**ChatGPT fait peur à Europol qui alerte sur les abus possibles.**

Les personnes mal attentionnées sont prêtes à tout, même a utilisé l’intelligence artificielle pour commettre des fraudes et d’autres cybercrimes, a averti l’agence de police européenne Europol.

**Qui est Europol ?**

Europol est à l’écoute des besoins des États membres de l’UE et analyse les tendances en matière de criminalité dans l’Union européenne. L’Agence soutient les enquêtes lancées par les États membres, mais ses agents n’arrêtent jamais les citoyens ni ne lancent d’enquêtes. Les travaux d’Europol consistent généralement à traiter les cas de criminalité qui nécessitent une approche internationale et la coopération entre plusieurs pays, au sein et en dehors de l’Union européenne.

L’exploitation potentielle de chatgpt offre « de sombres perspectives », met en garde Europol. Du phishing à la désinformation et aux logiciels malveillants, les capacités en évolution rapide des robots de conversation (chatbots) sont susceptibles d’être rapidement exploitées par des personnes mal intentionnées, a déclaré l’agence de police européenne dans un nouveau rapport.

Créé par la start-up américaine openAI, ChatGPT est apparu en novembre et a rapidement été pris d’assaut par des utilisateurs émerveillés par sa capacité à répondre clairement à des questions difficiles, à écrire des sonnets ou du code, et même à réussir des examens.

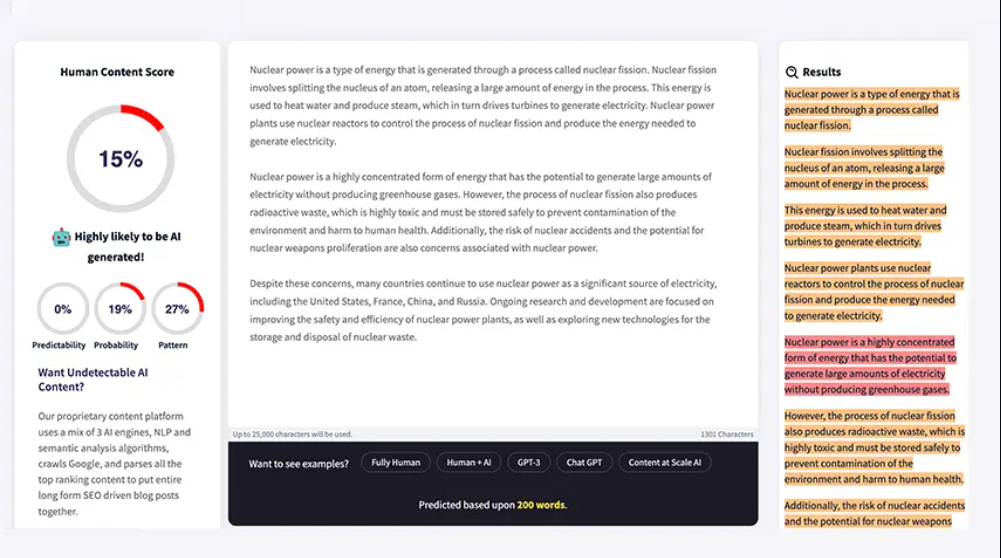
Les criminels pourraient alors ce concentré sur Chatgpt ou un autre système d’intelligence artificielle du même type, pour « *accélérer considérablement le processus de recherche* » dans des domaine ou le criminel n’y connais rien du tout, comme la rédaction d’un texte pour commettre une fraude ou donner des informations sur « *comment entrer par effraction dans une maison, cybercriminalité …* »

**Mode de défense mit en place**

Imaginons une facture a été créé à partir de chatgpt des mots, dates …, et que nous voulions mettre en place un système ou utiliser un système pour savoir si oui ou non cette facture a du texte qui appartient à Chatgpt.

Première méthode que j’ai trouvée est d’utiliser un détecteur IA, Content at Scale par exemple qui est un outil qui permet de détecter l’écriture de chatgpt. Simple d’utilisation on rentre dans le champ de saisie le texte provenant de ce dernier puis on lance la recherche, selon le degré de prévisibilité et de complexité du texte, des résultats apparaitront.

Voici un exemple ci-dessous :



On peut voire à droite les lignes en orange sont les mots de Chatgpt détecter par l’IA.

La deuxième méthode trouvée qui n’a pas encore vu le jours mais c’est pour bientôt, des rumeurs racontent que OpenAI prévoit d’ajouter « un code secret » appelé jeton, exclusif au modèle GPT-3, serait inséré n’importe ou dans le texte généré pour ajouter un filigrane.

