ISSN: 2150-8402 Volume 15, Number 1, February 2023

Scientific
Research
Publishing

Journal of Intelligent Learning Systems and Applications

ISSN: 2150-8402
9 772150 840003 01

https://www.scirp.org/journal/jilsa

Journal Editorial Board

ISSN Print: 2150-8402   ISSN Online: 2150-8410
https://www.scirp.org/journal/jilsa

Editor-in-Chief

Prof. Wei-Chiang Hong	Asia Eastern University of Science and Technology, Chinese Taipei

Editorial Advisory Board

Prof. Rob Law	Hong Kong Polytechnic University, China
Dr. Steve S. H. Ling	University of Technology Sydney, Australia
Prof. Yi Pan	Georgia State University, USA
Prof. Xin Xu	National University of Defense Technology, China

Editorial Board

Prof. Ala Al-Fuqaha	Western Michigan University, USA
Prof. George Angelos Anastassiou	Department of Mathematical Sciences, The University of Memphis, USA
Prof. Tao Ban	National Institute of Information and Communications Technology, Japan
Prof. Mokhtar Beldjehem	University of Ottawa, Canada
Dr. Rifai Chai	University of Technology, Sydney, Australia
Dr. Kit Yan Chan	Curtin University, Australia
Dr. Longbiao Chen	Xiamen University, China
Dr. Ying Guo	Commonwealth Scientific and Industrial Research Organization, Australia
Prof. Kao-Shing Hwang	National Sun Yat-sen University, Chinese Taipei
Prof. Fakhreddine Karray	University of Waterloo, Canada
Dr. Michail G. Lagoudakis	Technical University of Crete, Greece
Dr. Hak-Keung Lam	King’s College London, UK
Prof. Shutao Li	Hunan University, China
Prof. Yiguang Liu	Sichuan University, China
Prof. Kin Huat Low	Nanyang Technological University, Singapore
Dr. Vincenzo Pacelli	University of Foggia, Italy
Prof. George Panayiotakis	University of Patras, Greece
Dr. Phyo Phyo San	Nanyang Technological University, Singapore
Prof. Steven Weidong Su	University of Technology Sydney, Australia
Prof. Theodore B. Trafalis	The University of Oklahoma, USA
Dr. Jue Wang	Chinese Academy of Sciences, China
Prof. Xuesong Wang	China University of Mining and Technology, China
Prof. Simon X. Yang	University of Guelph, Canada
Prof. Shi Ying	Wuhan University, China
Prof. G. Peter Zhang	Georgia State University, USA
Dr. Yudong Zhang	Columbia University, USA
Dr. Yi Zhao	Harbin Institute of Technology Shenzhen Graduate School, China
Prof. Wei Zhong	University of South Carolina Upstate, USA
Prof. Dingxuan Zhou	City University of Hong Kong, China
Prof. Antanas Zilinskas	Institute of Mathematics and Informatics, Lithuania

Scientific Research Publishing

Journal of Intelligent Learning Systems and Applications, 2023, 15, 1-56
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Table of Contents

Volume 15 Number 1 February 2023

Journal of Intelligent Learning Systems and Applications (JILSA)
Journal Information

SUBSCRIPTIONS

The *Journal of Intelligent Learning Systems and Applications* (Online at Scientific Research Publishing, https://www.scirp.org/) is published quarterly by Scientific Research Publishing, Inc., USA.

Subscription rates:
Print: $89 per copy.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

COPYRIGHT

Copyright and reuse rights for the front matter of the journal:

Copyright for individual papers of the journal:

Reuse rights for individual papers:
Note: At SCIRP authors can choose between CC BY and CC BY-NC. Please consult each paper for its reuse rights.

Disclaimer of liability

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: jilsa@scirp.org

Scientific Research Publishing

Journal of Intelligent Learning Systems and Applications, 2023, 15, 1-23
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Architecting the Metaverse: Blockchain and the Financial and Legal Regulatory Challenges of Virtual Real Estate

James Hutson[1], Gaurango Banerjee[2], Naresh Kshetri[3], Kurt Odenwald[4], Jeremiah Ratican[5]

[1]Department of Art History and Visual Culture, Lindenwood University, St. Charles, USA
[2]Department of Finance, Lindenwood University, St. Charles, USA
[3]Department of Cybersecurity, Lindenwood University, St. Charles, USA
[4]Department of Business Law, Lindenwood University, St. Charles, USA
[5]Department of Game Design, Lindenwood University, St. Charles, USA
Email: GBanerjee@lindenwood.edu, NKshetri@lindenwood.edu, JHutson@lindenwood.edu, KOdenwald@lindenwood.edu, JRatican@lindenwood.edu

How to cite this paper: Hutson, J., Banerjee, G., Kshetri, N., Odenwald, K. and Ratican, J. (2023) Architecting the Metaverse: Blockchain and the Financial and Legal Regulatory Challenges of Virtual Real Estate. *Journal of Intelligent Learning Systems and Applications*, **15**, 1-23.
https://doi.org/10.4236/jilsa.2023.151001

Received: December 6, 2022
Accepted: February 4, 2023
Published: February 7, 2023

Abstract

There has been disagreement over the value of purchasing space in the metaverse, but many businesses including Nike, The Wendy's Company, and McDonald's have jumped in headfirst. While the metaverse land rush has been called an "illusion" given underdeveloped infrastructure, including inadequate software and servers, and the potential opportunities for economic and legal abuse, the "real estate of the future" shows no signs of slowing. While the current virtual space of the metaverse is worth $6.30 billion, that is expected to grow to $84.09 billion by the end of 2028. But the long-term legal and regulatory considerations of capitalizing on the investment, as well as the manner in which blockchain technology can secure users' data and digital assets, has yet to be properly investigated. With the metaverse still in a conceptual phase, building a new 3D social environment capable of digital transactions will represent most of the initial investment in time in human capital. Digital twin technologies, already well-established in industry, will be ported to support the need to architect and furnish the new digital world. The return on and viability of investing in the "real estate of the future" raises questions fundamental to the success or failure of the enterprise. As such this paper proposes a novel framing of the issue and looks at the intersection where finance, technology, and law are converging to prevent another Dot-com bubble of the late 1990s in metaverse-based virtual real estate transactions. Furthermore, the paper will argue that these domains are technologically feasible, but the main challenges for commercial users remain in the legal and regulatory arenas. As has been the case with the emergence of online commerce, a legal assessment of the metaverse indicates that courts will look to traditional and established legal principles when addressing issues until the enactment of federal and/or state statutes and accompanying regulations. Lastly, whereas traditional regulation of real estate would involve property law, the current legal framing of ownership of metaverse assets is governed by contract law.

Keywords

Blockchain, Digital Real Estate, Digital Retail, Digital Twin, Digital Content, Finance, Metaverse

1. Introduction

The word "Metaverse" is a finalist for the word of the year for Oxford University Press, and the past year has demonstrated why. Formerly Facebook, Meta invested billions into building an immersive reality platform, while Decentraland and The Sandbox worked to create Web3 versions of 3D social environments [1]. With films like *Ready Player One* (2018), immersive virtual environments are at the forefront of popular imagination, and seemingly within reach. Ball [2] defined the metaverse in terms of an interoperable and scalable real-time network that can be experienced in 3D virtually and synchronously. There would be no limit on the number of users and all interactions, including social, financial, and so on, would take place within this persistent virtual community. However, the version of the metaverse that we currently see with companies like Meta is not interoperable or scalable. Instead of being inside a virtual or 3D version of the internet, there are many portals to enter different virtual spaces on different platforms. Therefore, we actually have a multiverse of metaverses existing simultaneously [3]. One limitation on interoperability is that technology alone cannot be leveraged to build the metaverse. A stable and trustworthy economy is also required, and the current digital economy is still based around centralized organizations, such as banks and companies, who own digital properties instead of users [4]. Blockchain-related technologies have been cited as a potential solution for decentralization and a step closer to a true metaverse [5].

First introduced as bitcoin cryptocurrency, blockchain technology laid the foundation of today's digital currency market given the ability to create a shared economy [6]. As a ledger that stores committed transactions, blockchain facilitates digital asset tracking and security in a financial network. Each transaction is stored as a block that is linked together using cryptographic methods or hashing mechanisms to ensure the immutability of the ledger. The process allows for the potential for securing sharing even when the commercial environment is insecure. Most importantly, blockchain can operate without a centralized authority to oversee it and thus operates on a decentralized ledger [7]. Given that proof of work is employed as the consensus mechanism in blockchain, the process is ideal for e-commerce platforms. When considering the metaverse, blockchain can provide accountability in an unproven and unstable digital, financial eco-system. For example, since the ledgers are plainly visible to all involved, trust is not necessary between different parties as there is improved cybersecurity and protection against data manipulation with security provided by the network of participants and the technology itself [8] [9] [10]. Intermediaries that provided trust and security, such as central banks, and central counterparties would be obsolete [11] [12]. At the same time, the potential benefits of blockchain do not exist without inevitable challenges. Current world authorities have found applying technology-neutral regulation to distributed ledger technology (DLT) in finance. The emergence of cryptocurrencies also has the ability to bypass existing legislation regarding money laundering and has been used to facilitate illegal activity [13] [14] [15], which resulted in calls to action and further regulation [16] [17] [18].

In addition to blockchain disrupting financial models and potential challenges to the centralized banking system to monitor transactions, the potential social, professional, and economic changes to be wrought by the forthcoming metaverse should be equally weighed. In fact, the two will work in tandem to form a new manner in which the global economy will exchange information and currency in the near future. Businesses must now operate in three dimensions of commerce. Whereas having a brick-and-mortar establishment was sufficient to remain competitive until the rise of smartphones and web shopping, so was having a website domain with a company's name until the rise of the metaverse. Every professional, let alone every business is expected to have a website to be seen as credible, legitimate, and also to provide a portal to review and purchase potential goods and services.

Retailers are expected to have high-quality photographic reproductions of their products for consumers to peruse. The value of the website domain (and specific name) can thus be quantifiable, just as the value of a physical store shoppers purchase goods and services. Many companies have been using the metaverse to move beyond the two-dimensional display of products to "retail theater" and three-dimensional products that consumers can interact with as avatars in a 3D virtual world. The latest round of companies is continuing the trend since Second Life launched in 2003. Therefore, from the perspective of marketing and advertising, the investment in "real estate of the future" in the metaverse can be equated to the value of a company website. In the same fashion, a digital twin of a store in the metaverse must be designed, built, hosted on a server, and maintained. But how will businesses move past mere advertising in the metaverse to consider digital twins of retail establishments economically viable?

But what about investments that seem to outpace that value proposition? Given the rise of digital locations that only exist online, digital real estate has provided the ability for creators to sell directly to others in the metaverse. As an example, owning digital real estate and selling tickets to virtual events or digital art as non-fungible tokens (NFT). There are many opportunities to prospect land in the metaverse with more and more platforms appearing. In the real estate market, new construction of luxury homes is being marketed by pairing a digital twin of the new construction in the metaverse. ONE Sotheby's International Realty has recently aligned with a general contractor and NFT collector to offer a virtual counterpart to its real-world mansion in The Sandbox metaverse platform. The Purchaser of the NFT asset, which will be transacted on the Ethereum blockchain, will also acquire the rights to physical property [19]. Firms have begun to spend millions to purchase space in the metaverse, seemingly banking on the value of the blockchain-supported technology becoming something more than a 3D website. For instance, in 2021, an investment firm purchased 2000 acres of virtual real estate for around $4 million. The virtual real estate existed only in the metaverse platform known as The Sandbox. The firm then owned the equivalent of 1,200 city blocks in the virtual space which was paid for via 792 nonfungible tokens (NFTs) on Ethereum blockchain [20]. Another example can be found with the investment fund Republic Realm, which in June, 2021 bought a parcel of digital land in Decentraland for more than US$900,000. The fund has plans to develop the virtual plot of land into a virtual mall named Metajuku, which is designed after the Harajuku district in Tokyo [21]. As such, various firms are most interested in platforms with specific market caps because limiting the number of parcels of virtual real estate limits supply. As with Decentraland and The Sandbox, which both have market caps, as the demand for digital real estate rises, so does the price per parcel of land.

At the moment, however, the nature of ownership in the metaverse is still being negotiated [22]. What remains unclear still is which platforms will become the most popular. And because of that, it can be challenging to determine what real estate in the metaverse will be worth a significant amount in the future. In these and other examples, the spaces and even the products that they sell are either hosted on privately owned servers, or products sold in the metaverse as NFTs still limit control of the digital asset. With this in mind, what will be the major legal considerations with such an approach in the future? To allow economic interactions in the metaverse, the platform must allow for currency, goods, and services to be traded. While blockchain has the potential to address these challenges, immediate obstacles will be regulatory and not technological. For instance, even though there are claims that virtual ownership is guaranteed, the current legal framing of ownership of metaverse assets is not governed by property law at all, but rather by contract law. Ownership of land has always been a fundamental right with accompanying privileges in the United States. Indeed, the right to vote in most states was limited to freeholders upon the ratification of the Constitution. Without legislative change, the statutory protections attendant to the ownership, possession, and sale of real property will not apply to metaverse property, which is not land—but a digital asset. This study seeks to investigate potential future research and consider the forthcoming regulatory issues with metaverse-based virtual real estate transactions.

2. Literature Review

2.1. Building Regulation: Blockchain Finance and the Metaverse

A potential solution to regulation in a decentralized financial system is embedded supervision. Auer [23] argues that the rise of blockchain in finance will improve efficiency regarding supervision. Embedded supervision would provide a specific regulatory framework where a specific ledger would no longer need to be verified given the automatic monitoring built in with blockchain. However, the conditions required are currently hypothetical and would need to consider the following [24]. As outlined by Auer [23], embedded supervision must:

1) Be supported by a regulatory framework and an effective legal system;

2) Applied to achieve economic finality-once a transaction is not profitable to undo;

3) Be designed for economic consensus, knowing that the market will be automatically supervised; and

4) Promote low-cost compliance to be equitable for both large and small firms

The benefits of such an approach are evident in that, like blockchain, there is no need for oversight, which is also the primary point detractors point out. Such embedded supervision could easily be adapted to the current e-commerce strategies from the gaming industry. In fact, the combination of blockchain technologies and gaming has already led to play-to-earn games with tokens that use their own economy, commerce, and so on. Blockchain developers have drawn inspiration from gamification, which can currently be seen in Decentralized Finance (DeFi) and GameFi [25].

Moreover, developments in the gaming industry and blockchain continue to run parallel, leading to inevitable integration. With respect to the metaverse, blockchain is well-suited as a decentralized, financial solution for the following reasons, as outlined by Turdialiev [26]:

1) Digital proof of ownership: Through digital wallets, ownership can be demonstrated with regards to any asset on the blockchain.

2) Digital collectability: Using NFTs, entirely unique assets can be created that can be collected, reflecting practices in the real world.

3) Transferable value: Current multiplayer games online can transfer value between users. Such an approach can be adopted with blockchain as more currency is exchanged in the metaverse.

4) Governance: In a decentralized system, blockchain can replace centralized authority and ensure rules are adhered to instead of elected officials.

5) Accessibility: Instead of limiting who can open an account, as in a standard bank, digital wallets are open to the public to create with blockchain.

6) Interoperability: Developers are already creating custom blockchains that are interoperable, such as Polkadot (DOT) and Avalanche (AVAX). In a true metaverse, interoperability will be key and blockchain has demonstrated potential in this area.

Since the metaverse is envisioned as a parallel plane for human activity, the relative success of the enterprise will depend upon a strong, robust and secure economy. While this new virtual economy may seem a far cry from that in current use, Ball [2] points out that the metaverse economy will follow real-world patterns. The attributes that contribute to a thriving economy include competition, profitable businesses, agreed upon "rules" and sense of "fairness," along with consistent consumer spending and rights.

2.2. Payment Rails

However, there is one major factor that will shape the exchange of currency for goods or services in this new digital realm, and that is payment rails. There have been a number of new payment rails created thanks to communication technologies. In fact, the use of cash as a method of transaction has been dramatically declining. As Ball [2] relates, from 2010 to 2021, the share of US transactions that used cash dropped from 40% to nearly 20%. Today, the most common payment rails in the US are CHIPS (Clearing House Interbank Payment System), Fedwire (formerly the Federal Reserve Wire Network), ACH (Automated Clearing House), along with various credit cards and peer-to-peer payment applications such as PayPal, Venmo and others.

In considering how these transactions were used to purchase rights to software, one would imagine that the growth in the virtual world would have led to advances that were more flexible and forward thinking. In 2021, consumers spent over $50 billion on digital-only video games and the GDP of this virtual world quintupled since 2005. With that being said, payment rails of the virtual economy are more restrictive than in the real-world due to forced bundling of services, such as PlayStation's wallet, Apple's Apple Pay, and in-app payment services. Consoles such as Xbox and PlayStation allow consumers to download a version of a game, but only for use on their hardware. In 2003, Valve launched Steam as a PC alternative to the console economy. As many multiplayer online games were moving to a "games-as-service" model anyway, Valve was able to handle game updates and install internally with a "game launcher" that indexed and centrally managed the game installer files. The approach also handled a user's rights to the games, allowing automatic download when desired. The economic model still ensured 30% ongoing revenues for Valve as every sale kept that amount as with console game platforms. The 30% payment rails also govern Apple and Google and their app stores, which additionally restrain virtual world platforms.

2.3. Rise of Metaverse Retail

While seemingly only recently dominating financial headlines, the term "metaverse" has been applied to retail for almost two decades. Bourlakis and Papagiannidis [27] investigated the emergence of metaverse retailing following the release of the first smartphones and divided the evolution into three phases: traditional, electronic, and metaverse. The study focused on the new strategies retailers needed to adopt to operate in three different, but intertwined spaces. Key promotional aspects are highlighted with the different challenges faced by traditional (brick-and-mortar retailers), e-retailers (utilizing the Internet), and metaverse retailers. For the study, the researchers analyzed the "metaversephenomenon" of Second Life, but also note *World of Warcraft*, *Ever Quest*, *Eve Online*, and *Star Wars Galaxies*, as ground zero for the "third dimension of commerce". As evinced by the examples cited, most metaverses began as games, or, more specifically, massively multiplayer online role play games (MMORPGs). These would quickly evolve into alternate worlds that extended players virtual and electronic spaces. Given the sheer number of consumers spending large swaths of time (in some cases 12 hours straight) in these virtual environments, new social and business environments grew to accommodate with larger spaces. With such economic and social exchanges taking place between the players in these games and metaverses, greater crossover became common between physical businesses in the real-world and those e-businesses in the virtual. These developments have led to a new multi-faced, multi-spaced economic environment that has vastly increased in complexity. The intertwined nature of this new business environment, electronic, virtual, and physical space must be mapped out in order to conceptualize the economic, social and policy implications [28]. In order to be successful, Bourlakis and Papagiannidis [27] recommend a holistic promotional strategy that operates in all three arenas.

