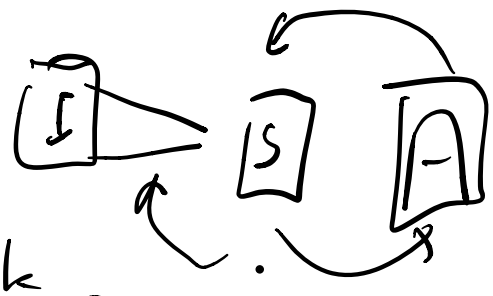
Learning where to see

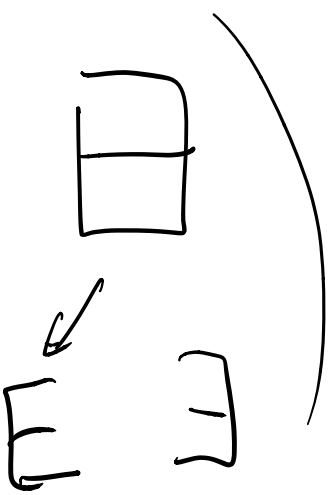
① Intro

A. Deep learning = feed-forward alone que la perception est active.

Bottleneck Deep Gaze □ → □



cf. transformer network



B. State of the art

→ Najemnik

. Priebe

· papier Frontiers + friction → lien sur les coûts

Attention network

→ identify gap.

c) outline

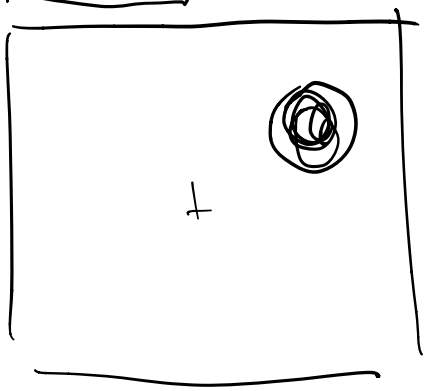
→ on se focalise sur l'invariance à la position u

