

FRUISE

Guide gratuit · École Fruise

Ta première vidéo IA en 7 étapes – le guide gratuit

La vue d'ensemble de la méthode SerialKiwix, nos vraies astuces et un plan d'action pour un week-end

Par l'équipe de SerialKiwix, la série
fabriquée avec l'IA.
fruisse.com/ecole · octobre 2026



Ce guide, c'est la carte. Il te montre tout le chemin, de l'idée à la vidéo publiée, avec les astuces qui nous ont fait gagner le plus de temps et un plan d'action pour fabriquer ta première vidéo en un week-end.

Il vient d'une expérience bien réelle : **SerialKiwix**, notre série d'animation IA. Kiwix y est le barman souriant du bar à jus « Pressé ! » le jour, et la nuit il presse les fruits qui l'ont mérité (jamais de sang : uniquement du jus, de la pulpe et des pépins). Six épisodes publiés sur TikTok, Instagram, Facebook et YouTube, fabriqués par KALEXEX avec Claude. Le cinquième, « La douche », a été fait **entièrement en local**, sur un PC : 60 images et 56 vidéos, environ 67 minutes de calcul sur une RTX 5080 (16 Go), 0 crédit cloud, pour 2 min 07 et 42 plans.

Tout ce qui suit est tiré de cette production : nos réglages, nos prompts, nos chiffres et, surtout, nos erreurs.

Avant de commencer : ton matériel, ton budget, tes licences

Il y a deux chemins, et un mélange des deux.

Scénario	Matériel	Coût mensuel indicatif	Pour qui
0 €	PC avec carte graphique ≥ 12-16 Go de VRAM déjà possédé	électricité seulement	tout en local : WanGP (Qwen Image, MiniMax H3), voix OmniVoice, musiques libres
≈ 25-35 €/mois	même PC	abonnement voix/musique (ElevenLabs Creator, prix à vérifier)	local + voix expressives et bruitages
≈ 100-200 €/mois	n'importe quel ordinateur	crédits vidéo cloud (Higgsfield, Kling, Veo via fal.ai : ≈ 0,3-3 \$ par clip) + voix	sans carte graphique, ou pour quelques plans clés

QUE FAIT TOURNER TA VRAM ?

Pour l'IA en local, ce qui compte, c'est la mémoire de la carte graphique (la VRAM). Ce qu'on a mesuré, et ce que dit la doc, palier par palier. Jauges à l'échelle d'une carte de 32 Go.

8 Go

POUR COMMENCER

doc de WanGP · non testé par nous

- « 8 Go suffisent pour commencer » (README de WanGP)
- MiniMax H3 en 480p : 15 s de vidéo dans **≈ 5-6 Go** (doc WanGP)
- Plus lent que nos temps : mesure sur 2-3 plans d'abord.

12 Go

JOUABLE

un chiffre mesuré, un de la doc

- Images Krea 2 Turbo : pic **≈ 11 Go** en 768x768 (mesuré sur notre PC)
- MiniMax H3 en 1080p : **≈ 11 Go** pour 15 s (doc WanGP)
- Profil mémoire 4 de WanGP, et reste en 480p.

16 Go

NOTRE MACHINE

RTX 5080 · mesuré

- Qwen Image 2.1 + Pruna 8 pas, 864x1536 : **≈ 20 s / image**
- MiniMax H3 + Pika merge lite, 480x832 : **≈ 50 s / plan** — le modèle pèse 20,6 Go : WanGP l'envoie par morceaux depuis la RAM (64 Go chez nous)
- FastH3 Omni (plusieurs personnages) : **≈ 1 min 40 / plan**
- Un épisode entier (E05) : **≈ 67 min**

Cartes de 16 Go : **RTX 5080** à partir de 1 099 €, **RTX 5070 Ti** à partir de 1 243 € (relevé du 3/08/2026, à vérifier)

24 Go +

PLUS DE MARGE

non testé par nous

- Résolutions plus hautes, plans plus longs, moins d'attente.
- La méthode ne change pas : le 480p reste le plus rapide.

RTX 5090 (32 Go) à partir de 3 899 € (relevé du 3/08/2026, à vérifier)

AVANT D'ACHETER

Tape **nvidia-smi** : la ligne « MiB » donne ta VRAM (16 303 MiB = 16 Go chez nous).

NOTRE CONSEIL

Vise **12 à 16 Go** pour une série. Et ferme les autres logiciels : ils grignotent la VRAM.

Les chiffres de la doc de WanGP (relevés le 31/10/2026) sont ceux de son auteur, avec ses réglages d'économie : des planchers, pas une promesse de confort.

Sources : nos mesures sur RTX 5080, README de WanGP, prix relevés par DropReference.

fruisse.com

Nos repères : un épisode d'une minute trente fabriqué tout en cloud nous coûtait environ **200 crédits Higgsfield** ; l'épisode 5, tout en local, **0 crédit**. Depuis, le local est notre méthode par défaut, et le cloud sert à deux ou trois plans clés quand il y a des crédits.

