École d'Ingénieurs du Littoral Côte d'Opale, Calais
Étude bibliographique, ING3 informatique
spécialité Data Science
Keepzy - Lille

Systèmes d'Intelligence Artificielle pour l'authentification des documents locatifs

Mots clés : Document verification, Blockchain, Artificial Intelligence, Machine learning for document verification, Real estate document verification

Nombre de mots : 2531

Etudiante : NIANG Fatou
Encadrant Chercheur : Mr. Mourad BOUNEFFA
Tuteur Entreprise : Mr. Oussama HAMZAOUI

Année universitaire : 2023/2024

Table des matières

1 Méthodologie de recherche

Pour réaliser cette revue bibliographique, nous avons suivi une méthodologie rigoureuse. Nous avons débuté par une recherche minutieuse dans des bases de données académiques renommées telles que Google Scholar, PubMed, IEEE Xplore et ScienceDirect, axée sur l'authentification des documents, notamment dans le domaine immobilier, et mettant en lumière l'utilisation de l'intelligence artificielle et/ou de la technologie blockchain. Les articles sélectionnés ont été évalués selon leur pertinence et leur qualité académique, en privilégiant ceux publiés dans des revues à comité de lecture ou des conférences réputées. Chaque article retenu a été minutieusement analysé pour extraire les méthodes, résultats et recommandations concernant l'authentification des documents, avec un focus particulier sur l'intégration de l'intelligence artificielle et de la blockchain. En synthétisant ces informations, nous avons dressé un portrait exhaustif de l'état de l'art dans ce domaine, en mettant en avant les tendances émergentes et les perspectives futures.

Mots clés : Document verification, Blockchain, Artificial Intelligence, Machine learning for document verification, Real estate document verification

2 Introduction

Dans un monde de plus en plus numérisé où les transactions immobilières et locatives se multiplient, l'authentification des documents locatifs est devenue une préoccupation majeure. La falsification de contrats de location et d'autres documents pertinents constitue une menace sérieuse pour la sécurité juridique des parties impliquées. Pourtant, avec l'avènement de technologies innovantes telles que la blockchain et les systèmes d'intelligence artificielle (IA), de nouvelles solutions émergent pour relever ce défi.

2.1 Contexte et Problématique

La falsification de documents d'identité et d'images est devenue omniprésente de nos jours, largement alimentée par les avancées technologiques accessibles à tous. Cette pratique illégale comprend la falsification de divers documents tels que les cartes d'identité, les passeports et les bulletins de paie, dans le but de tromper les autorités ou d'obtenir indûment certains avantages. Les fraudeurs réussissent souvent à contourner les systèmes de sécurité en ajoutant simplement un champ de texte sur le document original, altérant ainsi l'identifiant unique et attribuant une fausse identité qui peut passer inaperçue à l'œil humain. Ces manipulations menacent la sécurité des transactions et des informations personnelles, soulignant ainsi l'urgence de solutions fiables et sécurisées pour authentifier les documents et prévenir la fraude [1].

L'introduction de la technologie blockchain par Satoshi Nakamoto a été une avancée majeure dans la quête de solutions contre la fraude documentaire. Dans son article publié en 2008, Satoshi Nakamoto a proposé un système de paiement électronique basé sur des contrats intelligents, éliminant ainsi la nécessité d'intermédiaires entre les transactions peer-to-peer. L'avènement de la cryptomonnaie bitcoin en 2009 a ensuite popularisé la technologie blockchain, lui conférant une reconnaissance et une confiance mondiales.

Depuis lors, de nombreuses applications innovantes de la technologie blockchain ont émergé, notamment dans le domaine de la vérification des documents. Cette technologie offre un potentiel prometteur pour sécuriser et authentifier les documents de manière transparente et inviolable, ouvrant ainsi la voie à de nouvelles approches dans la lutte contre la fraude documentaire [2].

2.2 Objectifs de l'étude

Les recherches effectuées mettent en évidence l'importance croissante de l'utilisation de la Blockchain pour la vérification et l'authentification des documents. Par conséquent, les objectifs de notre revue bibliographique sont les suivants :

1. Examiner les fondements théoriques de la technologie blockchain, en mettant en lumière ses forces, ses faiblesses et ses applications spécifiques à l'authentification des documents.
2. Analyser les cas d'études et les exemples concrets d'utilisation de la blockchain pour l'authentification des documents, en mettant en évidence les modèles d'intelligence artificielle associés.

