

# Modèle numérique pour l'évaluation objective de la qualité d'images médicales 2D

Lu ZHANG<sup>1</sup>, Christine CAVARO-MÉNARD<sup>1</sup>, Patrick LE CALLET<sup>2</sup>

<sup>1</sup>LISA, EA 4094, University of Angers  
62 avenue Notre Dame du Lac, 49000 Angers, France

<sup>2</sup>IRCCyN, UMR-6597, University of Nantes  
Rue Christian Pauc, BP50609, 44306 Nantes cedex 3, France

lu.ge@etud.univ-angers.fr,  
christine.menard@univ-angers.fr,  
patrick.lecallet@univ-nantes.fr

**Résumé** – Dans ce papier, un nouveau modèle numérique (le PCJO pour Perceptually relevant Channelized Joint Observer) est proposé pour l'évaluation objective de la qualité d'IRMs cérébrales 2D. Afin d'améliorer la pertinence clinique des modèles numériques, le PCJO est capable d'évaluer la performance des tâches de détection et de localisation de plusieurs lésions potentielles avec paramètres variables (amplitude, orientation, taille, emplacement et nombre). Les résultats des tests de performance des observateurs humains ont montré que la performance du modèle proposé est proche de celle des radiologues pour certains réglages du modèle.

**Abstract** – In this paper, a new numerical observer (PCJO, Perceptually relevant Channelized Joint Observer) is proposed for the objective quality assessment of 2D cerebral MR images, to improve existing models' clinical relevance. The model is able to evaluate the task performance of the detection-localization of multiple lesions with variable parameters (amplitude, orientation, size, position and number). The results of a human observer study showed that the model's performance is close to that of radiologists with certain parameter settings.

## 1 Introduction

L'imagerie médicale est devenue incontournable dans la prise en charge d'un patient que ce soit au cours du dépistage ou dans la phase diagnostique et dans le suivi thérapeutique. Un aspect important en science de l'imagerie médicale est l'évaluation de la qualité de ces images pour un diagnostic fiable et précis.

De nombreuses études font état de l'utilisation de critères numériques simples tels que le rapport signal-bruit (SNR) afin d'évaluer la qualité d'images résultats d'un nouveau protocole d'acquisition, d'un système de codage (compression, tatouage), etc. Néanmoins, il est plus logique de définir la qualité de l'image en fonction de l'objectif recherché. Par exemple, si une image est utilisée pour une tâche diagnostique (détection, localisation, ou caractérisation), la qualité d'image doit idéalement être jugée selon la performance de la tâche effectuée sur cette image par un observateur. Une telle approche est connue sous le nom de l'*approche basée sur la tâche* (en anglais, task-based approach) [1], et est largement admise aujourd'hui.

En pratique, l'évaluation de la performance d'un système d'imagerie réalisée par des observateurs humains est longue et fastidieuse. De plus, des variances intra- et inter-observateurs existent dans leurs réponses. Afin de résoudre ces problèmes, des observateurs numériques ont été développés [1] et permettent une première évaluation rapide et reproductive. Néanmoins,

l'observateur humain demeure l'agent principal des décisions diagnostiques et reste nécessaire pour la validation de tout modèle numérique.

La problématique principale des modèles numériques existants réside dans le fait que la gamme des tâches ainsi que la gamme des paramètres caractérisant la lésion sont limitées [5]. La plupart des modèles existants ne gèrent que la tâche de détection d'un *signal-known-exactly* (SKE) où l'observateur numérique a une connaissance exacte du signal. Or en routine clinique l'observateur humain ne connaît pas *a priori* les propriétés du signal recherché<sup>1</sup>. Afin de prendre en compte l'incertitude du signal, il est naturel de considérer la tâche de détection d'un *signal-known-statistically* (SKS) dont les paramètres du signal ne sont spécifiés que par des lois de probabilité. Peu de modèles numériques SKS ont été développés et les modèles existants ne traitent que la détection de lésion unique centrale [1].

