

# **Segment Anything Model Fine-tuning on Electron Microscopic Images of Cells' Mitochondria**

**Amr MOHAMED  
Anh Thu DOAN  
Gwendal AUPHAN**

**ING3 - Option IA - Group 2  
AI Based Image Processing  
06/11/2023**

# Sommaire

## 1. Introduction

## 2. Methods

2.1. Dataset Description

2.2. Data Preparation

2.3. SAM Model

2.3.1. Definition of SAM Model

2.3.2. SAM's features

2.3.3. SAM architecture

2.3.4. SAM model fine tuning

2.4. Optimizer and Loss function

2.5. Model Evaluation

## 3. Results

## 4. Discussion

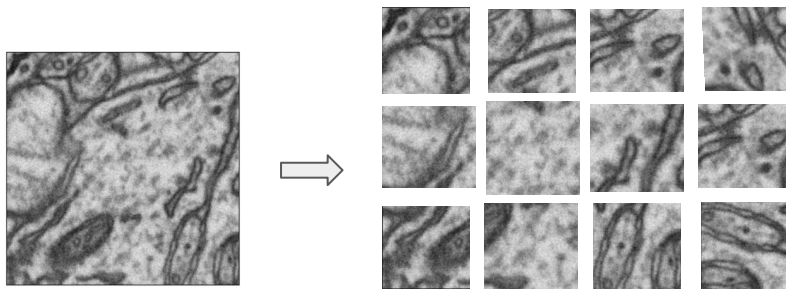
## 5. References

# Introduction

- Motivation : **Difficultés** liées à l'analyse **d'images médicales**, en particulier dans le domaine de la recherche sur les **mitochondries**, en raison de l'importance des mitochondries dans la fonction cellulaire et la recherche sur les maladies.
- Limites de SAM : SAM manque de **précision** dans l'analyse d'images médicales, en particulier avec des **structures complexes** comme les mitochondries.
- Nécessité de fine-tune: Il est essentiel d'ajuster le modèle pour répondre aux exigences spécifiques de ce domaine.

# Methods

- Dataset: Le dataset représente une section de 5x5x5μm prélevée dans la région de l'hippocampe CA1 du cerveau.
- Data preparation:
  - **Augmenter la diversité** de l'ensemble de dataset. Diviser une image en patchs (Patch d'image). Passant ainsi de **165** images de taille (768, 1024) à **1980** images de taille (256, 256)



- Rescale: **Mask [0,1]**: Aligne les valeurs du masque sur la tâche de segmentation, **0** représentant **background** et **1** signifiant les **mitochondries** d'intérêt.
- Supprimer les patchs qui ne contient pas les mitochondries (1980 → 1642)

# Qu'est-ce que Segment Anything ?

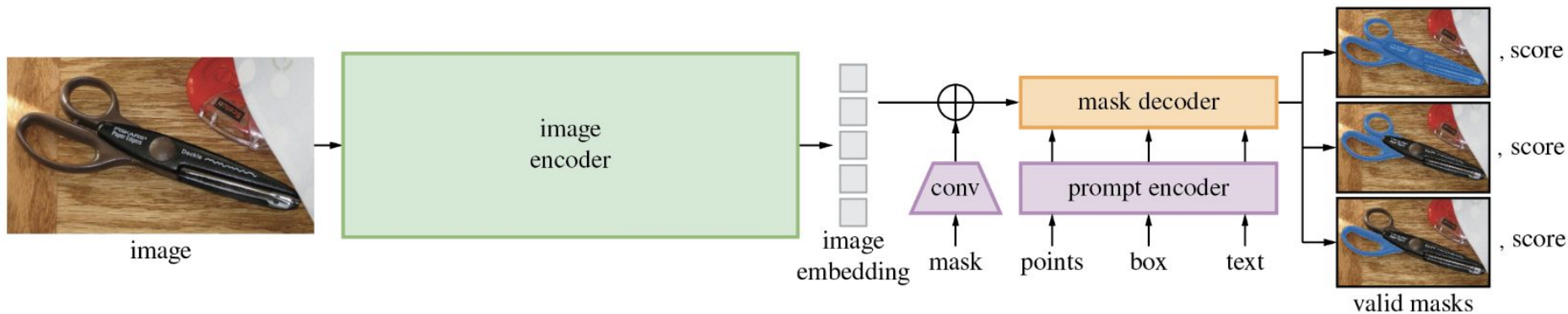
- Segment Anything Model (SAM) : un nouveau modèle d'IA de Meta AI qui peut "**découper**" n'importe quel objet, dans n'importe quelle image.
- Il a été entraîné sur **11 millions** d'images, **1B+ masques**.
- Il est conçu pour accepter des **prompt de humaines**, sous forme de **points**, de **boîtes** de délimitation ou même de **texte** décrivant ce qui doit être segmenté.