⚠ Piège : les licences (lues le 31/10/2026). « Gratuit à télécharger » ne veut pas dire « libre d'usage ». La licence **MiniMax H3 Community** (qui couvre aussi Pika H3 Flex et son LoRA « merge lite ») **exclut de son territoire l'Union européenne**, le Royaume-Uni, la Corée du Sud et les États-Unis ; **Qwen Image 2.1** est sous licence de recherche **non commerciale** ; le modèle de voix **OmniVoice** est en **CC-BY-NC** (pas d'usage commercial, pas de clonage sans autorisation). Alternatives : **LTX-2.5** (vidéo et son, usage commercial gratuit sous 10 M\$ de chiffre d'affaires annuel, pas encore testé par nous) ou les **services cloud**, dont les conditions commerciales sont à lire chez chacun. Nous sommes en France : on fait le point sur notre propre chaîne. Lis toujours la licence en entier avant de bâtir la tienne.

Les 7 étapes en un coup d'œil

ÉCOLE FRUISE · MÉTHODE

LA CHAÎNE EN 7 ÉTAPES

De l'idée à la vidéo publiée : la méthode qu'on a rodée épisode après épisode sur SerialKiwix.

1

💡 IDÉE

gratuit

Le cadre : thème, ton, public, limites, plateforme (9:16).

SerialKiwix « Un kiwi barman le jour, tueur de fruits la nuit. Jamais de sang : du jus. »

2

🔍 PERSONNAGE

quelques minutes

Fiche, bloc canonique en anglais, image de référence, voix.

SerialKiwix Kiwix : gros yeux verts, moustache en guidon, fourrure brun doré.

3

📄 SCRIPT

gratuit

Chaîne « cause → effet », plans de 2 à 4 s, accroche, cliffhanger.

SerialKiwix E05 « La douche » : 42 plans, ouverture sur le climax.

4

🖼️ IMAGES CLÉS

≈ 20 s / image

Décor maître → plan large → gros plans.

En local : Qwen Image 2.1 + Pruna 8 pas

SerialKiwix La salle de bain de Pêche sert de référence à tous ses plans.

5

🎬 ANIMATION

≈ 50 s / plan

Une image de départ + l'action décrite ; plans parlants avec la réplique.

En local : MiniMax H3 + LoRA Pika merge lite, 480x832

SerialKiwix Kiwix grimpe le long de l'écharpe jusqu'à la fenêtre.

6

🎤 VOIX & SON

quelques centimes

Répliques, bruitages et musiques, calés sur l'image.

Cloud : ElevenLabs v3 (émotions : [whispers], [sighs]...)

SerialKiwix Le « pop » de la bulle de Mangue tombe pile quand elle éclate.

7

✂️ MONTAGE & PUBLICATION

1-2 min

Montage écrit en Python + ffmpeg : 720x1280, sous-titres, -14 LUFS. Puis miniature 3:4, légende, publication 18 h-21 h.

SerialKiwix « SerialKiwix · Épisode N » + une question à choix + 5 hashtags fixes.

Bilan de l'épisode 5, fait 100 % en local (29/10/2026) :

60

images

56

vidéos

≈ 67

min de
machine

0

crédit cloud

Source : notre chaîne de production et le retour d'expérience de l'E05.

fruisse.com

#	Étape	Ce que tu produis	Outil (local / cloud)	Ordre de grandeur
1	Idée et format	une phrase, des limites, un format	ta tête, un carnet	30 min
2	Personnage et fiche	fiche, bloc canonique, image de référence, voix	Qwen Image 2.1 / Higgsfield ; ElevenLabs	≈ 20 s par image en local
3	Script et découpage	10 à 15 plans de 2 à 4 s, chaîne logique	un document	gratuit
4	Images clés	une image par plan, relue en planche	Qwen Image 2.1 + Pruna 8 pas / nano_banana_pro	≈ 20 s par image en local ; 2 crédits dans le cloud
5	Animation	un clip par plan	MiniMax H3 + LoRA Pika / Grok, Kling...	≈ 50 s par plan en local
6	Voix, son, musique	répliques, plans parlants, bruitages, musique	ElevenLabs, outils locaux	quelques minutes
7	Montage et publication	la vidéo, la miniature, la légende	ffmpeg ou ton logiciel de montage	une à deux heures la première fois

Ce qui saute aux yeux : **la génération n'arrive qu'à l'étape 4**. Sur l'épisode 5, environ 40 % du temps machine a servi à corriger des erreurs de logique qu'un découpage mieux pensé aurait évitées. Les trois premières étapes ne coûtent rien et économisent presque tout.

Étape 1 – L'idée et le format

Une idée tient en une phrase. SerialKiwiX : « Le jour, il vous sert un jus. La nuit, il vous presse. »

Un personnage, un contraste, une promesse. Si tu ne peux pas résumer ta vidéo en une phrase qu'on a envie de répéter, ce n'est pas encore une idée.