Dans ce papier, nous proposons un modèle SKS capable d'évaluer la performance des tâches de détection et localisation de plusieurs signaux avec paramètres (amplitude, orientation, taille, emplacement et nombre) variables. Les résultats d'une étude avec des observateurs humains montrent que la performance du modèle proposé est proche de celle des radiologues.

---

<sup>1</sup>Tel que l'amplitude, la taille, l'orientation, l'emplacement et le nombre

## 2 Modèle pour la détection

Nous nous plaçons dans un contexte de détection d'un signal d'intérêt sur un fond bruité. Soit la validation d'une des deux hypothèses exclusives suivantes:

$\mathcal{H}_0$  : le signal  $\mathbf{x}_\alpha$  est absent, seul le fond bruité  $\mathbf{b}$  est observé;

$\mathcal{H}_1$  : le signal  $\mathbf{x}_\alpha$  est présent sur le fond bruité  $\mathbf{b}$ .

où le vecteur  $\alpha = \{\mathbf{q}, \theta, b, \sigma^2, a\}$  contient les variables inconnues paramétrant un signal recherché ( $\mathbf{q}, \theta, b, \sigma^2, a$  spécifient respectivement l'emplacement, l'orientation, la forme, la taille et l'amplitude du signal).

Pour l'évaluation de la performance de la tâche de la détection, nous avons proposé un modèle nommé CJO (en anglais, Channelized Joint Observer), basé sur la théorie de la détection-estimation conjointe. Dans ce cas, l'estimateur  $\hat{\alpha}$  (caractérisant la lésion) et  $\widehat{\mathcal{H}}_k$  (caractérisant la présence ou l'absence d'une lésion) sont choisis de façon conjointe pour maximiser la probabilité conjointe:

$$\begin{aligned} (\hat{\alpha}, \widehat{\mathcal{H}}_k) &= \arg \max_{\alpha, \mathcal{H}_k} P(\alpha, \mathcal{H}_k | \mathbf{g}) \\ &= \arg \max_{\alpha, \mathcal{H}_k} P(\mathbf{g} | \alpha, \mathcal{H}_k) P(\alpha) P(\mathcal{H}_k) \end{aligned} \quad (1)$$

où l'indépendance statistique de  $\alpha$  et  $\mathcal{H}_k$  a été exploitée.

Sous l'hypothèse d'un fond gaussien corrélé de moyenne nulle, l'estimateur  $\hat{\alpha}$  peut s'écrire :

$$\hat{\alpha} = \arg \max_{\alpha} \frac{1}{\left\| \mathbf{U}_0 (\mathbf{A}'_{\alpha})^t \right\|_F^2} (\mathbf{x}'_0)^t (\Sigma'_b)^{-1} \left( \mathbf{A}'_{\alpha} \mathbf{g}' - \frac{1}{2} \mathbf{x}'_0 \right) \quad (2)$$

le test statistique  $\lambda$ , scalaire considéré comme la note de confiance de la détection, peut alors s'écrire :

$$\lambda = \frac{1}{\left\| \mathbf{U}_0 (\mathbf{A}'_{\hat{\alpha}})^t \right\|_F^2} (\mathbf{x}'_0)^t (\Sigma'_b)^{-1} \left( \mathbf{A}'_{\hat{\alpha}} \mathbf{g}' - \frac{1}{2} \mathbf{x}'_0 \right) \quad (3)$$

où  $\Sigma_b$  désigne la covariance du fond dans le domaine multi-canaux<sup>2</sup> [5] ;  $\mathbf{U}_0$  est une matrice fixe indépendante de  $\alpha$  ;  $\mathbf{A}'_{\alpha}$  permet de transformer le signal multi-canaux  $\mathbf{x}'_{\alpha}$  en signal standard multi-canaux  $\mathbf{x}'_0$  :  $\mathbf{A}'_{\alpha} \mathbf{x}'_{\alpha} = \mathbf{x}'_0$ .

