École Nationale des Sciences Appliquées
d'Oujda

Mini Projet : Indexation et Recherche d'Images par le Contenu (CBIR)

Auteur : Salma Barrak

Filière : GI4

Module : Machine Learning et Traitement d'Image

Responsable : Pr. Khalid EL ASNAOUI

Année Universitaire : 2025/2026

Table des matières

1 Introduction **1**
1.1 Contexte et Enjeux 1
1.2 Objectifs du Projet 1

2 Méthodologie **2**
2.1 Environnement Technique 2
2.2 Phase 1 : Indexation Hors Ligne 2
2.2.1 Préparation et Standardisation des Données 2
2.2.2 Extraction des Descripteurs de Couleur 3
2.2.3 Stockage des Descripteurs 3
2.3 Phase 2 : Recherche en Temps Réel 4
2.3.1 Interface Utilisateur Moderne 4
2.3.2 Processus de Recherche 5
2.4 Mesures de Similarité Implémentées 5
2.4.1 Intersection d'Histogrammes (Swain & Ballard) 5
2.4.2 Distance Euclidienne 6
2.4.3 Distance Chi-Carré 6
2.4.4 Mesure par Corrélation 6

3 Résultats et Analyse **7**
3.1 Visualisation des Descripteurs 7
3.2 Analyse de l'Impact du Nombre de Bins 8
3.2.1 Analyse Détaillée des Niveaux de Quantification 10
3.2.2 Compromis Performance/Précision 11
3.3 Évaluation des Espaces Couleur 11
3.3.1 Analyse Comparative 12
3.4 Comparaison des Mesures de Similarité 13
3.4.1 Analyse des Performances 14
3.5 Configuration Optimale 15

4 Conclusion **16**
4.1 Synthèse des Contributions 16
4.2 Limites Identifiées 16

5 Perspectives **17**
5.1 Améliorations Immédiates 17
5.2 Évolutions à Moyen Terme 17
5.3 Perspectives à Long Terme 17

Chapitre 1

Introduction

1.1 Contexte et Enjeux

Dans l'ère du numérique, la prolifération des bases d'images dans divers secteurs comme la santé, l'industrie et la recherche scientifique pose d'importants défis en matière de gestion et de recherche d'information visuelle. Les approches traditionnelles basées sur les métadonnées textuelles montrent leurs limites face à la croissance exponentielle des données multimédia.

La Recherche d'Images par le Contenu (CBIR) émerge comme une solution innovante qui s'affranchit des annotations textuelles pour se concentrer sur l'analyse directe des caractéristiques visuelles. Cette approche permet de retrouver des images similaires en se basant sur leur contenu intrinsèque plutôt que sur des descriptions externes.

1.2 Objectifs du Projet

Ce projet vise à concevoir et implémenter un système CBIR fonctionnel exploitant les caractéristiques couleur comme descripteurs principaux. Les objectifs spécifiques incluent :

- Développer un pipeline complet d'extraction et d'indexation des caractéristiques couleur
- Implémenter plusieurs algorithmes de mesure de similarité
- Évaluer l'impact des paramètres clés sur la performance du système
- Concevoir une interface utilisateur intuitive pour l'interaction avec le système
- Analyser les performances sur la base de données Corel-1000

Chapitre 2

Méthodologie

2.1 Environnement Technique

Le système a été développé en utilisant un environnement technique cohérent et moderne :

- **Langage de programmation :** Python 3
- **Bibliothèques principales :**
  - OpenCV pour le traitement d'images et le calcul d'histogrammes
  - NumPy pour les opérations matricielles et numériques
  - Matplotlib pour la visualisation des résultats et histogrammes
  - Tkinter pour le développement de l'interface graphique
  - JSON pour le stockage structuré des descripteurs

2.2 Phase 1 : Indexation Hors Ligne

2.2.1 Préparation et Standardisation des Données

La qualité des résultats en CBIR dépend fortement de la cohérence des données d'entrée. Pour garantir des comparaisons fiables, un prétraitement rigoureux a été appliqué à toutes les images de la base Corel-1000 :

Normalisation Dimensionnelle

Chaque image est redimensionnée à une taille fixe de 256×256 pixels. Cette standardisation élimine les variations de dimension qui pourraient fausser les calculs d'histogrammes. L'interpolation `INTER_AREA` est utilisée pour préserver la qualité des images lors de la réduction.

