Graph-based novelty detection in passive sonar

Matheus Bastos de Oliveira[1], **João Baptista de Oliveira e Souza Filho**[1]

[1]Programa de Engenharia Elétrica
Universidade Federal do Rio de Janeiro (UFRJ) – Rio de Janeiro, RJ – Brasil

{oliveiraa.maatheus,jbfilho}@poli.ufrj.br

Abstract. *Avanços recentes têm demonstrado o potencial dos grafos em tarefas de detecção de novidades, mas a sua aplicação em sinais de sonar passivo no contexto de Reconhecimento em Conjunto Aberto (OSR) permanece inexplorada. Este artigo avalia experimentalmente o uso de técnicas de grafos para este problema, com foco na geração de redes e atributos nodais, especialmente o uso de Grafos de Visibilidade Espectral (SVG). Os resultados mostram que a extração estrutural baseada em SVG, integrada à penalização de conexões de longo alcance, atinge 68,52% de AUROC. O método proposto supera abordagens baseadas em caminhadas aleatórias em 17 pontos percentuais e se equipara a redes neurais profundas com um terço do tempo de processamento.*

Abstract. *Recent advances have demonstrated the potential of graph-based methods in novelty detection tasks, but their application to passive sonar signals within the context of Open Set Recognition (OSR) remains unexplored. This paper experimentally evaluates the use of graph-based techniques for this problem, focusing on the generation of networks and nodal features, especially the use of Spectral Visibility Graphs (SVG). Results show that structural extraction based on SVGs, integrated with the penalization of long-range connections, achieves an AUROC of 68.52%. The proposed method outperforms random walk-based approaches by 17 percentage points and performs comparably to deep neural networks at one-third of the processing time.*

Palavras-chave: Sonar Passivo, *Open Set Recognition*, Grafos de Visibilidade Espectral, Processamento de Sinais em Grafos, Enriquecimento de Atributos.

1. Introdução

A monitorização de embarcações por meio de sistemas de sonar passivo é essencial para a vigilância marítima. Em cenários operacionais, esses sistemas enfrentam o problema de Reconhecimento em Conjunto Aberto (*Open Set Recognition* - OSR). Neste regime, o modelo deve classificar de forma acurada as assinaturas acústicas de classes conhecidas (*In-Distribution*) e, simultaneamente, detectar classes não vistas ao longo do treinamento (*Out-of-Distribution*).

Tradicionalmente, a detecção de novidades em sonar passivo explora classificadores de classe única, em particular por meio de técnicas como o *One-Class* SVM [Moura Junior 2018], e, mais recentemente, redes auto-codificadoras (*autoencoders*) convolucionais [Honorato et al. 2020] e *ensembles* hierárquicos de classificadores k-NN [Honorato et al. 2022].

Trabalhos recentes [Ma et al. 2023] têm mostrado que a modelagem topológica é altamente eficaz para isolar instâncias anômalas devido à sua capacidade de mapear dependências relacionais e explorar vieses indutivos estruturais que os espaços vetoriais explorados por técnicas clássicas comumente não capturam.

No campo da modelagem de sinais, o uso de grafos foi impulsionado por [Lacasa et al. 2008], que introduziu os Grafos de Visibilidade para mapear séries temporais em redes. Quanto ao domínio espectral, [Yela et al. 2019] propuseram os Grafos de Visibilidade Espectral (SVG) para extrair similaridades harmônicas. Em paralelo, o uso de Redes Neurais Convolucionais em Grafos (GCN) [Kipf and Welling 2016] e Redes de Isomorfismo em Grafos (GIN) [Xu et al. 2018] permitiu a agregação indutiva de atributos nodais. Recentemente, a aplicação dessas arquiteturas ao monitoramento de navios em conjunto aberto foi explorada por [Yichen et al. 2025] utilizando aprendizado de protótipos com o uso de GCNs. Contudo, enquanto [Yichen et al. 2025] constrói grafos baseados nos sinais de ruído no domínio do tempo, há uma carência de abordagens em OSR que explorem as relações topológicas extraídas diretamente do domínio espectral, portanto capazes de evidenciar as similaridades harmônicas.