3. Évaluer les performances et les limitations des solutions existantes basées sur la blockchain pour l'authentification des documents, en se basant sur une analyse critique des résultats de recherche.

En concentrant notre revue sur ces objectifs, nous cherchons à fournir un aperçu approfondi et actuel de l'utilisation de la Blockchain dans le processus d'authentification des documents, ainsi que des orientations pour la recherche future et les pratiques professionnelles dans ce domaine.

3 Fondements théoriques

3.1 Technologie Blockchain

La blockchain est une base de données distribuée qui permet à un réseau d'ordinateurs de maintenir une liste d'enregistrements en croissance continue, appelés blocs. Chaque bloc contient un horodatage et un lien vers le bloc précédent, formant ainsi une chaîne. Cette architecture permet de gérer la base de données de manière transparente et décentralisée, car tous les participants au réseau ont une copie de l'ensemble de la blockchain.

L'application la plus connue de la technologie blockchain est la crypto-monnaie Bitcoin, mais elle présente de nombreuses autres utilisations potentielles. Par exemple, elle peut être utilisée pour créer des systèmes de vote sécurisés et transparents, des systèmes de gestion de la chaîne d'approvisionnement, et même pour vérifier l'authenticité de documents.

La technologie blockchain repose sur les principes de décentralisation, de transparence et de sécurité. Elle utilise des techniques cryptographiques pour garantir que les données stockées sur la blockchain sont infalsifiables et ne peuvent pas être modifiées sans laisser de traces. Ainsi, elle est particulièrement adaptée aux situations où la confiance et la sécurité sont des préoccupations majeures [3].

3.2 Points forts de la technologie Blockchain

La technologie Blockchain offre un avantage significatif en permettant une base de données partagée sans nécessiter d'administrateur central. Contrairement à une logique d'application centralisée, les transactions sur la Blockchain sont validées et autorisées de manière autonome. Ainsi, la Blockchain agit comme un mécanisme de consensus, assurant une indépendance dans la vérification et l'autorisation des transactions [4].

3.3 Faiblesses de la technologie Blockchain

La technologie Blockchain présente également des faiblesses importantes. Tout d'abord, la vérification de la signature, bien que complexe, peut devenir un goulot d'étranglement sur le réseau, car elle nécessite des calculs intensifs. En revanche, les bases de données centralisées n'effectuent cette vérification qu'une seule fois, sans nécessiter de vérifications supplémentaires pour les demandes ultérieures. De plus, les mécanismes de consensus nécessaires pour garantir l'accord au sein du réseau peuvent entraîner une communication excessive, car l'atteinte d'un consensus implique la participation de plusieurs membres du réseau. Enfin, la redondance des traitements de chaque transaction par chaque nœud de la Blockchain entraîne un besoin accru de ressources, ce qui rend le processus plus coûteux et complexe par rapport aux bases de données centralisées [2].

3.4 Applications de la technologie Blockchain

3.4.1 Blockchain pour les crypto-monnaies

Le Bitcoin est la première crypto-monnaie à avoir implémenté la technologie de la blockchain. Inventée en 2009, elle a depuis gagné en popularité et en attrait auprès des entreprises qui ressentent le besoin d'un modèle de confiance distribué pour leurs transactions financières[2].

3.4.2 Blockchain pour l'identité numérique

L'utilisation de la blockchain pour la gestion de l'identité numérique représente une avancée majeure en termes d'intégrité des données d'identification des utilisateurs. Pour les entreprises, cela

signifie qu'elles n'ont plus à assumer les coûts liés au stockage des données d'identification des utilisateurs dans leur propre base de données. Ce modèle bénéficie également aux utilisateurs, car il leur offre un accès facile à leurs données d'identité et leur donne davantage de contrôle sur la gestion de leurs informations personnelles [2].

3.4.3 Blockchain pour l'immobilier

Le secteur immobilier implique de nombreuses parties prenantes nécessitant un haut niveau de confiance mutuelle. La blockchain peut introduire un élément de confiance en enregistrant les biens immobiliers avec leur historique dans un registre transparent et immuable. Cela fournit toutes les informations nécessaires sur la propriété, ce qui facilite les transactions et réduit les risques de fraude [2].

