



HAL
open science

Avancées Algorithmiques et Accélération Matérielle: Catalyseurs de la Révolution de l'Intelligence Artificielle

Moez Krichen

► **To cite this version:**

Moez Krichen. Avancées Algorithmiques et Accélération Matérielle: Catalyseurs de la Révolution de l'Intelligence Artificielle. 2024. hal-04446522

HAL Id: hal-04446522

<https://hal.science/hal-04446522>

Preprint submitted on 8 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Avancées Algorithmiques et Accélération Matérielle : Catalyseurs de la Révolution de l'Intelligence Artificielle

Moez Krichen

Laboratoire ReDCAD, Université de Sfax, Tunisie
moez.krichen@redcad.org

Résumé. Cet article explore les progrès significatifs réalisés dans les domaines des algorithmes d'intelligence artificielle (IA) et de l'accélération matérielle, soulignant leur rôle crucial dans la transformation des capacités de calcul et l'efficacité des applications IA. Nous examinons les développements récents dans l'apprentissage profond, l'apprentissage par renforcement, les modèles génératifs, et les algorithmes évolutifs, ainsi que l'impact de l'accélération matérielle à travers les unités de traitement graphique (GPUs), les unités de traitement tensoriel (TPUs), les arrays à portes programmables (FPGAs), les circuits intégrés spécifiques à une application (ASICs), et l'informatique quantique. En intégrant des approches interdisciplinaires et en exploitant des technologies d'accélération matérielle, cette étude démontre comment ces avancées soutiennent une amélioration significative des performances et de l'efficacité dans une variété de domaines d'application de l'IA.

1 Introduction

Le développement de l'apprentissage automatique (ML) et de l'intelligence artificielle (AI) a entraîné des changements significatifs dans plusieurs secteurs, menant à l'automatisation des emplois, à l'extraction de perspectives à partir de vastes ensembles de données et à la facilitation de processus de prise de décision avancés. L'utilisation des systèmes d'AI est de plus en plus répandue dans de nombreux domaines, englobant l'identification d'images, le traitement du langage naturel, les voitures autonomes et les recommandations personnalisées (72; 25). Néanmoins, la demande pour améliorer l'efficacité de ces systèmes d'AI a émergé comme un champ d'étude crucial, étant donné leur sophistication croissante et leurs besoins en ressources (50; 64).

L'avènement de l'intelligence artificielle (IA) représente l'une des transformations technologiques les plus significatives de notre époque, offrant des perspectives révolutionnaires dans presque tous les secteurs de la société. Ces progrès remarquables sont principalement attribuables à deux facteurs clés : les avancées dans les algorithmes d'IA et les progrès dans l'accélération matérielle (40; 67; 5; 66; 12; 46; 65; 2; 37; 44; 6). Les algorithmes d'IA, en particulier, ont connu des développements rapides, permettant des améliorations substantielles en termes d'efficacité, de précision, et de capacité à

traiter et interpréter des volumes massifs de données (58; 39; 47; 48; 35; 74; 7; 3; 11; 53; 20; 8; 45; 54). Simultanément, l'accélération matérielle a permis de surmonter les limitations des architectures informatiques traditionnelles, ouvrant la voie à des calculs plus rapides et à une efficacité énergétique accrue (61; 41; 49; 60; 4; 16; 15; 10; 27). Cette synergie entre les avancées algorithmiques et matérielles a été cruciale pour réaliser les promesses de l'IA, depuis la reconnaissance vocale et la vision par ordinateur jusqu'à la prise de décision autonome et au-delà (59; 42; 43; 38; 36).

Cette étude vise à explorer en profondeur ces développements, en mettant en évidence non seulement les progrès individuels dans les domaines des algorithmes et de l'accélération matérielle mais aussi la manière dont leur intégration favorise une nouvelle ère d'innovation et d'application de l'IA. Nous débutons par une revue des avancées algorithmiques clés, notamment dans l'apprentissage profond, l'apprentissage par renforcement, les modèles génératifs et les algorithmes évolutifs. Chacun de ces domaines a contribué à élargir les capacités de l'IA, permettant des applications plus sophistiquées et performantes.

Parallèlement, nous examinons les progrès dans l'accélération matérielle, y compris les développements des unités de traitement graphique (GPUs), des unités de traitement tensoriel (TPUs), des arrays à portes programmables (FPGAs), des circuits intégrés spécifiques à une application (ASICs) et de l'informatique quantique. Ces technologies ont joué un rôle crucial dans l'amélioration de la vitesse de calcul et de l'efficacité énergétique, permettant ainsi l'entraînement de modèles d'IA plus complexes et leur déploiement à une échelle sans précédent.

