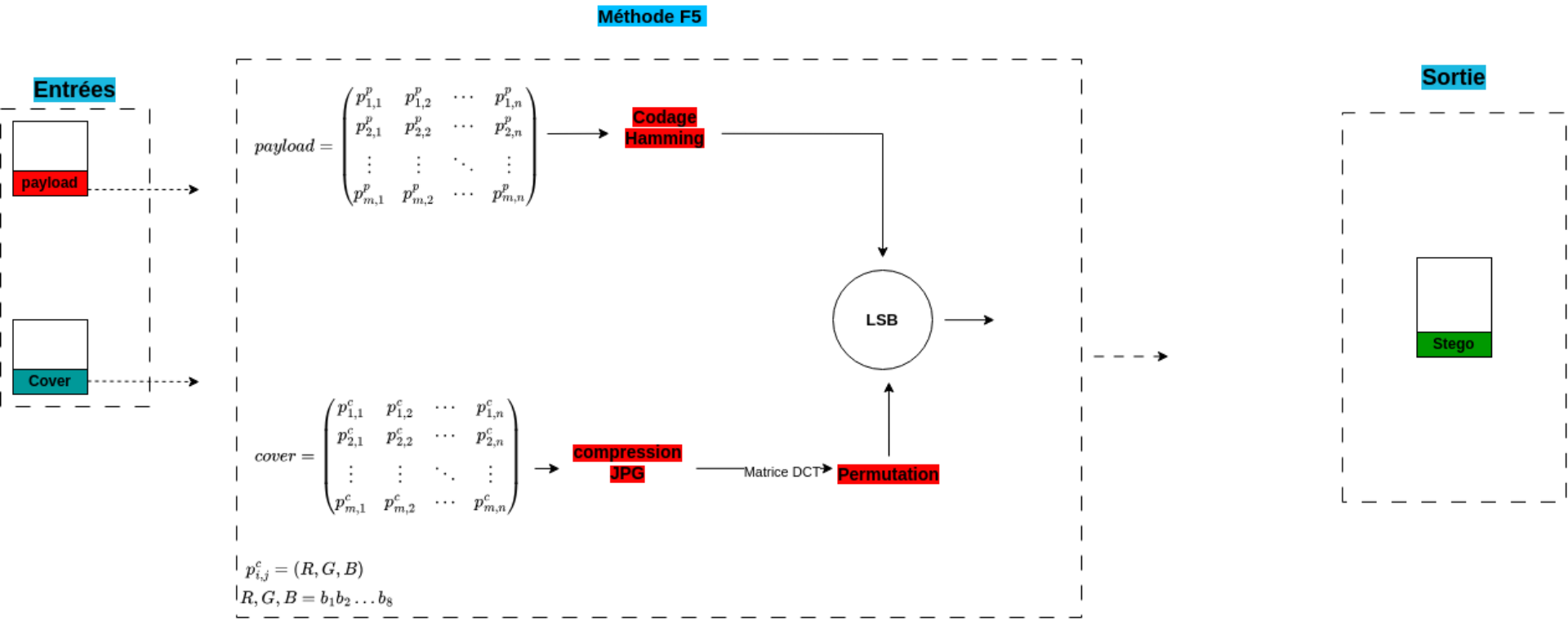


Figure 10 :
Illustration de la méthode F5

Dissimulation

Critère de dissimulation

Critères	Mesures	spécifications
Invisibilité / imperceptibilité	• PSNR (Peak Signal-to-Noise Ratio) • SSIM (Structural Similarity Index)	• Visibilité a l'œil humain • Plus la différence est faible, meilleure est la qualité
Capacité d'insertion	Bits par pixel (bpp)	C = nombre de pixels × bits par pixel T = S/C × 100
Résistance à la stéganalyse	Taux de détection (TP/FP)	• Analyse statistique

Figure 11 :
Critère de dissimulation

Dissimulation

Mesure sur le premier dataset

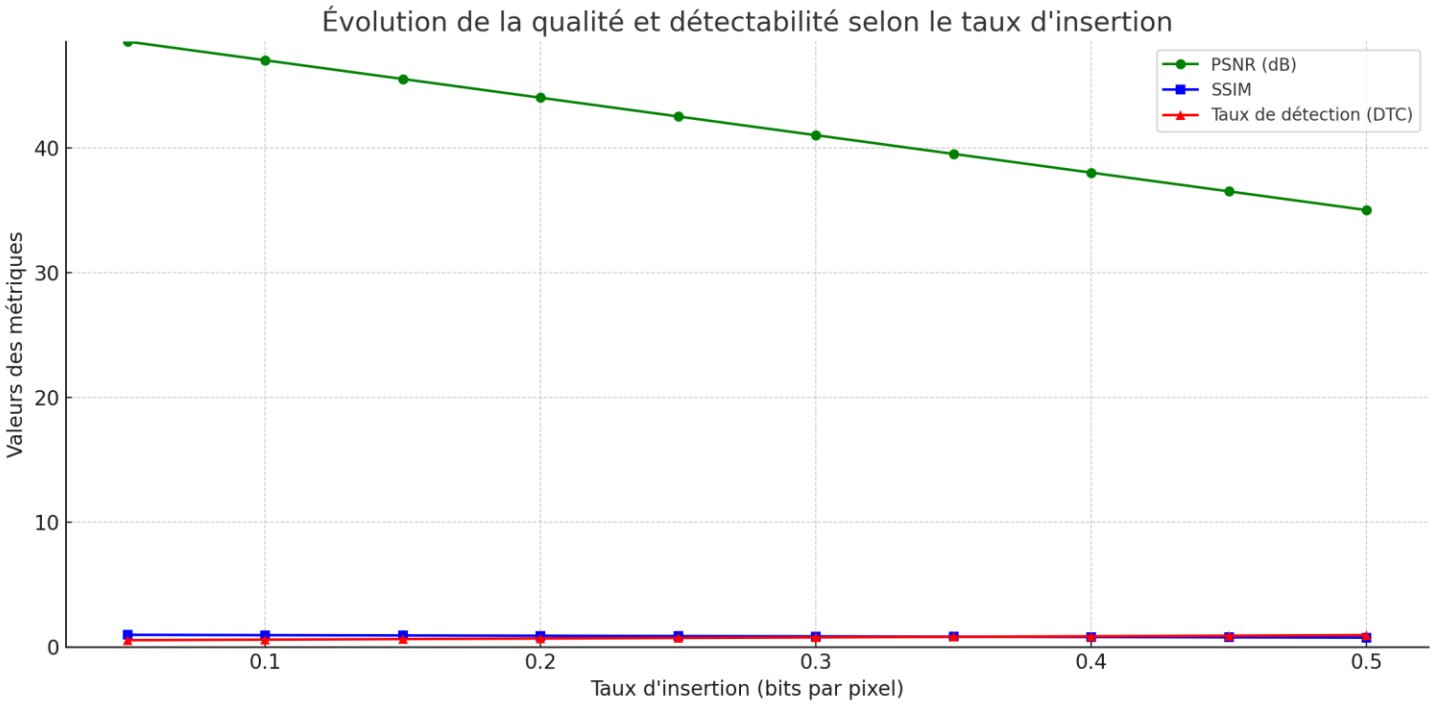


Figure 12 :

Graphe capacité de dissimulation (taille du jeu de données = 10 000 images)

Base de données

Organisation de la base de données

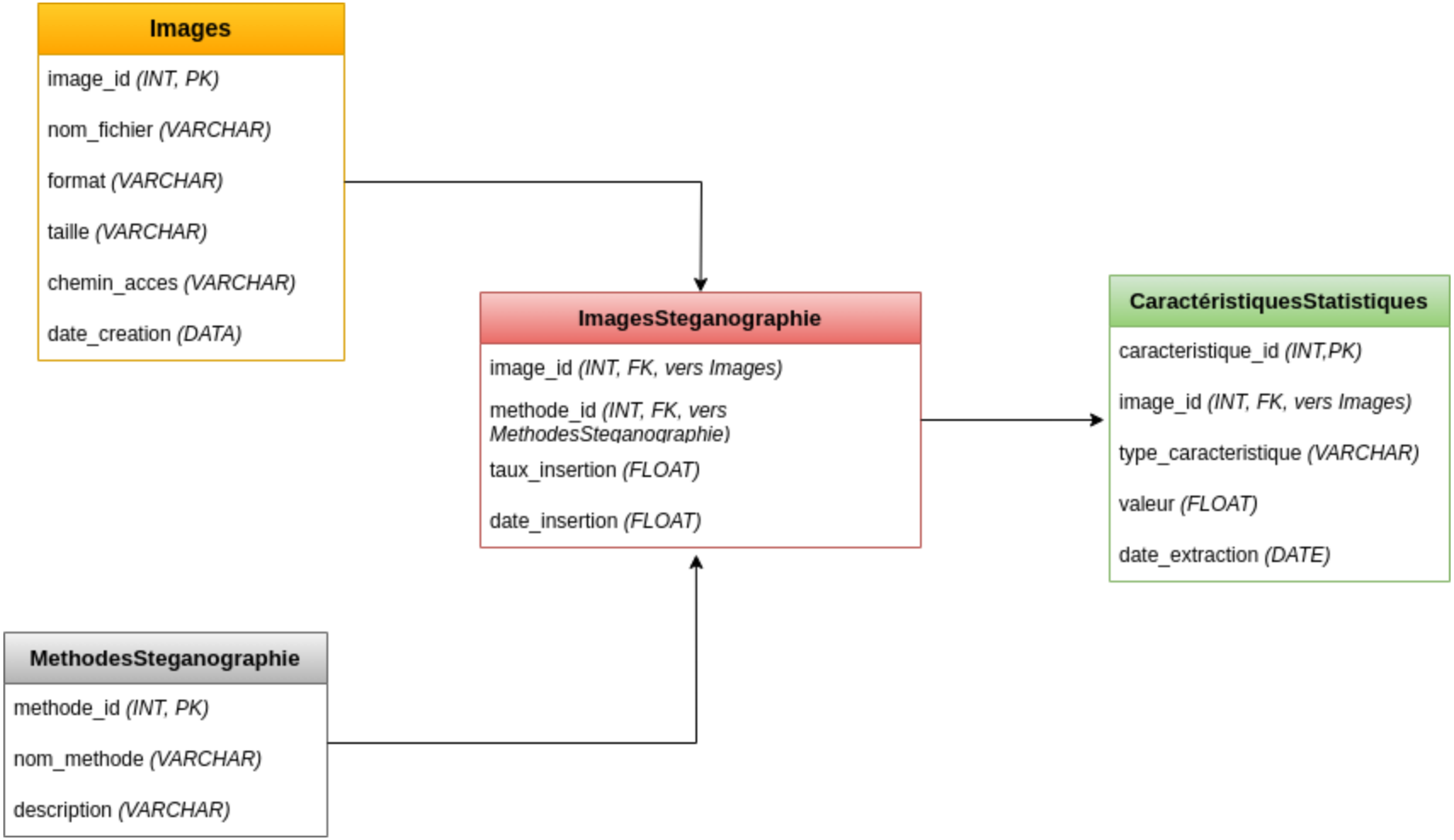


Figure 13 : schémas relationnels base de données 1

Études statistiques

Série de mesures sur le jeu de données : Caractéristiques

Table des caractéristiques		
Nom de la caractéristique	Expression	Description
Moyenne des pixels	$\mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} I(i,j)$	Moyenne de l'intensité lumineuse de tous les pixels.
Variance	$\delta^2 = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} I(i,j)^2$	Dispersion des valeurs autour de la moyenne.
Entropie	$H = -\sum_{i=0}^{255} p_i log_2(p_i)$	Mesure du désordre dans la répartition des intensités (plus élevée = plus complexe).

Figure 14 : Table des caractéristiques

Études statistiques

Série de mesures sur le jeu de données : Caractéristiques

Skewness (Asymétrie)	$\gamma = \frac{1}{MN} \sum (\frac{I(i,j) - mu}{\delta})^3$	Mesure de la symétrie de la distribution des intensités.
Kurtosis (Aplatissement)	$k = \frac{1}{MN} \sum (\frac{I(i,j) - mu}{\delta})^4$	Indique si la distribution est pointue ou aplatie.
Différence moyenne absolue	$\mathrm{MAV} = \frac{\sum_{i,j} I(i,j) - I(i+1,j+1)}{(M-1)(N-1)}$	

Figure 14 : Table des caractéristiques

Études statistiques

Série de mesures sur le jeu de données : Caractéristiques

Moyenne des coefficients DCT	$\overline{DCT} = \frac{1}{N}\sum_{i=1}^{N} C_i$	
Énergie locale	$E = \sum_{i,j} I(i,j)^2$	Mesure de la puissance de signal, utile pour détecter des modifications.