Afin d'estimer l'orientation et la taille du signal dans le domaine multi-canaux, les canaux doivent être *modulables* pour l'orientation et pour l'échelle<sup>3</sup>. Nous construisons donc  $\mathbf{A}'_{\theta}$  et  $\mathbf{A}'_{\sigma}$  de la manière suivante :

$$\begin{aligned} [\mathbf{A}'_{\theta}]_{m,n} &= \frac{1}{K} \frac{\sin(\pi(m-n) - \theta K)}{\sin(\pi(m-n)/K - \theta)} \quad m, n = 1, \dots, K; \\ [\mathbf{A}'_{\sigma}]_{m,n} &= \text{sinc}(((m-n) - (-\log_2 \sigma))) \quad m, n = 1, \dots, J; \end{aligned} \quad (4)$$

<sup>2</sup>La technique de décomposition en multi-canaux par une série de filtres est souvent utilisée par les modèles numérique pour ramener la charge calculatoire à un niveau raisonnable.

<sup>3</sup>en anglais, ce sont des canaux *steerable* et *scalable*

où  $[\cdot]_{m,n}$  désigne l'élément de la  $m$ -ème ligne, et  $n$ -ème colonne de la matrice sous-jacente;  $K$  et  $J$  désignent respectivement le nombre de canaux modulables pour l'orientation du signal ( $K$ ) et pour l'échelle (ou la taille) du signal ( $J$ ). Pour le cas où l'amplitude, l'orientation et la taille sont à la fois inconnus :

$$\mathbf{A}'_{\alpha} = \mathbf{A}'_{a,\theta,\sigma} = \frac{1}{a} \mathbf{A}'_{\sigma} \otimes \mathbf{A}'_{\theta} \quad (5)$$

où  $\otimes$  désigne le produit de Kronecker de deux matrices.

L'implémentation du CJO comprend deux étapes : une étape d'entraînement où un template est calculé à partir d'images d'entraînement (lésion-présente et lésion-absente) ; et une étape de test où le test statistique  $\lambda$  est calculé en utilisant le template et l'image de test.

## 3 Modèle pour la localisation

Pour traiter une tâche diagnostique plus complexe et plus pertinente cliniquement, nous proposons un modèle nommé PCJO (en anglais, Perceptually relevant CJO) pour l'évaluation de la performance des tâches de détection et de localisation de plusieurs signaux.

### 3.1 Structure

La structure générale du PCJO, illustrée Fig. 1, présente deux étapes principales:

1. La première étape permet de définir les zones candidates (dimension  $\sqrt{M} \times \sqrt{M}$ ) présentant une anomalie potentielle qui mérite davantage d'investigation. La carte de probabilité de détection est calculée par un modèle du système visuel humain (SVH) dont les entrées sont une coupe cérébrale sans anomalie (en tant qu'image de référence) et la même coupe avec des lésions simulées superposées (en tant qu'image de distorsion). L'intensité, l'orientation l'échelle, le nombre et l'emplacement des lésions sont aléatoires et cohérents avec la pathologie étudiée. Plusieurs images de référence et avec lésions simulées sont nécessaires en entrée pour permettre le calcul de la figure de mérite.
2. La seconde étape utilise le CJO pour obtenir un test statistique par image test. Le template résulte de l'entraînement sur des régions homogènes extraites de coupes cérébrales sans anomalie. Dans une partie de ces images homogènes est ajoutée une lésion simulée dont l'intensité, l'orientation et l'échelle sont aléatoires, mais dont l'emplacement est fixé au centre de l'image. Toutes les images d'entraînement (lésion-présente et lésion-absente) sont de même dimension ( $\sqrt{M} \times \sqrt{M}$ ).