Le jeton pourra être un mot, une ponctuation ou juste une partie d’un mot, le résultat pourrait être utilisés pour déterminer si quelque chose a été crée par le modèle GPT-3 tout en restant cohérant et grammaticalement correct. Cette méthode est comparable à l’ajout d’un filigrane a une image numérique, il sera invisible à l’œil nu mais peut être utilisé pour déterminer l’origine de l’image.

Sortie image chatGPT ?

Source explorec : <https://aiexplorer.io/actualites-ia/comment-detecter-un-texte-redige-par-chatgpt/#:~:text=Le%20d%C3%A9tecteur%20IA%20de%20Content,texte%2C%20vous%20obtiendrez%20des%20r%C3%A9sultats>.

Source leParisien : https://www.leparisien.fr/high-tech/fraude-desinformation-chatgpt-fait-peur-a-europol-qui-alerte-sur-les-abus-possibles-27-03-2023-R3PFOQTOGBGPRDHYAWSY447YDE.php

**Google met en place un outil qui permet de vérifier une image**

***Technologie :****Pour éviter la diffusion de fausses informations, notamment par le biais d’images trop réalistes générées par IA, Google propose une nouvelle fonctionnalité de vérification des images.*

**Recherche sur YOLO pour la détection de logo**

Recherche sur la précédente veille techno deux questions son ressortit est-ce que le système pour détecter les phrases par gpt ont été mis en place ? chatgpt est capable de renvoyer des images celons la demande du client ?

Pour la première question, chatgpt n’a pas encore mit en place un système qui permet de détecter un « jeton » dans la phrase mais cela devrait arriver.

Ensuite pour ce qui est de possibilité de chatgpt à renvoyer des images selon la demande du client, et bien, oui mais pas directement car chatgpt est reconnu pour faire du texte to texte cette à dire qu’en entrée il ne prend qu’une question ou une demande au format texte et renvoie une réponse en format texte uniquement.

Avec la nouvelle mise a jour ChatGPT se met au « dessin » grâce à DALL-E 3 « *DALL-E, est sur le point de recevoir une immense mise a jour. C’est ce qu’a annoncé Sam Altman, dans un message twitter le 20 septembre 2023 il a ajouté que sont déploiement est une question de semaines* »

Il est accessible mais pas a tout le monde il faudra posséder un abonnement a chatGPT (20 euros/mois) ou avoir chatgpt entreprise.

Un petit exemple de l’évolution entre DALL-E 2 et DALL-E 3 :



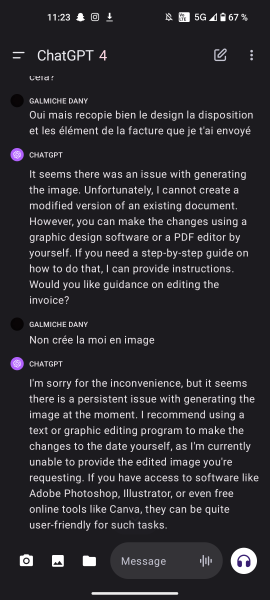
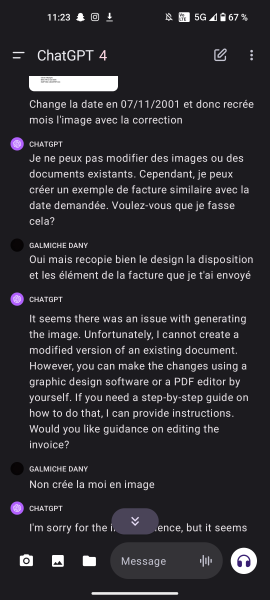
Là est le point fort de l’évolution de DALL-E 3 hormis l’évolution des graphismes, c’est qu’il a été nativement concu sur chatgpt, en clair il est possible d’utiliser Chatgpt comme partenaire de brainstorming pour affiner les instructions à envoyer à DALL-E 3.

« Lorsque vous soumettez une idée, notre OpenAI, Chatgpt génère des prompts personnalisé et détaillés pour DALL-E 3, automatiquement qui lui donnent vie. Si vous aimez une image particulière mais qu’elle ne vous convient pas on peut demander à Chatgpt d’y apporter des modifications en quelques mots »

Je suis aussi en discussion avec un ami qui a un abonnement chatgpt et je lui et demander quelques requêtes sur un exemple de facture (qui n’ont rien à voir avec notre projet).

Voici quelques exemples :

Prompt : peut tu me modifier cette image, change la date 29/01/2019 en 07/11/2001.



On peut voir que pour lui ce n’est pas encore possible de faire cela, mais on ne doit quand même pas écarter le sujet avec les nouvelles performances qui peuvent voir le jour nous ne sommes pas à l’abri que DALL-E un jour pourra modifier une facture.

