Redes neurais para detecção de séries temporais fora da distribuição em aprendizado de única classe

José Gilberto Barbosa de Medeiros Júnior[1], **Diego Furtado Silva**[1]

[1]Universidade de São Paulo (ICMC-USP) - São Carlos, São Paulo, Brasil

gilberto.barbosa@usp.br, dfsilva@usp.br

Resumo. *A classificação de séries temporais apresenta desafios únicos, especialmente em domínios onde rotular dados tem alto custo ou onde há grande desbalanceamento dos rótulos. Diante desses problemas, o aprendizado de única classe (one-class learning, OCL) surge como alternativa, focando exclusivamente no aprendizado de séries temporais pertencentes a uma única classe, também chamada de classe de interesse. Os métodos existentes de OCL frequentemente sofrem com instabilidades durante o treinamento, dependência de contraexemplos e adaptações inadequadas para capturar dependências temporais. Esta dissertação propõe um novo framework de OCL para séries temporais, integrando mecanismos de redes neurais e técnicas de detecção de dados fora da distribuição (OOD). Para aprimorar a extração de características temporais, são apresentados dois novos componentes: (i) a LeakySineLU, uma função de ativação semi-periódica projetada para séries temporais, e (ii) Convoluções Deformáveis, que capturam dependências não contínuas e de longo alcance entre observações de uma série. Esses mecanismos são incorporados à TGNet, método de OCL proposto que utiliza Modelos de Mistura de Gaussianas (GMM) para modelar a distribuição da classe de interesse e identificar instâncias fora dela. Foram conduzidos 33.110 experimentos em 112 conjuntos de dados do repositório UCR, demonstrando que a TGNet supera abordagens tradicionais e recentes de OCL, alcançando a maior F1-macro média e o melhor ranque médio entre os métodos comparados (*3.16*), com superioridade estatística confirmada por testes de Wilcoxon-Holm*

Informações da Candidatura:

Instituição onde o título foi obtido: Universidade de São Paulo

Orientador: Prof. Dr. Diego Furtado Silva, Universidade de São Paulo, São Paulo, Brasil

Trilha: Mestrado

Data da defesa da Dissertação: 15 de Abril de 2025

Membros da banca examinadora:

1. Elaine Parros Machado de Sousa, Universidade de São Paulo, São Paulo, Brasil
2. Gustavo Enrique de Almeida Prado Alves Batista, University of New South Wales, Nova Gales do Sul, Austrália
3. Vinicius Mourão Alves de Souza, Pontifícia Universidade Católica do Paraná, Paraná, Brasil

Link para a dissertação completa: https://teses.usp.br/teses/disponiveis/55/55134/tde-29072025-141957/pt-br.html

Links das publicações derivadas da dissertação (congressos e periódicos): [Barbosa de Medeiros Júnior et al. 2026, Golo et al. 2025, Barbosa de Medeiros Júnior et al. 2024, Silva et al. 2023]

Link para o perfil do Google Scholar: https://scholar.google.com/citations?hl=pt-BR&user=OvRan0cAAAA

Link para o perfil do DBPL: https://dblp.uni-trier.de/pid/379/9866.html

Link para o Lattes: http://lattes.cnpq.br/9862934822037317

1. Introdução

Com o avanço de tecnologias como a Internet das Coisas (IoT), cidades inteligentes e veículos autônomos, houve um aumento massivo na coleta de dados temporais através de sensores cada vez mais precisos e acessíveis [Yeh et al. 2018]. Entretanto, a classificação supervisionada tradicional de séries temporais enfrenta desafios significativos, pois a rotulação de cada instância exige esforço especializado, sendo frequentemente uma tarefa custosa e demorada [Gao et al. 2020]. Além disso, em muitos domínios práticos, ocorre um desbalanceamento severo de classes, onde a classe de interesse é significativamente mais prevalente que as demais [Cabral and de Oliveira 2014].

Nesse cenário, o Aprendizado de Única Classe (OCL) surge como uma alternativa viável e eficiente, focando no treinamento de modelos utilizando exclusivamente instâncias da classe de interesse [Fukunaga et al. 1987]. Esta abordagem é particularmente útil em aplicações como a saúde digital, onde um modelo pode ser treinado apenas com padrões de eletrocardiogramas (ECGs) de indivíduos saudáveis para identificar desvios que indiquem patologias, ou no monitoramento ambiental para identificar espécies específicas de insetos através de sensores.

Apesar do seu potencial, os métodos de OCL existentes para séries temporais apresentam limitações críticas. Abordagens baseadas em distância frequentemente utilizam dados não rotulados durante o treinamento, o que contradiz a premissa fundamental do OCL e pode levar ao colapso do modelo. Além disso, muitas técnicas baseadas em redes neurais, como autoencoders, não são devidamente adaptadas para capturar a natureza sequencial e as dependências temporais dos dados, além de sofrerem com instabilidades como a explosão ou desaparecimento de gradientes.