Vise petit pour la première. Un personnage principal (deux au maximum), un ou deux lieux, une action simple et visible, une chute claire. Pas de texte à lire dans l'image : les modèles écrivent mal.

Le format : vertical 9:16 (720×1280 suffit pour TikTok, Reels et Shorts), **30 à 60 secondes**, découpées en **plans de 2 à 4 secondes**. Pour comparaison, nos épisodes 1 à 5 durent de 1 min 24 à 2 min 31, avec 33 à 47 plans chacun.

Écris tes limites. Les nôtres : le corps est le fruit entier avec de petits bras et de petites jambes, jamais de sang, des armes qui sont des ustensiles de cuisine, une police fictive (la BCM), rien de sexualisé. Elles donnent une identité, créent l'humour (le cliché d'horreur qui finit en

smoothie) et t'évitent les filtres des services en ligne.

💡 Astuce : le hook se joue dans les deux premières secondes. Sur TikTok, une vidéo est d'abord montrée à un « lot test » de spectateurs (environ 150 à 200 vues pour nos premières vidéos, fin octobre 2026). Notre épisode 2, ouvert sur un néon de discothèque et un meurtre spectaculaire, a fait environ 13 % de j'aime en une heure ; l'épisode 1, ouvert au bar, environ 3 % (relevé du 29/10/2026). Ouvre sur ton moment fort, en plan serré : ton personnage occupe au moins un tiers de la hauteur dès la première seconde.

Étape 2 – Le personnage et sa fiche

L'IA n'a aucune mémoire : chaque image repart de zéro. Si tu décris ton personnage avec des mots un peu différents à chaque fois, tu obtiens un personnage un peu différent à chaque plan. La parade tient en deux documents.

La fiche (en français, pour toi) : personnalité, tic (Kiwix essuie toujours le même verre en parlant), façon de parler, phrase signature (« Aucun pépin. »), échelle, tenues, voix.

Le bloc canonique (en anglais, pour l'IA), recopié mot pour mot dans chaque prompt. Celui de Kiwix :

```
an anthropomorphic kiwi fruit: egg-shaped body covered in scruffy matted  
golden-brown kiwi fur, large bulging round eyes with bright green irises,  
heavy furrowed brows, thick black handlebar mustache, wide unsettling toothy  
grin, thin furry brown arms, stubby furry legs; warm golden-brown fur,  
not olive, no beard
```

ANATOMIE D'UN BLOC CANONIQUE

Une description en anglais, copiée **mot pour mot** dans chaque prompt : c'est elle qui garde Kiwix identique d'un plan à l'autre.



- 1 CORPS**
La forme exacte : le fruit entier, petits bras, petites jambes. Jamais un corps humain.
- 2 YEUX**
Taille, forme, couleur : ce que l'IA change en premier.
- 3 SIGNES DISTINCTIFS**
Ce qui le rend reconnaissable en une seconde.
- 4 COULEURS**
Écrites noir sur blanc, nuance comprise.
- 5 TENUE**
Copiée elle aussi, mot pour mot.
- 6 INTERDITS**
Ce que l'IA ajoute toute seule si on ne dit rien.

LE VRAI BLOC DE KIWIX

tenue de nuit

an anthropomorphic kiwi fruit: egg-shaped body covered in scruffy matted golden-brown kiwi fur ①, large bulging round eyes with bright green irises, heavy furrowed brows ②, thick black handlebar mustache, wide unsettling toothy grin ③, thin furry brown arms, stubby furry legs ①; warm golden-brown fur ④, not olive, no beard ⑥, wearing a worn torn beige trench coat, black fedora hat, white latex gloves ⑤

Tenue de jour, au bar « Pressé ! » : lime-green apron over a white short-sleeved shirt, small dark green bow tie

- A** Copie le bloc **mot pour mot** dans chaque prompt, tenue comprise. Ne le réécrits jamais « à peu près ».
- B** Ajoute 2-3 images de référence et « **exact same face, colors and eye size as <image2>** ».
- C** Pas d'adjectif extrême seul (« **creepy** ») : décris l'expression, comme « **wide unsettling toothy grin** ».

Source : fiche de Kiwix et prompts réels des épisodes SerialKiwix.

fruisse.com

Trois astuces qui valent de l'or :

- **Termine par des garde-fous.** « Not olive, no beard » : le modèle a déjà donné à Kiwix une fourrure vert olive et une barbe. Chaque défaut qui revient devient un interdit écrit.

- **Bannis les adjectifs d'humeur.** « Creepy » ou « psychopath » font exagérer le modèle, différemment à chaque image. Décris ce qu'on voit : « wide unsettling toothy grin ».
- **Une image de référence propre, jamais remplacée.** Sans texte ni logo : une de nos références recopiait l'inscription « FRESH SQUEEZE » dans les images suivantes, on l'a rayée pour de bon.