The precursor to the contemporary metaverse is often cited to frame many of these discussions. Launched in 2003, Second Life allowed users to monetize their efforts in the virtual world. Copyright for content created by users on the platform belonged to users, who were then able to monetize said content. Predating NFTs and blockchain technology, Linden Lab's Second Life (https://www.lindenlab.com/) allowed content creators to protect their creations using a system of three options either allowing or blocking owners to copy, modify or transfer their creations or purchases [29]. The system also tracked items and their creators, as well as functions in similar capacity creating digital items that are unique and identifiable. All transactions in Second Life were based on the Linden Dollar. These can be exchanged in-world for goods or services but can also be transferred for real currency to benefit creators in the real world. Given that a business license is not required to operate in the virtual platform as an entrepreneur, identifying how many are trading and how successful they are is not possible. The Positive Linden Dollar Flow (PMLF), however, is used to estimate the more than 66 million "business owners" on the platform. Businesses perceived the value of reaching customers in such a platform and began operating in the space [30]. Examples of these real-world firms span many different markets and industries, including ABN AMRO, Adidas, American Apparel, Dell, Harvard Law School, IBM, Microsoft, Pontiac, Reuters, Sony Ericsson, the Swedish Government, Toyota, and others. While most examples here use the platform merely for marketing purposes, some have announced intentions of actually trading in Second Life. Mainstream adoption still remains unrealized [31].

The major difference in traditional or e-retailing from that found in the metaverse is how customers expect to interact with the brands they encounter. In traditional two-dimensional marketing and advertising, print, images, videos, and music may be introduced to provide some idea of the product being sold, but in the metaverse customers expect to interact with it in a three-dimensional simulation. The practice is what Harris, Harris, and Baron [32] predicted with the rise of "retail theater". Papagiannidis and Bourlakis [33] argue that some retailers in Second Life designed experiences where potential customers could interact with products to lead to sales. Still other retailers prefer to develop a sense of community or belonging among their consumer bases. Even though this affords even greater access and the ability to customize marketing to specific customers, Haig [34] warned, and Bourlakis and Papagiannidis [27] reiterated, that businesses should be wary of the effects of overly bombarding potential customers. For example, when visiting busy locations in a metaverse platform, an automated system may deliver messages or notecards with information for products or services. While users have the option of muting a bot or automated agent, doing so repeatedly can result in frustration [35]. At the same time, if experiences are judiciously designed, the augmented 3D information provided can positively affect sales and the retail shopping experience. Virtual retail affords the ability to combine augmented reality (AR) and virtual reality (VR) to create a seamless shopping experience. Virtual objects and digital information can be viewed and reviewed within a virtual space. The benefits are self-evident since instead of viewing products on a flat screen, billboard, or piece of paper, the product could be transported out of the catalog and placed in the real environment or clothes modeled for size and fit. Such product experiences could assist in a purchasing decision and lead to greater assuredness among customers [36].

2.4. Law and Finance in the Metaverse

New legislation and regulation rose in response to the internet age [37]. Similarly, the potential economic and social change on the horizon with the metaverse will require addressing the disruptive influences on current law [38]. There is a precedent for the metaverse and that is with artificial intelligence (AI), which will modify the legal role of behavior and require new antitrust or contract laws [39]. The major consideration is how the metaverse and XR will change how objects interact in real or virtual space, and, in turn, how humans interact with and use them. For instance, augmented reality (AR) applications overlay digital objects onto the real world; virtual reality (VR) immerses users in a completely virtual environment and, using avatars, these users interact with others in virtual spaces and with virtual objects. As with AI, AR and VR have the potential to disrupt legal categories by way of the distinction between a real and virtual object and issues of ownership [40]. With an alternate comprehension of virtual and social surroundings, the potential for legal disruption is high. The use of an avatar, which can be digitally altered to look like virtually anything real or imagined, also complicates the matter. As humans interact in virtual or augmented environments, the potential for legal problems via relationships and legal expectations will arise that have not been considered until now. As virtual objects, such as NFTs, become closer to or combined with physical objects, the more legal expectations of ownership will blur. Furthermore, the more time spent in virtual spaces for more social and business interactions, the more questions will rise regarding legal complexities [41]. Dwivedi *et al.* [38] pose two key questions for further consideration: 1) How will the personality in the metaverse in avatar form be protected considering data protection laws and the mutability of the avatar and individual behavior in a virtual environment? and 2) How will concepts of property law need to develop to address virtual land and real estate in relation to blockchain technology?

2.5. Purchasing Land in the Metaverse

The history of purchasing land in the metaverse varies from incarnation to incarnation. Early metaverse contender Second Life had digital land "ownership" built right in the paid tier of gameplay. While users can play for free if a user opts for the paid subscription, they are awarded a small parcel of land that they can develop on. This digital land ownership has been the focus on several court cases with Linden Labs, eventually leading to the removal of the term "owned" from the marketing materials. This process of land "ownership" led to the rise of one of the most well-known business owners and real estate moguls in Second Life, Anshe Chung. Anshe Chung is the avatar of Ailin Graef and was featured on the cover of BusinessWeek magazine and has been referred to as the "Rockefeller of Second Life" by CNN. Anshe Chung was reported by Fortune magazine as the "first virtual millionaire" through purchasing a renting virtual real estate and charging land taxes [42].

Another virtual platform that began as a game, MindArk's Entropia Universe, originally Project Entropia before it's metaverse expanded into multiple planets, has broken several Guiness World records for owning "the most expensive virtual item" in reference to a digital property [43]. In 2005, NEVERDIE, an avatar of Jon Jacobs, purchased an asteroid space resort in a public auction for $100,000 USD or 1,000,000 PED. Much like Second Life, Entropia Universe has a currency exchange rate with the Project Entropia Dollar's exchange rate being 10:1 or rather it takes 10 PED to equal $1 US. This record would be passed in 2009 by the sale of Crystal Palace space station for 330,000 USD and again a year later when the planet Calypso was sold for $6 million [44]. LAND also uses NFTs, which measure 16 × 16 meters in parcels of land which can be purchased with the MANA cryptocurrency on the Decentraland platform. The purpose of such a purchase is for owners to build on these virtual spaces and earn money from them through rent or other means, creating a complex crypto economy (Bitlo, 2022). There are a growing number of platforms on which such virtual land can be bought and sold, including Decentraland, The Sandbox, Somnium Space, OVR, SuperWorld and Axie Infinity, Bloktopia, Next Earth [45] [46]. In order to purchase land on these platforms, a digital wallet must be created first. Once the wallet has been created, these companies can be searched, and the desired plots of virtual property may be purchased using specific cryptocurrency of that platform via the digital wallet. Many companies have seen the value and have opened stores, such as Samsung in Decentraland. The value, and thus cost, of virtual land often increases the closer to real-world regions (e.g. Paris or New York). The value may also increase depending on the features on the land, its size, or other objects contained within [47].

There have been notable detractors that point to the early limitations and volatility of virtual real estate prospecting, such as with Decentraland. The value proposition with Decentraland lies in the purchase of land on the platform, but the process is complicated. For instance, future virtual landowners cannot purchase tokens directly with standard currency. Even ether (ETH), the most popular bitcoin alternative, cannot be used to purchase virtual real estate. In the case of Decentraland, like other crypto projects, a cryptocurrency unique to the platform called MANA (ERC-20 token) must be used [21]. The most affordable plots of virtual land on the platform sell for around 4000 MANA, or the equivalent of nearly $2489 (down more than two thirds in value in the last year). Since the virtual land is non-fungible, the owner of a plot of land owns it until another user wishes to purchase from them. Alternatively, MANA can be sold to other users who may have needed to purchase land and be exchanged between users on the platform [21]. Given the volatility of the crypto market, the cost and value of land can be influenced relatively easily and quickly by several factors. Therefore, the value of virtual storefronts to generate revenue is unpredictable and in some cases has quintupled in value in about a month and then dropped dramatically [21].

2.6. Ownership and NFTs

While blockchain has the potential to undergird the metaverse and replace existing payment rails, concepts such as ownership cannot be readily transposed from the traditional economy into the new virtual world. Even when using blockchain to decentralize digital assets, reviewing the terms of service of the specific metaverse platform these were purchased on is still necessary [20]. The prevailing belief of those that support crypto currency is that true ownership of NFTs is possible due to decentralization and interoperability. With such an understanding, owners believe that tokens provide non-fungible proof of ownership of a digital asset that can be used across metaverse environments [48]. Because of decentralization, the ability to buy and sell virtual items on the blockchain is believed even without an individual or company providing permission [41] [49] However, despite claims of ownership, the situation is more complicated given that current ownership of metaverse assets is governed, not by property law, but contract law. As noted above, Marinotti [20] clarified the nature of ownership is different in the physical and virtual worlds and consumers may be misled. When an item is purchased in the metaverse, the transaction is recorded on a blockchain, which is a decentralized, digital ledger where such records cannot be deleted or altered [49]. As outlined, the process assigns ownership of an NFT in the user's digital, crypto wallet that can only be accessed by the owner. Since access is only possible via the "wallet's" private key, the NFT appears to be inaccessible to anyone other than the owner. However, a distinction needs to be made between the NFT and the digital asset because owning digital objects in a virtual world is not the same as in the physical world [41].

The distinction made here with "ownership" is outlined in the terms and conditions of service. Upon first joining any metaverse platform, users are required to agree to the terms of service, terms of use, or end-user license agreement. Since these are legally binding documents, the legal rights of users are defined. Most users do not read these terms of service. One study concluded that only 1.7% of users were able to locate and then question the "child assignment clause" which is embedded in a terms-of-service document, giving away their firstborn child [50]. Not surprisingly, the legal nuances of ownership are outlined in these long and dense documents, and unlike blockchain, the terms of service for each platform are centralized. Given that legal ownership is controlled by a single company, and that existing multiverse of metaverses is not connected, a user is unable to move an avatar or other digital asset between virtual worlds. Platforms are still connecting specific NFTs to proprietary digital assets. Therefore, according to the terms of service, those NFTs purchased on a metaverse platform and the digital goods they represent are rarely the same thing. While NFTs exist on the blockchain, the digital assets and real estate are stored and only exist on private servers on inaccessible databases [51]. With companies owning the servers on which digital assets are stored, they also have the ability to delete links and decouple to disallow use from owners. These platforms also reserve the right to amend their terms of service at any time and are often not required to provide notice to users [52]. In order to know if one is compliant with the terms of service, users would need to refresh and then reread the terms to ensure any language has been added that would lead to their banishment from the platform and deletion of their assets.

3. Analysis

3.1. Financial Considerations

The following analysis considers the previous sections, potential and volatility of virtual real estate in terms of financial, cyber security, and legal implications. In analyzing the financial implications for virtual real estate, one should first consider the costs and benefits of the application of block chain technology in this sector. Since the medium of exchange in this virtual sector would involve cryptocurrency, a brief review of the evolution and potential challenges in using digital currency would also be prudent. The decentralized finance system underlying blockchain technology has some clear advantages over the current centralized system. There would be increased efficiency in remittances in real time by avoiding any delay whatsoever in transaction receipts through traditional financial intermediaries in a centralized banking system [53]. The use of digital currencies in virtual real estate transactions could also ensure a more democratic process in terms of enabling easier access to users without a traditional bank account, while maintaining security in terms of digital identity validated through the sequential coding technology in blockchain transactions. As a result, transaction costs would fall dramatically due to the low cost of digital payments [54].

On the other hand, the potential challenges in using blockchain technology in transactions in the virtual real estate space should also be considered. A fundamental requirement of a medium of exchange in a market transaction between two parties involves whether that action is a good store of value and universally acceptable medium of exchange for the involved parties in the market. These two properties are currently lacking in the cryptocurrency market. One of the main problems with bitcoin's usage has been related to its extremely high volatility in market value thereby increasing the financial risk of use and limiting acceptability as a standard medium of exchange. High volatility in value for cryptocurrency could spill over to the value of virtual real estate using blockchain technology. The situation could potentially create a "virtual real estate crisis" in a decentralized financial system.

The costs and benefits of the application of blockchain technology can be compared but would instead be a matter of "when" and not "if" decentralized finance becomes the dominant financial system with the support of the younger tech savvy generation. Central banks across the world are aware of the acceptance of the decentralized financial system by the future generations and are presently designing central bank digital currencies (CBDC) that may be regulated to provide more stability to the system [55]. Since store of value is a fundamentally important property of an acceptable medium of exchange, we could foresee the application of blockchain technology in virtual real estate transactions in a more regulated cryptocurrency market in the future.

3.2. Infrastructure and Cybersecurity Considerations

Turning to the infrastructure and cybersecurity considerations for virtual real state, the measure of security over the web and trust represented is paramount. Managing transparency, trust along with the satisfaction of customers and citizens is needed to improve the efficiency of public service delivery [56]. There is an overwhelming aversion to centralizing authority in today's cyber world [57]. Rich user interaction and user involvement can be defined via metaverse in its digital representation. The technology behind the world's popular cryptocurrencies (the disruptive Blockchain technology) has numerous applications, and among them some of the major advantages can certainly benefit virtual world of real state in the metaverse [58]. The limitations of only investing in physical property will sooner or later come to an end because the new world of virtual real estate is growing and blooming [59].

In recent days, blockchain technology has been applied beyond finance and metaverse, including healthcare, public service, governance, currency exchange, food supply chain, e-voting, music royalty tracking, personal identification security on web, and elsewhere. Blockchain can also be considered in agricultural supply chain (popular as agribusiness) where there is lack of customer trust or traceability [60]. Blockchain can promise several advantages such as product traceability, efficiency enhancement, improving quality, benefiting farmers, and building customer's trust over traditional supply chains in agribusiness. The efficiency in supply chain management (SCM) can be improved and delivered in real-time to all members (especially to the farmers) that can change the product inventory and product price.

The infrastructure and analysis for metaverse architecting and regulatory challenges rely on the type of blockchain used. Although we have three popular types as public blockchain (permissionless BC), private blockchain (permissioned BC) and hybrid blockchain, depending upon the mode of peer participation, financial perspectives can be slightly different. Blockchain types based on financial perspectives with respect to business and currency can be categorized as C2C (Type One), B2C (Type Two) and B2B (Type Three) types [61] (**Figure 1**). Type One C2C is the Only Cryptocurrency blockchain type with High-Node scalability (and low-performance scalability). Type Three B2B, on the contrary, is the Only Business type with Low-Node scalability (and high-performance scalability). Type Two B2C is the Cryptography + Business type with High-Node scalability (and low-performance scalability).

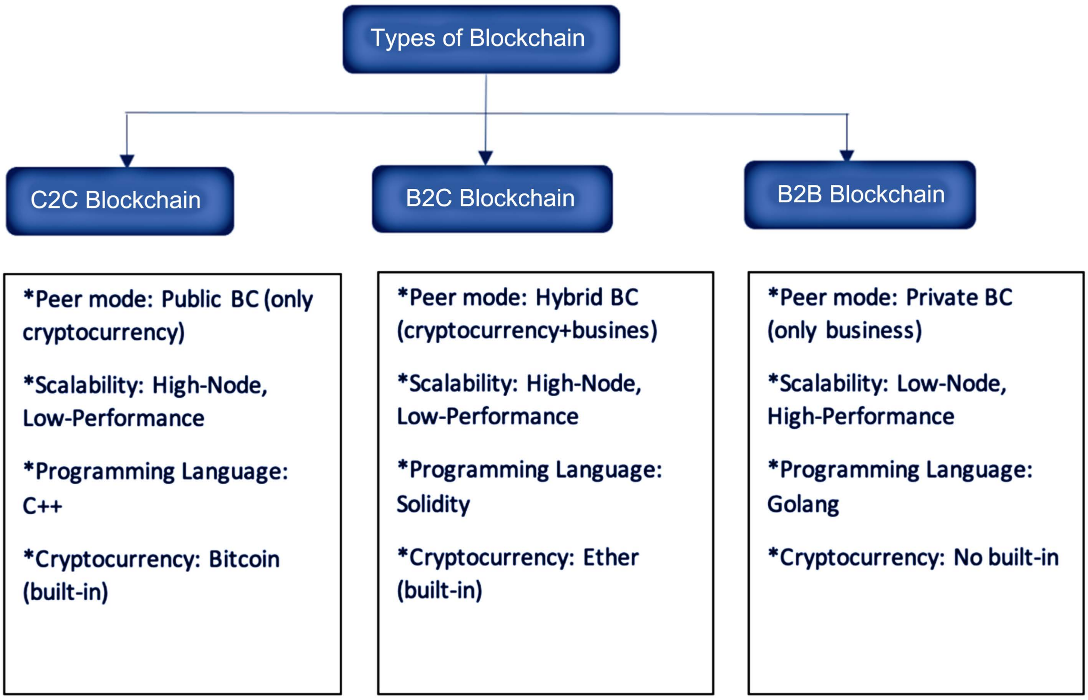


Figure 1. Three Blockchain types with mode of peer participation, scalability, programming language used, and built-in cryptocurrency of that blockchain type.

The financial and legal perspectives while architecting the metaverse are incomplete without cybersecurity and cyber defense considerations. Since the rise of the internet age, cyber war has been a concern. With more recent technological advancements, the global rise of online users and devices multiplied and become ever more complex, leaving new avenues for cyber-attack. These threats can also be seen in the financial and military sectors [62]. The upcoming new internet infrastructure will be at greater risk if the issues are not dealt with more effectively.

3.3. Legal Considerations

As noted above, the virtual world of the metaverse greatly surpasses the internet as we have known and presents significant challenges for persons and investors seeking protection for their investments-often sizeable-in this new frontier. As Lehot notes, "While the real-world property has years of established laws behind it, metaverse virtual land is the new Wild West…" A fundamental problem is how the law will recognize-or not recognize—the tokenization of real property as a digital asset. What does one acquire when purchasing an NFT? As noted by certain commentators, the concept of tokens being tied to property rights is not a new concept. Negotiable instruments, securities, deeds, and bills of lading are some examples where a document or certificate served as the basis for establishing a property right in a physical asset But the relationship of an NFT to real world assets is questionable. To determine what rights are obtained when acquiring an NFT, one must review the terms of service and associated agreements of those entities enabling the creation and transfer of NFTs. Moringello and Odinet [63] did just that when reviewing the service documents of eight platforms. Consistent in their findings was a disconnect between the broad statements suggesting the conveyance of property rights which would include the right of ownership, possession and control of the underlying asset and the actual terms of service, which were more analogous to licensing agreements that often reserve considerable rights to the platform to remove access to the NFT's on their sites and remove the user's assets from the site. In fact, the binding Terms of Service reviewed for Superworld placed quotation marks around the word "purchase" and reserved Superworld the right to terminate a user's access to its metaverse in certain circumstances [63].

Adding to the illusion of property rights in an NFT is the fact that the NFTs are not tethered to a physical thing—unlike a deed which has a legal connection to land, a title which has a legal connection to an automobile, or negotiable instruments which grant the party in possession of the instrument certain rights in an underlying debt. In these examples, an underlying body of law establishes the underlying control that is inherent in a property right [63]. Acknowledging the deficiencies of existing laws to address these conflicts, a committee was appointed to draft amendments to the Uniform Commercial Code to address emerging and past-emerged technologies. Any amendments approved and suggested by the American Law Institute and the Uniform Law Commission must be enacted by the state legislatures before having the force of law.

Currently, the most critical issue to be addressed regarding NFTs and real estate in the metaverse is found in the representations made by the platforms as to the rights received upon obtaining an NFT and the reality of what has been received. A major consumer protection issue in the metaverse is misrepresentation. NFT platforms often directly promote NFTs as being capable of conveying more than what the law will allow or send mixed messages about what is being offered and what the buyer will obtain [63]. Concurrent with claims of misrepresentation is the duty of performing due diligence. Sufficiently evaluating investments in the metaverse will require an understanding of the custody of the digital asset and the terms and conditions of the platform. Failure to perform the required due diligence to become familiar with those terms and conditions can severely undermine a claim of misrepresentation when the person who thought he or she acquired ownership of 100 parcels in the metaverse as evidenced by the NFT acquired from the platform one day learns that he or she no longer has access to that platform's metaverse. Courts will apply traditional principles of contract law to address contract disputes—even those disputes arising in the metaverse. Why? Because those disputes will be anchored to the contracts entered between the purchaser and the platform when acquiring the NFT that represents the purchaser's "ownership" of assets within the metaverse.