$$P(i \mid I, x) = P(y \mid I) \times P(i \mid x)$$

indépendance de la classe y par rapport

$$p(y, u \mid \tilde{x}) \sim p(y \mid \tilde{x})\, p(u \mid \tilde{x})$$

$$\int p(u \mid \tilde{x}, y)\, dy = p(u \mid \tilde{x})$$



modèle génératif

$\tilde{x} = T_\rho(x)$

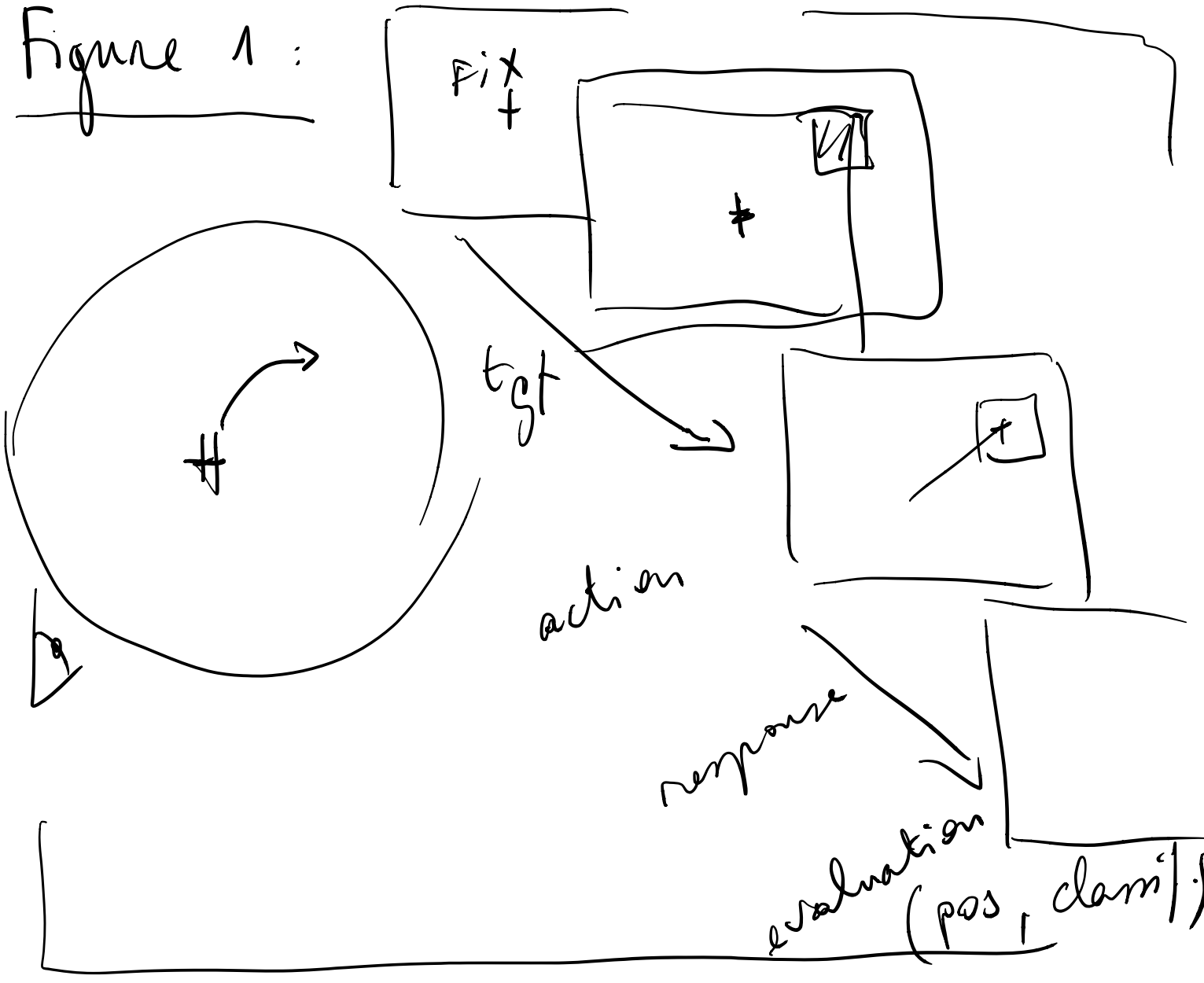
↑ observé ↑ caché

supervision $p(y \mid x)$ → classifieur

$p(y \mid T_u(x), u)$ → accuracy

. Yaka → outline -

Figure 1 :

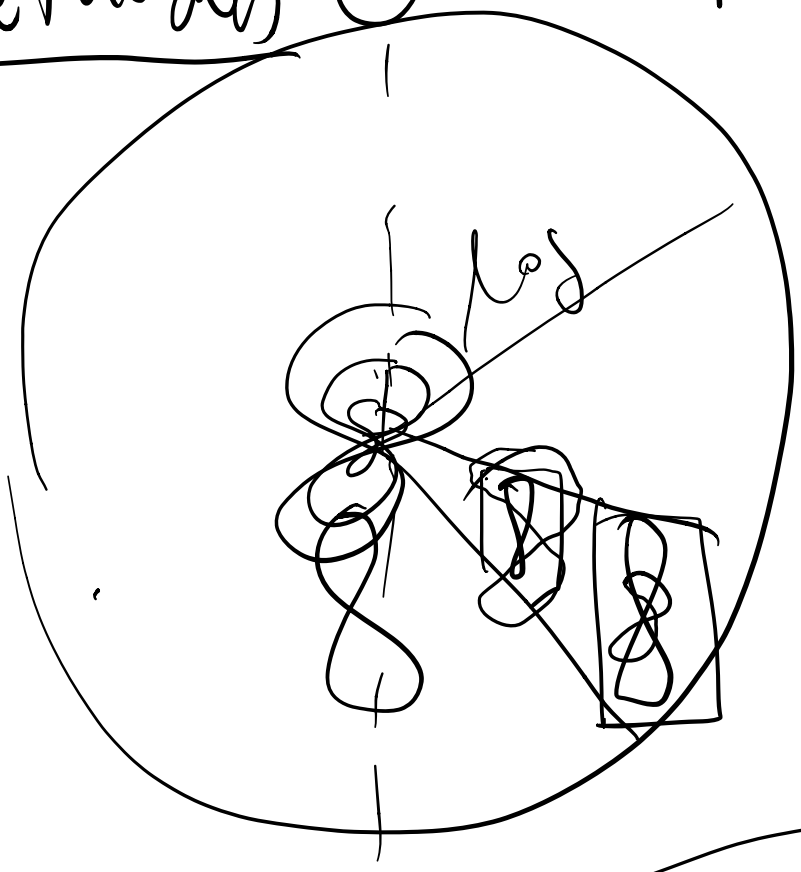


Problem setting :

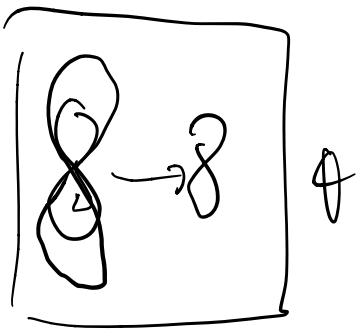
TODO → Apprendre la carte d'accumulation avec du bruit ?

