

Projet NeSS

Neuromorphic Semantic Segmentation

Jean MARTINET, Dalia HAREB

UniCA-CNRS i3s

19 novembre 2024

Plan

- © Contexte
Informatique neuromorphique, vision par ordinateur
- © Focus sur le projet **Neuromorphic Semantic Segmentation**
- © La suite

Informatique neuromorphique, vision par ordinateur

- © Matériels et logiciels inspirés de la structure et du fonctionnement du cerveau biologique
- © Apprentissage automatique neuro-inspiré pour la vision
- © **Perception** : caméras événementielles (rétines silicone)
- © **Traitement** : réseaux de neurones impulsionnels

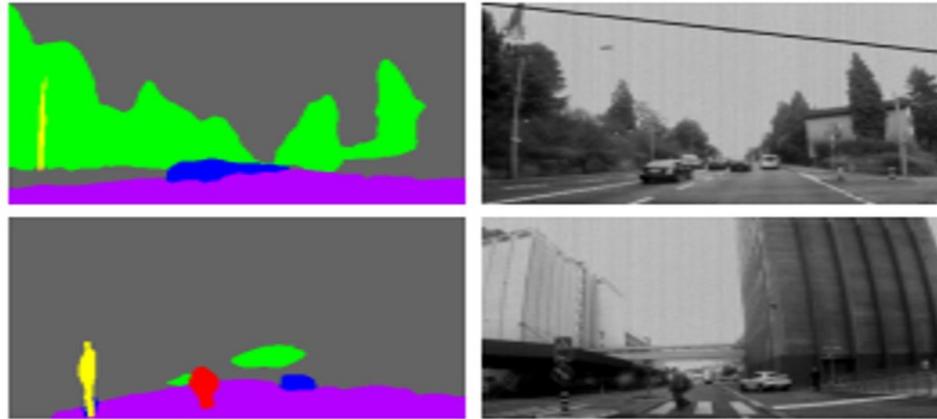
Plan

- © Contexte
Informatique neuromorphique, vision par ordinateur
- © Focus sur le projet Neuromorphic Semantic Segmentation
- © La suite

Segmentation sémantique

Tâche importante en vision par ordinateur

→ Classer chaque pixel selon sa classe sémantique



Exemple de segmentation sémantique
Inigo et al. (CVPRw 2018)

ANN "standards": consommation énergétique

ChatGPT: consomme 0.5 MWh/jour pour gérer ~200M requêtes ≈ L'énergie consommée par 180,000 foyers USA

ResNet: nécessite ~ 4.3 GJ (1.2 MWh) ≈ L'énergie consommée par un appartement de 2 personnes sur 40 jours.

Conséquences:

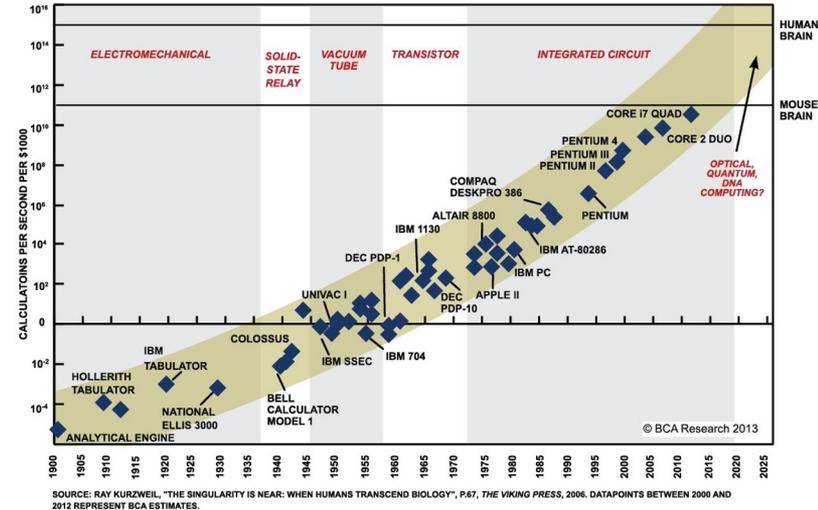
- Problèmes **Écologiques**: Data centers responsable de 4% de GHG (2020) (transport aérien ≈ 2.5%).
- **Restriction** de l'utilisation des applications de l'IA sur du matériel embarqué doté de puissance limitée: IoT, automobile, réalité augmentée et virtuelle ... etc.