A motivação central deste trabalho é a hipótese de que a modelagem em grafos de sinais acústicos, explorando aspectos como a conectividade e a ressonância espectral, contexto em que os harmônicos atuam como *hubs* estruturais, possa favorecer a derivação de descritores mais discriminativos para tarefas de OSR. Assim, este trabalho propõe avaliar estratégias de detecção de novidades em sonar passivo baseadas em representações topológicas dos dados. Diferentes aspectos são contemplados, desde o processo de geração das redes, realizado por Grafos de Visibilidade Espectral (SVG) e grafos de vizinhança (k-NN), combinado aos descritores locais e *embeddings* considerados, avaliando-se o uso de redes neurais do tipo GCN e GIN. Neste processo, uma atenção especial é dedicada ao processo de formação dos grafos e de representação dos nós, que pode explorar uma limiarização geométrica de conectividade (*cut-off*), um processo de filtragem espacial não-linear (SDW - *Spatial Decay Weighting*), bem como uma codificação posicional de nós logarítmica (*Positional Encoding*).

2. Materiais e métodos

O método proposto envolve três etapas principais para a avaliação da efetividade do uso de grafos para a produção de representações enriquecidas dos sinais de sonar, em contraposição à identificação de novidades realizada diretamente sobre os espectros apenas pré-processados, comumente adotada na literatura ([Moura Junior 2018] e [Honorato et al. 2020]). A Figura 1 ilustra a conversão do espectro acústico pré-processado em grafo, a extração e normalização dos atributos, e a geração da assinatura enriquecida que alimenta o detector de novidades.

2.1. Construção do grafo

Envolve a conversão da assinatura espectral, gerada conforme [Muniz et al. 2019], em um grafo. Modelar os *bins* de frequência como nós de uma rede permite capturar explicitamente dependências entre diferentes frequências, tais como harmônicos, padrões de ressonância e relações de energia entre bandas. Duas estratégias de modelagem foram consideradas neste trabalho:

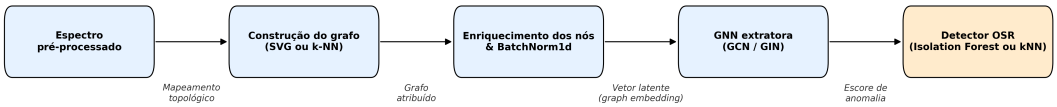


Figura 1. Arquitetura geral da metodologia proposta.

- **Geométrica** (**kNN-*graph***): neste caso, cada *bin* de frequência corresponde a um nó, que se conecta aos seus k vizinhos mais próximos em um espaço bidimensional cujas coordenadas são definidas pelo valor desta frequência e sua magnitude. Para a construção destas redes topológicas, a restrição de vizinhança foi fixada em $k = 5$. Ambas as coordenadas são submetidas a uma normalização min-max prévia para evitar a dominância de uma delas.
- **Visibilidade (SVG)**: é adotado o Grafo de Visibilidade Espectral [Yela et al. 2019], que estabelece uma aresta entre dois nós caso exista uma "linha de visão" direta e desobstruída entre suas magnitudes. Esta regra garante que os picos de energia do sinal (frequências dominantes) atuem naturalmente como *hubs* estruturais de alta conexão, conforme ilustrado na Figura 2.

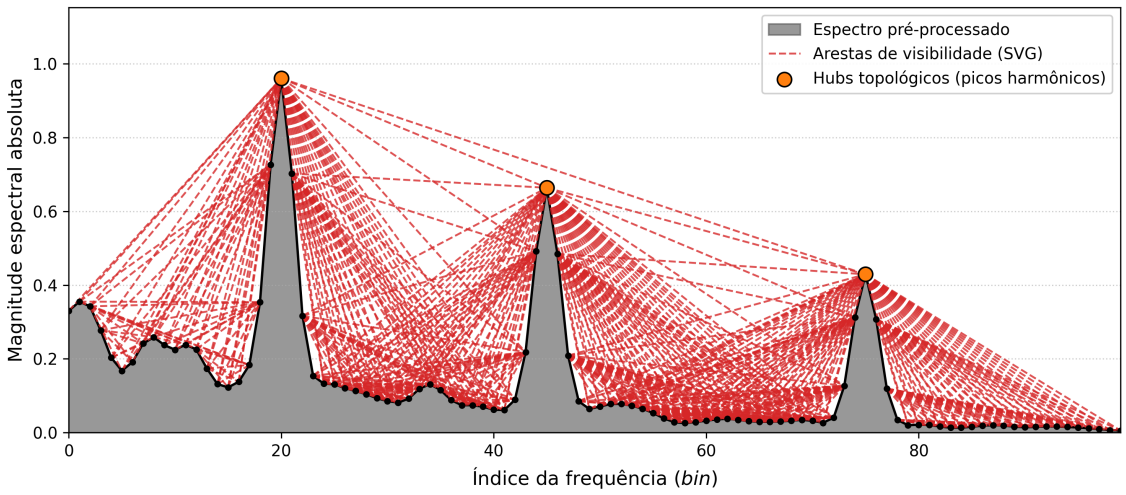


Figura 2. Ilustração das linhas de visão no Grafo de Visibilidade Espectral (SVG).