Pense aussi à la **voix** dès maintenant : une description anglaise, trois propositions, et des voix contrastées entre personnages. Kiwix : « Deep, velvety, low voice, very calm, slow and elegant diction, almost whispering, polite but chilling ».

Étape 3 – Le script et le découpage en plans

C'est l'étape la plus rentable. Chaque plan de 2 à 4 secondes reçoit sa fiche : cadrage, action, réplique, sons, prompt d'image. Chaque scène reçoit une **fiche de continuité** : décor, tenues, objet en main, échelle et, surtout, **l'état du décor après** la scène.

Puis tu écris la **chaîne logique « cause → effet »**. Celle de la salle de bain de l'épisode 5, telle qu'elle a fini par être validée :

L'écharpe va du panier à la fenêtre → Kiwix grimpe et entre par la fenêtre → mixeur ON → l'écharpe tombe dedans et le casse → Pêche crie, le rideau tombe sur elle → Mangue entend → Kiwix fuit par l'écharpe → Mangue entre : pièce vide.

Sans elle, notre v1 montrait un mixeur posé au sol puis tenu en l'air, une Mangue qui arrivait sans avoir entendu le cri et une Pêche qui interagissait avec un Kiwix qu'elle n'était pas censée voir. Tout se voyait sur le papier ; rien ne se voyait dans un prompt.

💡 Astuce : relis ton découpage d'une traite en te demandant « **pourquoi ?** » à chaque plan. Pourquoi il est là, pourquoi le personnage réagit, pourquoi l'objet est à cet endroit. « Parce que c'est drôle » ne suffit pas : il faut aussi un « parce que » dans l'histoire.

Quatre règles de découpage : toute réplique qui annonce une action a le plan qui la montre ; toute réaction a sa cause à l'écran ; au moins un plan à deux avant les champs-contrechamps ; ce qu'un personnage regarde est visible.

Étape 4 – Les images clés

Une image fixe par plan, qu'on animera ensuite. En local, on utilise **Qwen Image 2.1** dans WanGP avec l'accélérateur **Pruna 8 pas**, en **864×1536** : environ **20 secondes par image** sur une RTX 5080 (nos mesures : 21 images en 6 min 20). Dans le cloud, notre équivalent était nano_banana_pro sur Higgsfield, à 2 crédits l'image.

L'ordre qui change tout : le décor maître d'abord, puis le plan large avec le personnage, puis les gros plans **générés à partir du plan large**. Le gros plan hérite de la lumière, du fond et des accessoires ; généré de zéro, il « ne vient pas de là ».

ÉCOLE FRUISE · COHÉRENCE

PLAN LARGE D'ABORD,
GROS PLANS ENSUITE

La méthode qui garde le même décor dans tous les plans d'une scène. Exemple réel : la salle de bain de Pêche, épisode 5.

<image1>

1

LE DÉCOR MAÎTRE

Le lieu, généré **une seule fois**. Il devient la référence **<image1>** de tous les plans de la scène.

↓

2

LE PLAN LARGE DE LA SCÈNE

Toute l'action : qui est où, quels objets, quelle lumière. Ici, l'écharpe rose part du panier vers la fenêtre.

PROMPT

Same bathroom as <image1>. Wide shot...

↓

plan large



3

LES GROS PLANS,
À PARTIR DU PLAN LARGE



Le rideau opaque : une silhouette floue, rien d'autre.



Kiwix et le mixeur, dans la même pièce.

Référence = le plan large + l'image du personnage.

PROMPT

TIGHT close-up, face and shoulders only, filling the frame

CE QUI RESTE IDENTIQUE D'UN PLAN À L'AUTRE

✓ la lumière

✓ le fond

✓ les accessoires

✓ l'état du décor

X

Sans cette méthode, chaque gros plan réinvente son décor : autre lumière, autre fond, le panier et l'écharpe disparaissent.

Astuce

Jusqu'à 4 références par image. Le décor d'abord (<image1>), les personnages ensuite (<image2>, <image3>...).

Source : guide de production locale et images réelles de l'E05.

fruisse.com

École Fruise · fruisse.com

10 / 24

La formule de prompt :

```
Same <décor> as <image1>. <Cadrage>. The <personnage> from <image2>
(<bloc canonique>, <tenu>) <action>, <expression>, MOUTH CLOSED, eyes open.
<détails des objets>. <bloc de style> No text, no letters, no logo, no watermark.
```

Trois formules qui sauvent des plans :

- **un vrai gros plan** : « TIGHT close-up, face and shoulders only, filling the frame » ;
- **une place précise** : Kiwix est sorti debout *sur* son comptoir jusqu'à ce qu'on écrive « standing on the FLOOR BEHIND the counter... legs hidden » ;
- **un objet précis** : donne-le en image de référence, sinon le modèle en invente un autre (un épluche-légumes à la place de notre mixeur).