Without question, the federal government has the authority to investigate and pursue legal action in claims of deceptive trade practice against metaverse platforms. The Federal Trade Commission is empowered to police unfair and deceptive trade practices under Section 5 of the Federal Trade Commission Act. Also, many states authorize their state attorney general to act against unfair and deceptive trade practices under similar statutes. Missouri for example authorizes investigations and action by the state attorney general under the Missouri Merchandising Practices Act. But reliance on government to address claims of deceptive trade practices is not always practical. To be sure, one train of thought rejects the idea of governmental involvement when claims arise in the NFT market and suggests that such disputes are best left to be resolved by the private parties involved. Given a strong line of cases issued in recent years by the United States Supreme Court, purchasers of NFT real estate and other digital assets in the metaverse who seek to litigate claims against the metaverse platforms in courts of law will encounter a major obstacle commonly found in contracts today—mandatory arbitration and class action waiver provisions.

Nearly all the NFT minting platforms contain mandatory arbitration and class action waiver provisions in their service contracts [63]. Beginning with *AT & T Mobility LLC v. Concepion*. 563 U.S. 333 (2011), continuing with *Epic Systems Inc. v. Lewis*, 584 U.S., 138 S. Ct. 1612, 200 L. Ed. 2D 889 (2018) and most recently in *Viking River Cruises Inc. v. Angie Moriana*,. 596 U.S., 142 U.S. 1906 (2022), the Supreme Court has consistently reaffirmed the validity and enforceability of such provisions under the Federal Arbitration Act (FAA), "which makes arbitration agreements 'valid, irrevocable and enforceable, save upon such grounds as exist in law or in equity for the revocation of any contract.'" *Viking River Cruises* 142 U.S. at 1917, citing 9 U.S.C. Section 2. The prolific use of these mandatory clauses in contracts today has brought increased judicial scrutiny of such clauses. Such scrutiny has critically reviewed and sometimes voided mandatory arbitration and class action waiver clauses on common law contract principles (most often because the agreement lacked consideration, or because the terms of the agreement were found to be unconscionable), Like any contract, arbitration agreements may be invalidated by generally applicable contract defenses such as fraud, duress or unconscionability. *Rent-A-Center West*, *Inc. v. Jackson*, 130 S.Ct, 2772 (2010). But this same scrutiny has also reinforced the pre-emptive force of the FAA. Specifically, the FAA generally requires courts of both federal and state jurisdiction to uphold such provisions, and likewise curbs the power of state legislatures to enact legislation either limiting or invalidating mandatory arbitration and class action waiver provisions contained within an otherwise valid contract. Given the Supreme Court's recurrent judicial pronouncements reaffirming the pre-emptive force of the FAA, and absent congressional action amending it, conflict and disputes within the metaverse seem destined to be addressed within the more private and confidential arena of arbitration—an arena exempt from procedural and evidentiary rules and not limited by the boundaries of judicial precedent.

Generally, arbitration exists as an alternative form of dispute resolution—allowing parties to seek redress of claims outside of litigation in the courts. The parties choose an arbitrator who will conduct a hearing, take evidence, and make a binding decision on them. An arbitration award is final and is subject to appeal in only limited circumstances. While arbitration proceedings may be subject to the rules of organizations such as the American Arbitration Association or the International Council for Commercial Arbitration, to name just a few, arbitration proceedings are conducted outside of the oversight or supervision of the courts.

Until statutes and regulations are enacted to address transactions and conduct within the NFT metaverse, traditional application of law will be the primary means to try to tame the Wild West. Even applying the common law principles of contract law provides a limited safety net for persons and entities within the metaverse given the prevalence of mandatory arbitration and class action waivers and the judicial enforcement of such provisions. While common law principles of torts such as fraud and misrepresentation may provide a means for parties to avoid mandatory arbitration and pursue their claims in a court of law, the broad scope of many mandatory arbitration provisions applying to "any and all claims related to a transaction" will encompass even tort claims. Moreover, courts do not countenance arguments predicated upon a party's failure to know of the existence of contract term when the party could have learned of the term by reading the contract. A person signing an agreement has a duty to read it and may not avoid the consequences of the agreement. By claiming he or she did not know its contents. *Chochorowski v. Home Depot U.S.A*. 440 S.W. 3D 220 (Mo.

banc 2013). Neither the length nor complexity of the terms of service of an NFT platform will exempt a party from its legal obligation to read the terms of the contract—no matter what may have been said or promised in prior discussions, brochures, or negotiations. As a purchaser of real estate in the metaverse, you will be held to have agreed to the terms expressly set forth in the written agreement—whether you read them or not. Another common argument presented by those parties seeking to avoid mandatory arbitration or participate in class action litigation is that such terms were non-negotiable given their lack of bargaining power and, therefore, the mandatory provisions should be void as a contract of adhesion. Missouri has codified the principle that agreements to arbitrate obtained through a contract of adhesion are invalid (Section 435.020 RSMo). In states not codifying the invalidity of arbitration agreements resulting from contracts of adhesion, the doctrine of adhesion does not automatically invalidate an arbitration agreement but is a factor in determining if a contract is so unconscionable that it will not be enforced. A contract of adhesion is manifested by a form contract that is created and imposed by the stronger party of the relationship and presented on a "take it or leave it" proposition. But evidence that parties did not negotiate contract terms is not sufficient proof that the contract's terms were not negotiable. *State ex rel. Vincent v. Schneider*, 194 S. W. 3D 853, 857-858) (Mo banc 2006). Claims that mandatory arbitration provisions should be voided as contracts of adhesion have had limited success in the courts when the party seeking to avoid arbitration is a sophisticated party, The fundamental principle of freedom of contract prevails in arm's length transactions between sophisticated parties. Agreements negotiated by sophisticated parties are generally enforced according to the terms of the agreement. Absent any countervailing public policy concerns, there is no reason to relieve the parties of the consequences of their bargain. *159 MP Corp. v. Redbridge Bedford LLC*, 33 N. Y. 3D 353, 128 N. E. 3D 128 (2019). A reasonable argument can be made that parties doing business in the metaverse are not average unsophisticated consumers and will have significant challenges in voiding a mandatory arbitration or class action waiver clause.

4. Conclusion

The volatility in the market can be evinced by recent events. NFTs sales witnessed a dramatic downturn at the outset of October 2022. Reuters reported a 60% drop in the third quarter from the second [64]. Directly after the report, and following American Express, Visa (V) filed two trademark applications for digital wallets and non-fungible tokens to operate within the metaverse. The application includes a management system for digital transactions and the use of a digital currency wallet and storage service. Additionally, using blockchain technology, Visa will also allow consumers to purchase "non-downloadable virtual goods" and collectible NFTs in a virtual environment [65]. Two weeks later, Apple banned NFT functionality on all iOS devices, including iPhone and iPad in order to avoid continued revenue losses [66]. Mere days later on November 11, 2022, the cryptocurrency company FTX, who had partnered with Alameda Research, filed for Chapter 11. The collapse led to calls for more regulation in crypto exchange and illustrates the pendulum that continues to swing between centralization and decentralization in the metaverse [67]. If past examples are heeded as cautionary tales, as this paper argues, the technology industry and regulators need to consider these inevitable scenarios from the outset. Additionally, the legal precedents, along with those emerging, must be considered when determining the best regulatory course. A clear legal understanding of the regulatory undergirding of the metaverse will be crucial. Technology alone will not pave the way for true ownership of digital assets in the metaverse. NFTs cannot bypass the centralized control that metaverse platforms currently have and will continue to have under their contractual terms of service. These terms of service themselves present a number of issues as the courts better define how ownership in a metaverse will work and be enforced. Future research should include a consideration of the impact of blockchain and contractual issues to regulate terms of service as in other industries, such as communication. Looking ahead, the metaverse is inevitable but the question remains whether it will be decentralized or centralized within existing corporation control. In the end, technological innovation must be accompanied by legal reform in order to ensure a free, open, and interoperable metaverse can exist.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Quiroz-Guitierrez, M. (2022) This Week in the Metaverse: ‘Metaverse’ Could Be Word of the Year, Johnnie Walker Launches an NFT, and Mattel Creates an NFT Marketplace. Fortune Crypto. https://fortune.com/crypto/2022/11/25/this-week-in-the-metaverse-johnnie-walker-oxford-rockstar-games/

[2] Ball, M. (2022) The Metaverse and How It Will Revolutionize Everything. Liveright Publishing, W.W. Norton & Company, New York.

[3] Valaskova, K., Machova, V. and Lewis, E. (2022) Virtual Marketplace Dynamics Data, Spatial Analytics, and Customer Engagement Tools in a Real-Time Interoperable Decentralized Metaverse. *Linguistic and Philosophical Investigations*, **21**, 105-120. https://doi.org/10.22381/lpi2120227

[4] Hollensen, S., Kotler, P. and Opresnik, M.O. (2022) Metaverse—The New Marketing Universe. *Journal of Business Strategy*. https://doi.org/10.1108/JBS-01-2022-0014

[5] Berg, C., Davidson, S. and Potts, J. (2019) Blockchain Technology as Economic Infrastructure: Revisiting the Electronic Markets Hypothesis. *Frontiers in Blockchain*, **2**, Article 22. https://doi.org/10.3389/fbloc.2019.00022

[6] Gadekallu, T.R., Pham, Q.-V., Nguyen, D.C., Maddikunta, P.K.R., Deepa, N., Prabadevi, B., Pathirana, P.N., Zhao, J. and Hwang, W.-J. (2022) Blockchain for Edge of Things: Applications, Opportunities, and Challenges. *IEEE Internet of Things*

Journal, **9**, 964-988. https://doi.org/10.1109/JIOT.2021.3119639

[7] Yang, Q., Zhao, Y., Huang, H., Xiong, Z., Kang, J. and Zheng, Z. (2022) Fusing Blockchain and AI with Metaverse: A Survey. *IEEE Open Journal of the Computer Society*, **3**, 122-136. https://doi.org/10.1109/OJCS.2022.3188249

[8] Avital, M., King, J.L., Beck, R., Rossi, M. and Teigland, R. (2016) Jumping on the Blockchain Bandwagon: Lessons of the Past and Outlook to the Future. *Proceedings of the* 37*th International Conference on Information Systems*, Dublin, 11-14 December 2016, 1-6.

[9] Rückeshäuser, N. (2017) Do We Really Want Blockchain-Based Accounting? Decentralized Consensus as Enabler of Management Override of Internal Controls. *Proceedings of the* 13*th International Conference on Wirtschaftsinformatik*, St. Gallen, 12-15 February 2017, 16-30.

[10] Weber, I., Xu, X., Riveret, R., Governatori, G., Ponomarev, A. and Mendling, J. (2016) Untrusted Business Process Monitoring and Execution. In: La Rosa, M., Loos, P. and Pastor, O., Eds., *Business Process Management*. *BPM* 2016. *Lecture Notes in Computer Science*, Vol. 9850, Springer, Cham, 329-347. https://doi.org/10.1007/978-3-319-45348-4_19

[11] Nakamoto, S. (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. Decentralized Business Review, 21260.

[12] Alt, R. and Puschmann, T. (2016) Bankmodell. In: *Digitalisierung der Finanzindustrie*, Springer Gabler, Berlin, 55-117. https://doi.org/10.1007/978-3-662-50542-7_2

[13] Möser, M., Böhme, R. and Breuker, D. (2013) An Inquiry into Money Laundering Tools in the Bitcoin Ecosystem. *Proceedings of* 2013 *APWG eCrime Researchers Summit*, San Francisco, 17-18 September 2013, 1-14. https://doi.org/10.1109/eCRS.2013.6805780

[14] Foley, S., Karlsen, J. and Putniņš, T. (2018) Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies? *Review of Financial Studies*, *Forthcoming*. https://doi.org/10.2139/ssrn.3102645

[15] Fanusie, Y. and Robinson, T. (2018) Bitcoin Laundering: An Analysis of Illicit Flows into Digital Currency Services. Center on Sanctions & Illicit Finance, Washington DC.

[16] Carstens, A. (2018) Money in the Digital Age: What Role for Central Banks? Lecture at the House of Finance, Goethe University, Frankfurt.

[17] Landau, J.-P. and Genais, A. (2018) Les Crypto-Monnaies. Rapport au Ministre de l'Économie et des Finances. Ministre de l'Économie et des Finances, Paris.

[18] Auer, R. and Claessens, S. (2018) Regulating Cryptocurrencies: Assessing Market Reactions. BIS Quarterly Review, September, 51-65.

[19] PROFILEmiami (2022) Take a First Look at 'Metareal' Mansion to Be Sold via NFT as Real Home and Digital Twin in the Metaverse. https://profilemiamire.com/miamirealestate/2022/8/11/take-a-first-look-at-metareal-mansion-to-be-sold-via-nft-as-real-home-and-digital-twin-in-metaverse

[20] Marinotti, J. (2022) Can You Truly Own Anything in the Metaverse? A Law Professor Explains How Blockchains and NFTs Don't Protect Virtual Property. Market Watch. https://www.marketwatch.com/story/can-you-truly-own-anything-in-the-metaverse-a-law-professor-explains-how-blockchains-and-nfts-dont-protect-virtual-property-11650559724#:~:text=When%20you%20buy%20an%20item,a%20unique%20string%20of%20bits

[21] Ravenscraft, E. (2021) The Metaverse Land Rush Is an Illusion. Wired. https://www.wired.com/story/metaverse-land-rush-illusion/

[22] Tidy, J. (2022) Billions Being Spent in Metaverse Land Grab. BBC News. https://www.bbc.com/news/technology-63488059

[23] Auer, R. (2019) Embedded Supervision: How to Build Regulation into Blockchain Finance. Federal Reserve Bank of Dallas, Dallas. https://doi.org/10.24149/gwp371

[24] Elm, M. and Montez, T. (2022) The Banking Head of Digital's Five-Year Innovation Strategy: A Playbook for Open Finance, AI, the Metaverse and beyond. Insider Intelligence eMarketer, New York.

[25] Chao, C.-H., Ting, I.-H., Tseng, Y.-J., Wang, B.-W., Wang, S.-H. and Wang, Y.-Q. (2022) The Study of DeFi in Virtual Community and Financial Supervision. *Proceedings of the 9th Multidisciplinary International Social Networks Conference*, Matsuyama, 29-31 October 2022, 74-80. https://doi.org/10.1145/3561278.3561297

[26] Turdialiev, M.A. (2022) The Legal Issues of the Metaverse and Perspectives of Establishment of International Financial Center in Metaverse. *Oriental Renaissance: Innovative, Educational, Natural and Social Sciences*, **2**, 239-249.

[27] Bourlakis, M., Papagiannidis, S. and Li, F. (2009) Retail Spatial Evolution: Paving the Way from Traditional to Metaverse Retailing. *Electronic Commerce Research*, **9**, 135-148. https://doi.org/10.1007/s10660-009-9030-8

[28] Hackl, C. (2020) The Metaverse Is Coming and It's a Very Big Deal. Forbes. https://www.forbes.com/sites/cathyhackl/2020/07/05/the-metaverse-is-coming--its-a-very-big-deal/?sh=1f4c347a440f

[29] Rospigliosi, P.A. (2022) Metaverse or Simulacra? Roblox, Minecraft, Meta and the Turn to Virtual Reality for Education, Socialisation and Work. *Interactive Learning Environments*, **30**, 1-3. https://doi.org/10.1080/10494820.2022.2022899

[30] Strobel, J. and Hawkins, C. (2022) Designing in Second Life: Identity Construction and Learning in a Virtual Informal Environment. *Journal of Online Engineering Education*, **1**, Article 2.

[31] Fernandes, A. and Chatterjee, S. (2022) Possibilities of Metaverse: The Second Life. *International Journal of Engineering and Management Research*, **12**, 79-82.

[32] Harris, K., Harris, R. and Baron, S. (2001) Customer Participation in Retail Service: Lessons from Brecht. *International Journal of Retail and Distribution Management*, **29**, 359-369. https://doi.org/10.1108/09590550110396845

[33] Papagiannidis, S. and Bourlakis, M.A. (2010) Staging the New Retail Drama: At a Metaverse near You! *Journal of Virtual Worlds Research*, **2**, 425-446. https://doi.org/10.4101/jvwr.v2i5.808

[34] Haig, M. (2001) Talking to the Teen Generation. In: *Brand Strategy*, Henry Stewart Publications, London, 30.

[35] Ahn, S.J., Kim, J. and Kim, J. (2022) The Bifold Triadic Relationships Framework: A Theoretical Primer for Advertising Research in the Metaverse. *Journal of Advertising*, **51**, 592-607. https://doi.org/10.1080/00913367.2022.2111729

[36] Popescu, G.H., Valaskova, K. and Horak, J. (2022) Augmented Reality Shopping Experiences, Retail Business Analytics, and Machine Vision Algorithms in the Virtual Economy of the Metaverse. *Journal of Self-Governance & Management Economics*, **10**, 67-81. https://doi.org/10.22381/jsme10220225

[37] Jha, S.K. and Kumar, S.S. (2022) Cybersecurity in the Age of the Internet of Things: An Assessment of the Users' Privacy and Data Security. In: Jeena Jacob, I., Gonzalez-Longatt, F.M., Kolandapalayam Shanmugam, S. and Izonin, I., Eds., *Expert*

Clouds and Applications. *Lecture Notes in Networks and Systems*, Vol. 209, Springer, Singapore, 49-56. https://doi.org/10.1007/978-981-16-2126-0_5

[38] Dwivedi, Y.K., Hughes, L., Baabdullah, A.M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M.M., *et al*. (2022) Metaverse beyond the Hype: Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy. *International Journal of Information Management*, **66**, Article ID: 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

[39] Koos, S. (2021) Artificial Intelligence as Disruption Factor in the Civil Law: Impact of the Use of Artificial Intelligence in Liability, Contracting, Competition Law and Consumer Protection with Particular Reference to the German and Indonesian Legal Situation. *Yuridika*, **36**, 235-262. https://doi.org/10.20473/ydk.v36i1.24033

[40] Greenstein, S. (2022) Preserving the Rule of Law in the Era of Artificial Intelligence (AI). *Artificial Intelligence and Law*, **30**, 291-323. https://doi.org/10.1007/s10506-021-09294-4

[41] Belk, R., Humayun, M. and Brouard, M. (2022) Money, Possessions, and Ownership in the Metaverse: NFTs, Cryptocurrencies, Web3 and Wild Markets. *Journal of Business Research*, **153**, 198-205. https://doi.org/10.1016/j.jbusres.2022.08.031

[42] Parloff, R. (2006) Anshe Chung: First virtual Millionaire. Fortune. https://fortune.com/2006/11/27/anshe-chung-first-virtual-millionaire/

[43] Johnson, M.L. (2022) Where in the World Are These Worlds? In: *Social Virtual Worlds and Their Places*, Palgrave Macmillan, Singapore, 55-123. https://doi.org/10.1007/978-981-16-8626-9_2

[44] Demchenko, O. (2022) Electronic Commerce in the Gaming Industry. The Applicability of Consumer Protection Framework and Legal Challenges. University of Pécs, Pécs.