Em ambas as abordagens, para priorizar as componentes de maior energia no grafo, o peso da aresta entre dois nós é definido pelo produto de suas magnitudes ($W_{u,v} = M_u \cdot M_v$). Essa formulação favorece o acoplamento entre frequências dominantes e reduz a influência do ruído de fundo na extração de atributos.

2.2. Extração e enriquecimento de atributos

Cada nó possui como representação inicial um *embedding* constituído por quatro atributos: Magnitude, Grau Ponderado, Coeficiente de Aglomeração Ponderado (*Clustering*) e Grau de Centralidade.

Diferentes estratégias foram exploradas para a geração do *embedding* do grafo resultante, o qual corresponde à assinatura enriquecida base da identificação de novidades:

- **Descritores locais (*Structural Mean Pooling*)**: corresponde à agregação global (média) dos atributos nodais, definindo uma representação base (*baseline*).
- **Descritores globais (*Graph Structural*)**: define um vetor de características constituído por seis métricas globais: densidade da rede, grau médio, grau máximo, coeficiente de agrupamento médio global, magnitude média e magnitude máxima do espectro.
- **Representações nodais clássicas (*Spectral Embedding* e *Node2Vec*)**: O *Spectral Embedding* produz representações nodais pela decomposição espectral do Laplaciano, enquanto o *Node2Vec* é baseado em passeios aleatórios. Quanto ao último, buscou-se priorizar a exploração de vizinhanças, adotando-se uma penalidade de retorno p igual a $2,0$ e de avanço q igual a $0,5$. Foi ainda fixado um limiar de ocorrência unitário (*min_count* = 1) e a janela de contexto foi limitada a um terço do passeio ($w = 5$). A geração do *embedding* do grafo foi realizada por *Average Pooling*.
- **Aprendizado Profundo em grafos (GNNs):** os experimentos consideraram as arquiteturas convolucional (GCN) e de isomorfismo (GIN). Para a GCN, foram adotadas três camadas convolucionais ocultas parametrizadas com 32 dimensões ($hidden_channels = 32$), definindo-se um campo receptivo de três saltos. Para a GIN, utilizaram-se duas camadas de isomorfismo baseadas em MLP, também com uma dimensionalidade oculta fixada em 32. Em ambos os casos, a dimensão do vetor latente de saída (d) foi tratada como um hiperparâmetro sintonizável. Ambas as configurações foram dimensionadas para explorar a informação da vizinhança harmônica sem a introdução de *over-smoothing*. As GNNs foram treinadas via *Cross-Entropy* para discriminar as classes conhecidas, descartando-se o cabeçote de classificação após a convergência. Na inferência, a agregação dos *embeddings* dos nós para a geração da representação do grafo utilizou *Max Pooling*, seguida de uma padronização estatística via normalização Z-score (*StandardScaler*).

2.3. Detectores de novidade e calibração OSR

Dois modelos amplamente utilizados de identificação de novidades foram avaliados neste trabalho: o *k-Nearest Neighbors* (k-NN) e o *Isolation Forest* [Liu et al. 2008]. Bons resultados têm sido reportados para o k-NN em sinais de sonar passivo [Honorato et al. 2022]. A técnica *Isolation Forest* é amplamente utilizada na literatura em problemas de outros domínios de aplicação e ainda não tinha sido explorada neste contexto.

O número de vizinhos do k-NN considerado nos experimentos seguiu a regra empírica $k \approx \sqrt{N_{\text{treino}}}$ para bases densas [Duda et al. 2000]. A escolha dos melhores hiperparâmetros de ambos os detectores foi baseada na estratégia de busca em grade, cujos conjuntos de valores são detalhados na Seção 3, adotando-se como critério de otimização a maximização da Área Sob a Curva ROC (AUROC) [Fawcett 2006] computada sobre o conjunto de validação interna. Cabe destacar que o escore OSR final de cada assinatura submetida ao k-NN foi definido como a mediana das distâncias euclidianas em relação aos seus k vizinhos mais próximos no conjunto de treinamento *In-Distribution*.