Notons qu'afin de proposer un modèle numérique pertinent, le PCJO intègre différents aspects de la vision humaine spécifiques à l'interprétation d'images médicales :

- En imitant le processus diagnostique des radiologues qui dans une première étape font une recherche globale des

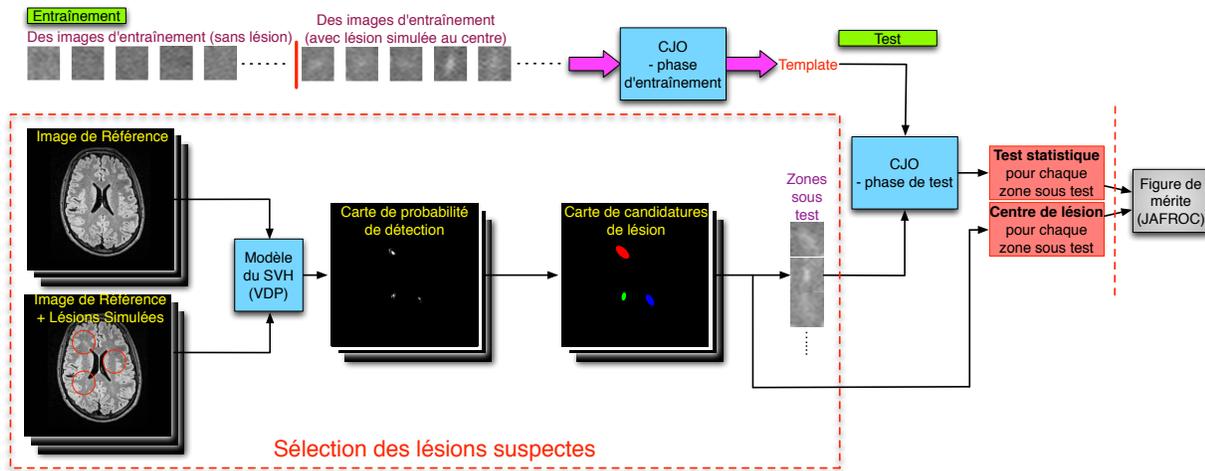


FIG. 1: Structure générale du PCJO pour détecter et localiser sur une coupe plusieurs signaux avec paramètres variables.

zones lésionnelles potentielles dans l'image, puis dans une deuxième étape réalisent une analyse cognitive des zones lésionnelles potentielles suivie d'une prise de décision.

- En se servant d'un modèle du SVH<sup>4</sup> pour modéliser la sensation (la première étape de la vision humaine). Le modèle VDP [3] a été choisi suite à une étude présentée en [4] qui montre que la performance de sensation du VDP s'approche le plus de celle des radiologues (comparée à la performance des modèles du SVH les plus connus).
- En adoptant la phase d'entraînement du CJO pour modéliser en partie la perception, en particulier celle liée à l'expérience. En effet, l'expérience ou la formation des radiologues aboutit à une différence cohérente de perception, ce qui explique l'influence importante de l'expertise dans la tâche de détection. Notre modèle utilise la phase d'entraînement pour se former comme les radiologues.

Le modèle PCJO est détaillé dans [5].

## 3.2 Validation du modèle proposé

### 3.2.1 Pathologie et modalité d'étude

Pour l'étude de validation, nous avons choisi la Sclérose En Plaque (SEP) comme pathologie d'étude. La raison principale de ce choix est que pour cette pathologie, l'orientation, la taille, l'emplacement et le nombre des lésions sont tous considérés comme des critères diagnostiques. Elle est donc adaptée à notre étude des modèles numériques. De plus, les lésions SEP sont subtiles et difficiles à détecter. Elle est donc une pathologie appropriée pour évaluer la performance des tâches de détection et de localisation.

Le diagnostic de la SEP est largement basé sur l'analyse d'IRM cérébrales. Dans certaines circonstances un diagnos-

tic de la SEP peut être confirmé avec un seul examen d'IRM. Nous avons donc choisi l'IRM comme modalité d'étude.

