

Méthodes de représentation et de déconvolution des couleurs en imagerie bio-médicale

Wassim Chikhi

Master 2 Vision et Machine Intelligente – 2025/2026

Les notebooks Python utilisés pour ce TP sont disponibles sur GitHub : <https://github.com/vvazzim/Tp-VMI-Wassim>

1. Objectifs du TP

L'objectif de ce TP est d'explorer de manière pratique la **représentation de la couleur** et la **déconvolution colorimétrique** pour des images naturelles et des images histologiques. Plus précisément, on cherche à :

- comparer les espaces **RGB** et **CIELab** sur des images tests (Normandie, clown, mandrill) ;
- analyser la distribution des couleurs dans le cube RGB sur des images de biologie ;
- mettre en œuvre la **déconvolution couleur** de Ruifrok & Johnston pour séparer les colorants H&E et H-DAB sur des coupes histologiques ;
- étudier l'impact de matrices de déconv *correctes* vs *incorrectes* sur la qualité de la séparation des colorants.

L'idée générale est de passer d'une simple "image couleur jolie à l'écran" à une vision beaucoup plus géométrique : chaque pixel devient un vecteur dans un espace 3D, que l'on peut tourner, projeter et décomposer pour isoler des structures biologiques.

2. Données et outils

2.1. Images utilisées

Les jeux de données proviennent du répertoire pédagogique du cours :

- **normandy.jpg** : image naturelle (mer + campagne), utilisée pour rappeler le rôle des canaux RGB ;
- **clown.png** : image RGB utilisée pour l'étude des histogrammes et du passage en CIELab ;
- **mandrill.tif** : image test classique très colorée, idéale pour comparer RGB et Lab ;
- **sampleHE.tif** : coupe histologique H&E (hématoxyline + éosine) ;
- **ZoomBrownish.tif** : zoom sur une région brunâtre (coloration DAB) ;
- **ytma10_010704_benign1_ccd.tif** : échantillon histologique plus large pour tester la robustesse de la déconv.

2.2. Environnement logiciel

Toutes les manipulations ont été réalisées dans un notebook **Kaggle** (Python 3) avec les bibliothèques :

- **numpy** pour les calculs matriciels,
- **imageio** pour le chargement des images (.jpg, .tif),
- **matplotlib** pour la visualisation et les histogrammes,
- **opencv-python** (cv2) pour les conversions d'espaces couleur (RGB ↔ CIELab).

Les scripts **color.py** et **color_decon.py** fournis avec le TP ont été reformulés en fonctions Python afin d'être exécutés directement dans le notebook.

3. Méthodologie

3.1. Pipeline général

Le TP se décompose en deux grandes parties :

1. Représentation de la couleur :

- split des canaux R, G, B ;
- histogrammes par canal et histogramme couleur global (pour clown/mandrill) ;
- calcul d'une luminance physique $Y = 0,299R + 0,587G + 0,114B$;
- passage en CIELab et affichage séparé des canaux L^* , a^* , b^* .

2. Déconvolution couleur :

- passage des images H&E et H-DAB en densité optique (OD) ;
- application d'une matrice de taches (vecteurs H, E, DAB) ;
- reconstruction de canaux mono-tache par inverse de la matrice ;
- test de matrices de déconv alternatives et analyse visuelle des artefacts.

3.2. Fonctions de base en Python

Les fonctions suivantes, inspirées de Ruifrok & Johnston, sont utilisées pour la déconvolution :

```
def rgb_to_od(rgb):
    rgb = rgb.astype(np.float32)
    rgb[rgb == 0] = 1
    return -np.log(rgb / 255.0)

def od_to_rgb(od):
    rgb = np.exp(-od) * 255.0
    return np.clip(rgb, 0, 255).astype(np.uint8)
```

et

```
def color_deconvolution(img, stain_matrix):
    M = stain_matrix / np.linalg.norm(stain_matrix, axis=0)
    M_inv = np.linalg.inv(M)
    od = rgb_to_od(img).reshape((-1, 3)).T # 3 x N
    C = M_inv @ od # 3 x N
    stains = []
    for i in range(3):
        od_i = np.zeros_like(od)
        od_i[i, :] = C[i, :]
        od_img = od_i.T.reshape(img.shape[0], img.shape[1], 3)
        stains.append(od_to_rgb(od_img))
    return stains
```

Cette implémentation suit la logique de `color_deconvolution` mais reste lisible et réutilisable dans le notebook.