2.4. Refinamento de atributos de representação dos grafos

Devido à susceptibilidade dos grafos de visibilidade a ruídos de banda-larga, que podem levar a efeitos de *over-smoothing*, e à perda de localização absoluta das frequências das

assinaturas espectrais causada pela invariância à permutação, inerente às convoluções em grafos, foram avaliadas as seguintes estratégias de modificação dos grafos explorados no treinamento e operação das GNNs:

- **GCN padrão**: consistiu na adoção do grafo original (*baseline* profundo).
- **GCN + *cut-off* (limiarização rígida de distância)**: neste caso foi imposto como raio de conexão máxima o valor de 8 *bins*. Este valor foi baseado em um estudo que considerou o percentil 90 da Largura a Meia Altura (*Full Width at Half Maximum* - FWHM) dos picos de energia, estimado sobre as assinaturas da Classe A do conjunto de treinamento (primeiro *fold*).
- **GCN + SDW (*Spatial Decay Weighting*)**: consistiu na aplicação de uma penalidade aos pesos do grafo, de forma inversamente proporcional ao quadrado da distância entre os nós, definida por $W_{u,v} = \dfrac{M_u \cdot M_v}{d_{u,v}^2}$, onde u e v representam os nós, M é a energia e $d_{u,v}$ é a distância geométrica em *bins*. Este decaimento quadrático visa preservar o acoplamento estrutural entre as vizinhanças imediatas e atenuar, substancialmente, as conexões espúrias de longa distância.
- **GCN + PE (*Positional Encoding* Logarítmico)**: corresponde à inserção de contexto espacial no vetor de atributos locais do i-ésimo nó, fornecendo uma indexação absoluta à rede. Trata-se de um mapeamento logarítmico que minimiza os efeitos de cauda ruidosa associados às altas frequências. As coordenadas posicionais são dadas por $PE_i = \ln(1+i)/\ln(N)$ para $N > 1$, sendo 0 caso contrário, onde $i \in \{0, 1, \ldots, N-1\}$ é o índice discreto do *bin* e N a dimensão do espectro.
- **Configurações combinadas**: trata-se de arranjos que integram simultaneamente a codificação posicional logarítmica (PE) com restrições de conectividade geométrica (*cut-off* ou SDW).

Por sua vez, nas representações globais dos grafos geradas pela estratégia *Graph Structural*, a informação da localização espacial das frequências é suprimida durante a agregação. Para resgatar essa localização em nível global, propõe-se um sétimo descritor para o grafo, denominado Centro de Massa Topológico (C_{massa}), que visa identificar onde a densidade de nós altamente conectados está concentrada, e é calculado por $C_{\text{massa}} = \sum_{i=1}^{N}(f_{\text{norm}_i} \cdot \text{grau}_i)/\sum_{i=1}^{N}\text{grau}_i$, onde f_{norm_i} representa a frequência normalizada e grau_i denota o grau ponderado do i-ésimo nó na rede.

3. Base de dados e protocolo experimental

O conjunto de dados do estudo foi fornecido pela Marinha do Brasil, sendo constituído por 29.277 assinaturas espectrais pré-processadas de 28 embarcações distribuídas em oito classes: Classe A (2.432 amostras), Classe B (3.432), Classe C (4.797), Classe D (3.072), Classe E (7.075), Classe F (2.934), Classe G (2.143) e Classe H (3.392). A base de dados possui acesso restrito por questões de segurança, o que impossibilita o compartilhamento público das assinaturas bem como dos códigos das aplicações. A geração das assinaturas espectrais-base para a avaliação dos métodos considerou o pré-processamento descrito em [Muniz et al. 2019]. Em síntese, após a remoção do ruído de ambiente, são obtidas assinaturas espectrais dos sinais coletados com energia unitária, as quais são armazenadas

em vetores de 557 dimensões cobrindo a faixa de frequência de 0 a 3 kHz. Consequentemente, cada vetor espectral foi modelado como um grafo contendo 557 nós.

A métrica de desempenho foi a Área Sob a Curva ROC (AUROC), em conformidade com estudos correlatos de [Honorato et al. 2022]. A avaliação adotou uma Validação Cruzada Aninhada (*Nested Cross-Validation*) em dois níveis. A formação dos *folds* considerou uma estratificação por navios, com o objetivo de mitigar o vazamento de dados. O particionamento externo operou com 10 *folds*; portanto, 10% dos dados foram reservados para o teste final em cada rodada. Já o nível interno aplicou uma divisão simples (*hold-out*), utilizando 9% dos dados para a validação e os 81% restantes para o treinamento.

