



École Nationale Supérieure d'Informatique et d'Analyse des Systèmes

Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

AVIS DE SOUTENANCE DE THÈSE DE DOCTORAT

Madame Wahiba ABOU-ZBIBA

Soutiendra publiquement sa thèse de Doctorat en Informatique

Le 09 Octobre 2026 à 15h00 au Grand Amphi à l'ENSIAS

Intitulé de la thèse

**Unified Image Forgery Detection: Investigations, Mechanisms, and
Forensic Artifacts**

Président :

Pr. Sanaa EL FKIH, PES, ENSIAS, Université Mohammed V, Rabat

Directeur de thèse :

Pr. Houda BENBRAHIM, PES, ENSIAS, Université Mohammed V, Rabat

Co-encadrant de thèse :

Pr. Driss BENHADDOU, Professeur et doyen, College of Engineering, Alfaisal University,
Riyadh, Saudi Arabia

Rapporteurs :

Pr. Amina RADGUI, PES, INPT, Rabat

Pr. Mohammed RZIZA, PES, FSR, Université Mohammed V, Rabat

Pr. Abdelalim SADIQ, PES, FSK, Université Ibn Tofail, Kénitra

Examineur(s) :

Pr. Hanan EL BAKKALI, PES, ENSIAS, Université Mohammed V, Rabat

Pr. Mahmoud EL HAMLAOUI, MCH, ENSIAS, Université Mohammed V, Rabat.

Résumé :

Depuis des décennies, les images ont servi d'archivistes fidèles de la vérité, agissant pendant si longtemps comme des témoins silencieux de la réalité. Dans le cas des images numériques, on croyait que la vérité s'exprimait à travers chaque pixel. Outre la communication visuelle qu'elles visent à transmettre, les images numériques ont acquis une importance considérable grâce à leur rôle de preuve cruciale dans de nombreux domaines, allant du journalisme et des enquêtes judiciaires à la documentation publique et aux informations sociales quotidiennes. Néanmoins, au cours des dernières années, les images numériques ont malheureusement perdu progressivement leur crédibilité spontanée, en raison de la propagation des outils de retouche d'images pratiques et de leur accessibilité aisée pour la plupart des gens. Parallèlement, les chercheurs dans le domaine de l'analyse forensique d'images numériques s'efforcent de résoudre le problème de la détection des falsifications d'images, dans le but d'apporter des solutions garantissant la fiabilité des images.

Dans cette thèse, nous nous intéressons à la détection de la falsification d'images, plus précisément à deux types de manipulation, à savoir la manipulation par copier-déplacer et la manipulation par épissage d'images. À travers ces travaux, nous visons à combler plusieurs lacunes dans la littérature.

Tout d'abord, nous menons une revue systématique de la littérature (SLR) afin d'étudier et de comparer l'utilisation de modèles unifiés, destinés à détecter à la fois les manipulations par copier-déplacer et par épissage d'images, à celle de modèles distincts. Ensuite, en nous appuyant sur les conclusions de cette revue, nous proposons SLIF, un modèle unifié et léger permettant de détecter ces deux types de manipulations. Ce modèle associe des éléments conçus manuellement à une architecture de réseaux neuronaux convolutifs, révélant des résultats compétitifs tout en conservant une structure légère. Ainsi, afin de fournir des informations complémentaires dans ce domaine, nous menons des recherches supplémentaires sur l'impact des principaux choix de prétraitement sur les performances de détection. Les conclusions de ces recherches ont conduit au développement d'un nouveau concept d'artefacts, nommé Resizing Residuals Artifacts (RRA), autrement dit, artefacts résiduels de redimensionnement, suivi de la proposition du modèle RRA-Net.