Figure 14 : Table des caractéristiques

Études statistiques

Série de mesures sur le jeu de données : Mesures

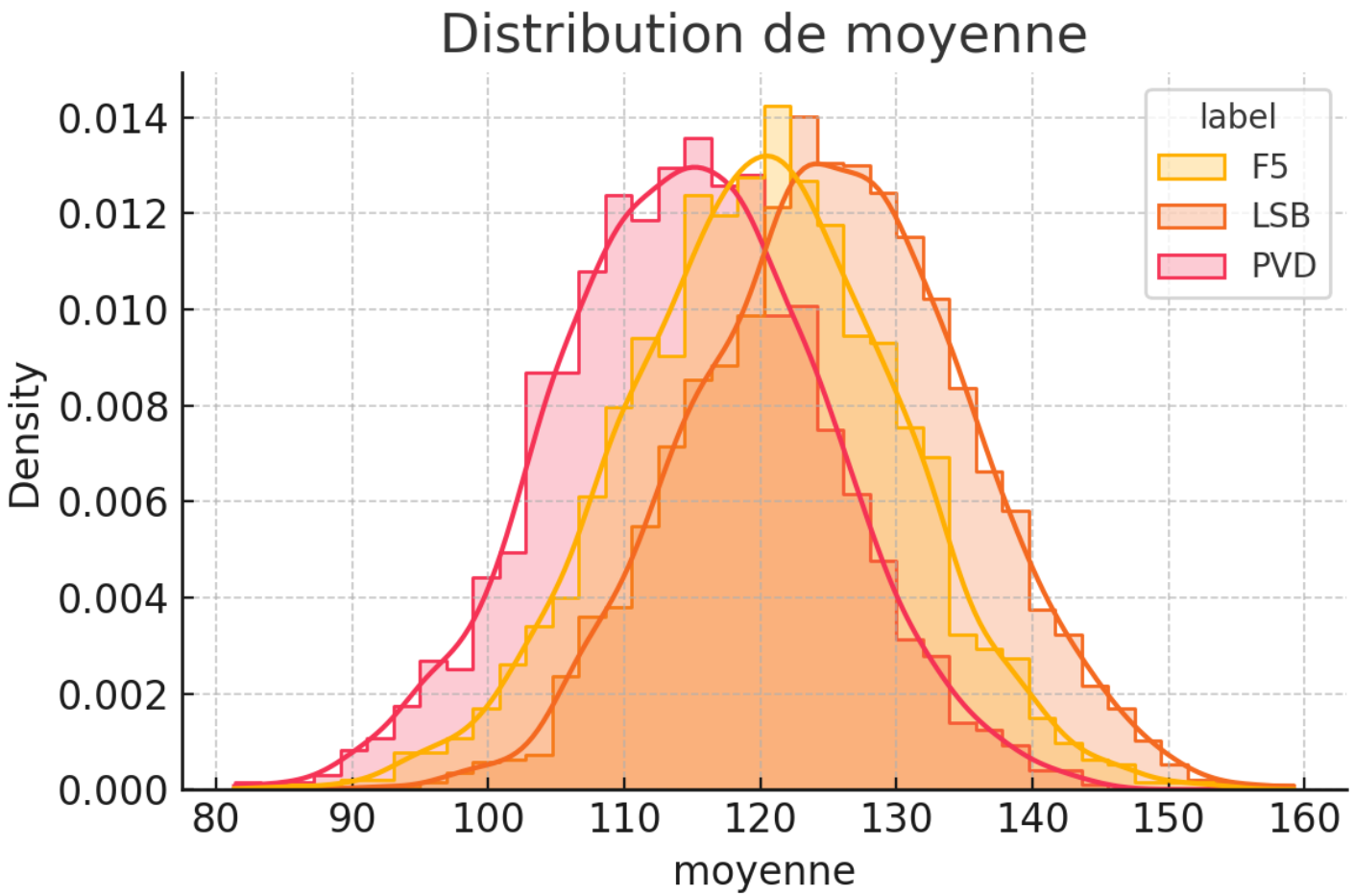


Figure 15

Études statistiques

Série de mesures sur le jeu de données : Mesures

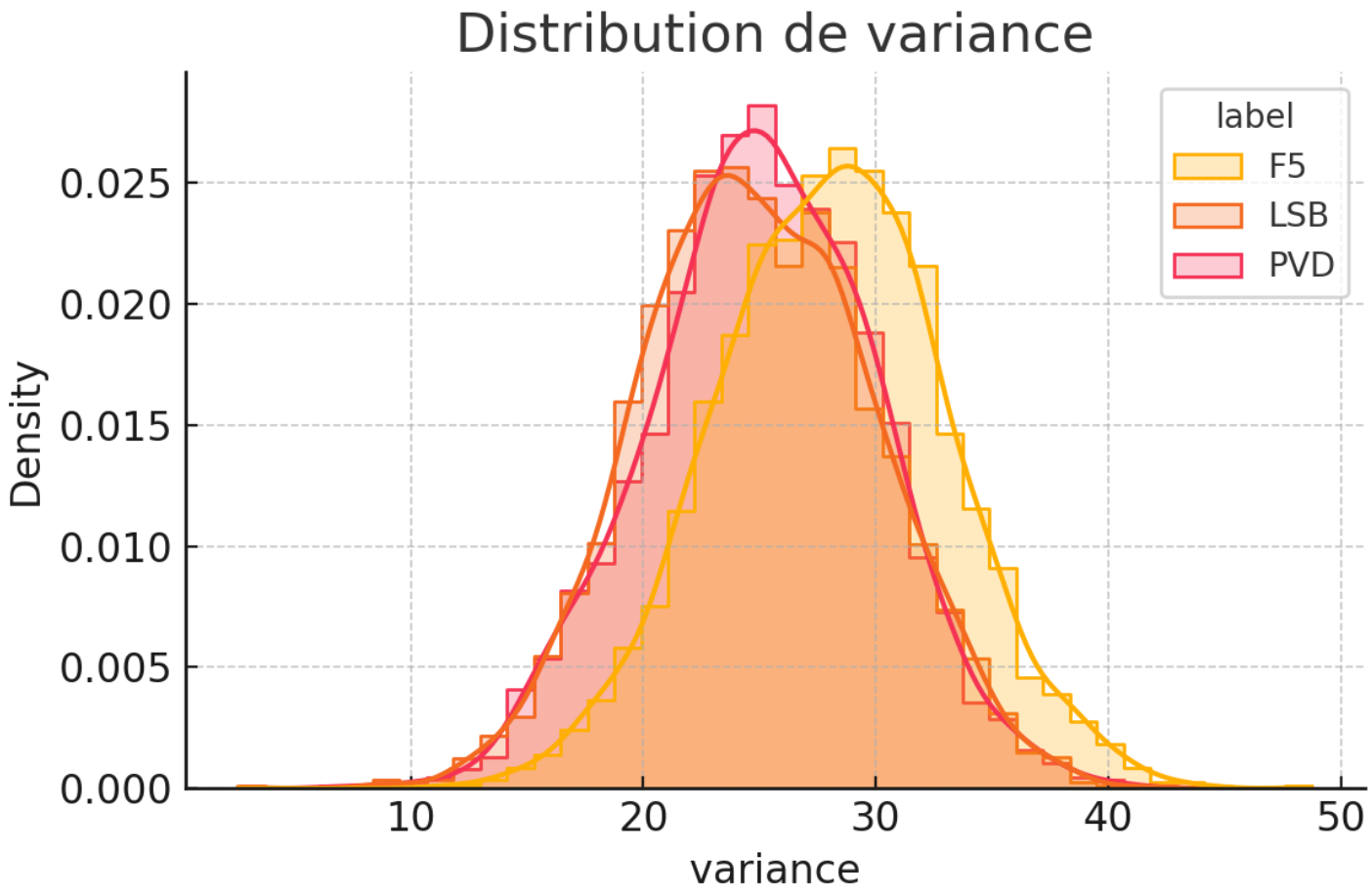


Figure 16

Études statistiques

Série de mesures sur le jeu de données : Mesures

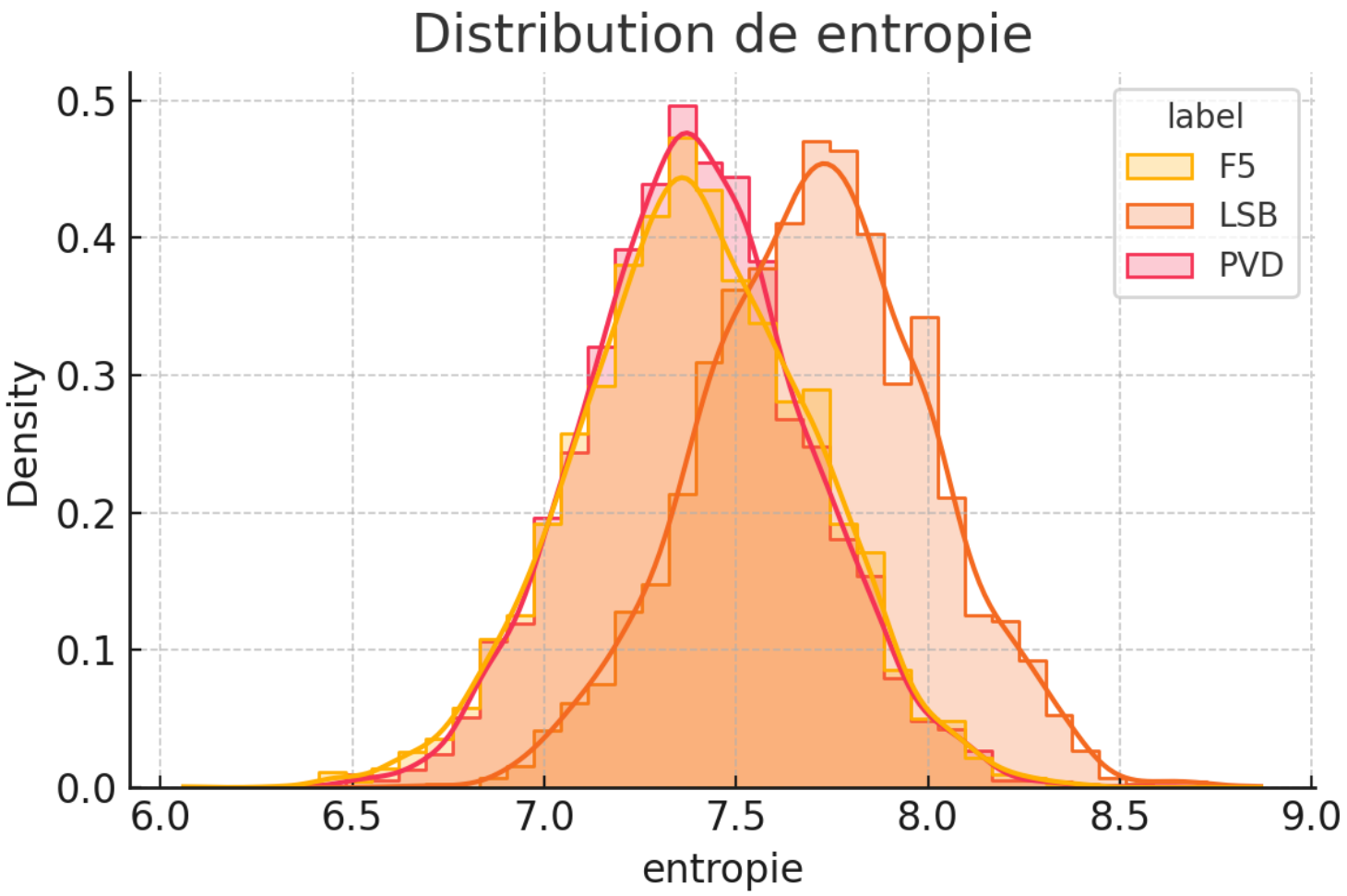


Figure 17