ANN "standards" : loi de Moore

Le nombre de transistors double tous les 18 mois grâce à la miniaturisation.

Elle Heurte des **limites physiques**:

1. Taille du transistor à l'échelle atomique.
2. La chaleur dissipée.



Conséquences:

- **Ralentissement** de l'amélioration des performances des processeurs.

Réseaux de Neurones Impulsionnels (SNNs)

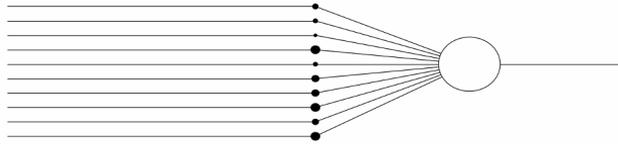
SNN

ANN

Information Spatio-Temporelle

Spatial information

Leaky Integrate and Fire Neuron



Accumulation

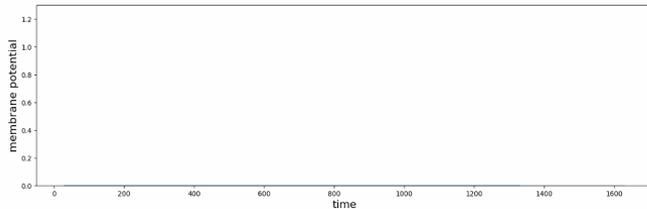
Accumulation + multiplication

Binaire et éparses

Décimaux et dense

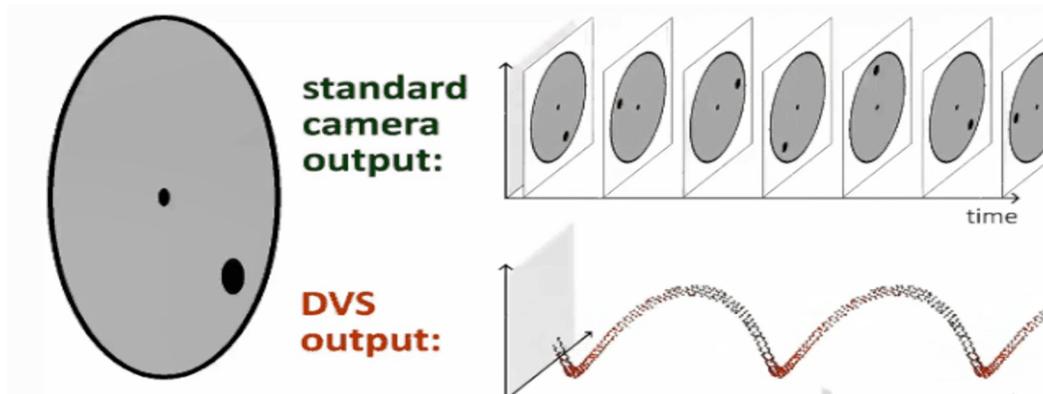
Difficile à entraîner

Facile à entraîner



Caméras événementielles

- ⊙ Résolution temporelle élevée ($\sim 1\mu\text{s}$)
- ⊙ Faible consommation d'énergie (mW versus W)
- ⊙ Faible latence
- ⊙ Absence de floutage dû au mouvement



Différence entre une caméra classique et une caméra événementielle.
Scaramuzza et al. (ECCV 2018)

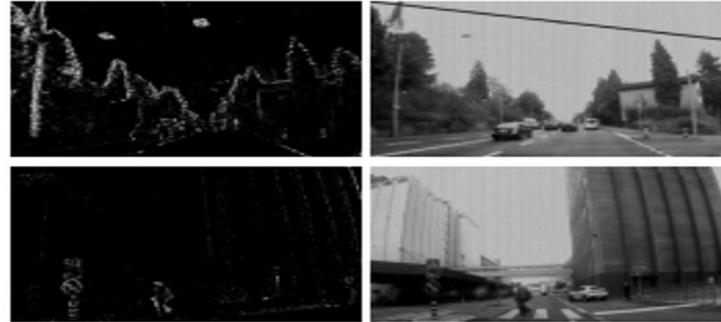
Proposition 1 - Modèle

EvSegSNN: Segmentation sémantique neuromorphique

- Modèle: SNN
- Données: événements

Valeurs ajoutées :