Na formulação do problema de Reconhecimento em Conjunto Aberto (OSR), foram assumidas as Classes A, B, C, F e G como conhecidas (*In-Distribution*) e as Classes D, E e H como novidades (*Out-of-Distribution*), em conformidade com [Honorato et al. 2020]. O protocolo experimental consistiu das etapas apresentadas na Tabela 1. Na etapa 1, o extrator GNN foi treinado para discriminar as classes conhecidas A, B, C e F. Nas etapas seguintes, essas classes conhecidas definiram a classe negativa. Na etapa 2, a Classe G assumiu o rótulo de classe desconhecida (*proxy* de novidade), visando à sintonia dos hiperparâmetros dos detectores. Por fim, na etapa 3, que consiste do teste OSR, as classes não vistas D, E e H foram agrupadas e associadas à classe positiva (novidades). Esta estratégia produziu um conjunto de teste balanceado, contendo 13.595 instâncias negativas e 13.539 positivas.

Tabela 1. Resumo do protocolo experimental adotado.

Etapa	Objetivo principal	Classes conhecidas	Novidades
1. Pré-treinamento	Aprendizado do extrator (GNN) sob regime multiclasse.	A, B, C e F	Não aplicável
2. Calibração	Sintonia dos hiperparâmetros dos detectores.	A, B, C e F	G (*proxy*)
3. Teste OSR	Avaliação final do modelo.	A, B, C e F	D, E e H

Para a sintonia dos hiperparâmetros, o espaço de busca para a dimensão do vetor de atributos nos métodos *Spectral Embedding*, *Node2Vec* e GNNs foi definido como $d \in \{8, 16, 32\}$. Para os detectores, foram avaliados o número de vizinhos $k \in \{10, 50, 100\}$ no k-NN e a taxa de contaminação $c \in \{0,01; 0,05; 0,10\}$ no *Isolation Forest*.

A otimização dos pesos das redes empregou o algoritmo *AdamW* com uma taxa de decaimento (*weight decay*) de 1×10^{-3} para regularização. A taxa de aprendizado máxima $(0,005)$ foi precedida por um aquecimento linear (*linear warmup*) de 5 épocas, visando à estabilização inicial dos gradientes. Após esse período, a taxa foi modulada pelo escalonador *Cosine Annealing* para refinar a convergência, com limite de 1.000 épocas. Durante o pré-treinamento do extrator GNN, a convergência da rede foi monitorada por meio do conjunto de treinamento, sendo o processo interrompido quando a variação da função de perda *Cross-Entropy* se mantinha abaixo de 1×10^{-4} por 30 épocas consecutivas.

Para garantir a reprodutibilidade dos experimentos, todas as fontes de pseudoaleatoriedade foram fixadas globalmente (*seed* = 42). Devido a restrições computacionais, as avaliações foram conduzidas utilizando uma única semente. A comparação estatística dos

modelos utilizou o teste não-paramétrico de Friedman [Han et al. 2022] sobre os valores da AUROC, enquanto as comparações pareadas entre modelos fizeram uso do teste *post-hoc* de Nemenyi [Nemenyi 1963, Demšar 2006], adotando-se, em ambos os casos, um nível de significância de 5% ($\alpha = 0,05$). Como ferramenta para a análise e visualização dos resultados, foi utilizado o Diagrama de Diferença Crítica (CDD) [Demšar 2006].

O treinamento dos modelos utilizou uma estação de trabalho equipada com um processador Intel Core i7-4790K (4 núcleos, 8 *threads*, 4,00 GHz) e uma GPU NVIDIA GeForce RTX 2080 Ti (11 GB de VRAM), sob o ecossistema PyTorch 2.7.1 e CUDA 11.8. As arquiteturas GCN e GIN foram implementadas via *PyTorch Geometric* (PyG). Para os detectores (k-NN e *Isolation Forest*) e para as métricas de avaliação, foi utilizada a biblioteca *Scikit-Learn*. Por fim, para evitar a ociosidade da GPU (*starvation*) ante o elevado custo da modelagem topológica, a conversão dos espectros em grafos foi paralelizada assincronamente em 8 *threads* de CPU por meio do pacote *Joblib*.

4. Resultados

A Tabela 2 apresenta a moda dos hiperparâmetros selecionados na etapa de calibração. As arquiteturas GCN e GIN mantiveram estabilidade dimensional para as duas modalidades de formação dos grafos ($d = 8$ e $d = 16$, respectivamente), justificando manter a GCN com $d = 8$ nos experimentos de refinamento. Em contrapartida, para as representações *Spectral Embedding* e *Node2Vec*, a formação SVG está associada a espaços latentes de maior dimensão. Quanto aos classificadores, a topologia SVG permitiu ao k-NN discriminar a classe *proxy* com vizinhanças locais ($k = 10$), enquanto o kNN-*graph* exigiu um escopo estendido ($k = 100$). O *Isolation Forest* manteve a fronteira de decisão constante em $c = 0,01$, imune ao critério de formação do grafo.