# Caractéristiques de SAM

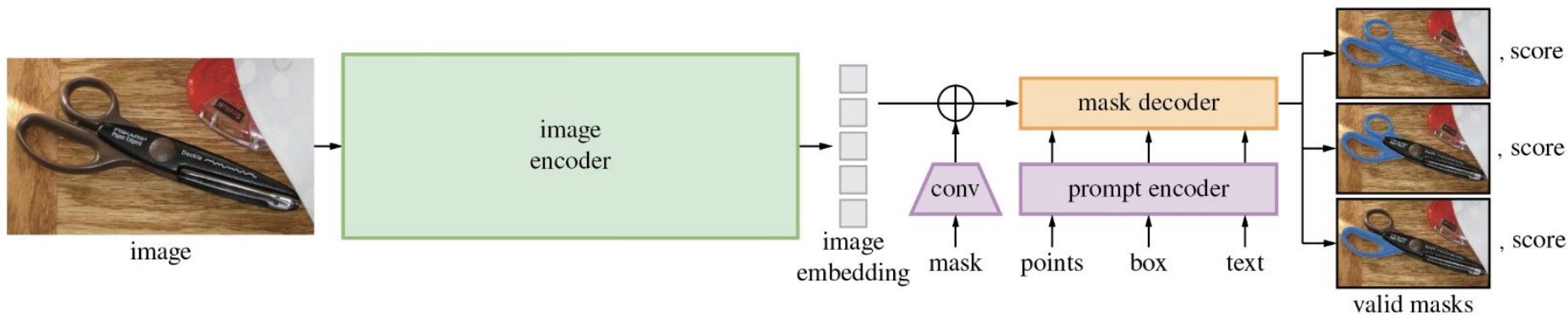
- **Zero-shot generalization:** SAM peut segmenter des objets qu'il **n'a jamais vus** auparavant, sans nécessiter de train supplémentaire.
- **Flexible prompting:** SAM peut être sollicité par une **variété d'entrées**, y compris des points, des boîtes et des descriptions textuelles.
- **Real-time mask computation:** SAM peut générer des masques pour les objets en **temps réel**. SAM est donc idéal pour les applications où il est nécessaire de **segmenter rapidement** les objets, comme la conduite autonome et la robotique.
- **Ambiguity awareness:** SAM est conscient de **l'ambiguïté** des objets dans les images. Cela signifie que SAM peut générer des masques pour des objets même lorsqu'ils sont **partiellement cachés** ou qu'ils se **overlap** avec d'autres objets.

# Model Architecture



- Image Encoder : **Encodage de l'image** dans une représentation vectorielle à haute dimension (**Image embedding** par ViT).
- Prompt Encoder : Encodage de l'invite (**prompt**) à partir d'une **représentation spatiale** en une **représentation vectorielle** à **additionner** avec l'**image embedding**
- Mask Decoder : **génère un masque** à l'aide de le **prompt encoder** et l' **image embedding**, Il mis a jour l'embedding par une mécanisme de **self et cross attention**. Après, un MLP à partir de l'embedding pour **sortir un likelihood** (**linear classifier**).