① Méthodes ⓐ Transfo visuelle

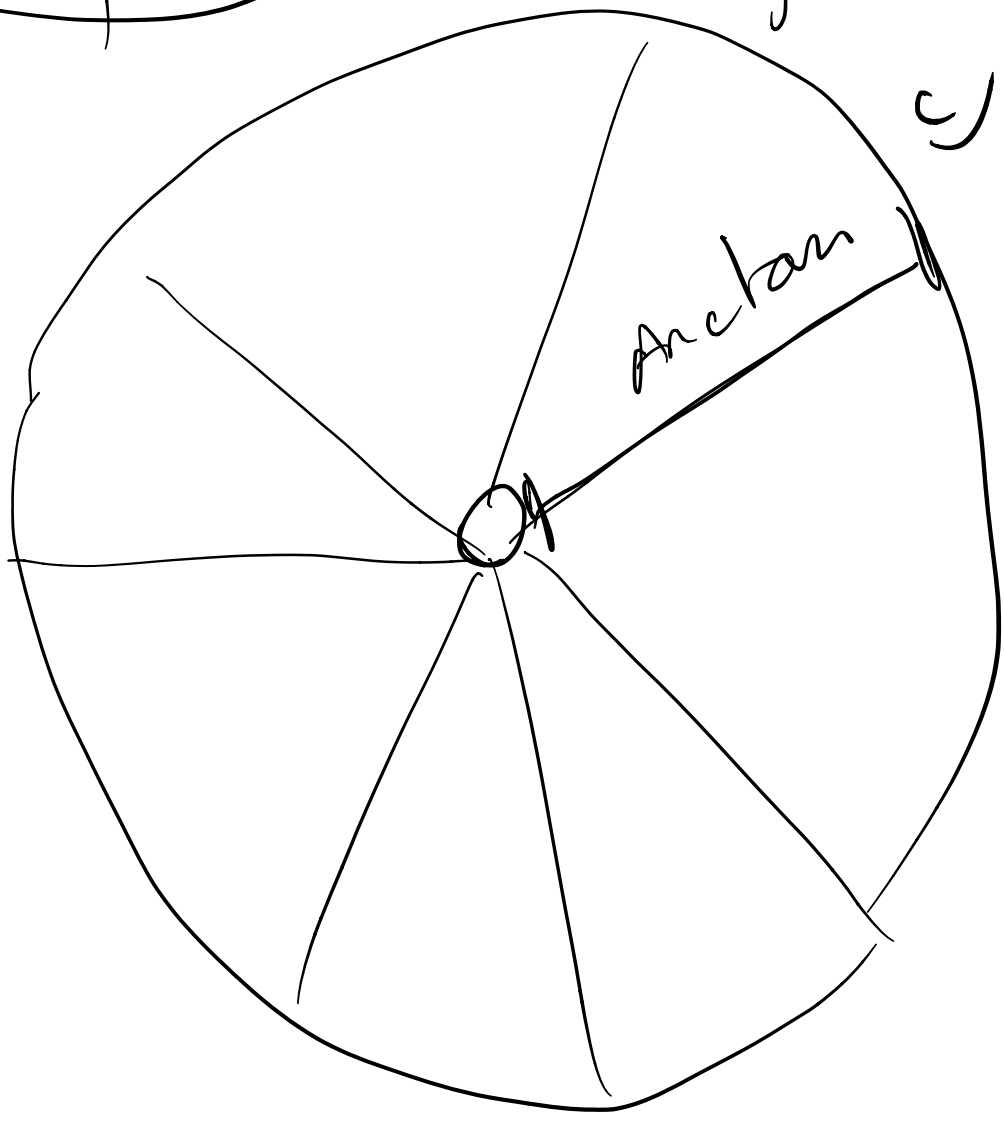
a) Wavelet

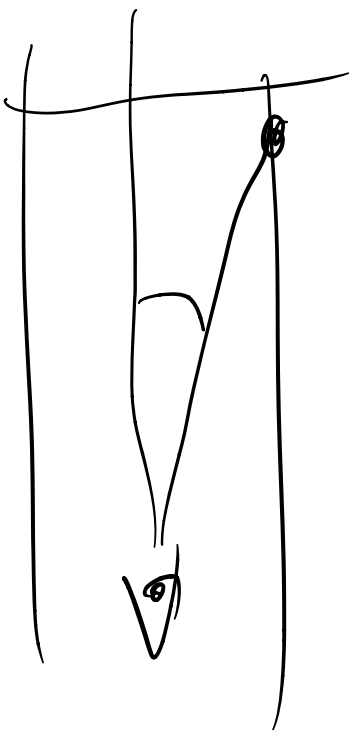


b)



c)