3.4.4 Blockchain pour la vérification des documents

La vérification des documents est cruciale dans la lutte contre la contrefaçon. La blockchain offre un environnement partagé sécurisé, où la manipulation ou la falsification des documents devient extrêmement difficile. Ainsi, la vérification des documents devient rapide, efficace et fiable, contribuant à renforcer la confiance dans les processus de vérification et d'authentification [2].

4 Analyse Documentaire

4.1 Cas d'études et exemples concrets d'utilisation de la Blockchain

Plusieurs initiatives interconnectées ont exploré l'intégration de la technologie blockchain dans le domaine de la vérification des certificats et des titres de compétences.

1. Projet BlockCert [5] : Ce projet de recherche, basé sur les fondements établis par le réseau Bitcoin, a été lancé en partenariat entre le Massachusetts Institute of Technology's Media Lab Training Project et Understanding Engine [5]. Il utilise des fonctions cryptographiques sur les données des certificats et des transactions Bitcoin pour stocker de manière sécurisée les valeurs de hachage générées à partir des données des certificats. La légitimité du certificat est vérifiée en comparant les résultats des opérations de hachage avec les valeurs enregistrées dans les transactions Bitcoin.
2. Projet Lifelong Learning Passport (LLP) [6] : Initié par Wolfgang Gräther et son équipe, ce projet se concentre sur l'utilisation de la technologie blockchain dans le domaine de l'éducation. Il utilise le système de fichiers interplanétaires (IPFS) pour le stockage sécurisé et à long terme des informations sur les bénéficiaires liées aux certificats. Pour gérer toutes les données des certificats, ils ont développé un système de gestion des documents centralisé, et pour l'enregistrement et l'authentification des certificats, ils utilisent des contrats intelligents.
3. Projet SPROOF [7] : Ce projet sert de plateforme pour l'émission et la vérification de documents sur une blockchain publique. Il intègre la technologie de portefeuille HD pour renforcer la sécurité des clés, utilise une table de hachage distribuée (DHT) pour stocker les données des certificats [8] et accorde une importance considérable à la protection de la vie privée des utilisateurs à chaque étape du processus de certification.

4.2 Modèles d'Intelligence Artificielle pour l'authentification

Dans le domaine de l'authentification, les modèles d'intelligence artificielle (IA) jouent un rôle crucial dans la détection des faux documents et des images altérées. Plusieurs recherches ont exploré l'utilisation de réseaux neuronaux convolutifs (CNN) pour cette tâche, montrant des résultats prometteurs [9].

Le domaine de la vérification des images numériques propose divers outils et techniques, comme en témoignent plusieurs travaux de recherche. Pour justifier l'authenticité des images, la majorité des approches se concentrent sur le format et les métadonnées des données d'image [10, 11]. Malgré les progrès dans ce domaine, la détection des images falsifiées reste un défi en raison de l'accès aux mêmes techniques de traitement d'image avancées pour les attaquants. Des organisations de réseaux sociaux réputées telles que Facebook et Microsoft ont récemment lancé des programmes de détection des faux pour relever ce défi [12].

Une approche traditionnelle pour la détection des images falsifiées est l'analyse du domaine fréquentiel, qui se concentre sur les données de compression de l'image révélant différentes valeurs pour les images manipulées. Les images falsifiées contiennent généralement des anomalies par rapport aux vraies images dans le domaine fréquentiel, et les expériences avec des détecteurs d'images deepfake ont donné de bons résultats sur le spectre des images [13]. Pourtant, les détecteurs d'images falsifiées basés sur le domaine fréquentiel ont du mal à traiter les bords lisses et sophistiqués des images.

D'autres chercheurs se sont concentrés sur les caractéristiques de texture de l'image pour différencier les images falsifiées des images réelles. Cependant, concevoir un détecteur d'images falsifiées basé sur l'information de texture semble critique, car l'information de texture globale seule ne peut pas différencier les images falsifiées des images réelles. Une approche plus récente combine la matrice de Gram avec Resnet pour capturer l'information de texture [14].

