

**FRONTEIRAS EMERGENTES NO APRENDIZADO PROFUNDO EM 2024-2025:
ARQUITETURAS TRANSFORMADORAS APLICADAS A DIVERSOS DOMÍNIOS
DO CONHECIMENTO**

 <https://doi.org/10.56238/arev7n4-183>

Data de submissão: 16/03/2025

Data de publicação: 16/04/2025

Danilo Gualberto Zavarize

Engenheiro Ambiental, Doutorando em Ciência dos Materiais e Especialista em Estatística Aplicada
Universidade Federal do Maranhão – UFMA e Instituto Federal do Maranhão – IFMA
<https://orcid.org/0009-0000-0714-7062>
<http://lattes.cnpq.br/0771028976036909>
danilo.zavarize@discente.ufma.br

Samuel Gonçalves Alves

Contador, Cientista da Computação e Especialista em Estatística Aplicada as Ciências
Centro de Ciências Exatas e Tecnologia, Universidade Federal do Maranhão, Imperatriz – 65080-
805, Brasil.
Universidade Federal do Maranhão – UFMA
<https://orcid.org/0009-0007-7936-6919>
<http://lattes.cnpq.br/0956882647769768>
samuelshmyt@gmail.com

Antônio Duda Oliveira da Silva

Matemático e Administrador, Especialista em Matemática Financeira e Estatística
Centro de Ciências de Imperatriz, Universidade Federal do Maranhão, Imperatriz – 65915-240,
Brasil
Universidade Federal do Maranhão – UFMA
<https://orcid.org/0009-0008-3674-3515>
<http://lattes.cnpq.br/3299470618963821>
antonioduda-sitio@hotmail.com

Joelson Gomes de Oliveira

Biólogo, Mestrando em Agricultura e Ambiente
Centro de Ciências de Imperatriz, Universidade Federal do Maranhão, Imperatriz – 65915-240,
Brasil
Universidade Federal do Maranhão – UFMA
<https://orcid.org/0009-0006-8718-3250>
<http://lattes.cnpq.br/2710429991772101>
joe.biologo@gmail.com

Haroldo Gomes Barroso Filho

Engenheiro de Computação e Doutor em Ciência da Computação

Centro de Ciências de Imperatriz, Universidade Federal do Maranhão, Imperatriz – 65915-240,
Brasil

Universidade Federal do Maranhão – UFMA

<https://orcid.org/0000-0003-3256-302X>

<http://lattes.cnpq.br/5326436049008095>

haroldo.gbf@ufma.br

Daniel Duarte Costa

Centro de Ciências de Imperatriz, Universidade Federal do Maranhão, Imperatriz – 65915-240,
Brasil

Universidade Federal do Maranhão – UFMA e Instituto Federal do Maranhão – IFMA

<https://orcid.org/0009-0000-0714-7062>

<http://lattes.cnpq.br/5830574741456955>

daniel.dc@ufma.br

RESUMO

O aprendizado profundo continua a revolucionar indústrias ao abordar desafios complexos por meio de arquiteturas neurais inovadoras, refletindo um dos campos mais dinâmicos e promissores da ciência da computação moderna. Desde os primeiros modelos de redes neurais artificiais até as sofisticadas estruturas que dominam atualmente o cenário tecnológico, o aprendizado profundo tem evoluído rapidamente, impulsionado pela disponibilidade de grandes volumes de dados, avanços em hardware computacional e a crescente demanda por soluções automatizadas e inteligentes. Neste trabalho de revisão, conduzido nas bases de dados Google Acadêmico, arXiv, Scopus, e SciELO, exploramos cinco modelos emergentes em 2024 e início de 2025 – *Temporal Convolutional Networks*, *Kolmogorov-Arnold Networks*, *Quantum-Inspired Recurrent Networks*, *Deep Reinforcement Learning* e *Generative Adversarial Networks* para apontar mecanismos de funcionamento e analisar o impacto transformador dessas ferramentas em diversos campos do conhecimento científico. Ao agrupar essas arquiteturas de acordo com suas aplicações em saúde, criação de conteúdo, sistemas autônomos, análise de séries temporais e detecção de anomalias, fornecemos uma visão abrangente de suas capacidades, forças e possíveis limitações. Assim, elucida-se as principais características utilizadas na seleção das técnicas mais adequadas para necessidades específicas, além de destacar oportunidades para avanços futuros.

Palavras-chave: Arquiteturas Neurais. Aplicações Multidomínio. Inovação Tecnológica.

1 INTRODUÇÃO

O aprendizado profundo emergiu como uma das áreas mais transformadoras da inteligência artificial (IA), impulsionando avanços em uma ampla gama de setores. Desde a automação industrial até a medicina personalizada, as redes neurais profundas têm demonstrado capacidade sem precedentes para resolver problemas complexos que antes eram considerados intratáveis (GANAIE *et al.* 2022; MU e ZENG, 2019). Essa evolução foi possibilitada por avanços em hardware, disponibilidade de grandes volumes de dados e o desenvolvimento de novas arquiteturas neurais especializadas (KIM *et al.* 2021). No entanto, à medida que as demandas das indústrias se tornam mais sofisticadas, também cresce a necessidade de modelos que sejam não apenas precisos, mas também adaptáveis, escaláveis e interpretáveis (ALYAMI *et al.* 2024).

Entre 2024 e os primeiros meses de 2025, tem-se observado um salto qualitativo no design de arquiteturas de aprendizado profundo, com foco em abordagens híbridas e inspiradas por princípios de outras áreas, como física quântica, teoria de grafos e probabilidade (KIM *et al.* 2021; ALYAMI *et al.* 2024). Essas inovações refletem uma tendência clara: a criação de modelos que combinam alta performance com flexibilidade para lidar com desafios específicos de domínio (ALYAMI *et al.* 2024). Além disso, houve um aumento significativo no uso de técnicas como aprendizado federado, que permite que modelos sejam treinados em dados descentralizados sem comprometer a privacidade dos usuários (DEMELIUS *et al.* 2025), particularmente relevante em setores como saúde, onde a proteção de dados sensíveis é crucial. Outra tendência importante é o desenvolvimento de modelos mais eficientes e escaláveis, implicações que são essenciais para a adoção em larga escala de soluções de IA em diferentes indústrias (SOORI *et al.*, 2024).

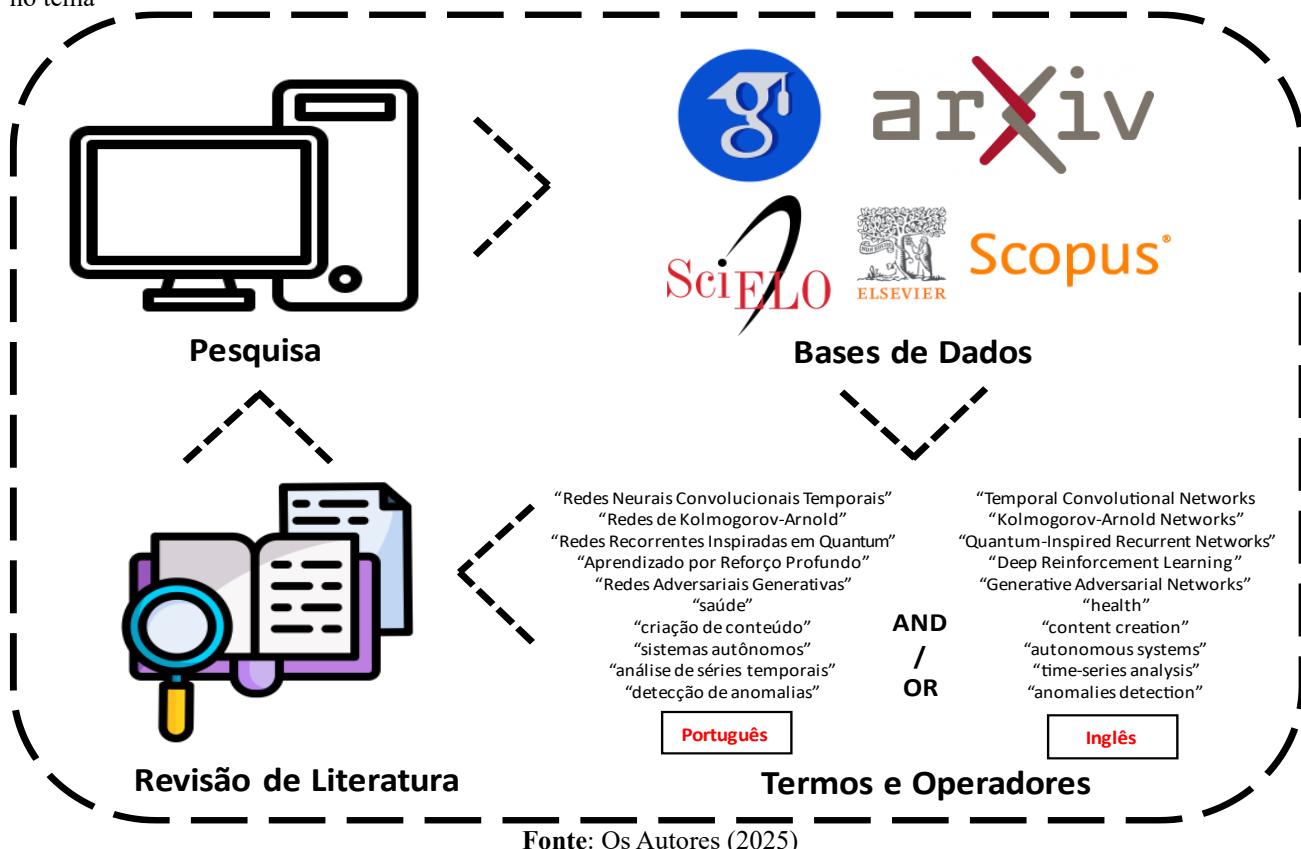
Nesse contexto, a explicabilidade dos modelos de aprendizado profundo e a junção da IA com a “computação de borda” também ganharam dentro desse período com o crescimento da complexidade dos modelos e a necessidade de transparência, compreensão e otimização para os diferentes tipos de usuários (HOSAIN *et al.*, 2024). A *Explainable AI* (ou “explicabilidade da IA”) visa justamente isso: tornar as decisões dos modelos mais claras e interpretáveis, algo especialmente importante em setores como saúde e finanças, onde as decisões tomadas pelos modelos podem ter impactos significativos na vida das pessoas (SAARELA e PODGORELEC, 2024). Além disso, o *Edge AI* (ou “computação na borda”) também está se tornando mais popular ao trazer a inteligência diretamente para os dispositivos, com o objetivo de reduzir a latência, melhorar a privacidade e diminuir os requisitos de largura de banda (GILL *et al.*, 2024), permitindo que esses dispositivos tomem decisões de forma mais rápida e eficiente (GILL *et al.*, 2024; ALAM *et al.*, 2024).