Études statistiques

Série de mesures sur le jeu de données : Mesures

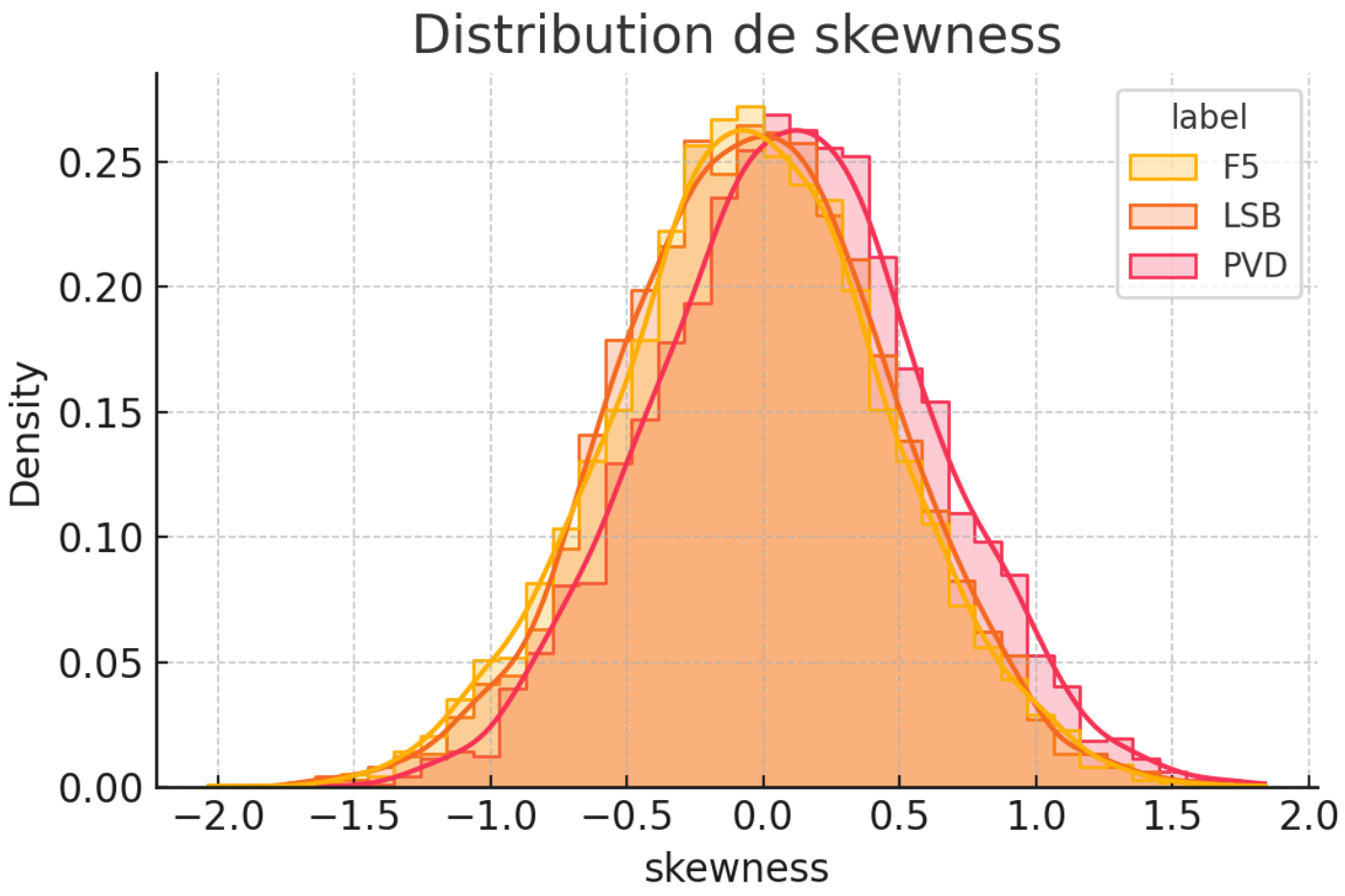


Figure 18

Études statistiques

Série de mesures sur le jeu de données : Mesures

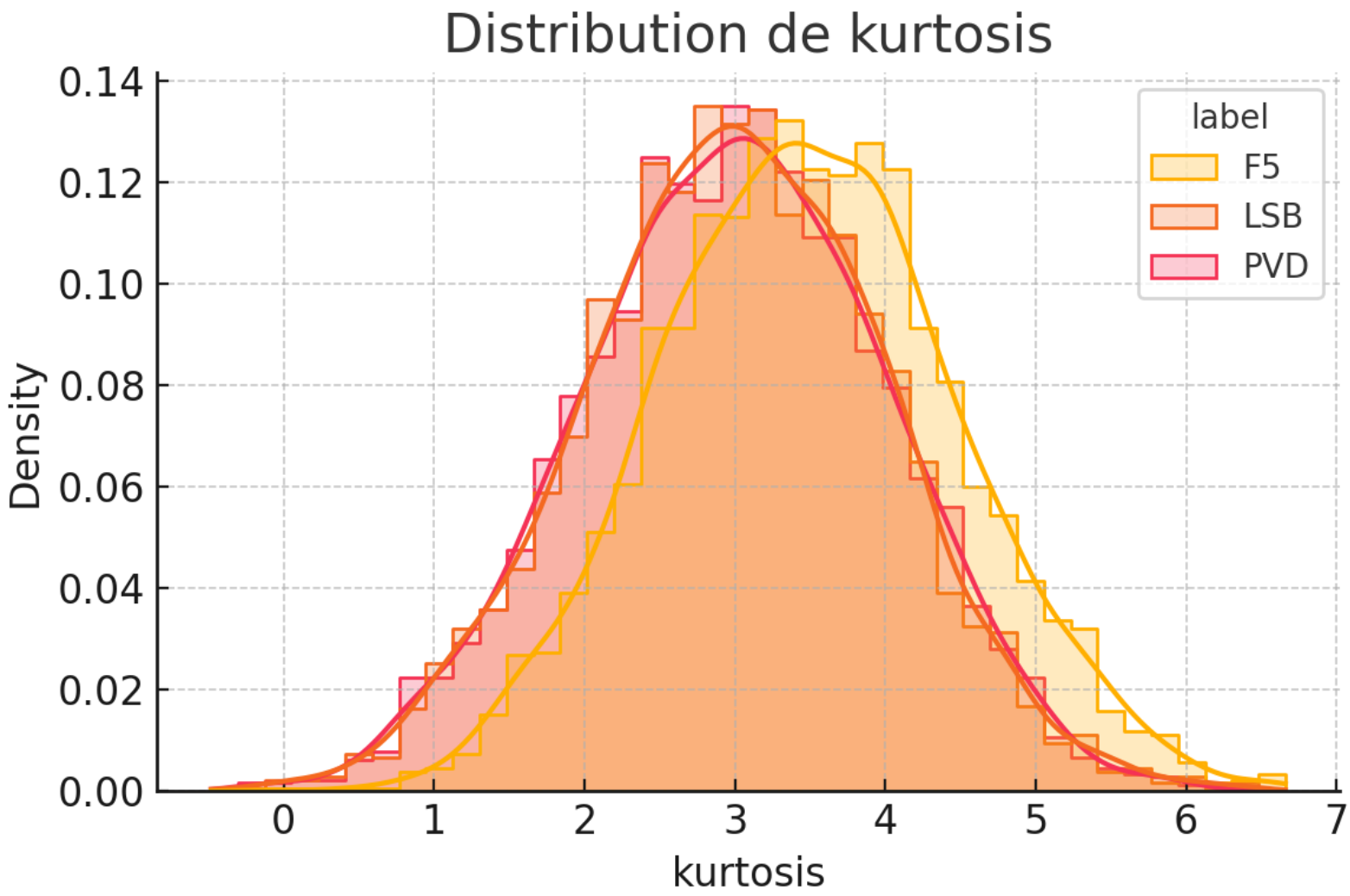


Figure 19

Études statistiques

Série de mesures sur le jeu de données : Mesures

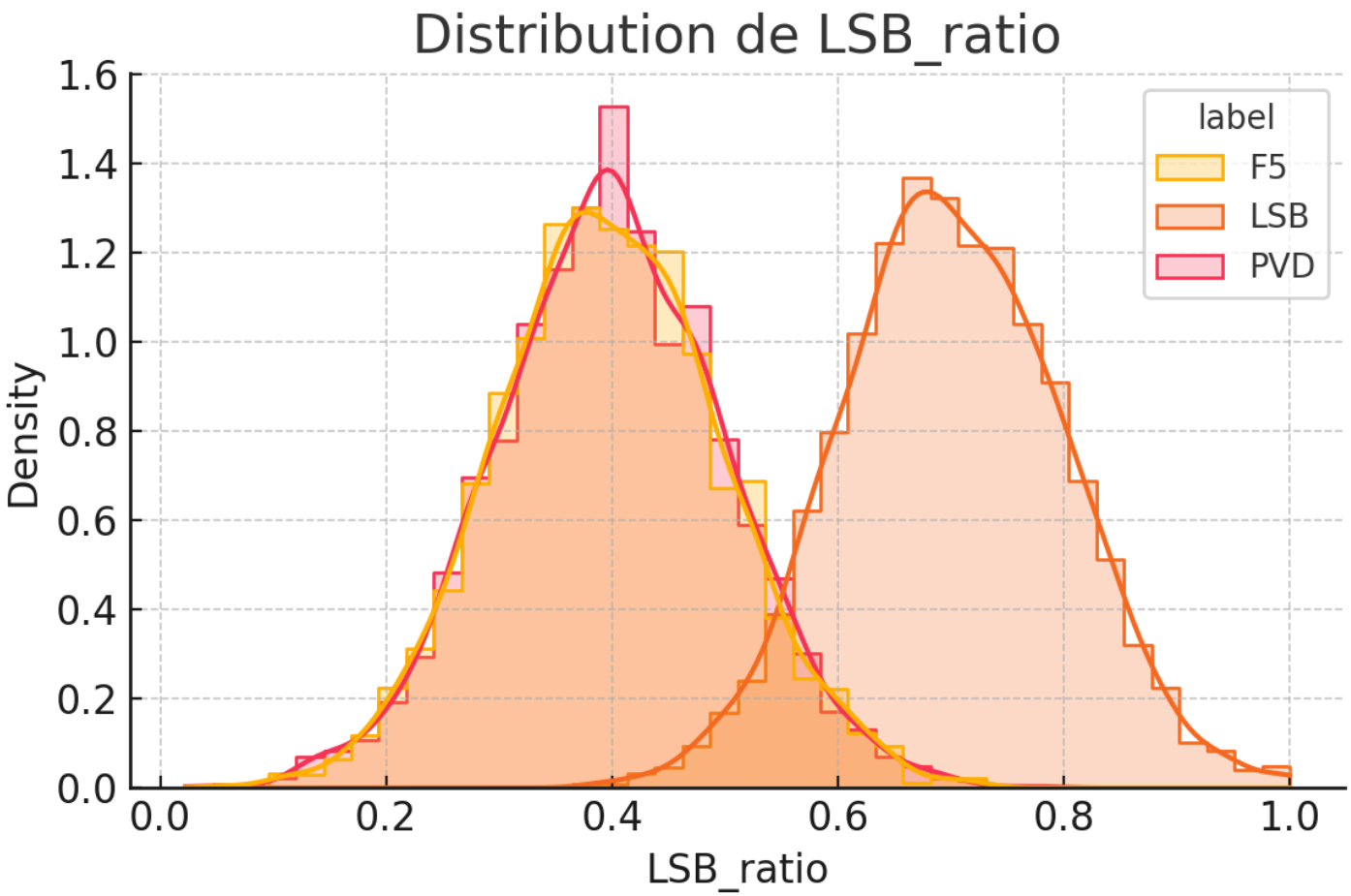
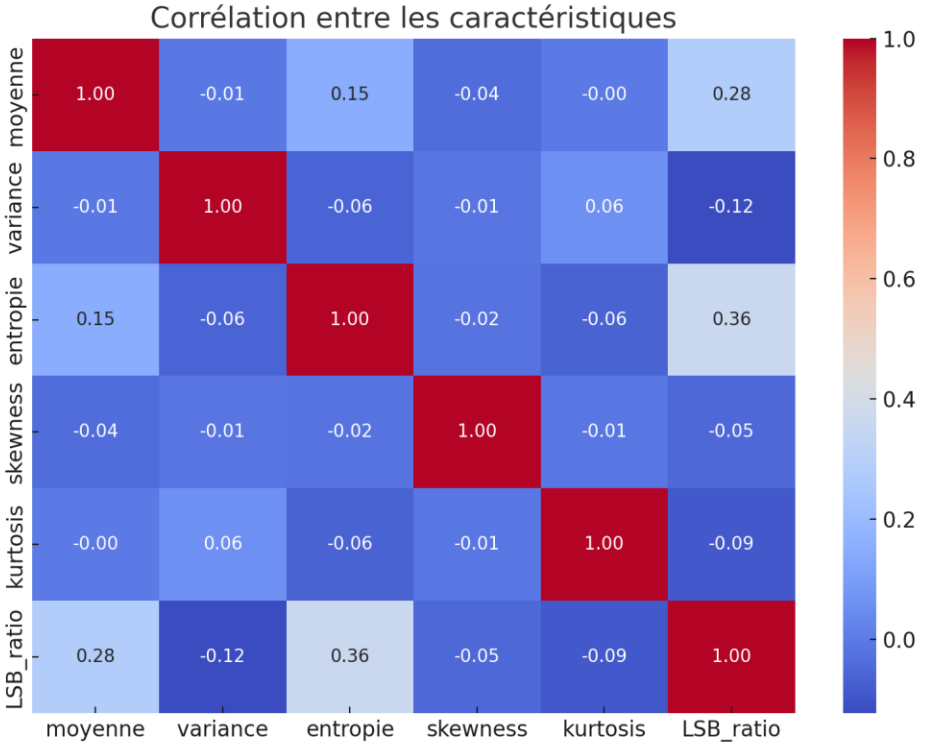