[45] Kılınç, Ü. (2022) Metaverse üzerinden arsa satılanplatformlar. https://www.webtekno.com/metaverse-evreninden-arsa-satin-alabileceginiz-platformlar-h119400.html

[46] Eryılmaz, H. (2021) Metaverse Arsa Alma Gerçek Oluyor—Bloktopia Arsa Satışlarına Başlıyor. Bitcoin Yorum. https://bitcoinyorum.com/2021/10/14/metaverse-arsa-alma-gercek-oluyor-bloktopia-arsa-satislarina-basliyor/

[47] Türk, T. (2022) The Concept of Metaverse, Its Future and Its Relationship with Spatial Information. *Advanced Geomatics*, **2**, 17-22.

[48] Christodoulou, K., Katelaris, L., Themistocleous, M., Christodoulou, P. and Iosif, E. (2022) NFTs and the Metaverse Revolution: Research Perspectives and Open Challenges. In: Lacity, M.C. and Treiblmaier, H., Eds., *Blockchains and the Token Economy*. *Technology*, *Work and Globalization*, Palgrave Macmillan, Cham, 139-178. https://doi.org/10.1007/978-3-030-95108-5_6

[49] Au, A. (2022) Cryptocurrencies and the Promise of Individual Economic Sovereignty in an Age of Digitalization: A Critical Appraisal. *The Sociological Review*. https://doi.org/10.1177/00380261221127858

[50] Obar, J. and Oeldorf-Hirsch, A. (2020) The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services. *Information*, *Communication & Society*, **23**, 128-147. https://doi.org/10.1080/1369118X.2018.1486870

[51] Marinotti, J. (2020), Tangibility as Technology. 37(3) Georgia State University Law Review 671 (2021), Indiana Legal Studies Research Paper No. 445, 671-738. https://doi.org/10.2139/ssrn.3682090

[52] Rustad, M.L. and Onufrio, M.V. (2011) Reconceptualizing Consumer Terms of Use for a Globalized Knowledge Economy. *University of Pennsylvania Journal of Business Law*, **14**, 1085-1190.

[53] Osmani, M., El-Haddadeh, R., Hindi, N., Janssen, M. and Weerakkody, V. (2020) Blockchain for Next Generation Services in Banking and Finance: Cost, Benefit, Risk and Opportunity Analysis. *Journal of Enterprise Information Management*, **34**, 884-899. https://doi.org/10.1108/JEIM-02-2020-0044

[54] Demirgüç-Kunt, A., Klapper, L., Singer, D. and Ansar, S. (2022) The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19. World Bank Publications, Washington DC. https://doi.org/10.1596/978-1-4648-1897-4

[55] Agur, I., Ari, A. and Dell'Ariccia, G. (2022) Designing Central Bank Digital Currencies. *Journal of Monetary Economics*, **125**, 62-79. https://doi.org/10.1016/j.jmoneco.2021.05.002

[56] Kshetri, N. (2021) Blockchain Technology for Improving Transparency and Citizen's Trust. In: Arai, K., Ed., *Advances in Information and Communication*. *FICC* 2021. *Advances in Intelligent Systems and Computing*, Vol. 1363, Springer, Cham, 716-735. https://doi.org/10.1007/978-3-030-73100-7_52

[57] Çubuk, E.B.S., Zeren, H.E. and Demirdöven, B. (2023) The Role of Data Governance in Cybersecurity for E-Municipal Services: Implications from the Case of Turkey. In: Saeed, S., Almuhaideb, A.M., Zaman, N. and Zikria, Y.D., Eds., *Handbook of Research on Cybersecurity Issues and Challenges for Business and FinTech Applications*, IGI Global, Pennsylvania, 410-425. https://doi.org/10.4018/978-1-6684-5284-4.ch020

[58] Wiederhold, B.K. and Riva, G. (2021) Virtual Reality Therapy in the Metaverse: Merging VR for the Outside with VR for the Inside. *Annual Review of CyberTherapy and Telemedicine*, **19**, 3-9. https://vrphobia.com/wp-content/uploads/2022/06/ARCTT_2021.pdf#page=23

[59] Tayal, S., Rajagopal, K. and Mahajan, V. (2022) Virtual Reality Based Metaverse of Gamification. 2022 6*th International Conference on Computing Methodologies and Communication* (*ICCMC*), Erode, 29-31 March 2022, 1597-1604. https://doi.org/10.1109/ICCMC53470.2022.9753727

[60] Kshetri, N., Bhusal, C.S. and Chapagain, D. (2021) BCT-AA: A Survey of Blockchain Technology-Based Applications in Context with Agribusiness. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3834004

[61] Sabry, S.S., Kaittan, N.M. and Majeed, I. (2019) The Road to the Blockchain Technology: Concept and Types. *Periodicals of Engineering and Natural Sciences*, **7**, 1821-1832. https://doi.org/10.21533/pen.v7i4.935

[62] Kshetri, N. (2022) The Global Rise of Online Devices, Cyber Crime and Cyber Defense: Enhancing Ethical Actions, Counter Measures, Cyber Strategy, and Approaches. University of Missouri, St. Louis.

[63] Moringiello, J.M. and Odinet, C.K. (2022) The Property Law of Tokens. 74 Florida Law Review 607, U Iowa Legal Studies Research Paper No. 2021-44, Widener Law Commonwealth Research Paper, 607-671. https://doi.org/10.2139/ssrn.3928901

[64] Howcroft, E. (2022) NFT Sales Plunge in Q3, down by 60% from Q2. Reuters. https://www.reuters.com/technology/nft-sales-plunge-q3-down-by-60-q2-2022-10-03/

[65] Bellusci, M. (2022) Visa Files Trademark Applications for Crypto Wallets, NFTs and the Metaverse. CoinDesk.

https://www.coindesk.com/business/2022/10/28/visa-files-trademark-applications-for-crypto-wallets-nfts-and-the-metaverse/

[66] Nelson, D. (2022) Apple Refuses to Exempt NFTs from App Store's 30% Fee. CoinDesk. https://www.coindesk.com/business/2022/10/24/apple-refuses-to-exempt-nfts-from-app-stores-30-fee/

[67] Allen, M. (2022) Can Switzerland Benefit from the FTX Crypto Crash? SWI. https://www.swissinfo.ch/eng/business/can-switzerland-benefit-from-the-ftx-crypto-crash-/48068900

Journal of Intelligent Learning Systems and Applications, 2023, 15, 24-35
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

A Proposed Meta-Reality Immersive Development Pipeline: Generative AI Models and Extended Reality (XR) Content for the Metaverse

Jeremiah Ratican[1], **James Hutson**[2], **Andrew Wright**[3]

[1]Department of Art, Production and Media, College of Arts and Humanities, Lindenwood University, St. Charles, USA
[2]Department of Art History and Visual Culture, College of Arts and Humanities, Lindenwood University, St. Charles, USA
[3]iTeacher, EDU Metaverse, Sydney, Australia

Email: Jratican@lindenwood.edu, JHutson@lindenwood.edu, info@iteacher.net.au

How to cite this paper: Ratican, J., Hutson, J. and Wright, A. (2023) A Proposed Meta-Reality Immersive Development Pipeline: Generative AI Models and Extended Reality (XR) Content for the Metaverse. *Journal of Intelligent Learning Systems and Applications*, **15**, 24-35.
https://doi.org/10.4236/jilsa.2023.151002

Received: December 7, 2022
Accepted: February 4, 2023
Published: February 7, 2023

Abstract

The realization of an interoperable and scalable virtual platform, currently known as the "metaverse," is inevitable, but many technological challenges need to be overcome first. With the metaverse still in a nascent phase, research currently indicates that building a new 3D social environment capable of interoperable avatars and digital transactions will represent most of the initial investment in time and capital. The return on investment, however, is worth the financial risk for firms like Meta, Google, and Apple. While the current virtual space of the metaverse is worth $6.30 billion, that is expected to grow to $84.09 billion by the end of 2028. But the creation of an entire alternate virtual universe of 3D avatars, objects, and otherworldly cityscapes calls for a new development pipeline and workflow. Existing 3D modeling and digital twin processes, already well-established in industry and gaming, will be ported to support the need to architect and furnish this new digital world. The current development pipeline, however, is cumbersome, expensive and limited in output capacity. This paper proposes a new and innovative immersive development pipeline leveraging the recent advances in artificial intelligence (AI) for 3D model creation and optimization. The previous reliance on 3D modeling software to create assets and then import into a game engine can be replaced with nearly instantaneous content creation with AI. While AI art generators like DALL-E 2 and DeepAI have been used for 2D asset creation, when combined with game engine technology, such as Unreal Engine 5 and virtualized geometry systems like Nanite, a new process for creating nearly unlimited content for immersive reality is possible. New processes and workflows, such as those proposed here, will revolutionize content creation and pave the way for Web 3.0, the metaverse and a truly 3D social environment.

Keywords

AI Content Generator, Metaverse, Development Pipeline, AI Art Generator, 3D Asset Creation, Unreal Engine 5, Nanite

1. Introduction

The use of artificial intelligence (AI) to generate content for art and design is not new [1]. The infancy of art history computing can be traced to the 1980s as artists experimented with the potential of "digitized" and "digital" iterations of algorithmic art [2]. As art and design have historically adopted emerging technologies, the rapid spread of AI art generators available today was inevitable [3]. However, such generative content produced using text prompts with DALLE-2, Midjourney, Jasper Art, Stable Diffusion, DeepAI, and many more tools, has been largely limited to two-dimensional output and has yet to disrupt the 3D modeling and digital twin development pipelines [4]. These pipelines are currently cumbersome and expensive, requiring specialized technical knowledge that often requires years of training in industry [5]. In order to create content for extended reality (XR), such as augmented reality (AR), mixed reality (MR), and virtual reality (VR), 360 photography, photogrammetry, or 3D modeling software are used. Whether producing a digital twin using the Matterport system or creating a 3D model using Autodesk 3Ds Max, Maya or Blender, there are many limitations. For instance, 3D cameras like the Matterport MC250 Pro2 are able to improve upon previous photogrammetric processes using 134 megapixels and 100k points per second and 1.5 million per scan, and then process for immersive viewing. The digital twin, however, is bound to proprietary software and is not interoperable [6]. The 3D models created are also not optimized and are not created for rendering and are often incomplete. On the other hand, models that are created using 3D modeling software, while optimized for rendering, are cumbersome to create, though are device agnostic and can be exported to file formats that are interoperable, geometry definition file formats such as OBJ. Modelers often save time and search for existing models in asset stores but are limited by what already exists that can be altered and reconfigured to the desired specifications. After completing modeling and alterations, XR content creators can import the compatible files into platforms such as Virbela's FRAME or Spatial [7]. Both development pipelines have limitations that can be resolved through a combination of emerging technologies.

As such this paper proposes a new meta-reality immersive development pipeline to address the current limitations of content creation for virtual and immersive environments in the metaverse. With advances in natural language processing

(NLP) and visual content generation AI, text prompts can be used to generate 3D models in interoperable formats that can support animation (e.g. OBJ, GLTF, GLB, etc.) [8]. Much like image generation with DALLE-2 and Stable Diffusion, 3D content generators like Interactive Pattern Generator and AI Material Designer allow for the creation of tileable (modular or repeatable) materials for assets through text prompts [9]. Using the generators AI can be trained to generate 3D elements such as patterns and texture. Whereas previous limitations included file size to ensure low latency in XR experiences, virtualized geometry systems can now be used to compress files for real-time rendering [10]. After generating content using AI, game engines can be used to edit the file and systems like Nanite to optimize for efficient run times. The proposed pipeline will not only remove the need for specialist technical knowledge and training but allows for unprecedented asset creation. AI-generated art is becoming increasingly common and accepted. XR designers and developers are producing virtual exhibitions to illustrate AI-generated art, such as Andrew Wright's *AI Art Exhibition*: *The Other Us* (2022) (https://framevr.io/theotherus) (**Figure 1** and **Figure 2**). The next logical step is to combine the technologies to generate XR content for an immersive space. The new process and workflow proposed here overcome existing limitations for 3D content creation and paves the way for Web 3.0, the metaverse and a truly 3D social environment.

Figure 1. Andrew Wright, *AI Art Exhibition*: *The Other Us.* FRAME. Virbela. (2022) (Detail 1).

Figure 2. Andrew Wright, *AI Art Exhibition*: *The Other Us.* FRAME. Virbela. (2022) (Detail 2).

2. Literature Review

First developed in the 1960s, 3D modeling involves the creation of a three-dimensional digital visual representation of an object using computer software. There are many varieties of 3D models, including wireframe, surface, and solid, the most computationally demanding. Each is suited to a particular task of capturing information about a three-dimensional object, such as design, size, appearance, texture, and even weight, density and gravity. The use of such models began in industry and industrial design and expanded in the late 1990s to video games and entertainment [11] [12]. The demand for 3D asset creation has only grown with the advent of the metaverse and the consumption of media, games, and entertainment in immersive environments [13]. 3D modelers often use software like Sketchfab, Blender, Maya, and Autodesk 3ds Max to create an asset, which can be viewed in these applications. Otherwise, for greater optimization and interactivity, assets are also imported into game engines, such as Unreal Engine 5 (Epic Games) and Unity (Unity Technologies) and situated within virtual environments [14]. These game engines even have marketplaces where 3D assets can be readily downloaded and sold, allowing developers without 3D modeling experience access to 3D models. However, as noted, the number of resources in these stores is finite and developers are limited by existing assets that need be altered and reconfigured to the desired specifications. The demand for such high-quality, editable and reconfigurable assets will only continue to rise in many industries, especially immersive content creation for the metaverse [15].

2.1. Game Engines and Cinematics

Industries such as film have seen an increase in the use of 3D models and game engines as details are now crossing the uncanny valley and becoming indistinguishable from real life, leading to advances in experimental filmmaking, including VR cinematics [16]. The mainstream film industry has also seen a rise in the use of traditional game development software like Unreal Engine 5, which has become a tool used by visual effects artists and filmmakers to create realistic worlds in real-time. The American space western television series *The Mandalorian* [17] was a front runner in standardizing the use of game engine technology as it provided filmmakers the ability to create realistic virtual sets that would dynamically change based on needs and camera position, and also reflect accurate lighting information [18]. The development pipeline saved countless hours in post-processing as much of the effects were done in camera and on set [19]. The process also has been praised by the actors as they can see the world they are acting in as opposed to the previous method of working in front of a green screen [20]. The need for 3D models and virtual sets will only rise as more and more film projects are relying on this technology. As more films are also shot in virtual reality (VR), the impact of these technologies has the potential to radically change how directing in the film industry operates [21] [22].

2.2. 3D Modeling Process

With all of the advances made possible using game-engine technology, 3D modeling remains limited to specialists. The development pipeline for 3D modeling explains why [23]. In traditional 3D modeling, the modeler starts with simple geometry, such as a polygon, which can be as simple as a triangle comprised of three vertices existing in three-dimensional space represented in the Cartesian coordinate standard of X, Y, Z. A minimum of three vertices are required to generate the surface geometry, and today's game models can easily be comprised of 20,000 polygons, which would be the equivalent to 40,000 triangles and those are models that have been optimized for performance. Because of these considerations, the modeling pipeline is heavily reliant on several specialized skills and techniques, such as re-topologizing meshes to improve animation and performance, rendering tools such as normal maps to produce fine details, as well as post-processing effects using specialized shaders [24]. The requirements can lead to the need to be able to process and render thousands, if not millions of polygons on screen at 60 frames per second. The following will investigate the two variables of the asset creation pipeline: the creation of models and the optimization and usability of those models in a game engine or other real-time applications.

2.3. 3D Scanning and Photogrammetry

The acquisition or creation of 3D models has seen a few developments in the past several years, the most well-known of which is 3D scanning. This technology is not new but continues to improve in both quality and adaptation, as well as accessibility. Whereas previous iterations were large and cumbersome, the latest generation of scanners are handheld (e.g. Artec 3D) or even free applications for smartphones that use LiDAR (e.g. Scaniverse and Polycam) [25]. These scanners and their associated software applications allow users to create 3D models of objects large and small and can even scan entire areas [26]. The technology is used in many industries beyond entertainment including engineering and even law enforcement (Chenoweth *et al.* 2022). Alternatively, another related method of model creation is photogrammetry, which is the process of generating 3D models from photographs or other data. This process is also not new and began with the creation of 2D information, not only from photographs but also from sonar and radar and has been used to create topographic maps [27]. These processes have expanded to other fields including entertainment and these techniques have been used in films such as *The Matrix* [28] and video games [29]. Taken together, both 3D scanning and photogrammetry can produce 3D models and assets.

2.4. AI-Generated Content

Technology continues to develop, and AI has been used with photogrammetry to improve models and fill in details that were not present in photographs

(Amaro *et al.* 2022). However, recently AI has begun to cross over from being a tool for helping with art creation to a method for generating art. AI art generators like the DALL-E 2 have recently made headlines by creating interesting and imaginative works of art [30]. While these early models often would make mistakes that seem obvious to human eyes, with each iteration the AI improves, and newer models have started to make photo-realistic art [30]. These examples are for creating 2D art but can clearly show an evolution of quality and sophistication with each newer algorithmic iteration. An exciting aspect about using AI is its ability to improve and learn from previous versions. Artists are using these generated images to inspire them and as starting points for conceptualizing ideas and designs, and many speculate it will not be long before AI can do the largest part of concept design work [3].

The move from 2D art generation to 3D content generation is a natural progression. This has led to several AI systems that can take a prompt, some of which can be as simple as a text description and transform that into a 3D model. Such SOTA models may be used to train AI to understand 3D space using image language models [31]. The open-world 3D scene understanding task is a 3D vision-language task that also includes open-set classification. The limitations of the tasks are that the AI does not currently have enough data. Unfortunately, existing 3D datasets are not varied enough in comparison to 2D counterparts to train AI to generate content [32]. Admittedly, the creation of a 3D model from a text prompt is still early in development and is not ideal for asset creation, but other solutions do currently exist.

Just as photometry uses data to generate pictures, some AI systems are using 2D pictures as inputs to generate 3D content. By loading images as reference, AI systems like Nvidia's Instant NeRF and Kaedim can generate 3D models. Kaedim is a newer image to 3D model AI tool aimed at 3D artists and those that need 3D models created from concept design. The tool is still in development and currently needs human reviewers to ensure quality of output. The software reviews images of a concept design from all angles and creates a 3D model. Kaedim is one of the few AI 3D model generating tools that takes the technical requirements of the models in mind but does require the user to specify the complexity of the model [33]. This process does require the user to be aware of the specific requirements of the platform or real-time application the model is developed for and experience to know how many polygons would be appropriate [34]. Nvidia's NeRF (Neural Radiance Field) uses a process based on a concept called inverse rendering, which essentially inverts the concept of normal rendering and attempts to recreate how light reacts to objects in the real world. Instead of normal baked lighting, NeRF instead uses AI to analyze a collection of 2D images and constructs a 3D model from them. The process can create a full 3D scene in a short amount of time [35]. These methods of 3D model creation are becoming increasingly easy to use, and with accessibility becoming as simple as an application accessed on a smartphone, the hurdle of creating a 3D model has all but been removed.

