

# Compte-rendu hebdomadaire

## Compte-rendu 1

Valentin Noyé

20 janvier 2025 - 24 janvier 2025

## 1 Problématique

Je suis parti des bases posées par le précédent stagiaire pour comprendre le fonctionnement de l'obscurité par translation de caractéristiques d'une classe à une autre, à partir d'une clé de chiffrement. Pour cela, le principe consiste à utiliser un auto-encodeur entraîné sur la base de données MNIST.

## 2 Travail réalisé

### 2.1 Classifieur

Un classifieur à deux couches a été implémenté afin de donner une mesure objective et sans référence, à l'inverse du PSNR et du SSIM. Ce modèle atteint une précision de 99,29% et une certitude de 99,26% sur MNIST avec 60 000 données d'entraînement et 10 000 de test, ce qui le rend apte à classer les chiffres générés par la suite.

### 2.2 Auto-encodeur

Deux auto-encodeurs ont été implémentés :

- Un auto-encodeur simple utilisant l'entropie croisée binaire comme fonction de perte.
- Un auto-encodeur variationnel utilisant la divergence de Kullback-Leibler comme fonction de perte.

Leurs performances ont été mesurées sur la qualité de la reconstruction des données du dataset MNIST à partir du PSNR, du SSIM ainsi que de la classification définie par la paire (classe, certitude), dans la figure 1.

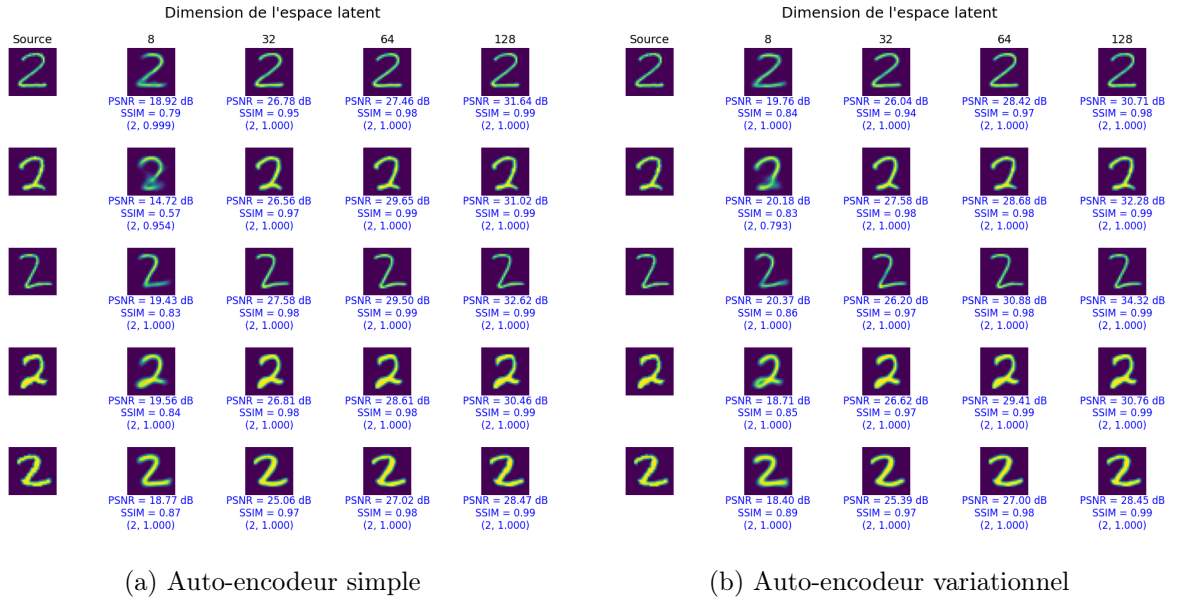


Figure 1: Analyse de la reconstruction par auto-encodeur sur le chiffre 2

Les performances des deux modèles sur ce dataset sont approximativement égales en ce qui concerne la qualité de reconstruction des données. Avec un vecteur latent de taille 128, les valeurs moyennes du PSNR et du SSIM suite à la reconstruction des données tendent respectivement vers 31,55 dB et 0,9913. Quant à la classification, celle-ci classe correctement les images reconstruites 99,52% des fois avec une certitude moyenne de 99,37%, ce qui est relativement élevé.