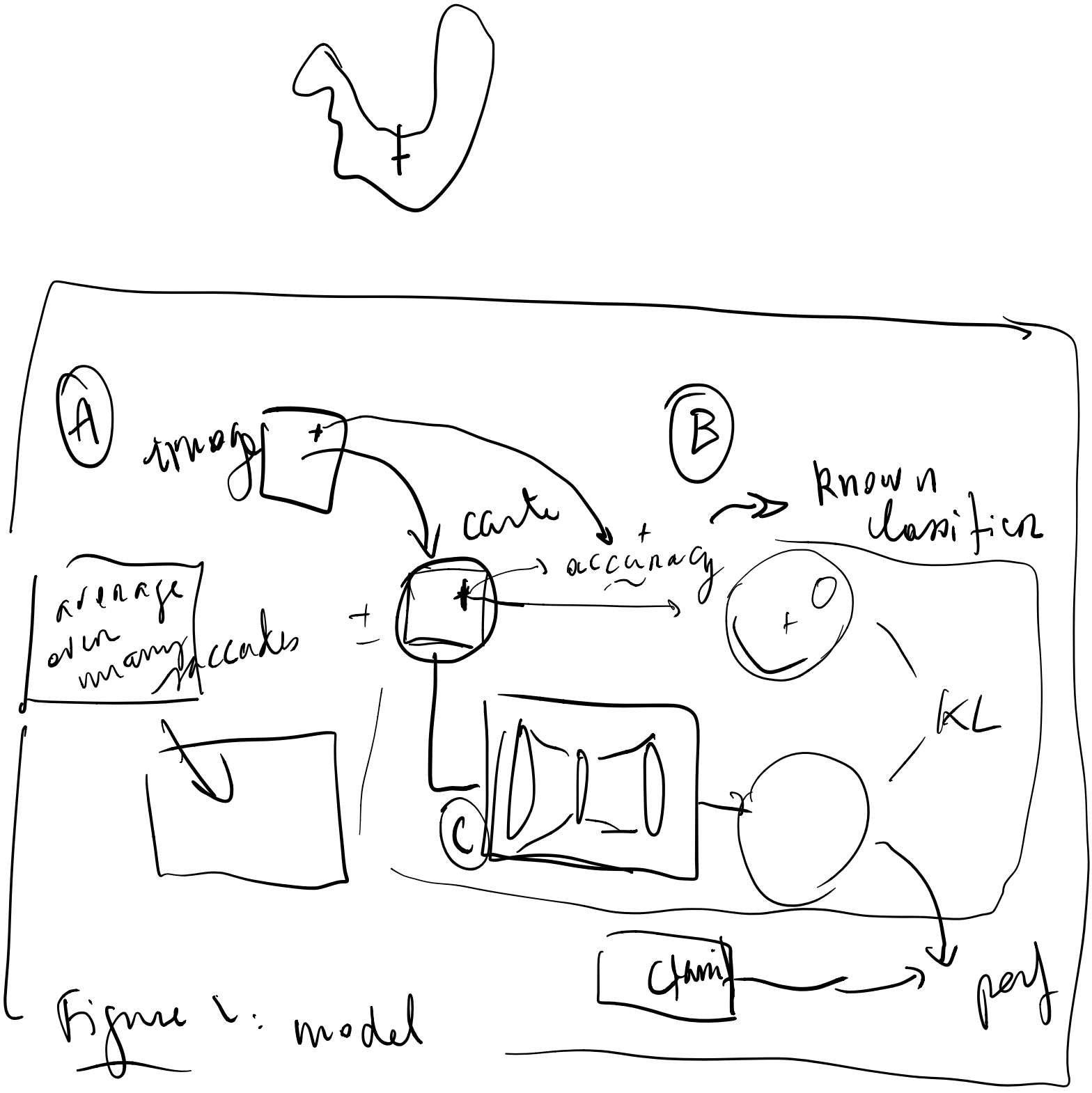
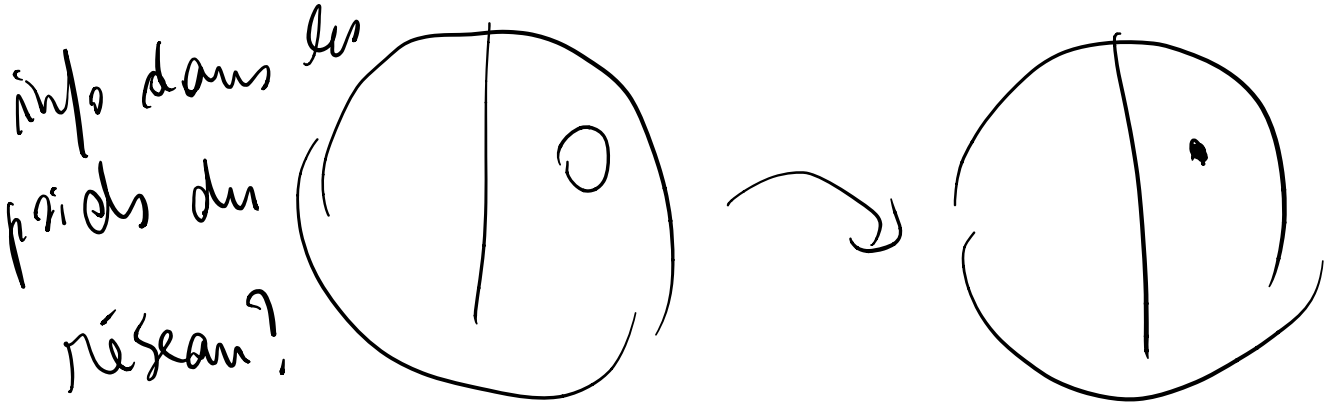