Source DALL-E 3: https://www.numerama.com/tech/1508138-chatgpt-va-generer-des-images-grace-a-la-toute-nouvelle-ia-generative-dall-e-3.html

Source Midjourney une autre IA générative: https://www.numerama.com/tech/1194728-comment-utiliser-midjourney-et-son-discord.html

Recherche sur le cas de retrouver des éléments modifier sur un PDF.

Il existe bien sûr Photoshop pour modifier le texte etc...

Mais un autre logiciel existe aussi qui permet de modifier directement un PDF c’est pdfelement qui payant.

Le processus est simple il suffit d’ouvrir le fichier PDF et de le modifier à sa guise.

Il est possible de rajouter beaucoup de chose

Zone de texte, des notes sur le PDF … .

***AutoEncodeur***

Si nous modifions une ou plusieurs images, exemple insérer un carré blanc pour faire disparaitre du texte et y ajouter du texte par-dessus cela peut faire apparaitre du bruit malgré le faite que cette modification soit invisible à l’œil nu.

Encodages des couleurs : Même si le carré blanc a exactement la même couleur que le fond, des variations dans l’encodage des couleurs peuvent se produire lors de la création du PDF. Cela pourra résulter en de petite différence de valeurs de pixel.

Compression PDF : Les fichiers PDF peuvent être compressés pour réduire leur taille, et cette compression peut créer des artefacts ou des variations subtiles dans l’image.

Certains algorithmes de compression peuvent réagir aux régions uniformes de couleur, comme celle créée par un carré blanc.

Donc si ces carrés blancs sont invisibles à l’œil nu une approche potentielle pourrait être d’utiliser des modèles d’apprentissage profond (Deep Learning) spécifiquement conçus pour la détection de modification d’anomalie subtiles dans les images. Ces modèles peuvent être formés sur des ensembles de données.

Apprentissage profond : On pourrait par exemple utiliser l’apprentissage non supervisé telles que des autoencodeurs. Ces modèles apprennent à représenter les caractéristiques les plus importantes des données sans étiquettes explicites. Si une modification a eu lieu le modèle devrais pouvoir détecter l’anomalie.

Réseaux de neurones siamois : Ces réseaux sont conçus pour apprendre à mesurer la similarité entre deux entrées. On pourrait former un réseau de neurones siamois pour comparer différentes parties de l’images avant et après l’ajout des carrés blancs.

Apprentissage profond serait la méthode la plus pratique détaillons en générale les différentes étapes à suivre pour le bon fonctionnement.

**1. COLLECTE DES DONNEES D’ENTRAINEMENT :**

Rassemblez un ensemble de données PDF qui incluent des factures vraies et des factures fausses

**2. PRE-TRAITEMENT DES DONNES :**

Convertir PDF en image puis les normalisées pour que le modèle puisse apprendre des caractéristiques générales.

**3. CREATION DU MODELE AUTOENCODEUR :**

Concevoir un autoencodeur qui est composé d’un encodeur et d’un décodeur.

L’encodeur réduit la dimension de l’image d’entrée a une représentation latente tandis que le décodeur tente de reconstruire l’image d’origine à partir de cette représentation.

(Une représentation latente fait référence a la représentation interne d’une donnée apprise par un modèle pendant le processus d’entrainement).

**4. ENTRAINEMENT DU MODELE :**

Divisez l’ensemble de données en données d’entrainement et de validation.

Entrainez le modèle autoencodeur sur les exemples normaux, en minimisant la différence entre l’entrée et la sortie reconstruite.

**5. DETECTION D’ANOMALIE :**

Evaluer les performances du modèle sur l’ensemble des données de validation.

Tester le modèle avec des images ou il y’a des carrés blancs pour voir comment le modèle réagit.

**6. OPTIMISATION ET REAJUSTEMENT :**

Selon les résultats, il faudra optimiser le modèle en ajustant certain hyperparamètres. Ou en augmentant les données.

Pour augmenter les données certaines techniques peuvent être mis en place, comme faire une rotation sur les images, les déplacez verticalement et horizontalement ce qui peut simuler un changement de position, faire des zoom, méthode flipping qui effectue une symétrie horizontale …

***FraudGPT***

Qu’Est-ce que FraudGPT ?

FraudGPT est un produit vendu sur le darkweb ou sur Telegram qui fonctionne de la même manière que ChatGPT, ce dernier serait capable de crée du contenu pour faciliter les cyberattaques.

Ce qu’il faut savoir c’est que FraudGPT n’a aucune limite, il ne dispose pas de contrôle et de limitation comme ChatGPT qui lui doit se restreindre au niveau des réponses fournit, il serait indiqué également que FraudGPT serait mit a jour toutes les une a deux semaines et utiliserais des modèle d’intelligence artificielle.

Comment fonctionne-t-il ?

