

Création d'un générateur d'images de peignage moléculaire

Problématique:

1. Besoin d'une grande base de données des images de scanner et de microscope pour l'apprentissage d'un modèle d'IA.
2. L'annotation des images réelles demande du temps et de l'effort, et n'est pas précise.

Objectif:

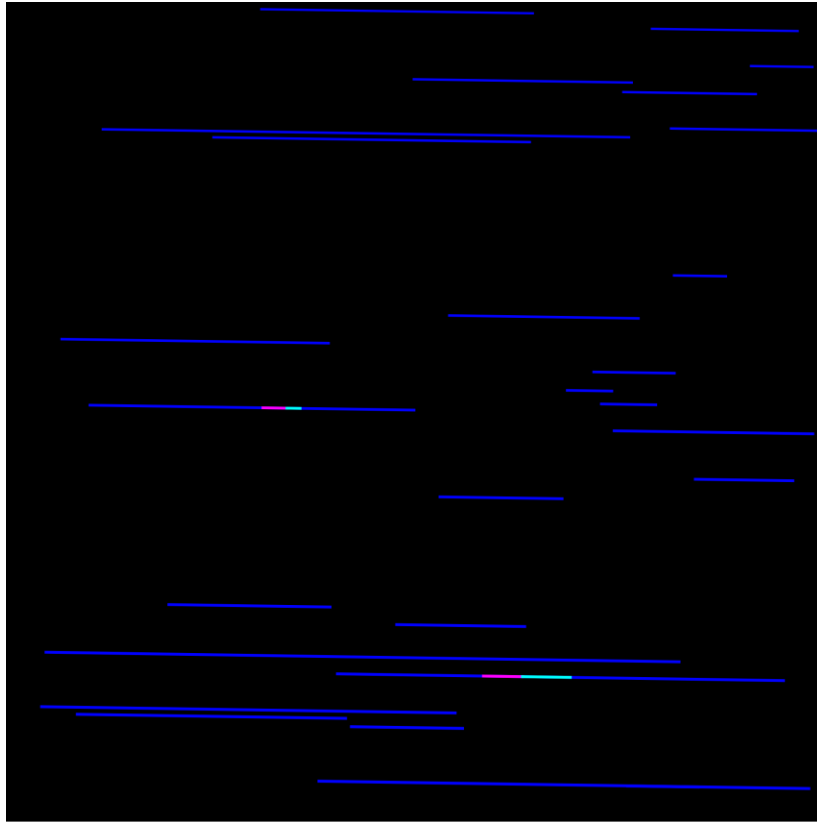
Créer une grande base de données des images simulées de peignage moléculaire, ayant les mêmes caractéristiques que les images réelles faites par le microscope ou le scanner de lames.

Plan:

1. créer des images parfaites avec des fibres d'ADN et des analogues de nucléotide, ces images seront utilisées comme masque ou sortie pour le modèle d'IA.
2. Ajouter tous les types de types de bruit pouvant exister sur les images réelles (Bruit électronique, poussière, bruit biologique...) à travers l'analyse des images de microscope et de scanner.

Générateur des images parfaites:

1. Créer des images RGB contenant des fibres d'ADN (lignes bleues):
 - i. Nombre des fibres dans chaque image et leurs coordonnées sont choisis aléatoirement sous conditions;
 - ii. Les pentes des lignes fluctuent autour d'une moyenne paramétrée.
2. Dessiner les analogues sur certaines fibres choisis aléatoirement (lignes rouges et verts collées sur les fibres bleues)
3. Les paramètres du générateur sont modifiables par l'utilisateur à travers un fichier de configuration.
4. Enregistrer les images dans un dossier afin de les utiliser comme masque pour le modèle de la segmentation.
5. Enregistrer les coordonnées des fibres dans un fichier csv pour les utiliser dans la partie 'Ajout de bruit'.



Générateur de bruit:

Après analyse des images de microscope et de scanner, les types de bruit que l'on peut ajouter aux images simulés sont les suivants:

1. Bruit biologique --> (fonction: Add_Biologic_Noise)

a) Le perlage: créer des discontinuités sur les fibres à partir des coordonnées des fibres.