Figure 1 : model



(B)

• Accuracy

online → environment ...

offline → { invariance translation + ergodicity

est-ce un posterior ?

→ TODO notebook montrant le succès de différentes sacades.

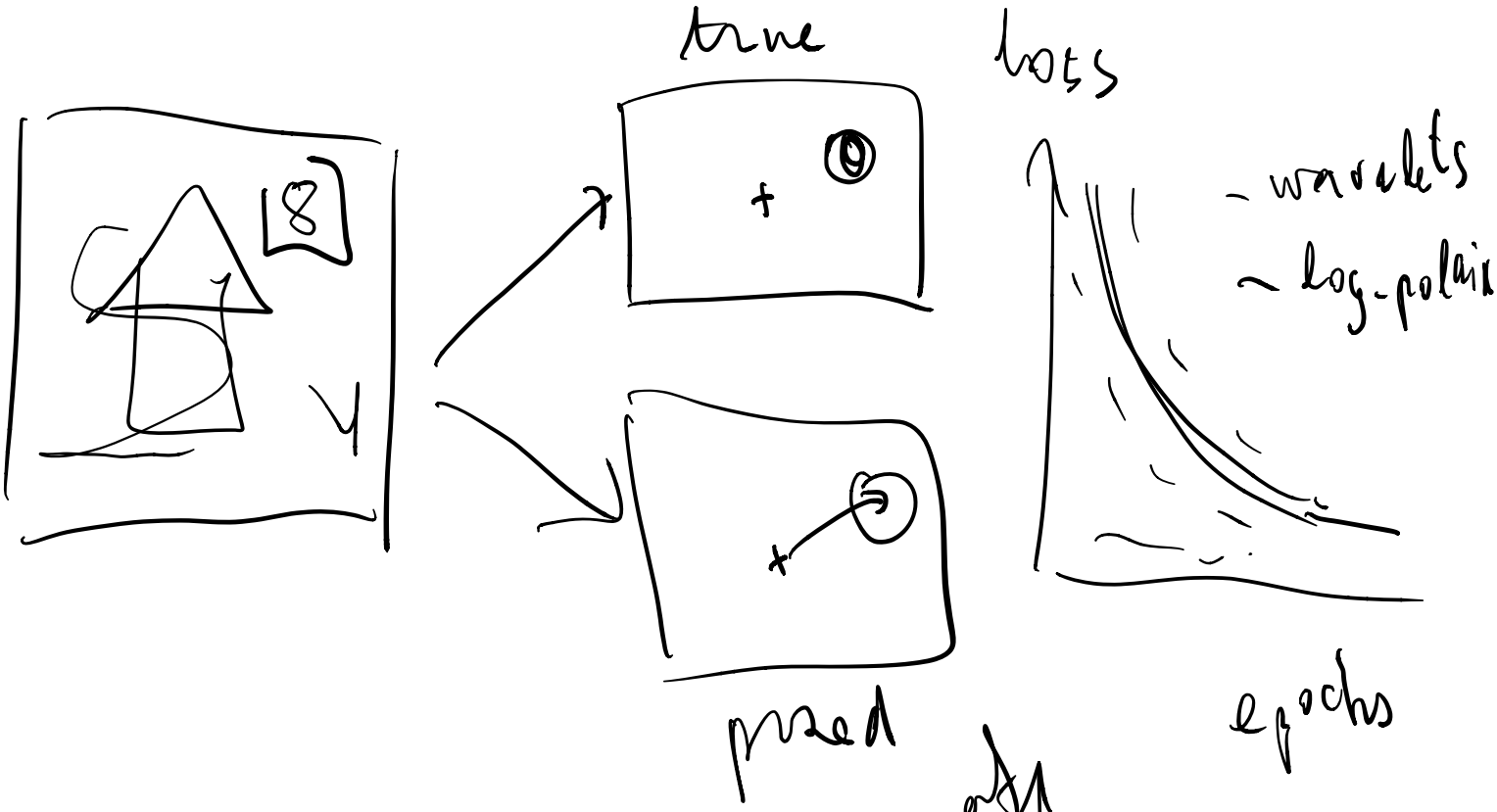
(C) • Network architecture -

N_hidden 1

N_hidden 2 ↑ Adam

Fonctions RELU_tanh -

Results

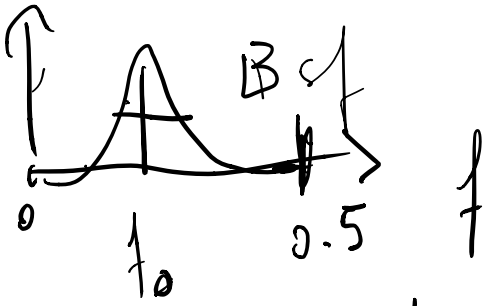


Figures Result1:

4.

https://github.com/laurentperrinet/WhereIsMyMNIST/blob/master/2018-11-13-Where%20recap%20(clutter)%20-%20offset%2030.ipynb