Esta dissertação propõe superar essas lacunas através do desenvolvimento de um framework robusto baseado em redes neurais e técnicas de detecção de dados fora da distribuição (OOD). O trabalho apresenta três contribuições principais:

1. LeakySineLU: Uma nova função de ativação semi-periódica projetada para capturar padrões sazonais e tendências em séries temporais, evitando problemas de estagnação de gradientes;
2. Convoluções Deformáveis para séries Temporais: Uma adaptação de mecanismos convolucionais que permite capturar dependências de longo alcance e não contínuas através de deslocamentos aprendíveis;

3. Um novo método de OCL que integra Modelos de Mistura de Gaussianas (GMM) para modelar a densidade de probabilidade da classe de interesse e identificar instâncias anômalas de forma estatisticamente fundamentada.

2. Trabalhos relacionados

O estado da arte em séries temporais tem focado predominantemente em tarefas de classificação supervisionada multi-classe e binária. No entanto, o Aprendizado de Única Classe (OCL) e a detecção de Dados Fora da Distribuição (OOD) apresentam desafios que as arquiteturas tradicionais não endereçam plenamente.

2.1. Aprendizado de Única Classe

As abordagens clássicas de OCL para séries temporais dividem-se majoritariamente em métodos baseados em distância e métodos baseados em redes neurais. Muitos métodos recentes baseados em distância violam a premissa fundamental do OCL ao utilizarem instâncias não rotuladas durante o treinamento, enquadrando-se tecnicamente em Positive-Unlabeled Learning (PUL) [Provotar et al. 2019, Ruff et al. 2018]. Quando esses modelos são restritos apenas à classe de interesse, eles frequentemente sofrem colapso. Outros métodos como o DAGMM [Zong et al. 2018] e o DeepSVDD [Ruff et al. 2018] não são adaptados para capturar dependências temporais e quando utilizados podem levar a explosões dos gradientes. Além disso, algoritmos tradicionais como Isolation Forest e One-Class SVM (OCSVM) falham ao não capturar as dependências temporais intrínsecas dos dados.

2.2. Arquiteturas Neurais e Convoluções

Arquiteturas consolidadas como Fully Convolutional Networks (FCN) [Wang et al. 2017], ResNet [Wang et al. 2017] e InceptionTime [Ismail Fawaz et al. 2020] demonstraram excelência na classificação de séries temporais, mas operam com grades de kernels fixas. Essa rigidez limita a capacidade do modelo de extrair características de dependências de longo alcance ou padrões não contínuos sem um aumento proibitivo no número de parâmetros. Embora mecanismos de atenção local e transformadores tentem mitigar essa falha, eles impõem um custo computacional elevado que dificulta sua aplicação em cenários de recursos limitados.

2.3. Funções de Ativações para séries Temporais

A literatura de Deep Learning frequentemente negligencia a escolha da função de ativação, adaptando funções projetadas para visão computacional. A função ReLU, a mais utilizada, sofre com o problema de ”dying ReLU” [He et al. 2015], resultando em perda de informação em valores negativos, que são fundamentais em diversos sinais temporais. Funções puramente periódicas, como o seno, embora tentem capturar sazonalidade, frequentemente estagnam em mínimos locais devido à sua natureza não monotônica [Ziyin et al. 2020].

2.4. Lacunas Cobertas por esta Dissertação

Esta dissertação preenche as lacunas mencionadas ao: (i) propor o framework TGNet, que utiliza modelagem probabilística por GMM integrada para evitar a dependência de contraexemplos e o colapso do modelo; (ii) introduzir Convoluções Deformáveis

1D [Barbosa de Medeiros Júnior et al. 2026], permitindo a captura adaptativa de padrões temporais esparsos com baixo custo computacional; e (iii) desenvolver a função LeakySineLU [Barbosa de Medeiros Júnior et al. 2024], que combina propriedades semi-periódicas e fluxo de gradiente contínuo para melhor reconstrução e classificação de sinais. Além disso, o trabalho resolve problemas de instabilidade numérica comuns em modelos que utilizam GMMs através de uma nova função de perda e da função explu [Golo et al. 2025].

3. Metodologia

A metodologia central deste trabalho é o desenvolvimento do framework TGNet (Time Series Gaussian Mixture Model Network), uma arquitetura de rede neural fim-a-fim projetada especificamente para o Aprendizado de Única Classe (OCL) em séries temporais. O framework combina as capacidades de extração de características de um autoencoder com a modelagem probabilística de Modelos de Mistura de Gaussianas (GMM), visando identificar instâncias fora da distribuição (OOD) sem a necessidade de contraexemplos durante o treinamento.

3.1. Função de Ativação LeakySineLU

A LeakySineLU é uma das principais inovações deste trabalho, desenvolvida para endereçar lacunas críticas no processamento de sinais temporais por redes neurais profundas. A escolha da função de ativação é frequentemente negligenciada na literatura de séries temporais, que costuma adaptar funções projetadas para visão computacional, como a ReLU [Ismail Fawaz et al. 2019]. A Figura 1 mostra a comparação entre diferentes funções de ativação aplicadas à uma saída de camada que processou uma série temporal.