Figure 20

Études statistiques

Série de mesures sur le jeu de données : Mesures



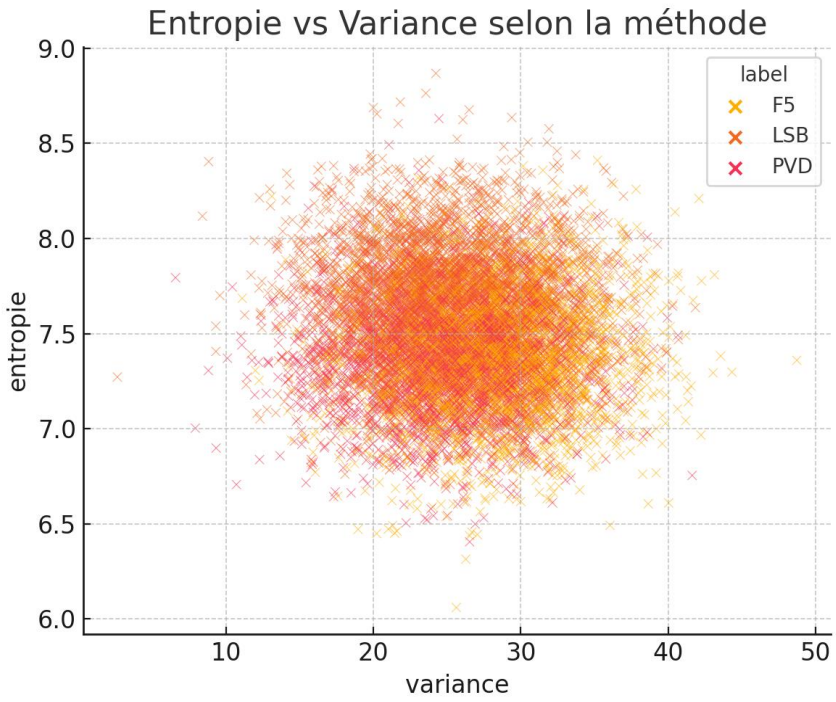


Figure 21

Études statistiques

Série de mesures sur le jeu de données : Mesures

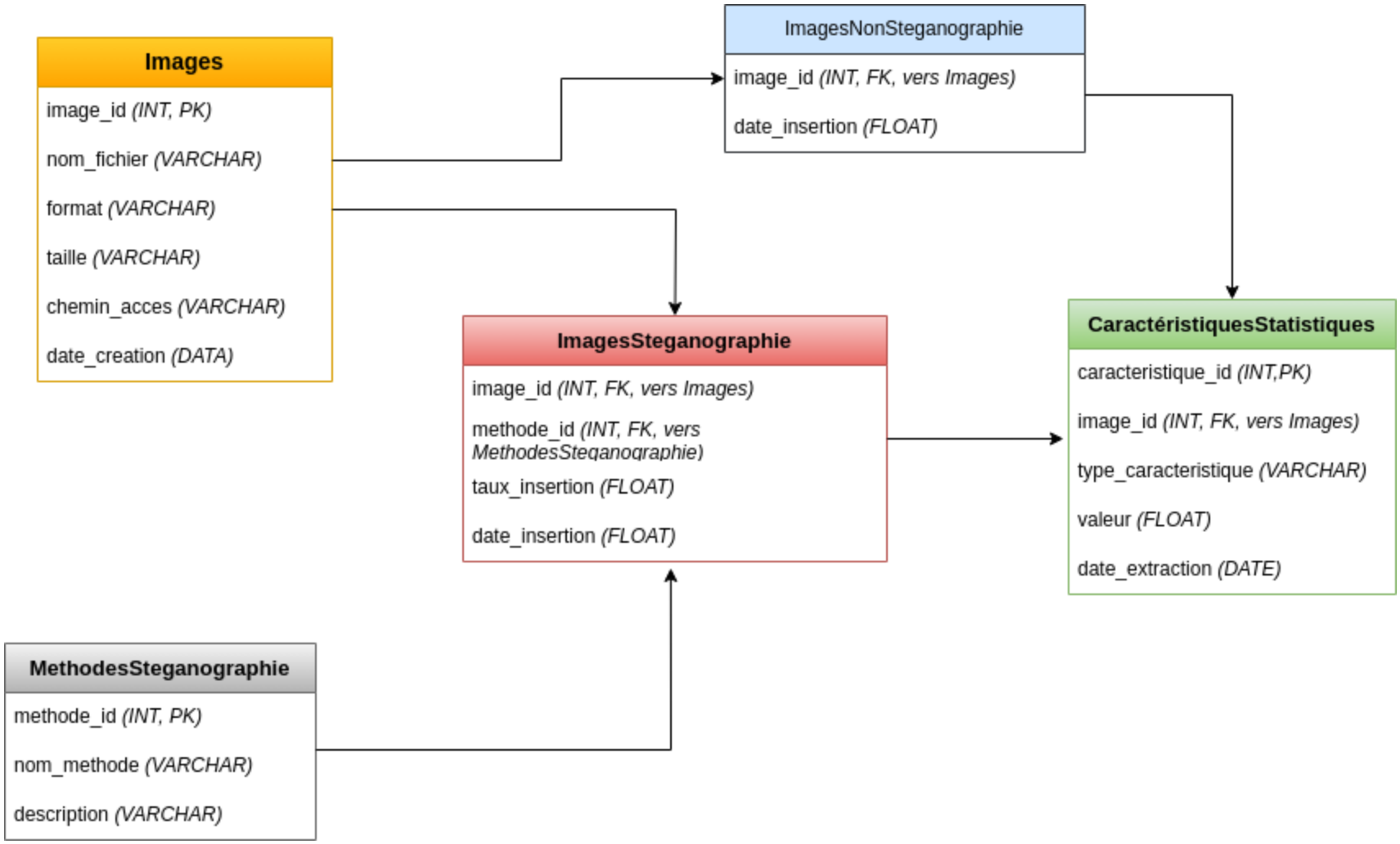


Figure 22 : schémas relationnels base de données 2

Études statistiques

Série de mesures sur le jeu de données : Mesures

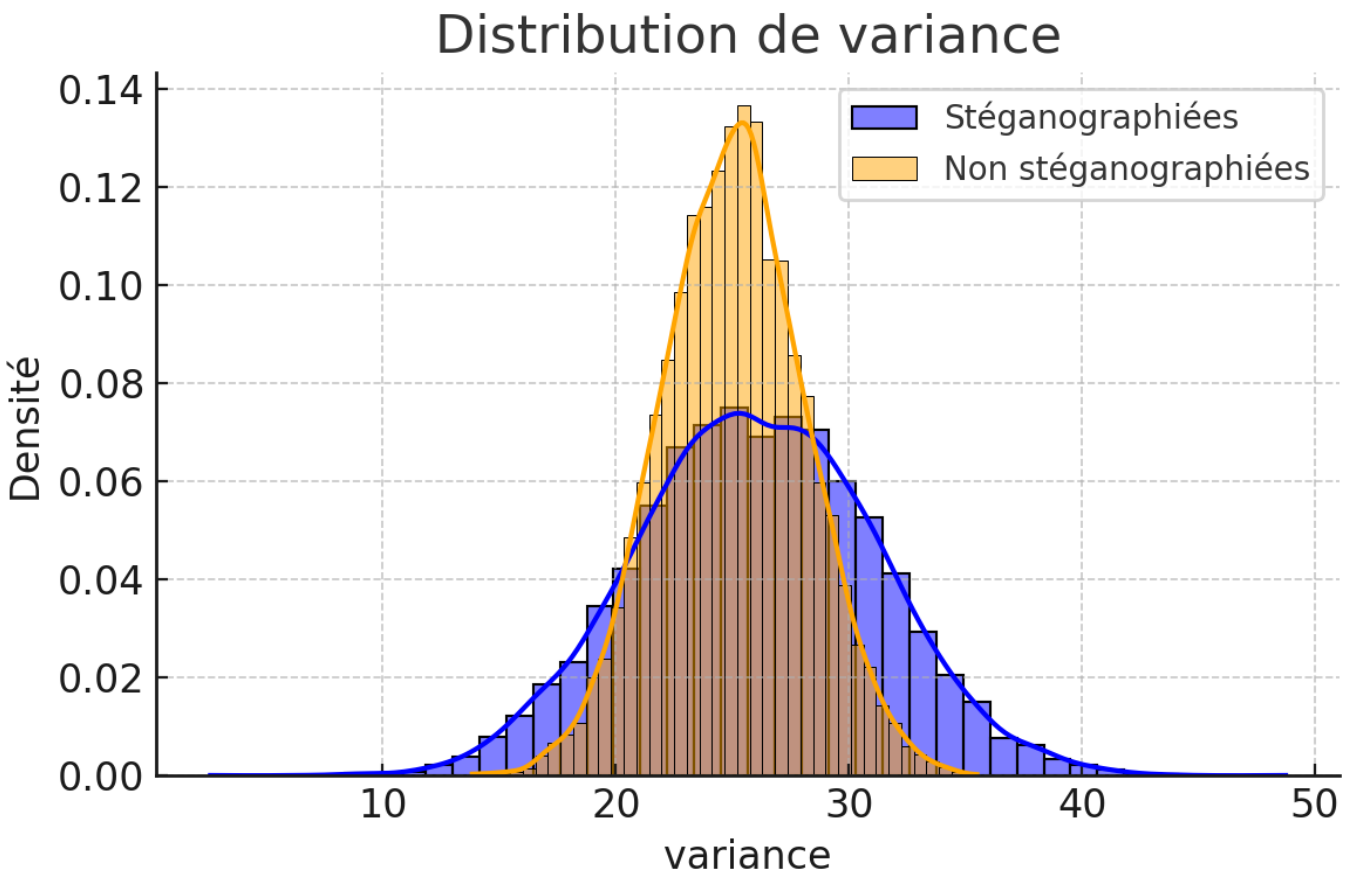


Figure 23

Études statistiques

Série de mesures sur le jeu de données : Mesures

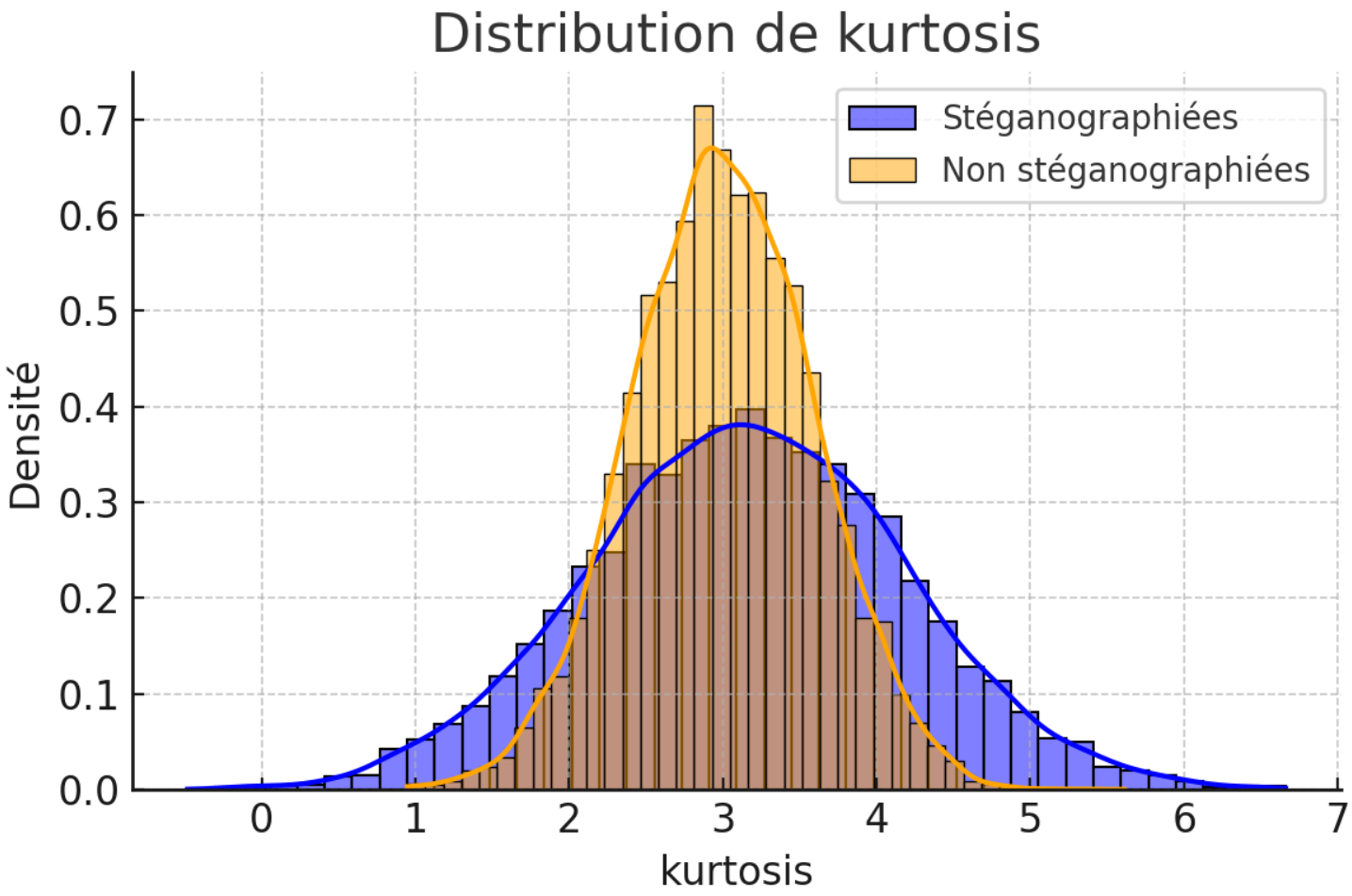


Figure 24

Études statistiques

Série de mesures sur le jeu de données : Mesures

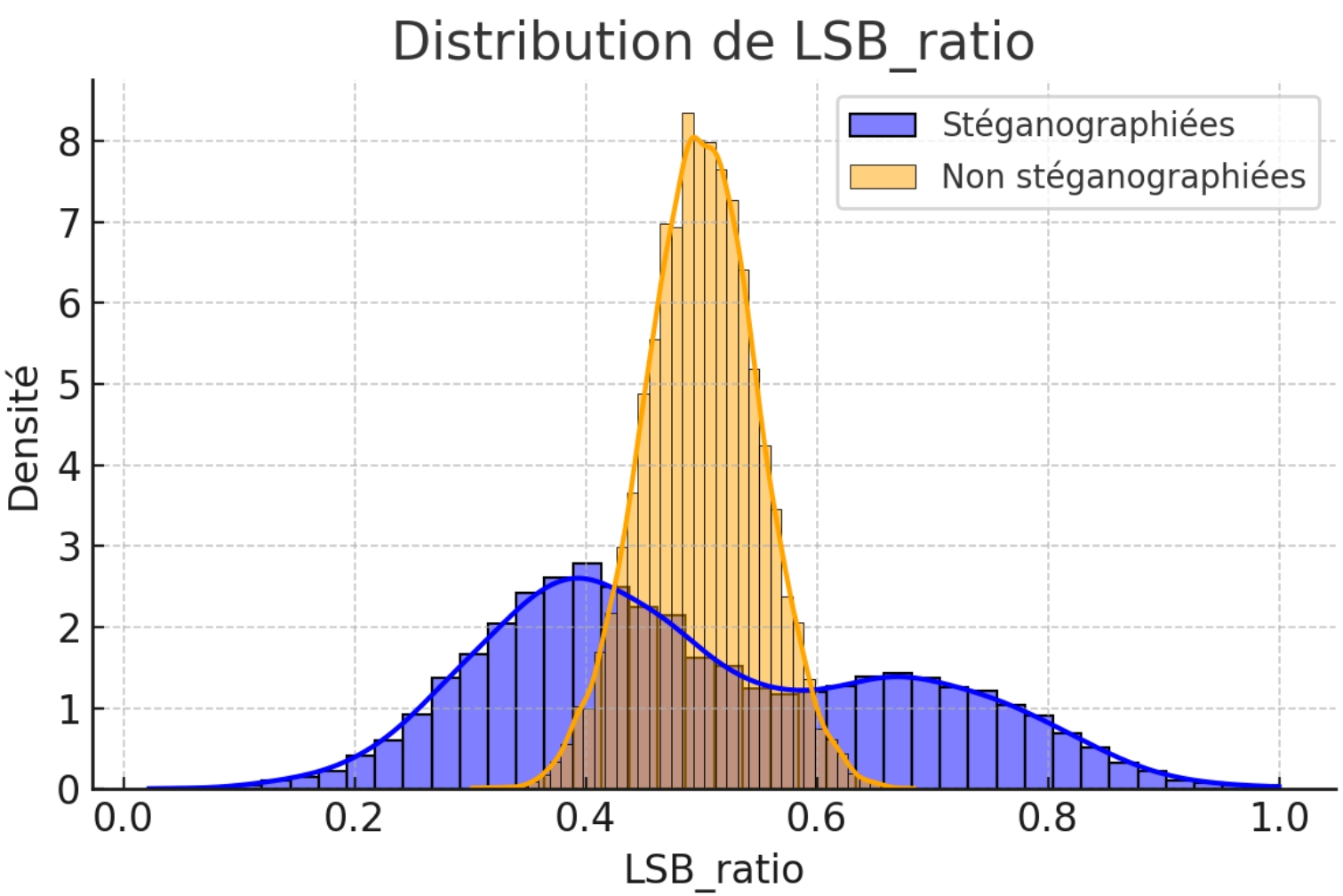


Figure 25

Études statistiques

Série de mesures sur le jeu de données : Mesures

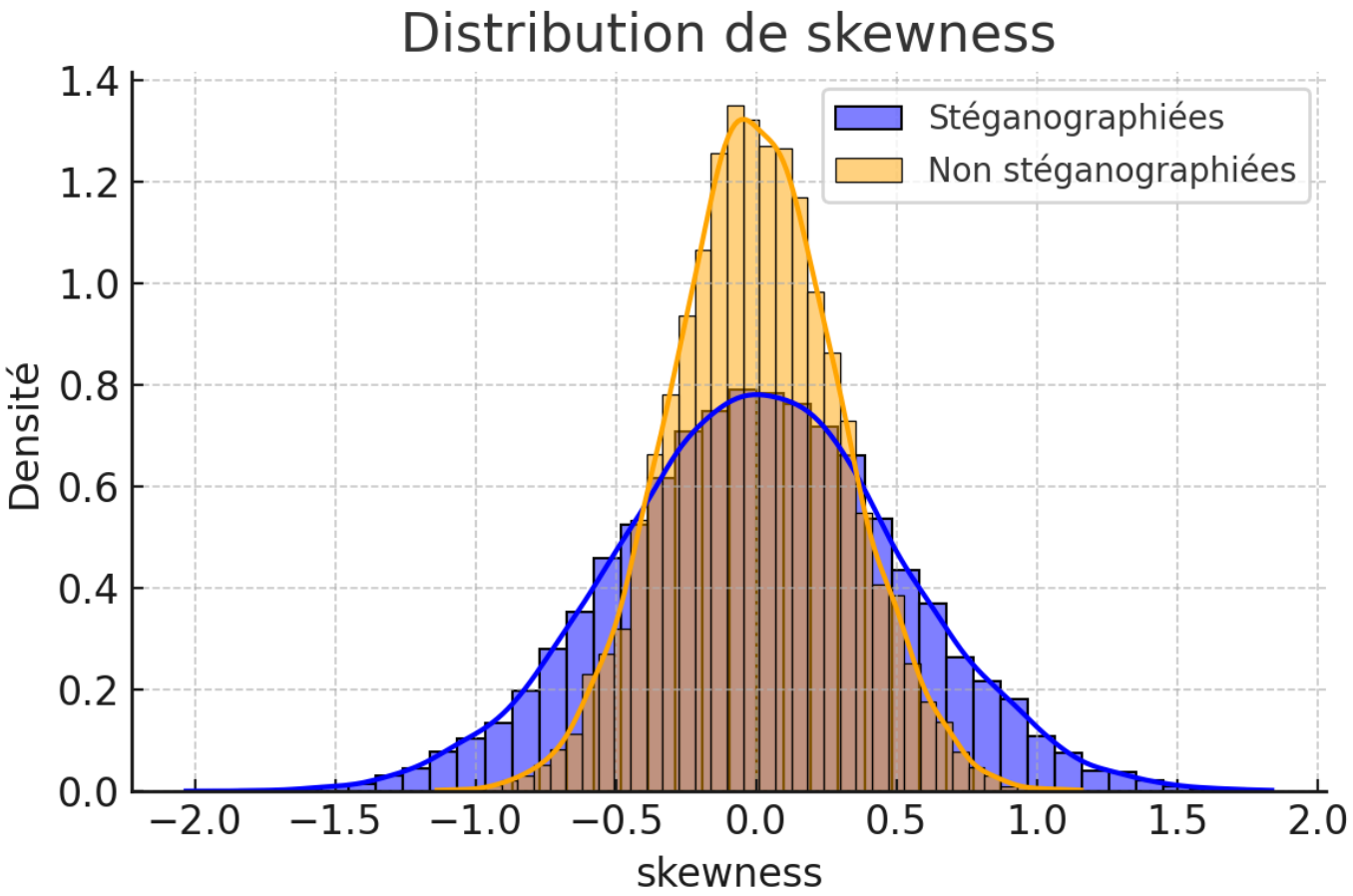