Tabela 2. Valores predominantes dos hiperparâmetros selecionados, correspondentes à moda (maioria dos *folds*) na etapa de calibração.

Método de representação (*embedding*)	**SVG**			**kNN-*graph***		
	Dim. (d)	**Detectores**		**Dim. (d)**	**Detectores**	
		k-NN (k)	**IF (c)**		**k-NN (k)**	**IF (c)**
GCN	8	10	0,01	8	100	0,01
GIN	16	10	0,01	16	100	0,01
Spectral	16	10	0,01	8	100	0,01
Node2Vec	32	10	0,01	16	10	0,01
Graph Structural	–	10	0,01	–	100	0,01
Structural Mean Pooling	–	10	0,01	–	10	0,01

4.1. Avaliação comparativa das abordagens baseadas em grafos

A Tabela 3 consolida os valores de AUROC e o tempo computacional de cada arranjo experimental baseado em grafos. Os descritores estruturais diretos superaram as abordagens profundas e espectrais, enquanto o arranjo SVG + *Structural Mean Pooling* + *Isolation Forest* atingiu o melhor desempenho global com custo computacional inferior ao típico das redes profundas (19,17 contra $\sim$60 minutos). Este resultado supera em aproximadamente 17 pontos percentuais a abordagem *Node2Vec*, que obteve o menor desempenho e o maior custo computacional, superior a 18 horas por *fold*.

Tabela 3. Resultados da avaliação experimental OSR em 10 *folds* aninhados.

Construção do grafo	Método de representação (*embedding*)	AUROC (%)		Tempo médio por *fold* (min)
		k-NN	*Isolation Forest*	
SVG	***Structural Mean Pooling***	65,02 ± 0,76	**68,25 ± 0,94**	19,17
SVG	*Graph Structural*	66,34 ± 0,98	67,85 ± 0,95	19,01
SVG	**GCN**	**67,34 ± 1,97**	65,97 ± 1,67	64,20
SVG	GIN	60,41 ± 2,42	59,36 ± 2,39	62,64
SVG	*Spectral*	57,21 ± 0,73	57,65 ± 0,90	21,72
SVG	*Node2Vec*	46,85 ± 2,16	50,89 ± 1,38	1.103,90
kNN-*graph*	*Graph Structural*	66,07 ± 0,60	67,20 ± 0,80	13,18
kNN-*graph*	*Structural Mean Pooling*	65,85 ± 0,84	65,67 ± 0,82	13,26
kNN-*graph*	GIN	63,82 ± 2,57	64,64 ± 2,50	53,33
kNN-*graph*	GCN	61,40 ± 2,31	63,11 ± 1,80	55,14
kNN-*graph*	*Spectral*	53,42 ± 1,09	51,58 ± 0,98	30,93
kNN-*graph*	*Node2Vec*	51,30 ± 0,68	51,36 ± 0,65	1.123,04

Na Figura 3 são apresentados os boxplots associados à métrica AUROC para as diferentes comparações experimentais. É possível observar uma menor disparidade, face às menores distâncias interquartis, entre as representações baseadas em descritores relacionais diretos (SVG, *Structural Mean Pooling* e *Isolation Forest*) do que entre as de aprendizado profundo (GCN e GIN).

Em relação aos resultados, o teste de Bartlett [Montgomery and Runger 2010] indicou a rejeição da hipótese nula de homocedasticidade ($\chi^2(23) = 97,41$, $p < 0,001$), enquanto o teste não-paramétrico de Friedman ($\chi^2(23) = 213,15$, $p < 0,001$) indicou diferenças de desempenho estatisticamente significativas entre os métodos. A Figura 4 apresenta o CDD desta análise. É possível perceber um melhor ranqueamento da combinação SVG+*Structural Mean Pooling*+*Isolation Forest*, porém seus resultados são estatisticamente equivalentes a outras onze combinações de métodos.

