

# Workflow en IA générative comme substitue à la rotoscopie

---



**Eloi Champagne**  
Stratège en technologies  
de production

# But des tests

Explorer certains outils liés à l'intelligence artificielle générative pour voir s'il serait possible de simplifier et accélérer (voire remplacer) la rotoscopie pour un projet documentaire ayant des séquences animées.

# Les outils IA testés

- [EbSynth](#)
- [Stable Video Diffusion](#)
- [Runway Research Gen1 et Gen2](#)
- [ComfyUI](#) (stable diffusion avec « custom workflows »)

# EbSynth

## Constat

- Pas de l'IA (Frame Interpolation)
- Une option possible, pas idéale
- EbSynth donne un résultat très "CG" (déplacement des pixels)
- Demande beaucoup d'interventions de la part d'un artiste (chaque fois que l'image change un peu trop, voir exemple ci-joint)
- Gratuit!



# Stable Diffusion (Video)

## Constat

- Stable Video Diffusion est limité à 14 ou 25 frames d'animation, après l'image se transforme.
- Beaucoup de « flickering »
- Manque de contrôle.



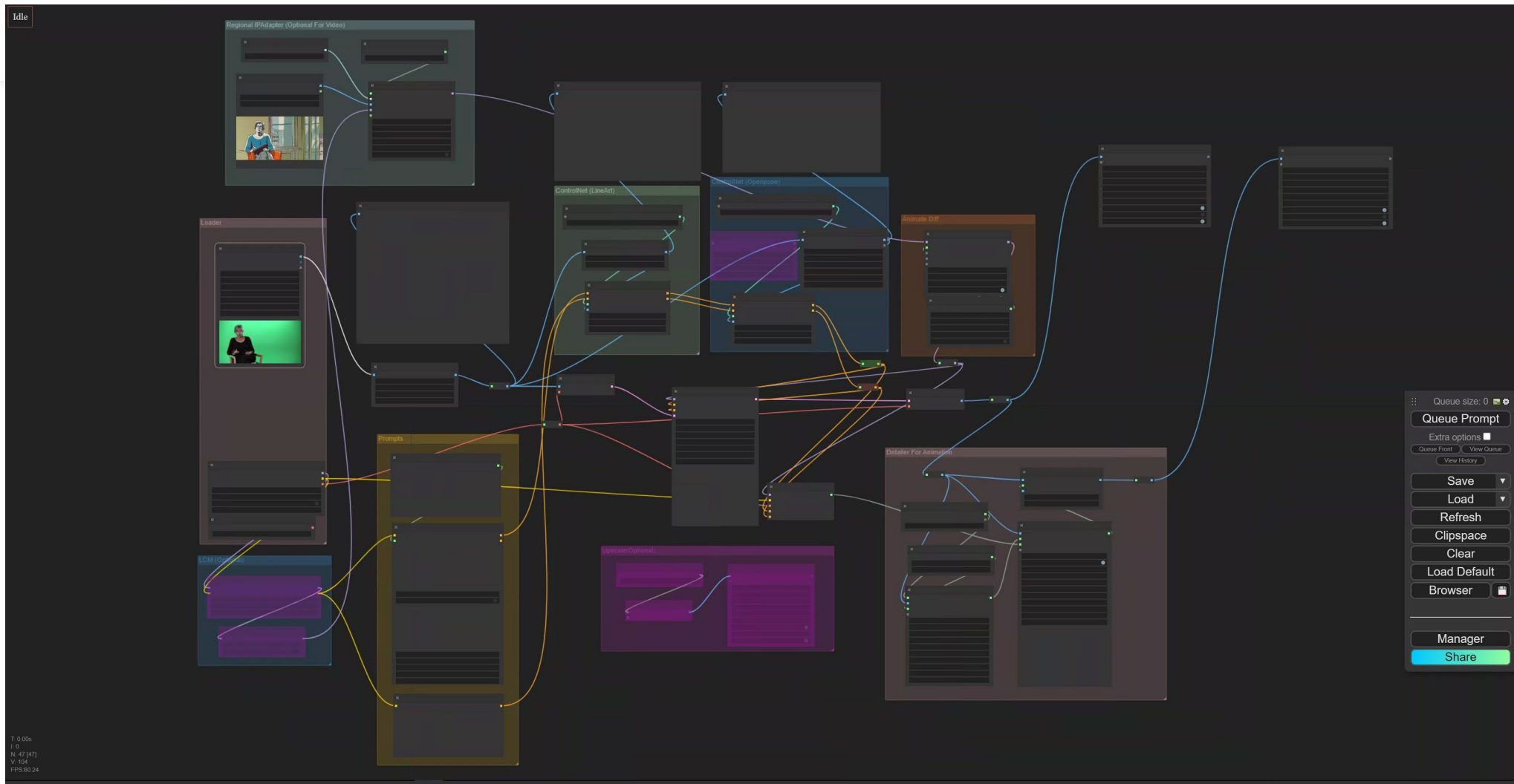


# Runway Research Gen1/Gen2

## Constat

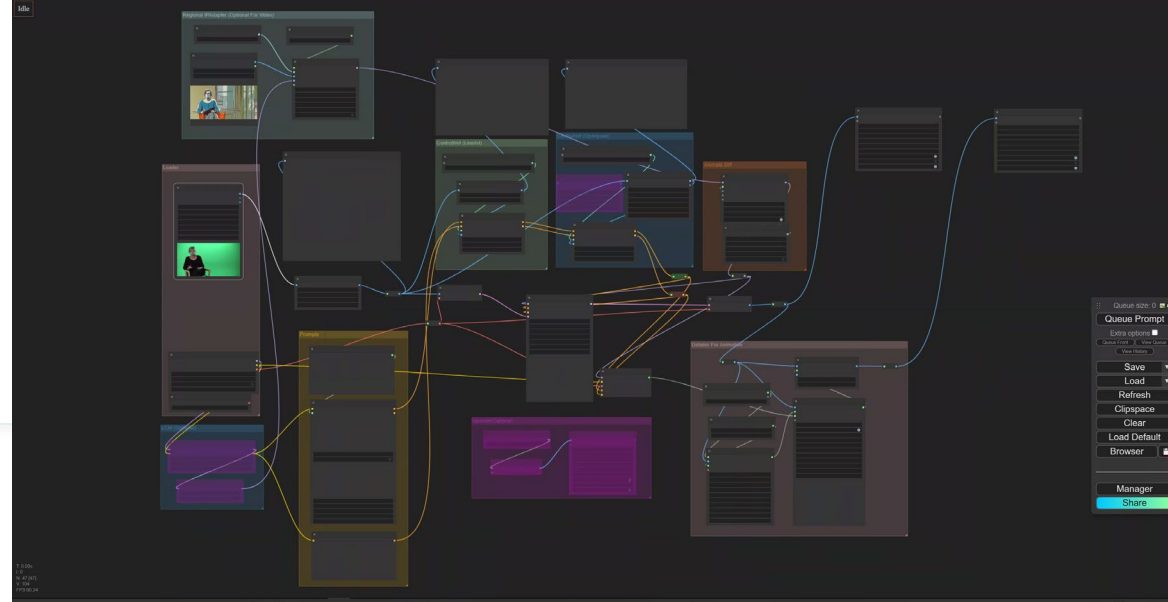
- Limité à 4 secondes de vidéo.
- Beaucoup de « flickering »
- Très peu de contrôle de l'animation et de la qualité des résultats (prompt seulement)
- Le lip sync est un problème
- Demande un abonnement (\$)

# ComfyUI





# ComfyUI



## Constat

- Comme Stable Video Diffusion (qui est le modèle de base), limité à la base à 14 ou 25 frames d'animation avant que l'image se transforme. Cependant, nos tests indiquent qu'il est possible de contourner cette limitation grâce à des outils donnant plus de contrôle.
- Encore un peu de « flickering », mais controlable et peut être corrigé en post.
- Permet la création de “Workflow” custom, donnant plus de contrôle et de possibilités artistiques avec ControlNet, IPA et LoRAs.
- Les outils “ControlNet” permettent plus de control, et ouvrent pleins de possibilités pour aider au compositing ou créer des effets.