4. Résultats : représentation des couleurs

4.1. Image de Normandie (rappel sur le RGB)

Avant de passer aux images tests classiques (clown, mandrill), une première expérience a été réalisée sur l'image naturelle `normandy.jpg` (paysage mer + campagne), afin d'illustrer la signification réelle des canaux R, G, B.

La Figure 1 présente les trois canaux séparés. Cette décomposition montre que :

- le canal **G** ne représente pas uniquement la végétation : la mer, le ciel et une partie du relief y contribuent fortement ;
- le canal **B** n'est pas "la mer" : on retrouve des structures terrestres qui possèdent elles aussi une composante bleue non négligeable ;
- chaque canal est une *projection spectrale* d'un détecteur, et non une couleur "pure" d'objet. Il est donc impossible d'associer naïvement un objet de la scène à un seul des trois canaux.



FIGURE 1. Image de Normandie : canaux R, G, B en niveaux de gris.

Pour illustrer la vulnérabilité de notre perception, la Figure 2 présente la même image après simple permutation des canaux (G,R,B). Bien que physiquement incorrecte, l'image reste visuellement "cohérente" : les couleurs changent mais les structures demeurent identifiables. Cette expérience reproduit le *Color Display Paradox* présenté en cours.



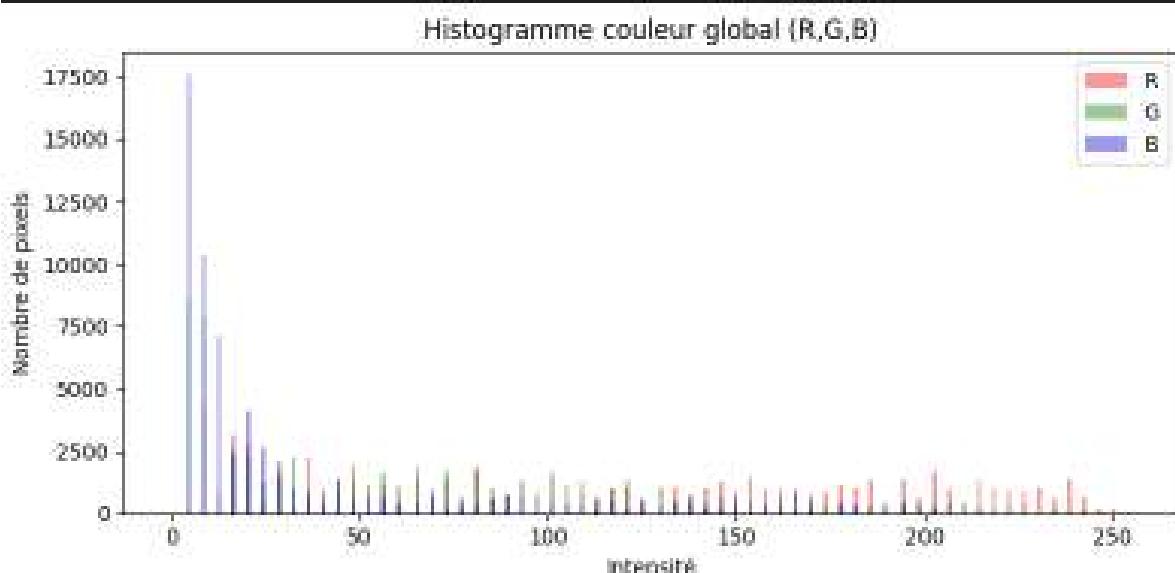
FIGURE 2. Effet du "Color Display Paradox" : permutation des canaux (G,R,B). L'image reste interprétable malgré une correspondance couleur totalement fausse.