De plus, afin d'approfondir notre compréhension du problème abordé, et en nous appuyant sur les résultats des travaux antérieurs, nous menons une étude sur le rôle des mécanismes d'attention dans la décision de classification des modèles unifiés de détection de falsification d'images basés sur les réseaux neuronaux convolutifs, entraînés sans supervision au niveau des pixels. À cette fin, nous proposons un module d'attention à double focalisation (DFAM) dans lequel l'étude a été menée, et celle-ci s'est déroulée dans le cadre de l'IA explicable. Par ailleurs, les artefacts RRA proposés ayant démontré des résultats compétitifs en tant qu'indices d'analyse, nous nous sommes intéressés à une éventuelle identification de la tâche de détection de falsification d'images comme une tâche de détection d'anomalies séquentielles, sur la base des caractéristiques RRA. Ensuite, en nous inspirant du concept de RRA, nous menons une analyse systématique à travers une analyse qualitative des résidus, une analyse statistique quantitative et une évaluation comparative approfondie basée sur les caractéristiques à l'aide de réseaux siamois, afin d'étudier l'impact différentiel des artefacts de reconstruction d'images sur les images authentiques et falsifiées, dans le cadre de deux modèles de génération d'images représentant deux paradigmes génératifs: les réseaux antagonistes génératifs (GAN) et les auto-encodeurs variationnels (VAE). Les conclusions de cette analyse systématique apportent de nouvelles perspectives sur la détection de la falsification d'images selon deux angles: les aspects forensiques et anti-forensiques.

Enfin, nous présentons SI-CAR, une nouvelle architecture de Mobile CrowdSensing pour collecter des données liées à la sécurité routière. Celle-ci sert de cas d'usage intégrant le modèle proposé SLIF pour assurer une collecte de données fiables. À travers ce cas d'usage, notre objectif est de montrer l'utilité pratique de notre travail et de souligner l'importance de garantir la fiabilité des images.

Mots-clés:

Forensique numérique, Détection de falsification d'images, Copier-déplacer, Épissage d'images, Modèle unifié et léger, Traitement d'images, Apprentissage profond.

Abstract:

For decades, images have served as loyal archivists of truth, acting for so long as silent witnesses to reality. In the case of digital images, truth was believed to be spoken through each pixel. Alongside the visual communication that digital images aim to convey, they have gained considerable importance thanks to their role as crucial evidence in many fields, ranging from journalism and legal investigations to public documentation and daily social news. Nevertheless, over the last years, digital images have unfortunately lost their spontaneous credibility gradually, as a result of the widespread user-friendly image editing tools and their easy accessibility to most people. Meanwhile, researchers in the field of digital forensics are committed to addressing the issue of image forgery detection, with the aim of providing solutions to ensure image reliability.

In this thesis, we address the task of image forgery detection, focusing on two types of manipulation, namely, copy-move and image splicing. Through this work, we aim to address several gaps in the literature. First, we conduct a Systematic Literature Review (SLR) to study and compare the use of unified models to detect both image splicing and copy-move forgeries to separate models. Subsequently, and building upon the conclusions of the SLR, we propose SLIF, a unified lightweight model to detect both these forgeries. This model combines handcrafted features with a Convolutional Neural Network (CNN) architecture, demonstrating competitive results while maintaining a lightweight aspect. Additionally, to provide further insights in this field, we carry out complementary investigations on the impact of main pre-processing choices on the detection performance. The findings of these investigations motivated the idea of our proposed artifacts, Resizing Residuals Artifacts (RRA), followed by the proposition of the RRA-Net model.

Furthermore, to gain more insights regarding the addressed problem, and motivated by the findings of the previous work, we perform an investigation on the role of attention mechanisms in the classification decision of unified image forgery detection CNN-based models trained without pixel-level supervision. For this purpose, we propose a Dual Focus Attention Module (DFAM) within which the study is conducted, under the framework of explainable AI. With the proposed RRA artifacts demonstrating competitive results as a forensic cue, we explored whether the image forgery detection task could be identified as a sequential anomaly detection task, based on the RRA features. In addition, and influenced by the concept of RRA, we

conduct a systematic analysis through a qualitative residual analysis, a quantitative statistical analysis, and a deep feature-based comparative evaluation using Siamese networks, in order to investigate the differential impact of image reconstruction artifacts on authentic and forged images, under two image generation models representing two generative paradigms; Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). The findings of this systematic analysis provide insights into the image forgery detection task with respect to two aspects: the forensic and the anti-forensic perspectives.

Finally, we present SI-CAR, a new Mobile CrowdSensing architecture for data collection regarding road safety concerns, as a use case featuring the proposed model SLIF to ensure reliable data collection. Through this use case, we aim to highlight the practical use and emphasis on ensuring image reliability.

Keywords:

Digital Forensics, Image Forgery Detection, Copy-move, Image Splicing, Unified Lightweight Model, Image Processing, Deep Learning.