Diante dessas inovações, esta revisão busca fornecer um compilado das arquiteturas de IA mais relevantes lançadas ou significativamente avançadas em 2024 e primeiro semestre de 2025. O objetivo, portanto, foi explorar sobre a teoria desses modelos e como eles são aplicados em diferentes áreas do conhecimento, desde aplicações na saúde humana até sistemas autônomos em automóveis e robôs, e na criação de mídias. Para isso, organizamos a discussão desse estudo em categorias temáticas, que destacou as contribuições específicas de cada arquitetura explorada e suas implicações práticas. Ao final, esperou-se não apenas abranger o estado atual do campo, como também identificar lacunas e oportunidades para futuros desenvolvimentos.

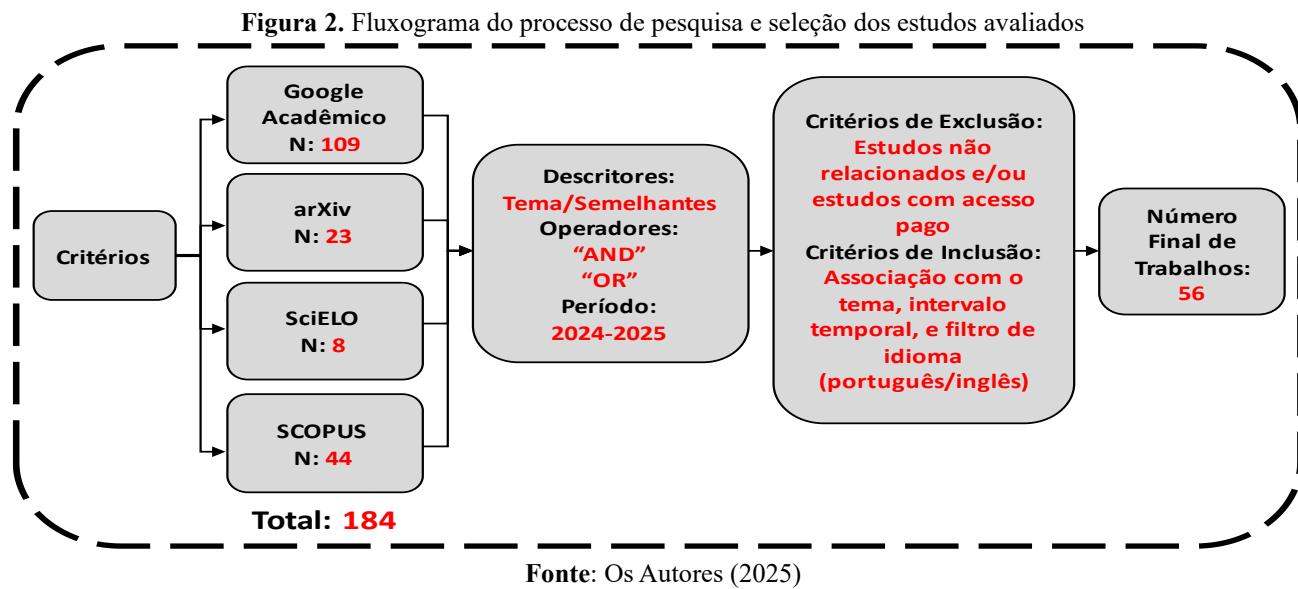
2 METODOLOGIA

Realizou-se pesquisa bibliográfica (SNYDER, 2019) de natureza qualitativa e de característica não-sistemática (FORNARI e PINHO, 2022) nas bases *Google Acadêmico*, *arXiv*, *Scopus* e *SciELO* de modo a embasar o presente estudo desenvolvido na forma de revisão narrativa (ROTHER, 2007), como ilustrado no esquema da Figura 1.

Figura 1. Esquema de condução da pesquisa e critérios de seleção de estudos nas bases de dados para a revisão de literatura no tema



Para a pesquisa, empregou-se os operadores Booleanos “AND” e “OR” e os termos associados as ferramentas de *deep learning* em português e inglês, assim como as aplicações nos campos do conhecimento selecionados, conforme indicado na Figura 2.



3 RESULTADOS E DISCUSSÃO

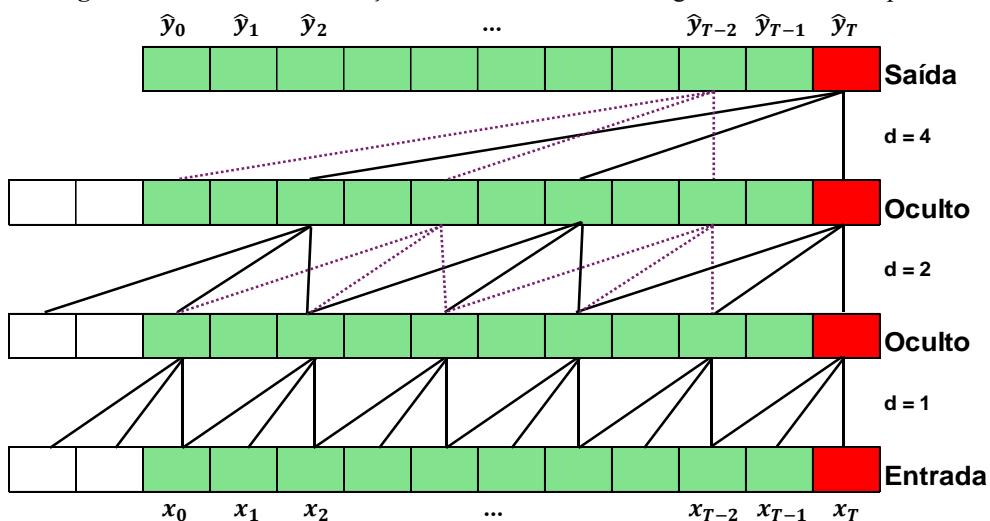
A seguir apresentamos os aspectos teóricos das arquiteturas de aprendizado profundo emergentes entre 2024 e início de 2025 e suas aplicações em diversos campos do conhecimento, abordando ainda questões sobre as vantagens, desvantagens, limitações e projeções futuras.

3.1 REDES CONVOLUCIONAIS TEMPORAIS (TCN)

3.1.1 Aspectos Teóricos da Arquitetura

As TCN são arquiteturas de aprendizado profundo projetadas especificamente para modelar dependências temporais em dados sequenciais (ZEGHINA *et al.*, 2024). Ao contrário das redes neurais recorrentes tradicionais, que processam dados sequencialmente e usualmente enfrentam dificuldades com dependências de longo prazo por problemas como gradientes evanescentes ou explosivos, as TCN utilizam uma abordagem convolucional para melhor processar informações temporais. Essa arquitetura emprega convoluções unidimensionais ao longo da dimensão temporal (Figura 3), o que permite fazer a captura de padrões e tendências em dados obtidos na perspectiva de séries temporais com maior eficiência e estabilidade (ZEGHINA *et al.*, 2024; THAPA e LEE, 2024).

Figura 3. Modelo de convoluções unidimensionais ao longo da dimensão temporal



Fonte: Os Autores, 2025 – Adaptado de ZEGHINA *et al.* (2024)

A fundamentação teórica das TCN está na capacidade das operações convolucionais em extrair características locais mantendo o contexto global pela expansão do campo receptivo (ZEGHINA *et al.*, 2024), decorrente do empilhamento de múltiplas camadas convolucionais com taxas de dilatação crescentes que permite à rede "visualizar" eventos mais distantes no passado sem perder resolução ou aumentar significativamente a complexidade computacional (ZEGHINA *et al.*, 2024). Ressalta-se que o uso de conexões residuais nessas arquiteturas mitiga problemas com gradientes evanescentes que impactam no desempenho geral, o que permite e facilita o treinamento de redes naturalmente mais profundas (THAPA e LEE, 2024).

No entanto, as arquiteturas TCN possuem limitações notáveis que tem direcionado as melhorias atuais: (a) exigem grandes volumes de dados para alcançar um desempenho ótimo por conta do alto número de parâmetros em camadas convolucionais, e (b) o processamento conduzido em paralelo pode impedir a incorporação de informações futuras, o que é desvantajoso em tarefas que demandam um contexto bidirecional (ZEGHINA *et al.*, 2024; THAPA e LEE, 2024).

Desse modo, um dos principais avanços das TCN em 2024-2025 foi o uso de convoluções causais (LI *et al.*, 2023; THAPA e LEE, 2024), que garantem que previsões no instante t se baseiem apenas em informações de t ou momentos anteriores para preservar os aspectos de causalidade temporal. Ainda, a aplicação de convoluções dilatadas (LI *et al.*, 2023) para expansão do campo receptivo da rede sem aumentar o número de parâmetros possibilita a captura mais eficaz de dependências de longo prazo. A combinação dessas técnicas permite que a arquitetura TCN processe melhor as sequências em paralelo, resultando em tempos de treinamento mais rápidos em comparação com modelos baseados em redes neurais recorrentes tradicionais (LI *et al.*, 2023; THAPA e LEE, 2024).

3.1.2 Aplicações Práticas de Arquiteturas TCN

Na prática, as arquiteturas TCN tem demonstrado potencial significativo em diversas áreas do conhecimento, como na saúde (WANG *et al.*, 2025), em sistemas autônomos (SADID e ANTONIOU, 2024) e detecção de anomalias (REN *et al.*, 2024).

No campo da saúde, Wang *et al.* (2025) desenvolveram um novo algoritmo chamado CS-TCN (*Channel Selection-Based Temporal Convolutional Network*) para a detecção personalizada de crises epilépticas em pacientes a partir de sinais de eletroencefalograma (EEG). Essa estratégia combina um método de seleção de canais baseado em escore expandido de Fisher com a arquitetura TCN, capaz de captar eficazmente as dependências temporais nos sinais de EEG, ao mesmo tempo que também reduz a complexidade do modelo e os requisitos computacionais. Essa abordagem dos autores possibilitou alcançar métricas de desempenho notáveis, com sensibilidades de 98,56% e 98,88% nos conjuntos de dados CHB-MIT e Siena, respectivamente, que ainda destacam o fato desse novo algoritmo superar os métodos existentes de detecção de crises com poucos canais. Isto é, a capacidade do algoritmo CS-TCN de manter um alto desempenho mesmo com um número reduzido de canais de EEG (de 1 a 5) é particularmente relevante no contexto das aplicações de saúde, pois torna essa abordagem relevante para o desenvolvimento de dispositivos portáteis e vestíveis que possam detectar crises mais facilmente e serem usados em ambientes domésticos. Os desafios e limitações abordados pelos autores foram a variabilidade entre pacientes e a capacidade de detecção em tempo real, que pavimentam o caminho para soluções futuras mais personalizadas e adaptativas de cuidados de saúde.