Une équipe de recherche (Netenrich) a pu acheter et testé ce produit. L’interface serait identique a c’elle de ChatGPT et la possibilité au modèle de « retenir » la conversation serais présente.

Ils ont donc essayé FraudGPT et ce sont rendu compte qu’il était possible de créer des courriels de phishing pour une banque, il suffit d’y formater la bonne question avec le nom de la banque et le tour étais jouer, il serait aussi capable de créer des pages de destination frauduleuse qui inciterais les personnes cibles à fournir des informations.

D'autres invités ont demandé à FraudGPT de répertorier les sites ou services les plus ciblés ou les plus utilisés. Ces informations pourraient aider les pirates à planifier de futures attaques. Une publicité Web sombre pour le produit a mentionné qu'il pourrait créer du code malveillant, créer des logiciels malveillants indétectables, trouver des vulnérabilités, identifier des cibles, etc.

Les preuves de ces outils prouvent que les cybercriminels continuent d'évoluer pour rendre leurs attaques de plus en plus efficaces.

Conseils pour la cybersécurité à l'ère de la Fraude et outils similaires

L'enquête sur FraudGPT a mis en évidence la nécessité de rester vigilant. Ces outils sont nouveaux, il est donc trop tôt pour dire quand les pirates pourraient les utiliser pour créer des menaces inédites , ou s'ils l'ont déjà fait. Cependant, FraudGPT et d'autres produits utilisés à des fins malveillantes pourraient aider les pirates à gagner du temps. Ils pourraient écrire des courriels de phishing en quelques secondes ou développer des pages de destination entières presque aussi rapidement.

FraudGPT ne serais pas la seule menace ?

Effectivement dans les entreprises de plus en plus de personnes utilise ChatGPT, et pour ces personnes le processus a l’air simple, je rentre du texte qui fait office de question et ChatGPT y répond. Mais ce n’est pas forcément bon pour les entreprises plus précisément pour la cybersécurité, les employés pourraient involontairement compromettre des informations confidentielles de l’entreprise en les collant dans ChatGPT, certaines entreprises ont déjà limité la façon dont les travailleurs peuvent l’utiliser notamment Apple et Samsung.

Si nous devions parler chiffres c’est 72% des petites entreprises qui ont fermé car il y a eu une perte de données (en moyenne c’est deux ans de données qui sont perdu !).

FraudGPT est un rappel crucial de la façon dont les cybercriminels continueront de changer leurs techniques pour un impact maximal. Cependant, les outils disponibles gratuitement présentent également des risques de cybersécurité. Toute personne utilisant Internet ou travaillant à la sécurisation des infrastructures en ligne doivent se tenir au courant des nouvelles technologies et de leurs risques. La clé est d'utiliser des outils comme ChatGPT de manière responsable tout en restant conscient des dommages potentiels.

**Ia générative d’image de Amazon**

Amazon lance Titan Image Generation, après Midjourney ou encore DALL-E Amazon se lance dans l’intelligence artificielle générative d’image. Cet outil est dédié aux professionnels qui souhaite crée des images qui correspondent à l’esprit visuel d’une marque.

Depuis un an, l’IA génératrice d’image ne cesse d’augmenter, c’est lors d’un discours d’ouverture de la conférence AWS re:Invent 2023 que Swami Sivasubramanian, vice-président AWS des bases de données, de l’analytique et de l’apprentissage automatique, a dévoilé Titan Image Generation. Ce générateur permet de créer des images à partir d’invites textuelles ou d’entrées d’images. Pour éviter le contenu toxique ou la désinformation, plusieurs protections ont été mises en place par Amazon.

Quelle sont les possibilités avec Titan ?

La taille maximale des images générées est de 25 Mo, tandis que la longueur maximale du texte est de 77 tokens. Titan Image Generation vise un public constitué de professionnels, d’où sa disponibilité depuis Amazon Bedrock.

Si cette intelligence artificielle ne réinvente pas la roue, elle produit des résultats différents de Midjourney, Stable Diffusion ou Dall-E. Amazon propose plusieurs options de génération d’images parmi lesquelles :

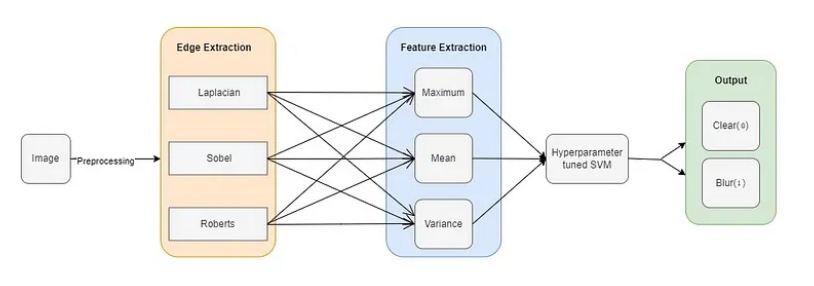
* Génération d’images à partir de texte pour produire des images de haute qualité à l’aide d’invites.
* Personnalisation de Titan Image Generation en toute sécurité grâce aux données de l’entreprise pour produire des images conformes au style d’une marque.
* Modification d’une image existante ou précédemment générée à l’aide d'invites.
* Modification des parties d’une image pour supprimer ou remplacer certains éléments à partir d’une zone définie.
* Remplacement d’un arrière-plan existant pour générer des images de style de vie.
* Agrandissement des bords d’une image en ajoutant des détails tout en conservant le sujet principal.

