

Vision artificielle neuromimétique et localisation de cible visuelle

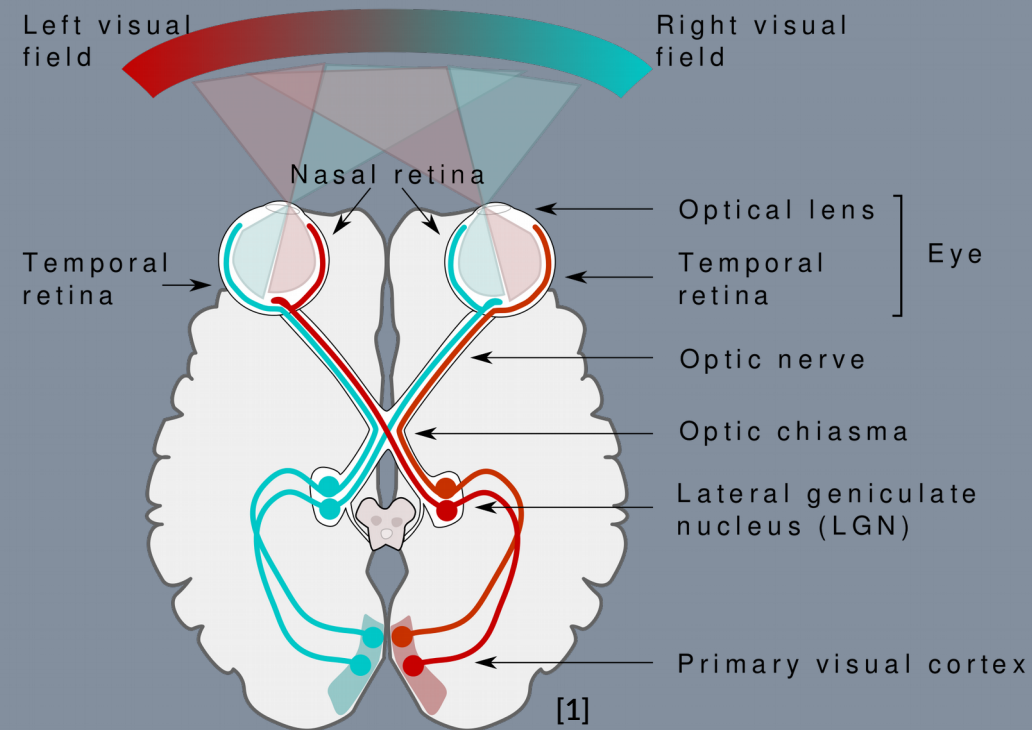
ALBIGES Pierre (M2 NIC)

Superviseurs:

DAUCE Emmanuel (INS) – PERRINET Laurent (INT)



Vision naturelle?



→ Appréhender à la fois rapidement et avec précision notre environnement

Vision naturelle ?

Vitesse & Précision

→ Acuité visuelle variable

+

→ Saccades oculaires

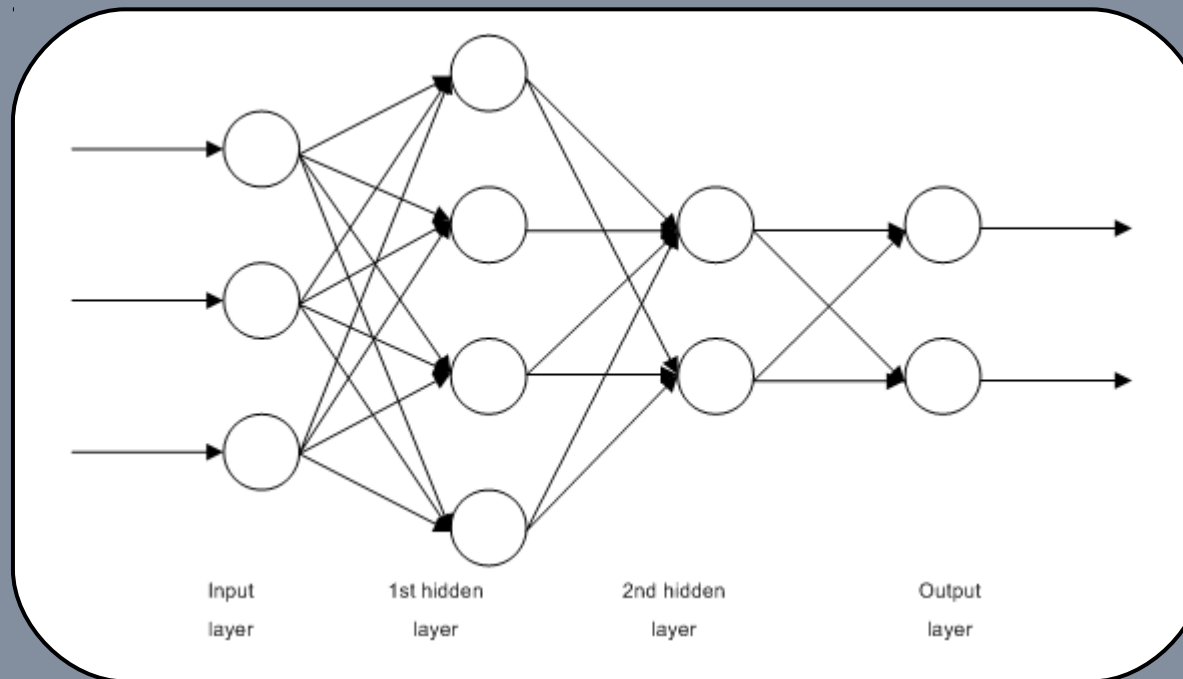
Vision artificielle ?