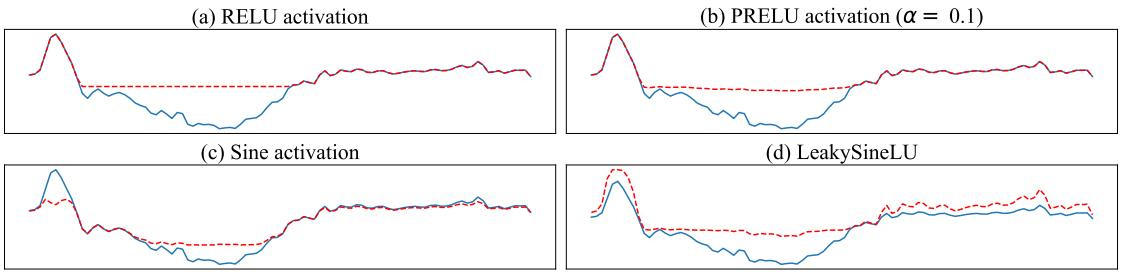


Figura 1. Comparação entre diferentes funções de ativação aplicadas a saída de uma camada de rede neural aplicada em uma série temporal.

Diferente da ReLU, que zera valores negativos e causa a perda de informações essenciais em sinais oscilatórios (problema conhecido como dying ReLU [He et al. 2015]), a LeakySineLU mantém inclinações em todo o seu domínio. Ela é definida matematicamente como uma função semi-periódica:

$$\sigma(x) = \begin{cases} \sin^2(x) + x, & \text{se } x > 0, \\ \frac{sin^2(x)+x}{2} & \text{caso contrário.} \end{cases} \tag{1}$$

O termo "semi-periódica" refere-se ao fato de que a periodicidade não reside na função em si (que é crescente e ilimitada), mas sim em sua derivada, que oscila de forma

controlada. Ao contrário de funções como Sigmoid ou TanH, a LeakySineLU não possui limites superiores ou inferiores ($+\infty$ e $-\infty$), o que atenua o problema do desaparecimento do gradiente (vanishing gradient) durante o treinamento de redes profundas.

Funções puramente periódicas, como o seno, frequentemente estagnam em mínimos locais devido a derivadas não monotônicas. A LeakySineLU supera isso ao garantir que sua derivada seja sempre positiva, facilitando a convergência de algoritmos baseados em gradiente descendente. Embora a derivada possua uma descontinuidade em $x = 0$, o trabalho utiliza o conceito de sub-derivadas (valores $\{1, 0.5\}$) para permitir a propagação suave de gradientes via *backpropagation*, seguindo a mesma estratégia de aplicada em funções como PReLU e ReLU.

3.2. Convoluções Deformáveis para séries Temporais

As convoluções padrão, amplamente utilizadas em arquiteturas como FCN, ResNet [Wang et al. 2017] e InceptionTime [Ismail Fawaz et al. 2020], operam sobre uma grade fixa de amostragem. Essa rigidez limita a capacidade do modelo de extrair características de dependências temporais de longo alcance ou de padrões que sofrem variações de fase e distorções, a menos que se utilize filtros muito grandes ou mecanismos de atenção global, ambos computacionalmente custosos. É possível observar na Figura 2 os diferentes mecanismos aplicados à uma série temporal.

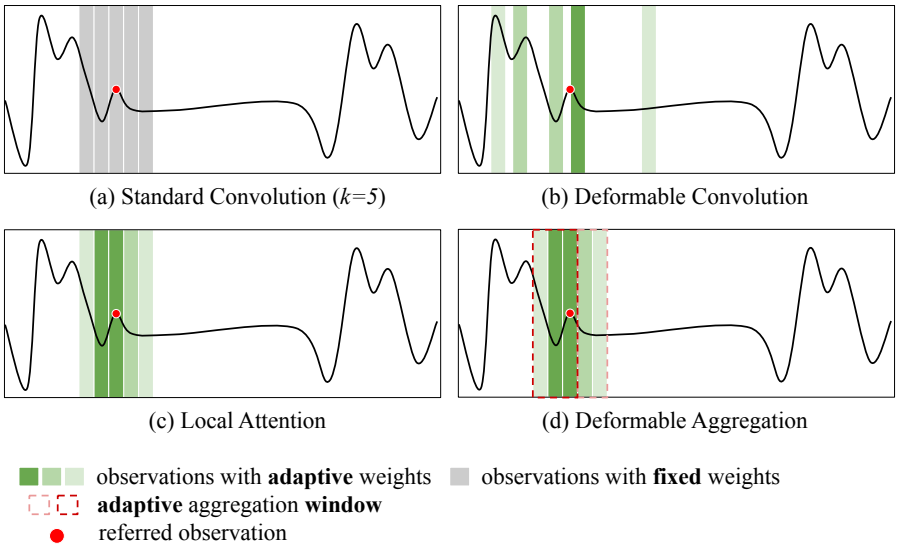


Figura 2. Comparação de diferentes mecanismos utilizados para extrair padrões temporais entre as observações de uma série temporal.