Cette expérience met en évidence que la couleur est réellement un vecteur 3D (R,G,B) et non une étiquette sémantique ("bleu = mer", "vert = herbe").

4.2. Image du clown

La Figure 3 montre la décomposition RGB et CIELab de l'image `clown.png`. Les histogrammes des trois canaux RGB et la luminance Y sont rapportés Figure 4.

Décomposition RGB - Image du Clown

**FIGURE 3.** Clown : canaux R, G, B (ligne du haut) et L^* , a^* , b^* (ligne du bas).**FIGURE 4.** Clown : histogramme couleur global (R,G,B) et luminance Y .

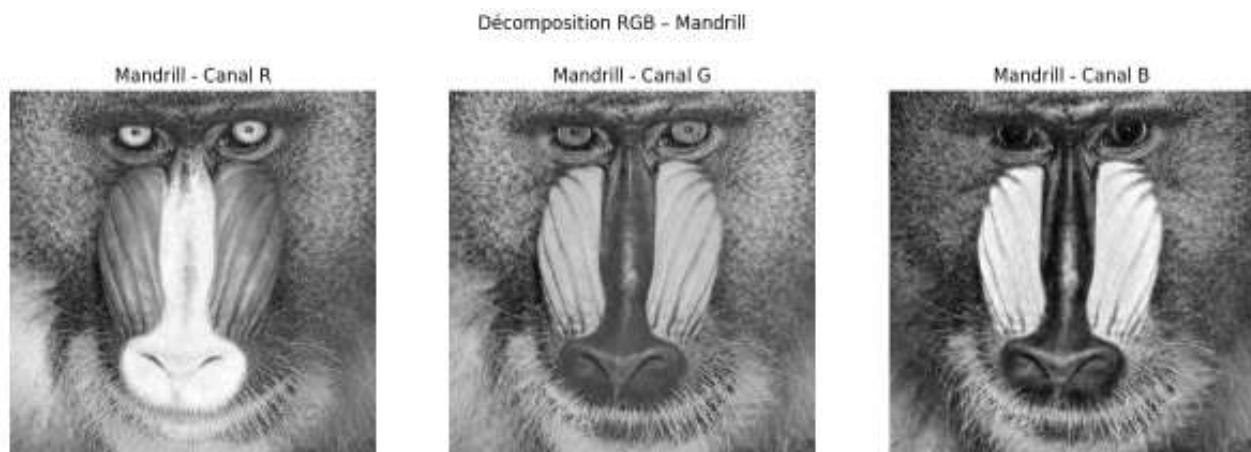
On observe que :

- le canal **B** est concentré sur les basses intensités, ce qui est cohérent avec l'éclairage très chaud de la scène ;
- la luminance Y donne une image en niveaux de gris plus équilibrée que le simple canal rouge, même si les structures principales restent déjà bien visibles dans R ;
- le canal L^* du CIELab isole la structure (clarté/ombre) de façon plus nette, tandis que a^* et b^* portent la chrominance (axes vert–rouge, bleu–jaune).