No campo dos sistemas autônomos, Sadid *et al.* (2024) propuseram um modelo baseado em uma rede neural gráfica espaço-temporal dinâmica (STGCN) que integra arquiteturas TCN na captura e previsão de dependências temporais das trajetórias dos veículos autônomos. Assim, ao combinar um mecanismo de convolução gráfica espacial que modela as interações entre os veículos em uma cena de tráfego, Sadid e colaboradores puderam extrair padrões temporais a partir de observações históricas de movimento, o que resultou em melhorias significativas na precisão da previsão de trajetórias quando comparada a modelos existentes durante experimentos realizados com o conjunto de dados *HighD*. Os autores ainda destacam: (a) a capacidade do algoritmo em processar simultaneamente as trajetórias de todos os veículos em uma cena de tráfego, algo que métodos baseados em *Long Short-Term Memory* (LSTM), por exemplo, geralmente não conseguem fazer de forma eficiente; (b) os campos receptivos flexíveis e as conexões residuais, que permitem ao modelo manter uma memória de longo prazo enquanto opera com alta eficiência computacional; e (c) a integração do algoritmo TCN com Redes Convolucionais Gráficas (GCN) para capturar as interações espaciais entre os veículos e também extraer informações temporais relevantes na previsão de estados futuros.

No campo da detecção de anomalias, Ren *et al.* (2024) propuseram um *framework* inovador chamado *Mirror Temporal Graph Autoencoder* (MTGAE) que explora o uso de TCN no contexto da detecção de anomalias em séries temporais de tráfego com foco em sistemas de transporte inteligentes. Esse framework surgiu a partir da combinação de TCN adaptativas e Redes Convolucionais Gráficas (GCN) capaz de capturar correlações espaço-temporais complexas em redes de tráfego. O principal diferencial está na criação de um Módulo de Convolução Temporal Espelhada (MTCM) que amplia o campo perceptivo e adapta-se melhor às características específicas dos dados de tráfego. O framework ainda inclui o módulo *GCGRU CELL*, que utiliza funções de *kernel* gaussiano para mapear os dados em um espaço de alta dimensão e permitir a identificação de anomalias ocultas nas interdependências complexas da rede de tráfego. De acordo com os autores, os testes experimentais no conjunto de dados NYC demonstraram que MTGAE supera outros modelos avançados de detecção de anomalias, vista a capacidade do MTCM em lidar com as limitações tradicionais das TCN quando se trata de redes de transporte dinâmicas. Em outras palavras, os autores sugerem que enquanto TCN tradicionais podem ter dificuldades na captura de informações contextuais globais e nas dinâmicas com campo perceptivo limitado, o MTCM se sobressai ao explorar associações ocultas entre nós e aprender relacionamentos complexos existentes e dependências nos dados de tráfego, tornando-se mais robusto na detecção de anomalias tanto espaciais quanto temporais.

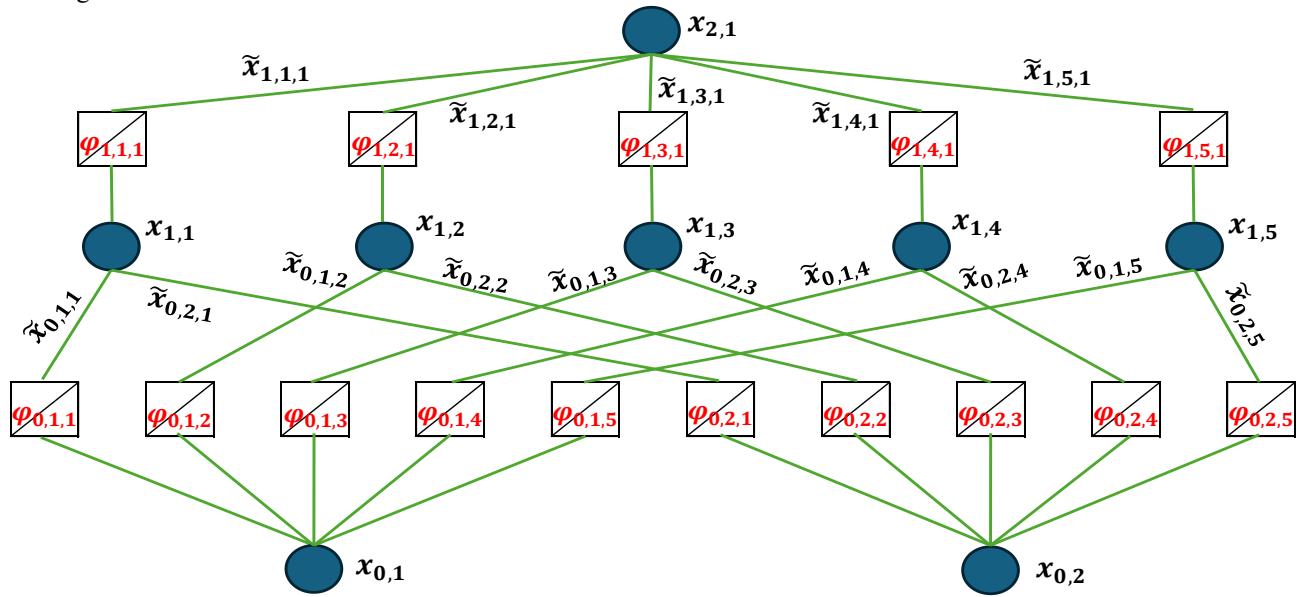
3.2 REDES DE KOLMOGOROV-ARNOLD (KAN)

3.2.1 Aspectos Teóricos da Arquitetura

KAN são arquiteturas inspiradas no teorema da representação de Kolmogorov-Arnold (KILANI, 2024), em que uma função multivariada contínua qualquer pode ser decomposta em uma soma finita de funções univariadas e operações de adição. Portanto, KAN são ideais para a construção de redes capazes de aproximar funções complexas com auxílio de componentes simplificados (KILANI, 2024; ALONSO, 2025). Ao contrário das Redes Neurais Multicamadas (MLP), que utilizam funções de ativação fixas em nós (neurônios), as KAN usam funções de ativação aprendíveis (ex.: *B-splines*) em arestas (também denominados pesos).

Ou seja, nessa arquitetura cada parâmetro de peso é substituído por uma função univariada, que elimina a necessidade de matrizes de pesos lineares, ao passo que os nós realizam apenas operações de soma e as arestas participam ativamente do aprendizado, permitindo que a rede Kolmogorov-Arnold faça aproximações de funções de alta dimensão fazendo uso de menos parâmetros e conferindo maior interpretabilidade (BASINA *et al.*, 2024) (Figura 4).

Figura 4. Fluxo de decomposição de funções multivariadas em somas finitas de funções univariadas na rede neural de Kolmogorov-Arnold



Fonte: Os Autores, 2025 - Adaptado de KILANI (2024)

Dentre as vantagens dessa arquitetura, ressalta-se que essa capacidade para a decomposição de funções multivariadas em somas de funções univariadas mitiga problemas de dimensionalidade em tarefas voltadas ao machine learning (BASINA *et al.*, 2024; GAO e TAN, 2024), bem como promove mais precisão com redes menores, e favorece maior interpretabilidade em razão de suas funções de ativação aprendíveis que podem ser visualizadas e analisadas, tornando essa arquitetura útil em vários domínios científicos onde a transparência é crítica (BASINA *et al.*, 2024).

Contudo, arquiteturas KAN enfrentam limitações, como o fenômeno de *overhead* computacional causado pelas funções de ativação baseadas em *splines*, o que resulta em tempos de treinamento e inferência significativamente mais lentos se comparados às MLP (BASINA *et al.*, 2024). Além disso, uma outra desvantagem está associada a escalabilidade da arquitetura para entradas de alta dimensão, já que a profundidade e complexidade da rede crescem exponencialmente com a dimensionalidade da função (BASINA *et al.*, 2024; IBRAHUM *et al.*, 2024).

3.2.2 Aplicações Práticas de Arquiteturas KAN

Na prática, arquiteturas KAN já foram aplicadas em análises de séries temporais (XU *et al.*, 2024), na predição no aprendizado de padrões em sinais fisiológicos multivariados (DONG *et al.*, 2024), e na detecção de anomalias em sistemas industriais ou tráfego de rede por aproximação de distribuições multivariadas irregulares (ABUDUREXITI *et al.*, 2025).

Na análise de séries temporais, Xu *et al.* (2024) introduz duas variantes: T-KAN e MT-KAN, que são capazes de combinar poder preditivo com interpretabilidade, desafios que são comuns em

métodos tradicionais de previsão de séries temporais. A arquitetura T-KAN foi projetada para detectar mudanças ou deslocamentos nos conceitos (ou relações) em séries temporais univariadas, utilizando regressão simbólica para explicar as relações não-lineares entre previsões e os passos temporais anteriores. Por outro lado, a arquitetura MT-KAN focou em explorar interdependências complexas entre variáveis em séries temporais multivariadas, melhorando significativamente a precisão preditiva. Os experimentos realizados pelos autores em bancos de dados financeiros demonstraram que ambas as variantes, mesmo com arquiteturas simples (2 camadas e 5 neurônios ocultos), alcançam desempenho comparável ou até mesmo superior àqueles modelos baseados em MLP tradicionais, oferecendo maior transparência no processo de decisão. O diferencial dessas arquiteturas, portanto, reside na substituição de pesos lineares tradicionais por funções univariadas parametrizadas por *splines* que permitiu que as KAN aprendessem padrões dinâmicos nos dados e se adaptassem a mudanças no ambiente.

Na área da saúde, Dong *et al.* (2024) propuseram o modelo TCKAN para a previsão do risco de mortalidade por sepse em pacientes críticos, que utiliza arquitetura KAN na captura de características dinâmicas e não-lineares de dados médicos e integra informações temporais, constantes e diagnósticas em um vetor de características unificado. Os experimentos realizados nos conjuntos de dados MIMIC-III e MIMIC-IV demonstraram que o TCKAN supera significativamente os métodos existentes, com diferencial importante na capacidade do modelo de lidar com séries temporais irregulares por meio do módulo GRU-D que melhora a interpretabilidade e a capacidade de reconhecimento de características. Além disso, estudos de ablação conduzidos pelos autores confirmaram a importância das KAN no modelo ao mostrar que sua substituição por MLP leva a uma queda no desempenho.