- Pas de problème avec le lip sync ou les mouvements
- Très puissant, avec beaucoup de possibilités et de configuration possible, ce qui veut aussi dire que cela peut prendre beaucoup de temps pour trouver les bonnes options.

# ComfyUI (suite)

## Image de référence de style vs prompt

Une seule image de référence pour le style n'est pas suffisante. Il faut entraîner un checkpoint, ou à tout le moins un LoRA\* sur une banque d'images de références suffisante.

Utiliser des prompts seulement, ou un mélange prompts et image de référence est, à court terme, la meilleure option en combinaison avec des outils ControlNet et des LoRA.

\*LoRA: Low-Rank Adaptation of Large Language Models

# ControlNet

ControlNet est un type de modèle pour contrôler les modèles de diffusion d'image en conditionnant le modèle avec une image d'entrée supplémentaire.

## Quelques Exemples (30 secondes)



# Conclusion

La solution ComfyUI, bien qu'imparfaite, est une bonne base de départ. Elle peut permettre de sauver un temps considérable qui serait passé à faire de la rotoscopie à la main. L'intervention d'un animateur/artiste reste nécessaire cependant pour faire des corrections et ajustements à la main dans un logiciel d'animation "raster" (tel que TVPaint ou Photoshop).

Un artiste de compositing sera possiblement aussi nécessaire. Surtout s'il y a tournage devant écran vert ou si intégration de l'animation à un décor rendu séparément.

# Prochaines étapes

- Continuer l'exploration avec ComfyUI
  - Entraîner un LoRA avec un style particulier
  - Ajouter plus de fonctions de ControlNet
- Faire un test de workflow complet avec une séquence éditée.
- Tester les fonctions « face swap » au besoin.

## **Autres outils en développement à explorer éventuellement**

- [Boximator](#): (Rich and Controllable Motions for Video Synthesis)
- [ConsistI2V](#) (Papier de l'université de Waterloo)
- [SORA](#) de OpenAI

**Merci !**

**onf.ca**