1. Bio-plausibilité
2. Réduction de la complexité du modèle



Exemples des données événementielles et les images de niveau de gris correspondant

Proposition 1 - Résultats

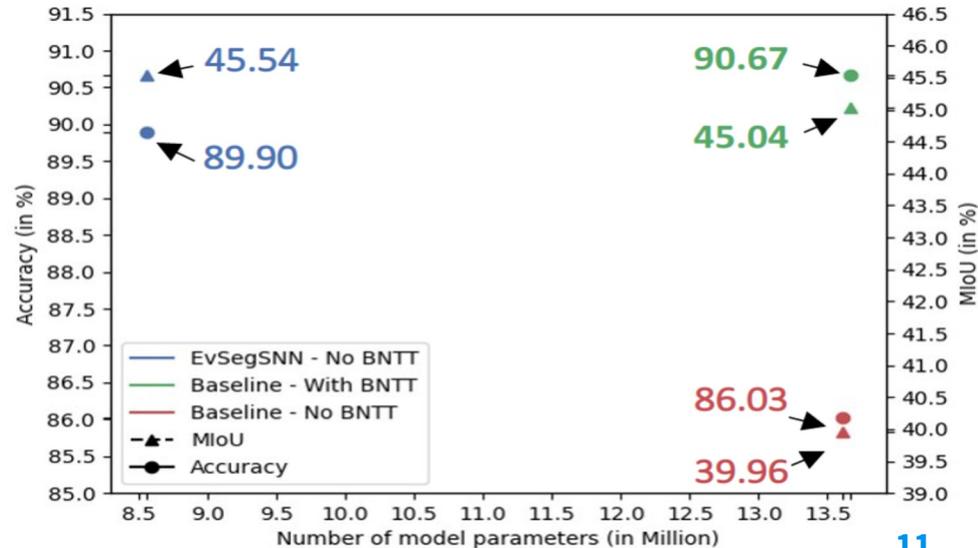
Réduction du nombre de paramètres par un facteur de 1.6

- Meilleure performance de EvSegSNN vs baseline sans “batch normalisation” (bio-plausible)
- Performance équivalente à la baseline avec “batch normalisation” (non bio-plausible)

Métriques:

- $MIoU = TP / Total$
- $Accuracy = (TP + TN) / Total$

(IJCNN'2024)



Proposition 2 - Modèle

Combinaison des caméras événementielles et à frames (RGB)