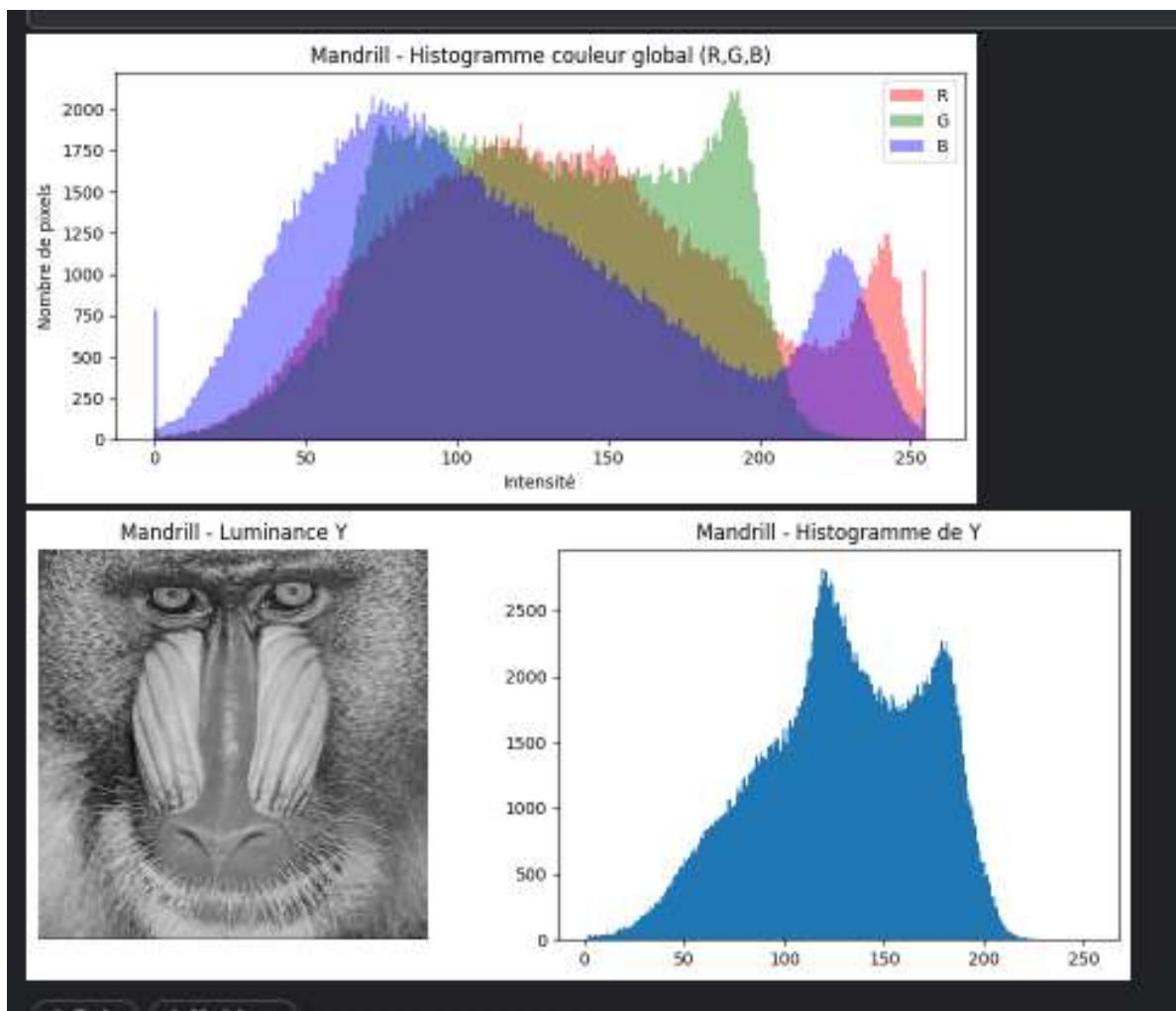
Ce qui m'a surpris ici, c'est à quel point on peut "changer de lunettes" : la même image racontait surtout une histoire de rouge/orange en RGB, et tout à coup en Lab on se retrouve avec un canal purement structurel et deux cartes de couleurs abstraites.

4.3. Image du mandrill

La même procédure appliquée à `mandrill.tif` (Figure 5) met en évidence la saturation extrême de certaines zones (nez, joues, pelage).

**FIGURE 5.** Mandrill : décomposition RGB (haut) et CIELab (bas).

Les histogrammes couleur (Figure 6) sont beaucoup plus étalés que pour le clown, avec des recouvrements importants entre les canaux, ce qui rend la séparation par simple seuillage RGB difficile. En revanche, les canaux a^* et b^* présentent des contrastes nets sur les zones très rouges (nez) et très bleues (bandes latérales), ce qui est intéressant pour une segmentation basée sur les axes de couleur.

**FIGURE 6.** Mandrill : histogramme couleur global (R,G,B) et luminance Y .

5. Résultats : déconvolution couleur

5.1. Déconvolution H&E sur sampleHE.tif

La matrice H&E standard utilisée est :

$$H = (0.650, 0.704, 0.286), \quad E = (0.072, 0.990, 0.105), \quad C = (0.268, 0.570, 0.776).$$

L'application de cette matrice à `sampleHE.tif` donne la Figure 7.