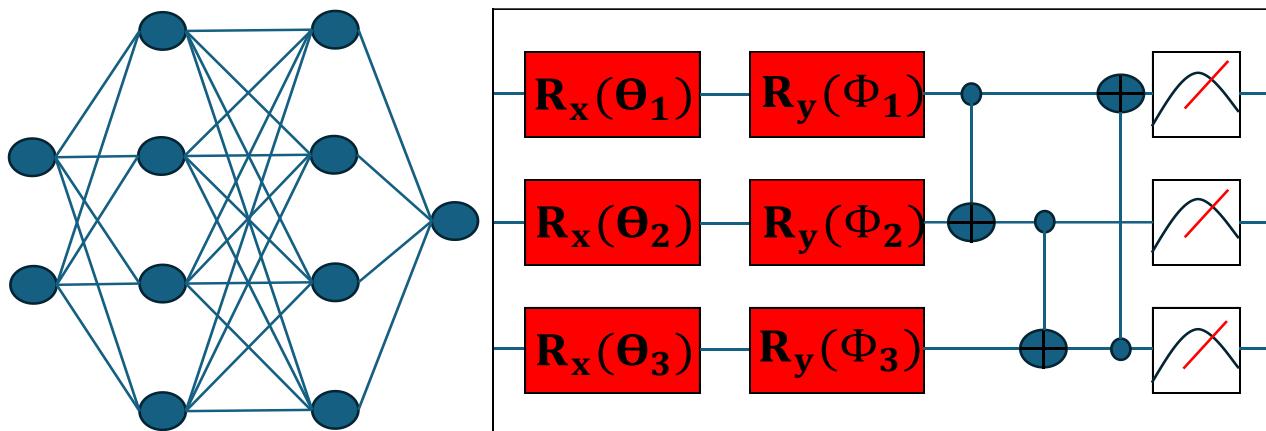
Na detecção de anomalias, Abudurexiti *et al.* (2025) propuseram um *framework* específico não-supervisionado e explicável para detecção de anomalias em sistemas de Internet Industrial das Coisas (IIoT), integrando arquiteturas KAN à *Variational Auto-Encoders* (VAE) para melhorar a capacidade de captura de padrões complexos em séries temporais multivariadas, conferindo maior robustez e interpretabilidade. Essa característica de captura de padrões complexos é particularmente relevante no contexto de IIoT, visto que os dados frequentemente apresentam anomalias relacionadas a ruído e alta dimensionalidade. A combinação da arquitetura KAN com o *Convolucional Block Attention Module* (CBAM) integrado ao VAE melhorado para extração de características locais é um outro diferencial reportado pelos autores, pois permite que o modelo identifique eficientemente anomalias em dados de séries temporais mesmo em cenários com poucos ou nenhum dado rotulado.

3.3 REDES RECORRENTES INSPIRADAS EM QUANTUM (QINN)

3.3.1 Aspectos Teóricos da Arquitetura

QINN representam uma nova classe de arquiteturas de aprendizado profundo ao integrarem os princípios da computação quântica (MALHOTRA, 2024), como a superposição, o emaranhamento e o paralelismo quântico, com capacidade de processamento sequencial similar àquela de redes neurais recorrentes (RNN). Arquiteturas QINN utilizam mecanismos quantum para melhoraria da eficiência computacional, maior retenção de memória, e maior capacidade de modelar dependências temporais complexas em dados sequenciais (PANDURANGAN *et al.*, 2024) (Figura 5).

Figura 5. Esquema de superposição, emaranhamento e o paralelismo quântico dos algoritmos QINN



Fonte: Os Autores, 2025 – Adaptado de MALHOTRA (2024)

Em outras palavras, arquiteturas QINN empregam estados quânticos para codificar informações por meio da superposição durante o processamento simultâneo de múltiplas entradas, possibilitando a exploração de um vasto espaço de soluções de forma mais eficiente. Ainda, o emaranhamento quântico atua como facilitador das correlações entre dados distantes, ampliando a capacidade de captura das dependências de longo alcance (PANDURANGAN *et al.*, 2024; CHEN e LUO, 2024). Essa habilidade está atrelada aos componentes *Quantum-Adaptive-Computation-Time* e *Quantum-Dynamic-Routing*, que ajustam a dinâmica dos passos de processamento em virtude da complexidade das entradas e, por isso, viabiliza a otimização do fluxo de informações que melhora a adaptabilidade às sequências de comprimento variável (CHEN e LUO, 2024). Tais princípios também são observados em redes neurais dotadas tensores em rede (TIWARI *et al.*, 2024), que são capazes de decompor dados de alta dimensão em componentes de menor dimensão, diminuindo custo computacional enquanto preserva informações críticas. Arquiteturas QINN também mitigam problemas com gradientes evanescentes, comum em RNN clássicas, por meio de mecanismos de

gating inspirados em *quantum*, que estabilizam o fluxo de gradientes durante o treinamento (TIWARI *et al.*, 2024).

As vantagens de arquiteturas QINN estão no desempenho superior às redes neurais tradicionais, principalmente em tarefas que exigem retenção de memória e processamento paralelo, como previsão de séries temporais e processamento de linguagem natural (AHMAD e JAS, 2025). Contudo, o desafio atrelado a dependência de longo alcance em simulações quânticas introduz *overhead* computacional, o que tende a tornar o treinamento mais lento em comparação aos modelos clássicos. Pontua-se ainda o desafio da escalabilidade de arquiteturas QINN para grandes conjuntos de dados ainda, limitado em virtude do crescimento exponencial das dimensões dos estados quânticos, bem como restrições no que tange a interpretabilidade do modelo, visto que operações quânticas podem complicar a análise de representações internas (LABHANE *et al.*, 2024).

3.3.2 Aplicações Práticas de Arquiteturas QINN

Apesar dessas limitações, as arquiteturas QINN apresentam aplicações onde modelos clássicos frequentemente falham. Como exemplos, pode-se citar a predição de séries temporais (HONG *et al.*, 2024) e, na saúde, por meio da criação de robôs humanoides autônomos visando o cuidado compassivo de pacientes hospitalares (HERNANDEZ, 2024).

Hong *et al.* (2024) propuseram um modelo híbrido que combinou arquitetura QINN com redes neurais profundas (CNN paralelas e LSTM), otimizado por meio de um algoritmo de Otimização por Enxame de Partículas Quânticas (QPSO), aplicados na análise de séries temporais voltadas à previsão antecipada de velocidade dos ventos. Esse modelo híbrido faz uso de pesos complexos na arquitetura QINN capazes de expandir a dimensionalidade do espaço de busca e melhorar a precisão das previsões em comparação com modelos tradicionais que utilizam pesos reais. Os resultados alcançados pelos autores mostram que o modelo proposto supera abordagens similares como CNN-LSTM-FC e CNN-LSTM-CVNN, em termos de métricas como MAE, RMSE e R², que validam a eficácia da QINN para problemas de previsão espacial e temporal. Além disso, o uso de três camadas CNN paralelas no modelo para extração de características dos dados de velocidade dos ventos mostrou-se mais eficaz do que uma única camada CNN, destacando a importância da arquitetura proposta. O papel do QPSO, por sua vez, também foi crucial ao ajustar os hiperparâmetros do modelo, que conferiu convergência mais rápida e resultados melhores do que aqueles obtidos em algoritmos de otimização clássicos.

No campo da saúde, Hernandez (2024) propôs um modelo de simulação multissensorial que integra conceitos quânticos de emaranhamento e superposição que aprimoraram as interações humano-robô em ambientes hospitalares, no que tangem as necessidades emocionais e físicas dos pacientes,

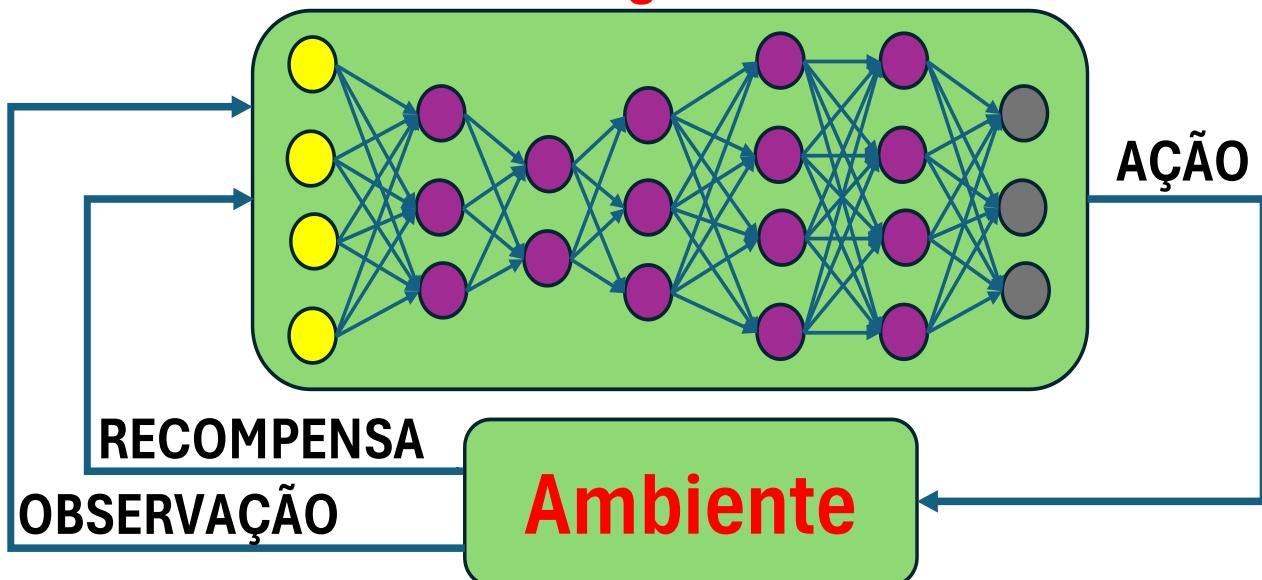
visando desenvolver robôs humanoides autônomos para cuidados compassivos e uma comunicação mais empática e personalizada. O autor ilustra como o sistema utiliza um esquema de memória circular com capacidade para 100 experiências e uma taxa de aprendizado adaptativa para ajustar suas ações com base em interações passadas, capaz de incorporar incertezas para refletir dinâmicas do mundo real. O diferencial do estudo é, portanto, a forma como os princípios quânticos são aplicados ao modelo de comunicação dinâmica entre o sistema de saúde, o robô e o paciente. Isto é, uma matriz de densidade quântica usada para avaliar a fidelidade das informações recebidas pelos pacientes e garantir maior precisão e confiabilidade nas interações. A simulação conduzida pelo autor e seu grupo de pesquisa também demonstrou que o uso de arquitetura QINN permite que os robôs processem múltiplas entradas simultaneamente e respondam de maneira adaptativa às variações culturais, preferências individuais e necessidades específicas de cuidado.