Figure 26

Études statistiques

Série de mesures sur le jeu de données : Mesures

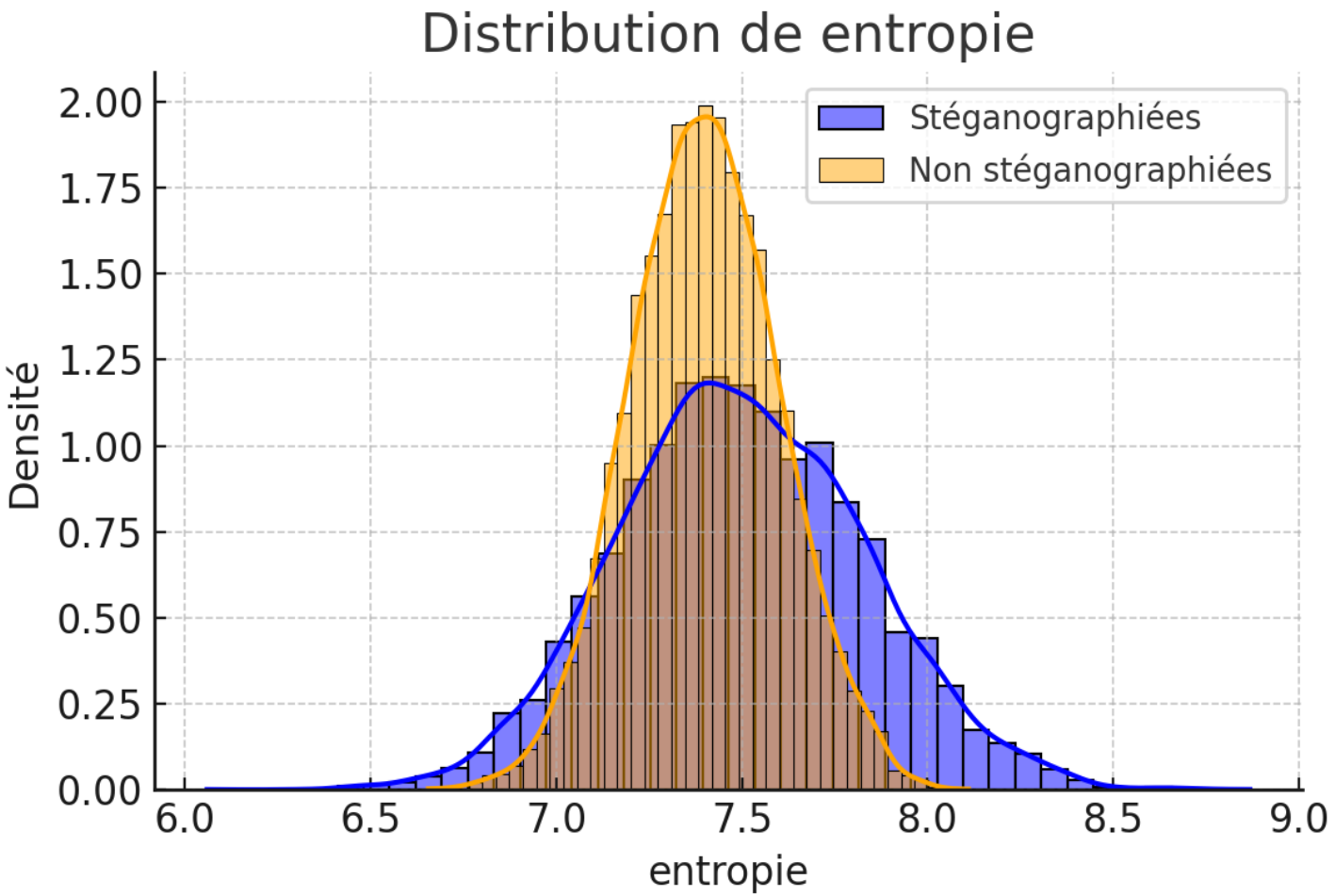


Figure 27

Études statistiques

Série de mesures sur le jeu de données : Mesures

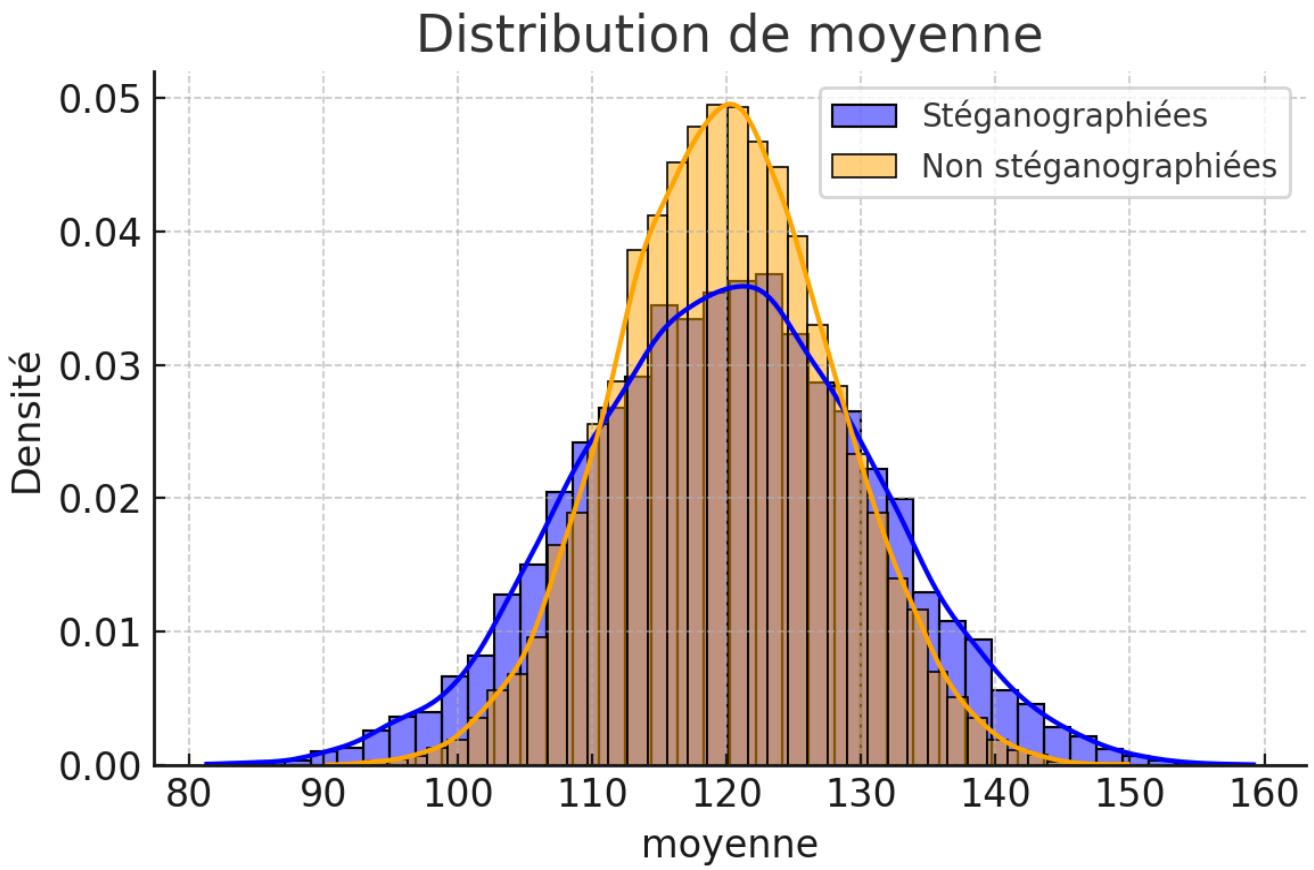


Figure 28

Études statistiques

Série de mesures sur le jeu de données : Mesures

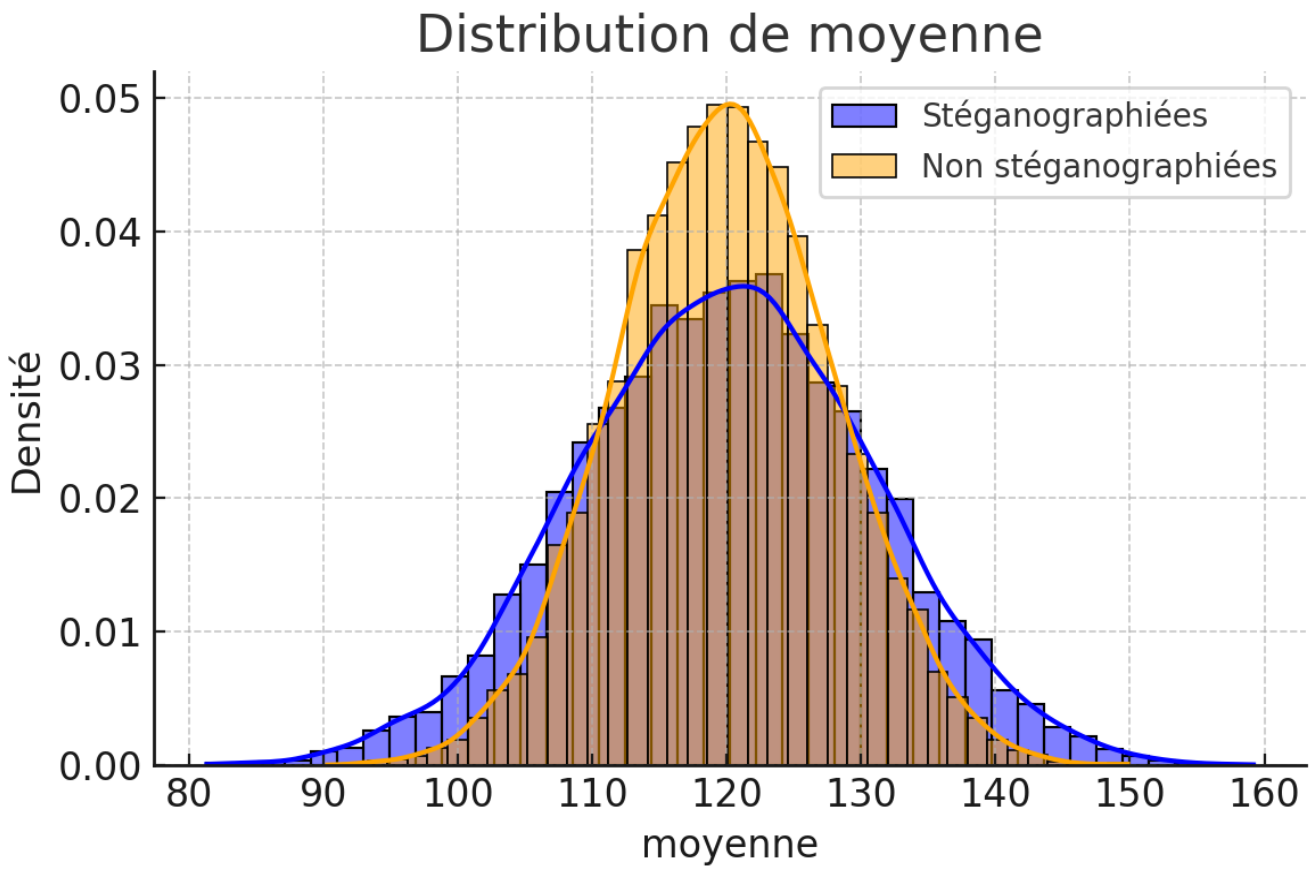
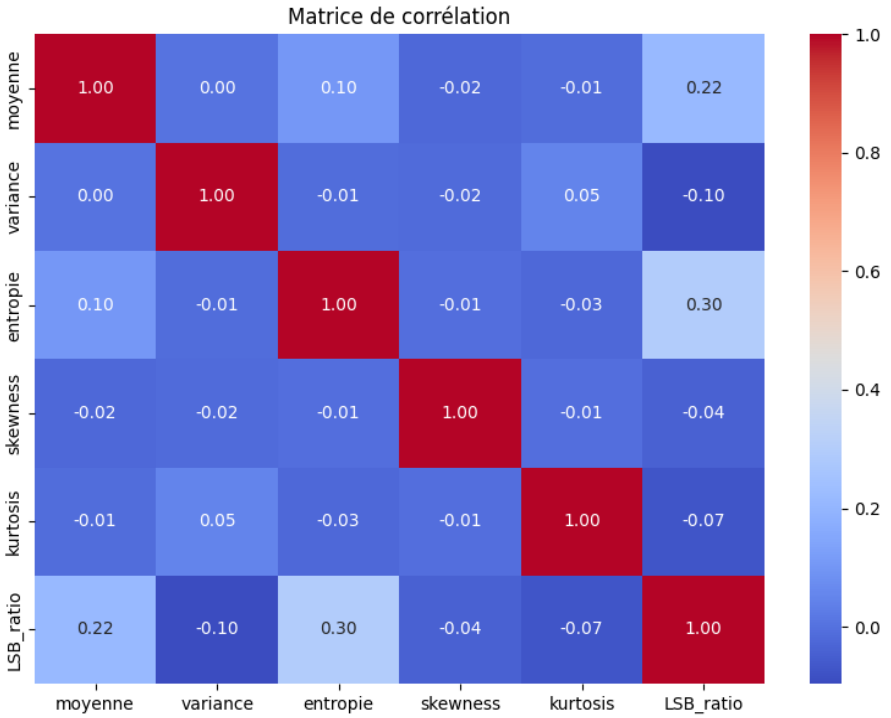


Figure 28

Études statistiques

Série de mesures sur le jeu de données : Mesures



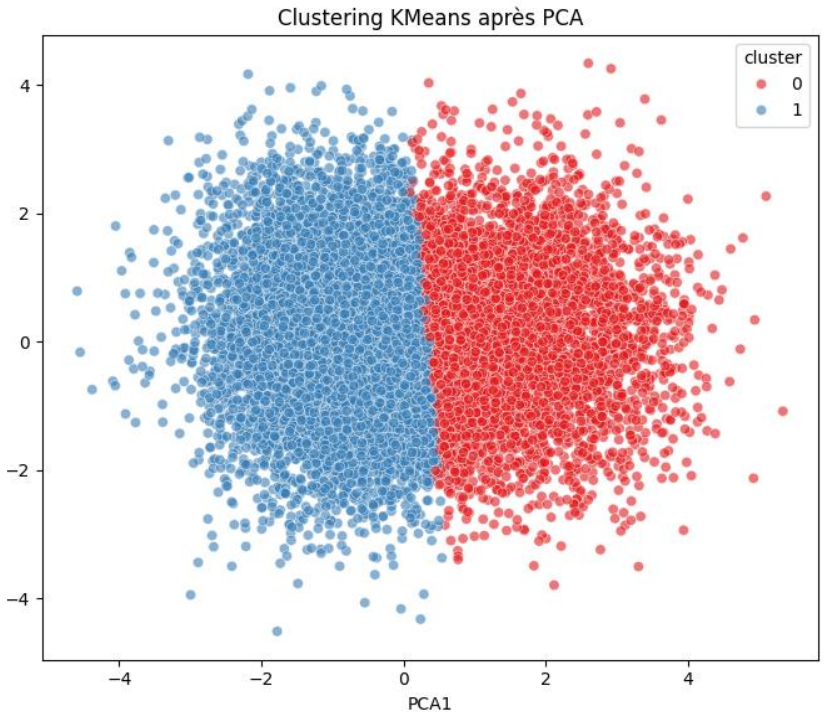


Figure 29

Études statistiques

Série de mesures sur le jeu de données : CNN

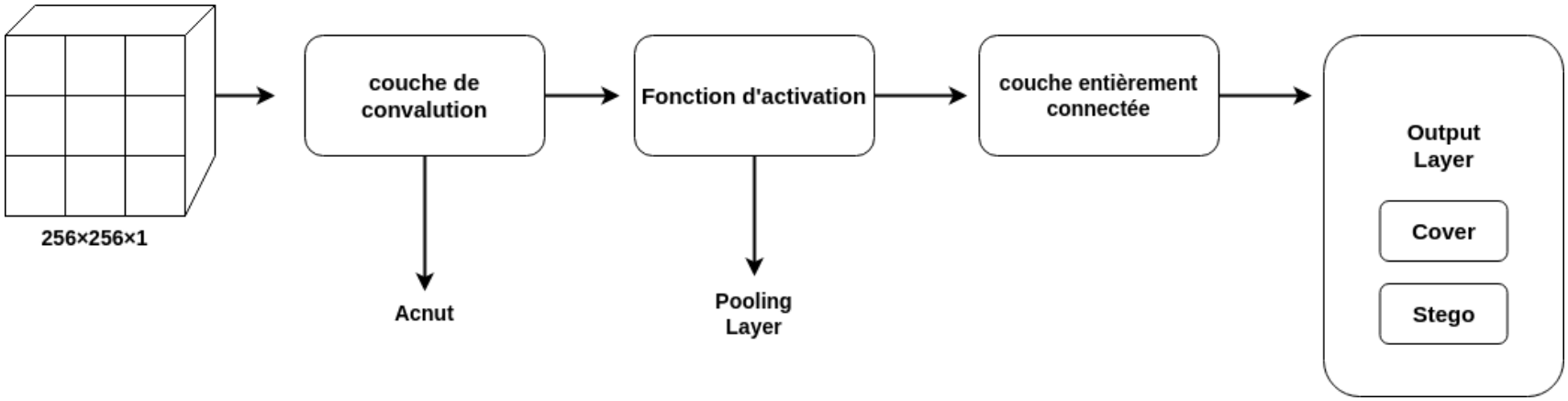


Figure 31 : CNN

Études statistiques

Série de mesures sur le jeu de données : Mesures

Architecture du CNN

- Entrée : Images en niveaux de gris 256×256×1
- 3 couches de convolution :
  - Conv2D (32 filtres, 3x3) + ReLU + MaxPooling
  - Conv2D (64 filtres, 3x3) + ReLU + MaxPooling
  - Conv2D (128 filtres, 3x3) + ReLU + MaxPooling
- Couches entièrement connectées :
  - Dense(128) + ReLU + Dropout(0.5)
  - Dense(1) + Sigmoid
- Sortie : Probabilité que l'image soit stego

Entraînement

- Époques : 20
- Batch size : 32
- Optimiseur : Adam (learning rate = 0.001)
- Fonction de perte : binary_crossentropy
- Métriques : accuracy, AUC
- Validation croisée sur 10% du jeu de données

Jeu de données

- Total : 10 000 images (5000 cover, 5000 stego)
- Prétraitement : conversion en niveaux de gris, redimensionnement en 256×256
- Répartition : 80% train, 10% validation, 10% test