Conversion des Espaces Couleur

OpenCV charge naturellement les images dans l'espace BGR. Une conversion systématique en RGB est effectuée pour établir une base de travail cohérente. Cette étape prépare les images pour les conversions ultérieures vers d'autres espaces couleur selon les besoins de l'utilisateur.

2.2.2 Extraction des Descripteurs de Couleur

Concept d'Histobine

Un histogramme couleur complet avec 256 niveaux par canal générerait un vecteur de 256^3 éléments, ce qui est computationnellement prohibitif. L'approche retenue utilise le concept d'histobine qui agrège les informations couleur en un nombre réduit de compartiments (bins).

La fonction mathématique de l'histogramme pour un canal c est donnée par :

$$H_c(i) = \frac{1}{N} \sum_{p \in I_c} [b(p) = i]$$

où H_c représente l'histogramme du canal c, $b(p)$ désigne la bin correspondant au pixel p, et N est le nombre total de pixels.

Niveaux de Quantification

Le système supporte cinq niveaux de quantification : 8, 16, 32, 64 et 256 bins par canal. Ce choix permet d'étudier le compromis entre précision et efficacité computationnelle.

Normalisation des Descripteurs

Chaque histogramme est normalisé par canal pour obtenir une distribution de probabilité. Cette étape cruciale rend les descripteurs invariants à la taille des images et facilite les comparaisons.

2.2.3 Stockage des Descripteurs

Architecture de Stockage

Tous les descripteurs sont stockés dans un fichier JSON unique (`descriptors.json`) utilisant une structure de dictionnaire imbriquée. Cette organisation permet un accès rapide aux descripteurs pendant la phase de recherche.

Indexation Multi-Configurations

Pour permettre l'analyse comparative, le système génère et stocke simultanément plusieurs configurations de descripteurs en variant :

- Les espaces couleur : RGB, HSV, Lab
- Le nombre de bins : 8, 16, 32, 64, 256

2.3 Phase 2 : Recherche en Temps Réel

2.3.1 Interface Utilisateur Moderne

Conception de l'Interface

L'interface Tkinter a été conçue pour offrir une expérience utilisateur fluide et intuitive. Son design moderne intègre :

- Une organisation claire des contrôles de configuration
- Des éléments visuels attrayants (icônes, couleurs)
- Une disposition logique facilitant l'interaction
- Une palette de couleurs moderne (#f5f6f7 pour le fond, #4CAF50 et #2196F3 pour les boutons)