⚠ Piège : demander du texte à l'IA. Journal, panneau, contrat, compte à rebours : tout texte se compose à part (logiciel de dessin, petit script), et dans le prompt on écrit « blank paper, no writing ». Lisible, gratuit, sans lettres inventées.

Avant d'animer, assemble toutes tes images en **planche contact** et compare-les côte à côte : même visage, bonne tenue, bon objet, bon état du décor, aucune lettre parasite. Ce qui rate se refait avec une autre graine.

Étape 5 – L'animation

Chaque plan part de **son image clé validée** ; le prompt vidéo ne décrit que ce qui bouge. En local, notre réglage par défaut est **MiniMax H3** (mode « image de départ ») avec le **LoRA « Pika merge lite »** : 480×832, 4 pas, environ **50 secondes par plan**. Ce LoRA garde mieux le visage d'un bout à l'autre du plan (Kiwix y garde ses gros yeux fous).

📺 Vidéo : Un plan de Kiwix fabriqué en local, de l'image clé à la vidéo, avec son making-of — à regarder sur fruisse.com/ecole/

Les astuces qui comptent :

- **Reste en 480×832** et agrandis en 720×1280 au montage : en 720×1280, le calcul est environ 2,5 fois plus lent pour un gain invisible sur téléphone.
- **Le modèle sort au moins 4,5 secondes** : tu coupes au montage et tu gardes le meilleur passage.
- **Écris l'action dans l'ordre, caméra fixe** sauf mouvement demandé : « Animated film style. . Static camera unless stated. »

- **Une sortie de champ doit être définitive** : « drops straight DOWN and disappears... he never comes back ». Sans ça, notre Kiwix revenait dans la fenêtre qu'il venait de quitter.
- **Les inserts n'ont pas besoin d'IA** : un zoom lent sur l'image clé, au montage, fait un très bon plan gratuit.

💡 Astuce : un plan peut démarrer juste et **dérivé à la fin**. Pour chaque plan d'action important, lance deux graines, et relis chaque vidéo sur cinq images : à 0,2 / 0,8 / 1,6 / 2,6 et 3,6 secondes. C'est ainsi qu'on a attrapé un rideau uni qui se dessinait un visage.

Sans carte graphique : sur Higgsfield, Grok 1.5 Lite coûte environ 1 crédit par seconde pour l'action et Wan 2.7 environ 1,5 crédit par seconde pour les plans parlants ; sur OpenRouter, nos plans d'action de l'épisode 6 en Grok Imagine 1.5 Lite (720p) ont coûté environ 0,10 à 0,13 \$ pièce, soit environ 4 \$ pour l'épisode (relevé le 30/10/2026, à vérifier). Les services cloud ont des filtres : deux refus, on change d'image de départ.

LOCAL OU CLOUD ?

Ton PC avec une carte graphique, ou des services en ligne payants : ce qu'on a vraiment vécu en fabriquant SerialKiwix.



€ COÛT

✓ avantage

0 crédit. Seulement l'électricité du PC.
L'épisode 5 n'a rien coûté.

≈ **200 crédits** Higgsfield pour un épisode de 1 min 30.
E01 à E04 : ≈ 185-216 crédits chacun.

🕒 VITESSE

ça dépend

≈ **20 s** par image, ≈ **50 s** par plan vidéo.
Un épisode : ≈ 40 min à 1 h de machine.

Rien à installer, ton PC reste libre.
Délai selon le service et sa file d'attente.

👁 QUALITÉ

✓ avantage

Proche de Higgsfield, personnages reconnaissables.
Vidéo en 480p, agrandie en 720p au montage.

Le plus net pour **2-3 plans clés**.
Jusqu'en 1080p, voire 4K selon le modèle.

🛡 FILTRES

✓ avantage

Aucun refus : Pêche passe partout.
Tu gardes tes propres limites (rien de sexualisé).

Filtre « nsfw » : **3 refus** sur Pêche dans l'E04.
Règle : 2 refus → change d'image de départ.

📄 LICENCES

à vérifier

MiniMax H3 / Pika : licence qui **exclut l'Union européenne**. Qwen Image 2.1 : **non commercial**.
Textes lus le 31/10/2026 : à lire avant de publier.

Les conditions du service **et de ton abonnement**.
Usage commercial : à vérifier avant de monétiser.

NOTRE RÈGLE

Local gratuit par défaut, cloud pour 2-3 plans clés quand il reste des crédits.

Source : épisodes 1 à 5 de SerialKiwix (octobre 2026).

fruisse.com

Étape 6 – Voix, son et musique

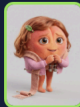
Les voix. Sur la série, elles viennent d'ElevenLabs (modèle eleven_v3), qui comprend des balises de jeu dans le texte : [whispers], [sighs], [shouting], [nervous], [softly] ... Des voix gratuites existent en local (OmniVoice, par exemple) ; on ne les a pas encore utilisées pour un épisode, et la licence du modèle OmniVoice exclut l'usage commercial.