→ Permettre à des agents artificiels de percevoir, analyser et appréhender des images et des vidéos

Vision artificielle neuromimétique ?

→ Créer un modèle de vision artificielle s'inspirant de l'acuité variable et des saccades oculaires biologiques pour percevoir et décrire son environnement

Machine learning ?



[2]

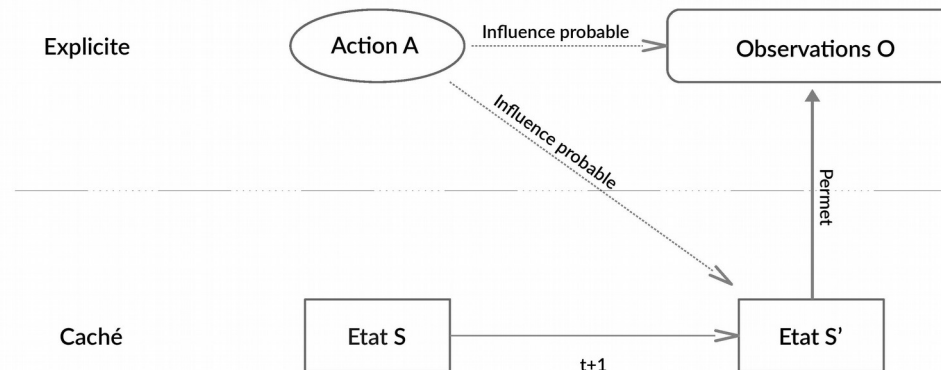
Implémentation

Stage A

Stage B

→ Python + TensorFlow

→ Processus de décision markoviens

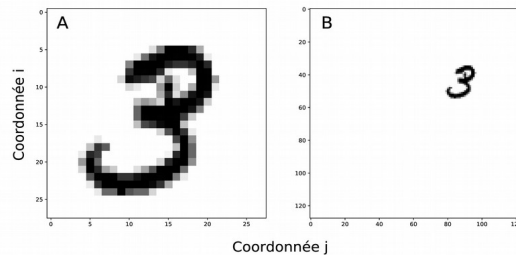


Implémentation

Stage A

Stage B

→ Base de données MNIST



→ Objectif de l'agent : détecter les coordonnées de la cible dans l'espace

Implémentation

Stage A

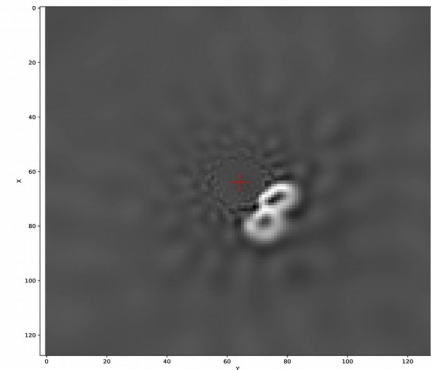
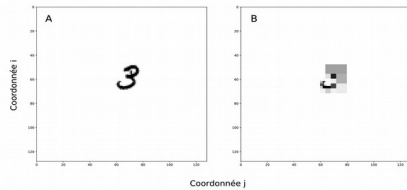
Stage B

Filtre rétinien ?