A inovação proposta nesta seção é a adaptação de convoluções deformáveis para o domínio unidimensional das séries temporais [Barbosa de Medeiros Júnior et al. 2026]. Diferente da convolução tradicional, a operação deformável introduz deslocamentos aprendíveis (*offsets*) Δp_n, e um componente de modulação Δm_k. Matematicamente, a operação é definida como:

$$y(p_0) = \sum_{p_n \in \mathcal{P}_k} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_k, \tag{2}$$

onde p_0 é a posição central, p_n enumera as posições na grade padrão e Δp_n representa o deslocamento que deforma essa grade. A modulação Δm_k funciona como um amplificador de amplitude para as características mais relevantes. Como os deslocamentos calculados são geralmente valores fracionários, a camada utiliza interpolação linear entre as observações vizinhas para estimar o valor do sinal na posição deformada.

Para que a rede aprenda onde "olhar" no sinal temporal, os offsets são gerados por uma camada adicional de Convolução Separável em Profundidade (DWSC) [Ismail-Fawaz et al. 2022] aplicada ao mapa de características anterior. Essa escolha técnica visa reduzir o número de parâmetros extras, mantendo a eficiência do modelo.

Um ponto crucial da metodologia é a aplicação da camada deformável ao final dos blocos convolucionais (conforme a arquitetura Deformable FCN ou DefFCN [Barbosa de Medeiros Júnior et al. 2026]). Experimentos de ablação demonstraram que posicionar a camada deformável onde o campo receptivo já é amplo permite que o modelo capture dependências de longo alcance de forma muito mais eficaz do que nas camadas iniciais.

3.3. Integração com o TGNet

O TGNet (Time Series Gaussian Mixture Model Network) é um framework ponta-a-ponta que unifica a extração de características temporais via *autoencoders* com a modelagem probabilística por Modelos de Mistura de Gaussianas (GMM). Diferente de métodos em dois estágios, onde a representação é aprendida de forma agnóstica à classificação, o TGNet otimiza ambos os componentes simultaneamente, garantindo que o espaço latente seja estruturado especificamente para a detecção de instâncias fora da distribuição (OOD).

A arquitetura baseia-se em um autoencoder profundo que incorpora as inovações das seções anteriores: três blocos convolucionais com ativação LeakySineLU, sendo o último bloco composto por Convoluções Deformáveis. O codificador ($e(\cdot)$) mapeia a série temporal de entrada para uma representação latente de baixa dimensão que captura a estrutura fundamental do sinal (z_c). O decodificador ($d(\cdot)$) reconstrói o sinal original (z_r). A partir dessa reconstrução, calcula-se um vetor de métricas de erro contendo o Erro Quadrático Médio (MSE) e a Similaridade de Cosseno entre a entrada e a saída.

O diferencial do TGNet é utilizar um vetor concatenado da similaridade de cosseno calculada e a representação latente ($z = [z_c, z_r]$), que fornece ao modelo probabilístico tanto as características estruturais do sinal quanto informações sobre a qualidade da sua reconstrução, facilitando a diferenciação da classe de interesse.

Um desafio crítico em modelos que integram GMMs a redes neurais (como o DAGMM [Zong et al. 2018]) é a instabilidade numérica, que frequentemente leva à explosão de gradientes ou ao colapso do modelo quando a matriz de covariância apresenta valores muito pequenos. Para mitigar isso, este trabalho propõe duas soluções técnicas:

- No cálculo da log-verossimilhança, a função exponencial tradicional é substituída pela explu ($\sigma(x) = \exp(x)$ para $x \leq 0$ e $x + 1$ caso contrário). Isso impede o crescimento descontrolado dos valores de energia quando $x > 0$, mantendo a continuidade e a magnitude necessárias para o aprendizado;

- A função de perda inclui um termo de penalização que evita soluções triviais onde a diagonal da matriz de covariância ($\hat{\Sigma}$) se aproxima de zero, garantindo a estabilidade durante todo o treinamento.

Durante a fase de teste, o TGNet estima a densidade de probabilidade da instância através da energia da amostra baseada na log-verossimilhança. Para evitar a necessidade de definição manual de parâmetros, o framework utiliza um limiar automático definido como um percentil das energias observadas no conjunto de treinamento. Amostras que apresentam energia superior a esse limiar (baixa probabilidade de pertencerem à distribuição aprendida) são classificadas como fora da distribuição. A Figura 3 mostra uma visão geral da arquitetura do TGNet.

