

Spiking Neural Networks Sound Detection and Classification

COURREGÉ Téo
GANDEEL Loïc

https://github.com/LGPolytech/Project_S9

February 22, 2024

Introduction

1 Introduction

2 Rappels

- Spikes et Spike encoding
- Neurones spikant
- Réseau de neurones spikants

3 SNN - Behind the scene

- Entraîner un SNN
- Dead Neuron Problem
- Surrogate gradient

4 Classification et résultats

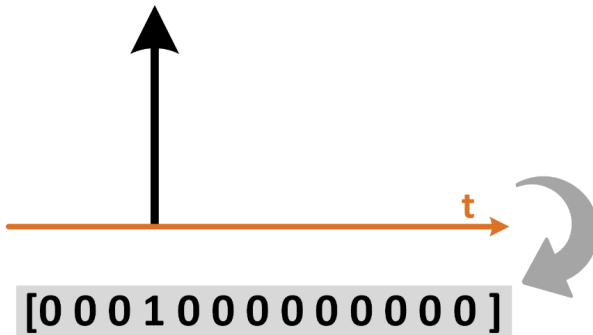
- MNIST
- Dataset pour la classification sonore
- Résultats sur notre dataset

Spike ou impulsions électriques

Impulsion électrique de courte durée (I_{in})

Modèle de communication entre neurones

Permet de transmettre de l'information



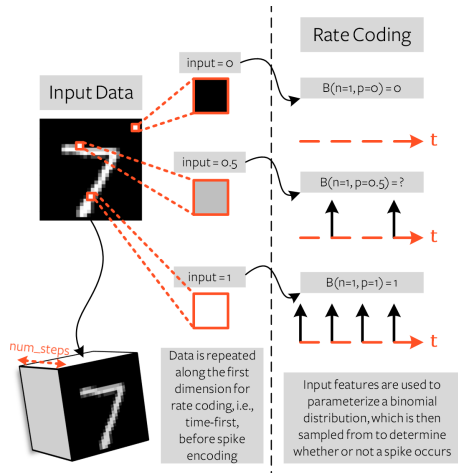
Encodage de l'information en spikes

Manière de représenter l'information

Approximation de l'information par des impulsions électriques

Variable temporelle ou fréquentielle

Encodage temporel, fréquentiel, ou mixte



Exemple avec MNIST

Modèle Leaky Integrate and Fire (LIF)

$$U[t + 1] = \underbrace{\beta U[t]}_{\text{decay}} + \underbrace{WX[t + 1]}_{\text{input}} - \underbrace{\beta S[t] U_{\text{thr}}}_{\text{soft reset}}$$

$$S[t] = \begin{cases} 1, & \text{if } U[t] > U_{\text{thr}} \\ 0, & \text{otherwise} \end{cases}$$

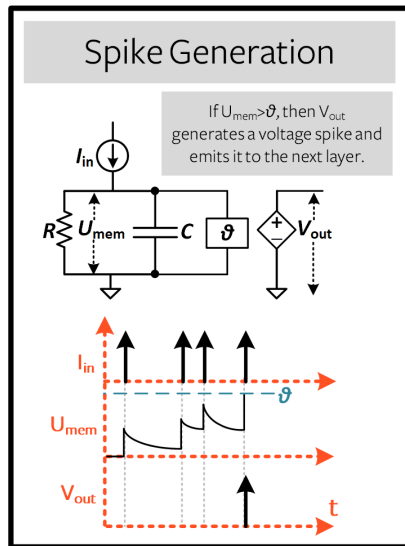
U : Potentiel de membrane

W : Poids du réseau

X : Entrée du réseau (des spikes)

S : Fonction de spike

$\beta \in]0, 1[$: Facteur de décharge



Définition d'un réseau de neurones

Convolutional Spiking Neural Networks

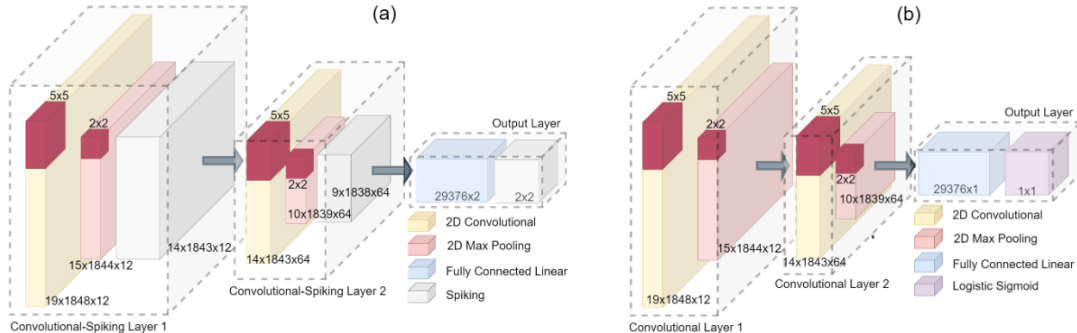


Figure 2. Schematic of neural network architectures. (a) CSNN. (b) CNN.