2.5. 3D Model Optimization

The other variable to consider in the proposed immersive development pipeline is the optimization of 3D models for use in real-time applications. While the methods outlined above can create models often with a high number of details, the models themselves are not in a usable format consisting of dense meshes that would lead to poor performance and be unsuitable for animation [36]. This is an issue even current processes face, as 3D modeling for entertainment often uses a technique called digital sculpture [37]. Many of the most popular modeling applications use these techniques to spectacular effect; the most popular software application for digital sculpting is ZBrush. The process where an artist uses digital sculpture to create a model can also result in a dense unusable mesh that must be re-topologized. Retopology can be a time-consuming process where a lower resolution and optimized version of a model is made, the version that would work well in a game engine, for instance, and then the high-resolution details are added back in during rendering [38]. The process usually includes baking the higher-resolution details into color images where the surface of the high-resolution mesh is represented in a texture where the three channels of red, green, and blue are controlling the XYZ information of how light reacts to the surface of the model. These textures are called normal maps and have been standard practice since the early 2000s [39].

The process involved with retopology can be rather involved and lengthy, therefore, software developers have been working on ways to make it easier, such as including auto-retopology tools to popular 3D software applications [40]. One of the foremost pioneers of graphics technology in this area is the graphics card manufacturer Nvidia. The AI computing company has been sponsoring and developing new graphics technologies for decades and has an annual technology conference where new graphics technologies are showcased. As so much of the processing and rendering of immersive realities and real-time applications rely on the hardware, Nvidia is also involved in improving the performance of those functions [41]. Two of Nvidia's recent innovations that are relevant for this study are the Nvidia NGX and aforementioned Nvidia NeRF. Both technologies approach rendering in new and dynamic ways. The Nvidia NGX requires an Nvidia RTX video card and uses that hardware combined with AI and deep learning to improve performance and improve the graphic output. The NGX can load an entire 3D scene from an previous design iteration and create a modern lighting and rendering solution. The process can transform older, lower resolution graphics and ensure that they appear crisp and clear on modern systems and resolutions [42]. While this process can make older graphics appear more modern it is, however, focused on the textural graphics and not the resolution of the models themselves making the solution ideal for lower resolution models. Higher resolution models would still require optimized geometry like that which is done through proper retopology [43].

In order to address the retopology issue, Epic Games's engineers working on

Unreal Engine 5 have used a new approach of virtualized geometry systems that aims to render retopology unnecessary all together. The Nanite system, which is built on earlier iterations and research on how to process and render high-resolution geometry in real-time, allows the user to bring in non-optimized high-resolution 3D models directly from a 3D scan or highly detailed digital sculpt. These files are able to be altered in size to ensure fast processing and rendering [44] [45]. Nanite does this in several novel ways, including dynamically making level of detail (LOD) in real time, which consist of multiple copies at varied polygonal complexity that decrease in resolution as the viewer moves away from an object and increases in complexity as a viewer moves closer. The virtualized geometry system also automatically occludes polygons that are not visible from the viewer's perspective, making them non-rendered or non-processed. Adjusting the resolution of objects that are out-of-frame from a viewer ensures better performance for real-time applications since render engines often process geometry even when not seen [46]. Nanite also works to improve rendering by using a system of virtualized textures also automatically generated [47]. All of these optimizations are done by the system, removing the technical obstacles that could impact performance and allowing the developer to focus on viewer experience. This system, and others that are sure to follow, entirely removes many of the most time-consuming and technically challenging aspects of 3D asset creation.

3. Recommendations

While both 3D AI asset generators and virtualized geometry systems are currently in use in the market, combining them to create a new development pipeline of 3D asset creation has not been explored. A theoretical framework for using these technologies in tandem is proposed and is an alternative method for designing, creating and producing content for immersive environments in XR. Such an immersive development pipeline would involve generating 2D concept designs via an AI art generator, such as DALLE-2 or Stable Diffusion. These designs can then be rendered as 3D models using software such as Nvidia's Instant NeRF and Kaedim. Finally, these 3D models can be imported into a game engine, such as Unreal Engine 5 where the virtualized geometry system Nanite can optimize for appropriate resolution to ensure low latency in a virtual environment such as when using a head-mounted display (HMD). As for implementing those models, Nanite is proving to be a novel solution on how to use non-optimized models and has taken much of the asset creation process and removed or automated it. Recommendations for next steps in research include implementing the proposed development workflow and pipeline.

4. Conclusion

Developments in AI generative content witnessed unprecedented strides in 2022 [48]. New technologies are opening alternative methods of 3D asset generation. While this study is the first to examine these technologies within the lens of 3D asset creation, the proposed development pipeline shows that AI 3D art generation programs continue to grow and be developed in a range of industries. As seen in Andrew Wright's *AI Art Exhibition*: *Yeti* (2022) (https://framevr.io/theotherus) (**Figure 1** and **Figure 2**), AI-generated content is now within the reach of the general public. Each image within the exhibition was generated using natural language prompts in an AI generator. At the moment, each image requires trial and error with thousands of text prompt permutations to arrive at the desired effect. But advances are being made rapidly, as evinced by the number of options of generators now freely available. The rise of AI-driven content, and the increased accessibility to formally specialized and technically challenging 3D designers and developers means massive disruption in the field is quickly approaching the horizon. As the combination of these technical achievements creates an alternate possible 3D asset creation pipeline wherein a developer could use commonplace technology, such as a mobile phone to scan in objects, or use an AI system to generate 3D content either from a 2D concept design, which could also have been generated by AI (Bouchard, 2022), or a simple text prompt [31]. These developments democratize the 3D design and modeling field and create more opportunities for users to make the models required without an experienced artist or designer, rendering current design and development pipelines obsolete.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Mondal, B. (2020) Artificial Intelligence: State of the Art. In: Balas, V.E., Kumar, R. and Srivastava, R., Eds., *Recent Trends and Advances in Artificial Intelligence and Internet of Things*, Springer, Berlin, 389-425. https://doi.org/10.1007/978-3-030-32644-9_32

[2] Zweig, B. (2015) Forgotten Genealogies: Brief Reflections on the History of Digital Art History. *International Journal for Digital Art History*, No. 1.

[3] Mikalonytė, E.S. and Kneer, M. (2022) Can Artificial Intelligence Make Art? Folk Intuitions as to whether AI-Driven Robots Can Be Viewed as Artists and Produce Art. *ACM Transactions on Human-Robot Interaction*, **11**, Article No. 43. https://doi.org/10.1145/3530875

[4] Li, S. and Zhou, Y. (2022) Pipeline 3D Modeling Based on High-Definition Rendering Intelligent Calculation. *Mathematical Problems in Engineering*, **2022**, Article ID: 4580363. https://doi.org/10.1155/2022/4580363

[5] Burova, A., Mäkelä, J., Heinonen, H., Palma, P.B., Hakulinen, J., Opas, V. and Turunen, M. (2022) Asynchronous Industrial Collaboration: How Virtual Reality and Virtual Tools Aid the Process of Maintenance Method Development and Documentation Creation. *Computers in Industry*, **140**, Article ID: 103663. https://doi.org/10.1016/j.compind.2022.103663

[6] Tchomdji, L.O.K., Park, S.J. and Kim, R. (2022) Developing Virtual Tour Content

for the Inside and Outside of a Building Using Drones and Matterport. *International Journal of Contents*, **18**, 74-84.

[7] Harrington, M.C., Jones, C. and Peters, C. (2022) Course on Virtual Nature as a Digital Twin: Botanically Correct 3D AR and VR Optimized Low-Polygon and Photogrammetry High-Polygon Plant Models. *ACM SIGGRAPH* 2022 *Courses*, Vancouver, 7-11 August 2022, 1-69. https://doi.org/10.1145/3532720.3535663

[8] Oussalah, M. (2021) AI Explainability. A Bridge between Machine Vision and Natural Language Processing. In: *International Conference on Pattern Recognition*, Springer, Cham, 257-273. https://doi.org/10.1007/978-3-030-68796-0_19

[9] Beer, R.D. (1995) A Dynamical Systems Perspective on Agent-Environment Interaction. *Artificial Intelligence*, **72**, 173-215. https://doi.org/10.1016/0004-3702(94)00005-L

[10] Salehi, M., Hooli, K., Hulkkonen, J. and Tolli, A. (2022) Enhancing Next-Generation Extended Reality Applications with Coded Caching.

[11] Schmidt, R., Isenberg, T., Jepp, P., Singh, K. and Wyvill, B. (2007) Sketching, Scaffolding, and Inking: A Visual History for interactive 3D Modeling. *Proceedings of the* 5*th International Symposium on Non-Photorealistic Animation and Rendering*, San Diego, 4-5 August 2007, 23-32. https://doi.org/10.1145/1274871.1274875

[12] Madhavadas, V., Srivastava, D., Chadha, U., Raj, S.A., Sultan, M.T.H., Shahar, F.S. and Shah, A.U.M. (2022) A Review on Metal Additive Manufacturing for Intricately Shaped Aerospace Components. *CIRP Journal of Manufacturing Science and Technology*, **39**, 18-36. https://doi.org/10.1016/j.cirpj.2022.07.005

[13] Cheng, R., Wu, N., Chen, S. and Han, B. (2022) Reality Check of Metaverse: A First Look at Commercial Social Virtual Reality Platforms. 2022 *IEEE Conference on Virtual Reality and* 3*D User Interfaces Abstracts and Workshops* (*VRW*), Christchurch, 12-16 March 2022, 141-148. https://doi.org/10.1109/VRW55335.2022.00040

[14] Medina, J.J., Maley, J.M., Sannapareddy, S., Medina, N.N., Gilman, C.M. and McCormack, J.E. (2020) A Rapid and Cost-Effective Pipeline for Digitization of Museum Specimens with 3D Photogrammetry. *PLOS ONE*, **15**, e0236417. https://doi.org/10.1371/journal.pone.0236417

[15] Korbel, J.J., Siddiq, U.H. and Zarnekow, R. (2022) Towards Virtual 3D Asset Price Prediction Based on Machine Learning. *Journal of Theoretical and Applied Electronic Commerce Research*, **17**, 924-948. https://doi.org/10.3390/jtaer17030048

[16] Zhang, Y. (2021, December 10) Cross-Platform Methods in Computer Graphics That Boost Experimental Film Making. Thesis, Institute of Technology, Rochester. https://scholarworks.rit.edu/cgi/viewcontent.cgi?article=12190&context=theses

[17] Favreau, J. (2019) The Mandalorian. [Film] Disney.

[18] Turnock, J.A. (2022) Conclusion. Unreal Engine. In: *The Empire of Effects*, University of Texas Press, Austin, 216-222. https://doi.org/10.7560/325308

[19] Anderson, K. (2020, August 24) The Mandalorian's Virtual Sets Are Wild. Nerdist. https://nerdist.com/article/mandalorian-virtual-production-star-wars

[20] Dent, S. (2021, May 13) See How "The Mandalorian" Used Unreal Engine for Its Real-Time Digital Sets. Engadget. https://www.engadget.com/2020-02-21-mandalorian-ilm-stagecraft-real-time-digital-sets.html

[21] Ronfard, R. (2021, May) Film Directing for Computer Games and Animation. *Computer Graphics Forum*, **40**, 713-730. https://doi.org/10.1111/cgf.142663

[22] Cohen, J.L. (2022) Film/Video-Based Therapy and Trauma: Research and Practice.

Taylor & Francis, Abingdon-on-Thames. https://doi.org/10.4324/9781315622507

[23] Kuusela, V. (2022) 3D Modeling Pipeline for Games. Turku University of Applied Sciences, Turku.

[24] Hevko, I.V., Potapchuk, O.I., Lutsyk, I.B., Yavorska, V.V., Hiltay, L.S. and Stoliar, O.B. (2022) The Method of Teaching Graphic 3D Reconstruction of Architectural Objects for Future IT Specialists. *AET* 2020, **1**, 119-131. https://doi.org/10.5220/0010921800003364

[25] Sharif, M.M., Haas, C. and Walbridge, S. (2022) Using Termination Points and 3D Visualization for Dimensional Control in Prefabrication. *Automation in Construction*, **133**, Article ID: 103998. https://doi.org/10.1016/j.autcon.2021.103998

[26] Abboud, A. and Nasrullah, M. (2022) 3D Modeling for the Largest Stone in the World Using Laser Scanning Technique. Doctoral Dissertation, Полоцкий государственный университет имени Евфросинии Полоцкой.

[27] Karami, A., Menna, F. and Remondino, F. (2022) Combining Photogrammetry and Photometric Stereo to Achieve Precise and Complete 3D Reconstruction. *Sensors*, **22**, 8172. https://doi.org/10.3390/s22218172

[28] Wachowski, L. and Wachowski, L. (1999) *The Matrix* [Film] Warner Bros. Pictures.

[29] Mamani, A.R.C., Polanco, S.R.R. and Ordonez, C.A.C. (2022) Systematic Review on Photogrammetry, Streaming, Virtual and Augmented Reality for Virtual Tourism. *HCI International* 2022—*Late Breaking Papers*: *Interacting with Extended Reality and Artificial Intelligence*: 24*th International Conference on Human-Computer Interaction*, *HCII* 2022, Vol. 13518, 46.

[30] OpenAI (2022, April 14) Dall·E 2. https://openai.com/dall-e-2

[31] Blance, A. (2022, July 27) Using AI to Generate 3D Models, Fast! Medium. https://towardsdatascience.com/using-ai-to-generate-3d-models-2634398c0799

[32] Ha, H. and Song, S. (2022) Semantic Abstraction: Open-World 3d Scene Understanding from 2d Vision-Language Models. 6*th Annual Conference on Robot Learning*, Cornell University, Ithaca.

[33] McKenzie, T. (2022, August 17) Kaedim: A Cool AI for Turning 2D Images into 3D Models. https://80.lv/articles/kaedim-a-cool-ai-for-turning-2d-images-into-3d-models

[34] Bebeshko, B., Khorolska, K., Kotenko, N., Desiatko, A., Sauanova, K., Sagyndykova, S. and Tyshchenko, D. (2021) 3D Modelling by Means of Artificial Intelligence. *Journal of Theoretical and Applied Information Technology*, **99**, 1296-1308.

[35] Salian, I. (2022, August 12) Nerf Research Turns 2D Photos into 3D Scenes. NVIDIA Blog. https://blogs.nvidia.com/blog/2022/03/25/instant-nerf-research-3d-ai

[36] Touloupaki, E. and Theodosiou, T. (2017) Performance Simulation Integrated in Parametric 3D Modeling as a Method for Early Stage Design Optimization—A Review. *Energies*, **10**, 637. https://doi.org/10.3390/en10050637

[37] Raitt, B. and Minter, G. (2000) Digital Sculpture Techniques. *Interactivity Magazine*, **4**, 1-13. https://doi.org/10.14733/cadconfP.2019.392-396

[38] Rossoni, M., Barsanti, S.G., Colombo, G. and Guidi, G. (2020) Retopology and Simplification of Reality-Based Models for Finite Element Analysis.

[39] Tasdizen, T., Whitaker, R., Burchard, P. and Osher, S. (2003) Geometric Surface Processing via Normal Maps. *ACM Transactions on Graphics* (*TOG*), **22**, 1012-1033. https://doi.org/10.1145/944020.944024

[40] Jean-Pierre, J. (2021) The Invasion: Applying the Aesthetics of Horror Films in a

Virtual Reality Gaming Environment. Doctoral Dissertation, Florida Atlantic University, Boca Raton.

[41] Choquette, J., Gandhi, W., Giroux, O., Stam, N. and Krashinsky, R. (2021) NVIDIA A100 Tensor Core GPU: Performance and Innovation. *IEEE Micro*, **41**, 29-35. https://doi.org/10.1109/MM.2021.3061394

[42] Nvidia Developer (2019, December 3) NVIDIA NGX Technology—AI for Visual Applications. https://developer.nvidia.com/rtx/ngx

[43] Wu, J., Dick, C. and Westermann, R. (2015) A System for High-Resolution Topology Optimization. *IEEE Transactions on Visualization and COMPUTER Graphics*, **22**, 1195-1208. https://doi.org/10.1109/TVCG.2015.2502588

[44] Haar, U. and Aaltonen, S. (2015) GPU-Driven Rendering Pipelines—Real-Time Rendering. http://advances.realtimerendering.com/s2015/aaltonenhaar_siggraph2015_combined_final_footer_220dpi.pdf

[45] Barrier, R. (2022, May 17) What Unreal Engine 5 MEANS for the Games Industry. IGN. https://www.ign.com/articles/what-unreal-engine-5-means-for-the-games-industry

[46] El-Wajeh, Y.A., Hatton, P.V. and Lee, N.J. (2022) Unreal Engine 5 and Immersive Surgical Training: Translating Advances in Gaming Technology into Extended-Reality Surgical Simulation Training Programmes. *British Journal of Surgery*, **109**, 470-471. https://doi.org/10.1093/bjs/znac015

[47] Beardsall, R. (2020, September 4) Unreal Engine 5 and Nanite Virtualized Geometry—What Does It Mean for Content Creators? Medium. https://medium.com/xrlo-extended-reality-lowdown/unreal-engine-5-and-nanite-virtualized-geometry-what-does-it-mean-for-content-creators-b4106accd306

[48] Grba, D. (2022) Deep Else: A Critical Framework for AI Art. *Digital*, **2**, 1-32. https://doi.org/10.3390/digital2010001

Scientific Research Publishing

Journal of Intelligent Learning Systems and Applications, 2023, 15, 36-56
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

A Comparison of PPO, TD3 and SAC Reinforcement Algorithms for Quadruped Walking Gait Generation

James W. Mock, Suresh S. Muknahallipatna*

Department of Electrical Engineering and Computer Science, University of Wyoming, Laramie, Wyoming, USA
Email: jmock2@uwyo.edu, *sureshm@uwyo.edu

How to cite this paper: Mock, J.W. and Muknahallipatna, S.S. (2023) A Comparison of PPO, TD3 and SAC Reinforcement Algorithms for Quadruped Walking Gait Generation. *Journal of Intelligent Learning Systems and Applications*, **15**, 36-56.
https://doi.org/10.4236/jilsa.2023.151003

Received: December 19, 2022
Accepted: February 25, 2023
Published: February 28, 2023

Abstract

Deep reinforcement learning (deep RL) has the potential to replace classic robotic controllers. State-of-the-art Deep Reinforcement algorithms such as Proximal Policy Optimization, Twin Delayed Deep Deterministic Policy Gradient and Soft Actor-Critic Reinforcement Algorithms, to mention a few, have been investigated for training robots to walk. However, conflicting performance results of these algorithms have been reported in the literature. In this work, we present the performance analysis of the above three state-of-the-art Deep Reinforcement algorithms for a constant velocity walking task on a quadruped. The performance is analyzed by simulating the walking task of a quadruped equipped with a range of sensors present on a physical quadruped robot. Simulations of the three algorithms across a range of sensor inputs and with domain randomization are performed. The strengths and weaknesses of each algorithm for the given task are discussed. We also identify a set of sensors that contribute to the best performance of each Deep Reinforcement algorithm.

Keywords

Reinforcement Learning, Machine Learning, Markov Decision Process, Domain Randomization

1. Introduction

Robots have become extremely common within the past few decades. From manufacturing to healthcare, robots play an important integral role in the modern world and with likely be even more integral in the future. However, biomimetic robots, such as humanoid and quadruped robots, are significantly less common. This is primarily due to the limitation of their control algorithms. Many of these controllers utilize sophisticated kinematic and dynamic models separated into submodules so they are easier to manage [1]. These models are difficult and time-consuming to develop and require expertise in both robotics and walking locomotion. Furthermore, these controllers routinely fail to achieve the performance of their biological counterparts. Though more difficult to control, walking robots offer an attractive alternative to typical locomotion systems. Walking robots are more suited for efficiently moving over uneven terrain due to their ability to select where they make contact with the terrain. They also possess an edge with regard to navigation since they are cable of stepping or jumping over obstacles that wheeled or tracked vehicles could not pass [1] [2].