En intégrant des perspectives interdisciplinaires et en tirant parti des dernières innovations en matière de technologie matérielle, cette étude offre une vue d'ensemble de l'état actuel de l'IA et envisage son avenir. Nous discutons de l'impact de ces avancées sur la recherche et les applications industrielles, ainsi que des défis et des opportunités qui se profilent à l'horizon. Notre objectif est de fournir aux chercheurs, aux praticiens, et aux décideurs une compréhension approfondie des moteurs de la révolution de l'IA, soulignant comment ils peuvent être exploités pour adresser des problèmes complexes et ouvrir de nouvelles voies pour l'innovation et le progrès.

2 Avancées Algorithmiques

Les avancées dans les algorithmes ont entraîné une transformation significative dans un large éventail d'applications de l'IA, conduisant à des niveaux améliorés d'efficacité, de précision, de sécurité et de robustesse. Cette étude examine les développements algorithmiques significatifs dans plusieurs secteurs de l'IA, soulignant leur influence et leurs conséquences potentielles (34; 26; 69; 70).

2.1 Apprentissage Profond

Le domaine de l'IA a vu l'émergence de l'apprentissage profond (Deep Learning, DL) comme un paradigme clé, permettant des progrès notables dans la vision par ordinateur, le traitement du langage naturel et la reconnaissance vocale. Les modèles de DL, tels que les réseaux neuronaux convolutionnels (CNNs) et les réseaux neuronaux récur-

rents (RNNs), ont considérablement transformé les tâches de reconnaissance d'image et de parole grâce à leur capacité à acquérir de manière autonome des représentations hiérarchiques à partir des données. La structure hiérarchique inhérente aux modèles de DL leur permet de capturer efficacement des motifs complexes et des interdépendances, résultant en une performance exceptionnelle dans divers domaines. L'utilisation du DL a facilité des avancées significatives dans le traitement du langage naturel, permettant la réalisation de nombreuses tâches, y compris la traduction automatique, l'analyse des sentiments et la réponse aux questions. Le succès des algorithmes de DL peut être attribué à la présence de vastes ensembles de données étiquetées et aux capacités de traitement offertes par la technologie contemporaine.

2.2 Apprentissage par Renforcement

Des progrès significatifs ont été réalisés dans le développement des algorithmes d'apprentissage par renforcement (RL), qui permettent de former des agents à prendre des décisions optimales dans des environnements dynamiques. Les algorithmes de RL apprennent par interaction avec un environnement, recevant des retours sous forme de récompenses ou de pénalités pour leurs actions. Les algorithmes de RL notables incluent Q-learning, les gradients de politique et les approches acteur-critique. Ces algorithmes ont été appliqués avec succès à divers problèmes, y compris les jeux, la robotique et les systèmes autonomes. Le RL a démontré des réalisations remarquables, telles qu'AlphaGo, qui a battu les champions du monde au jeu de Go, et plus récemment, les modèles DALL-E et CLIP d'OpenAI, qui montrent le potentiel du RL dans la génération de contenu créatif et l'apprentissage multimodal.

2.3 Modèles Génératifs

L'objectif principal des modèles génératifs est de développer un modèle qui capture précisément la distribution de probabilité sous-jacente d'un ensemble de données donné, permettant la génération de nouveaux exemples inédits. L'attention accordée aux modèles génératifs ces dernières années est attribuée à leur capacité à produire des sorties réalistes et de haute qualité. Notamment, les réseaux antagonistes génératifs (GANs) ont émergé comme une catégorie puissante de modèles qui acquièrent la capacité de générer de nouveaux échantillons en formant un réseau générateur en opposition à un réseau discriminatoire, favorisant une dynamique compétitive. Les GANs ont démontré des réalisations notables dans diverses disciplines, englobant la synthèse d'images, le transfert de style et l'augmentation de données. Les autoencodeurs variationnels (VAEs) constituent une catégorie remarquable de modèles génératifs qui peuvent acquérir une représentation latente des données et ensuite créer de nouveaux échantillons en échantillonnant à partir de l'espace latent acquis. Les modèles génératifs possèdent un potentiel significatif dans divers domaines, y compris la génération de contenu, la synthèse de données et la simulation.