Le problème de la backpropagation

Expression générale de la backpropagation :

$$\frac{\partial \mathcal{L}}{\partial W} = \frac{\partial \mathcal{L}}{\partial S} \underbrace{\frac{\partial S}{\partial U}}_{\{0, \infty\}} \frac{\partial U}{\partial I} \frac{\partial I}{\partial W}$$

\mathcal{L} : la loss function

W : les poids du réseau

S : fonction qui génère un spike

U : Le potentiel de la membrane

$I = WX$: L'entrée du réseau (input)

Un challenge, plusieurs solutions

Le challenge : La non différentiabilité de sorties spikantes (dead neuron problem)

Les solutions :

Shadow training : Transformer un ANN en SNN

Surrogate Gradient (ou dérivée approchée)

$$\tilde{S}(U) \approx \frac{1}{1 + e^{-(U - U_{thr})}}$$

$$\tilde{S}(U) \approx \frac{U}{1 + k|U|}$$

Gradient général

Le gradient de la loss function générale \mathcal{L} par rapport aux poids W est donné par :

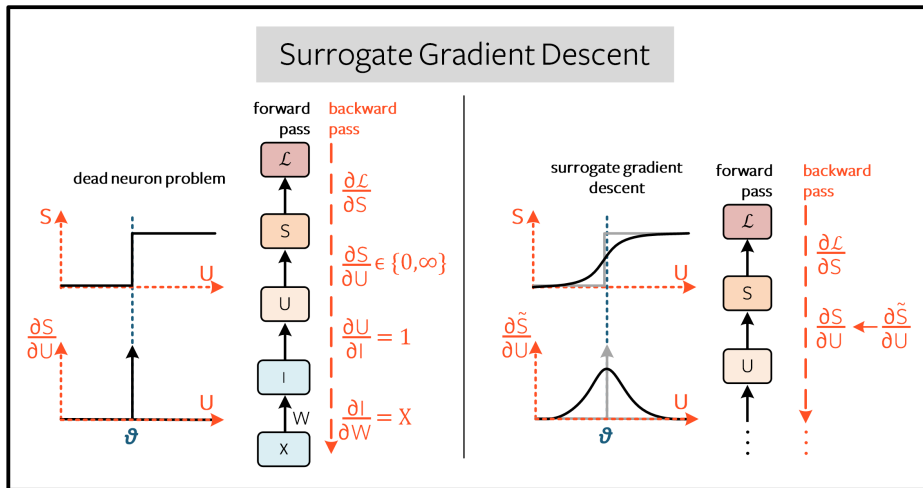
$$\frac{\partial \mathcal{L}}{\partial W} = \sum_t \sum_{s \leq t} \frac{\partial \mathcal{L}[t]}{\partial W[s]}$$

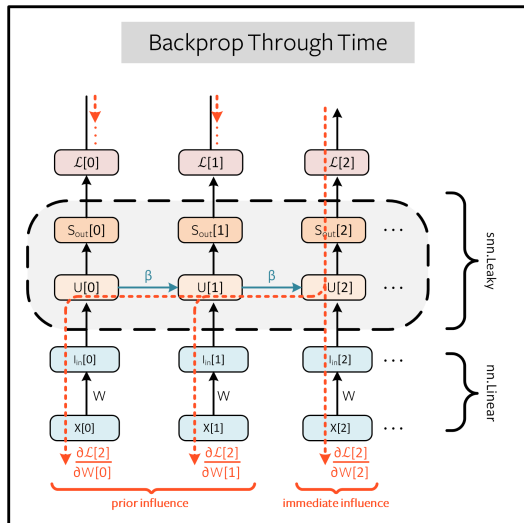
Chain rule

$$\frac{\partial \mathcal{L}[t]}{\partial W[t-1]} = \frac{\partial \mathcal{L}[t]}{\partial S[t]} \frac{\partial \tilde{S}[t]}{\partial U[t]} \frac{\partial U[t]}{\partial U[t-1]} \frac{\partial U[t-1]}{\partial I[t-1]} \frac{\partial I[t-1]}{\partial W[t-1]}$$

Influence du poids précédent

$$\frac{\partial \mathcal{L}[t]}{\partial W[t-1]} = \frac{\partial \mathcal{L}[t]}{\partial S[t]} \frac{\partial \tilde{S}[t]}{\partial U[t]} \cdot \beta \cdot X[t-1]$$





Les spikes sont des événements asynchrones

Les événements temporels à générer

Les poids à ajuster

Tout cela multiplié par le nombre de neurones et la complexité des données

Résultats sur MNIST

Dataset pour la classification sonore

STFT

$$\{x[n]\} \equiv X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}$$

Echelle Mel

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right)$$

DCT

$$S_{u,v} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} s(m, n) e^{-2i\pi \left(\frac{um}{M} + \frac{vn}{N} \right)}$$

Du signal audio vers le MFCC

Résultats sur notre dataset

Guitare

Le fonctionnement des SNNs

Les challenges de l'entraînement

Les résultats

Malgrès la durée d'entraînement, sparsité et efficacité

Utile pour les systèmes embarqués

Applications intéressantes :

Détection de sons

Classification en temps réel