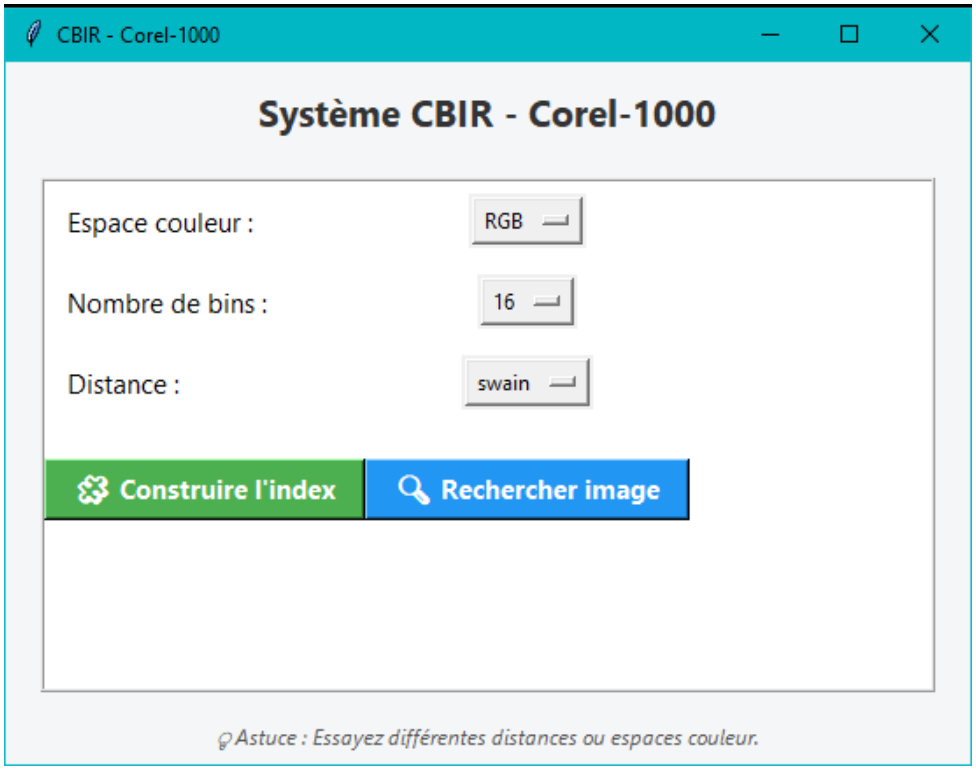

FIGURE 2.1 – Interface graphique du système CBIR développée avec Tkinter

Fonctionnalités Principales

L'interface permet à l'utilisateur de :

- Sélectionner dynamiquement l'espace couleur désiré (RGB, HSV, Lab)
- Choisir le niveau de précision via le nombre de bins (8, 16, 32, 64, 256)
- Spécifier l'algorithme de similarité à utiliser (Swain, Euclidienne, Chi2, Corrélation)
- Visualiser les histogrammes de l'image requête
- Consulter les résultats sous forme de grille organisée

— Construire l'index de la base de données en un clic
— Charger une image requête via un dialogue de fichier

Architecture Interactive

L'interface implémente un système réactif où :

— Les paramètres sont modifiables en temps réel
— Les résultats se mettent à jour automatiquement
— Les messages d'information guident l'utilisateur
— La visualisation des histogrammes est intégrée directement

2.3.2 Processus de Recherche

Traitement de la Requête

L'image requête subit exactement le même traitement que les images indexées :

1. Redimensionnement à 256×256 pixels
2. Conversion vers l'espace couleur sélectionné
3. Calcul de l'histobine avec le nombre de bins choisi
4. Normalisation des histogrammes

Comparaison et Classement

Le système compare le descripteur de la requête avec l'ensemble des descripteurs indexés en utilisant la mesure de distance sélectionnée. Les résultats sont triés selon la pertinence et les 10 meilleures correspondances sont présentées à l'utilisateur.

2.4 Mesures de Similarité Implémentées

2.4.1 Intersection d'Histogrammes (Swain & Ballard)

Cette mesure calcule le degré de chevauchement entre deux distributions de couleur :

$$S_{swain} = \frac{\sum_{i=1}^{N} \min(H_Q[i], H_I[i])}{\sum_{i=1}^{N} H_Q[i]}$$

Elle est particulièrement efficace pour détecter les similarités globales mais peut être sensible aux distributions déséquilibrées.

2.4.2 Distance Euclidienne

Mesure géométrique classique calculant la distance directe entre les vecteurs d'histogrammes :

$$D_{euclidienne} = -\sqrt{\sum_{i=1}^{N}(H_Q[i] - H_I[i])^2}$$

Sa simplicité en fait un choix populaire, mais elle est sensible aux variations d'échelle.