L'apprentissage profond est également une autre approche récente dans la détection des faux visages. Les réseaux neuronaux convolutionnels sont utilisés pour les méthodes de contrefaçon de visage telles que FaceForensics, où la convolution et le pooling maximal sont les deux étapes majeures du processus d'apprentissage supervisé. Dans cette catégorie de méthodes, les réseaux adverses génératifs à auto-attention sont une autre technique qui génère des ensembles de données de faux visages raffinés [15].

Certaines approches dans la littérature, telles que l'outil forensique MMC image, utilisent la qualité JPEG et les métadonnées [16]. Mais le défi réside dans les images fournies par la technique GAN, car même si les métadonnées elles-mêmes sont modifiées ou cachées par l'attaquant, ces outils échouent à identifier les images falsifiées.

En résumé, la vérification des images numériques est un domaine en constante évolution, avec des avancées dans diverses techniques, mais les défis persistent en raison de la sophistication croissante des techniques de contrefaçon et de l'adaptabilité des attaquants.

4.3 Resultats

La technologie blockchain émerge comme un élément crucial dans la vérification de l'authenticité des documents, offrant sécurité et intégrité des données. En exploitant la blockchain, le système peut garantir un processus sûr et transparent de validation des documents, tels que les cartes d'identité, les licences, les passeports, et bien d'autres. En évitant les soumissions répétées de documents, cette technologie optimise l'efficacité et la sécurité du processus.

Parallèlement, l'intégration du traitement d'images avec la blockchain assure la précision et l'efficacité de la vérification des documents. La blockchain sécurise les données à l'aide d'algorithmes robustes comme SHA, tandis que les techniques de traitement d'images, telles que l'analyse du niveau d'erreur et l'utilisation de réseaux neuronaux, renforcent la fiabilité de la validation de l'authenticité des documents [1].

Ces avancées marquent un tournant significatif dans le domaine de la vérification des documents, ouvrant la voie à des processus plus efficaces, sécurisés et précis. L'adoption croissante de ces technologies promet de transformer radicalement les normes actuelles de vérification des documents, avec des implications profondes dans une multitude de secteurs.

4.4 Évaluation des Performances et Limitations

Certains systèmes étudiés offrent des avancées significatives dans divers domaines. Il est remarquable de constater que l'évolution de la technologie blockchain a ouvert de nouvelles perspectives, notamment dans le domaine académique, où les universités exploitent ces systèmes pour sécuriser et rendre infalsifiables les documents académiques. Par exemple, [17] l'Université de Nicosie a été pionnière en 2014 dans l'enregistrement de certificats académiques pour un cours en ligne sur la blockchain Bitcoin. Leur approche, décentralisée, sans permission et transparente, enregistre un hachage d'un document d'index sur la blockchain, contenant une liste de hachages de tous les certificats pour un semestre spécifique. Cependant, cette approche ne permet pas l'intégration d'une vérification émetteur et ne valide pas l'exhaustivité des certificats académiques déposés. De même, l'intégration de la blockchain avec l'intelligence artificielle permet également le développement de systèmes de vérification de documents de plus en plus intelligents et sophistiqués.

Néanmoins, cette évolution technologique présente également des défis. Les fraudeurs, profitant de ces avancées, utilisent des techniques de falsification de plus en plus avancées pour contourner les systèmes de sécurité mis en place. Cette course à l'armement technologique entre les dévelop-

peurs de solutions de vérification et les fraudeurs souligne la nécessité d'une vigilance continue et de l'amélioration constante des technologies de sécurité.

	Caractéristiques				
Éléments	**Décentralisé**	**Sans autorisation**	**Transparent**	**Vérification**	**Exhaustivité**
Université de Nicosie	Oui	Oui	Oui	Non	Non
Blockcerts	Oui	Oui	Oui	Non	Non
LLP	Oui	Non	Oui	Oui	Non
SPROOF	Oui	Oui	Oui	Oui	Oui

Table 1 – Tableau comparatif des méthodes

5 Conclusion

En conclusion, notre étude démontre que l'intégration de la technologie blockchain dans l'authentification des documents présente des avantages significatifs en termes de sécurité et de transparence. Cependant, des défis persistent, notamment en ce qui concerne l'intégration de l'intelligence artificielle et la gestion de l'identité numérique. Pour progresser dans ce domaine, des recherches futures sont nécessaires pour améliorer l'efficacité et l'adaptabilité de ces systèmes. Dans l'ensemble, la technologie blockchain offre un potentiel prometteur pour renforcer la confiance dans les transactions documentaires, mais son adoption réussie nécessitera une approche holistique et collaborative entre les chercheurs, les praticiens et les décideurs.