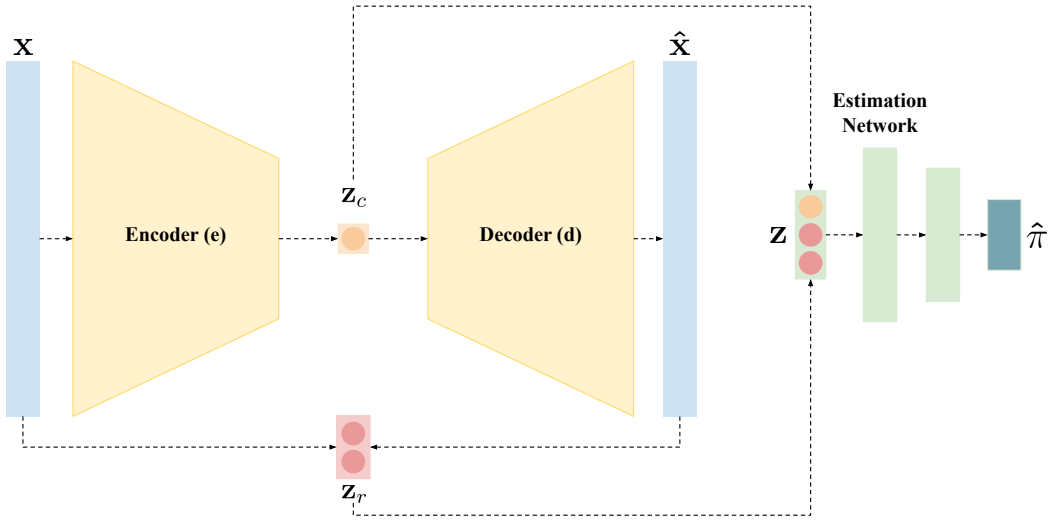


Figura 3. Visão geral da arquitetura do TGNet.

A função de perda total do TGNet ($\mathcal{L}_t$) é composta por três termos principais, equilibrados por metaparâmetros (λ_1 e λ_2), visando otimizar a reconstrução, a modelagem de densidade e a estabilidade das matrizes de covariância.

$$\mathcal{L}_t(\theta_e, \theta_d, \theta_m) = \frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_r(x_i, \hat{x}_i) + \frac{\lambda_1}{N}\sum_{i=1}^{N}E(z_i) + \lambda_2 \cdot P(\hat{\Sigma}), \tag{3}$$

onde, $\mathcal{L}_r$ é o erro de reconstrução e garante que o *autoencoder* aprenda a representar fielmente os padrões da classe de interesse, $E(z_i)$ é a energia da amostra z_i baseada na log-verossimilhança e utilizando a função explu [Golo et al. 2025], este termo minimiza a energia das instâncias de treinamento, e $P(\hat{\Sigma})$ é a regularização da matriz de covariância que penaliza valores excessivamente pequenos na diagonal da matriz de covariância, evitando que o modelo colapse em soluções triviais onde a variância se torna nula (problema de colapso da hiperesfera que acontece no DeepSVDD [Ruff et al. 2018]).

4. Resultados Relevantes

A avaliação experimental deste trabalho foi abrangente, utilizando 112 conjuntos de dados do repositório UCR para garantir a diversidade de domínios. No total, para a validação

do framework TGNet, foram realizados 33.110 experimentos, considerando cada classe de cada dataset como um problema independente de aprendizado de única classe.

4.1. Experimentos da LeakySineLU

A função proposta obteve o 1º lugar em ranque médio tanto em arquiteturas MLP (ranque $4,24$) quanto FCN (ranque $3,04$). Testes de Wilcoxon confirmaram que a LeakySineLU vence a ReLU e a PReLU na maioria dos datasets, demonstrando ser uma alternativa superior para a extração de características em sinais temporais. O diagrama de diferença crítica da Figura 4 resume os resultados dos experimentos de forma visual com MLPs, enquanto a Figura 5 resume os resultados utilizando a FCN, neste caso com as ativações que tiveram piores resultados retiradas da comparação.

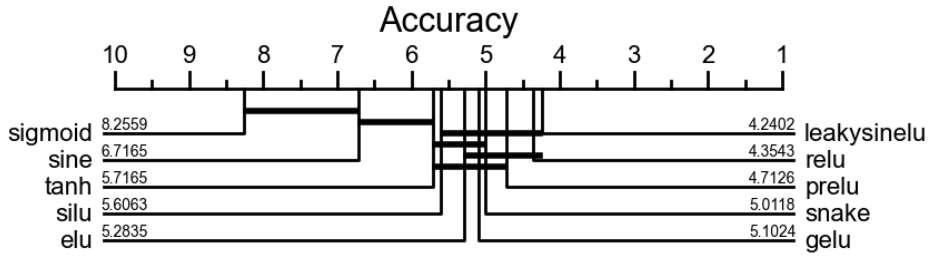


Figura 4. Diagrama de diferença crítica entre as ativações comparadas utilizando como base uma rede MLP [Wang et al. 2017].

4.2. Desempenho da Convolução Deformável

O modelo Deformable FCN alcançou o 2º lugar no ranqueamento geral $(2,55)$, superando estatisticamente a FCN tradicional, a ResNet e o Inception. Em estudos de caso com dados de saúde (ECG), o DefFCN superou o InceptionTime utilizando apenas 30% dos parâmetros, provando ser uma arquitetura altamente eficiente e capaz de capturar dependências de longo alcance fundamentais para diagnósticos médicos [Barbosa de Medeiros Júnior et al. 2026].