La synchronisation labiale, en local. Quatre réglages, et ça marche :

1. une image de départ en gros plan, **bouche fermée** (« MOUTH CLOSED, eyes open, static camera ») ;
2. la réplique donnée au modèle comme bande-son ;
3. la réplique **rallongée de silence jusqu'à 5,6 secondes** (sans ça, le modèle invente une suite, souvent en anglais) ;
4. dans le prompt, une seule phrase : « is speaking with natural lip movements synchronized, then stays silent with the mouth closed. » Jamais les mots de la réplique.

LA SYNCHRO LABIALE EN LOCAL

Faire parler un personnage avec MiniMax H3 sur ton PC, sans qu'il invente la suite. Toutes les pistes sont à la même échelle.



PÊCHE, ÉPISODE 5

« J'ai rien vu... mais on a mixé mon tricot ! »

0 s 1 s 2 s 3 s 4 s 5 s 6 s

1 LA RÉPLIQUE ELEVENLABS



2 RALLONGÉE DE SILENCE JUSQU'À 5,6 s



3 MINIMAX H3 REÇOIT TROIS CHOSES

IMAGE DE DÉPART

« MOUTH CLOSED, eyes open, static camera »

AUDIO_GUIDE

la réplique rallongée à 5,6 s

LE PROMPT

il décrit le jeu, pas le texte

<personnage> is speaking with natural lip movements synchronized, then stays silent with the mouth closed.

Ne jamais écrire les mots de la réplique.

4 LA VIDÉO DURE ≈ 4,5 s

107 images minimum : H3 ne fait jamais plus court.



5 COUPE AU MONTAGE

On garde jusqu'à la fin de la réplique : après la voix, le personnage change d'expression.



× fin de la réplique



Sans la rallonge de silence, H3 invente une suite à la réplique... souvent en anglais.

Source : guide de production locale et retour d'expérience de l'E05.

fruisse.com

Les bruitages. Chacun calé sur l'instant exact de l'image : le « pop » de la bulle de chewing-gum de Mangue tombait à 2,0 s alors que la bulle éclate à 2,25 s, et le gag tombait à plat. Aucun son sans image. Rien de strident : notre première douche était pleine d'aigus ; refaite

douce et coupée vers 2,8 kHz, plus personne ne grimace. Pour l'eau, la pluie et la vapeur, un passe-bas vers 3 kHz.

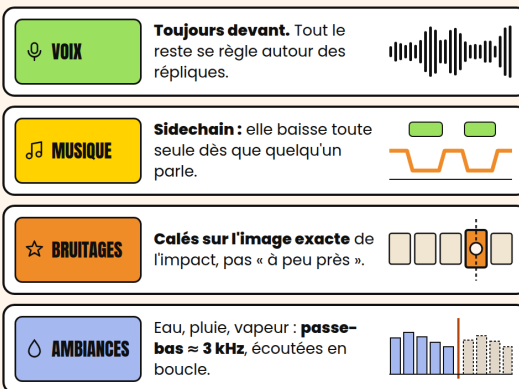
La musique. ElevenLabs en génère ; Lyria 3 de Google, via OpenRouter, coûtait environ 0,04 \$ le morceau de 30 secondes et 0,08 \$ le morceau long (testé le 30/10/2026). Écris toujours « Instrumental only, no vocals ».

Le mixage. Musique baissée sous les voix, limiteur, environ **-14 LUFS**, pic sous -0,3 dB, audio en **48 kHz** (sur l'épisode 5, un export en 192 kHz créait des pics à +3,7 dB).

UN SON DE PRO

Ce qui fait passer une vidéo de « on entend » à « on y croit ». Nos réglages, testés et corrigés sur SerialKiwix.

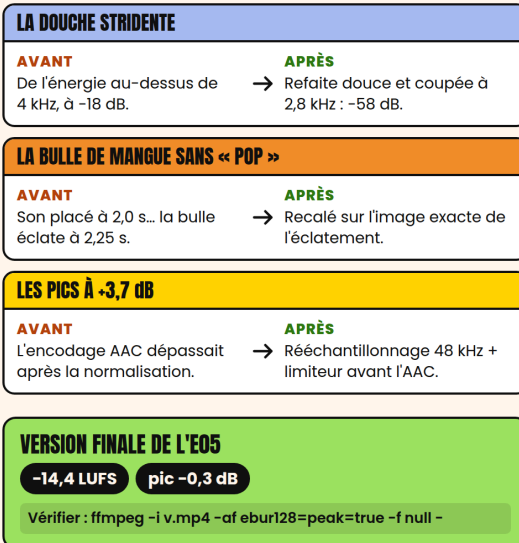
4 PISTES, 4 TRAITEMENTS



AUCUN SON SANS IMAGE

Chaque bruitage correspond à quelque chose qu'on voit à l'écran, à l'instant où on le voit.