Wavelets

ou

LogPolaire



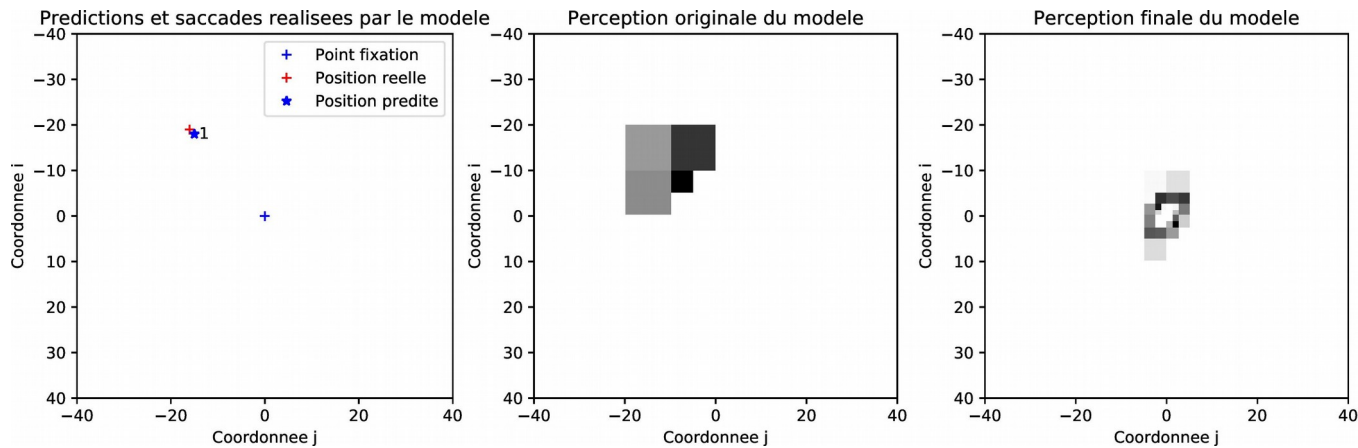
Entrée du réseau

Comportement du modèle

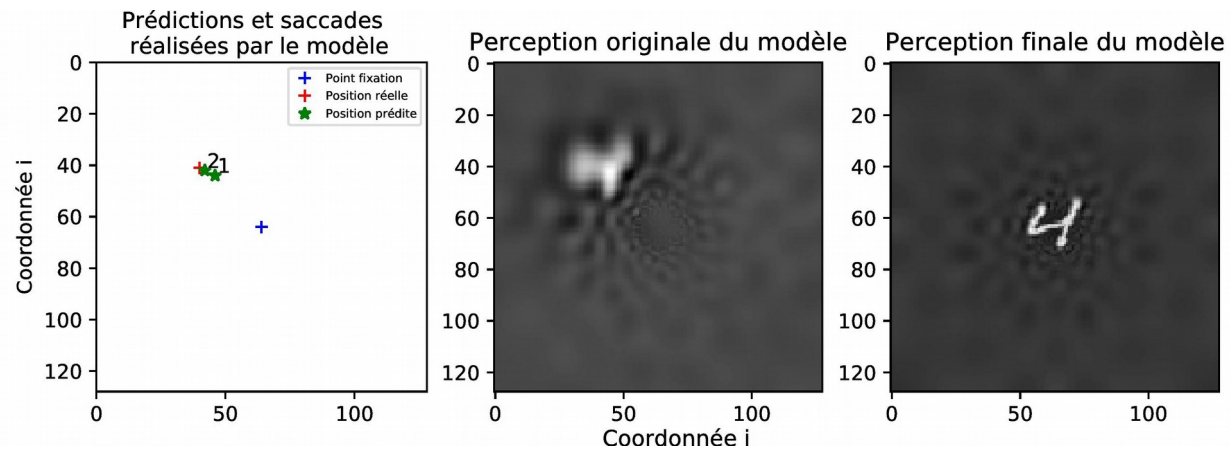
Stage A

Stage B

Wavelets



LogPolaire



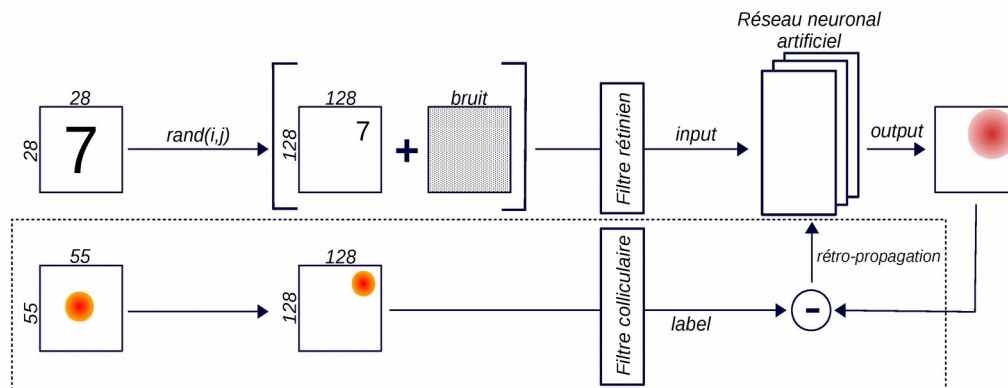