# Model Fine Tuning



- Freeze: l' image encoder et le prompt encoder
- Mask decoder: autoriser la mise à jour des poids.



# Optimizer et Loss function

- Optimizer: Adam
  - Learning rate: 1e-5
  - Weight decay: 0
  - Effectuer des **ajustements minimaux** tout en **préservant les connaissances pré-entraînées**. Cela permet des mises à jour subtiles des paramètres sans régularisation supplémentaire de la **weight decay**.
- Loss function: Dice + Cross Entropy Loss
  - Garantit que le modèle excelle dans la **précision au niveau des pixels** et dans la **séparation des objets**, en trouvant un équilibre entre la classification des pixels et la délimitation des objets.

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))$$

CE Loss

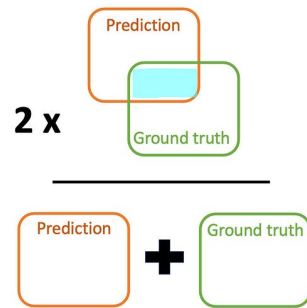
$$\mathbf{L}_{dice} = \frac{2 * \sum p_{true} * p_{pred}}{\sum p_{true}^2 + \sum p_{pred}^2 + \epsilon}$$

Dice Loss

# Evaluation

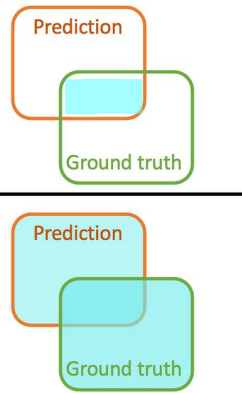
- Dice coefficient:

$$\text{Dice} = \frac{2 \times \text{Area of overlap}}{\text{Total area}} =$$



- IoU score:

