Pixel Patrol

Louis Jean

Master 1 IMAGINE
Faculté des Sciences, Université de Montpellier

23 avril 2024

Détection de falsifications dans des images

Sommaire

Section Overview

Introduction

Contexte et enjeux

- On estime à 760 milliards le nombre d'images en circulation sur internet
- L'authenticité visuelle est une monnaie de confiance
- Les images sont très puissantes pour façonner l'opinion publique
- L'évolution de la technologie a grandement simplifié la création de contenus falsifiés convaincants
- Ces manipulations posent des risques importants : propagation de fausses informations, atteintes à la réputation, conséquences sécuritaires
- Être capable de dêmeler le vrai du faux dans les images est donc primordial pour maintenir l'intégrité informationnelle

Section Overview

État de l'art

- Types de falsifications et enjeux.
- Bases de données utilisées pour le test.
- Méthodes traditionnelles et utilisation des réseaux de neurones convolutifs.

Les méthodes de falsification

- **Copy-move :** duplication d'une partie de l'image pour masquer ou cloner des objets
- **Splicing :** combinaison de morceaux de différentes images pour créer nouvelle image
- **Removal :** suppression d'un élément d'une image
- **Deepfakes :** utilisation de l'intelligence artificielle pour remplacer les visages ou simuler des vidéos
- **Quelques exemples parmi tant d'autres :**
  - Photos retouchées de célébrités influençant les normes esthétiques.
  - Falsifications d'images de surveillance utilisées pour incriminer des innocents.
  - ...

Les méthodes de falsification : exemples

Figure 1 – Falsification par copy-move

Figure 2 – Falsification par splicing

Figure 3 – Falsification par removal

Figure 4 – Falsification par deepfake

Méthodes de détection actuelles

Pour détecter les falsifications, plusieurs approches ont été développées :

- Cohérence des pixels : couleur, bruit, motifs, et estimation de la source de lumière.
- Analyse en fréquence : utilisation de la `FFT`, filtres passe-haut, et amplification des artefacts.
- Examen des métadonnées et techniques de tatouage numérique (watermarking).

Les méthodes contemporaines s'appuient de plus en plus sur l'intelligence artificielle :

- Réseaux de neurones (`CNN`) pour la reconnaissance de motifs
- Machines à vecteurs de support (`SVM`) pour la classification

Ces méthodes plus récentes montrent une efficacité accrue par rapport aux techniques traditionnelles.

Bases de données disponibles

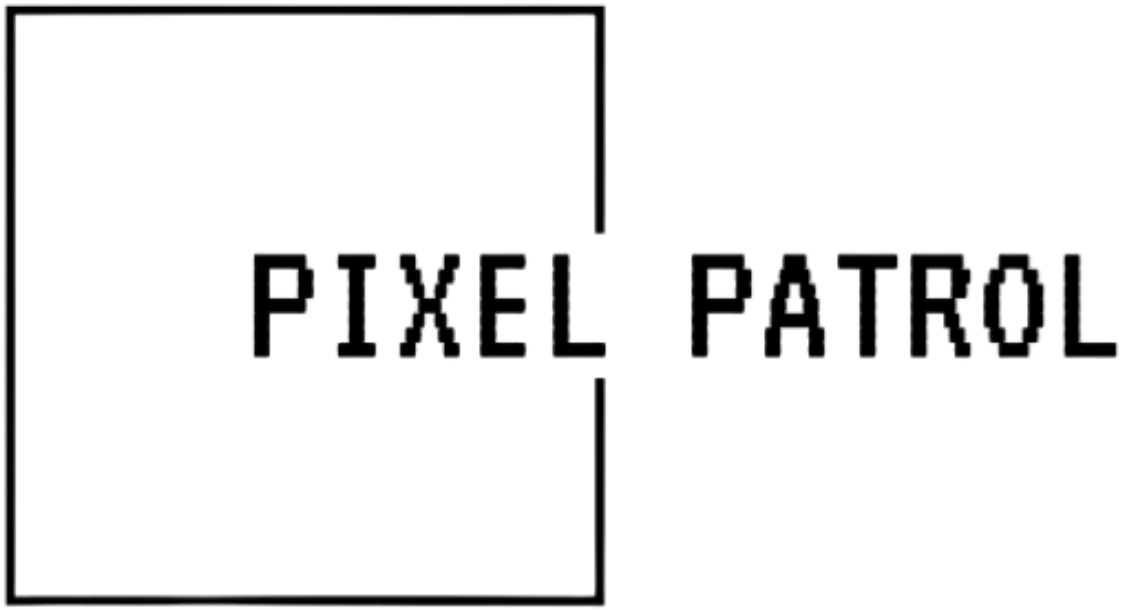
Il existe des bases de données spécialement conçues pour la détection de falsifications des images. En voici une liste non-exhaustive :

- **Splicing, copy-move, removal :** CASIA 1.0, CASIA 2.0, NIST, ...
- **Splicing :** Wild Web, MISD, VIPP, ...
- **Copy-move :** CoMoFoD, COVERAGE, ...