Taille vec. latent	PSNR (dB)	SSIM	Précision	Certitude
8	19,9448	0,8503	96,69%	98,52%
32	26,6700	0,9726	99,38%	99,33%
64	29,8304	0,9870	99,50%	99,37%
128	31,6039	0,9914	99,52%	99,38%

Table 1: Performances de l'auto-encodeur simple

Taille vec. latent	PSNR (dB)	SSIM	Précision	Certitude
8	20,1281	0,8604	96,54%	98,52%
32	26,9417	0,9738	99,36%	99,30%
64	29,7995	0,9868	99,52%	99,37%
128	31,5282	0,9912	99,53%	99,37%

Table 2: Performances de l'auto-encodeur variationnel

Pour chacune des classes, le vecteur latent moyen est calculé à partir d'un échantillon des données encodées de cette classe. Nous calculons ce vecteur latent moyen par exemple avec 100,00%, 1,00% ou encore 0,02% (soit 2) vecteurs latents d'images encodées parmi 7 000 par classe, et cela sur différentes tailles de vecteurs latents.

Nous en tirons de cela, dans la figure 2 que seulement 1,00% à 10,00% sont nécessaires afin de reconstruire le vecteur latent moyen, quelle que soit la dimension de l'espace latent.

Par ailleurs, une dimension de l'espace latent plus faible fait visuellement apparaître des résultats un peu plus affûtés sur l'image moyenne, même si la précision de classification est rela-

tivement élevée.