Despite limited use outside of research applications, many walking robots have been developed. Examples of humanoid robots include NASA's Valkyrie [3], Boston Dynamics' Atlas, and Agility Robotics' Cassie. Prominent quadruped robots include Boston Dynamics' Spot, MIT's Mini Cheetah [4] and Robotic Systems Lab's ANYmal [5]. All of these robots are sophisticated enough to perform incredible feats of agility but lack the control systems require to operate at their peak performance. To address this shortfall, machine learning (ML) is employed to develop more complex and robust control systems.

Reinforcement learning (RL) is a sub-field of machine learning that learns through interacting with an environment rather than from large datasets as with supervised and unsupervised learning. The goal of RL is to map states to actions with an artificial neural network (ANN) through a trial-and-error process. There are two primary components in RL, the agent and the environment. **Figure 1** depicts the agent environment interaction. The agent is responsible for making decisions based on the current state of the environment. The environment is anything the agent cannot change arbitrarily. In robotic applications, it would be natural to assume that the robot is the agent. However, the robot's actuators, links, sensors, etc. are considered to be part of the environment since the agent cannot explicitly change them. Therefore, the agent is not the robot but actually the control algorithm for the robot.

The fundamental basis for modern reinforcement learning algorithms is the Markov Decision Process (MDP). A MDP is a discrete time state transition model which consists of four components: state space, action space, state transition probabilities and reward. This is usually represented as the tuple ($\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}$).

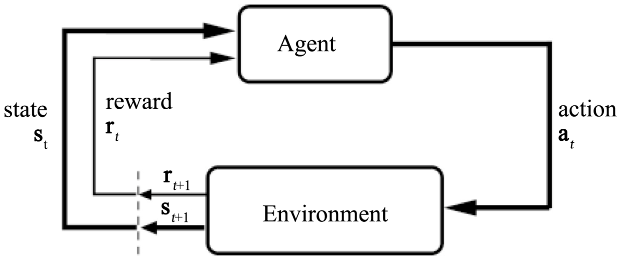


Figure 1. Agent-environment interaction. Image credit: [6].

The state space is defined by a set of observations of the robot and its environment. An observation at timestep t is given as $\mathbf{s}_t \in \mathcal{S}$. In virtual environments many observations can be obtained directly from the physics engine. For real robots observations usually come in the form of a variety of sensors such as IMUs, motor encoders and cameras. In robotic applications the action space is usually defined by the range of each actuator. The action taken at timestep t is represented as $\mathbf{a}_t \in \mathcal{A}$. $\mathcal{P}$ represents the probability density of the next state $\mathbf{s}_{t+1}$ given the current state $\mathbf{s}_t$ and action $\mathbf{a}_t$. Reward $\mathbf{r}_t \in \mathcal{R}$ is received after transitioning from state $\mathbf{s}_t$ to state $\mathbf{s}_{t+1}$, due to action $\mathbf{a}_t$. The reward is always a real scalar value. The function that provides the reward is defined be the system designer to achieve some goal (e.g. walking). The return is defined as the discounted sum of rewards $J_t = \sum_{t=0}^{T} \gamma^t \mathbf{r}_t$ where $\gamma \in (0,1]$ is the discount factor determining the priority of long term rewards. Values of γ closer to 0 will cause the agent to prioritize short term rewards over long term rewards.

A solution to an MDP is defined as policy π, which maps each state to an action to take in this state to return the highest average reward. RL methods specify how the agent updates its policy as a result of its experience to maximize the return. Additionally, most RL algorithms involve estimating value functions. These functions estimate either the value of being in a particular state or the value of taking particular action. The state-value function, $V_\pi(\mathbf{s}_t)$, for policy π is the expected return when starting in s_t and following π. $V_\pi(\mathbf{s}_t)$ is formally defined as

$$V_\pi(\mathbf{s}_t) = \mathbb{E}_\pi[J_t \mid \mathbf{s}_t] = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty} \gamma^k \mathbf{r}_{t+k+1} \mid \mathbf{s}_t\right]. \quad (1)$$

The action-value function, also know as the Q-function, $Q_\pi(\mathbf{s}_t, \mathbf{a}_t)$, is the expected return starting from $\mathbf{s}_t$, taking the action $\mathbf{a}_t$, and thereafter following policy π. $Q_\pi(\mathbf{s}_t)$ is defined as

$$Q_\pi(\mathbf{s}_t, \mathbf{a}_t) = \mathbb{E}_\pi[J_t \mid \mathbf{s}_t, \mathbf{a}_t] = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty} \gamma^k \mathbf{r}_{t+k+1} \mid \mathbf{s}_t, \mathbf{a}_t\right]. \quad (2)$$

Both value functions can be estimated from experience. A policy π is defined to be better than or equal to a policy π' if its expected return is greater than or equal to that of π' for all states. The optimal policy is defined as the policy with state-value function $V^*(\mathbf{s}_t) = \max_\pi V_\pi(\mathbf{s}_t)$ and action-value function $Q^*(\mathbf{s}_t, \mathbf{a}_t) = \max_\pi Q_\pi(\mathbf{s}_t, \mathbf{a}_t)$.

Using reinforcement learning over traditional control methods has three potential advantages over traditional controller designs. The first and most significant advantage is the ability to create more sophisticated and robust control algorithms for walking robots. Currently, even mildly rough terrain would pose a serious challenge for most walking robots. The second advantage is a reduction in human effort to develop complex control algorithms. In many cases it may take months or even years to develop control schemes for walking robots. A robot that could learn to walk on its own could drastically reduce the time it takes to develop a suitable control algorithm. The third advantage is the possibility of creating more complex robots. Currently, robots are intentionally simplified so they are easier to control. For example, all of the quadruped robots mentioned previously use the same three degrees of freedom (DOF) per leg configuration. A real dog has at least six DOF per leg excluding toes. This anatomical simplification is likely contributing to the limited capabilities of bio-mimetic robots.

This paper is organized as follows: Section 2 provides a brief summary of the related work. In Section 3, an in-depth mathematical background of the three RL algorithms is presented. Section 4 discusses the experimental setup, training and performance metrics. In Section 5, the simulation results of the training are presented. Section 6 presents the conclusion and future work.

2. Related Work

It has been widely shown that RL algorithms can produce highly sophisticated control policies for tasks in simulations [7] [8] [9]. Three of the top performing algorithms often used for robotics task are Twin Delayed Deep Deterministic Policy Gradient (TD3), Proximal Policy Optimization (PPO) and Soft Actor Critic (SAC). Nevertheless, few performance comparisons of the RL algorithms for robotics applications can be found in the literature, and the few existing comparisons exhibit contradictory performance results. For example, in Fujimoto *et al.* [9], TD3 is shown be the top performing algorithm in several robotic walking tasks, including the "HalfCheetah" and "Ant" walking tasks, compared to PPO and SAC. However, this is contradicted in Haarnoja *et al.* [8] where SAC is shown to be the top performing algorithm for the same tasks compared to TD3 and PPO. Such contradictions make it difficult to ascertain which algorithm is suitable for a particular robot application.

This work seeks to clearly demonstrate how each algorithm performs on a simulated quadruped robotic walking task. Additionally, the algorithms are compared over a variety of sensory inputs. Lastly, each algorithm is tested with domain randomization which is essential for transfer learning of real robots.

3. Overview of Algorithms

RL algorithms are roughly separated into two categories, model-based and model-free. The key distinction between the two is whether or not the agent uses a model of the environment to predict state transitions and rewards. Model-based RL is a deductive approach for solving a problem. The agent uses its understanding of the system to select a best action. Model-based algorithms may learn or be given the environment model. Model-free RL is an inductive approach for solving a problem. The agent uses its past experience to estimate the value of its action. Since model-free algorithm does not rely on the transition probabilities of the MDP in order to find a policy. This is ideal for robots with high dimension, continuous state and action spaces.

The most common type of reinforcement learning used in robotic applications are model-free actor-critic algorithms. Actor-critic methods are time difference (TD) methods that have a separate structures to explicitly represent the policy independent of a value function. The policy structure is known as the *actor*, because it is used to select actions, and the estimated value function is known as the *critic*, because it criticizes the actions made by the actor. The critique takes the form of a TD error which is shown in Equation (3).

$$\delta_t = \mathbf{r}_t + \gamma V(\mathbf{s}_{t+1}) - V(\mathbf{s}_t) \tag{3}$$

where $\mathbf{r}_t$ is the reward at time *t* given state $\mathbf{s}_t$ and action $\mathbf{a}_t$, γ is the discount factor and V is the value function implemented by the critic at time *t*. This scalar signal is the only output of the critic and drives all learning in both actor and critic. **Figure 2** shows the actor-critic architecture. Typically, the critic is a state-value function. After each action selection, the critic evaluates the new state to determine whether things have improved or worse than expected. If the TD error is positive, the action is encouraged in the future, whereas if the TD error is negative, it becomes adverse to it.

Model-free actor-critic algorithms can be subdivided into two groups, on-policy and off-policy. On-policy methods, also known as policy optimization, only uses data collected while acting according to the most recent version of the policy to make updates to the policy. Policy optimization also usually involves learning an approximator $V_\phi(\mathbf{s}_t)$ for the on-policy value function $V^\pi(\mathbf{s}_t)$, which is used to update the policy. Q-Learning methods learn an approximator $Q_\theta(\mathbf{s}_t, \mathbf{a}_t)$ for the optimal action-value function, $Q^*(\mathbf{s}_t, \mathbf{a}_t)$. This optimization is usually performed off-policy. Meaning that each update can use data collected at any point during training, regardless of how the agent was choosing to explore the environment when the data was obtained. These methods often make use of memory buffers that store state-action-state tuples. The primary strength of on-policy methods is that they tend to be stable and reliable. By contrast, off-policy methods tend to be less stable but substantially more sample efficient, because they can reuse data more effectively than on-policy techniques. Both methods have shown good performance in robotic tasks [7] [8] [9].

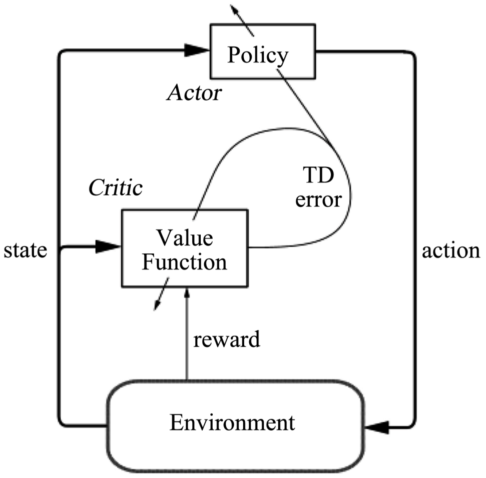


Figure 2. Actor-critic architecture. Image credit: [6].

Three state-of-the-art continuous control policy learning algorithms were chosen to benchmark the gait learning and performance. Proximal Policy Optimization, Twin Delayed Deep Deterministic Policy Gradient and Soft Actor-Critic are consistently shown to be the top performing model-free actor-critic algorithms used for robotic tasks.

PPO is an on-policy RL algorithm that attempts to improve the on Trust Region Policy Optimization (TRPO) algorithm [10]. TRPO attempts to control the policy updates through a Kullback-Leibler (KL) divergence constraint, which quantifies how much a probability distribution differs from another [10]. A major disadvantage of this approach is that it's computationally expensive. PPO clips the objective function to prevent large updates to the policy [7]. This make PPO easier to implement and computationally faster. The clipped objective function is shown in Equation (4).

$$L^{CLIP}(\theta)=\hat{\mathbb{E}}_t\left[\min\left(\frac{\pi_\theta(\mathbf{a}_t\mid\mathbf{s}_t)}{\pi_{\theta_{old}}(\mathbf{a}_t\mid\mathbf{s}_t)}\hat{A}_t,\operatorname{clip}\left(\frac{\pi_\theta(\mathbf{a}_t\mid\mathbf{s}_t)}{\pi_{\theta_{old}}(\mathbf{a}_t\mid\mathbf{s}_t)},1-\varepsilon,1+\varepsilon\right)\hat{A}_t\right)\right] \quad (4)$$

where π_θ is a stochastic policy. The clipping function limits the lower and upper value of the probability ratio $\frac{\pi_\theta(\mathbf{a}_t\mid\mathbf{s}_t)}{\pi_{\theta_{old}}(\mathbf{a}_t\mid\mathbf{s}_t)}$ and ε is the hyperparameter that sets clip range. The larger the the value of ε, larger the potential policy changes. $\hat{A}_t$ is the advantage function shown in Equation (5).

$$\hat{A}_t=\delta_t+(\gamma\lambda)\delta_{t+1}+\cdots+(\gamma\lambda)^{T-t+1}\delta_{T-1} \quad (5)$$

where δ_t is the TD error defined in Equation (3), γ is the discount factor, and λ is the bias-variance trade-off factor for the generalized advantage estimator [11]. During each episode of training the actor collects T timesteps of data. Then the surrogate loss is computed over T timesteps and optimized with minibatch stochastic gradient descent for K epochs. **Algorithm 1** summarizes the training process for PPO. As training progresses the policy will try to exploit rewards that it has already found over exploration.

TD3 is an off-policy algorithm that significantly improves upon the deep deterministic policy gradient (DDPG) algorithm [12]. The primary downfall of DDPG is the overestimation bias of the critic network which leads to degraded performance. TD3 implements three key features to improve performance [9]. First, TD3 proposes the use of a clipped double Q-learning algorithm to replace

```
for iteration=1,2,... do
    for iteration=1,2,...,$N$ do
        Run policy $\pi_{\theta\text{old}}$ in environment for $T$ timesteps
        Compute advantage estimates $\hat{A}_1, \ldots, \hat{A}_T$
    end for
    Optimize surrogate $L$ wrt $\theta$, with $K$ epochs and minibatch size $M \le NT$
    $\theta_{old} \leftarrow \theta$
end for
```

Algorithm 1. PPO, actor-critic style.

the standard Q-learning found in DDPG. The second feature implemented is the use of action noise to reduce overfitting to narrow peaks in the value estimate, a problem often encountered with deterministic policies. The addition of action noise also results in target policy smoothing. For each timestep, both Q networks ($Q_{\theta_1}, Q_{\theta_2}$) are updated towards the minimum target value of actions selected by the target policy shown in Equation (6)

$$y = \mathbf{r}_t + \gamma \min_{i=1,2} Q_{\theta_i'}\left(\mathbf{s}_{t+1}, \pi_{\phi'}\left(\mathbf{s}_{t+1}\right) + \varepsilon\right). \tag{6}$$

where $\mathbf{r}_t$ is the reward at time *t*, γ is the discount factor and π_ϕ is a deterministic policy, with parameters ϕ, which maximizes the expected return. ε is the clipped Gaussian action noise added and is defined by Equation (7).

$$\varepsilon \sim \operatorname{clip}\left(\mathcal{N}\left(0, \sigma\right), -c, c\right) \tag{7}$$

The third feature of TD3 is to delay the policy updates by a fixed number of updates to the critic. This is done to suppress the value estimate variance caused by the accumulated TD-error. Parameters ϕ are updated according to the deterministic policy gradient shown in Equation (8).

$$\nabla_\phi J\left(\phi\right) = \mathbb{E}_{\mathbf{s}_t \sim p_\pi}\left[\nabla_a Q_\pi\left(\mathbf{s}_t, \mathbf{a}_t\right)\Big|_{\mathbf{a}_t = \pi_\phi(\mathbf{s}_t)} \nabla_\phi \pi_\phi\left(\mathbf{s}_t\right)\right]. \tag{8}$$

where Q_π is the action-value function defined in Equation (2). TD3 is summarized in **Algorithm 2**.

SAC is an off-policy actor-critic algorithm that seeks to maximize a trade-off between expected return and entropy. This encourages a high degree exploration compared to other algorithms. The entropy augmented objective is defined by Equation (9).

$$\pi^* = \arg\max_\pi \sum_t \mathbb{E}_{(\mathbf{s}_t, \mathbf{a}_t) \sim \rho_\pi}\left[\mathbf{r}_t + \alpha \mathcal{H}\left(\pi\left(\cdot \mid \mathbf{s}_t\right)\right)\right], \tag{9}$$

```
Initialize critic networks $Q_{\theta_1}, Q_{\theta_2}$, and actor network $\pi_\phi$ with random parameters $\theta_1, \theta_2, \phi$
Initialize target networks weights $\theta_1' \leftarrow \theta_1, \theta_1' \leftarrow \theta_1, \phi' \leftarrow \phi$
Initialize replay buffer $\mathcal{B}$
for $t = 1$ to $T$ do
    Select action with exploration noise $\mathbf{a}_t \sim \pi_\phi(\mathbf{s}_t) + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma)$
    Observe reward $\mathbf{r}_t$ and new state $\mathbf{s}_{t+1}$
    Store transition tuple $(\mathbf{s}_t, \mathbf{a}_t, \mathbf{r}_t, \mathbf{s}_{t+1})$ in $\mathcal{B}$
    Sample mini-batch of $N$ transitions $(\mathbf{s}_t, \mathbf{a}_t, \mathbf{r}_t, \mathbf{s}_{t+1})$ from $\mathcal{B}$
    $\tilde{\mathbf{a}} \leftarrow \pi_{\phi'}(\mathbf{s}_t) + \epsilon$, $\epsilon \sim \mathrm{clip}(\mathcal{N}(0, \tilde{\sigma}), -c, c)$
    $y \leftarrow \mathbf{r}_t + \gamma \min_{i=1,2} Q_{\theta_i'}(\mathbf{s}_{t+1}, \tilde{\mathbf{a}})$
    Update critics $\theta_i \leftarrow \mathrm{argmin}_{\theta_i} N^{-1} \sum (y - Q_{\theta_i}(\mathbf{s}_t, \mathbf{a}_t))^2$ for $i \in \{1, 2\}$
    if $t$ mod $d$ then
        Update $\phi$ by the deterministic policy gradient:
        $\nabla_\phi J(\phi) = N^{-1} \sum \nabla_a Q_{\theta_1}(\mathbf{s}_t, \mathbf{a}_t)|_{a=\pi_\phi(\mathbf{s}_t)} \nabla_\phi \pi_\phi(\mathbf{s}_t)$
        Update target networks:
        $\theta_i' \leftarrow \tau\theta_i + (1 - \tau)\theta_i'$ for $i \in \{1, 2\}$
        $\phi' \leftarrow \tau\phi + (1 - \tau)\phi'$
    end if
end for
```

Algorithm 2. TD3.