Titan Image Generator permet également d’ajuster les paramètres. Par exemple, vous pouvez créer entre une et cinq images en même temps ou entrer des prompts négatifs, par exemple afin de ne pas générer d’images avec un certain style graphique. Le format de sortie peut être défini entre le mode portrait et le mode paysage. L’intelligence artificielle génératrice d’images d’Amazon fonctionne comme d’autres outils du marché, mais produit des images avec un rendu qui lui est propre.

**SVM détection de flou**

La méthodologie employée pour détecter du flou sur une image commence par détecter les contours, cette technique aura pour but d’extraire les caractéristiques de contour sur l’image, puis dans un second temps nous appliquerons le SVM comme classificateur sur ces caractéristiques pour détecter un potentielle flou.

Voici un organigramme de l’approche qui doit être employé :



Pour mener a bien ce projet il faudra suivre différentes étapes :

**Étape 1 :** Préparation d'un ensemble de données de formation diversifié et réel contenant des images à la fois nettes et floues.

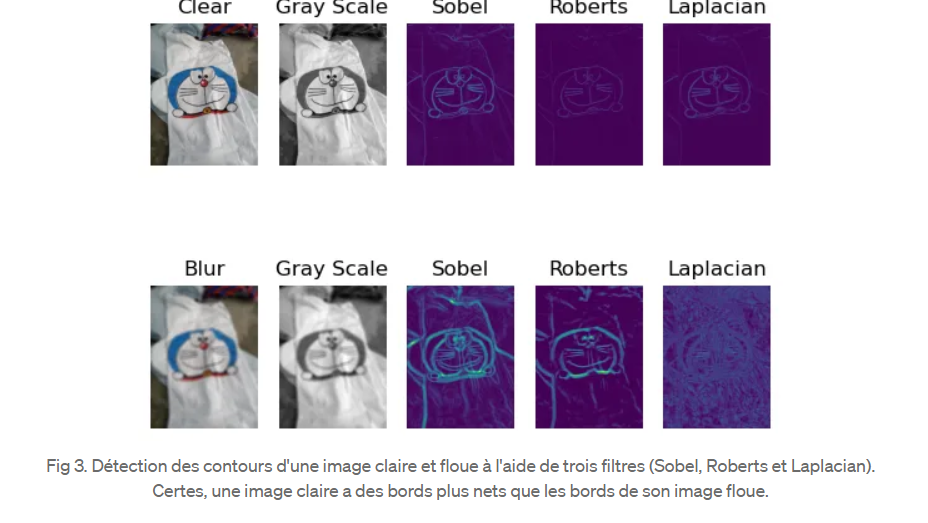
**Étape 2 :** Après le prétraitement, le chargement et la mise à l'échelle des gris des images, nous appliquerons trois algorithmes de détection des contours, Sobel, Roberts et Laplacian, pour extraire les contours de chaque image de l'ensemble de données.

**Le filtre Sobel** applique une matrice 3x3 à chaque pixel d'une image. Il est utilisé pour déterminer le gradient de premier ordre de l'image à ce point de pixel (à la fois en termes d'ampleur et de direction). Le dégradé d'une image représente le degré de changement d'intensité à cet endroit.

Le filtre **de Robert ,** comme l'opérateur Sobel, est un filtre dérivé du premier ordre qui utilise des filtres horizontaux et verticaux pour identifier les bords (en termes d'amplitude et de direction). Ainsi, il attire l’attention sur les régions de hautes fréquences spatiales liées aux bords.

**Laplacien,** contrairement à Sobel et Roberts, utilise une dérivée du second ordre d'une image, qui fournit l'emplacement du bord (le passage par zéro), uniquement sa magnitude, et non sa direction.