2.4 Algorithmes Évolutifs

Les algorithmes évolutifs sont basés sur les principes fondamentaux de l'évolution biologique et sont utilisés pour aborder des problèmes d'optimisation complexes (22; 51). Ces algorithmes emploient une population de solutions potentielles et les raffinent progressivement à travers des opérations de sélection, de croisement et de mutation. Les algorithmes génétiques, la programmation génétique et les techniques évolutives sont reconnus comme des instances notables d'algorithmes évolutifs. Ces techniques ont été efficacement utilisées dans plusieurs problèmes, tels que l'optimisation, la robotique et la conception. Les algorithmes évolutifs possèdent des avantages inhérents dans des situations caractérisées par la complexité, la non-linéarité, ou une compréhension limitée du domaine du problème. Ils ont la capacité d'explorer une gamme diversifiée d'alternatives et possèdent la capacité de découvrir des réponses innovantes et non traditionnelles qui peuvent échapper aux méthodes d'optimisation conventionnelles.

2.5 Approches Interdisciplinaires

La recherche en IA a progressivement adopté des méthodologies interdisciplinaires, intégrant des concepts et des méthodologies de diverses autres disciplines. Cette approche permet aux chercheurs de tirer parti des insights et des techniques d'autres domaines pour améliorer les algorithmes d'IA. Par exemple, les réseaux neuronaux graphiques sont apparus comme un outil puissant en incorporant des principes de la théorie des graphes, permettant une analyse et une prédiction efficaces dans les données structurées en graphes. Les algorithmes de jeu adversarial combinent des concepts de la théorie des jeux et de l'IA pour créer des agents capables de concurrencer des experts humains dans des jeux complexes. Les architectures cognitives s'inspirent de la science cognitive pour simuler des processus de pensée de type humain et permettre un comportement intelligent. De plus, l'apprentissage automatique informé par la physique combine des principes de la physique et de l'apprentissage automatique pour améliorer la modélisation et la prédiction des systèmes physiques. Les approches interdisciplinaires offrent des perspectives nouvelles et contribuent au développement d'algorithmes d'IA plus puissants et polyvalents capables de s'attaquer à des problèmes réels de plus en plus complexes.

Diverses technologies d'accélération matérielle, y compris les GPU, les TPU, les FPGA, les circuits intégrés spécifiques à une application (ASIC) et l'informatique quantique, ont été développées pour surmonter les contraintes des architectures traditionnelles et augmenter les capacités computationnelles et l'efficacité. Les méthodologies hybrides utilisent diverses techniques d'accélération matérielle pour améliorer à la fois la performance et l'efficacité. Les progrès dans l'accélération matérielle ont amélioré la formation, l'inférence et l'efficacité globale des modèles d'IA dans divers domaines.

3 Accélération Matérielle

Cette section explore l'importance de l'accélération matérielle dans l'accélération de l'avancement et du déploiement des modèles d'IA. La complexité croissante des

algorithmes d'IA et leur dépendance aux données peuvent limiter les architectures informatiques conventionnelles, conduisant à des contraintes de puissance de traitement et d'efficacité. Les approches d'accélération matérielle sont employées pour relever ces défis en tirant parti de composants et d'architectures matériels spécialisés expressément conçus pour les tâches d'IA (17; 21; 9; 31; 55).

3.1 Unités de Traitement Graphique (GPUs)

L'incorporation de GPUs a été cruciale pour accélérer les calculs d'IA. Initialement conçus pour la génération visuelle, les GPUs ont démontré une compétence exceptionnelle dans le traitement parallèle, les rendant bien adaptés pour effectuer des tâches d'entraînement et d'inférence en apprentissage profond (30; 68). Les GPUs démontrent une compétence exceptionnelle à exécuter plusieurs calculs simultanément, capitalisant sur leur architecture hautement parallèle pour améliorer la vitesse des opérations matricielles et des calculs de réseaux neuronaux. Le parallélisme inhérent des GPUs facilite le traitement efficace de vastes ensembles de données et l'entraînement de modèles d'apprentissage profond complexes. Actuellement, les GPUs ont gagné une traction significative tant dans la recherche académique que dans les applications industrielles. Cette adoption généralisée a facilité l'entraînement et le déploiement de modèles d'IA à grande échelle par les chercheurs et les praticiens.

3.2 Unités de Traitement Tensoriel (TPUs)

Les TPUs sont un type d'accélérateur matériel spécialisé spécifiquement développé par Google. Ces accélérateurs sont destinés avec l'objectif principal d'améliorer la performance globale des activités de ML (32; 23). Les TPUs sont des dispositifs de calcul conçus pour effectuer efficacement des opérations tensorielles, qui sont fondamentales pour les processus de calcul impliqués dans les calculs d'apprentissage profond. Les TPUs sont capables d'accélérer considérablement les charges de travail d'IA. Les TPUs présentent une aptitude exceptionnelle pour des opérations d'entraînement et d'inférence étendues à grande échelle, grâce à leurs remarquables capacités de calcul et leur économie d'énergie. En utilisant des TPUs, les chercheurs et les entreprises peuvent atteindre des durées d'entraînement accélérées, réduire les dépenses et améliorer l'efficacité des modèles dans leurs efforts d'IA.