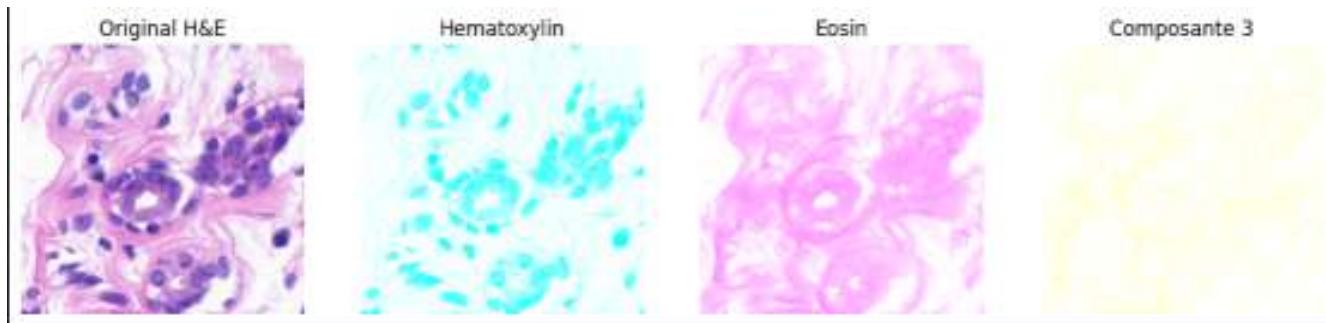


FIGURE 7. Déconvolution H&E sur `sampleHE.tif` : original, canal H, canal E et composante 3.

Les noyaux (Hématoxyline) sont clairement isolés dans le premier canal, tandis que le cytoplasme et la matrice extracellulaire sont concentrés dans le canal Éosine. La troisième composante contient peu d'information structurelle, ce qui confirme que la matrice est bien adaptée à une double coloration H&E.

5.2. Matrices alternatives et impact sur la séparation

Plusieurs matrices de déconv artificielles ont été testées, par exemple :

$$V_1 = (0.7, 0.2, 0.1), \quad V_2 = (0.1, 0.8, 0.1), \quad V_3 = (0.1, 0.2, 0.9).$$

La Figure 8 illustre le résultat de la déconv avec cette matrice (V_1, V_2, V_3).

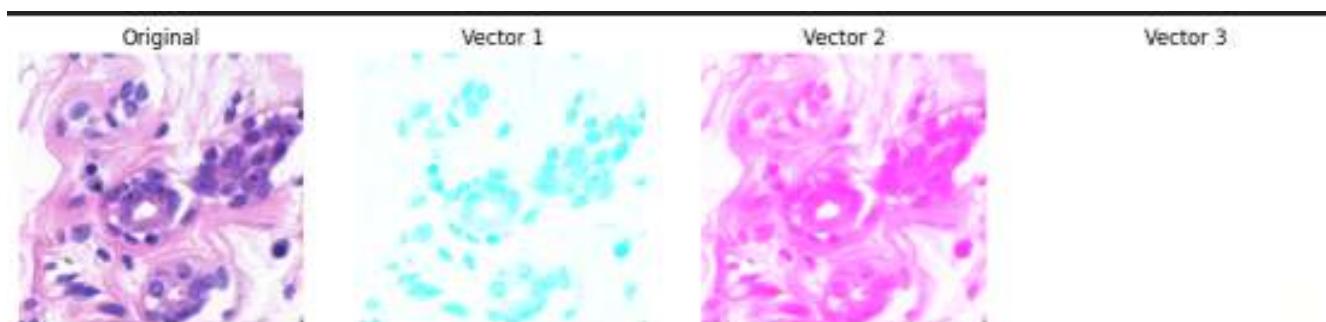


FIGURE 8. Déconvolution H&E avec une matrice de taches artificielle : apparition d'artefacts et de canaux peu interprétables.