CE QUE L'E05 NOUS A APPRIS



Source : méthode son et montage, retour d'expérience de l'E05.

fruisse.com

Étape 7 – Montage et publication

Les règles de montage : un plan dure 4 secondes au maximum ; une réplique ne déborde jamais sur le plan suivant et finit au moins 0,3 s avant un fondu ; la narration a sa propre couleur de sous-titre (le vert pour la voix off de Kiwix) ; cartons de lieu gros (environ 60 px) sur fond opaque ; un gros carton de fin avec un appel. N'importe quel logiciel de montage fait l'affaire pour commencer.

La revue : extrais une image par seconde, assemble-les en grilles, et regarde ta vidéo en entier en te demandant « pourquoi ? » à chaque plan. Sur l'épisode 6, cette relecture a attrapé un rideau de fer encore ouvert après sa fermeture et un carton « LE MÊME JOUR » qui devait dire « LA VEILLE ».

Le kit de publication :

- une **miniature 3:4** (1080×1440) avec le personnage reconnaissable et l'objet clé, le visage hors des zones de texte ;
- pour Instagram, une **couverture 9:16** dont les titres restent dans la zone centrale 3:4 (de y = 240 à y = 1680 sur 1080×1920), car la grille du profil la recadre ;
- une **légende** à structure fixe : « SerialKiwix · Épisode N », une accroche, une question à choix (« Team écharpe ou team mixeur ? »), et toujours les mêmes hashtags de base ;
- la **mention IA**, cochée au moment de publier (sur trois réseaux sur quatre, on n'a pu l'ajouter après coup que depuis l'appli mobile).

LE RYTHME DE PUBLICATION

Ce que nos premiers chiffres TikTok (compte @fruisse.com, octobre 2026) nous ont appris.

1 QUAND PUBLIER



Un épisode par jour, à heure fixe, entre 18 h et 21 h.

2 ENTRE DEUX ÉPISODES



L'extrait renvoie vers l'épisode complet. Il ne coûte rien : on le coupe dans nos montages.

3 LA LÉGENDE TYPE

@fruisse.com

SerialKiwix · Épisode 3
[ton accroche, en une ligne]
Team Kiwix ou team BCM ?

#SerialKiwix #fruitdrama
#histoiredefruits #fruitsIA #pomme

Commentaire épinglé
Il renvoie aux épisodes précédents et annonce le suivant.

TITRE CONSTANT
Toujours au début.

ACCROCHE
Une seule ligne.

QUESTION À CHOIX
Simple : « Team X ou team Y ? »

4 HASHTAGS FIXES
+ 1 lié à l'épisode.

4 SIX HABITUDES

Mêmes hashtags partout
Sinon, aucun fil pour retrouver la série.

Épingle l'épisode 1
Les nouveaux commencent par le début.

Réponds dans l'heure
À chaque commentaire, en vidéo si possible.

Épisodes de plus d'1 min
Ne te bride pas sur la durée.

Coche « contenu IA »
Sur TikTok, Facebook et Instagram.

Une question à choix
E02 : 0 commentaire malgré une question.

≈ **150-200**

vues : le premier « lot test » que TikTok donne à toute nouvelle vidéo. La suite dépend du temps de visionnage.

Source : analyse du compte TikTok du 29/10/2026 et méthode de publication.

fruisse.com

✓ À retenir : publie entre 18 h et 21 h, à heure fixe, épingle ton épisode 1, réponds aux commentaires dans l'heure, et attends-toi à 150-200 vues pour commencer. Le vrai indicateur, c'est le temps de visionnage, visible dans tes statistiques.

L'épisode 5 en chiffres

ÉCOLE FRUISE · RETOUR D'EXPÉRIENCE

L'ÉPISODE 5 EN CHIFFRES

« La douche » : le premier épisode de SerialKiwiX fabriqué 100 % en local, sur le PC de KALEXEX, le 29 octobre 2026.



Six images clés de l'épisode, toutes générées sur le PC.

60

IMAGES

Qwen Image 2.1 + Pruna 8 pas
≈ 20 s pièce

56

VIDÉOS

MiniMax H3, 480x832
≈ 45-50 s pièce

≈ 67

MINUTES

de calcul pour tout l'épisode,
corrections comprises

0

CRÉDIT CLOUD

aucun crédit Higgsfield (voix
et bruitages : ElevenLabs)



RTX 5080 - 16 Go

La carte graphique du PC (16 Go de VRAM) : tout a
tourné dessus.

3 VERSIONS DANS LA JOURNÉE

Temps de machine de chaque passe (sur ≈ 67 min au total)

V1

1 min 59

≈ 40 min

Fabrication complète : 37 images + 37 vidéos

V2

2 min 05

≈ 9 min

Passe de cohérence : 6 images + 8 vidéos

V3

2 min 07

≈ 18 min

Salle de bain refaite : 17 images + 11 vidéos

≈ 40 %

du temps machine a servi à **corriger des
erreurs de logique**. D'où notre checklist de
cohérence, à remplir **avant** de générer.