3.3 Field-Programmable Gate Arrays (FPGAs)

Les FPGAs offrent une approche flexible pour augmenter les capacités de calcul des applications d'IA via l'accélération matérielle (13; 33; 18). Contrairement aux GPUs et TPUs, les FPGAs possèdent l'attribut unique de reprogrammabilité, permettant l'implémentation d'architectures matérielles sur mesure optimisées pour certains workloads d'IA. Les FPGAs démontrent une performance exceptionnelle dans les situations nécessitant un faible délai temporel, et un traitement immédiat des données et ont été employés avec succès dans des domaines incluant la vision par ordinateur, le calcul en périphérie, et le streaming de données.

3.4 Circuits Intégrés Spécifiques à une Application (ASICs)

Les ASICs se réfèrent à des puces matérielles spécialisées méticuleusement conçues pour exécuter certaines tâches efficacement (19; 56). Les ASICs sont des circuits électroniques sur mesure spécifiquement adaptés pour optimiser la performance de calcul et l'efficacité énergétique des calculs d'apprentissage profond dans le domaine de l'IA (24; 28). Les ASICs ont la capacité d'atteindre des améliorations notables en vitesse et en économie d'énergie par rapport aux unités de traitement central (CPUs) conventionnelles et autres accélérateurs matériels en éliminant le besoin de calcul généraliste. Les ASICs sont importants dans les situations nécessitant une inférence d'IA efficace, en particulier dans des domaines tels que les véhicules autonomes, la robotique, et les centres de données.

3.5 Informatique Quantique

L'utilisation des phénomènes quantiques dans l'informatique quantique a la capacité de transformer significativement le domaine de l'IA en permettant l'exécution de calculs à une échelle sans précédent (57; 62; 73; 63; 71). Les ordinateurs quantiques emploient des bits quantiques, communément appelés qubits, qui possèdent la capacité de représenter et de manipuler des informations dans plusieurs états simultanément. Le parallélisme intrinsèque observé dans l'informatique quantique peut améliorer significativement l'efficacité de calcul de certaines activités d'IA, notamment l'optimisation, l'apprentissage automatique, et la simulation, en tirant parti d'une accélération exponentielle. L'utilisation d'ordinateurs quantiques est actuellement explorée pour tirer parti des méthodes d'apprentissage automatique quantique, qui ont le potentiel de révéler des motifs et des corrélations précédemment inconnus au sein des ensembles de données. Néanmoins, le domaine de l'informatique quantique est actuellement dans sa phase naissante, nécessitant d'importants efforts de recherche et de développement pour surmonter les obstacles technologiques et construire des systèmes quantiques fonctionnels et extensibles adaptés aux applications d'IA.

3.6 Approches Hybrides

Les méthodologies hybrides utilisent diverses techniques d'accélération matérielle pour optimiser à la fois la performance et l'efficacité. Les systèmes hybrides peuvent atteindre des capacités de calcul améliorées et optimiser l'allocation des ressources en exploitant efficacement les avantages offerts par diverses architectures matérielles (14; 52; 1; 29). Un exemple illustratif d'une stratégie hybride pourrait impliquer l'utilisation de GPUs pour exécuter des tâches de traitement parallèle, de TPUs pour effectuer des opérations tensorielles, et de FPGAs pour faciliter l'accélération matérielle sur mesure. Les combinaisons mentionnées peuvent être personnalisées pour répondre aux critères spécifiques des algorithmes d'IA, permettant une exécution rationalisée et une performance globale améliorée. Les techniques hybrides offrent également un degré d'adaptabilité pour accommoder les workloads d'IA dynamiques et tirer parti des dernières avancées en technologie d'accélération matérielle. La progression de la technologie d'IA est anticipée pour être accompagnée par l'utilisation prédominante

d'architectures hybrides, qui contribueront grandement à l'accélération de la recherche, du développement et de la mise en œuvre de l'IA.

4 Conclusion

Les avancées dans les algorithmes d'IA et l'accélération matérielle ont conjointement joué un rôle pivot dans l'élargissement des frontières de l'IA, permettant des progrès sans précédent dans divers domaines d'application. L'apprentissage profond, l'apprentissage par renforcement, les modèles génératifs, et les algorithmes évolutifs ont tous bénéficié de l'adoption de technologies d'accélération matérielle avancées, telles que les GPUs, TPUs, FPGAs, ASICs, et même l'informatique quantique. Ces développements offrent une plateforme pour des innovations continues, promettant d'accélérer la recherche et le déploiement de solutions d'IA plus avancées. Cependant, pour réaliser pleinement le potentiel de ces avancées, il est crucial de continuer à explorer des approches interdisciplinaires et à développer des architectures matérielles innovantes qui peuvent soutenir la complexité croissante des modèles d'IA. En conclusion, l'interaction entre les avancées algorithmiques et l'accélération matérielle reste au cœur de la révolution de l'IA, façonnant l'avenir de la technologie et son application à travers une multitude de domaines.