T_{-1}

$T (+50ms)$

Time



Image1

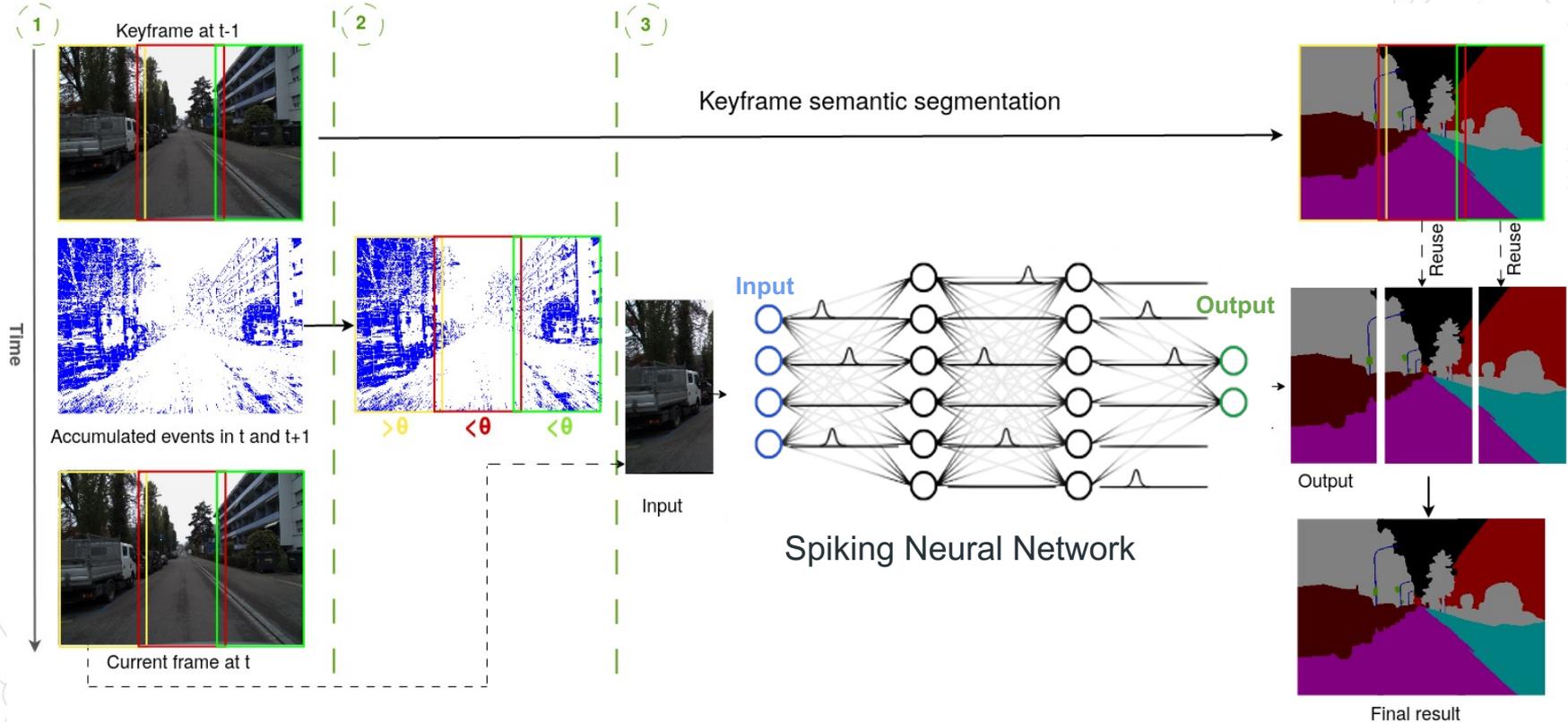


Events in $[T_{-1}, T]$



Image2

Proposition 2 - Modèle



Proposition 2 - Résultats

Threshold	MIoU [%]	FPS
Baseline	54.46	151.25
0 w/ split	54.19 (-0.30)	47.34 ($\times 0.27$)
4.00	53.05 (-1.41)	542.62 ($\times 3.57$)
4.75	52.71 (-1.75)	666.52 ($\times 4.47$)
5.25	52.52 (-1.95)	759.85 ($\times 5.07$)
6.00	52.19 (-2.27)	916.82 ($\times 6.08$)
8.00	51.56 (-2.90)	1261.00 ($\times 8.15$)
10.0	51.11 (-3.35)	1513.67 ($\times 10.05$)

Résultats de segmentation en termes de MIoU et Frame/s

Threshold	Baseline	4.75	10.0
#Flops	$5,99 \times 10^{12}$	$3,30 \times 10^{12}$ ($\times 1.81$)	$2,17 \times 10^{12}$ ($\times 2.75$)

Nombre de flops nécessaires pour évaluer l'ensemble des tests

Production scientifique (Nov 22 – Mars 25)

- © Dalia Hareb, Jean Martinet, Benoît Miramond. **Enhanced neuromorphic semantic segmentation latency through stream event**. International Conference on Neural Information Processing (ICONIP). Auckland, New Zealand, December 2024
- © Huyen Trang Nguyen, Laurent Sparrow, Jean Martinet. **Top-down and bottom-up attentions in event data: a survey**. International Conference on Neural Information Processing (ICONIP). Poster. Auckland, New Zealand, December 2024
- © Dalia Hareb, Jean Martinet. **EvSegSNN: Neuromorphic Semantic Segmentation for Event Data**. IJCNN 2024, Yokohama, Japan. July 2024
- © Amélie Gruel, Dalia Hareb, Antoine Grimaldi, Jean Martinet, Laurent Perrinet, Bernabé Linares-Barranco and Teresa Serrano-Gotarredona. **Stakes of Neuromorphic Foveation: a promising future for embedded event cameras**. Biological Cybernetics Special Issue: What can Computer Vision learn from Visual Neuroscience? 2023
- © Dalia Hareb, Jean Martinet. **EvSegSNN : Segmentation sémantique neuromorphique pour la vision événementielle**. ORASIS 2023, Carqueiranne, France. May 2023
- © Amélie Gruel, Jean Martinet. **Sélection neuromorphique simultanée d'objets saillants dans des événements**. ORASIS 2023, Carqueiranne, France. May 2023.
- © Amélie Gruel, Dalia Hareb, Jean Martinet, Bernabé Linares-Barranco, Teresa Serrano-Gotarredona. **Neuromorphic foveation applied to semantic segmentation**. CVPR 2022 workshop "NeuroVision: What can computer vision learn from visual neuroscience?", New Orleans, Louisiana, June 2022. BEST POSTER AWARD