### 3.2.2 Expérience subjective

Nous avons mené une expérience subjective à réponse libre [2], proche d'une analyse réaliste en routine clinique, afin de comparer la performance du modèle avec celle des radiologues et finalement valider le modèle.

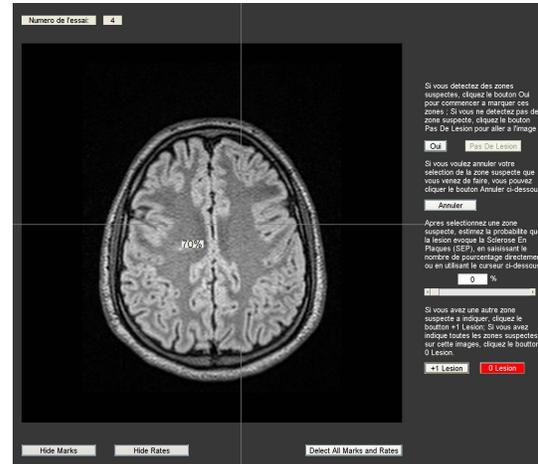


FIG. 2: Interface graphique de l'expérience subjective pour la validation du nouveau modèle d'observateur proposé

Un expert et quatre radiologues ont participé à l'expérience. La distance d'observation était de 40cm. L'interface graphique est montrée Fig. 2. Il était demandé aux radiologues de détecter et localiser des lésions de SEP sur une coupe cérébrale d'une séquence FLAIR T2 sans limite de temps. Les coupes à interpréter étaient soit normales (sans lésion), soit anormales avec lésion(s) simulée(s). Le nombre de lésions était aléatoire mais  $< 5$ . Il était demandé aux radiologues de marquer avec précision le centre de la zone suspecte. En effet, la tâche de localisa-

<sup>4</sup>La modélisation de diverses composantes et caractéristiques du SVH.

tion intègre un critère de proximité du signal marqué au signal simulée constituant le gold standard. Le critère peut être un rayon d'acceptation (pour des petits signaux, comme dans notre cas) ou un pourcentage de chevauchement (pour des grands signaux). Il était également demandé aux radiologues de définir pour chaque lésion potentielle la probabilité que la zone soit une lésion.

Lors de l'expérience, 160 images de taille  $1024 \times 1024$ , dont 90 images positives (avec lésions) et 70 images négatives (sans lésion), ont été montrées à chaque radiologue.

Les paramètres du CJO étaient configurés comme suit afin d'être cohérent avec la pathologie étudiée (ici la SEP) :

$[\sigma_{\min} \sigma_{\max}] = [1 \ 12]$ ,  $[\theta_{\min} \theta_{\max}] = [0 \ \pi]$ ,  $[a_{\min} a_{\max}] = [1 \ 255]$  (degré d'intensité du pixel),  $\sqrt{M} \times \sqrt{M} = 65 \times 65$ . Notons que la valeur de  $\sqrt{M}$  (cf. section 3.1) est choisie pour que la zone candidate puisse couvrir les plus grandes lésions de SEP. 1000 zones différentes, dont 500 zones avec lésion et 500 zones sans lésion, ont été utilisées pour la phase d'entraînement du CJO.

### 3.2.3 Résultats

TAB. 1: Valeurs de la figure de mérite JAFROC1 (FOMs) des radiologues

	JAFROC1 FOM	Ecart-type
Radiologue 1 ( <b>expert</b> )	<b>0.7672</b>	0.0326
Radiologue 2	0.4736	0.0378
Radiologue 3	0.4278	0.0399
Radiologue 4	0.4742	0.0419
Radiologue 5	0.4728	0.0369