Références

- [1] Mohamed S Abdalzaher, M Sami Soliman, and Sherif M El-Hady. Seismic intensity estimation for earthquake early warning using optimized machine learning model. *IEEE Transactions on Geoscience and Remote Sensing*, 2023.
- [2] Q Abu Al-Haija and M Krichen. A lightweight in-vehicle alcohol detection using smart sensing and supervised learning. *computers* 2022, 11, 121, 2022.
- [3] Qasem Abu Al-Haija, Moez Krichen, and Wejdan Abu Elhaija. Machine-learning-based darknet traffic detection system for iot applications. *Electronics*, 11(4) :556, 2022.
- [4] Qasem Abu Al-Haija and Moez Krichen. Analyzing malware from api call sequences using support vector machines. In *International Conference on Cybersecurity, Cybercrimes, and Smart Emerging Technologies*, pages 27–39. Springer International Publishing Cham, 2022.
- [5] Omar Azib Alkhudaydi, Moez Krichen, and Ans D Alghamdi. A deep learning methodology for predicting cybersecurity attacks on the internet of things. *Information*, 14(10) :550, 2023.
- [6] Hamoud Alshammari, Karim Gasmi, Ibtihel Ben Ltaifa, Moez Krichen, Lassaad Ben Ammar, and Mahmood A Mahmood. Olive disease classification based on vision transformer and cnn models. *Computational Intelligence and Neuroscience*, 2022, 2022.
- [7] Hamoud Alshammari, Karim Gasmi, Moez Krichen, Lassaad Ben Ammar, Mohamed Osman Abdelhadi, Ammar Boukrara, and Mahmood A Mahmood. Optimal

- deep learning model for olive disease diagnosis based on an adaptive genetic algorithm. *Wireless Communications and Mobile Computing*, 2022 :1–13, 2022.
- [8] Hashem Alyami, Wael Alosaimi, Moez Krichen, and Roobaea Alroobaea. Monitoring social distancing using artificial intelligence for fighting covid-19 virus spread. *International Journal of Open Source Software and Processes (IJOSSP)*, 12(3) :48–63, 2021.
- [9] Adam Auten, Matthew Tomei, and Rakesh Kumar. Hardware acceleration of graph neural networks. In *2020 57th ACM/IEEE Design Automation Conference (DAC)*, pages 1–6. IEEE, 2020.
- [10] Rubby Aworka, Lontsi Saadio Cedric, Wilfried Yves Hamilton Adoni, Jérémie Thouakessah Zoueu, Franck Kalala Mutombo, Charles Lebon Mberi Kimpolo, Tarik Nahhal, and Moez Krichen. Agricultural decision system based on advanced machine learning models for yield prediction : Case of east african countries. *Smart Agricultural Technology*, 2 :100048, 2022.
- [11] Wadii Boulila, Maha Driss, Eman Alshantiti, Mohamed Al-Sarem, Faisal Saeed, and Moez Krichen. Weight initialization techniques for deep learning algorithms in remote sensing : Recent trends and future perspectives. *Advances on Smart and Soft Computing : Proceedings of ICACIn 2021*, pages 477–484, 2022.
- [12] Zakaria Boulouard, Mariyam Ouaisa, Mariya Ouaisa, Farhan Siddiqui, Mutiq Almutiq, and Moez Krichen. An integrated artificial intelligence of things environment for river flood prevention. *Sensors*, 22(23) :9485, 2022.
- [13] Sérgio Branco, André G Ferreira, and Jorge Cabral. Machine learning in resource-scarce embedded systems, fpgas, and end-devices : A survey. *Electronics*, 8(11) :1289, 2019.
- [14] Maurizio Capra, Beatrice Bussolino, Alberto Marchisio, Muhammad Shafique, Guido Maserà, and Maurizio Martina. An updated survey of efficient hardware architectures for accelerating deep convolutional neural networks. *Future Internet*, 12(7) :113, 2020.
- [15] Lontsi Saadio Cedric, Wilfried Yves Hamilton Adoni, Rubby Aworka, Jérémie Thouakessah Zoueu, Franck Kalala Mutombo, Moez Krichen, and Charles Lebon Mberi Kimpolo. Crops yield prediction based on machine learning models : Case of west african countries. *Smart Agricultural Technology*, 2 :100049, 2022.
- [16] Oumaima Chakir, Abdeslam Rehami, Yassine Sadqi, Moez Krichen, Gurjot Singh Gaba, Andrei Gurtov, et al. An empirical assessment of ensemble methods and traditional machine learning techniques for web-based attack detection in industry 5.0. *Journal of King Saud University-Computer and Information Sciences*, 35(3) :103–119, 2023.
- [17] Lei Deng, Guoqi Li, Song Han, Luping Shi, and Yuan Xie. Model compression and hardware acceleration for neural networks : A comprehensive survey. *Proceedings of the IEEE*, 108(4) :485–532, 2020.
- [18] Javier Duarte, Philip Harris, Scott Hauck, Burt Holzman, Shih-Chieh Hsu, Sergio Jindariani, Suffian Khan, Benjamin Kreis, Brian Lee, Mia Liu, et al. Fpga-accelerated machine learning inference as a service for particle physics computing.