4.2. Avaliação do refinamento de atributos dos grafos

Embora a representação *Structural Mean Pooling* tenha o melhor ranqueamento absoluto na avaliação inicial, o teste de Nemenyi (CDD) comprovou sua equivalência estatística com o método *Graph Structural*. Sendo configurações estatisticamente empatadas, o estudo focado no refinamento de atributos dos grafos foi conduzido sobre esta última e as redes GCN. Essa escolha metodológica justifica-se pois métricas globais puras (como densidade e grau médio) são intrinsecamente mais suscetíveis à degradação por conexões espúrias de longo alcance. Portanto, avaliar os mecanismos de filtragem (como SDW e *cut-off*) sobre essa representação permite isolar e validar a eficácia empírica das penalizações sem o viés prévio da agregação de descritores locais. A Tabela 4 resume os resultados para as estratégias de refinamento dos atributos apresentadas na Subseção 2.4.

Na representação estrutural direta, o uso do peso de distâncias quadráticas (SDW) sobre os descritores globais aumentou a AUROC em 0,67 pontos percentuais em relação à alternativa original (de 67,85% para 68,52%), atingindo o melhor desempenho entre as abordagens em grafos avaliadas. Este ganho, contudo, custou mais do que o dobro do tempo de processamento por *fold* (de 19,01 para 45,88 minutos). Por sua vez, a restrição rígida de conexões (*cut-off* em 8) reduziu a métrica para 65,46%, indicando

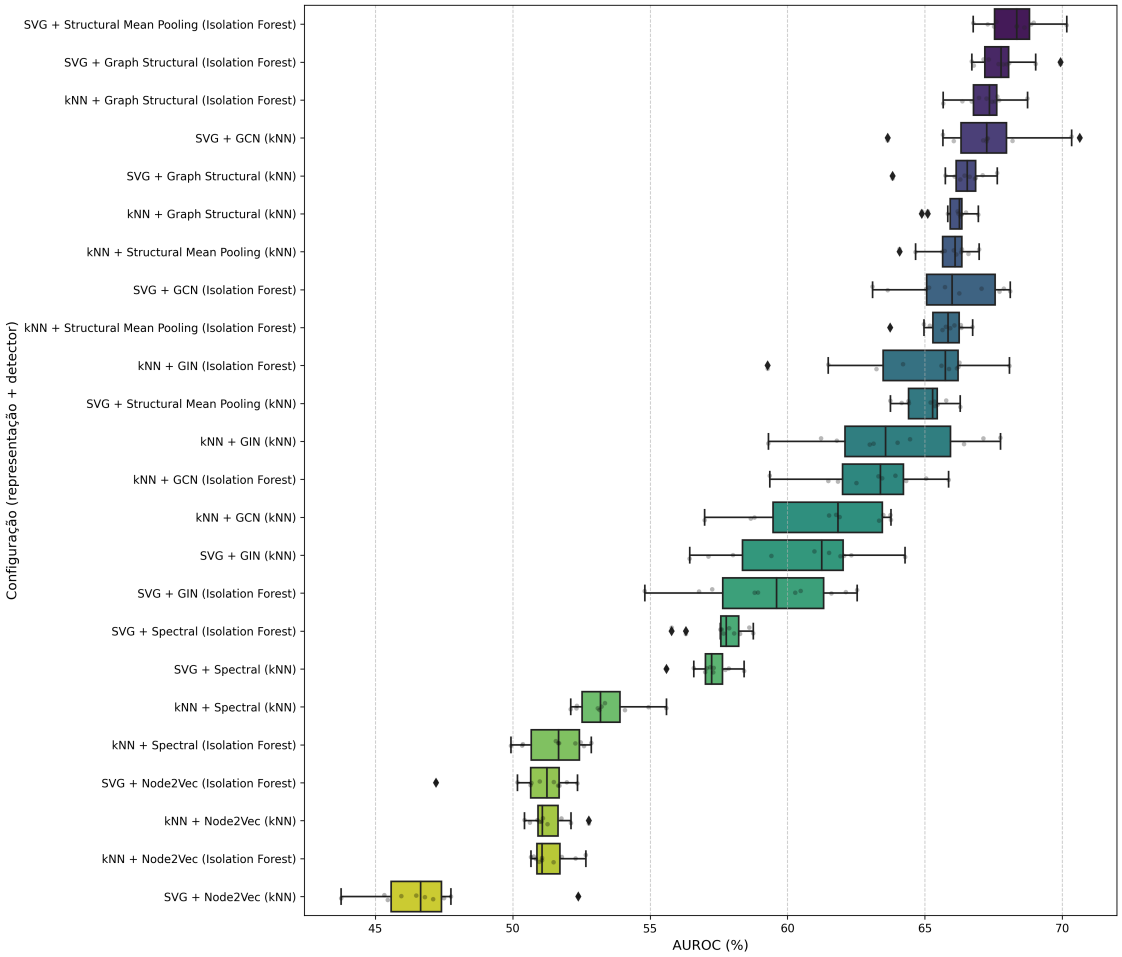


Figura 3. Distribuição da AUROC nos 10 *folds*, ordenada pela mediana, evidenciando a estabilidade dos descritores estruturais diretos frente à maior variância das arquiteturas profundas.