Implémentation

Stage A

Stage B

→ Python + ~~TensorFlow~~ Torch

→ Filtres LogPolaire + ~~Wavelets~~



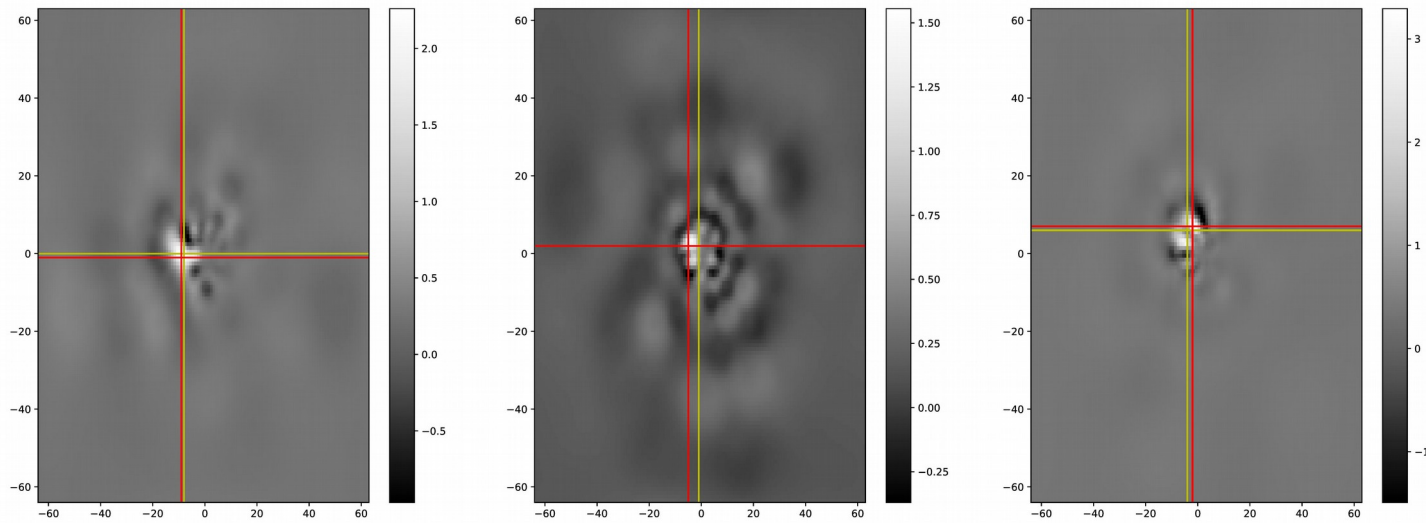
→ Objectif de l'agent : prédire, dans l'espace, la probabilité de la présence de la cible