3.4 APRENDIZADO POR REFORÇO PROFUNDO (DRL)

3.4.1 Aspectos Teóricos da Arquitetura

A arquitetura DRL é um paradigma que integra princípios de aprendizado por reforço (RL) com redes neurais profundas (DNN), permitindo que agentes assimilem melhor as políticas de tomada de decisão em ambientes que são mais complexos e dinâmicos (GAO e SCHEWEIDTMANN, 2024). Em seu núcleo, a arquitetura DRL opera com o *framework* de Processos de Decisão de Markov (MDP) (ZHANG *et al.*, 2024), em que um agente interage com o ambiente selecionando ações com base em estados observados, recebe recompensas e atualiza sua política para maximizar o retorno cumulativo ao longo do tempo (Figura 6).

Figura 6. Esquema do paradigma do aprendizado profundo por reforço
Agente



Fonte: Os Autores, 2025 - Adaptado de GAO e SCHEWEIDTMANN (2024)

A fundamentação teórica do DRL baseia-se na Equação de Optimalidade de Bellman (CHUNG *et al.*, 2024), que decompõe recursivamente o valor de um estado em recompensas imediatas e o valor descontado de estados subsequentes, permitindo a aproximação de retornos de longo prazo por meio de atualizações iterativas. Algoritmos como *Deep Q-Networks* (DQN) (LUO *et al.*, 2024), *Proximal Policy Optimization* (PPO) (PENG *et al.*, 2024) e *Twin Delayed Deep Deterministic Policy Gradient* (TD3) (SHU *et al.*, 2025) aplicam arquitetura DRL para aproximar funções de valor de ação ou políticas, característico do algoritmo PPO, superando o problema da questão sobre dimensionalidade em espaços de estados e ações de alta complexidade. Ainda, o DQN introduziu *replay* de experiência e redes-alvo para estabilizar a etapa de treinamento (LUO *et al.*, 2024), enquanto o TD3 veio para aprimorar a estabilidade em tarefas de controle contínuo, ação essa que mitiga o viés de superestimação por meio de *double Q-learning* e suavização da política-alvo (YAO *et al.*, 2024).

Um avanço crucial do DRL em 2024 foi a sua capacidade de aprender diretamente a partir de dados brutos e não estruturados (por exemplo: imagens, leituras de sensores), um avanço que eliminou a necessidade de engenharia manual de características, como demonstrado em aplicações como jogos Atari e controle robótico (YAO *et al.*, 2024). A arquitetura DRL, portanto, destaca-se em tarefas que exigem planejamento de longo prazo e adaptabilidade, em que os agentes devem equilibrar exploração e exploração para otimizar recompensas cumulativas sob incerteza.

No entanto, o principal desafio de arquiteturas DRL ainda é a ineficiência amostral, pelo fato de que o treinamento frequentemente requer uma quantidade massiva de interações ambientais,

tornando a implementação em cenários reais inviável sem um devido pré-treinamento em simulações reais. Isto é, o denominado “*sim-to-real gap*”, em que as políticas treinadas em simulações falham em generalizar para dinâmicas do mundo real, complica ainda mais a aplicação prática (YAO *et al.*, 2024). Além disso, restrições de segurança em domínios críticos como saúde, veículos autônomos etc., exigem garantias de robustez, algo que as arquiteturas DRL não tendem a suprir devido ao seu processo de aprendizado baseado em tentativa e erro (GAO e TAN, 2024; ZHANG *et al.*, 2024). A interpretabilidade, por sua vez, também tem sido um obstáculo, pois modelos de DRL frequentemente atuam como "caixas-pretas", prejudicando a confiança em aplicações de alto risco (GAO e TAN, 2024).

3.4.2 Aplicações Práticas de Arquiteturas DRL

Apesar dos desafios, DRL possuem potencial transformador em sistemas de aprendizado *end-to-end* (PAPADOPOULOS *et al.*, 2024), na otimização de planos de tratamento para doenças crônicas (YANG *et al.*, 2024a), e na identificação de eventos raros em sistemas industriais ou tráfego de rede ao recompensar desvios de padrões normais (YANG *et al.*, 2024b).

Papadopoulos *et al.* (2024) propuseram um modelo que combinou a arquitetura DRL com redes neurais profundas visando o gerenciamento de sistemas dinâmicos e não-lineares, em que os métodos tradicionais de aprendizado frequentemente enfrentam limitações. O modelo apresenta capacidade de aprender políticas de controle adaptativas que maximizam a eficiência operacional em tempo real, mesmo em ambientes com alta variabilidade e incerteza, como provado pelos experimentos realizados pelos autores, alcançando maior precisão e robustez na execução de tarefas automatizadas ao superar abordagens convencionais que se baseiam em controle preditivo por modelo (MPC). Outro diferencial importante do modelo é a integração de técnicas de exploração-explotação, que permite ao sistema equilibrar a busca por novas estratégias e a utilização daquelas já aprendidas. Além disso, a arquitetura hierárquica que divide o processo de tomada de decisão em múltiplas camadas facilita a interpretação das decisões e a adaptação a diferentes cenários operacionais. Os autores concluíram que esse avanço na arquitetura DRL é particularmente útil no setor de manufatura inteligente e otimização de cadeias de suprimentos, em que a capacidade de adaptação rápida a mudanças é essencial para manter a competitividade e a eficiência.

Yang *et al.* (2024a) propuseram o modelo *PrescDRL* que utiliza DRL para planejar prescrições herbáceas personalizadas no tratamento de doenças crônicas, como o diabetes, e se destaca por otimizar esquemas de diagnóstico e tratamento em medicina tradicional chinesa (MTC). Os autores avançaram com esse modelo pelo mesmo ter a capacidade de formular o problema de otimização de diagnóstico e tratamento como se fosse uma tarefa de aprendizado por reforço, ou seja: (a) o estado dos pacientes

é representado pela observações de sintomas, (b) as ações correspondem às prescrições herbáceas, e (c) as recompensas refletem a eficácia do tratamento ao longo do tempo. O modelo demonstrou um desempenho superior em relação a abordagens tradicionais, com melhorias significativas em precisão (40,5%) e recall (63%) na previsão de prescrições. Um outro diferencial importante se dá pelo foco em resultados de longo prazo, ao invés de maximizar recompensas imediatas, o que resulta em planos de tratamento mais razoáveis e eficazes para os pacientes. Além disso, o modelo utiliza um ambiente virtual offline baseado em dados clínicos sequenciais para simular transições de estado, permitindo uma adaptação dinâmica às condições individuais dos pacientes.

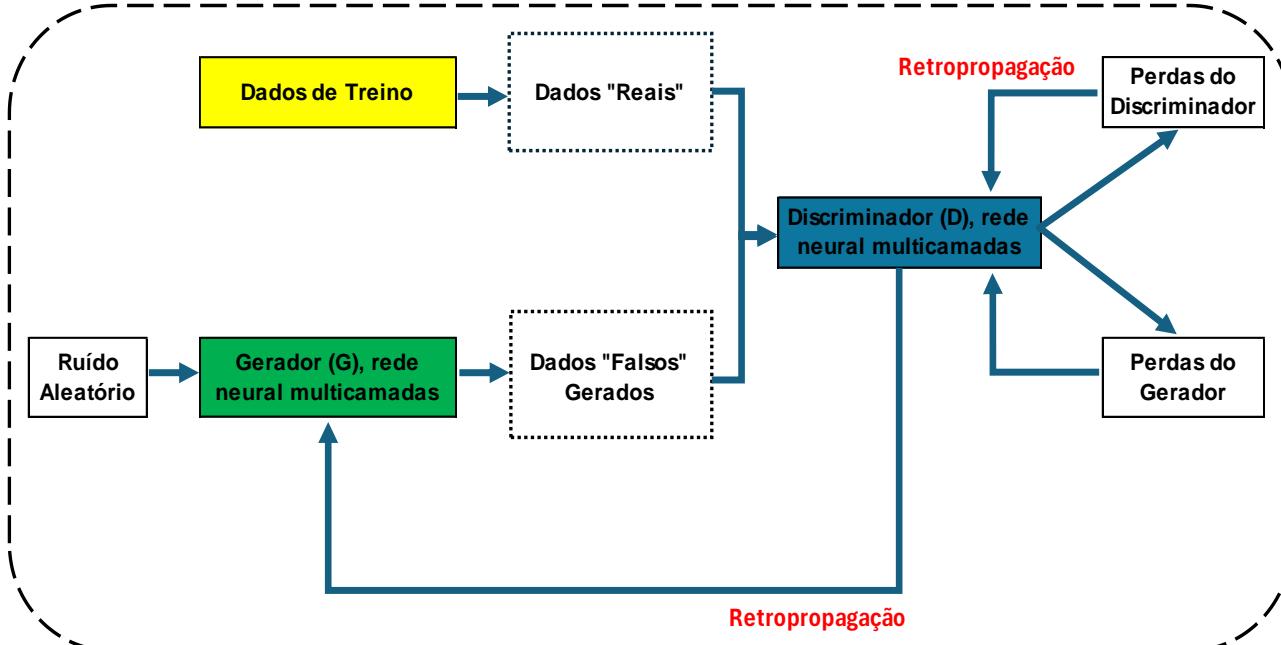
Yang *et al.* (2024b) desenvolveram um modelo DRL de detecção de anomalias em sistemas de monitoramento de tráfego de rede que aprende as políticas de detecção adaptativas para identificar comportamentos anômalos em tempo real, utilizando dados de séries temporais multivariadas. Isto é, a tarefa de tomada de decisão sequencial do modelo é formular o problema de detecção de anomalias enquanto o agente ajusta dinamicamente os limiares de detecção baseado em características observadas no tráfego de rede. Essa abordagem específica, segundo os autores, permite que o modelo se adapte a mudanças no ambiente e supere os métodos tradicionais de detecção baseados em limiares fixos ou modelos supervisionados. Um outro diferencial importante é a arquitetura que combina redes neurais convolucionais e redes neurais recorrentes para extraír características espaciais e temporais dos dados, enquanto o componente de DRL otimiza as decisões de detecção. O modelo demonstrou robustez em cenários com alto nível de ruído e variabilidade, com resultados experimentais mostrando melhorias significativas em métricas como precisão, recall e F1-score quando comparado a abordagens clássicas de detecção de anomalias. Os autores concluíram que o potencial do DRL em lidar com a complexidade inerente aos dados de rede é o ponto chave para o oferecimento de uma solução eficaz que identifique anomalias em tempo real em ambientes dinâmicos e não estacionários.

3.5 REDES ADVERSARIAIS GENERATIVAS (GAN)

3.5.1 Aspectos Teóricos da Arquitetura

Redes Adversariais Generativas é um modelo de aprendizado profundo introduzido ainda em 2014 por Goodfellow *et al.* (2014), projetado para aprender distribuições complexas de dados por meio da estratégia de treinamento adversarial. A arquitetura desse modelo consiste em duas redes neurais: um gerador (G) e um discriminador (D), isto é, o gerador sintetiza os dados a partir de ruído aleatório para produzir amostras indistinguíveis de dados reais, e o discriminador avalia a verossimilhança das entradas e distingue entre dados reais e dados gerados (Figura 7).