$\mathbf{r}_t$ is the reward at time t and α determines the relative importance of the entropy term, $\mathcal{H}\left(\pi\left(\cdot\,|\,\mathbf{s}_t\right)\right) = \log\left(\pi_\phi\left(\mathbf{a}_t\,|\,\mathbf{s}_t\right)\right)$, against the reward. π_ϕ is a deterministic policy, with parameters ϕ. SAC utilizes two soft Q-functions to mitigate positive bias in the policy improvement step. The soft Q-function parameters, θ, are trained to minimize the soft Bellman residual given in Equation (10).

$$J_Q\left(\theta\right) = \mathbb{E}_{\left(\mathbf{s}_t,\mathbf{a}_t\right)\sim\mathcal{B}}\left[\frac{1}{2}\left(Q_\theta\left(\mathbf{s}_t,\mathbf{a}_t\right) - \left(r_t + \gamma\mathbb{E}_{\mathbf{s}_{t+1}\sim p}\left[V_{\bar{\theta}}\left(\mathbf{s}_{t+1}\right)\right]\right)\right)^2\right], \tag{10}$$

where $Q_\theta\left(\mathbf{s}_t,\mathbf{a}_t\right)$ is the minimum of the two soft Q-functions and γ is the discount factor. The value function V_θ is the value function implicitly parameterized through the soft Q-function parameters via Equation (11).

$$V\left(\mathbf{s}_t\right) = \mathbb{E}_{\mathbf{a}_t\sim\pi}\left[Q\left(\mathbf{s}_t,\mathbf{a}_t\right) - \alpha\log\pi\left(\mathbf{a}_t\,|\,\mathbf{s}_t\right)\right] \tag{11}$$

The policy parameters are trained by minimizing the objective function in Equation (12).

$$J_\pi\left(\phi\right) = \mathbb{E}_{\mathbf{s}_t\sim\mathcal{B}}\left[\mathbb{E}_{\mathbf{a}_t\sim\pi_\phi}\left[\alpha\log\left(\pi_\phi\left(\mathbf{a}_t\,|\,\mathbf{s}_t\right)\right) - Q_\theta\left(\mathbf{s}_t,\mathbf{a}_t\right)\right]\right]. \tag{12}$$

Additionally, the temperature parameter α can be learned with the following objective function in Equation (13).

$$J\left(\alpha\right) = \mathbb{E}_{\mathbf{a}_t\sim\pi_t}\left[-\alpha\log\pi_t\left(\mathbf{a}_t\,|\,\mathbf{s}_t\right) - \alpha\bar{\mathcal{H}}\right] \tag{13}$$

The pseudo code for SAC is listed in **Algorithm 3**. SAC alternates between collecting experience from the environment with the current policy and updating the actor and critic network parameters using stochastic gradients from batches randomly sampled from a replay buffer [8].

4. Experimental Setup

This section covers the design and setup of the simulated quadruped robot and its environment.

```
Initialize parameters $\theta_1, \theta_2, \phi$
Initialize target networks weights $\theta_1' \leftarrow \theta_1, \theta_1' \leftarrow \theta_1, \phi' \leftarrow \phi$
Initialize replay buffer $\mathcal{B}$
for each iteration do
    for each environment step do
        Sample action from policy $\mathbf{a}_t \sim \pi_\phi(\mathbf{a}_t|\mathbf{s}_t)$
        Sample transition from environment $\mathbf{s}_{t+1} \sim p(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t)$
        Store transition tuple $(\mathbf{s}_t, \mathbf{a}_t, \mathbf{r}_t, \mathbf{s}_{t+1})$ in $\mathcal{B}$
    end for
    for each gradient step do
        Update the Q-function parameters $\theta_i \leftarrow \theta_i - \lambda_Q \hat{\nabla}_{\theta_i} J_Q(\theta_i)$ for $i \in \{1, 2\}$
        Update policy weights $\phi \leftarrow \phi - \lambda_\pi \hat{\nabla}_\phi J_\pi(\phi)$
        Adjust temperature parameter $\alpha \leftarrow \alpha - \lambda \hat{\nabla}_\alpha J(\alpha)$
        Update target network weights $\theta_i' \leftarrow \tau\theta_i + (1 - \tau)\theta_i'$ for $i \in \{1, 2\}$
    end for
end for
```

Algorithm 3. SAC.

4.1. System Identification

The simulated environment is constructed using the MuJoCo's physics simulator. MuJoCo is a free and open source physics engine that aims to facilitate research and development in robotics, biomechanics, graphics and animation [13]. MuJoCo makes it possible to scale up computationally-intensive techniques such optimal control, physically-consistent state estimation, system identification and automated mechanism design, and apply them to complex dynamical systems in contact-rich behaviors. It is well suited for training RL policies for robotic tasks. The simulation environment consists of a single 2 DOF quadruped robot and a ground plane. **Figure 3** shows the simulated robot. The robot is similar to the popular Mujoco "Ant" benchmark, but has more realistic actuator torques and includes a variety common sensors that can be found on real robots. These sensors include a body position, body quaternion, foot contact sensors, and actuator position, actuator velocity, actuator load (force), IMUs on the body and legs. The large yellow spheres at the end of the feet represent the feet contact sensors. The smaller yellow spheres on the legs represent the locations of the IMU sensors. An additional IMU is located at the center of the main body. The body position and quaternion are measured at the center of the main body. Gaussian noise was added to each sensor to imitate the imprecision of real sensors. The robot has eight actuators. No actuator or latency models were considered for the simulation.

4.2. State and Action Spaces

Actions $\mathbf{a}_t$ are actuator target positions mapped to values between −1 and 1. The state $\mathbf{s}_t$ consists of the most recent readings of various sensors. Seven sensor configurations were tested with each algorithm to identify the best possible level of sensory input for each algorithm. The sensors used in each configuration are listed in **Table 1**. The first configuration (v0) uses only the body quaternion for the state space and the last configuration (v6) utilizes all sensor data on the robot for the state space.

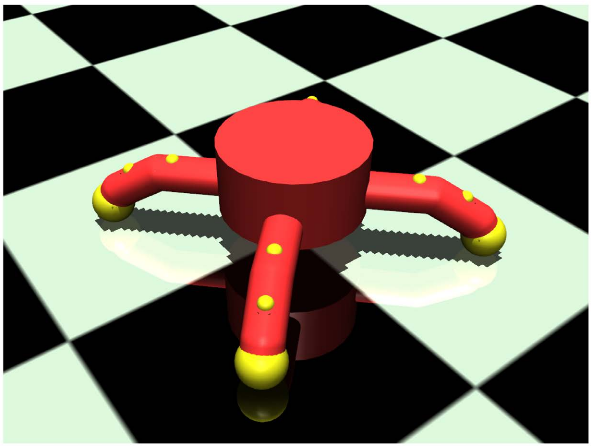


Figure 3. Simulated quadruped robot testbed.

Table 1. State space configurations.

Config	v0	v1	v2	v3	v4	v5	v6
Body Quaternion	✓	✓	✓	✓	✓	✓	✓
Actuator Position (Qty: 8)	-	✓	✓	✓	✓	✓	✓
Actuator Velocity (Qty: 8)	-	-	✓	✓	✓	✓	✓
Actuator Load (Qty: 8)	-	-	-	✓	✓	✓	✓
Foot Pressure Sensor (Qty: 4)	-	-	-	-	✓	✓	✓
3-axis Accelerometer (Qty: 9)	-	-	-	-	-	✓	✓
3-axis Gyro (Qty: 9)	-	-	-	-	-	-	✓

4.3. Reward Function

The reward function was designed to encourage a stable forward walking gait at a target velocity $\hat{v}_x$ with a target orientation $\hat{q}$. The reward function is given by the Equation (14)

$$\begin{aligned} R = {} & r_H + w_A * \sum(\mathbf{a}_t)^2 + w_V * \sum(\Omega_t)^2 + w_{TV} * |\hat{v}_x - v_x| + w_D * |v_y| \\ & + w_{TQ} * \text{sum}(|\hat{q} - q|), \end{aligned} \tag{14}$$

where r_H is the reward for not experiencing a catastrophic failure, such as flipping over. w_A is the weight that determines the importance of penalizing actions, $\mathbf{a}_t$. w_V is the weight that determines the importance of penalizing actuator velocities, Ω_t. w_{TV} is the importance weight for the target velocity and w_D is the importance weight for penalizing linear velocity in the lateral direction. w_{TQ} is the importance weight for deviation from the target quaternion. The final weights used are listed in Table 2.

4.4. Domain Randomization

The inevitable imperfections of physics simulations will automatically be exploited by any optimization method to achieve an improvement. However, since these exploits don't exist in the real world, policies transferred to the real world will not perform as expected. This is known as the simulation optimization bias (SOB) [14]. One method to combat SOB is to randomize parameters of the simulation. Unlike system identification which aims to carefully model the real world, domain randomization aims to randomize the visuals or system dynamics of a simulation to encourage generalization. System identification and domain randomization are often used together to achieve better results [1] [15] [16]. Early domain randomization techniques largely consisted of adding i.i.d. noise to observations and actions [14]. Newer techniques involve changing the appearance and core dynamics of a simulated environment. Vision based learning have a particularly wide reality gap because it is very difficult to generate sufficiently high-quality rendered images [16]. Additionally, simulated cameras fail to incorporate noise and optical distortions produced by real cameras [17]. For a vision based object manipulation tasks, Pinto *et al.* [18] randomized textures,

Table 2. Reward function parameters.

Parameter	$\hat{v}_x$	$\hat{q}$	r_H	w_A	w_V	w_{TV}	w_D	w_{TQ}
Value	0.5 m/s	[1, 0, 0, 0]	1.0	−0.05	−0.05	−1.0	−0.5	−0.5

lighting and the position of the camera. They found that policies trained without domain randomization failed to perform when transferred to the real robot. For non-vision based robots parameters like mass, friction coefficients and actuator behavior are randomized. Tan *et al.* [15] found that using inertia randomization when learning a quadruped gait significantly improves robustness at the cost optimality. Meaning that using domain randomization causes the simulated policy to have degraded performance but will perform better on the physical robot. Adversarial disturbances to the agent are another common form of domain randomization. Rudin *et al.* [19] implemented this idea by pushing the simulated robot every 10 seconds. The robots' base is accelerated up to ±1 m/s in both x and y directions. This results in a highly stable and dynamic walking gait which was successfully deployed on a real robot.

4.5. Training

The simulated environment is setup to recreate the agent-environment described previously. Every 50 ms in simulation time the agent reads in the current state of the robot which is described by the robot's sensors. The sensors that are used depend on which configuration is being tested. The agent then uses the state space to generate target motor positions. The updated motor positions are sent the robot. After 50 ms the state of the robot is read again and a reward is given based on the reward function described in Section 4.3. This process repeats for one thousand iterations. Upon completion of an episode of one thousand steps the simulation is reset. TD3 and SAC make updates following the end each episode while PPO makes updates at a fixed interval. Each algorithm was trained on each sensor configuration for three million steps. This was repeated five times for each algorithm configuration combination.

To evaluate if an algorithm is suitable for transfer learning to a real robot a second group of policies were trained under identical circumstances except with domain randomization. The group using dynamics randomization experienced random variations in robot's mass, inertia and friction coefficients as well as variations in actuator stiffness, friction loss, damping, and reflected inertia.

4.6. Models and Hyperparameters

To compare optimal performance of each algorithm the Stable-Baselines3 (SB3) implementation was used for all three algorithms. SB3 is a set of reliable implementations of reinforcement learning algorithms in PyTorch [20]. Several combinations of hyperparameters were tested for each algorithm. However, the default SB3 values were found to be the best. **Table 3** summarizes the ANN architectures and hyperparameters used for each algorithm.

Table 3. Hyperparameters for each RL algorithm.

Hyperparameter	PPO	TD3	SAC
Network Architecture	[64, 64]	[256, 256]	[256, 256]
Activation	ReLU	ReLU	ReLU
Optimizer	Adam	Adam	Adam
Learning Rate	0.0003	0.001	0.0003
Target Update Rate	2048 Steps	1 Episode	1 Episode
Batch Size	64	100	256
Epochs	10	-	-
Discount Factor (γ)	0.99	0.99	0.99
Replay Buffer Size	-	10^6	10^6
Clip Range (ε)	0.2	-	-
GAE (λ)	0.95	-	-
Soft Update Coefficient (τ)	-	0.005	0.005
Target Entropy (α)	-	-	Auto
Action Noise	-	$\mathcal{N}(0,0.1)$	-
Policy Delay	-	2	-

4.7. Performance Metrics

The performance of trained policies was evaluated by comparing the quantitative metrics of walking gait for the three algorithms. The metrics associated with the walking gait are the average forward velocity (m/s), average forward velocity variance, average lateral velocity (m/s), average lateral velocity variance, and quaternion root mean square deviation (RMSD). Ideally an agent should achieve an average forward velocity of 0.5 m/s, a lateral velocity of 0.0 m/s, no forward or lateral velocity variance and no deviation in the quaternion. The maximum reward per time step that can be achieved is 1.0. Performance was evaluated as the average of all five trials over one thousand steps or fifty seconds in simulation time.

5. Results

This section provides an analysis of the simulated agent's performance. It also offers a comparison of algorithm performance across sensor configurations and with domain randomization.

5.1. Training without Domain Randomization

Figures 4-6 show the average learning curve in terms of the reward of each robot configuration using PPO, TD3 and SAC respectively without domain randomization. Across all three algorithms configurations v2, v3, and v4 achieve the

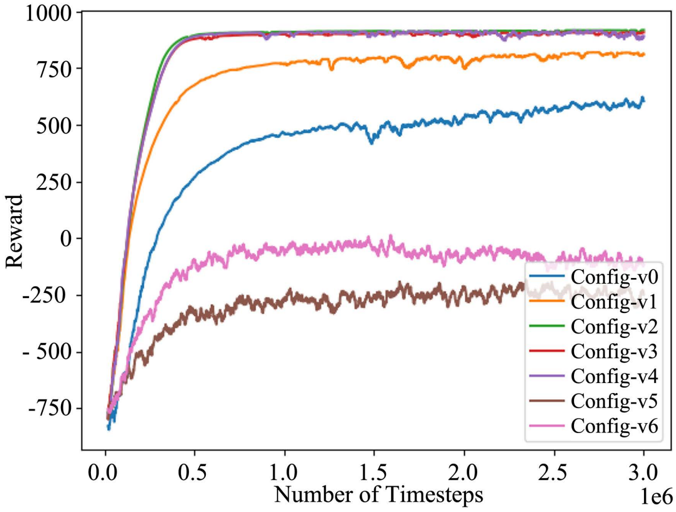


Figure 4. Average learning curve for each sensor configuration using PPO algorithm.

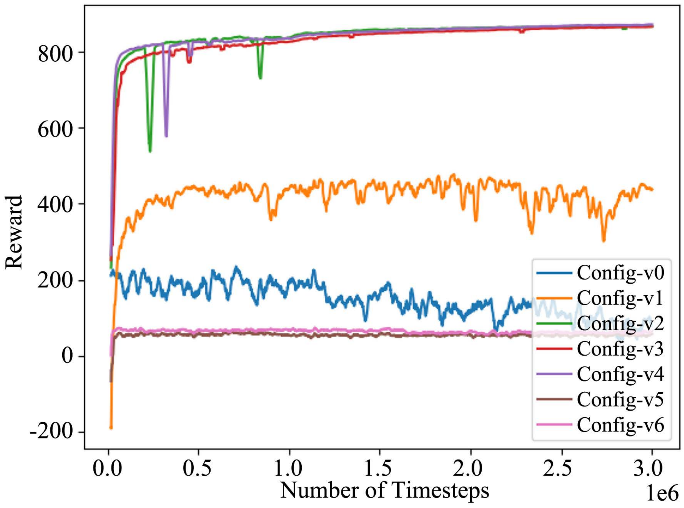


Figure 5. Average learning curve for each sensor configuration using TD3 algorithm.

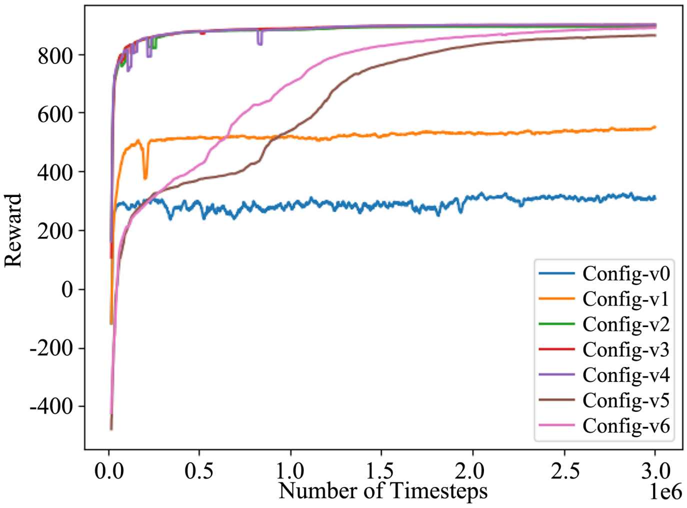


Figure 6. Average learning curve for each sensor configuration using SAC algorithm.

highest average reward. Configuration v2 is the first sensor configuration that includes actuator velocities. Unlike the body quaternion and actuator position, actuator velocity has a temporal relation which is likely the reason its inclusion in the state space significantly improves learning performance. The addition of actuator load in configuration v3 and contact sensors in configuration v4 do not improve the max reward beyond configuration v2. The addition of more sensors does not degrade learning till the IMU sensor data is added in configurations v4 and v5. SAC was the only algorithm to achieve a high reward for these configurations. However, training is significantly slower compared to other configurations. Configuration v6 achieve a reward comparable to configurations v2, v3, and v4 while configuration v5 has a slightly lower max reward. The fact that configuration v6 is higher than v5 would indicate that the policy is developing a form of sensor fusion for the IMU data. The inclusion of IMU data significantly increases the size of the state space. This indicates that for agents with large state spaces SAC may perform better. In contrast, PPO was the only algorithm to achieve an average reward higher than 500 using configurations v0 and v1, though they are still not on par with configurations v2, v3 and v4. This indicates that PPO performs better with smaller state spaces. It can also be seen that the TD3 and SAC agents have significantly steeper learning curves compared to PPO agents. This is due to the fact that off-policy algorithms make far more updates than on-policy algorithms. However, from **Figure 5** it can be seen that TD3 takes significantly longer to reach its max reward compared to the other two algorithms. PPO and SAC converge at their maximum reward before five hundred thousand steps, TD3 takes nearly two million steps to converge.

Tables 4-6 show the average quantitative metrics of the walking gaits achieved for each sensor configuration. **Table 4** shows the average performance of each sensor configuration using PPO. Configurations v2, v3 and v4 show the best performance with average forward velocities very close to the target velocity of 0.5 m/s. Additionally, all other metrics are close to zero as desired. Overall, it appears that configuration v3 performs the best as it was able to achieve an average velocity closest to the target velocity. Configurations v0 and v1 also seem

Table 4. Average performance of each sensor configuration trained with PPO algorithm.

Config	Avg Forward Velocity (m/s)	Forward Velocity Var	Avg Lateral Velocity (m/s)	Lateral Velocity Var	Quaternion RMSD
v0	0.37207	0.01229	0.01903	0.00649	0.06534
v1	0.44240	0.01486	0.00388	0.00136	0.03266
v2	0.49090	**0.00246**	−0.00108	0.00055	**0.01823**
v3	**0.49341**	0.00273	−0.00175	**0.00051**	0.01984
v4	0.49204	0.00267	**0.00023**	0.00053	0.01843
v5	0.00222	0.04943	0.01359	0.04511	0.80789
v6	0.07267	0.03297	−0.00192	0.03139	0.30946

Table 5. Average performance of each sensor configuration trained with TD3 algorithm.