Plan

- © Contexte

Informatique neuromorphique, vision par ordinateur

- © Focus sur le projet Neuromorphic Semantic Segmentation

- © **La suite**

Future directions

Neuromorphic attention models for event data (French-Swiss ANR NAMED),
Feb 2024-Jan 2028

- © Leverage top-down attention and bottom-up saliency
- © Ex : would you say that this family is rich?



Future directions

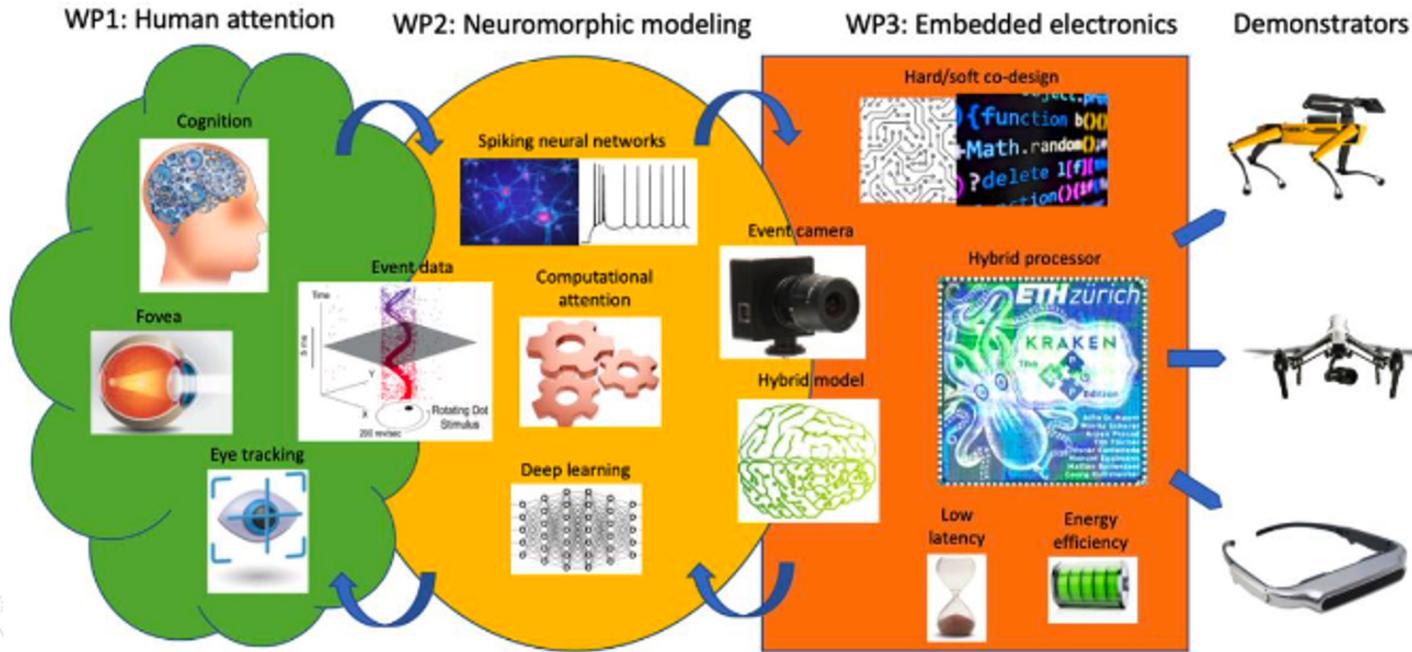
Example of task-driven cognitive attention (Yarbus, 1967)

- © Left: original image
- © Middle: Eye tracking trace (no task, free viewing)
- © Right: Eye tracking trace (Task: “estimate the material circumstances of the family”)



Future directions

Neuromorphic attention models for event data (French-Swiss ANR NAMED), Feb 2024-Jan 2028



Projet NeSS

Neuromorphic Semantic Segmentation

Jean MARTINET, Dalia HAREB

UniCA-CNRS i3s

19 novembre 2024