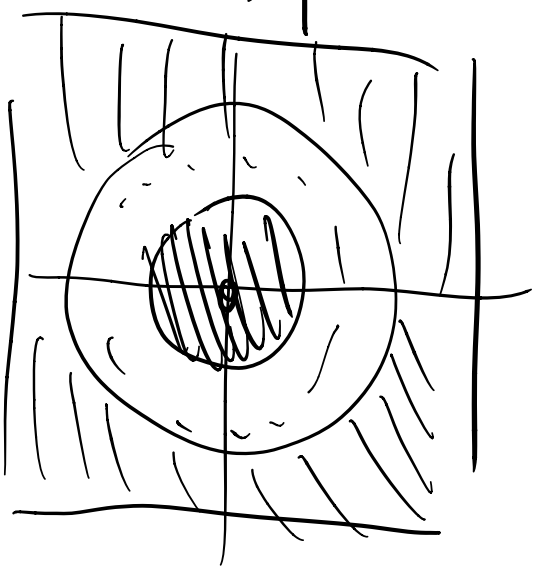
Results with different Clutters

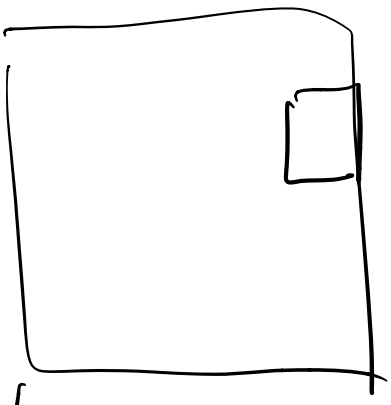


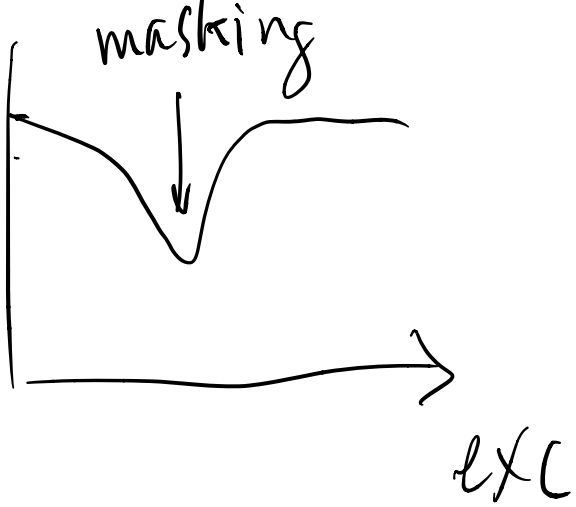
. MC $f = \emptyset$ cycles / px

Bsf

MC $\Delta f = 0$

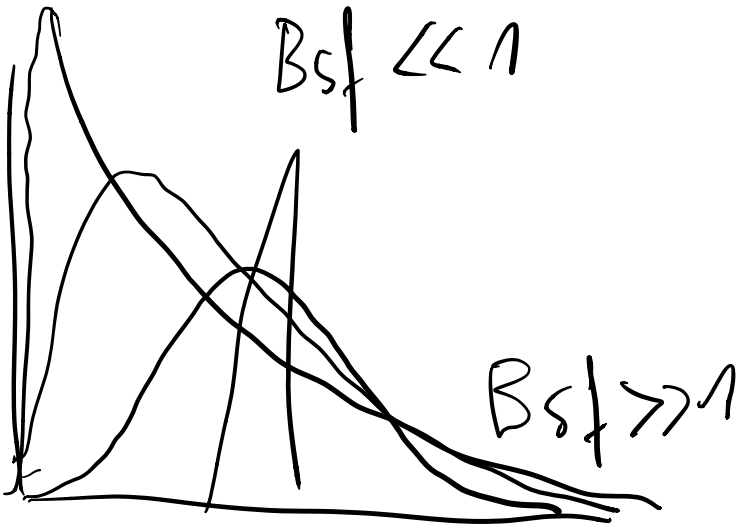




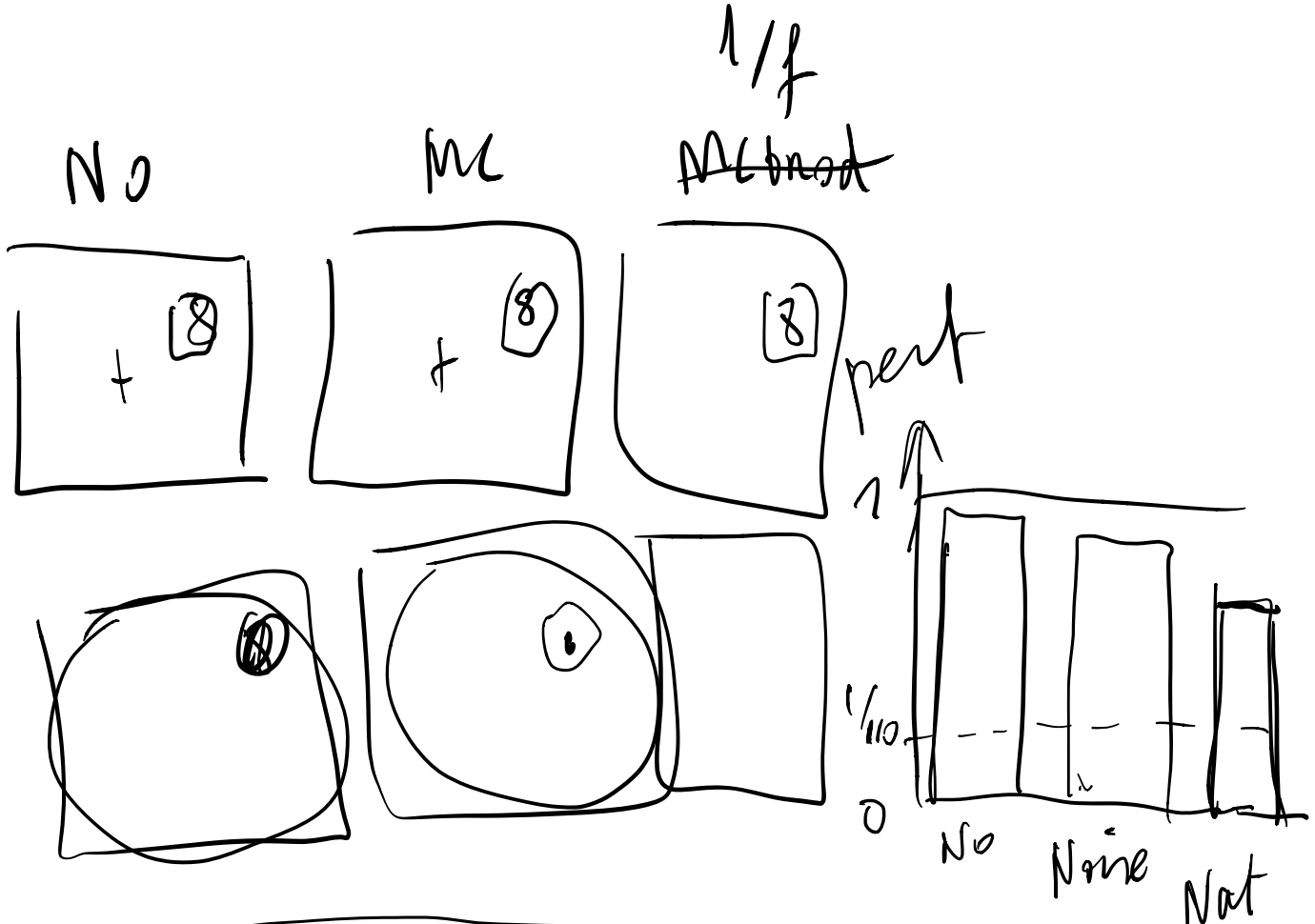