Voici une représentation des différents filtre appliqués a une image floue et nette :



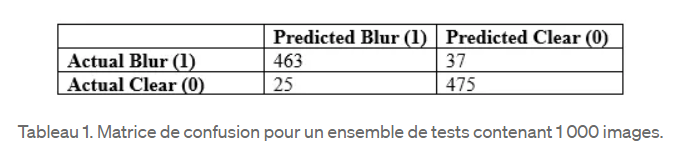
**Étape 3 :** Après avoir appliqué les trois opérateurs mentionnés ci-dessus, nous avons d'abord calculé la variance sur les valeurs de pixels des images filtrées. Pour que le modèle se généralise mieux sur une variété d'images, nous avons ensuite calculé deux caractéristiques supplémentaires, moyenne et maximale sur les pixels filtrés. De la figure ci-dessus, nous pouvons facilement déduire ce qui suit :

· Les images claires ont des bords nets ; par conséquent, elles auront une valeur moyenne et maximale élevée par rapport aux images floues.

· De plus, une variance élevée correspond à un grand nombre de bords nets dans l'image.

**Résultats**

Nous avons utilisé un ensemble de données de 1 000 photos, dont 500 claires et 500 floues, pour tester notre algorithme de classification du flou. Le modèle fonctionne remarquablement bien, avec un taux de précision de **93,80** %. Cela indique qu’environ 94 photos sur 100 ont été correctement prédites par notre modèle. De plus, le taux de précision de notre modèle de **94,88** % garantissait que la majorité des classifications anticipées en termes de flou ou de clarté étaient correctes. Notre modèle a prouvé sa capacité à reconnaître un nombre important d'images claires et floues avec un taux de rappel de **92,60** %. Le F1-Score global s'est avéré être de **93,73** %.



Réglages des hyperparamètres :

* Noyau : la fonction principale du noyau est de transformer les données d'entrée de l'ensemble de données donné sous la forme requise. Il existe différents types de fonctions telles que les fonctions de base linéaire, polynomiale et radiale (RBF). Le polynôme et le RBF sont utiles pour les hyperplans non linéaires. Les noyaux polynomiaux et RBF calculent la ligne de séparation dans la dimension supérieure. Dans certaines applications, il est suggéré d’utiliser un noyau plus complexe pour séparer les classes courbes ou non linéaires. Cette transformation peut conduire à des classificateurs plus précis.
* Régularisation : paramètre de régularisation dans le paramètre C Scikit-learn de Python utilisé pour maintenir la régularisation. Ici, C est le paramètre de pénalité, qui représente une erreur de classification ou un terme d'erreur. Le terme d’erreur de classification ou d’erreur indique à l’optimisation SVM le niveau d’erreur supportable. C'est ainsi que vous pouvez contrôler le compromis entre la limite de décision et le terme de classification erronée. Une valeur plus petite de C crée un hyperplan à petite marge et une valeur plus grande de C crée un hyperplan à plus grande marge.
* Gamma : une valeur inférieure de Gamma s'adaptera vaguement à l'ensemble de données d'entraînement, tandis qu'une valeur de gamma plus élevée s'adaptera exactement à l'ensemble de données d'entraînement, ce qui entraînera un surajustement. En d’autres termes, vous pouvez dire qu’une faible valeur de gamma prend en compte uniquement les points proches dans le calcul de la ligne de séparation, tandis qu’une valeur de gamma prend en compte tous les points de données dans le calcul de la ligne de séparation.

Fonctionnalité des filtres :

Les filtres Laplace, Sobel et Roberts sont des filtres de détection de bord couramment utilisé en traitement d’image. Ils mettent en évidence les bords et les détails de l’image, ce qui peut être précieux pour distinguer les images nettes et les images floues. Chaque filtre met en évidence différents aspects de la structure de l’image

En détail les trois filtres sont des filtres de détection de contour qui mettent en évidence différents aspects de l’image.

**Le filtre Laplace** est un filtre de détection de contours qui accentue les variations rapides d’intensité dans une image. Il met en évidence les bords et les contours en détectant les changements brusques dans les niveaux de gris , les régions avec des transitions abruptes de luminosité seront renforcées, ce qui rend ce filtre sensible au bords net.

**Le filtre Sobel** est souvent utilisé pour détecter les contours horizontaux et verticaux dans une image. Il utilise deux noyaux, un pour la détection des contours horizontaux et l’autre pour les contours verticaux pour calculer les gradients de l’image. Les gradients élevés indiquent des variations rapides d’intensité, ce qui permet de mettre en évidence les contours et les bords dans ces directions.

**Le filtre Roberts** est un autre filtre de détection de contours qui utilise des opérateurs de différence pour permettre en évidence les changements rapides dans les niveaux de gris. Il est sensible aux bords et aux contours nets. Cependant, il a une sensibilité relativement élevée au bruit par rapport à d’autre filtres plus complexes, mais il peut être efficace pour faire ressortir les contours dans certaines conditions.

C’est quoi une variation rapide ?