2.4.3 Distance Chi-Carré

Spécialement conçue pour la comparaison d'histogrammes, cette mesure pondère les différences relatives :

$$D_{\chi^2} = -\frac{1}{2}\sum_{i=1}^{N}\frac{(H_Q[i] - H_I[i])^2}{H_Q[i] + H_I[i] + \epsilon}$$

Elle excelle dans la distinction des nuances subtiles entre distributions similaires.

2.4.4 Mesure par Corrélation

Cette approche évalue la relation linéaire entre les histogrammes :

$$Corr = \frac{\sum(H_Q[i] - \bar{H_Q})(H_I[i] - \bar{H_I})}{\sqrt{\sum(H_Q[i] - \bar{H_Q})^2}\sqrt{\sum(H_I[i] - \bar{H_I})^2} + \epsilon}$$

Elle est robuste aux variations d'éclairage globales.

Chapitre 3

Résultats et Analyse

3.1 Visualisation des Descripteurs

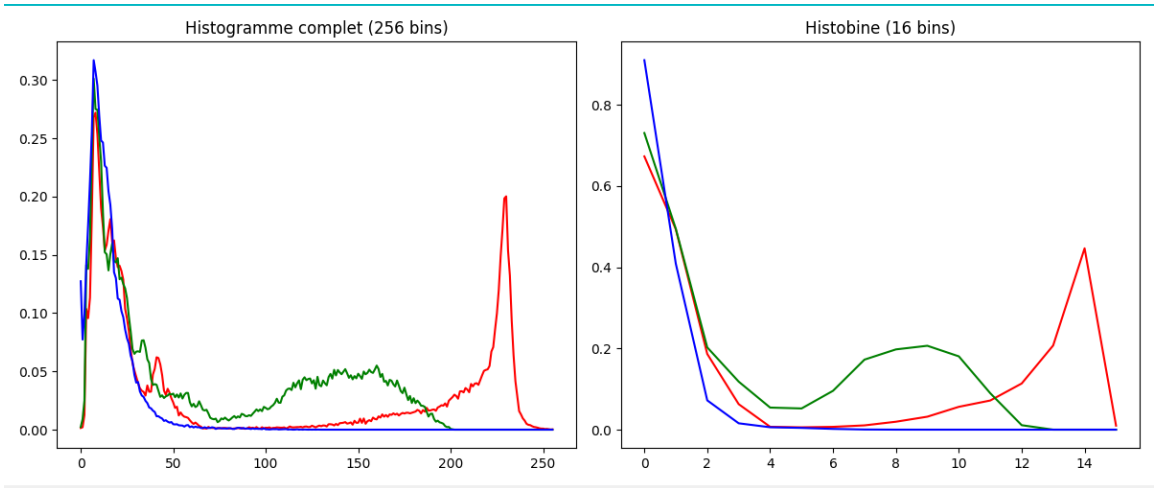

FIGURE 3.1 – Comparaison entre histogramme détaillé (256 bins) et histobine réduite (16 bins)

La visualisation des histogrammes permet de comprendre l'impact de la réduction dimensionnelle. L'histobine préserve les caractéristiques principales tout en éliminant le bruit et en réduisant la complexité computationnelle.