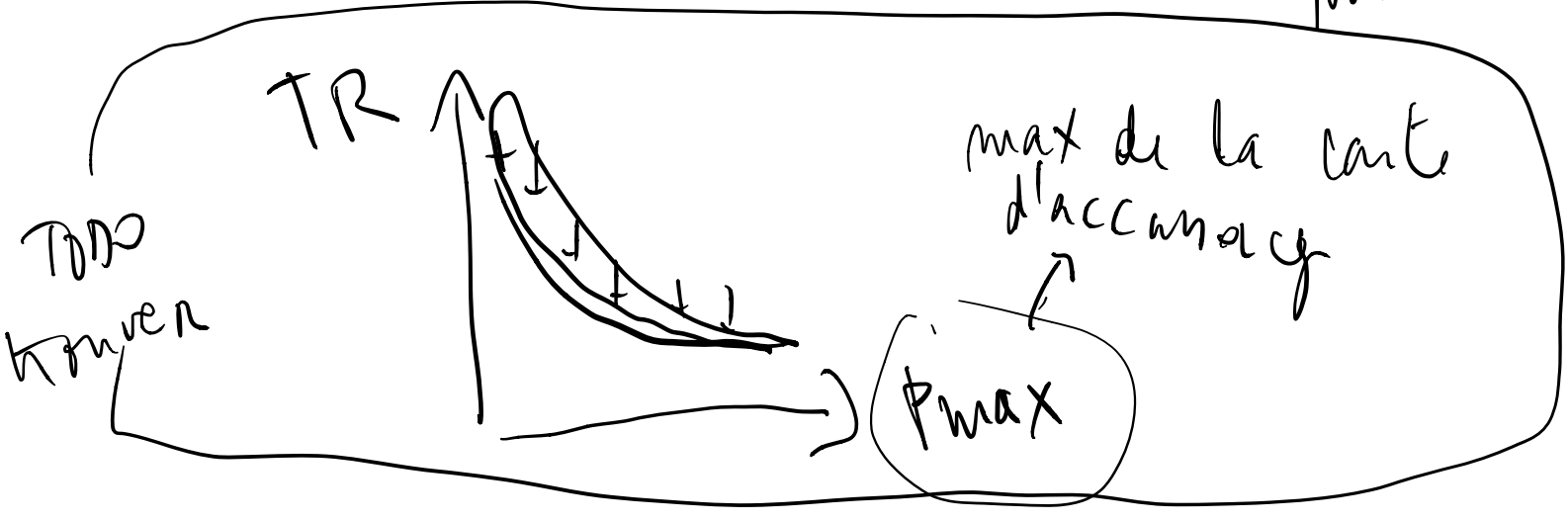
- Bsf $\to \infty$

$\frac{1}{f}$ → stats in naturelles



- image — outdoor
  ~~> f. Malek





4) Perspectives

(A). Summary

... more transformations ... ? = more dim of accuracy map

role de la vision périphérique / task dep.

(B) Limits

- reward? online

main limit = off. line