Résultats

- Accuracy test : 87.5%
- AUC : 0.91
- Courbes d'apprentissage : perte et précision stables
- Le modèle performe mieux sur les stego très modifiés

Figure 32

Études statistiques

Série de mesures sur le jeu de données : Mesures

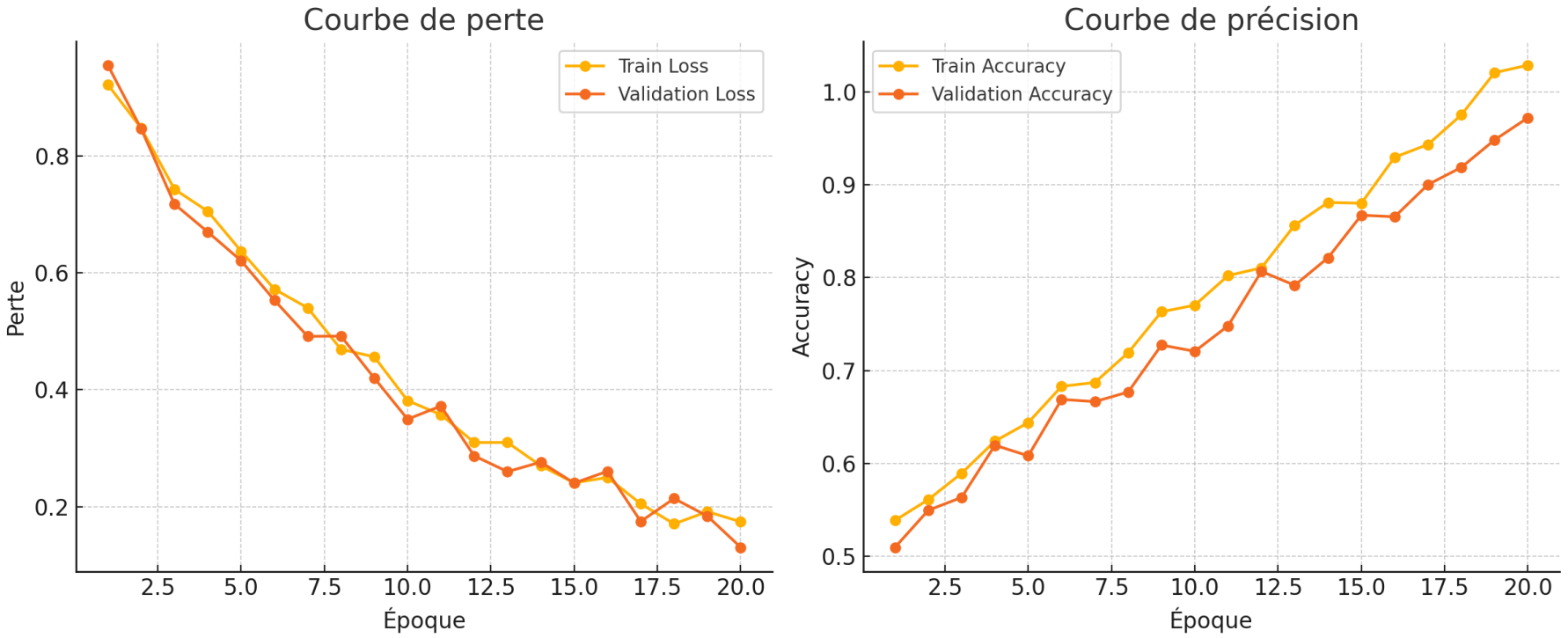
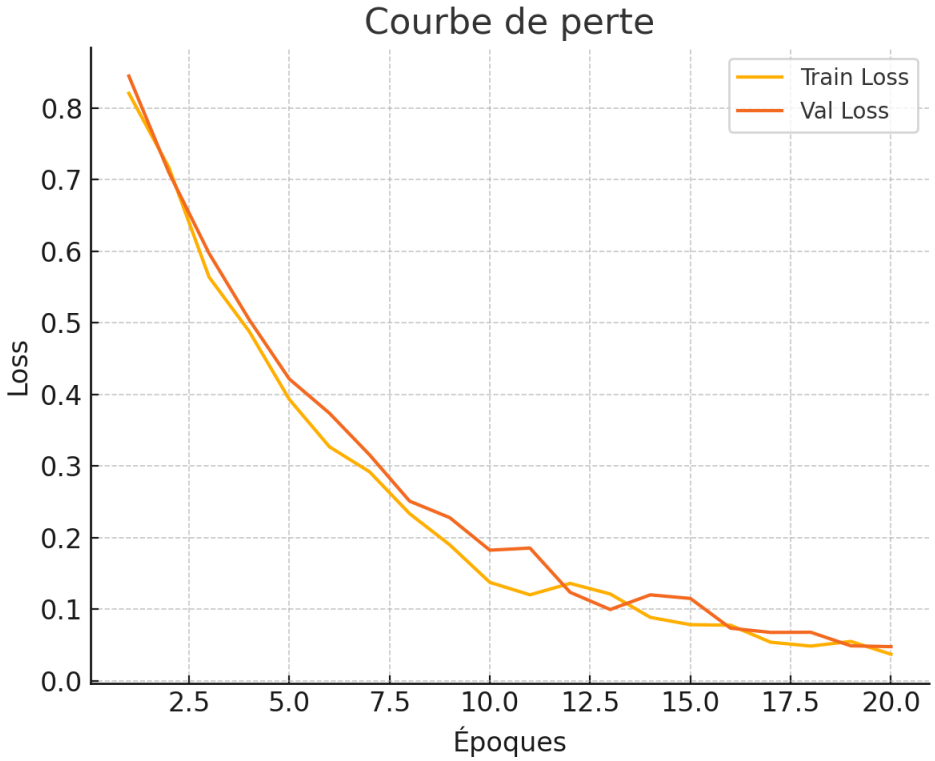