3.2 Analyse de l'Impact du Nombre de Bins

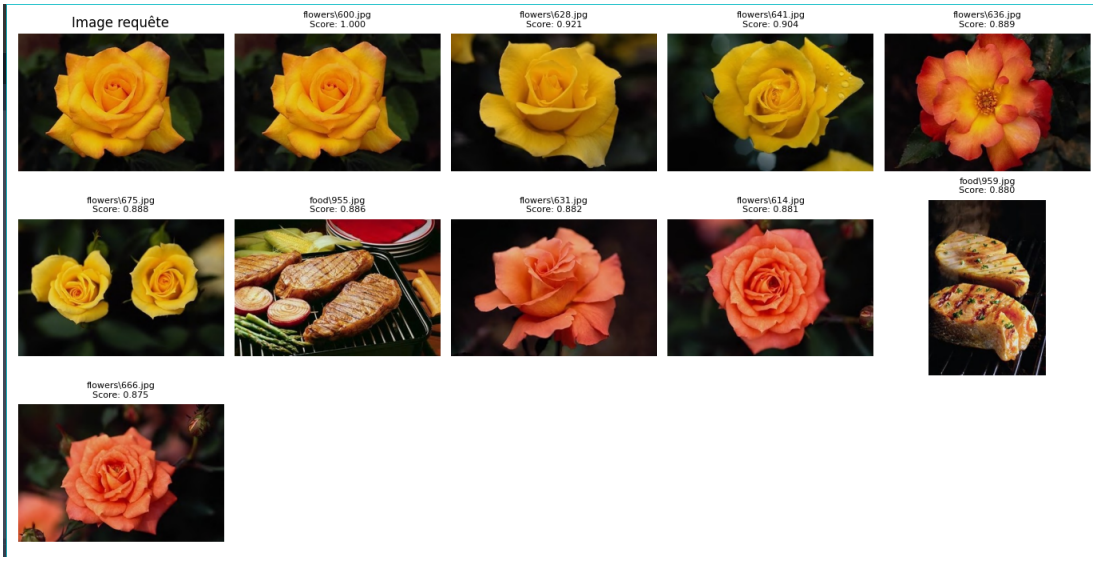

FIGURE 3.2 – Résultats avec quantification à 8 bins - Vitesse optimale, précision limitée

FIGURE 3.3 – Résultats avec quantification à 16 bins - Compromis équilibré performance/précision

FIGURE 3.4 – Résultats avec quantification à 32 bins - Précision élevée, coût computationnel accru

FIGURE 3.5 – Résultats avec quantification à 64 bins - Haute précision, temps de calcul significatif

FIGURE 3.6 – Résultats avec quantification à 256 bins - Précision maximale, charge computationnelle importante

3.2.1 Analyse Détaillée des Niveaux de Quantification

8 bins : Cette configuration offre la plus grande rapidité de traitement mais souffre d'une perte importante d'information. Les nuances colorimétriques subtiles sont perdues, ce qui peut conduire à des confusions entre images distinctes mais partageant des palettes de couleur similaires. Le vecteur descripteur de seulement 24 valeurs (3 canaux $\times$ 8 bins) permet des comparaisons extrêmement rapides.

16 bins : Représente le point d'équilibre optimal entre précision et performance. Cette granularité capture suffisamment d'information pour distinguer les images similaires des images différentes tout en maintenant des temps de calcul raisonnables. Avec 48 valeurs par descripteur, cette configuration convient à la plupart des applications pratiques.

32 bins : Offre une précision notable en capturant des nuances colorimétriques fines. L'augmentation à 96 valeurs par descripteur permet une différenciation plus fine mais entraîne une augmentation mesurable du temps de calcul. Recommandé pour des applications nécessitant une bonne précision.

64 bins : Avec 192 valeurs par descripteur, cette configuration fournit une représentation très détaillée de la distribution des couleurs. La précision est excellente mais le coût computationnel devient significatif, particulièrement pour les grandes bases de données. Adapté aux applications où la précision prime sur la vitesse.

256 bins : Représente l'histogramme complet sans réduction. Chaque descripteur contient 768 valeurs, offrant la représentation la plus fidèle mais au prix d'une charge computationnelle importante. Les temps de comparaison sont considérablement accrus et les bénéfices en précision peuvent être marginaux par rapport à 64 bins.