On obtient des canaux bruités, mal contrastés, dans lesquels les noyaux et le stroma sont fortement mélangés. Cette expérience montre que la déconv n'est pas un simple "truc numérique" arbitraire : elle repose sur des vecteurs de densité optique qui doivent correspondre aux spectres d'absorption réels des colorants.

5.3. Déconvolution H-DAB sur ZoomBrownish.tif

Pour la coloration H-DAB, on utilise les vecteurs :

$$H = (0.650, 0.704, 0.286), \quad DAB = (0.268, 0.570, 0.776),$$

et un troisième vecteur complémentaire obtenu par produit vectoriel. La Figure 9 montre la séparation obtenue sur `ZoomBrownish.tif`.

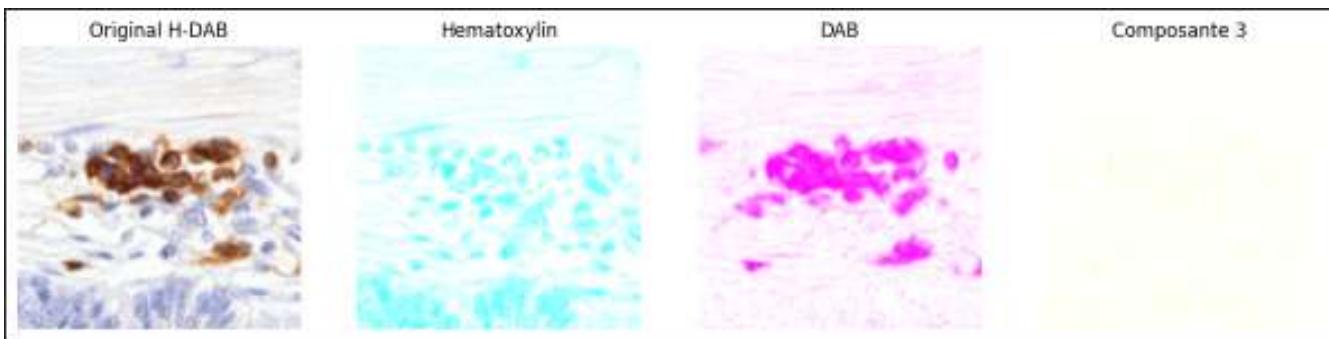


FIGURE 9. Déconvolution H-DAB sur ZoomBrownish.tif : séparation nette entre noyaux (H) et dépôts DAB.

On remarque que le dépôt DAB apparaît macroscopiquement comme une *tache marron*. D'un point de vue physico-numérique, il ne s'agit pas d'une couleur fondamentale : le "marron" est essentiellement un **orange très sombre**, produit par une forte absorption dans le bleu et le vert, combinée à une luminance relativement faible. En RGB, il se traduit par un triplet où le canal rouge domine mais avec une intensité globale basse. En Lab, la composante de luminance L^* est faible, tandis que b^* (axe bleu–jaune) reste nettement positif.

Autrement dit, il n'existe pas de "canal marron" : notre cerveau interprète comme brun une zone orangée plongée dans un contexte plus clair. La déconvolution permet de récupérer le signal DAB sans se laisser piéger par cette perception.

La Figure 10 illustre maintenant une matrice de taches volontairement mal choisie.