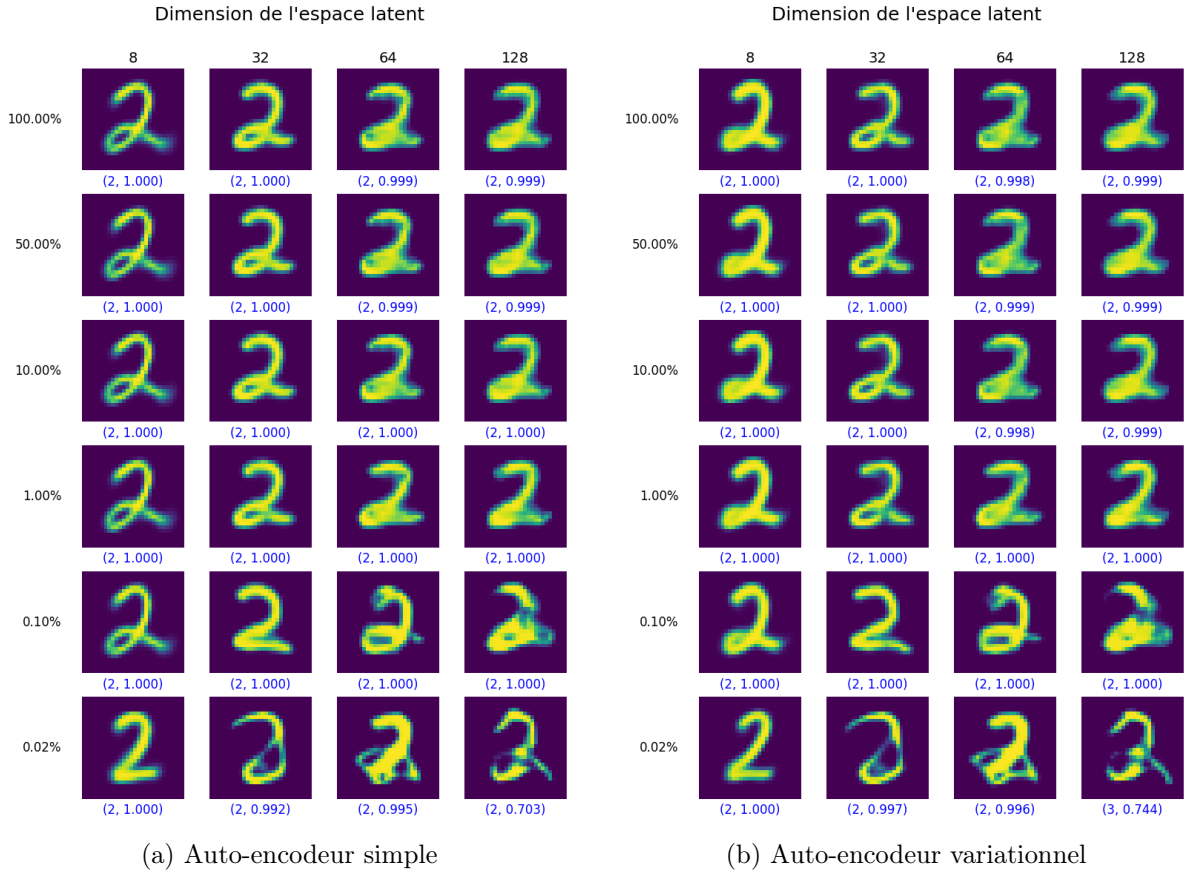


Figure 2: Reconstruction moyenne selon la taille de l'échantillon et du vecteur latent

### 2.3 Projection t-SNE

La projection t-SNE a pour objectif de réduire la dimensionnalité de l'espace latent généré par notre encodeur (ici, celui du VAE), afin de le rendre visualisable tout en préservant les relations locales entre les points de données. Cela explique l'apparition des différents clusters observés dans la figure 3, où chaque cluster représente un chiffre. Ces clusters et les voisins qui les entourent permettent d'établir des correspondances entre les caractéristiques des chiffres.

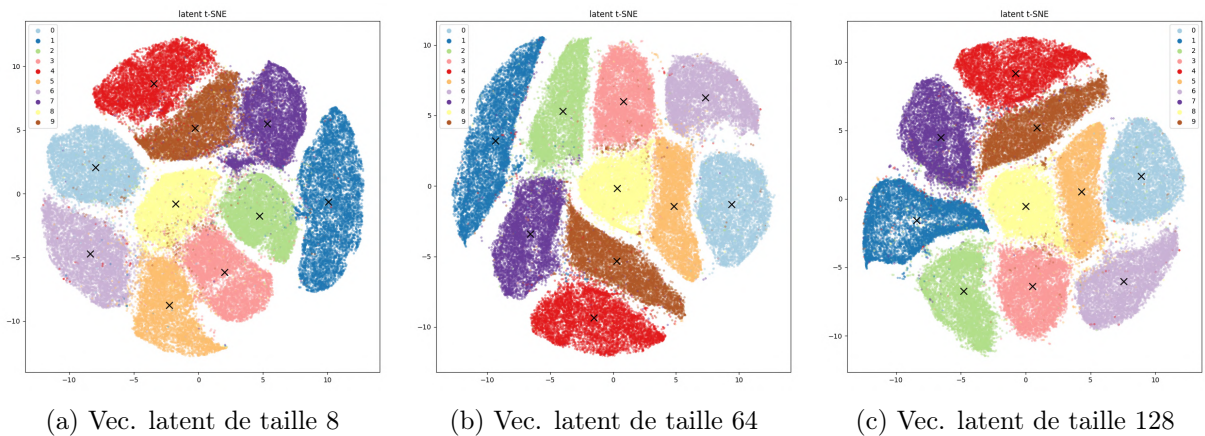


Figure 3: Reconstruction moyenne selon la taille de l'échantillon et du vecteur latent