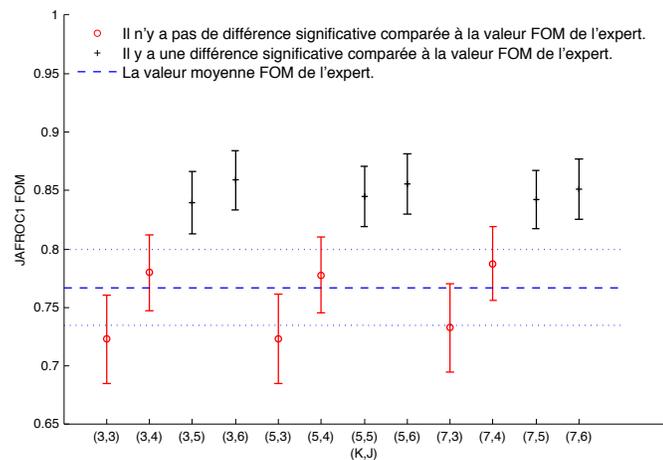


FIG. 3: Valeurs de la figure de mérite JAFROC1 (FOMs) du PCJO avec différents  $K$  (nombre de canaux d'orientation) et  $J$  (nombre de canaux d'échelle) configurés dans le CJO.

La performance du modèle a été comparée à celle des radiologues en utilisant la figure de mérite (FOM) JAFROC1 [2]. Notons que l'analyse JAFROC1 nécessite deux entrées : la décision de détection (désignée par le test statistique) pour

chaque image de test et l'emplacement de chaque lésion.

Les valeurs de la figure de mérite JAFROC1 (FOMs) des radiologues sont montrées dans le TAB. 1. La performance de l'expert apparaît significativement plus élevée que celle des autres radiologues.

Nous calculons ensuite les FOMs du PCJO avec différentes configurations de  $K$  (nombre des canaux modulables pour l'orientation) et  $J$  (nombre des canaux modulables pour l'échelle). Comme le montrent les résultats présentés Fig.3, le PCJO atteint une performance proche de celle de l'expert lorsque  $J \leq 4$ . De plus, nous observons que  $K$  influence peu la performance du PCJO, à l'inverse de  $J$ . Par ailleurs, les résultats illustrent que le PCJO peut atteindre une performance satisfaisante même avec un petit nombre de canaux (par exemple, pour  $K = 3$  et  $J = 4$ ), ce qui est favorable à la réduction de la charge calculatoire dans le PCJO.

## 4 Conclusion

Dans le cadre de l'évaluation objective de la qualité d'images médicales basée sur la tâche diagnostique, nous avons proposé un modèle numérique intégrant plusieurs facteurs du SVH impliqués dans l'analyse et l'interprétation des images (la sensation et la perception, l'expertise des radiologues, etc.). Tandis que la plupart des modèles existants gèrent uniquement la détection d'un seul signal fixé au centre de l'image, le PCJO a élargi la gamme des paramètres variables du signal ainsi que la gamme des tâches diagnostiques prises en compte (détection et localisation). La performance du modèle proposé est proche de celle des radiologues pour certains réglages des paramètres du modèle. Ce prometteur modèle est actuellement étendu aux images médicales volumétriques (3D) pour une modélisation plus réaliste des tâches de détection et de localisation de lésions.

## References

- [1] H.H. Barrett and K.J. Myerers. *Foundations of Image Science*. John Wiley and Sons, Inc., Hoboken, New Jersey, USA, 2004.
- [2] D. P. Chakraborty. Recent advances in observer performance methodology: jackknife free-response ROC (JAFROC). *Radiat Prot Dosimetry.*, 114:26–31, 2005.
- [3] S. Daly. *The visible differences predictor: an algorithm for the assessment of image fidelity*, volume 13, pages 179–206. MA: MIT Press, 1993.
- [4] L. Zhang, C. Cavaro-Menard, and P. Le Callet. Evaluation of HVS models in the application of medical image quality assessment. In *Proceedings of SPIE*, volume 8291, page 82910T, 2012.
- [5] L. Zhang, C. Cavaro-Ménard, P. Le Callet, and J. TAN-GUY. A perceptually relevant channelized joint observer (PCJO) for the detection-localization of parametric signals. *IEEE Transactions on Medical Imaging*, pp:1, 2012.