Références

[1] Shantanu Sarode, Utkarsha Khandare, Shubham Jadhav, Avinash Jannu, Vishnu Kamble, and Digvijay Patil. Document manipulation detection and authenticity verification using machine learning and blockchain. *International Research Journal of Engineering and Technology (IRJET)*, 07 :4758, May 2020. Impact Factor value : 7.529 | ISO 9001 :2008 Certified Journal.

[2] Eliutherius Juma Wanyonyi. Academic document authenticity verification using blockchain technology. 2019.

[3] Vishal Khetavat, Shubhendu Gupta, Pradip Bhor, Vimleshkumar Varma, Sairabanu Pansare, and Tejaswini Zope. Blockchain based document verification system. *Scandinavian Journal of Information Systems*, 5(1) :33–38, 2023.

[4] Issam (Sam) Andoni. Forbes technology council : Blockchain as an application platform, 2018.

[5] Abderahman Rejeb, Karim Rejeb, Horst Treiblmaier, Andrea Appolloni, Salem Alghamdi, Yaser Alhasawi, and Mohammad Iranmanesh. The internet of things (iot) in healthcare : Taking stock and moving forward. *Journal of Ambient Intelligence and Humanized Computing*, 2021.

[6] Pollard TD Ong JJ Elbadawi M McCoubrey LE Goyanes A Gaisford S Basit AW Awad A, Trenfield SJ. Connected healthcare : Improving patient care using digital health technologies. *Advanced Drug Delivery Reviews*, 178 :113958, Mar 1 2021.

[7] Joseph Kvedar, Molly Joel Coye, and William Everett. Leveraging blockchain technology to enhance supply chain management in healthcare : Best care at lower cost. *Blockchain in Healthcare Today*, 1, 2018.

[8] Aziz Sheikh, Harpreet Singh Sood, and David W Bates. Leveraging health information technology to achieve the "triple aim" of healthcare reform. *Journal of the American Medical Informatics Association*, 22(4) :849–856, Jul 2015.

[9] D. Jayaram, M. V. Gopalachari, S. Rakesh, J. S. Sai, and G. K. Kumar. Fake face image detection using feature network. *International Journal of Health Sciences*, 6(S5) :3027–3039, 2022.

[10] Zhengzhe Liu, Xiaojuan Qi, Jiaya Jia, and Philip Torr. Global texture enhancement for fake face detection in the wild. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8057–8066. IEEE, 2020.

[11] Xiaodan Li, Yining Lang, Yuefeng Chen, Xiaofeng Mao, Yuan He, Shuhui Wang, et al. Sharp multiple instance learning for deepfake video detection. In *Proceedings of the 28th ACM International Conference on Multimedia*, October 2020.

[12] Brian Dolhansky, Russ Howes, Ben Pflaum, Nicole Baram, and Cristian Canton Ferrer. The deepfake detection challenge (dfdc) preview dataset. *arXiv preprint arXiv :1910.08854*, 2019.

[13] Luca Guarnera, Oliver Giudice, Cristina Nastasi, and Sebastiano Battiato. Preliminary forensics analysis of deepfake images. *arXiv preprint arXiv :2004.12626*, 2020.

[14] S. Azarian-Pour, M. Babaie-Zadeh, and A. R. Sadri. An automatic jpeg ghost detection approach for digital image forensics. In *2016 24th Iranian Conference on Electrical Engineering (ICEE)*, pages 1645–1649. IEEE, 2016.

[15] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision (IJCV)*, 115 :211–252, 2015.

[16] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, CA, USA, 2019. PMLR.

[17] Clemens Brunner, Fabian Knirsch, and Dominik Engel. Sproof : A platform for issuing and verifying documents in a public blockchain. In *Proceedings of the 5th International Conference on Information Systems Security and Privacy (ICISSP 2019)*, Center for Secure Energy Informatics, Salzburg University of Applied Sciences, Puch bei Hallein, Austria, 2019. SCITEPRESS.