Comportement du modèle

Stage A

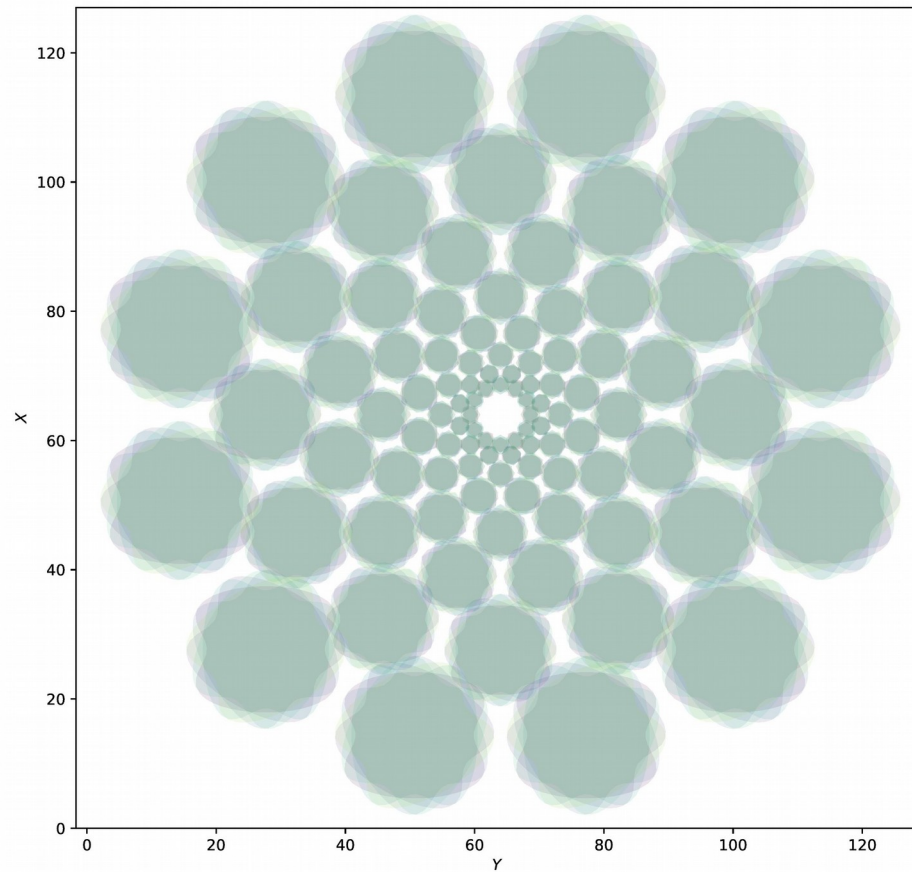
Stage B

Prédictions du modèle

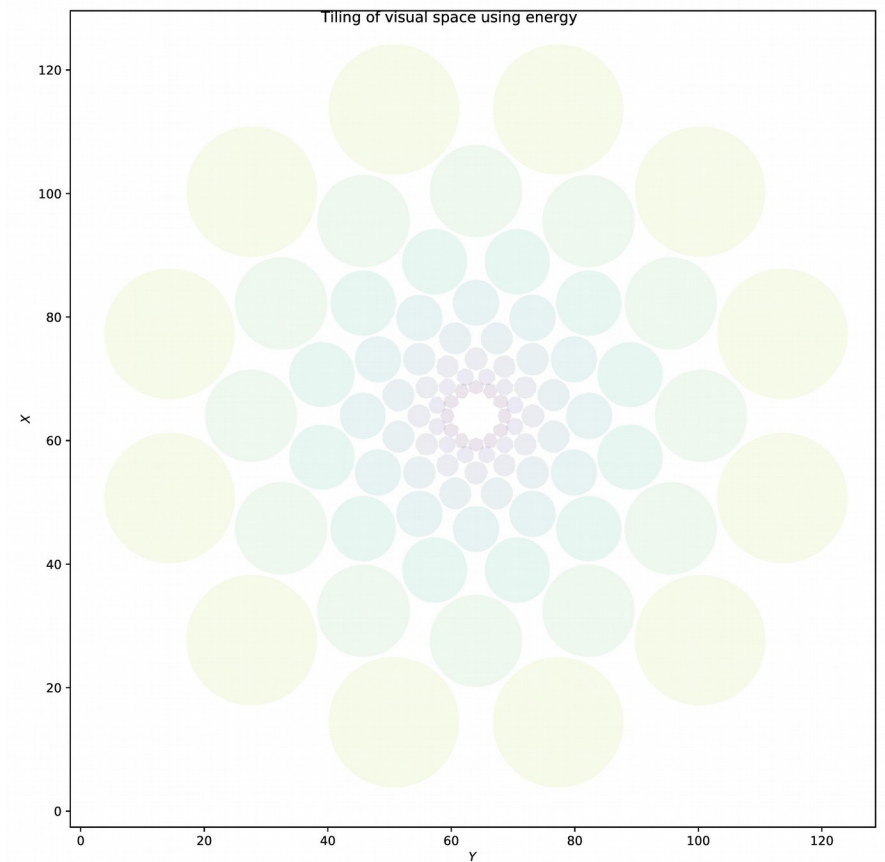


- Quantification les prédictions et leur robustesse
- Intégration du comportement saccadique
- Intégration d'une classification après avoir détecté la position de la cible (modèle « detect and identify »)
- Intégration d'une forme de mémoire et d'une inhibition de retour
- Intégration d'une influence top-down (sélection attentionnelle)
- Insertion du modèle dans un agent physique (caméra, drone, robot)