3.2.2 Compromis Performance/Précision

Nombre de bins	Taille descripteur	Précision	Vitesse	Usage recommandé
8 bins	24 valeurs	2/5	5/5	Applications temps réel
16 bins	48 valeurs	3/5	4/5	Usage général
32 bins	96 valeurs	4/5	3/5	Bonne précision
64 bins	192 valeurs	5/5	2/5	Haute précision
256 bins	768 valeurs	5/5	1/5	Recherche spécialisée

TABLE 3.1 – Compromis entre précision et performance selon le nombre de bins

3.3 Évaluation des Espaces Couleur

FIGURE 3.7 – Performance dans l'espace Lab - Cohérence perceptuelle optimale

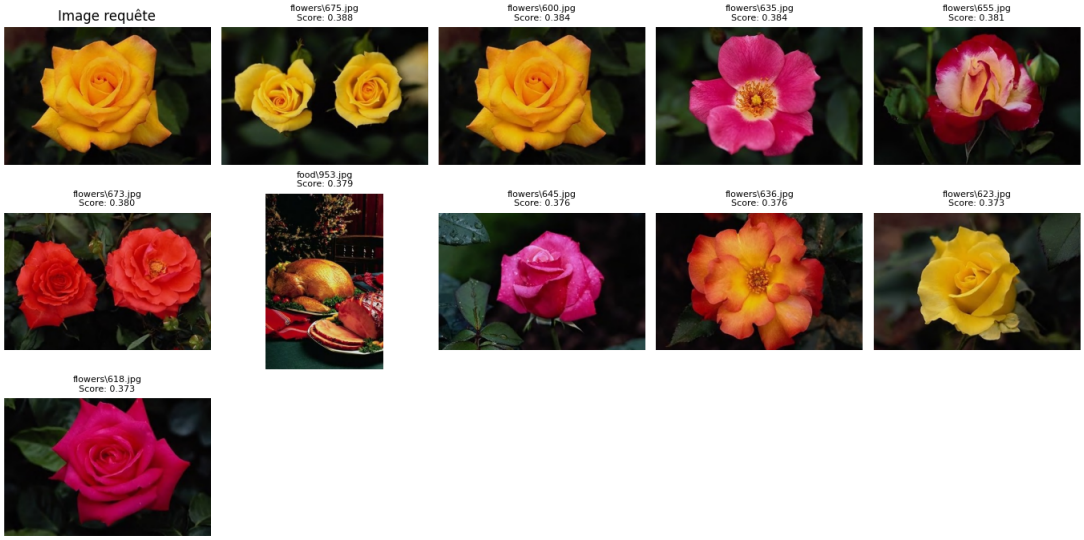

FIGURE 3.8 – Performance dans l'espace HSV - Robustesse aux variations lumineuses

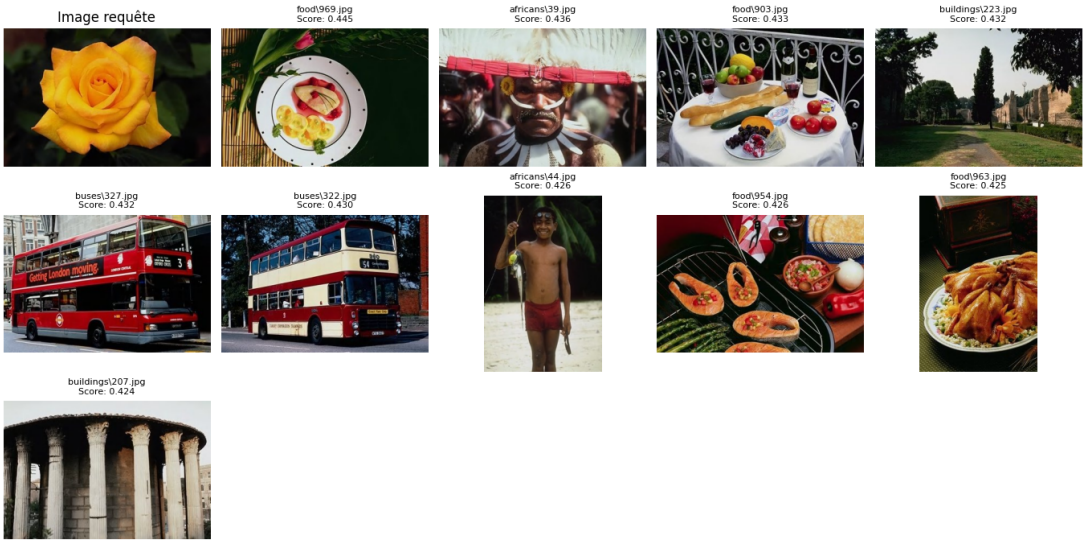

FIGURE 3.9 – Performance dans l'espace RGB - Sensibilité aux conditions d'éclairage