- Computing and Software for Big Science*, 3 :1–15, 2019.
- [19] Norman Einspruch. *Application specific integrated circuit (ASIC) technology*, volume 23. Academic Press, 2012.
- [20] Mourad Ellouze, Seifeddine Mechti, Moez Krichen, Vinayakumar Ravi, and Lamia Hadrich Belguith. A deep learning approach for detecting the behaviour of people having personality disorders towards covid-19 from twitter. *International Journal of Computational Science and Engineering*, 25(4) :353–366, 2022.
- [21] Deepak Ghimire, Dayoung Kil, and Seong-heum Kim. A survey on efficient convolutional neural networks and hardware acceleration. *Electronics*, 11(6) :945, 2022.
- [22] Sacha Gobeyn, Ans M Mouton, Anna F Cord, Andrea Kaim, Martin Volk, and Peter LM Goethals. Evolutionary algorithms for species distribution modelling : A review in the context of machine learning. *Ecological Modelling*, 392 :179–195, 2019.
- [23] Ashish Gondimalla, Noah Chesnut, Mithuna Thottethodi, and TN Vijaykumar. Sparten : A sparse tensor accelerator for convolutional neural networks. In *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, pages 151–165, 2019.
- [24] Neha Gupta. Introduction to hardware accelerator systems for artificial intelligence and machine learning. In *Advances in Computers*, volume 122, pages 1–21. Elsevier, 2021.
- [25] Michael Haenlein and Andreas Kaplan. A brief history of artificial intelligence : On the past, present, and future of artificial intelligence. *California management review*, 61(4) :5–14, 2019.
- [26] Gabriel Paes Herrera, Michel Constantino, Jen-Je Su, and Athula Naranpanawa. The use of icts and income distribution in brazil : A machine learning explanation using shap values. *Telecommunications Policy*, 47(8) :102598, 2023.
- [27] Olfa Hrizi, Karim Gasmi, Ibtihel Ben Ltaifa, Hamoud Alshammari, Hanen Karanti, Moez Krichen, Lassaad Ben Ammar, and Mahmood A Mahmood. Tuberculosis disease diagnosis based on an optimized machine learning model. *Journal of Healthcare Engineering*, 2022, 2022.
- [28] Aabha Jain and Neha Sharma. Accelerated ai inference at cnn-based machine vision in asics : A design approach. *ECS Transactions*, 107(1) :5165, 2022.
- [29] Conrad D James, James B Aimone, Nadine E Miner, Craig M Vineyard, Fredrick H Rothganger, Kristofor D Carlson, Samuel A Mulder, Timothy J Draelos, Aleksandra Faust, Matthew J Marinella, et al. A historical survey of algorithms and hardware architectures for neural-inspired and neuromorphic computing applications. *Biologically Inspired Cognitive Architectures*, 19 :49–64, 2017.
- [30] Abhinav Jangda, Sandeep Polisetty, Arjun Guha, and Marco Serafini. Accelerating graph sampling for graph machine learning using gpus. In *Proceedings of the Sixteenth European Conference on Computer Systems*, pages 311–326, 2021.
- [31] Pooja Jawandhiya. Hardware design for machine learning. *Int. J. Artif. Intell. Appl*, 9(1) :63–84, 2018.

