

# Single image to physically accurate origami pattern

Olivier Schwander

2024-2025

## Localisation et encadrement

Localisation :

— Équipe MLIA, ISIR, Sorbonne Université

Encadrants :

— Olivier Schwander <olivier.schwander@isir.upmc.fr>

## Contexte et objectifs

L'objectif du stage est de concevoir un modèle capable de générer le patron de pliage d'un origami à partir d'une image d'objet. Ce patron devra être pliable et donner un pliage ressemblant à l'objet sur l'image d'entrée. Pour la sortie, une possibilité est d'utiliser le format *crease pattern*, comme présenté Figure 1.

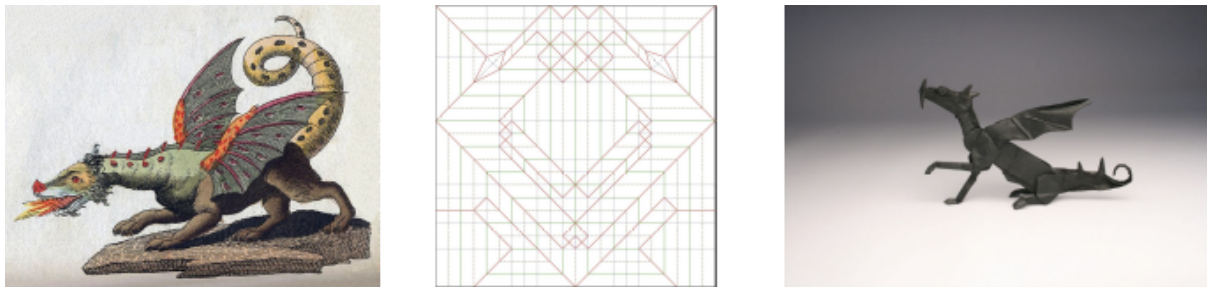


FIGURE 1 – Un exemple d'image d'entrée, de patron générée et de pliage final possible, à partir du site <https://langorigami.com/crease-patterns/>

En effet, les progrès récents des modèles multi-modaux ont permis de traiter des tâches complexes de compréhension et de génération de textes ou d'images associés. Une des difficultés à résoudre reste l'exactitude des contenus générés, par exemple les hallucinations avec des textes qui inventent des faits ou des images avec des parties irréalistes. Figure 2, on voit un patron d'origami généré par ChatGPT et Dall-E à partir d'une entrée textuelle, le dessin est crédible mais rien ne garantit que le pliage est possible et qu'il conduit à l'image associée.

On se propose dans ce stage d'étudier ce problème sous l'angle de l'exactitude physique des contenus générés. Au travers de la tâche de génération de patrons d'origami à partir d'images, on cherche à contraindre le modèle à générer une sortie réaliste au sens physique. Dans le cas des origamis, il s'agira de construire un patron qui puisse effectivement conduire à un pliage ressemblant à l'image de départ. Une approche possible est d'utiliser de l'apprentissage par renforcement en exploitant un simulateur (Ghassaei, Demaine, and Gershenfeld 2018) permettant de valider le patron. Une autre approche possible est de s'inspirer du RAG (Retrieval Augmented Generation) et d'utiliser une base de données de patrons valides comme par exemple dans (Lim and Shim 2024).

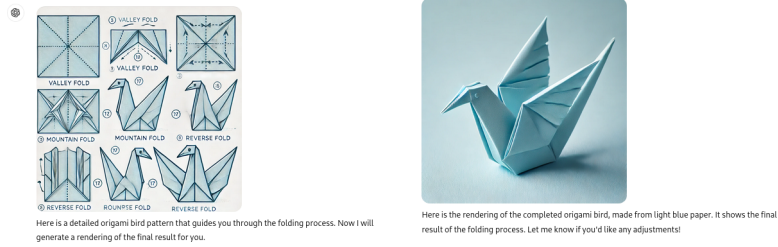


FIGURE 2 – Exemple de patron et de rendu avec ChatGPT et Dall-E avec une entrée textuelle. Même si visuellement les images semblent réussies, rien n’indique que le patron est physiquement correct.

Au-delà des origamis, ce travail permettra de réfléchir à la génération de contenu physiquement exact, par exemple pour des dessins techniques ou des corps humains plus réalistes.

## Bibliographie

- Geiger, Jeremia, Karolis Martinkus, Oliver Richter, and Roger Wattenhofer. 2023. “Automating Rigid Origami Design.” April 28, 2023. <https://doi.org/10.48550/arXiv.2211.13219>.
- Ghassaei, Amanda, E. Demaine, and N. Gershenfeld. 2018. “Fast , Interactive Origami Simulation Using GPU Computation.” In. <https://www.semanticscholar.org/paper/Fast-%2C-Interactive-Origami-Simulation-using-GPU-Ghassaei-Demaine/3d23781cba3a48e576d22742575ac9da05c06b8e>.
- Lim, Youngsun, and Hyunjung Shim. 2024. “Addressing Image Hallucination in Text-to-Image Generation Through Factual Image Retrieval.” arXiv.org. July 15, 2024. <https://arxiv.org/abs/2407.10683v1>.
- Liu, Lijuan, Xiangyu Xu, Zhijie Lin, Jiabin Liang, and Shuicheng Yan. 2023. “Towards Garment Sewing Pattern Reconstruction from a Single Image.” November 7, 2023. <https://doi.org/10.48550/arXiv.2311.04218>.
- Nakagaito, Chiaki, Takanori Nishino, and Kazuya Takeda. 2020. “Generation of Origami Folding Animations from 3D Point Cloud Using Latent Space Interpolation.” In *SIGGRAPH Asia 2020 Posters*, 1–2. SA '20. New York, NY, USA : Association for Computing Machinery. <https://doi.org/10.1145/3415264.3425450>.