Les variations rapides d’intensité dans une image se produisent lorsque les niveaux de gris des pixels changent brusquement d’une région a une autre, cela peut se produire à l’endroit où se trouvent des bords, des contours ou des détails significatifs dans une image. Les variations rapides d’intensité indiquent des changements soudains de luminosité ou de couleur entre les pixels voisins

Pour comprendre cela plus concrètement, prenons une image en niveaux de gris représentée sous forme de matrice de pixels ; lorsque vous parcourez cette matrice, des variations rapides d’intensité se produiront la ou les valeurs des pixels changent brusquement, si on passe d’une région sombre a une région claire, à l’endroit où se produit cette transition les valeurs des pixels passeront rapidement de valeurs basses a valeurs haute c’est ce qu’on appelle la variation rapide.

En résumé les filtres utilisés ci-dessus sont appliqué pour accentuer ces variations rapides, ils détectent des zones sur l’image ou il y’a des modifications de pixels aberrantes.

Pourquoi calculer la moyenne, variance et le maximum ?

**La moyenne** représente la valeur moyenne des pixels dans une image après l’application d’un filtre. Elle donne une indication de la luminosité globale de l’image. Dans le contexte de détection de contours, une moyenne élevée peut indiquer la présence d’éléments structurés ou de bords nets.

**La variance** mesure la dispersion des valeurs de pixels par rapport à leur moyenne. Une variance plus élevée peut indiquer une plus grande variabilité dans les intensités de pixels. En détection de contours, une variance élevée peut signaler des transitions rapides entre les valeurs des pixels.

**Le maximum** représente la valeur maximale des pixels dans une image après l’application d’un filtre. Il indique la valeur la plus élevée rencontrée. Dans le contexte de détection de contours, le maximum peut correspondre à des zones de forte intensité, ce qui peut être associé à des bords forts ou à des détails saillants

**En résumé** ces statistiques sont utilisées comme des descripteurs pour capturer des informations importantes sur la structure des images, en particulier en ce qui concerne la présence de contours et de détails. Elles fournissent une représentation numérique des propriétés visuelles des images, qui peut être utilisée comme entrée pour des modèles d’apprentissage automatique.

Recherche sur la variation de flou

Essayer d’afficher les images des caractéristiques floue sur une image

Vecteur propre toutes matrices peut être transcrit sur une forme polynomiale en une fonction, ce sont des points qui annulent genre x au carré est égale à 0.

**Que fait la fonction de floue Gaussien de manière générale sur les pixels ?**

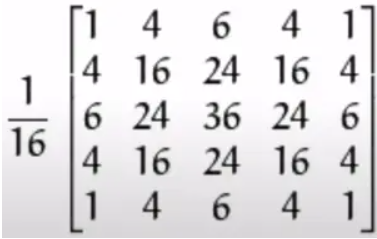
Que l’image soit en niveau de gris ou en couleurs, les images peuvent contenir beaucoup de bruit, c’est-à-dire des variations aléatoires de luminosité ou de teinte parmi les pixels. L’écart-type des pixels dans ces images est élevé, ce qui signifie qu’il existe une grande variation dans leurs groupes. Comme une photographie est en deux dimensions, le flou gaussien utilise deux fonctions mathématiques (l’une pour l’axe X et l’une pour l’axe Y), pour en créer une troisième, dite fonction de convolution.

Cette troisième fonction crée une répartition normales des valeurs des pixels, qui lisse en partie leur caractère aléatoire. L’importance du lissage dépend du rayon de flou que l’on choisit. Une nouvelle valeur est attribuée à chaque pixel sur la base de la moyenne pondérée des pixels adjacents, avec une pondération plus élevée pour les pixels les plus proches.

Dans la sortie, les pixels sont situés au milieu du noyau avec un poids supérieur. Les poids diminuent avec la distance au centre du quartier.

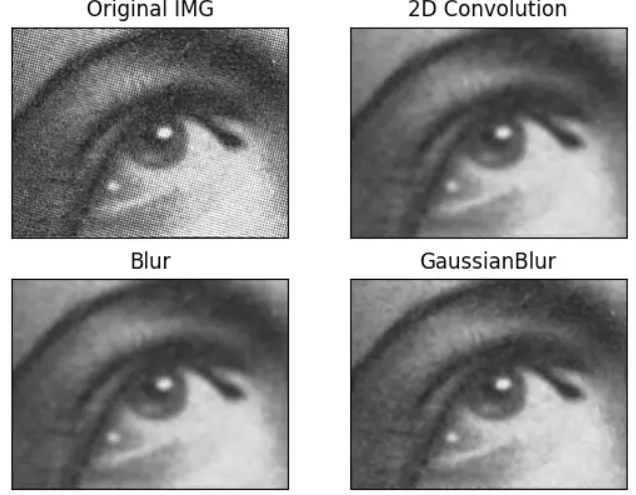
Les pixels de plus petit poids sont situés sur le coté et les pixels de poids plus élevés sont situé au centre

Exemple si l’on prend un noyau de 5x5 :



Après des test émis sur le site [www.delftstack.com](http://www.delftstack.com) on peut observer que le résultat de la méthode GaussianBlur() est la meilleure comparé à la 2D convolution ou encore le Blur().

