LA MAISON DES ARTISTES
L'association nationale des artistes des arts visuels

Protéger ses images contre l'entraînement des IA génératives : Glaze et Nightshade

Les modèles d'IA générative d'images s'entraînent sur des millions d'images collectées en ligne, le plus souvent sans l'accord des artistes-auteur·ices qui les ont créées. Glaze et Nightshade, deux outils gratuits développés par le SAND Lab de l'Université de Chicago, modifient discrètement vos fichiers avant publication pour limiter cette exploitation. Ce ne sont ni des solutions juridiques ni des garanties absolues, mais des contre-mesures à intégrer dans une stratégie plus large de protection de votre travail.

Pour un·e artiste, l'aspiration de ses images par les IA peut causer deux préjudices distincts :

- la reproduction de son style par un modèle affiné sur son travail (perte de commandes, dilution de sa signature visuelle…) ;
- l'utilisation de ses images, sans consentement ni rémunération, pour entraîner les grands modèles.

Glaze répond au premier, Nightshade au second.

GLAZE : PROTÉGER SON STYLE

Glaze applique à vos images un « voile » de perturbations invisibles à l'œil humain, qui fait percevoir à un modèle d'IA un style différent du vôtre. Un modèle entraîné sur des images « glazées » apprend donc un style déformé, qui ne correspond plus à votre travail.

Objectif : empêcher qu'un modèle soit affiné (fine-tuné) pour imiter le style personnel d'un·e artiste à partir de ses images.

- **Robustesse :** la protection résiste aux captures d'écran, au recadrage, à la compression, au filtrage et au redimensionnement.
- **Portée :** Glaze contre l'imitation ciblée de votre travail. Il n'empêche pas un modèle de reproduire un style déjà très présent dans ses données de base (impressionnisme, esthétique manga…).

♥ **Bon à savoir :** sans carte graphique compatible, vous pouvez utiliser WebGlaze, la version en ligne et gratuite de Glaze. Les images sont traitées sur les serveurs du projet, renvoyées par e-mail, puis supprimées. L'accès se fait sur invitation, à demander au compte X (Twitter) ou Instagram @TheGlazeProject.

NIGHTSHADE : RENDRE LE PILLAGE PLUS RISQUÉ

Nightshade transforme une image en échantillon d'entraînement « toxique ». À l'œil humain, elle ne change presque pas ; mais un modèle qui l'intègre associe le sujet représenté à des caractéristiques erronées. Exemple cité par les développeurs : une vache finit par être perçue comme un sac à main en cuir.

Objectif : rendre l'entraînement sur des œuvres non consenties coûteux et risqué, pour inciter les plateformes à acheter des licences plutôt qu'à aspirer les images.

- **Effet cumulatif et collectif :** plus un jeu de données contient d'images empoisonnées sur un même sujet, plus le modèle entraîné dessus se dégrade. Comme pour Glaze, l'effet résiste au recadrage, à la compression et aux captures d'écran.

⚠ **Attention :** dans sa version 1.0, Nightshade seul ne protège pas votre style. Il vise le jeu de données dans son ensemble, pas l'imitation de votre patte personnelle. Pour vous protéger vous-même, associez-le à Glaze.

GLAZE OU NIGHTSHADE ? LES DEUX

Le SAND Lab recommande d'appliquer les deux outils à toute image publiée en ligne, **toujours Nightshade en premier, puis Glaze**.

	Glaze	**Nightshade**
Ce qu'il fait	Protège votre style contre l'imitation	Rend l'image toxique pour l'entraînement
Bénéfice	Individuel : direct, pour l'image traitée	Collectif : croît avec le nombre d'images empoisonnées en circulation
Version en ligne	Oui, WebGlaze (sur invitation)	Non, logiciel à installer uniquement

MODE D'EMPLOI

→ AVANT DE COMMENCER :

- **Formats :** JPG ou PNG uniquement. Pour Glaze, 5 000 pixels maximum sur le plus grand côté, et aucun caractère accentué ou spécial dans le nom du fichier.
- **Système :** sur Mac, Glaze demande macOS 13.0 ou ultérieur (puce Intel ou Apple Silicon M1, M2, M3).
- **Carte graphique (recommandée) :** NVIDIA, avec plus de 4 Go de mémoire et présente dans la liste de compatibilité des développeurs. Avec une GTX 1660, 1650 ou 1550, installez d'abord le toolkit NVIDIA CUDA depuis le site officiel de NVIDIA.

- **Sans carte graphique compatible :** les deux outils fonctionnent aussi avec le processeur (mode CPU), sur Mac comme sur Windows, mais plus lentement. Fermez les autres applications pendant le traitement. Pour Glaze, WebGlaze est une alternative.
- **Premier lancement :** l'outil télécharge une seule fois ses bibliothèques et modèles, puis fonctionne hors connexion.