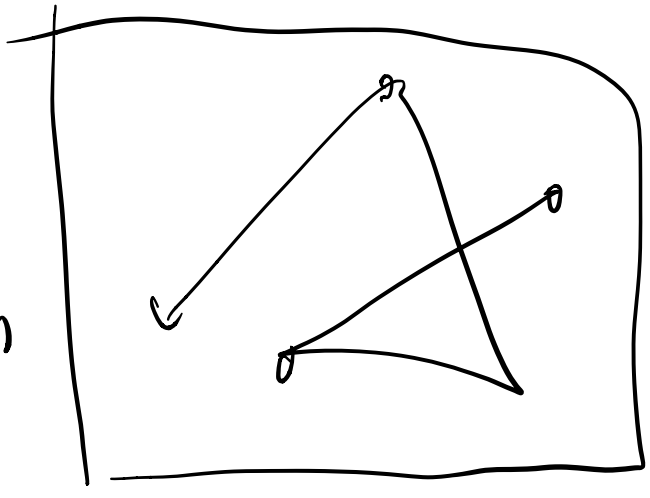
© ~~plusieurs~~ perspectives

• multiples saccades : dépendance

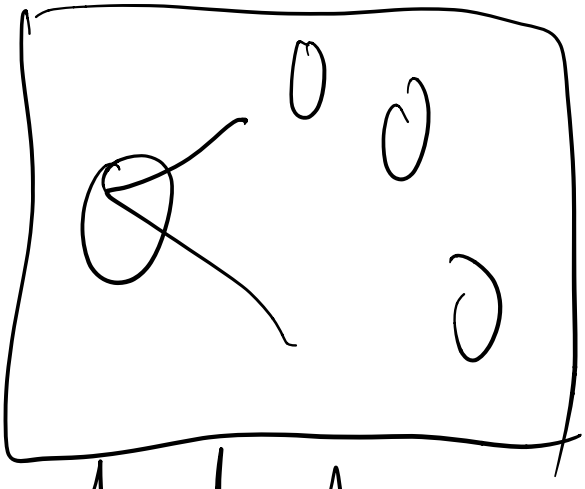
saliency

(→ scanpath) versus context

cf. Yarbus



find MNIST



find faces

Figure : perspectives

→ nécessité de pouvoir prédire la prochaine carte avant de voir la nouvelle image