3.3.1 Analyse Comparative

Espace Lab : Son principal avantage réside dans sa propriété de perception uniforme. Les distances calculées dans cet espace correspondent étroitement à la perception humaine des différences de couleur. La séparation de la luminance (composante L) des informations chromatiques (composantes a et b) le rend particulièrement robuste aux variations d'éclairage.

Espace HSV : La séparation explicite de la teinte (Hue) de la saturation et de la valeur rend cet espace efficace pour la recherche par couleur dominante. Il montre une bonne résistance aux changements d'illumination mais peut être affecté par les zones sous-exposées ou sur-exposées.

Espace RGB : Bien qu'intuitif et directement lié à la capture des images, cet espace est fortement couplé aux conditions d'acquisition. Les variations d'éclairage impactent simultanément les trois canaux, ce qui peut réduire la cohérence des résultats.

3.4 Comparaison des Mesures de Similarité

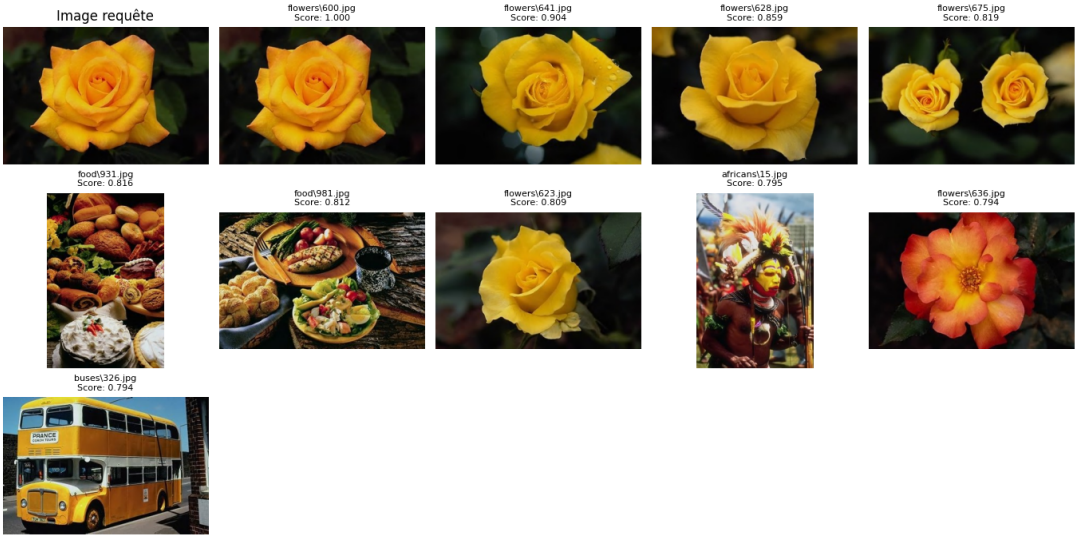

FIGURE 3.10 – Résultats avec intersection d'histogrammes - Rapidité avec risques de faux positifs

FIGURE 3.11 – Résultats avec distance euclidienne - Simplicité mais sensibilité aux échelles

FIGURE 3.12 – Résultats avec distance Chi-carré - Précision optimale pour les nuances subtiles

FIGURE 3.13 – Résultats avec corrélation - Stabilité face aux variations globales

3.4.1 Analyse des Performances

Distance Chi-carré : Excellente performance globale due à sa capacité à pondérer les différences en fonction de l'importance relative des bins. Particulièrement efficace pour distinguer des images ayant des distributions couleur similaires mais présentant des variations locales significatives.