$$\text{IoU} = \frac{\text{Area of overlap}}{\text{Area of union}} =$$

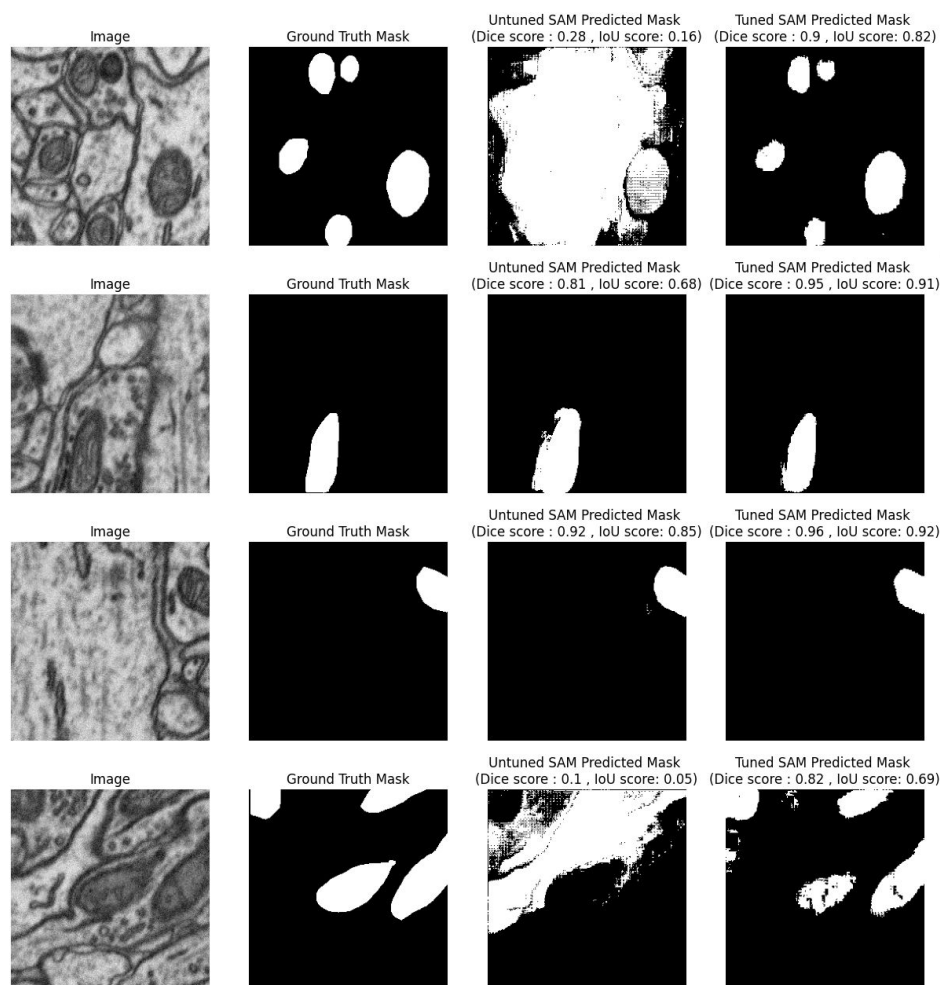


# Results

- Dice coefficient:
  - Base model:  $\sim 0.55$
  - Fine tuned model:  $\sim 0.85$
- IoU score:
  - Base model:  $\sim 0.47$
  - Fined tuned model:  $\sim 0.75$

# Results

- Capacité accrue du modèle fine-tuned à segmenter des images avec *plusieurs mitochondries*.



# Discussion

- La capacité de prédiction de segmentation du modèle SAM s'est considérablement améliorée après le “**fine-tuning**”, avec des améliorations notables des performances.
- Le score de Dice est passé de 0,55 en moyenne dans le modèle non ajusté à environ **0,85** dans le modèle ajusté, indiquant une amélioration considérable dans la segmentation des images cellulaires. Ceci correspond aux **positions réelles** des mitochondries dans les images cellulaires, démontrant une bonne précision dans la segmentation.
- Le score IoU est passé de 0,47 dans le modèle non ajusté à 0,764 dans le modèle ajusté, indiquant qu'environ **76,4 %** des masques prédits correspondent étroitement aux positions et formes des mitochondries.

# Discussion

- Les comparaisons visuelles des images de test ont montré que le modèle ajusté montrait de bien meilleurs résultats que le modèle de base, en particulier dans les images contenant **plusieurs mitochondries** à segmenter.
- Une segmentation précise des structures cellulaires a des implications significatives pour les applications scientifiques et médicales, **réduisant l'intervention manuelle** dans des tâches telles que la biologie cellulaire et le diagnostic de maladies.
- L'amélioration des performances du modèle SAM ajusté peut **économiser du temps** pour les spécialistes du domaine qui effectuent habituellement des annotations et des validations manuelles dans ces tâches.
- Les principales limitations de l'étude étaient liées aux ressources informatiques, suggérant des travaux futurs afin d'explorer l'impact des différents composants du modèle sur la puissance prédictive.

# References

- [1] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). "Segment Anything." *arXiv preprint, arXiv:2304.02643 [cs.CV]*. [Link](#)
- [2] Sørensen, T. (1948). "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons." *Kongelige Danske Videnskabernes Selskab*, 5(4), 1–34.
- [3] Dice, L. R. (1945). "Measures of the Amount of Ecologic Association Between Species." *Ecology*, 26(3), 297–302. doi:10.2307/1932409. JSTOR 1932409. S2CID 53335638.
- [4] Jaccard index. (2023, September 16). In Wikipedia. Retrieved from [Link](#)
- [5] "Group and Crowd Behavior for Computer Vision." (2020). Shuyu Sun and Tao Zhang. ISBN 978-0-12-820957-8. Gulf Professional Publishing. Retrieved from [Link](#)