- [32] Norman Jouppi, Cliff Young, Nishant Patil, and David Patterson. Motivation for and evaluation of the first tensor processing unit. *IEEE Micro*, 38(3) :10–19, 2018.
- [33] Kaan Kara, Dan Alistarh, Gustavo Alonso, Onur Mutlu, and Ce Zhang. Fpga-accelerated dense linear machine learning : A precision-convergence trade-off. In *2017 IEEE 25th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pages 160–167. IEEE, 2017.
- [34] Athanasios Karapantelakis, Pegah Alizadeh, Abdulrahman Alabassi, Kaushik Dey, and Alexandros Nikou. Generative ai in mobile networks : a survey. *Annals of Telecommunications*, pages 1–19, 2023.
- [35] Hajra Khan, Imran Fareed Nizami, Saeed Mian Qaisar, Asad Waqar, Moez Krichen, and Abdulaziz Turki Almaktoom. Analyzing optimal battery sizing in microgrids based on the feature selection and machine learning approaches. *Energies*, 15(21) :7865, 2022.
- [36] Moez Krichen. How artificial intelligence can revolutionize software testing techniques. In *International Conference on Innovations in Bio-Inspired Computing and Applications*, pages 189–198. Springer Nature Switzerland Cham, 2022.
- [37] Moez Krichen. Les méthodes formelles sont-elles applicables à l'apprentissage automatique et à l'intelligence artificielle. 2022.
- [38] Moez Krichen. Comment l'intelligence artificielle peut révolutionner les techniques de test de logiciels. 2023.
- [39] Moez Krichen. Convolutional neural networks : A survey. *Computers*, 12(8) :151, 2023.
- [40] Moez Krichen. Deep reinforcement learning. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–7. IEEE, 2023.
- [41] Moez Krichen. Generative adversarial networks. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–7. IEEE, 2023.
- [42] Moez Krichen. Renforcer la sécurité des contrats intelligents grâce à la puissance de l'intelligence artificielle. 2023.
- [43] Moez Krichen. Strengthening the security of smart contracts through the power of artificial intelligence. *Computers*, 12(5) :107, 2023.
- [44] Moez Krichen, Alaeddine Mihoub, Mohammed Y Alzahrani, Wilfried Yves Hamilton Adoni, and Tarik Nahhal. Are formal methods applicable to machine learning and artificial intelligence ? In *2022 2nd International Conference of Smart Systems and Emerging Technologies (SMARTTECH)*, pages 48–53. IEEE, 2022.
- [45] Seifeddine Mechti, Moez Krichen, Dhouha Ben Noureddine, and Lamia H Belguith. A decision system for computational authors profiling : From machine learning to deep learning. *Concurrency and Computation : Practice and Experience*, 34(7) :e5985, 2022.
- [46] Saeed Mian Qaisar, Nehal Alyamani, Asad Waqar, and Moez Krichen. Machine learning with adaptive rate processing for power quality disturbances identifica-

- tion. *SN Computer Science*, 3 :1–6, 2022.
- [47] Saeed Mian Qaisar, Dalila Say, Salah Zidi, and Krichen Moez. Automated categorization of multiclass welding defects using the x-ray image augmentation and convolutional neural network. 2023.
 - [48] Alaeddine Mihoub, Moez Krichen, Mohammad Alswailim, Sami Mahfoudhi, and Riadh Bel Hadj Salah. Road scanner : A road state scanning approach based on machine learning techniques. *Applied Sciences*, 13(2) :683, 2023.
 - [49] Alaeddine Mihoub, Hosni Snoun, Moez Krichen, Riadh Bel Hadj Salah, and Montassar Kahia. Predicting covid-19 spread level using socio-economic indicators and machine learning techniques. In *2020 first international conference of smart systems and emerging technologies (SMARTTECH)*, pages 128–133. IEEE, 2020.
 - [50] Madison Milne-Ives, Caroline de Cock, Ernest Lim, Melissa Harper Shehadeh, Nick de Pennington, Guy Mole, Eduardo Normando, and Edward Meinert. The effectiveness of artificial intelligence conversational agents in health care : systematic review. *Journal of medical Internet research*, 22(10) :e20346, 2020.
 - [51] Seyedali Mirjalili. Evolutionary algorithms and neural networks. In *Studies in computational intelligence*, volume 780. Springer, 2019.
 - [52] Sayed SR Moustafa, Mohamed S Abdalzaher, Farhan Khan, Mohamed Metwaly, Eslam A Elawadi, and Nassir S Al-Arifi. A quantitative site-specific classification approach based on affinity propagation clustering. *IEEE Access*, 9 :155297–155313, 2021.
 - [53] Pierre Stanislas Birame Ndong, Wilfried Yves Hamilton Adoni, Tarik Nahhal, Charles Kimpolo, Moez Krichen, Abdeltif EL Byed, Ismail Assayad, and Franck Kalala Mutombo. A face-mask detection system based on deep learning convolutional neural networks. In *Advances on Smart and Soft Computing : Proceedings of ICACIn 2021*, pages 273–283. Springer Singapore Singapore, 2021.
 - [54] Dhouha Ben Noureddine, Moez Krichen, Seifeddine Mechti, Tarik Nahhal, and Wilfried Yves Hamilton Adoni. An agent-based architecture using deep reinforcement learning for the intelligent internet of things applications. In *Advances on Smart and Soft Computing : Proceedings of ICACIn 2020*, pages 273–283. Springer Singapore, 2021.
 - [55] Jongse Park, Hardik Sharma, Divya Mahajan, Joon Kyung Kim, Preston Olds, and Hadi Esmaeilzadeh. Scale-out acceleration for machine learning. In *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture*, pages 367–381, 2017.
 - [56] Loukas Petrou, Kypros M Kossifos, Marco A Antoniadis, and Julius Georgiou. The first family of application-specific integrated circuits for programmable and reconfigurable metasurfaces. *Scientific reports*, 12(1) :5826, 2022.
 - [57] Frank Phillipson. Quantum machine learning : Benefits and practical examples. In *QANSWER*, pages 51–56, 2020.
 - [58] Saeed Mian Qaisar, Alaeddine Mihoub, Moez Krichen, and Humaira Nisar. Multi-rate processing with selective subbands and machine learning for efficient arrhyth-