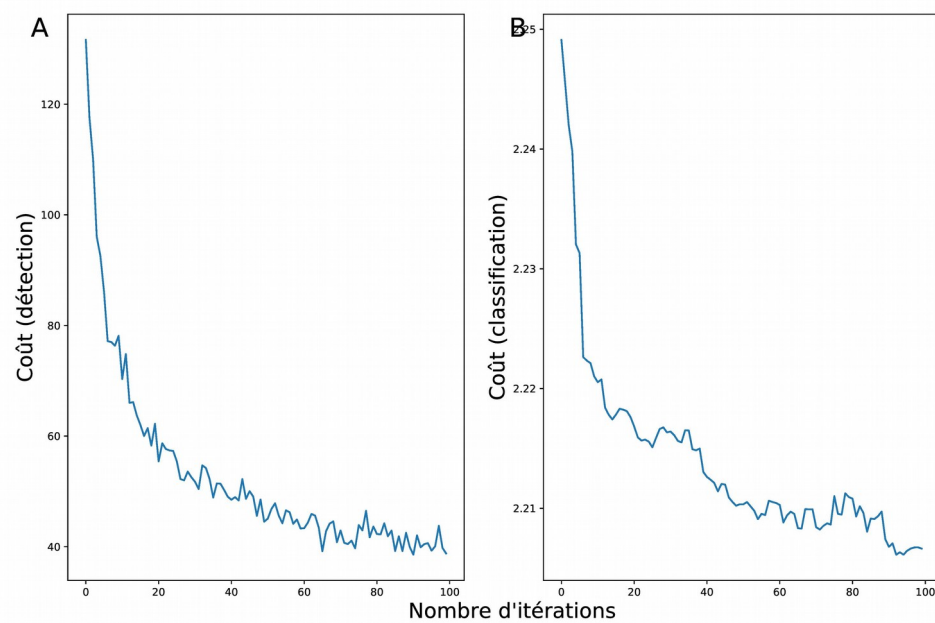
Filtre rétinien



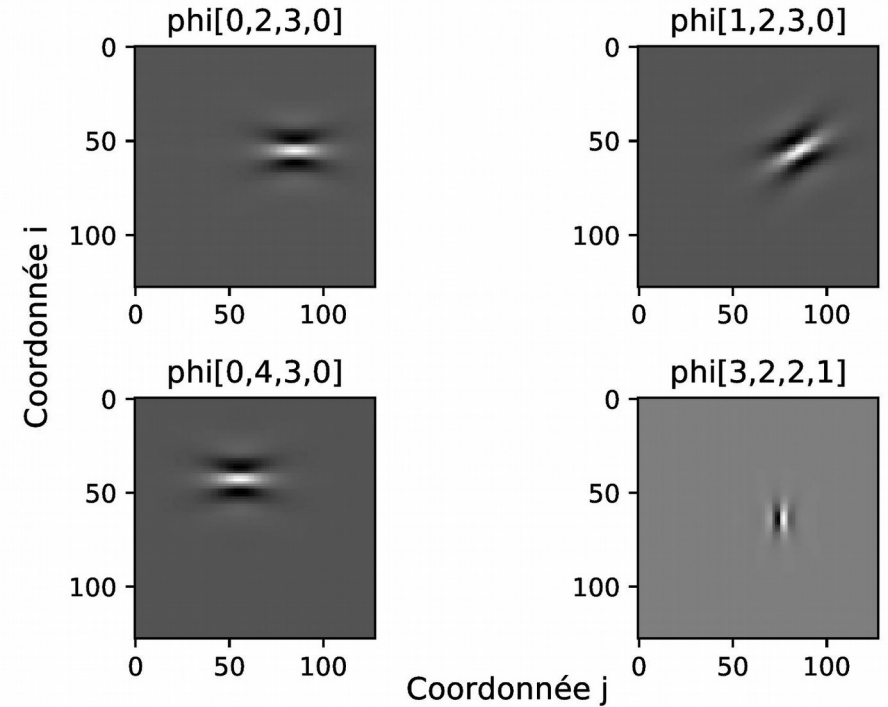
Filtre colliculaire



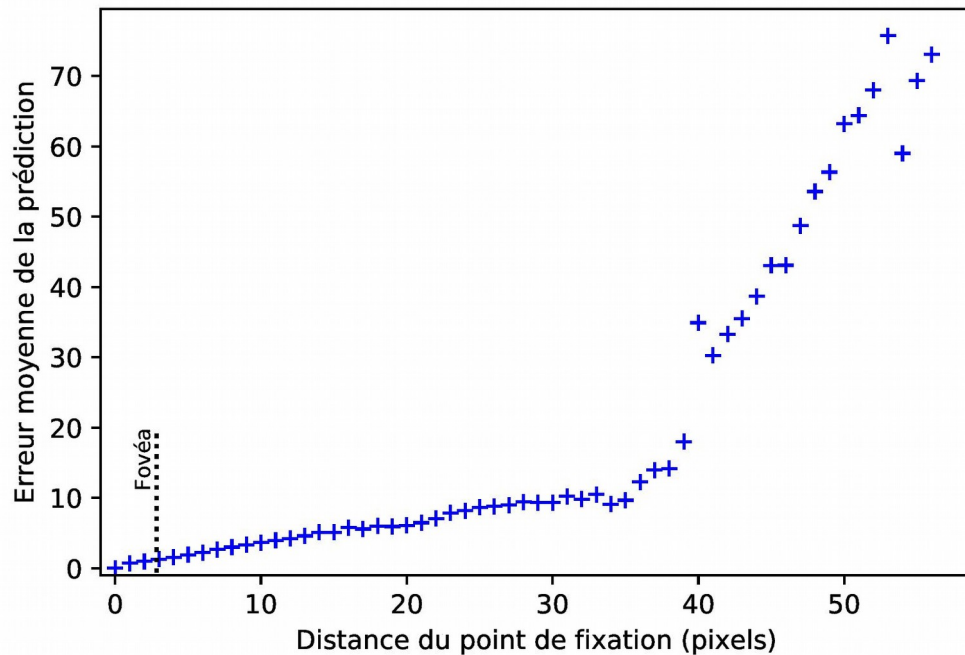
Apprentissage