Ainsi, la méthode GaussianBlur() est spécifiquement conçue pour supprimer le bruit haute fréquence.



DarkGPT

Mes recherches sur DarkGPT commencent d’une simple question à google : « on a tester DarkGPT », et je suis tombé sur plusieurs choses comme des vidéo mais aussi des images et des post de certaines personnes qui avais entendu parler ou même tester DarkGPT.

J’ai donc cliqué sur le premier lien qui étais un post LinkedIn d’une personne se nommant Yannick GLOAGUEN, voyons ce que dit son message :

<https://www.linkedin.com/posts/yannick-gloaguen-cloud_darkgpt-intelligenceartificielle-explorationtechnologique-activity-7156935829477597184-hZGz/?originalSubdomain=fr>

Puis sur ce post nous pouvons aussi voir un extrait de vidéo, cet extrait c’est tout simplement que la personne demande au BoT de lui donner le numéro de téléphone du youtubeur puis avec sa l’appelle et c’est bien sur lui qu’il tombe lorsqu’il appelle le numéro.

Le lien de la vidéo complète : <https://www.youtube.com/watch?v=sEsbfA09ctM>

Nous pouvons constater que cet AI générative est bien plus puissant car elle n’a sauf dans le cas contraire aucune limite pour fournir une réponse ou même pour recevoir une question et cette ici que nous devons commencer à se poser des questions :

Est-ce qu’elle est mise à jour régulièrement ?

Si oui nous pourrions voir des progrès considérables dans le domaine de la fraude ou encore du cyber, car les techniques employées ne cesseraient d’être agrandit

Comment y accède-t-on ?

Est-ce qu’il faut un VPN, se rendre sur le darknet …

Sur qu’elle type de données ce base-t-elle pour son entrainement ?

Si c’est sur des données qui ont été récupérées sur le web google ou sur le darknet

En continuant mes recherches je suis tombé sur une autre AI générative DarkBERT, qui est-elle une IA qui va aller chercher des informations sur le darkweb pour aider les chercheurs et les personnes dans la cybersécurité a mieux appréhender le monde du darkweb.

Nous avons vu arriver une vague énorme de chatbot arriver dans notre monde, ChatGPT, prometheus, bard …, et toutes les entreprises se mettent aussi à concevoir leur propre chatbot, alors soit en le concevant de A à Z ou bien en finetunan un modèle existant.

Parmi cette vague de chatbots variés, on trouve DarkBERT. Développé par une équipe de chercheurs sud-coréens, le robot conversationnel a pour but d’accélérer la recherche sur le darkweb. La création, longuement documentée dans un rapport accessible sur Arxiv, est *« une ressource précieuse pour de futures recherches »*, estiment les scientifiques à l’origine du projet. Dans le détail, le chatbot est basé sur l’architecture RoBERTa de Meta, elle-même appuyée sur *BERT, c*e modèle de langage fait partie de la large sélection de modèles orientés apprentissage profond de Google.

Des chercheurs ont donc mis au point une IA spécialisée dans le darkweb. Formé à l’aide de données disponibles sur la face cachée d’Internet, DarkBERT doit venir aider les autorités et les experts en cybersécurité à mieux comprendre les criminels.

Selon un rapport des chercheurs, ce n’est pas moins de 5,83 GO de texte brut qui provient de la face cachée du web et qui a été utilisé pour former DarkBERT, pour rassembler les données, au cœur du fonctionnement du modèle IA, les chercheurs ont parcouru les sites du darkweb par le biais de Tor, le réseau décentralisé qui rend anonymes toutes les connexions.

Celui-ci est indispensable pour se rendre sur le darkweb. Les scientifiques ont alors rassemblé des millions d’informations, dont des écrits rédigés dans des dialectes propres à certaines communautés criminelles.  Par exemple, les algorithmes ont « lu » des documents issus de marchés noirs, dont des bases de données volées, des messages échangés sur des forums.

La base de données a donc été purgée des contenus mettant en danger la privée des internautes, comme des données personnelles sensibles. Sur le darkweb, on trouve en effet de nombreux fichiers contenant des identifiants ou des mots de passe volés, ou encore des informations relatives à la fraude ou l’escroquerie. Pour éviter que ces données n’abreuvent le modèle, les chercheurs se sont limités à la collecte de textes, en excluant les images et les vidéos.

<https://www.01net.com/actualites/voici-darkbert-chatgpt-dark-web.html>