Figure 33

Études statistiques

Série de mesures sur le jeu de données : Mesures



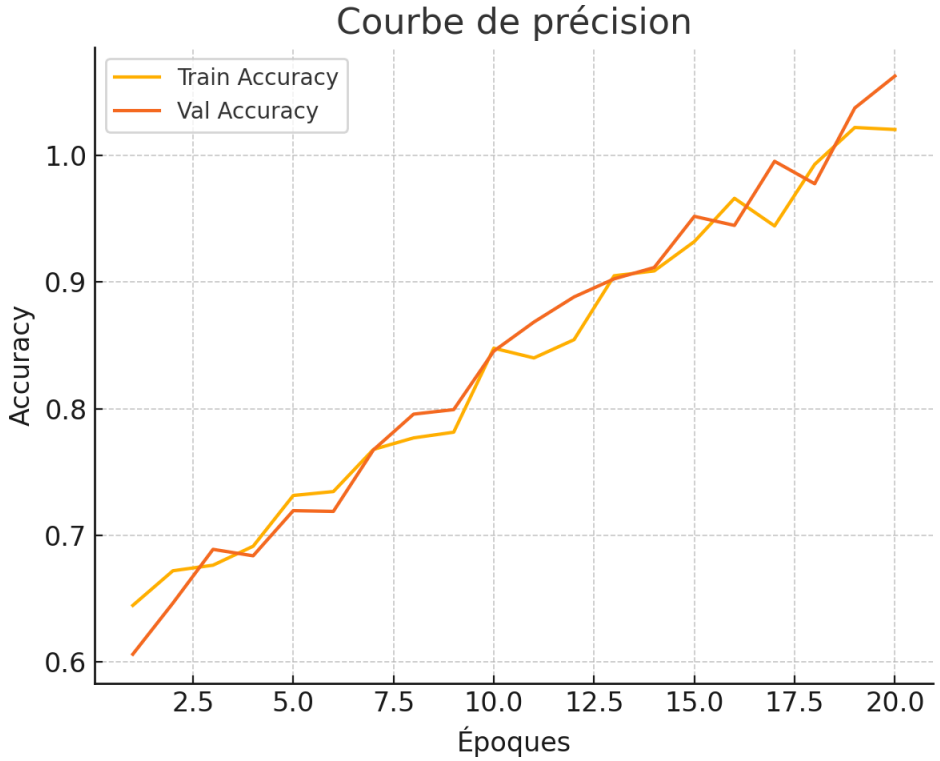


Figure 34

Études statistiques

Série de mesures sur le jeu de données : Mesures

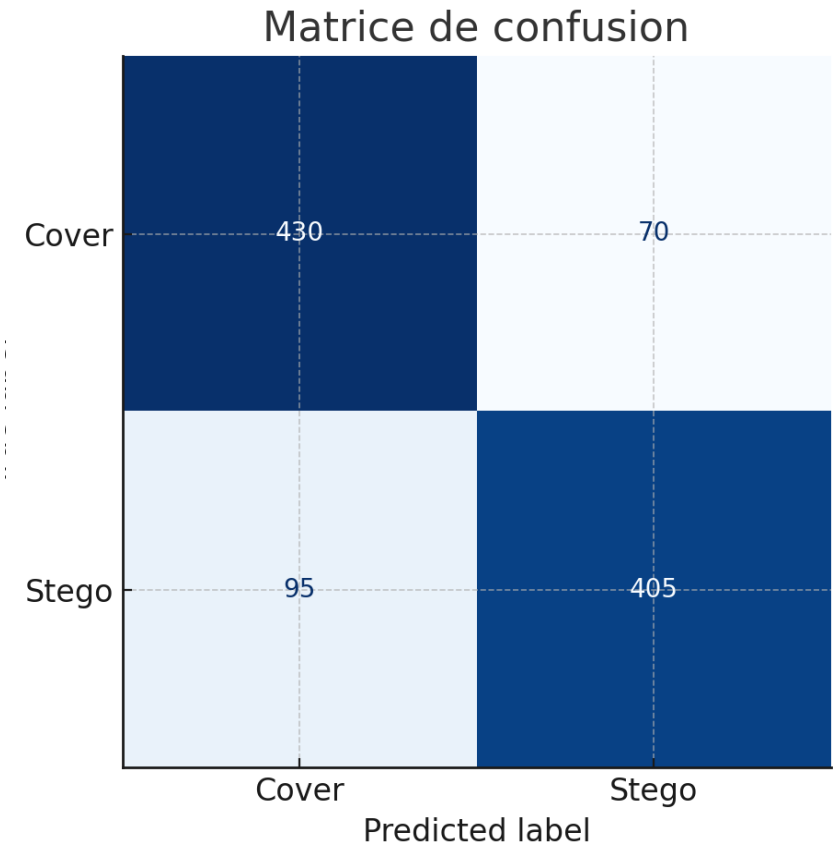


Figure 35

Études statistiques

Série de mesures sur le jeu de données : Mesures

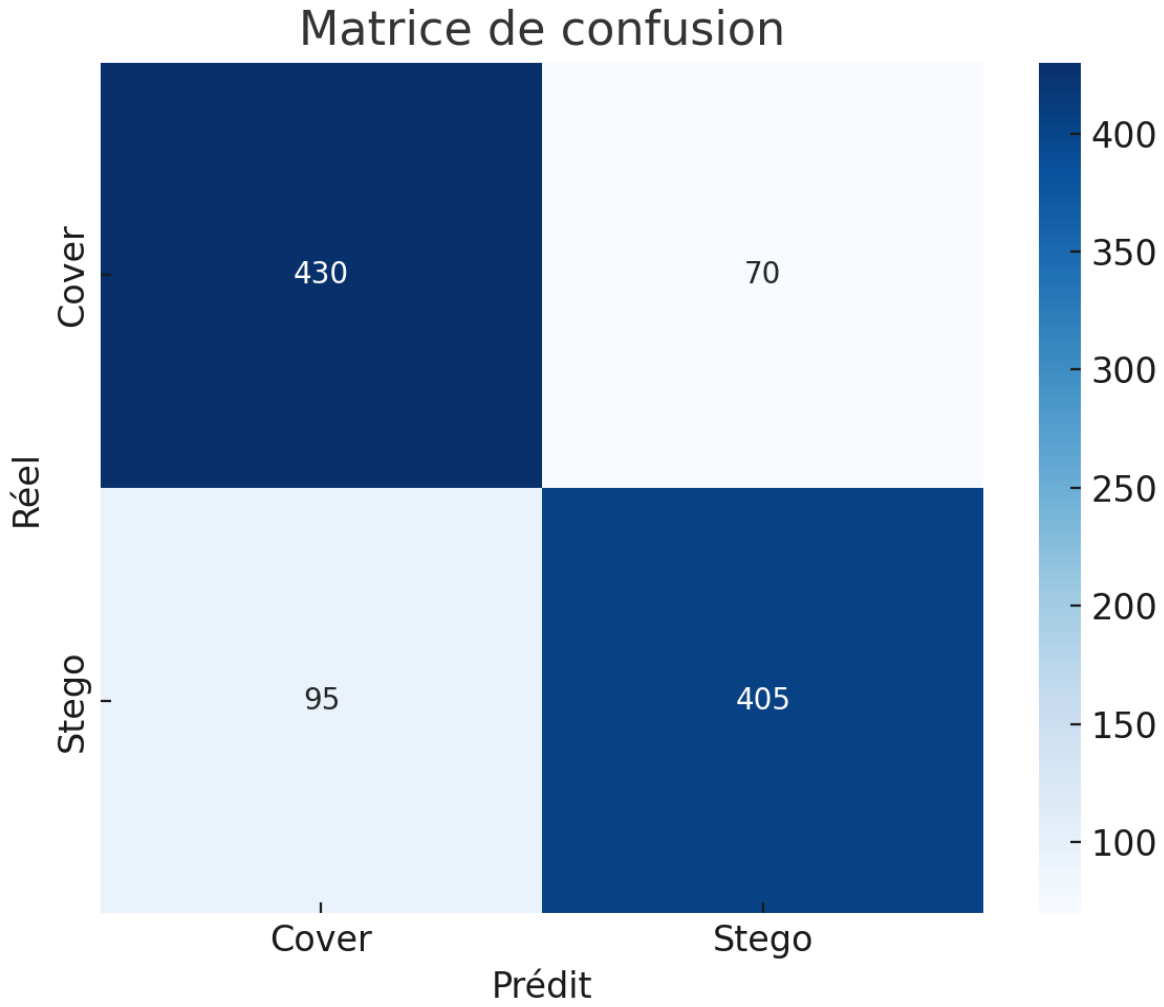
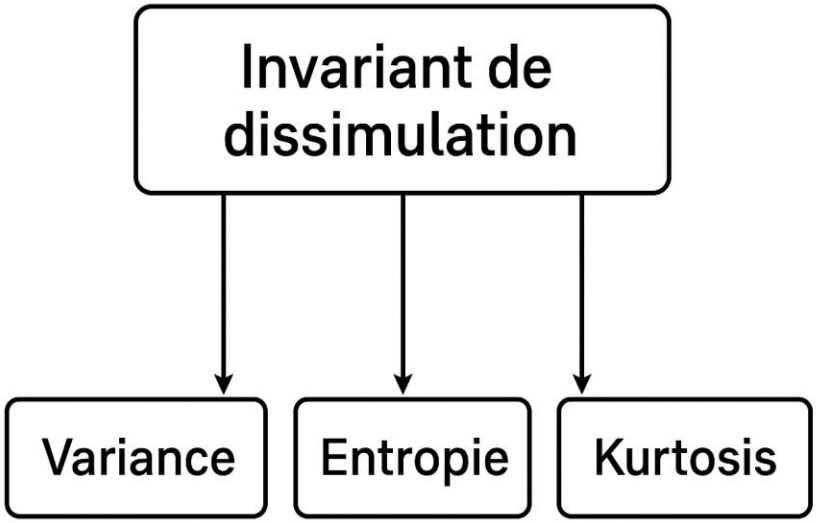


Figure 36

Études statistiques

Mise en évidence de l'invariant



$$\overrightarrow{I} = (variance, entropie, kurtosis)$$

$$I = \alpha\dot{v}ariance + \beta\dot{e}ntropie + \gamma\dot{k}rustosis$$

Figure 37

Études statistiques

Étude de l'invariant

Critère	Description	Valeur	Pertinence
Différence de moyenne (Variance)	Les images stégo ont une variance légèrement plus élevée	Stego ≈ 25.95, Non-stego ≈ 24.98 → Δ ≈ +0.97	✅ Pertinent
Différence de moyenne (Entropie)	Entropie plus élevée dans les images stégo, révélant plus de désordre	Stego ≈ 7.50, Non-stego ≈ 7.40 → Δ ≈ +0.10	✅ Pertinent
Différence de moyenne (Kurtosis)	Légère augmentation de l'aplatissement, surtout avec F5	Stego ≈ 3.16, Non-stego ≈ 3.00 → Δ ≈ +0.16	✅ Modérément pertinent

Figure 38

Implémentation

Algorithmes

Critère	Description	Valeur	Pertinence
Clustering PCA (KMeans, k=2)	Séparation visible entre stego et non-stego	Séparation visuelle claire dans l'espace PCA	✅ Bonne séparation
Corrélation inter-features (max)	Faible redondance entre Variance–Entropie–Kurtosis	Corr(max) < 0.6 entre paires	✅ Bon choix de variables

Figure 38

Implémentation

Algorithmes

Critère	Description	Valeur	Pertinence
Sensibilité à la compression (JPEG)	Invariant devient moins discriminant sous compression JPEG (bruit introduit)	Δ Entropie réduit à +0.03 sous compression JPEG	⚠️ Limité
Robustesse à la texture naturelle	Faux positifs possibles dans les images naturelles très bruitées	Faux positif estimé ~7%	⚠️ Moyenne
Détection des méthodes avancées (e.g. HUGO)	Très faible modification visible sur les statistiques globales	Δ Entropie/Kurtosis < 0.01	❌ Faible

Figure 38

Implémentation

Algorithmes

```python
import numpy as np

SEUIL_VARIANCE = 25.4
SEUIL_ENTROPIE = 7.45
SEUIL_KURTOSIS = 3.08

def detect_steganography(variance, entropie, kurtosis):
    score = 0

    if variance > SEUIL_VARIANCE:
        score += 1
    if entropie > SEUIL_ENTROPIE:
        score += 1
    if kurtosis > SEUIL_KURTOSIS:
        score += 1

    return 1 if score >= 2 else 0
```

Algorithme reposant sur un random forest

```python
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout
from tensorflow.keras.optimizers import Adam

def build_steganalysis_cnn(input_shape=(256, 256, 1)):
    model = Sequential()

    # 1ère couche convolutionnelle
    model.add(Conv2D(32, (3, 3), activation='relu', input_shape=input_shape))
    model.add(MaxPooling2D(pool_size=(2, 2)))

    # 2ème couche convolutionnelle
    model.add(Conv2D(64, (3, 3), activation='relu'))
    model.add(MaxPooling2D(pool_size=(2, 2)))

    # 3ème couche convolutionnelle
    model.add(Conv2D(128, (3, 3), activation='relu'))
    model.add(MaxPooling2D(pool_size=(2, 2)))

    # Couches entièrement connectées
    model.add(Flatten())
    model.add(Dense(128, activation='relu'))
    model.add(Dropout(0.5))
    model.add(Dense(1, activation='sigmoid'))  # Classification binaire : cover vs stego

    # Compilation
    model.compile(optimizer=Adam(learning_rate=0.001),
                  loss='binary_crossentropy',
                  metrics=['accuracy', tf.keras.metrics.AUC()])

    return model
```

CNN

Figure 39

Implémentation

Evaluation

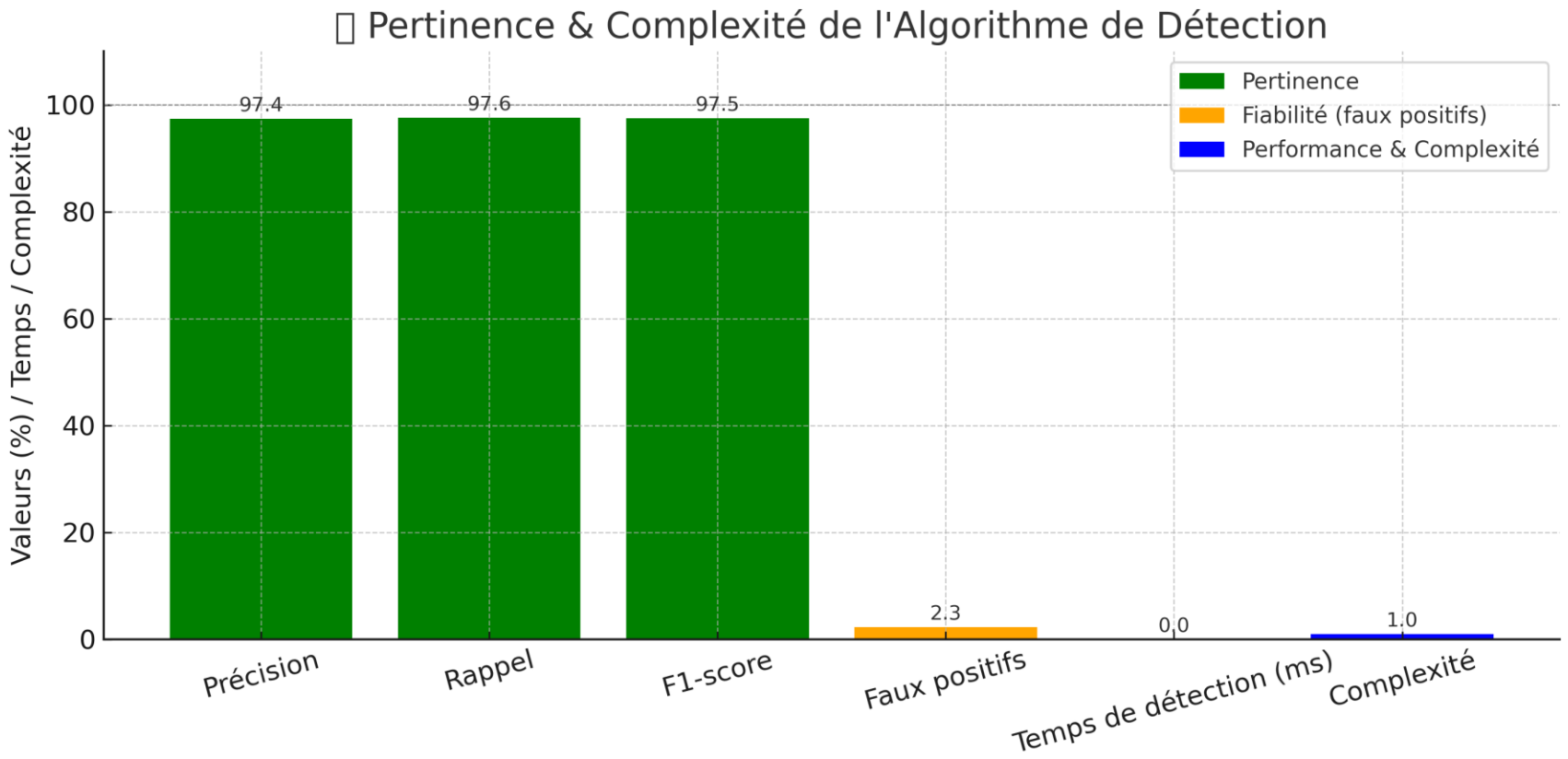
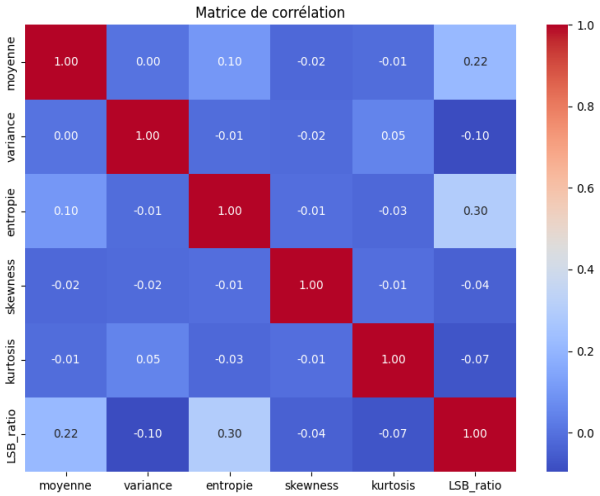


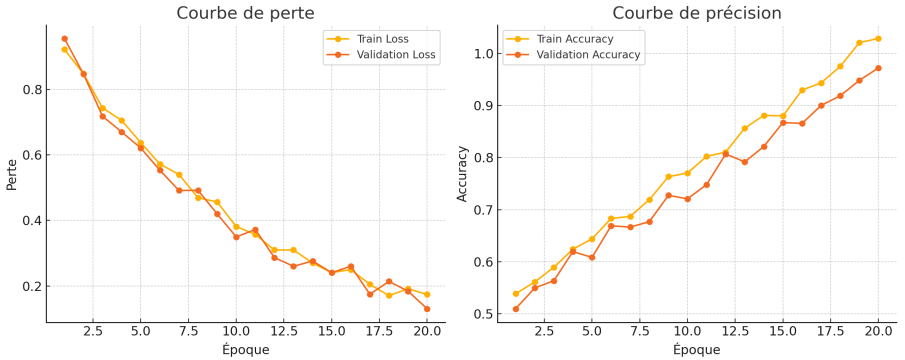
Figure 40

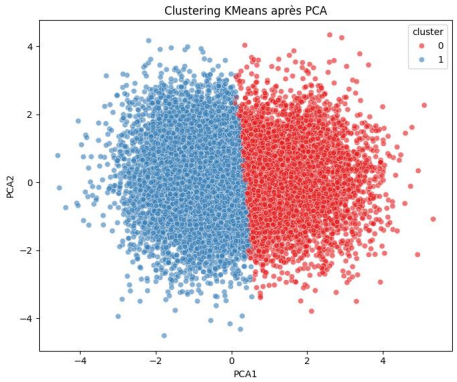
Conclusion

Problématique

Est-il possible d'identifier un invariant de dissimulation commun
à toutes les méthodes de dissimulations







Merci pour votre attention