→ ÉTAPES :

1. **Préparez vos images.** Redimensionnez-les, puis filigranez-les si vous le souhaitez : Glaze et Nightshade ne le font pas. Deux outils gratuits et sans création de compte traitent des lots d'images :
   - **iLoveIMG** : en ligne, fichiers supprimés automatiquement sous deux heures, certifié ISO 27001 et conforme au RGPD ;
   - **Watermarkly** : traitement dans le navigateur, les images ne sont jamais envoyées sur les serveurs de l'éditeur.

   Évitez en revanche les outils de design grand public type Canva, dont les conditions liées à l'IA peuvent permettre d'utiliser vos contenus pour entraîner leurs services.
2. **Importez-les dans Nightshade** par glisser-déposer ou sélection multiple.
3. **Réglez l'intensité et la qualité de rendu.** Plus l'intensité est élevée, plus la protection est robuste, mais plus les altérations deviennent visibles. La qualité de rendu détermine le temps de calcul.
4. **Vérifiez le tag (Nightshade uniquement).** L'outil propose un mot décrivant le sujet principal de l'image (« portrait », « paysage »…) : corrigez-le si besoin.
5. **Choisissez le dossier de sortie et lancez le traitement.** Comptez quelques minutes avec une carte graphique, de 20 minutes à plusieurs heures en mode CPU. Les fichiers gardent leur nom d'origine.
6. **Recommencez avec Glaze** sur les fichiers obtenus (étapes 2, 3 et 5).
7. **Exportez en PNG**, puis compressez en JPG seulement après, pour préserver la qualité. Vérifiez le rendu à l'écran avant de publier.

⚠ **Attention :** traitez vos images avant toute mise en ligne. Une image déjà partagée sans protection reste accessible ailleurs sur le web, même si vous en republiez ensuite une version protégée.

♥ **Bon à savoir :** ne signalez nulle part (légende, nom de fichier, métadonnées) qu'une image a été traitée avec Glaze ou Nightshade : cela facilite son repérage et son contournement par les systèmes de scraping.

✱ **Exemple :** Camille, illustratrice, prépare la mise en ligne d'une nouvelle série de dessins. Elle redimensionne et filigrane ses fichiers, les traite avec Nightshade (tag « portrait », intensité moyenne), puis avec Glaze (intensité moyenne également). Elle exporte en PNG, vérifie qu'aucune différence n'est visible à l'écran, compresse en JPG pour le web et publie sans jamais mentionner le traitement.

LIMITES : CE QUE CES OUTILS NE GARANTISSENT PAS

- **Pas d'effet rétroactif :** le traitement ne retire rien des jeux de données ni des modèles existants. Il ne protège que les images traitées avant leur mise en ligne.
- **Pas de garantie permanente :** les développeurs reconnaissent que de futures méthodes pourraient contourner la protection. Deux attaques contre Glaze ont échoué (attaque PEZ, outil AdverseCleaner) ; une troisième, dite du « noisy upscaler » (juin 2024), a fonctionné contre les versions antérieures, ~~d'où la sortie de Glaze 2.1.~~
- **Altérations plus visibles sur certains styles :** surtout sur les aplats de couleur et les fonds unis, même si Glaze 2.0 a nettement amélioré ce point.
- **Protection inégale contre les générateurs « image vers image » (img2img) :** seule une intensité élevée de Glaze (jusqu'à 5x) protège vraiment, au prix d'altérations parfois trop visibles.
- **Accès matériel inégal :** un traitement de qualité suppose une carte graphique compatible, ou un accès à WebGlaze, sur invitation et à capacité limitée.
- **Pas une protection juridique :** ces contre-mesures techniques ne remplacent ni l'affirmation de vos droits d'auteur, ni la mention de vos conditions d'utilisation, ni les démarches contre une exploitation constatée de votre travail.

En complément de cette fiche pratique, pour aller plus loin sur ces deux outils, nous vous proposons de consulter :

→ le site officiel de Glaze : glaze.cs.uchicago.edu,

→ le site officiel de Nightshade : nightshade.cs.uchicago.edu,

→ le guide utilisateur de Nightshade : Nightshade Software User Guide,

→ la FAQ de Glaze : Glaze FAQ,

→ pour demander un accès à WebGlaze et suivre les mises à jour des deux outils : @TheGlazeProject (X et Instagram).

La Maison des Artistes · 11 rue Berryer, 75008 Paris · 01 42 25 06 53 · **lamaisondesartistes.fr**