Section Overview

Pixel Patrol

Pixel Patrol est une application visant à



Méthodes implémentées dans Pixel Patrol

Pixel Patrol inclut trois méthodes différentes dans leurs implémentations et leurs objectifs.

- **Première méthode :** excellant dans la détection de falsification par copy-move, elle utilise les algorithmes `SIFT` et `RANSAC`.
- **Deuxième méthode :** spécialisée dans la détection de falsification par splicing et removal, elle fait appel à la `DCT`.
- **Troisième méthode :** ayant pour unique but de classifier les images authentiques ou falsifiées, elle utilise le `LBP`, la `DCT` et un `SVM`.

Première méthode : détection de copy-move

Cette méthode est tirée de l'article [1], qui propose une approche robuste utilisant `SIFT` et `RANSAC`.

- **Qu'est-ce que SIFT (Scale-Invariant Feature Transform) ?**
  - Algorithme mis au point en 1999 par David Lowe
  - Détecte et décrit les points d'intérêt locaux dans les images
  - Les caractéristiques sont invariantes à l'échelle, la rotation et partiellement au changement de perspective
- **Qu'est-ce que RANSAC (RAndom SAmple Consensus) ?**
  - Algorithme publié par Fischler et Bolles en 1981
  - Estime les paramètres d'un modèle mathématique à partir d'un ensemble de données observées
  - Sélectionne un sous-ensemble aléatoire des points, construit un modèle de transformation possible, puis vérifie le nombre de points s'adaptant à ce modèle
  - Répète le processus pour trouver le modèle le plus robuste

Choix de SIFT et RANSAC pour la détection de copy-move

- **Pourquoi SIFT ?**
  - **Robustesse :** `SIFT` détecte des points d'intérêt qui sont invariants à l'échelle, la rotation, et partiellement à la translation, le rendant idéal pour identifier des régions similaires dans différentes parties de l'image
  - **Précision :** les descripteurs obtenus fournissent une représentation détaillée des points-clés, permettant une comparaison précise entre les régions copiées et déplacées

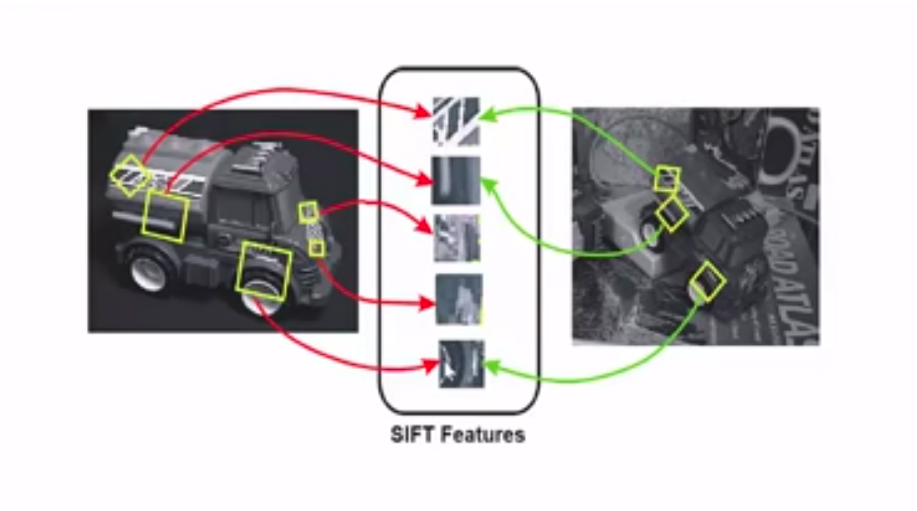


Figure 5 – Exemple de capture de caractéristiques `SIFT`

Choix de SIFT et RANSAC pour la détection de copy-move

- **Pourquoi RANSAC ?**
  - **Fiabilité :** `RANSAC` est robuste aux données aberrantes et peut donc trouver les transformations appliquées aux zones copiées même si certains points d'appariement sont incorrects
  - **Sélectivité :** le processus itératif garantit le meilleur modèle qui explique le plus grand nombre de points entre les régions suspectées d'être copiées, minimisant ainsi l'impact des erreurs de détection.

Figure 6 –

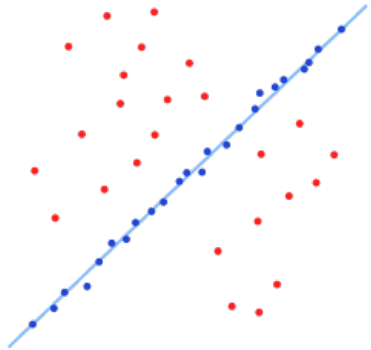


Figure 7 –

Section Overview

Conclusion

Section Overview

Références I

[1] Gonapalli Ramu et S. B. G. Thilak Babu. “Image forgery detection for high resolution images using SIFT and RANSAC algorithm”. In : *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*. 2017, p. 850-854. doi : 10.1109/CESYS.2017.8321205.