4.3. Resultados do TGNet

O framework final (TGNet com LeakySineLU e Convoluções Deformáveis) obteve o melhor ranque médio $(3,16)$ e a maior pontuação F1-macro média entre todos os competidores de OCL, incluindo Isolation Forest, OCSVM, DAGMM e outras redes. A análise do diagrama de diferença crítica confirmou a superioridade estatística do TGNet com um p-valor $\leq 10^{-4}$ em relação aos métodos tradicionais. A Figura 7 mostra o diagrama de

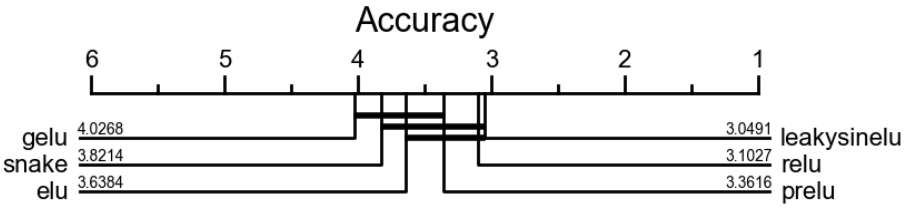


Figura 5. Diagrama de diferença crítica entre as ativações comparadas utilizando como base uma rede FCN [Wang et al. 2017].

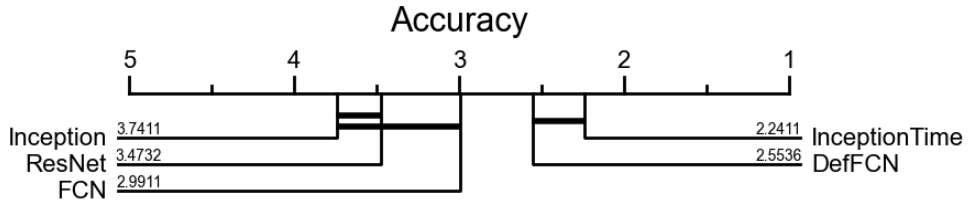


Figura 6. Diagrama de diferença crítica entre as redes para classificação de séries temporais e a DeformableFCN que utiliza convoluções deformáveis.

diferença crítica construído a partir do teste de Wilcoxon-Holm com $\alpha = 0.05$, contendo todos os métodos comparados.

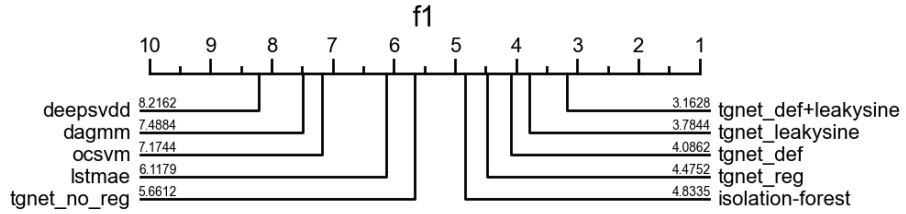


Figura 7. Diagrama de diferença crítica dos métodos de OCL. Os experimentos foram realizados nas variações de conjuntos de OCL a partir do repositório da UCR.

A partir do diagrama de diferença crítica exibido na Figura 8, é possível observar a influência e o impacto dos mecanismos adicionados, como a LeakySineLU e a convolução deformável, nos resultados do TGNet. Note que a versão com todos os mecanismos reflete o melhor resultado obtido. Vale ressaltar que a ausência de equivalências estatísticas se deve ao extensivo número de experimentos realizados, uma vez que, ao converter os 112 conjuntos de dados da UCR para problemas de OCL, o número de experimentos ultrapassa 30 mil execuções, como mencionado anteriormente.

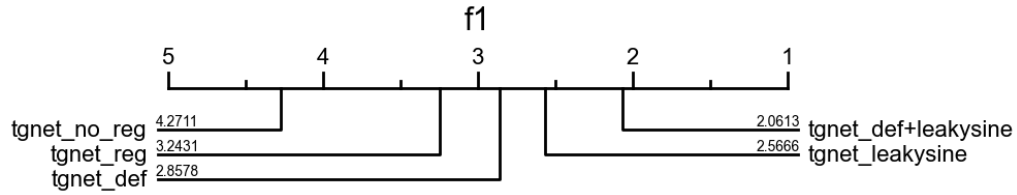


Figura 8. Diagrama de diferença crítica do estudo de ablação da influência de cada mecanismo agregado a arquitetura do TGNet.

4.4. Artigos Resultantes da Dissertação

A Tabela 1 mostra o resumo das publicações e produtos resultantes da dissertação. De acordo com a tabela, todas as publicações foram aceitas em importantes eventos e periódicos da área. Os algoritmos são todos de código aberto e estão disponíveis no github.