VERSION VALIDÉE

2 min 07

42 plans

18 répliques

-14,4 LUFS

aucun refus de filtre

Source : retour d'expérience S01E05 « La douche »
(29/10/2026).

fruise.com

	Images	Vidéos	Temps machine
v1 (fabrication complète)	37	37	≈ 40 min
v2 (passe de cohérence)	6	8	≈ 9 min
v3 (salle de bain refaite)	17	11	≈ 18 min
Total	60	56	≈ 67 min

Deux leçons dans ce tableau. La mauvaise nouvelle : les v2 et v3, soit environ 40 % du temps machine total, ont servi à corriger la logique, ce qu'une checklist de cohérence remplie avant de générer aurait évité. La bonne : en local, un plan raté se refait en une minute et demie et une scène entière en un quart d'heure environ. On peut se permettre de corriger.

Ton plan d'action pour un week-end

Objectif : une vidéo de 30 à 45 secondes, 10 à 12 plans, un personnage, un lieu, publiée dimanche soir. Les temps machine viennent de nos mesures sur une RTX 5080 ; les temps « toi » sont une estimation pour une première fois, à ajuster.

Quand	Ce que tu fais	Ce que tu dois avoir à la fin
Samedi matin (≈ 2 h)	étape 1 : la phrase, les limites, le format, le hook et la chute ; étape 2 : la fiche, le bloc canonique en anglais, la voix décrite	une fiche projet et une fiche personnage
Samedi midi (≈ 1 h)	l'image de référence (4 graines, environ 1 min 30 de calcul), puis le décor maître	une référence validée, sans texte ; un décor maître
Samedi après-midi (≈ 3 h)	étape 3 : le découpage en 10-12 plans, la fiche de continuité, la chaîne « cause → effet », le « pourquoi ? » à chaque plan ; enregistrer les répliques	un découpage complet et les répliques audio
Dimanche matin (≈ 3 h)	étape 4 : 15 à 20 images (environ 5 à 7 min de calcul), planche contact, reprises ; étape 5 : 12 à 16 plans vidéo (environ 10 à 13 min de calcul), relecture sur 5 images	une image validée et un clip validé par plan
Dimanche après-midi (≈ 3 h)	étape 6 : plans parlants, bruitages calés, musique ; étape 7 : montage, sous-titres, mixage à -14 LUFS	une vidéo complète
Dimanche soir (≈ 1 h)	revue en planches, miniature 3:4, légende, mention IA ; publication entre 18 h et 21 h	ta première vidéo en ligne

Sans carte graphique, le plan reste le même : remplace les temps de calcul par des crédits cloud, et vérifie les prix le jour même.

💡 Astuce : garde un fichier ouvert tout le week-end et note chaque raté avec sa correction. Dimanche soir, ces notes deviennent ton premier **retour d'expérience** : c'est exactement comme ça que nos 35 règles, de A à AI, sont nées.

Les erreurs qu'on a faites pour toi

- Générer avant d'avoir écrit la chaîne logique (40 % du temps machine de l'épisode 5 en corrections).
- Monter un hook avec un plan d'un autre épisode, parce qu'il « en jetait ».
- Générer les gros plans avant le plan large.
- Une référence avec du texte dedans.
- Des adjectifs d'humeur au lieu de traits visibles.
- Ne relire que la première image d'une vidéo.
- Un bruitage strident jamais écouté en boucle.
- Un son placé un quart de seconde trop tôt.
- Publier sans cocher la mention IA.

Pour aller plus loin

Ce guide te donne la carte. Le cours « **Ta première vidéo IA en 7 étapes** » de l'École Fruise te donne le terrain, étape par étape : les réglages complets de chaque modèle (tels qu'on les envoie à WanGP), des prompts entiers commentés, les gabarits à remplir (fiche projet, fiche personnage, découpage, fiche de continuité, retour d'expérience), les commandes ffmpeg testées pour le son, les inserts et les planches, et les licences de chaque outil avec leurs alternatives. Les deux premières leçons sont en accès libre ; les cinq suivantes s'ouvrent avec le Pass École.

Et si ton personnage change de tête d'un plan à l'autre, le cours « **Un personnage qui reste le même, plan après plan** » y est entièrement consacré.

Bonne fabrication. Et que ta première vidéo soit juteuse.

Va jusqu'au bout avec le Pass École

Ce guide te donne la carte. Les cours de l'École Fruise te donnent le pas-à-pas : nos vrais prompts et réglages, les infographies, les vidéos expliquées, les budgets et les projections chiffrées.

29 €

Tous les cours pendant 30 jours. Paiement unique, sans renouvellement automatique.

Commence par les leçons gratuites : fruisse.com/ecole

Le Pass : fruisse.com/ecole/pass