Corrélation : Montre une grande stabilité face aux variations globales de luminosité et de contraste. Sa nature statistique la rend robuste mais peut parfois manquer de sensibilité pour des différences localisées.

Intersection d'histogrammes : Algorithmiquement simple et rapide, mais peut pro-

duire des faux positifs lorsque des images partagent des couleurs dominantes sans similarité sémantique.

Distance euclidienne : Approche directe mais sensible aux normalisations. Peut être affectée par des différences d'échelle entre les histogrammes comparés.

3.5 Configuration Optimale

FIGURE 3.14 – Résultats avec la configuration optimale (Lab + 32 bins + Chi-carré)

La combinaison de l'espace Lab, de 32 bins et de la distance Chi-carré représente le meilleur compromis identifié. Cette configuration maximise la précision tout en maintenant une complexité computationnelle acceptable pour la base Corel-1000.

Chapitre 4

Conclusion

4.1 Synthèse des Contributions

Ce projet a permis de développer un système CBIR complet démontrant l'efficacité des histogrammes de couleur pour la recherche d'images par similarité. Les principaux apports incluent :

- Une implémentation robuste de l'extraction multi-configurations de descripteurs couleur
- Une analyse comparative approfondie des paramètres influençant la performance
- Une interface utilisateur moderne facilitant l'exploration interactive
- L'identification d'une configuration optimale pour la base Corel-1000

4.2 Limites Identifiées

Malgré les performances satisfaisantes, certaines limitations ont été observées :

- **Dépendance exclusive à la couleur** : L'absence de descripteurs de texture et de forme peut conduire à des confusions entre images partageant des palettes de couleurs similaires mais de contenu sémantique différent
- **Sensibilité aux variations géométriques** : Les histogrammes ignorent la distribution spatiale des couleurs, rendant le système sensible aux rotations et changements d'échelle
- **Caractère global des descripteurs** : L'approche par histogrammes ne capture pas les relations locales entre régions de l'image

Chapitre 5

Perspectives

5.1 Améliorations Immédiates

Plusieurs axes d'amélioration peuvent être envisagés à court terme :

- **Descripteurs de texture** : Intégration de méthodes comme LBP (Local Binary Patterns) ou les filtres de Gabor pour capturer les motifs et structures
- **Descripteurs de forme** : Utilisation de moments invariants ou de descripteurs de contours pour améliorer la discrimination
- **Segmentation couleur** : Implémentation d'algorithmes de segmentation pour capturer la distribution spatiale des couleurs

5.2 Évolutions à Moyen Terme

Des développements plus ambitieux pourraient inclure :

- **Apprentissage profond** : Utilisation de réseaux de neurones convolutionnels pour l'extraction automatique de caractéristiques
- **Indexation avancée** : Mise en œuvre de structures de données spatiales (KD-Tree, R-tree) pour accélérer la recherche
- **Recherche hybride** : Combinaison de plusieurs types de descripteurs dans une approche multimodale

5.3 Perspectives à Long Terme

À plus long terme, le système pourrait évoluer vers :

- **Apprentissage interactif** : Intégration de retours utilisateur pour améliorer progressivement les résultats
- **Adaptation contextuelle** : Développement de mécanismes d'adaptation aux caractéristiques spécifiques de différentes bases d'images
- **Intégration sémantique** : Combinaison avec des techniques de vision par ordinateur avancées pour capturer des aspects sémantiques

Ce projet constitue une base solide pour le développement de systèmes de recherche d'images plus intelligents et adaptatifs, ouvrant des perspectives intéressantes pour des applications dans des domaines variés tels que le commerce électronique, la médecine, la sécurité ou la gestion d'archives multimédias.