Tabela 1. Resumo das características das publicações.

Evento/Periódico	TA[a]/FI[b]	Ano	Qualis[c]	Downloads	Citações[d]
ACM SAC	23%	2026	A2	-	0
Information Sciences	6.0	2025	A1	-	3
BRACIS (best paper)	38%	2024	A4	-	2
CBMS	37.1%	2023	A3	-	3
Algoritmo	**Linguagem**	**Ano**	**Descrição**		
LeakySineLU	Python	2024	Ativação semi-periódica para séries temporais.		
Deformable FCN	Python	2026	Convoluções com offsets para séries temporais.		
TGNet	Python	2025	Framework para OCL em séries temporais.		
labicompare	Python	2026	Biblioteca para comparação de modelos.		

[a] Taxa de aceitação, [b] Fator de Impacto, [c] Qualis 2020, [d] Google Scholar

5. Conclusões

Esta dissertação desenvolveu um framework robusto para o Aprendizado de Única Classe (OCL) em séries temporais, superando limitações críticas de métodos anteriores, como a instabilidade numérica e a incapacidade de capturar dependências temporais complexas. O trabalho demonstrou que a eficácia do aprendizado profundo em dados sequenciais depende fundamentalmente da adaptação de mecanismos arquiteturais às propriedades específicas desses sinais. As três principais inovações técnicas: LeakySineLU, Convoluções Deformáveis, e o framework TGNet, representam avanços significativos para a área.

A LeakySineLU provou ser uma alternativa superior às funções de ativação tradicionais (como ReLU e PReLU), sendo capaz de capturar padrões sazonais e oscilatórios enquanto mantém o fluxo de gradiente estável em todo o domínio. As Convoluções Deformáveis permitiram que modelos leves capturassem dependências de longo alcance e padrões esparsos com alta precisão, atingindo desempenho comparável a arquiteturas muito mais densas (como o InceptionTime) utilizando apenas 30% de seus parâmetros. O framework TGNet consolidou essas inovações em uma solução ponta-a-ponta que utiliza Modelos de Mistura de Gaussianas (GMM) para uma detecção de dados fora da distribuição (OOD) estatisticamente fundamentada, evitando o colapso do modelo através de novas funções de perda e ativação.

A robustez das propostas foi validada através de uma das maiores baterias de testes em OCL para séries temporais registradas na literatura, englobando 33.110 experimentos em 112 conjuntos de dados. O TGNet alcançou o 1º lugar no ranqueamento estatístico médio, demonstrando superioridade tanto sobre métodos clássicos (Isolation Forest, OCSVM) quanto sobre modelos de Deep Learning de última geração.

Apesar dos avanços, o trabalho identificou oportunidades para pesquisas futuras: Os mecanismos propostos podem ser adaptados para capturar dependências não apenas no tempo, mas também entre as diferentes dimensões (canais) de séries multivariadas; A integração de técnicas como Normalizing Flows pode aumentar a flexibilidade do TGNet para lidar com dados que não seguem distribuições Gaussianas ideais; Pretende-se explorar a aplicação do framework em outros tipos de dados representados por vetores

numéricos, como texto e imagens, adaptando as camadas de extração de características.

Em suma, as contribuições desta dissertação tornam o uso de redes neurais mais viável e eficiente para aplicações reais de monitoramento e diagnóstico, especialmente em cenários de escassez de dados rotulados e recursos computacionais limitados.

Referências

Barbosa de Medeiros Júnior, J. G., de Mitri, A. G., and Silva, D. F. (2024). Semi-periodic activation for time series classification. In *Brazilian Conference on Intelligent Systems*, pages 76–90. Springer.

Barbosa de Medeiros Júnior, J. G., Guarnier de Mitri, A., Devanne, M., Forestier, G., and Furtado Silva, D. (2026). Deformable convolutions improve time series classification by capturing sparsed patterns. In *Proceedings of the 41st ACM/SIGAPP Symposium on Applied Computing*, pages 1065–1072.

Cabral, G. G. and de Oliveira, A. L. I. (2014). One-class classification for heart disease diagnosis. In *2014 IEEE International conference on systems, man, and cybernetics (SMC)*, pages 2551–2556. IEEE.

Fukunaga, K., Hayes, R. R., and Novak, L. M. (1987). The acquisition probability for a minimum distance one-class classifier. *IEEE Transactions on Aerospace and Electronic Systems*, (4):493–499.

Gao, L., Zhang, L., Liu, C., and Wu, S. (2020). Handling imbalanced medical image data: A deep-learning-based one-class classification approach. *Artificial intelligence in medicine*, 108:101935.

Golo, M. P. S., de Medeiros Junior, J. G. B., Silva, D. F., and Marcacini, R. M. (2025). One-class graph autoencoder: A new end-to-end, low-dimensional, and interpretable approach for node classification. *Information Sciences*, 708:122060.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034.