Figura 7. Fluxograma de funcionamento das Redes Adversariais Generativas



Fonte: Os Autores, 2025 – Adaptado de Goodfellow *et al.* (2014)

O treinamento ocorre pelo método *minimax* (MEGAHED e MOHAMMED, 2023), em que G é a função que minimiza a probabilidade de D identificar corretamente suas saídas como falsas, e D é a função que maximiza a precisão na detecção de fraudes. A função objetiva para o modelo é formulada da seguinte forma (Eq. 1):

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_{data}} [\log D(x)] + \mathbb{E}_{z \sim P_z} [\log (1 - D(G(z)))] \quad (\text{Eq. 1})$$

Teoricamente, essa função converge para o denominado equilíbrio de Nash (WANG *et al.*, 2024), no qual G captura a distribuição real dos dados, embora a convergência prática seja desafiadora devido a problemas como *mode collapse* (relacionada à falta de diversidade nas amostras) e a instabilidade no treinamento. Outras variantes, como as arquiteturas GAN de Wasserstein, tem abordado essas questões usando métricas de divergência alternativas (ex.: distância de Wasserstein) para estabilizar o processo de aprendizado profundo (QIU *et al.*, 2024).

Mesmo datada de 2014, arquiteturas GAN receberam um avanço significativo em 2024 – 2025 no que tange a capacidade de gerar dados sintéticos de alta fidelidade (imagens foto-realistas, vídeos e traduções texto-imagem) e, por isso, superaram os modelos generativos tradicionais como *Variational Autoencoders* (VAE) em qualidade de amostras (LYTVYNENKO *et al.*, 2024), ao exigir aprendizado com mínima supervisão ou até mesmo o aprendizado não supervisionado de estruturas complexas em dados (PURWANTO *et al.*, 2024).

Contudo, arquiteturas GAN também enfrentam algumas problemáticas e limitações (SHARMA *et al.*, 2024), como a instabilidade nos treinamentos que demandam ajustes de hiperparâmetros mais cuidadosos e restrições arquiteturais, o alto custo computacional especialmente para saídas de alta resolução e, ainda, as preocupações éticas, como a criação de *deepfakes*, reforçando a necessidade por métodos de detecção forense mais estratégicos.

3.5.2 Aplicações Práticas de Arquiteturas GAN

Na prática, arquiteturas GAN mostram potencial transformador na área da saúde ao sintetizarem dados de imagem médica para melhoria da precisão diagnóstica (MAMO *et al.*, 2024), na criação de conteúdo, impulsionando ferramentas como tradução texto-imagem (ex.: DALL-E) e transferência de estilo, práticas que tem revolucionado a produção artística e midiática (CHEN e ZHANG, 2023).

Mamo *et al.* (2024) propuseram um modelo GAN capaz de gerar dados realistas de pacientes e preservar a privacidade, na premissa da melhoria da qualidade e utilidade de dados médicos sintéticos. A principal contribuição dessa abordagem, portanto, está na capacidade do modelo de produzir dados sintéticos que mantêm as características estatísticas dos dados reais, ao mesmo tempo que mitigam riscos de reidentificação de pacientes. Essa estratégia visa mitigar os impactos da falta de acesso a grandes conjuntos de dados médicos, como HIPAA ou RGPD, limitados por regulamentações de privacidade. Os resultados apresentados pelos autores sugerem que os dados gerados pelo modelo são suficientemente realistas para análises subsequentes (previsão de desfechos clínicos e diagnósticos). No geral, o principal diferencial do modelo é a integração de técnicas avançadas de regularização e balanceamento no treinamento da arquitetura GAN, garantindo estabilidade e qualidade na geração de dados. Ainda, o modelo foi validado em múltiplos conjuntos de dados médicos com acesso irrestrito por condições de privacidade, que incluíram de registros eletrônicos de saúde a imagens médicas, o que mostra a versatilidade do modelo em diferentes tipos de dados.

Chen e Zhang (2023) propuseram o modelo inovador denominado CA-GAN, que desacopla os espaços de conteúdo e atributos por meio de mecanismos de atenção, permitindo a síntese de imagens com maior fidelidade e resolução (280 x 280) em comparação com métodos anteriores. Esse modelo representou uma contribuição significativa para o campo de Redes Adversariais Generativas aplicadas à geração de obras de arte, especialmente no contexto das pinturas tradicionais chinesas. É relatado ainda que essa abordagem resolve não apenas problemas comuns em trabalhos anteriores, como a perda de detalhes e a confusão de cores, como também preenche uma lacuna importante na literatura que historicamente concentrou-se mais na arte ocidental. Um dos diferenciais é como a introdução de

perdas especificamente designadas (*MS-SSIM* e *Charbonnier*) demonstra a personalização de funções de perda melhora significativamente a qualidade e a diversidade das obras geradas, aspectos cruciais para aplicações artísticas orientais. Outro avanço discutido pelos autores é sobre como estruturas de atenção podem ser eficazmente integradas às arquiteturas GAN, permitindo que o modelo foque em elementos específicos das obras, como flores e pássaros, que são característicos nas pinturas chinesas tradicionais. A abordagem proposta pelos autores tem implicações mais amplas para o campo das GAN aplicadas à arte, isto é, a estratégia de desacoplar conteúdo e estilo não apenas melhora os resultados para pinturas chinesas, mas também pode ser adaptada para outros estilos artísticos, abrindo novas possibilidades para a geração de obras com características específicas. A coleta e organização de um dataset substancial de pinturas tradicionais chinesas pelos autores também representa uma contribuição valiosa por viabilizar outras futuras pesquisas na área. No entanto, é importante considerar os limites desta abordagem, como a dependência de bancos de dados bem curados e a necessidade de ajustes especificamente para cada estilo artístico, bem como questões éticas relacionadas ao direito de uso e reprodução das obras artísticas exploradas.

4 DESAFIOS, PERSPECTIVAS FUTURAS E OUTRAS POTENCIAIS APLICAÇÕES

No Quadro 1 a seguir se evidencia os principais pontos sobre as técnicas abordadas nos tópicos anteriores e seus fundamentos, prós e contras, bem como as aplicações exploradas e/ou possíveis para essas arquiteturas.

Quadro 1. Comparativo das técnicas de *deep learning* e seus avanços em 2024-2025

Técnica	Fundamentos	Prós	Contras	Aplicações
<i>Redes Neurais Convolucionais Temporais</i> Ref.: a	Usa convoluções 1D com kernels causais/dilatados para modelagem de dados temporais.	Eficiência com dependências de longo alcance, processamento paralelo.	Requer grandes conjuntos de dados; não possui contexto bidirecional.	Análise de séries temporais, assistência médica, sistemas autônomos.
<i>Redes de Kolmogorov-Arnold</i> Ref.: b	Aprende funções multivariadas por meio de funções de ativação univariadas em arestas (inspirado no teorema de Kolmogorov-Arnold).	Eficiente em termos de parâmetros, e interpretável.	Alta sobrecarga computacional; desafios de escalabilidade.	Previsão de séries temporais, resolução de PDE, assistência médica.
<i>Redes Recorrentes Inspiradas em Quantum</i> Ref.: c	Integra princípios quânticos em RNN para retenção de memória.	Memória eficiente, processamento paralelo, modelagem de dependência complexa.	Sobrecarga computacional; dificuldades de transferência de simulação para realidade.	Sistemas autônomos, análise de séries temporais, detecção de anomalias.
<i>Aprendizado por Reforço Profundo</i> Ref.: d	Combina RL com DNN para aprender políticas por meio de Processos de	Adaptabilidade a ambientes complexos,	Ineficiência amostral, lacunas de generalização,	Sistemas autônomos, assistência médica, detecção de anomalias.

<i>Redes Adversariais Generativas</i> Ref.: e	Decisão de Markov (MDP).	engenharia de recursos mínimos.	treinamento instável.	Criação de conteúdo, assistência médica, detecção de anomalias.
	Treinamento adversário de redes geradoras-discriminadoras para sintetizar dados.	Dados sintéticos de alta fidelidade, supervisão mínima.	Instabilidade de treinamento, riscos éticos (por exemplo, deepfakes).	

Nota: a – Zeghina *et al.* (2024), b – Alonso (2025), c – Pandurangan *et al.* (2024), d – Gao e Scheweidtmann (2024), e – Goodfellow *et al.* (2014).

Como observado, cada técnica de aprendizado profundo abordada nesta revisão enfrenta desafios distintos. Redes Convolucionais Temporais apresentam limitações no contexto bidirecional e exigem grandes conjuntos de dados para um melhor desempenho, restringindo sua aplicação em cenários com escassez de dados. As Redes de Kolmogorov-Arnold podem apresentar alto custo computacional e limitações de escalabilidade no que tange entradas de alta dimensão, devido à dependência de funções de ativação univariadas nas arestas. A técnica de Redes Recorrentes Inspiradas em Quantum enfrenta a lacuna "*sim-to-real*" e *overhead* de simulações quânticas, algo que dificulta sua implementação em sistemas com segurança crítica. O Aprendizado por Reforço Profundo sofre com ineficiência amostral, instabilidade no treinamento e *gaps* de generalização, restringindo o uso em ambientes mais dinâmicos. Já as Redes Adversariais Generativas lidam com instabilidade no treinamento (o denominado *mode collapse*), além de riscos éticos (uso na produção de *deepfakes*) e os estritos requisitos de qualidade de dados, complicando sua aplicação em domínios sensíveis como saúde.

As soluções emergentes em 2024 e início de 2025 buscam resolver as limitações citadas, como as Redes Convolucionais Temporais com arquiteturas híbridas dotadas de processamento bidirecional ou mecanismos de atenção semelhantes aos *transformers* para aprimorar a percepção de contexto e reduzir demandas por dados. Por outro lado, Redes de Kolmogorov-Arnold se beneficiam de estruturas computacionais otimizadas ou modelos híbridos quântico-clássicos para melhorar a escalabilidade. As Redes Recorrentes Inspiradas em Quantum, por sua vez, aproveitam avanços em hardware quântico para viabilizar aplicações práticas, especialmente em sistemas multiagentes. No caso do Aprendizado por Reforço Profundo, há foco em algoritmos mais eficientes em termos de amostragem, aprendizado multitarefa, e adaptação em tempo real para mitigar os efeitos das lacunas "*sim-to-real*". As Redes Adversariais Generativas tem avançado no que tange o treinamento estável (ex.: GAN de Wasserstein com penalidades de gradiente) e salvaguardas relacionadas a práticas éticas, expandindo aplicações para medicina personalizada e cibersegurança.