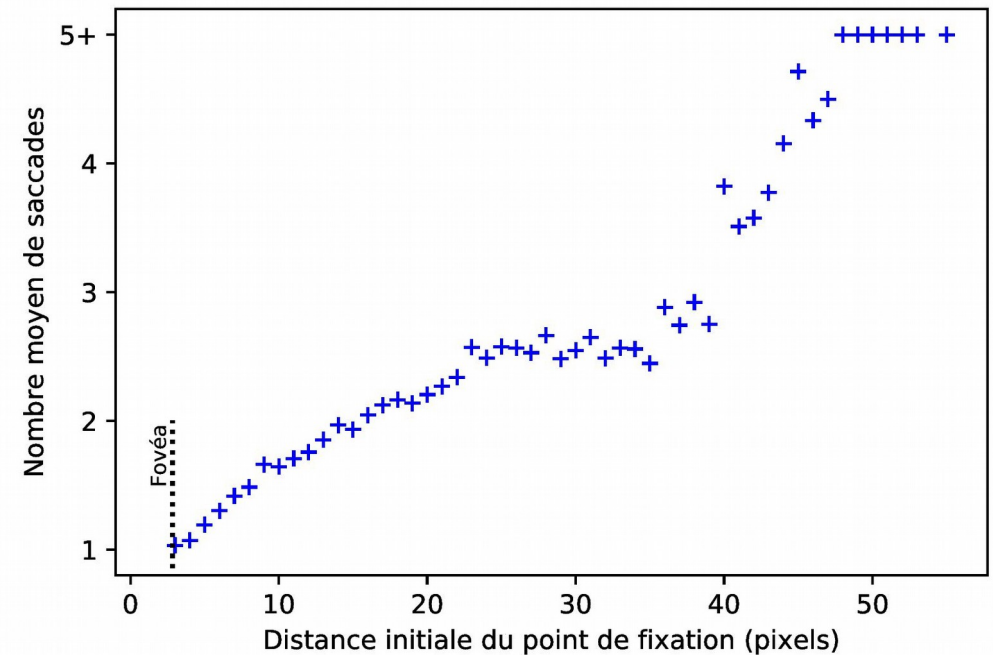
Filtres Gabor



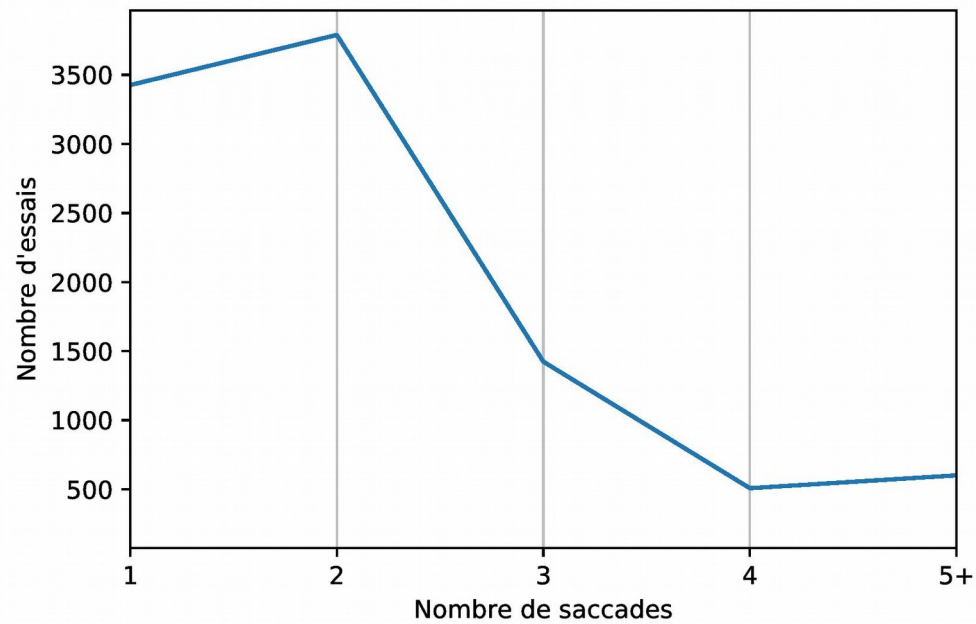
Rapport erreur/distance



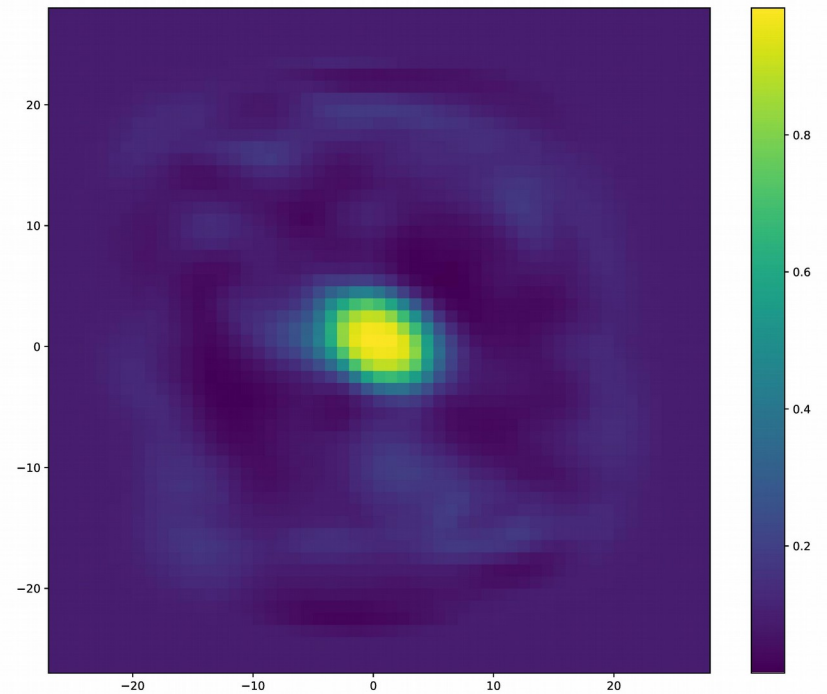
Rapport nombre saccades/distance initiale