- mia classification. *Sensors*, 21(4) :1511, 2021.
- [59] Shalli Rani, Ali Kashif Bashir, Moez Krichen, Abdulaziz Alshammari, et al. A low-rank learning based multi-label security solution for industry 5.0 consumers using machine learning classifiers. *IEEE Transactions on Consumer Electronics*, 2023.
- [60] Shashidhar Rudregowda, Sudarshan Patil Kulkarni, Gururaj HL, Vinayakumar Ravi, and Moez Krichen. Visual speech recognition for kannada language using vgg16 convolutional neural network. In *Acoustics*, volume 5, pages 343–353. MDPI, 2023.
- [61] Dalila Say, Salah Zidi, Saeed Mian Qaisar, and Moez Krichen. Automated categorization of multiclass welding defects using the x-ray image augmentation and convolutional neural network. *Sensors*, 23(14) :6422, 2023.
- [62] Maria Schuld and Francesco Petruccione. *Machine learning with quantum computers*. Springer, 2021.
- [63] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2) :172–185, 2015.
- [64] Terrence J Sejnowski. The unreasonable effectiveness of deep learning in artificial intelligence. *Proceedings of the National Academy of Sciences*, 117(48) :30033–30038, 2020.
- [65] Souhir Sghaier, Moez Krichen, Abir Othman Elfaki, and Qasem Abu Al-Haija. Efficient machine-learning based 3d face identification system under large pose variation. In *International Conference on Computational Collective Intelligence*, pages 273–285. Springer International Publishing Cham, 2022.
- [66] R Shashidhar, S Patilkulkarni, Vinayakumar Ravi, HL Gururaj, and Moez Krichen. Audiovisual speech recognition based on a deep convolutional neural network. *Data Science and Management*, 2023.
- [67] Wiem Souai, Alaeddine Mihoub, Mounira Tarhouni, Salah Zidi, Moez Krichen, and Sami Mahfoudhi. Predicting at-risk students using the deep learning blstm approach. In *2022 2nd International Conference of Smart Systems and Emerging Technologies (SMARTTECH)*, pages 32–37. IEEE, 2022.
- [68] Sijun Tan, Brian Knott, Yuan Tian, and David J Wu. Cryptgpu : Fast privacy-preserving machine learning on the gpu. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 1021–1038. IEEE, 2021.
- [69] João Vitorino, Isabel Praça, and Eva Maia. Towards adversarial realism and robust learning for iot intrusion detection and classification. *Annals of Telecommunications*, pages 1–12, 2023.
- [70] Ahmad Samer Wazan and Frédéric Cuppens. Cybersecurity in networking : adaptations, investigation, attacks, and countermeasures. *Annals of Telecommunications*, 78(3-4) :133–134, 2023.
- [71] Peter Wittek. *Quantum machine learning : what quantum computing means to data mining*. Academic Press, 2014.
- [72] Caiming Zhang and Yang Lu. Study on artificial intelligence : The state of the art

and future prospects. *Journal of Industrial Information Integration*, 23 :100224, 2021.

- [73] Yao Zhang and Qiang Ni. Recent advances in quantum machine learning. *Quantum Engineering*, 2(1) :e34, 2020.
- [74] Salah Zidi, Alaeddine Mihoub, Saeed Mian Qaisar, Moez Krichen, and Qasem Abu Al-Haija. Theft detection dataset for benchmarking and machine learning based classification in a smart grid environment. *Journal of King Saud University-Computer and Information Sciences*, 35(1) :13–25, 2023.