Config	Avg Forward Velocity (m/s)	Forward Velocity Var	Avg Lateral Velocity (m/s)	Lateral Velocity Var	Quaternion RMSD
v0	−0.00078	0.00404	0.00027	0.00148	0.09021
v1	0.36572	0.02110	0.01244	0.02461	0.06642
v2	0.49277	0.00189	−0.00063	0.00075	0.02226
v3	**0.49724**	0.00208	−0.00197	0.00074	**0.02036**
v4	0.49554	0.00183	0.00666	0.00084	0.02175
v5	−0.00001	0.00053	−0.00039	0.00024	0.06395
v6	0.00026	**0.00028**	**0.00019**	**0.00010**	0.06356

Table 6. Average performance of each sensor configuration trained with SAC algorithm.

Config	Avg Forward Velocity (m/s)	Forward Velocity Var	Avg Lateral Velocity (m/s)	Lateral Velocity Var	Quaternion RMSD
v0	0.05932	0.01301	−0.01809	0.01040	0.05622
v1	0.34406	0.01992	0.01017	0.02643	0.04924
v2	0.48994	0.00182	**0.00018**	**0.00046**	0.02003
v3	0.49091	**0.00166**	−0.00111	0.00056	0.01956
v4	**0.49191**	0.00182	−0.00114	0.00062	0.01939
v5	0.48757	0.00184	−0.00896	0.00074	0.01924
v6	0.49165	0.00185	0.00066	0.00054	**0.01824**

to generate stable walking gaits. However, they were unable to come as close to the target velocity as the previously mentioned configurations. Lastly, configurations v5 and v6 fail to generate any walking gaits.

From **Table 5**, it can be seen that the TD3 algorithm only generates walking gaits for configurations v1 through v4. Like PPO, the average forward velocity for configuration v1 does not reach the target velocity. Also similar to PPO, the best gait was achieved with configuration v3. For configurations v2, v3, and v4 TD3 achieves a higher average velocity than both PPO and SAC. However, the other four metrics seem to be worse as a result. For configurations v5 and v6 all agents learn to remain still to achieve a maximum reward.

Table 6 shows that SAC was able to generate walking gaits for all configurations except v0. Like PPO and TD3, the average velocity of configuration v1 fails to achieve to desired velocity. Configurations v4 and v6 are the top performing configurations in terms of average forward velocity.

5.2. Training with Domain Randomization

Figures 7-9 show the average learning curve in terms of the reward of each robot configuration using PPO, TD3 and SAC respectively using domain

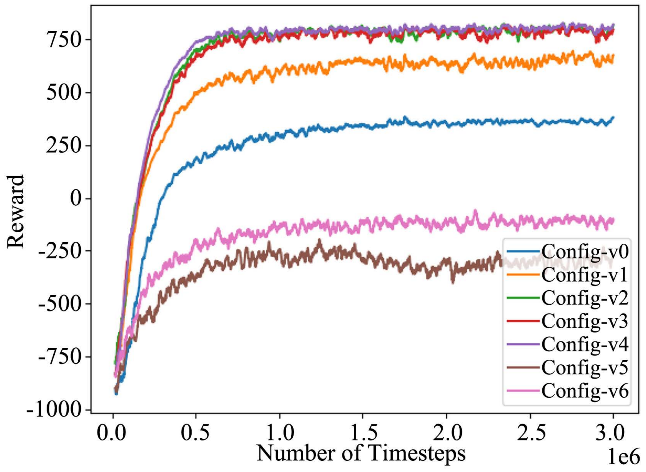


Figure 7. Average learning curve for each sensor configuration using PPO algorithm with domain randomization.

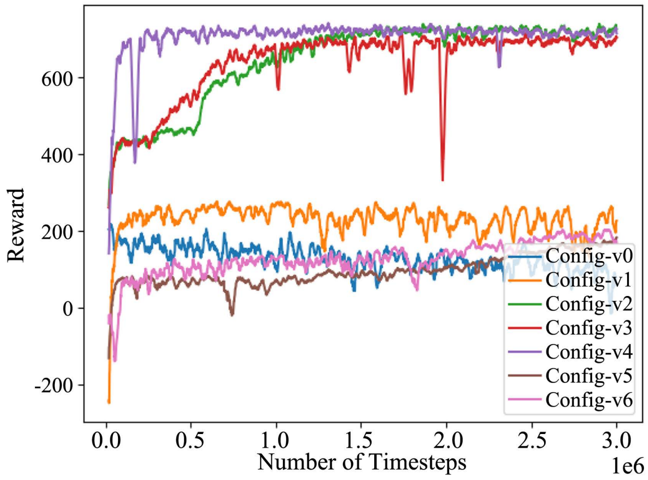


Figure 8. Average learning curve for each sensor configuration using TD3 algorithm with domain randomization.

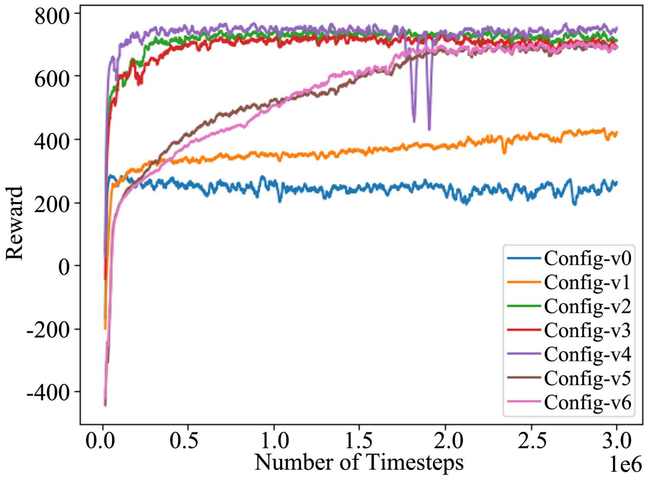


Figure 9. Average learning curve for each sensor configuration using SAC algorithm with domain randomization.

randomization. The learning curves are very similar to the simulations without domain randomization. The most significant difference being the lower maximum reward compared to agents trained without domain randomization. This is expected since randomization of dynamics makes walking much more difficult with the goal of achieving a more robust walking gait. Additionally, for some algorithm-configuration combinations, training is significantly slower. All configurations of PPO require approximately an additional three hundred thousand steps to reach their maximum reward. TD3 shows a notable preference for configuration v4 over v2 and v3 in terms of learning speed. Likewise, SAC also shows a slight preference for configuration v4. This indicates that foot sensors should be an important sensory input for real-world walking tasks.

Tables 7-9 show the average performance metrics for each algorithm-configuration combination. A notable difference that can be observed is the overshooting of the target forward velocity for PPO and TD3. TD3 especially overshoots be a considerable margin for configurations v2 and v3. Interestingly, SAC does not demonstrate the same overshooting behavior. It can also be seen that

Table 7. Average performance of each sensor configuration trained with PPO algorithm and domain randomization.

Config	Avg Forward Velocity (m/s)	Forward Velocity Var	Avg Lateral Velocity (m/s)	Lateral Velocity Var	Quaternion RMSD
v0	0.08164	0.00315	−0.00813	0.00479	0.04419
v1	0.51125	0.00537	−0.01126	0.00272	0.03613
v2	**0.50433**	0.00290	**0.00051**	0.00088	0.02525
v3	0.50494	0.00320	0.00125	**0.00082**	**0.02259**
v4	0.50726	**0.00281**	0.00185	0.00122	0.02613
v5	0.01854	0.04464	−0.02413	0.04370	0.84128
v6	0.03879	0.03025	−0.00069	0.02937	0.41497

Table 8. Average performance of each sensor configuration trained with TD3 algorithm and domain randomization.

Config	Avg Forward Velocity (m/s)	Forward Velocity Var	Avg Lateral Velocity (m/s)	Lateral Velocity Var	Quaternion RMSD
v0	−0.00034	**0.00129**	**−0.00062**	**0.00075**	0.10078
v1	0.00113	0.00293	0.00111	0.00092	**0.02000**
v2	0.51670	0.00318	0.00751	0.00170	0.02845
v3	0.51140	0.00414	0.01952	0.00348	0.03058
v4	**0.49742**	0.00388	0.00312	0.00205	0.02617
v5	0.11113	0.00282	0.00086	0.00145	0.08520
v6	0.11087	0.00166	−0.00901	0.00125	0.05135

Table 9. Average performance of each sensor configuration trained with SAC algorithm and domain randomization.

Config	Avg Forward Velocity (m/s)	Forward Velocity Var	Avg Lateral Velocity (m/s)	Lateral Velocity Var	Quaternion RMSD
v0	0.00680	0.01492	**0.00400**	0.01125	0.05964
v1	0.31937	0.02401	0.00403	0.04855	0.05884
v2	0.47342	0.00366	−0.02700	0.00293	**0.02600**
v3	0.48273	0.00452	0.01178	0.00190	0.02776
v4	0.48594	0.00348	0.00818	**0.00173**	0.02344
v5	0.49743	0.00390	−0.02039	0.00265	0.03122
v6	**0.49873**	**0.00239**	−0.00677	0.00327	0.03284

the performance of configuration v0 with PPO is significantly worse than without domain randomization. This can also be seen with configuration v1 with TD3. Lastly, SAC shows a significant performance increase for configurations v5 and v6 over other configurations.

5.3. Results Summary

- Actuator velocity is essential for generating stable walking gaits for all three RL algorithms.
- The performance of all three algorithms is very similar for configurations v2, v3 and v4 both with and without domain randomization.
- TD3 and SAC both learn significantly quicker than PPO.
- PPO excels with minimal state spaces but performs very poorly with the addition of IMU data.
- SAC was the only algorithm to generate a stable walking gait with IMU data both with and without domain randomization.
- TD3 does not perform well with minimal state spaces or with the addition of IMU data.
- Domain randomization does affect the performance of all three algorithms in a negative manner. However, in most cases the algorithms are still able to generate stable gaits comparable to policies trained without domain randomization.
- Contact sensors in the feet significantly improve performance in all three algorithms when using domain randomization.

6. Conclusion

In this paper, the performance of three state-of-the-art RL algorithms was compared with the walking gait of a quadruped robot. The performance of the three algorithms was studied on a quadruped robot simulated by modeling the robot using the MuJoCo's native MJCF modeling language. Each algorithm performance was evaluated in seven different state spaces along with addressing the simulation optimization basis (domain randomization). The performance results demonstrated that the performance of the three algorithms was dependent on the sensor configurations, *i.e.*, the state space. Without domain randomization, the SAC algorithm was able to generate walking gaits for all state spaces other than the state space which consisted only of the body quaternion. The PPO and TD3 algorithms were not able to generate walking gaits for the state spaces including the accelerometer and gyro data. The TD3 and PPO algorithms were noticed to have overshooting of the target velocity with domain randomization while SAC did not exhibit overshooting. Also, SAC had a significant performance improvement with the use of an accelerometer and gyro along with domain randomization. The performance results of the three algorithms do not present a clear winner. The results demonstrate the preference of the algorithms to state spaces. It can be seen that PPO tends to perform better with smaller state spaces while SAC excels with larger state spaces. Finally, it was shown that domain randomization does not significantly degrade policy performance in most cases for any algorithm. Even though all three algorithms can potentially be used for transfer learning on real robots, their performance needs to be evaluated on a real physical quadruped robot.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Hwangbo, J., Lee, J., Dosovitskiy, A., Bellicoso, D., Tsounis, V., Koltun, V. and Hutter, M. (2019) Learning Agile and Dynamic Motor Skills for Legged Robots. *Science Robotics*, **4**. http://arxiv.org/abs/1901.08652 https://doi.org/10.1126/scirobotics.aau5872

[2] Biswal, P. and Mohanty, P.K. (2021) Development of Quadruped Walking Robots: A review. *Ain Shams Engineering Journal*, **12**, 2017-2031. https://doi.org/10.1016/j.asej.2020.11.005 https://www.sciencedirect.com/science/article/pii/S2090447920302501

[3] Radford, N.A., Strawser, P., Hambuchen, K., Mehling, J.S., Verdeyen, W.K., Donnan, A.S., Holley, J., Sanchez, J., Nguyen, V., Bridgwater, L., Berka, R., Ambrose, R., Myles Markee, M., Fraser-Chanpong, N.J., McQuin, C., Yamokoski, J.D., Hart, S., Guo, R., Parsons, A., Wightman, B., Dinh, P., Ames, B., Blakely, C., Edmondson, C., Sommers, B., Rea, R., Tobler, C., Bibby, H., Howard, B., Niu, L., Lee, A., Conover, M., Truong, L., Reed, R., Chesney, D., Platt Jr., R., Johnson, G., Fok, C.-L., Paine, N., Sentis, L., Cousineau, E., Sinnet, R., Lack, J., Powell, M., Morris, B., Ames, A. and Akinyode, J. (2015) Valkyrie: Nasa's First Bipedal Humanoid Robot. *Journal of Field Robotics*, **32**, 397-419. https://doi.org/10.1002/rob.21560 https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21560

[4] Bledt, G., Powell, M.J., Katz, B., Carlo, J.D., Wensing, P.M. and Kim, S. (2018) MIT Cheetah 3: Design and Control of a Robust, Dynamic Quadruped Robot. 2018 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Madrid, 1-5 October 2018, 2245-2252. https://doi.org/10.1109/IROS.2018.8593885

[5] Hutter, M., Gehring, C., Jud, D., Lauber, A., Bellicoso, C. D., Tsounis, V., Hwangbo, J., Bodie, K., Fankhauser, P., Bloesch, M., Diethelm, R., Bachmann, S., Melzer, A. and Hoepflinger, M. A. (2016) ANYmal—A Highly Mobile and Dynamic Quadrupedal Robot. 2016 *IEEE*/*RSJ International Conference on Intelligent Robots and Systems*, Daejeon, 9-14 October 2016, 38-44. https://doi.org/10.1109/IROS.2016.7758092

[6] Sutton, R.S. and Barto, A.G. (2018) Reinforcement Learning: An Introduction. 2nd Edition, The MIT Press, Cambridge. http://incompleteideas.net/book/the-book-2nd.html

[7] Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O. (2017) Proximal Policy Optimization Algorithms. ArXiv: 1707.06347. http://arxiv.org/abs/1707.06347

[8] Haarnoja, T., Zhou, A., Abbeel, P. and Levine, S. (2018) Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. ArXiv: 1801.01290. http://arxiv.org/abs/1801.01290

[9] Fujimoto, S., van Hoof, H. and Meger, D. (2018) Addressing Function Approximation Error in Actor-Critic Methods. ArXiv: 1802.09477. http://arxiv.org/abs/1802.09477

[10] Schulman, J., Levine, S., Moritz, P., Jordan, M.I. and Abbeel, P. (2015) Trust Region Policy Optimization. ArXiv: 1502.05477. http://arxiv.org/abs/1502.05477

[11] Schulman, J., Moritz, P., Levine, S., Jordan, M.I. and Abbeel, P. (2016) High-Dimensional Continuous Control Using Generalized Advantage Estimation. 4*th International Conference on Learning Representations*, *ICLR* 2016, San Juan, 2-4 May 2016. http://arxiv.org/abs/1506.02438

[12] Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D. (2016) Continuous Control with Deep Reinforcement Learning. 4*th International Conference on Learning Representations*, *ICLR* 2016, San Juan, Puerto Rico, 2-4 May 2016. http://arxiv.org/abs/1509.02971

[13] Todorov, E., Erez, T. and Tassa, Y. (2012) Mujoco: A Physics Engine for Model-Based Control. 2012 *IEEE*/*RSJ International Conference on Intelligent Robots and Systems*, Vilamoura-Algarve, 7-12 October 2012, 5026-5033. https://doi.org/10.1109/IROS.2012.6386109

[14] Muratore, F., Gienger, M. and Peters, J. (2021) Assessing Transferability from Simulation to Reality for Reinforcement Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **43**, 1172-1183. http://arxiv.org/abs/1907.04685 https://doi.org/10.1109/TPAMI.2019.2952353

[15] Tan, J., Zhang, T., Coumans, E., Iscen, A., Bai, Y., Hafner, D., Bohez, S. and Vanhoucke, V. (2018) Sim-to-Real: Learning Agile Locomotion for Quadruped Robots. ArXiv: 1804.10332. http://arxiv.org/abs/1804.10332 https://doi.org/10.15607/RSS.2018.XIV.010

[16] Zhao, W., Queralta, J.P. and Westerlund, T. (2020) Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: A Survey. 2020 *IEEE Symposium Series on Computational Intelligence*, Canberra, 1-4 December 2020, 737-744. https://arxiv.org/abs/2009.13303 https://doi.org/10.1109/SSCI47803.2020.9308468

[17] Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W. and Abbeel, P. (2017) Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World. 2017 *IEEE*/*RSJ International Conference on Intelligent Robots and Systems*, Vancouver, 24-28 September 2017, 23-30. http://arxiv.org/abs/1703.06907

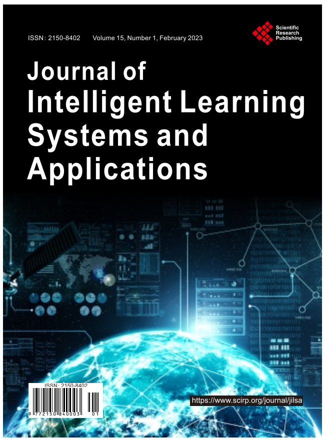
https://doi.org/10.1109/IROS.2017.8202133

[18] Pinto, L., Andrychowicz, M., Welinder, P., Zaremba, W. and Abbeel, P. (2017) Asymmetric Actor Critic for Image-Based Robot Learning. ArXiv: 1710.06542. http://arxiv.org/abs/1710.06542 https://doi.org/10.15607/RSS.2018.XIV.008

[19] Rudin, N., Hoeller, D., Reist, P. and Hutter, M. (2021) Learning to Walk in Minutes Using Massively Parallel Deep Reinforcement Learning. ArXiv: 2109.11978. https://arxiv.org/abs/2109.11978

[20] Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M. and Dormann, N. (2021) Stable-Baselines3: Reliable Reinforcement Learning Implementations. *Journal of Machine Learning Research*, **22**, 1-8. http://jmlr.org/papers/v22/20-1364.html

Call for Papers



Journal of Intelligent Learning Systems and Applications

ISSN Print: 2150-8402 ISSN Online: 2150-8410

https://www.scirp.org/journal/jilsa

The ***Journal of Intelligent Learning Systems and Applications*** **(JILSA)** is a peer reviewed international journal with a key objective to provide the academic and industrial community a medium for presenting original cutting-edge research related to intelligent learning systems and their applications. JILSA invites authors to submit their original and unpublished work that communicates current research on intelligent learning systems both in the theoretical and methodological aspects, as well as various applications in real-world applications.

Papers are invited on the topics including, but not limited to:

- Approximate Dynamic Programming
- Autonomic Computing
- Autonomous Learning Systems
- Bio-Inspired Learning Method
- Clustering
- Cyber-Physical Systems
- Evolutionary Computation
- Filter Bank
- Fuzzy and Rough Set
- General Theory on Intelligent Learning Systems
- Intelligent Applications
- Learning Control Systems
- Multi-Agent Learning
- Multiple Models Strategy
- Neural Networks
- Pattern Recognition Based on Learning Techniques
- Predictive Control
- Recommendation System
- Reinforcement Learning
- Robotics
- Smart Grid Techniques and Applications
- Social Network Analysis
- Statistical Learning Theory
- Supervised Learning
- Time-Frequency Representation
- Unsupervised Learning
- User Behavior Modelling
- Wavelet Theory

Editor-in-Chief

Prof. Wei-Chiang Hong	Asia Eastern University of Science and Technology, Chinese Taipei

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

https://www.scirp.org/journal/jilsa **E-Mail: jilsa@scirp.org**