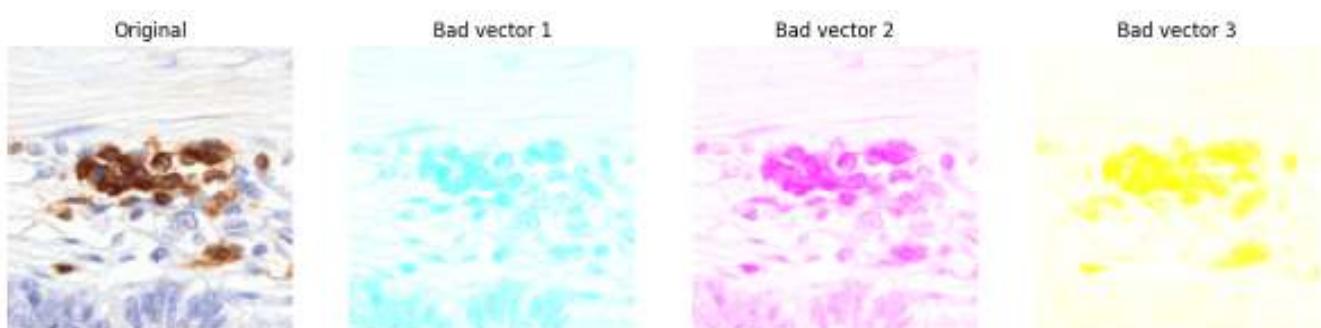


FIGURE 10. Déconvolution H-DAB avec vecteurs inadéquats : forte contamination croisée entre les canaux.

5.4. Application à une image histologique plus large

L'image *ytma10_010704_benign1_ccd.tif* sert de test sur un tissu plus complexe. La Figure 11 montre l'image originale avant déconv.

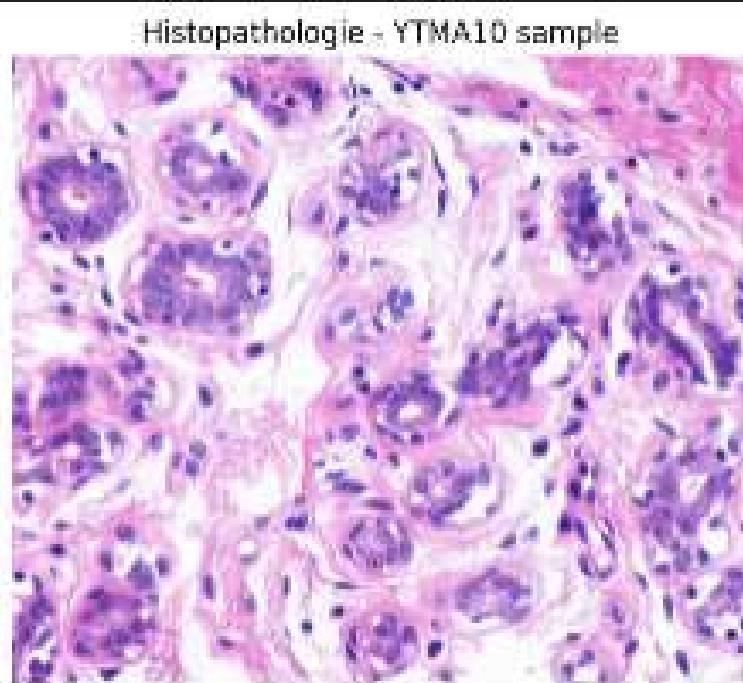


FIGURE 11. Image histologique originale de l'échantillon YTMA10 avant déconvolution H-DAB.

La même matrice H-DAB appliquée à cet échantillon donne la Figure 12. La méthode reste valable sur un échantillon plus complexe, même si quelques artefacts apparaissent en périphérie.

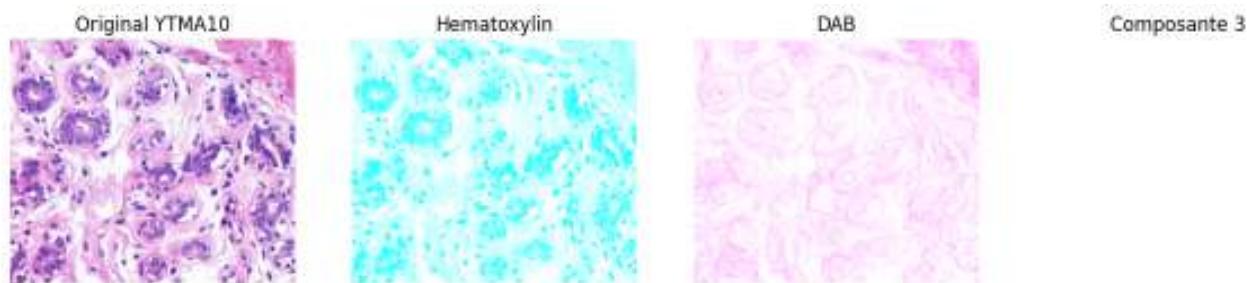


FIGURE 12. Déconvolution H-DAB sur l'échantillon histologique *ytma10_010704_benign1_ccd.tif*.