Ismail-Fawaz, A., Devanne, M., Weber, J., and Forestier, G. (2022). Deep learning for time series classification using new hand-crafted convolution filters. In *2022 IEEE International conference on big data (Big Data)*, pages 972–981. IEEE.

Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., and Muller, P.-A. (2019). Deep learning for time series classification: a review: H. ismail fawaz et al. *Data mining and knowledge discovery*, 33(4):917–963.

Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., Webb, G. I., Idoumghar, L., Muller, P.-A., and Petitjean, F. (2020). Inceptiontime: Finding alexnet for time series classification. *Data mining and knowledge discovery*, 34(6):1936–1962.

Provotar, O. I., Linder, Y. M., and Veres, M. M. (2019). Unsupervised anomaly detection in time series using lstm-based autoencoders. In *2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT)*, pages 513–517. IEEE.

Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., Müller, E., and Kloft, M. (2018). Deep one-class classification. In *International conference on machine learning*, pages 4393–4402. PMLR.

Silva, D. F., De M, J. G., Domingues, L. V., and Mazzu-Nascimento, T. (2023). Hemoglobin estimation from smartphone-based photoplethysmography with small data. In *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 75–78. IEEE.

Wang, Z., Yan, W., and Oates, T. (2017). Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pages 1578–1585. IEEE.

Yeh, C.-C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., Zimmerman, Z., Silva, D. F., Mueen, A., and Keogh, E. (2018). Time series joins, motifs, discords and shapelets: a unifying view that exploits the matrix profile. *Data Mining and Knowledge Discovery*, 32(1):83–123.

Ziyin, L., Hartwig, T., and Ueda, M. (2020). Neural networks fail to learn periodic functions and how to fix it. *Advances in Neural Information Processing Systems*, 33:1583–1594.

Zong, B., Song, Q., Min, M. R., Cheng, W., Lumezanu, C., Cho, D., and Chen, H. (2018). Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In *International conference on learning representations*.

Apêndice: Subprodutos (max 3 páginas)

Os subprodutos da dissertação são enumerados da seguinte forma:

1. **Pacotes de software** Todo o código fonte e dados utilizados no trabalho e artigos relacionados são distribuídos como código aberto:
    - `https://github.com/jose-gilberto/chronogam`
    - `https://github.com/jose-gilberto/leakysinelu`
    - `https://github.com/jose-gilberto/mhealth_cbms2023`
    - `https://github.com/jose-gilberto/defconv_time_series`
    - `https://github.com/jose-gilberto/labicompare`
2. **Publicações em Periódicos - JCR 6.0, Qualis A1:**
   [Golo et al. 2025]
   Este trabalho tem o objetivo de explorar a utilização de redes neurais de grafos para tarefas de classificação de única classe (OCL).
    - Ano de publicação: 2025
    - Information Sciences (link - `https://www.sciencedirect.com/journal/information-sciences`)
    - Citações do artigo: `https://github.com/GoloMarcos/OLGA` 3 citações;
    - Link: `https://doi.org/10.1016/j.ins.2025.122060`
3. **Publicação em Evento - Hindex 80, Qualis A2:**
   [Barbosa de Medeiros Júnior et al. 2026]
   Este trabalho propôs a utilização de convoluções deformáveis na tarefa de classificação de séries temporais, incorporando seu uso em redes já conhecidas para a tarefa e analisando seus resultados comparados com o estado-da-arte para classificação de séries temporais.
    - Ano do Evento: 2026
    - ACM/SIGAPP Symposium On Applied Computing (link do evento - por exemplo `https://www.sigapp.org/sac/sac2026/index.php`) 23% acceptance rate
    - Citações do trabalho: 0 citações;
    - Link: `https://doi.org/10.1145/3748522.3780002`
4. **Publicação em Evento - Hindex 25, Qualis A4:**
   [Barbosa de Medeiros Júnior et al. 2024]
   Este trabalho propôs a função de ativação semi-periódica LeakySineLU, utilizada para tarefas de classificação de séries temporais, incorporando seu uso em redes neurais já conhecidas e comparando em 112 conjuntos de dados com outras ativações da literatura. Foi premiado como *best paper* da conferência.
    - Ano do Evento: 2024

- Brazilian Conference on Intelligent Systems (link do evento - por exemplo https://link.springer.com/book/10.1007/978-3-031-79029-4) 38% acceptance rate
- Citações do trabalho: 2 citações;
- Link: https://doi.org/10.1007/978-3-031-79038-6_6

5. **Publicação em Evento - Hindex 45, Qualis A3:**
   [Silva et al. 2023]
   Este trabalho propôs a utilização de redes neurais para classificar sinais de fotopletismografia capturados utilizando uma câmera de *smartphone* e sua lanterna em pacientes para estimar a hemoglobina.
   - Ano do Evento: 2023
   - International Symposium on Computer-Based Medical Systems (link do evento - por exemplo https://ieeexplore.ieee.org/xpl/conhome/989/proceeding) 37.1% acceptance rate
   - Citações do trabalho: 3 citações;
   - Link: https://doi.org/10.1109/CBMS58004.2023.00195