- i. La probabilité d'avoir du perlage sur une fibre est choisie aléatoirement à partir des valeurs min et max saisi par l'utilisateur sur le fichier de configuration --> (min_Prob_perlage, max_Prob_perlage)
- ii. le nombre de perlage par fibre est proportionnel au nombre de pixel --> min_N_pixels_perlage: le nombre de pixel qui peuvent avoir du perlage.
- iii. la taille du perlage est choisie aléatoirement entre des plages de valeurs. --> max_lenght_perlage

b) Morceaux de fibre: Ajout des petits morceaux des fibres d'ADN et des analogues qui ne sont pas collés aux fibres.

c) Fibre en forme de U: dessiner un demi-cercle entre deux fibres choisis depuis les coordonnées des fibres enregistrés dans le fichier csv. --> Draw_curves

d) Courbes entre certaines fibres: dessiner une courbes de bézier entre deux fibres choisis depuis les coordonnées des fibres enregistrés dans le fichier csv. --> Draw_curves

e) Poussière --> dessiner des petits cube de taille et de couleurs différentes.

2. Bruit Numérique: --> (fonction: Add_Electronic_Noise)

Le bruit numérique représente toute fluctuation parasite ou dégradation que subit l'image à l'instant de son acquisition jusqu'à son enregistrement.

Les facteurs qui limitent la qualité des images du microscope et du scanner sont essentiellement :

a) Bruit électronique: lié à l'appareil --> modéliser par un bruit gaussien (sigma variable)

b) Bruit de photons: lié à l'émission de photons --> loi de poisson

c) Bruit de fond: présence de photons parasites dans le système ou changement du temps d'exposition --> ajouter une constante choisi aléatoirement à chaque canal de l'image.

d) le flou: PSF

Electronic_noise_functions --> fonction : degraded_fibers

Dégrader les fibres et les analogues sur chaque channel en ajoutant du bruit gaussien et en remplissant les valeurs inférieure d'une valeur données par 0 : appliquer mask $<255 == 0$

Electronic_noise_functions --> fonction : Add_channel_noise

Ajoute chaque type de bruit avec des paramètres variantes sur chaque canal:

- i. la quantité de bruit de fond (de la lumière parasite et d'autres sources ou le temps de pose): Parasites_ch
- ii. Salt: Ajouter des pixels de valeurs et tailles variables sur chaque canal avec des quantités différentes selon le canal dominant (avec plus de bruit).
- iii. Bruit gaussien
- iv. Blur

Electronic_noise_functions --> fonction : Add_PSF_to_channel

Générer une distribution gaussienne d'intensité:

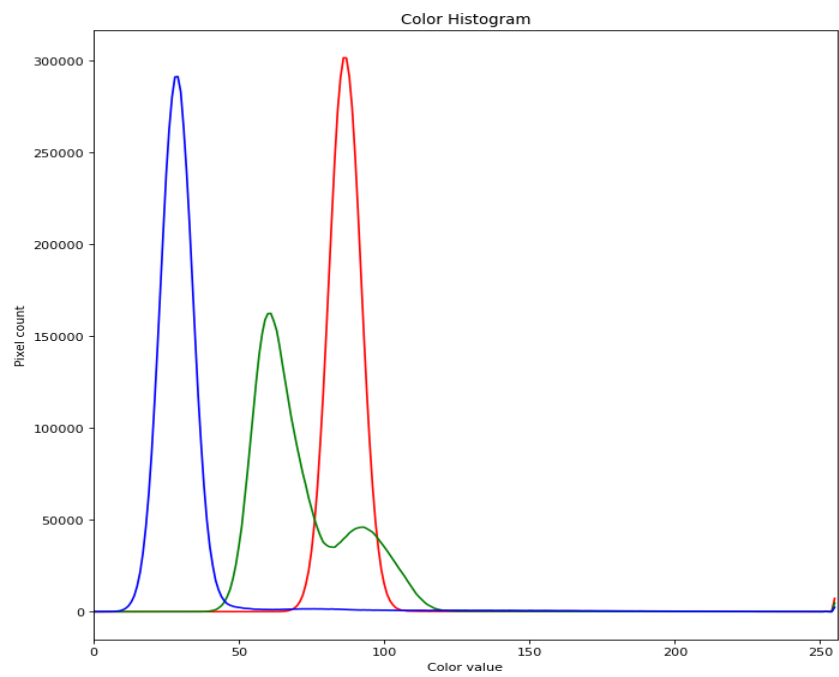
$$gauss_function = a * np.exp(-(x-m)**2/(2*s**2))$$

Electronic_noise_functions --> fonction : get_gradient_3d

Changer l'intensité sur certaine zone de l'image horizontalement et verticalement.



Image du générateur après ajout du bruit



Histogramme de l'image