Outras aplicações das arquiteturas abordadas são: (I) Redes Convolucionais Temporais na análise de comportamento social, com a finalidade de modelar padrões temporais em interações

humanas, como dinâmicas de multidões ou evolução de redes sociais, devido à sua eficiência em dependências de longo prazo; (II) Redes de Kolmogorov-Arnold aplicadas a modelagem financeira que preveja dinâmicas complexas de mercado por terem eficiência paramétrica e interpretabilidade adequadas para dados econômicos de alta dimensão; (III) Redes Recorrentes Inspiradas em Quantum em simulações de química quântica para otimização energética molecular, que aproveite o emaranhamento quântico para uma modelagem mais precisa; (IV) Aprendizado Profundo por Reforço em sistemas educacionais adaptativos para personalizar trajetórias de aprendizado, em que agentes otimizam engajamento e retenção de conhecimento via políticas baseadas em MDP; e (V) Redes Adversariais Generativas na prática de medicina personalizada para gerar dados sintéticos específicos de pacientes em descoberta de fármacos, preservando privacidade e acelerando testes clínicos.

5 CONCLUSÕES E PERSPECTIVAS FUTURAS

Esta revisão destacou os aspectos teóricos e o potencial transformador de cinco arquiteturas de aprendizado profundo emergentes em 2024 e início de 2025 em diversas áreas do conhecimento, como na saúde, em sistemas autônomos, na criação de conteúdo, na análise de séries temporais, e na detecção de anomalias. Essas arquiteturas apresentam vantagens únicas, contudo, os avanços são limitados por desafios significativos. Desse modo, ressalta-se a necessidade por pesquisas contínuas para a superação de barreiras técnicas e ampliação de aplicabilidades práticas. Dentre as perspectivas futuras, inclui-se a ampliação e aprimoramento de arquiteturas híbridas, modelos quânticos clássicos, arquiteturas mais eficientes em amostragem, *frameworks* mais estáveis em treinamento, e adequações éticas em modelos geratitivos. Ao fornecer uma comparação abrangente e identificar lacunas, esta revisão visou elucidar insights para seleção de ferramentas mais otimizadas para desafios específicos, apontando os avanços no aprendizado profundo. À medida em que evoluem, a integração dessas arquiteturas com tendências emergentes como a *Federated Learning*, *Edge AI* e *Explainable AI* se torna crucial para a escalada dos seus potenciais em cenários do mundo real.

REFERÊNCIAS

ABUDUREXITI, Yilixiati; HAN, Guangjie; ZHANG, Fan; LIU, Li. An explainable unsupervised anomaly detection framework for Industrial Internet of Things. *Computers & Security*, [S.L.], v. 148, p. 104130, jan. 2025. Elsevier BV. <http://dx.doi.org/10.1016/j.cose.2024.104130>.

AHMAD, Naushad; JAS, Shubham. Quantum-inspired neural networks for time-series air pollution prediction and control of the most polluted region in the world. *Quantum Machine Intelligence*, [S.L.], v. 7, n. 1, p. 42484, 17 jan. 2025. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s42484-024-00232-6>.

ALAM, Shahanur; YAKOPCIC, Chris; WU, Qing; BARNELL, Mark; KHAN, Simon; TAHA, Tarek M.. Survey of Deep Learning Accelerators for Edge and Emerging Computing. *Electronics*, [S.L.], v. 13, n. 15, p. 2988, 29 jul. 2024. MDPI AG. <http://dx.doi.org/10.3390/electronics13152988>.

ALONSO, Miquel Noguer I. The Mathematics of Kolmogorov-Arnold-Networks versus Artificial Neural Networks. *SSRN*, [S.L.], v. 26, n. 1, p. 5072413, 2025. Elsevier BV. <http://dx.doi.org/10.2139/ssrn.5072413>

ALYAMI, Sarah; LUQMAN, Hamzah; HAMMOUDEH, Mohammad. Reviewing 25 years of continuous sign language recognition research: advances, challenges, and prospects. *Information Processing & Management*, [S.L.], v. 61, n. 5, p. 103774, set. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.ipm.2024.103774>.

BASINA, Divesh; VISHAL, Joseph Raj; CHOUDHARY, Aarya; CHAKRAVARTHI, Bharatesh. KAT to KANs: a review of kolmogorov-arnold networks and the neural leap forward. *Arxiv*, [S.L.], v. 11, n. 24, p. 10622, 2024. ArXiv. <http://dx.doi.org/10.48550/ARXIV.2411.10622>.

CHEN, Zhuo; LUO, Di. Entangling Intelligence: ai-quantum crossovers and perspectives. 2024 IEEE 6Th International Conference On Trust, Privacy And Security In Intelligent Systems, And Applications (TpS-Isa), [S.L.], p. 516-519, 28 out. 2024. IEEE. <http://dx.doi.org/10.1109/tps-isa62245.2024.00069>.

CHEN, Zihan; ZHANG, Yi. CA-GAN: the synthesis of chinese art paintings using generative adversarial networks. *The Visual Computer*, [S.L.], v. 40, n. 8, p. 5451-5463, 16 out. 2023. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s00371-023-03115-2>.

CHUNG, Jaehoon; FAYYAD, Jamil; YOUNES, Younes Al; NAJJARAN, Homayoun. Learning team-based navigation: a review of deep reinforcement learning techniques for multi-agent pathfinding. *Artificial Intelligence Review*, [S.L.], v. 57, n. 2, p. 10670, 8 fev. 2024. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10462-023-10670-6>.

DEMELIUS, Lea; KERN, Roman; TRÜGLER, Andreas. Recent Advances of Differential Privacy in Centralized Deep Learning: a systematic survey. *Acm Computing Surveys*, [S.L.], v. 57, n. 6, p. 1-28, 10 fev. 2025. Association for Computing Machinery (ACM). <http://dx.doi.org/10.1145/3712000>.

DONG, Fanglin; LI, Shibo; LI, Weihua. TCKAN: a novel integrated network model for predicting mortality risk in sepsis patients. *Medical & Biological Engineering & Computing*, [S.L.], v. 24, n. 2, p. 03245, 19 nov. 2024. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s11517-024-03245-2>.

FORNARI, Lucimara Fabiana; PINHO, Isabel. Revisão da Literatura com apoio de ferramentas digitais: avanços e desafios. *New Trends In Qualitative Research*, [S.L.], v. 10, n. 2022, p. 1-15, 19 maio 2022. Ludomedia. <http://dx.doi.org/10.36367/ntqr.10.2022.e512>.

GANIAE, M.A.; HU, Minghui; MALIK, A.K.; TANVEER, M.; SUGANTHAN, P.N. Ensemble deep learning: a review. *Engineering Applications Of Artificial Intelligence*, [S.L.], v. 115, p. 105151, out. 2022. Elsevier BV. <http://dx.doi.org/10.1016/j.engappai.2022.105151>.

GAO, Yihang; TAN, Vincent Y. F. On the Convergence of (Stochastic) Gradient Descent for Kolmogorov--Arnold Networks. *SSRN*, [S.L.], v. 10, n. 24, p. 08041, 2024. ArXiv. <http://dx.doi.org/10.48550/ARXIV.2410.08041>.

GAO, Qinghe; SCHWEIDTMANN, Artur M. Deep reinforcement learning for process design: review and perspective. *Current Opinion In Chemical Engineering*, [S.L.], v. 44, p. 101012, jun. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.coche.2024.101012>.

GILL, Sukhpal Singh; GOLEC, Muhammed; HU, Jianmin; XU, Minxian; DU, Junhui; WU, Huaming; WALIA, Guneet Kaur; MURUGESAN, Subramaniam Subramanian; ALI, Babar; KUMAR, Mohit. Edge AI: a taxonomy, systematic review and future directions. *Cluster Computing*, [S.L.], v. 28, n. 1, p. 1-24, 18 out. 2024. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10586-024-04686-y>.

GOODFELLOW, Ian J.; POUGET-ABADIE, Jean; MIRZA, Mehdi; XU, Bing; WARDE-FARLEY, David; OZAIR, Sherjil; COURVILLE, Aaron; BENGIO, Yoshua. Generative Adversarial Networks. *Arxiv*, [S.L.], v. 14, n. 6, p. 2661, 2014. ArXiv. <http://dx.doi.org/10.48550/ARXIV.1406.2661>.

HERNANDEZ, Joannes Paulus Tolentino. Compassionate Care with Autonomous AI Humanoid Robots in Future Healthcare Delivery: a multisensory simulation of next-generation models. *Biomimetics*, [S.L.], v. 9, n. 11, p. 687, 11 nov. 2024. MDPI AG. <http://dx.doi.org/10.3390/biomimetics9110687>.

HONG, Ying-Yi; RIOFLORIDO, Christian Lian Paulo P.; ZHANG, Weina. Hybrid deep learning and quantum-inspired neural network for day-ahead spatiotemporal wind speed forecasting. *Expert Systems With Applications*, [S.L.], v. 241, p. 122645, maio 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2023.122645>.

HOSAIN, Md. Tanzib; JIM, Jamin Rahman; MRIDHA, M.F.; KABIR, Md Mohsin. Explainable AI approaches in deep learning: advancements, applications and challenges. *Computers And Electrical Engineering*, [S.L.], v. 117, p. 109246, jul. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.compeleceng.2024.109246>.

IBRAHUM, Ahmed Dawod Mohammed; SHANG, Zhengyu; HONG, Jang-Eui. How Resilient Are Kolmogorov–Arnold Networks in Classification Tasks? A Robustness Investigation. *Applied Sciences*, [S.L.], v. 14, n. 22, p. 10173, 6 nov. 2024. MDPI AG. <http://dx.doi.org/10.3390/app142210173>.

KILANI, Bochra Hadj. Kolmogorov-Arnold Networks: key developments and uses. *Qeios*, [S.L.], v. 21, n. 1, p. 04663521, 24 jun. 2024. Qeios Ltd. <http://dx.doi.org/10.32388/7nncaa>.

KIM, Sung Wook; KONG, Jun Ho; LEE, Sang Won; LEE, Seungchul. Recent Advances of Artificial Intelligence in Manufacturing Industrial Sectors: a review. International Journal Of Precision Engineering And Manufacturing, [S.L.], v. 23, n. 1, p. 111-129, 3 nov. 2021. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s12541-021-00600-3>.

LABHANE, Suchitra; RADHA, J.; POKKULURI, Kiran Sree; SOMASUNDARAM, R.; SHANKAR, R. Shiva; SRIVASTAVA, Prateek. Quantum-Inspired Deep Learning for Networked Data Analysis With Quantum Networked Discord and Allies. Advances In Computational Intelligence And Robotics, [S.L.], p. 13-29, 2 ago. 2024. IGI Global. <http://dx.doi.org/10.4018/979-8-3693-5832-0.ch002>.