6. Discussion

Plusieurs points ressortent de ce TP :

- **Espaces couleur** : RGB est idéal pour l'affichage mais peu adapté à l'analyse perceptuelle. L'espace CIELab, en séparant luminance et chrominance, offre une base plus naturelle pour la segmentation ou la normalisation de couleur.
- **Luminance** : utiliser uniquement le canal rouge comme image de luminance fonctionne "à peu près" pour le clown, mais la combinaison $Y(R, G, B)$ et surtout L^* donnent des résultats plus robustes.
- **Perception du marron** : le cas du DAB illustre bien que certaines couleurs courantes (marron, beige) n'existent pas comme "axes fondamentaux" de couleur. Ce sont des combinaisons d'orange/jaune avec une luminance réduite. C'est une bonne piqûre de rappel sur la différence entre *ce que l'on voit* et *comment les couleurs sont codées* dans les espaces RGB ou Lab.
- **Déconvolution** : la séparation H&E ou H-DAB est très sensible au choix des vecteurs de tache. Des vecteurs mal calibrés conduisent à une forte contamination croisée et à des canaux peu interprétables.
- **Généralisation** : la même matrice H-DAB appliquée à un autre tissu (YTMA10) fonctionne globalement bien, mais montre aussi les limites d'une approche "tout terrain" : en pratique, les vecteurs devraient être ré-estimés à partir d'échantillons mono-tache.

En résumé, ce TP m'a surtout appris qu'une image couleur de pathologie ne se résume pas à "du rose et du violet" : derrière, il y a une structure mathématique très rigide, qui permet de séparer proprement les colorants si on respecte la physique, et qui s'effondre dès qu'on triche sur les vecteurs.

7. Conclusion

Ce TP met en place une chaîne complète d'analyse couleur pour l'imagerie bio-médicale :

1. étude des canaux RGB, de la luminance et du passage vers CIELab sur des images tests ;
2. exploration des distributions de couleurs en 3D (cube RGB) sur des images de biologie ;
3. implémentation en Python de la déconvolution couleur selon Ruifrok & Johnston ;
4. application à des images H&E et H-DAB réelles, avec comparaison de matrices standard et de matrices expérimentales.

Au final, je retiens que la couleur n'est pas juste une décoration visuelle mais un signal riche qu'on peut re-paramétriser, décomposer et quantifier. Ce changement de point de vue est exactement ce qui fait passer d'un simple "regard de biologiste" à un "regard de vision par ordinateur" sur les mêmes données.

Annexe A – Liste des figures et fichiers

Figure	Fichier	Description
1	normandy_rgb_channels.png	Normandie : canaux R/G/B en niveaux de gris.
2	normandy_paradox_rgbswap.png	Normandie : RGB normal vs permutation des canaux (Color Display Paradox).
3	clown_rgb_lab.png	Clown : R,G,B et L^*, a^*, b^* .
4	clown_hist_rgb_y.png	Clown : histogramme R/G/B et luminance Y .
5	mandrill_rgb_lab.png	Mandrill : R,G,B et L^*, a^*, b^* .
6	mandrill_hist_rgb_y.png	Mandrill : histogramme R/G/B et luminance Y .
7	sampleHE_deconv_HE.png	Déconvolution H&E : original + H + E + composante 3.
8	sampleHE_deconv_bad.png	Déconv H&E avec matrice artificielle (mauvaise séparation).
9	ZoomBrownish_deconv_HDAB.png	Déconvolution H-DAB sur ZoomBrownish.
10	ZoomBrownish_deconv_bad.png	Déconv H-DAB avec vecteurs inadéquats.
11	ytma10_original.png	Image originale de l'échantillon YTMA10.
12	ytma10_deconv_HDAB.png	Déconvolution H-DAB sur l'échantillon YTMA10.

TABLE 1. Récapitulatif des figures et des fichiers d'images associés.