## 2.4 Translation d'une classe à une autre

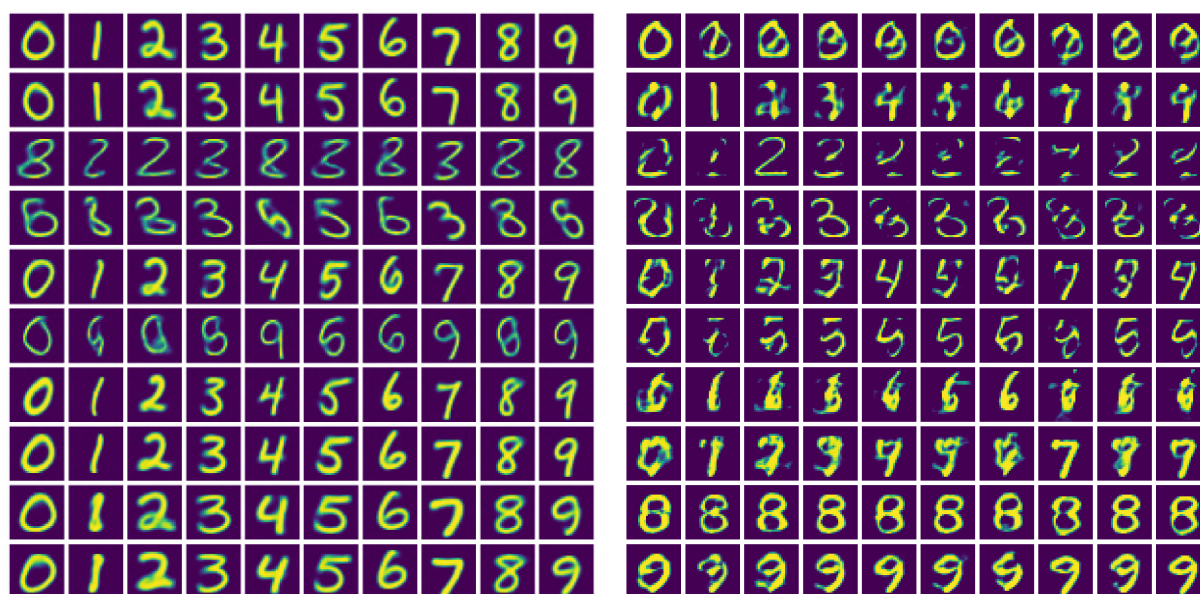
Le calcul de la moyenne des vecteurs latents de chaque classe nous permet d'en tirer le vecteur de translation d'une classe vers une autre. Ce vecteur peut donc être appliqué sur un chiffre de la classe source pour le transformer en un chiffre d'une classe de destination tout en maintenant certaines de ses caractéristiques clé, telles que l'épaisseur, le rayon des arcs, l'inclinaison, etc. Dix chiffres de chaque classe sont sélectionnés arbitrairement.



Nous dressons un tableau des translations de ces chiffres-ci vers une autre classe, dans la figure 4. Chaque ligne représente les translations de son chiffre correspondant (0 à 9) vers la classe de destination (0 à 9) dans chaque colonne.

Par exemple, la 3ème cellule de la seconde ligne correspond à la translation du chiffre 1 vers la classe 3. La diagonale correspond à ces chiffres encodés puis décodés sans translation effective.

Nous constatons que pour des petit vecteurs latents, la translation aboutit la majorité du temps en gardant des caractéristiques du chiffre source, mais nous observons d'autres cas où la translation est incorrecte. La translation se dégrade à mesure que nous augmentons la taille du vecteur latent.



(a) Vecteur latent de 8

(b) Vecteur latent de 128

Figure 4: Table des translations d'une classe à une autre

## 3 Travail à effectuer

- Analyser la variance ou l'écart du centroïde dans la projection t-SNE en fonction de la taille du vecteur latent;
- Implémenter un ré-encodage après un décodage des images pour régulariser le modèle et améliorer la reconstruction;
- Corriger la translation d'une classe vers une autre.

## 4 Activités

### Réunions :

- Réunion d'accueil - Lundi 20/01/25
- Réunion de mise au point - Mardi 21/01/25
- «*Analysis of the Edge Length Distribution of Crypto-Compressed 3D Objects* » par Khelian Larvet - Jeudi 23/01/25
- «*Re-identifiable Image Obscuration for Face Anonymization* » par Norman Hutte - Jeudi 23/01/25