LI, Min; TIAN, Hangwei; CHEN, Qinghui; ZHOU, Mingle; LI, Gang. A hybrid prediction method for short-term load based on temporal convolutional networks and attentional mechanisms. Iet Generation, Transmission & Distribution, [S.L.], v. 18, n. 5, p. 885-898, 12 mar. 2023. Institution of Engineering and Technology (IET). <http://dx.doi.org/10.1049/gtd2.12798>.

LUO, Zhenglong; CHEN, Zhiyong; WELSH, James. Multi-agent Reinforcement Learning with Deep Networks for Diverse Q-Vectors. Arxiv, [S.L.], v. 24, n. 6, p. 1-8, 2024. ArXiv. <http://dx.doi.org/10.48550/ARXIV.2406.07848>.

LYTVYNENKO, Volodtmyr; OLSZEWSKI, Serge; OSYPENKO, Volodymyr; LURIE, Iryna; DEMCHENKO, Violeta; DONTSOVA, Dariia. Comparative Analysis of the Use of Variational Autoencoders and Generative Adversarial Networks for Anomalous Image Detection Tasks. 2024 Ieee 7Th International Conference On Actual Problems Of Unmanned Aerial Vehicles Development (Apuavd), [S.L.], p. 270-275, 2024. IEEE. <http://dx.doi.org/10.1109/apuavd64488.2024.10765911>.

MALHOTRA, P. A. Quantum-Inspired Neural Networks: A Paradigm Shift in AI Computation. International Journal of Sustainable Development in Computer Science Engineering, 10(10). 2024. <https://journals.threws.com/index.php/IJSDCSE/article/view/331>

MAMO, Abiy Abinet; GEBRESILASSIE, Bealu Girma; MUKHERJEE, Aniruddha; HASSIJA, Vikas; CHAMOLA, Vinay. Advancing Medical Imaging Through Generative Adversarial Networks: a comprehensive review and future prospects. Cognitive Computation, [S.L.], v. 16, n. 5, p. 2131-2153, 13 jun. 2024. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s12559-024-10291-3>.

MEGAHED, Mohammed; MOHAMMED, Ammar. A comprehensive review of generative adversarial networks: fundamentals, applications, and challenges. Wires Computational Statistics, [S.L.], v. 16, n. 1, p. 1629, 2 ago. 2023. Wiley. <http://dx.doi.org/10.1002/wics.1629>.

MU, R.; ZENG, Xiaoqin. A Review of Deep Learning Research. KSII Transactions On Internet And Information Systems, [S.L.], v. 13, n. 4, p. 1738-1764, 2019. <http://dx.doi.org/10.3837/tiis.2019.04.001>.

PANDURANGAN, Kamaleswari; PRIYADHARSHINI, A.; TASEEN, Rakheeba; GALEBATHULLAH, B.; YASEEN, Haseeba; RAVICHANDRAN, P.. Quantum-Inspired Algorithms for AI and Machine Learning. Advances In Mechatronics And Mechanical Engineering, [S.L.], p. 79-92, 11 out. 2024. IGI Global. <http://dx.doi.org/10.4018/979-8-3693-7076-6.ch004>.

PAPADOPOULOS, George; BASTAS, Alevizos; VOUROS, George A.; CROOK, Ian; ANDRIENKO, Natalia; ANDRIENKO, Gennady; CORDERO, Jose Manuel. Deep reinforcement learning in service of air traffic controllers to resolve tactical conflicts. *Expert Systems With Applications*, [S.L.], v. 236, p. 121234, fev. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2023.121234>.

PENG, Yiming; CHEN, Gang; ZHANG, Mengjie; XUE, Bing. Proximal evolutionary strategy: improving deep reinforcement learning through evolutionary policy optimization. *Memetic Computing*, [S.L.], v. 16, n. 3, p. 445-466, 2024. <http://dx.doi.org/10.1007/s12293-024-00419-1>.

PURWANTO, Agus; KUSRINI; UTAMI, Ema; AGUSTRIAWAN, David. A Comprehensive Literature Review on Generative Adversarial Networks (GANs) for AI Anime Image Generation. *2024 Ieee International Conference On Artificial Intelligence And Mechatronics Systems (Aims)*, [S.L.], p. 1-6, 21 fev. 2024. IEEE. <http://dx.doi.org/10.1109/aims61812.2024.10513308>.

QIU, Yixuan; GAO, Qingyi; WANG, Xiao. Adaptive Learning of the Latent Space of Wasserstein Generative Adversarial Networks. *Journal Of The American Statistical Association*, [S.L.], p. 1-13, 19 nov. 2024. Informa UK Limited. <http://dx.doi.org/10.1080/01621459.2024.2408778>.

REN, Zhiyu; LI, Xiaojie; PENG, Jing; CHEN, Ken; TAN, Qushan; WU, XI; SHI, Canghong. Graph autoencoder with mirror temporal convolutional networks for traffic anomaly detection. *Scientific Reports*, [S.L.], v. 14, n. 1, p. 1247-1261, 13 jan. 2024. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41598-024-51374-3>.

ROTHER, Edna Terezinha. Revisão sistemática X revisão narrativa. *Acta Paulista de Enfermagem*, [S.L.], v. 20, n. 2, p. 1-2, 2007. <http://dx.doi.org/10.1590/s0103-21002007000200001>.

SAARELA, Mirka; PODGORELEC, Vili. Recent Applications of Explainable AI (XAI): a systematic literature review. *Applied Sciences*, [S.L.], v. 14, n. 19, p. 8884, 2024. <http://dx.doi.org/10.3390/app14198884>.

SADID, Hashmatullah; ANTONIOU, Constantinos. Dynamic Spatio-temporal Graph Neural Network for Surrounding-aware Trajectory Prediction of Autonomous Vehicles. *Ieee Transactions On Intelligent Vehicles*, [S.L.], p. 1-14, 2024. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tiv.2024.3406507>.

SHARMA, Preeti; KUMAR, Manoj; SHARMA, Hitesh Kumar; BIJU, Soly Mathew. Generative adversarial networks (GANs): introduction, taxonomy, variants, limitations, and applications. *Multimedia Tools And Applications*, [S.L.], v. 83, n. 41, p. 88811-88858, 26 mar. 2024. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s11042-024-18767-y>.

SHU, Man; LÜ, Shuai; GONG, Xiaoyu; AN, Daolong; LI, Songlin. Episodic Memory-Double Actor-Critic Twin Delayed Deep Deterministic Policy Gradient. *Neural Networks*, [S.L.], p. 107286, fev. 2025. Elsevier BV. <http://dx.doi.org/10.1016/j.neunet.2025.107286>.

SNYDER, Hannah. Literature review as a research methodology: an overview and guidelines. *Journal Of Business Research*, v. 104, p. 333-339, 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.jbusres.2019.07.039>.

SOORI, Mohsen; JOUGH, Foad Karimi Ghaleh; DASTRES, Roza; AREZOO, Behrooz. AI-Based Decision Support Systems in Industry 4.0, A Review. *Journal Of Economy And Technology*, [S.L.], v. 5, n. 8, p. 1-35, ago. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.ject.2024.08.005>.

THAPA, Nikhil; LEE, Joonwhoan. Dual-Path Beat Tracking: combining temporal convolutional networks and transformers in parallel. *Applied Sciences*, [S.L.], v. 14, n. 24, p. 11777, 17 dez. 2024. MDPI AG. <http://dx.doi.org/10.3390/app142411777>.

TIWARI, Rajesh Kumar; HAMID, Abu Bakar Bin Abdul; NYAMASVISVA, Tadiwa Elisha. Potential in Quantum Machine Learning for Real-World Problems. 2024 5Th International Conference On Recent Trends In Computer Science And Technology (Icrcst), [S.L.], p. 695-700, 9 abr. 2024. IEEE. <http://dx.doi.org/10.1109/icrcst61793.2024.10578519>.

WANG, Guangming; LEI, Xiyuan; LI, Wen; LEE, Won Hee; HUANG, Lianchi; ZHU, Jialin; JIA, Shanshan; WANG, Dong; ZHENG, Yang; ZHANG, Hua. Channel-Selection-Based Temporal Convolutional Network for Patient-Specific Epileptic Seizure Detection. *IEEE Transactions On Cognitive And Developmental Systems*, v. 17, n. 1, p. 179-188, 2025. <http://dx.doi.org/10.1109/tcds.2024.3433551>.

WANG, Yong; ZHANG, Qian; WANG, Gai-Ge; CHENG, Honglei. The application of evolutionary computation in generative adversarial networks (GANs): a systematic literature survey. *Artificial Intelligence Review*, v. 57, n. 7, p. 10818, 2024. <http://dx.doi.org/10.1007/s10462-024-10818-y>.

XU, Kunpeng; CHEN, Lifei; WANG, Shengrui. Kolmogorov-Arnold Networks for Time Series: bridging predictive power and interpretability. *SSRN*, [S.L.], v. 24, n. 06, p. 02496, 2024. ArXiv. <http://dx.doi.org/10.48550/ARXIV.2406.02496>.

YANG, Kuo; YU, Zecong; SU, Xin; ZHANG, Fengjin; HE, Xiong; WANG, Ning; ZHENG, Qiguang; YU, Feidie; WEN, Tiancai; ZHOU, Xuezhong. PrescDRL: deep reinforcement learning for herbal prescription planning in treatment of chronic diseases. *Chinese Medicine*, [S.L.], v. 19, n. 1, p. 13020, 16 out. 2024a. Springer Science and Business Media LLC. <http://dx.doi.org/10.1186/s13020-024-01005-w>.

YANG, Xue; HOWLEY, Enda; SCHUKAT, Michael. ADT: time series anomaly detection for cyber-physical systems via deep reinforcement learning. *Computers & Security*, [S.L.], v. 141, p. 103825, jun. 2024b. Elsevier BV. <http://dx.doi.org/10.1016/j.cose.2024.103825>.

YAO, Huan; LOU, Hao; WANG, Dan; CHEN, Yijun; YAN, Junkun. A Resource Scheduling Algorithm for Multi-Target 3D Imaging in Radar Network Based on Deep Reinforcement Learning. *Remote Sensing*, [S.L.], v. 16, n. 23, p. 4472, 28 nov. 2024. MDPI AG. <http://dx.doi.org/10.3390/rs16234472>.

ZEGHINA, Assaad; LEBORGNE, Aurélie; BER, Florence Le; VACAVANT, Antoine. Deep learning on spatiotemporal graphs: a systematic review, methodological landscape, and research opportunities. *Neurocomputing*, v. 594, p. 127861, 2024. <http://dx.doi.org/10.1016/j.neucom.2024.127861>.

ZHANG, Ye; ZHAO, Wang; WANG, Jingyu; YUAN, Yuan. Recent progress, challenges and future prospects of applied deep reinforcement learning: a practical perspective in path planning. *Neurocomputing*, v. 608, p. 128423, 2024a.
<http://dx.doi.org/10.1016/j.neucom.2024.128423>.