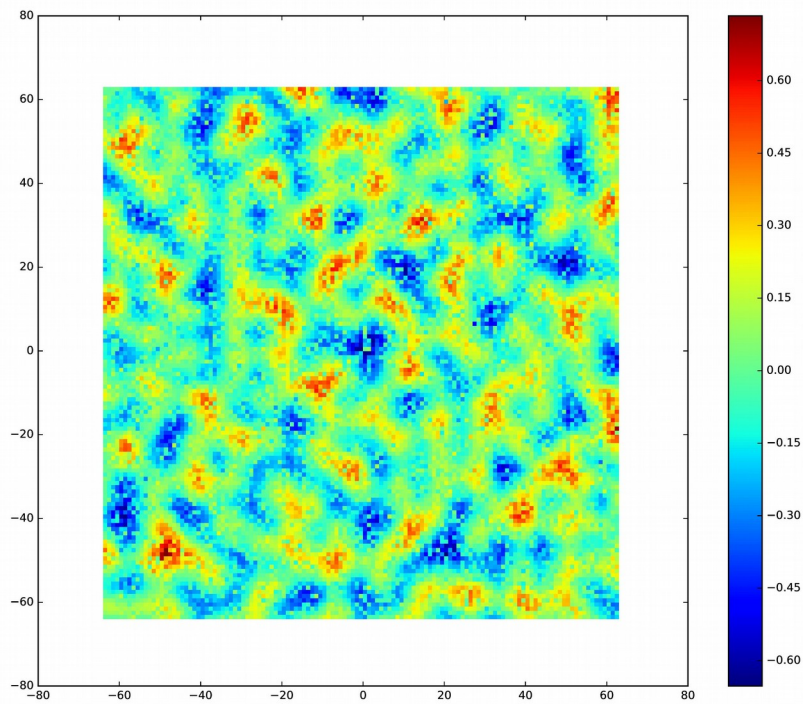
Quantification nombre de saccades



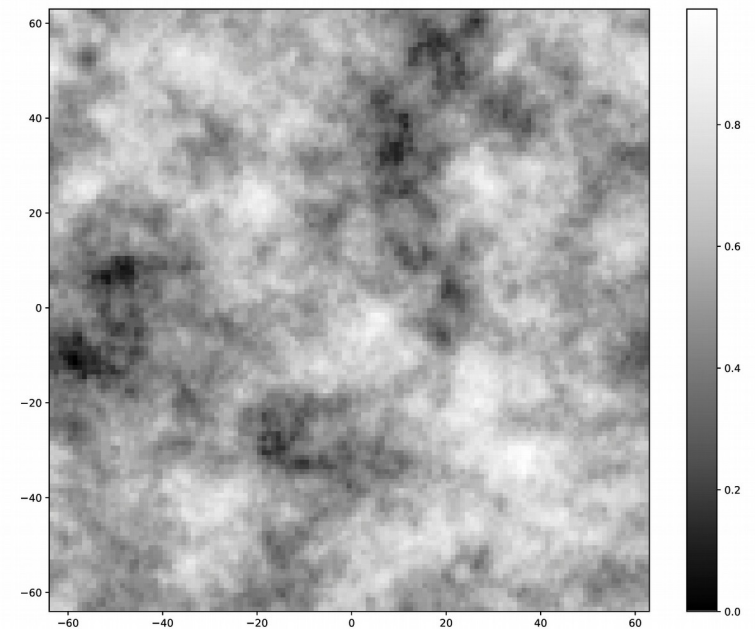
Carte de certitude (label)



Bruit Perlin



Bruit MotionCloud



Sources

[1] https://upload.wikimedia.org/wikipedia/commons/thumb/b/bf/Human_visual_pathway.svg/2000px-Human_visual_pathway.svg.png

[2] https://upload.wikimedia.org/wikipedia/commons/3/30/Multilayer_Neural_Network.png

L'ensemble des figures, scripts et notes relevant du travail présenté sont disponibles à l'adresse suivante : <https://github.com/pierreallbige/ActiveVision>