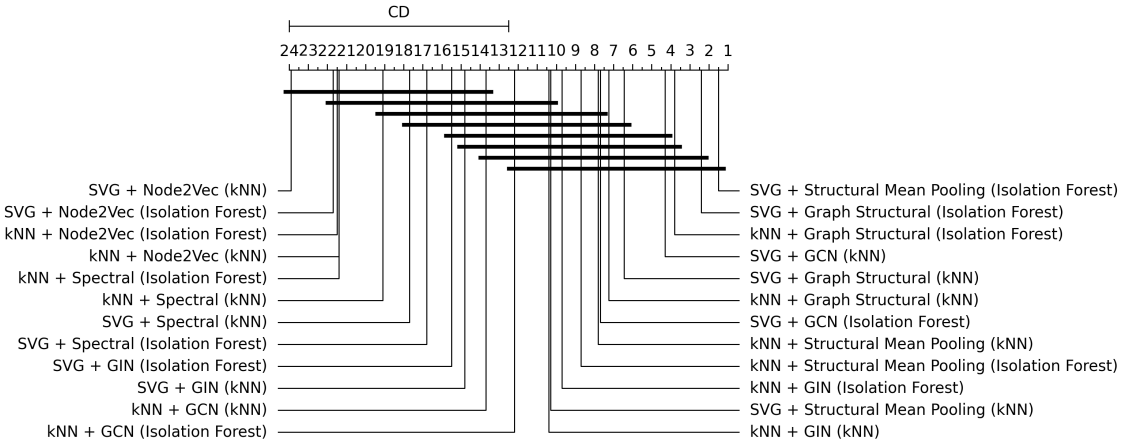


Figura 4. Diagrama de Diferença Crítica (CDD) para a AUROC dos experimentos. Linhas horizontais espessas conectam configurações sem diferença estatística significativa.

Tabela 4. Avaliação do impacto dos mecanismos de enriquecimento estrutural-espacial (refinamento de atributos) considerando o classificador *Isolation Forest*, com resultados ordenados por desempenho decrescente de AUROC. O tempo reflete o custo computacional médio por *fold*.

Construção do grafo	Método de representação (*embedding*)	AUROC (%)	Tempo médio por *fold* (min)
Representação estrutural direta			
SVG + SDW	***Graph Structural***	**68,52 ± 1,17**	45,88
SVG	*Graph Structural*	67,85 ± 0,95	19,01
SVG + SDW	*Graph Structural* + Centro de massa	68,16 ± 0,87	46,70
SVG	*Graph Structural* + Centro de massa	67,57 ± 1,16	46,26
SVG + *cut-off* (dist. máx. = 8)	*Graph Structural*	65,46 ± 0,89	44,11
Aprendizado Profundo em grafos			
SVG	GCN	65,97 ± 1,67	64,20
SVG + *cut-off* (dist. máx. = 8)	GCN	65,59 ± 3,19	209,13
SVG + *cut-off* (dist. máx. = 8)	GCN + *Positional Encoding* (PE)	65,03 ± 4,26	292,78
SVG + SDW	GCN	64,79 ± 2,40	316,08
SVG	GCN + *Positional Encoding* (PE)	61,93 ± 1,68	352,67
SVG + SDW	GCN + *Positional Encoding* (PE)	57,05 ± 2,88	334,28

que a limitação artificial de arestas suprime propriedades estruturais relevantes para a caracterização acústica. Adicionalmente, a incorporação do Centro de Massa Topológico não produziu ganhos relevantes quando combinada ao SDW.

Em relação às arquiteturas profundas, a injeção de *Positional Encoding* (PE) na GCN não contribuiu para uma maior capacidade discriminativa. O uso simultâneo de PE e SDW reduziu o desempenho para 57,05%, com um aumento da ordem de cinco vezes o tempo requerido por *fold*. Para todas as variações associadas ao enriquecimento, a GCN apresentou uma maior dispersão inter-*fold* (desvios superiores a 4%). Neste cenário de variabilidade, notou-se que o *fold* 8 impôs o maior desafio de classificação para a ampla maioria dos métodos, embora o conjunto de estratégias de refinamento topológico tenha demonstrado capacidade de mitigar essa dificuldade específica.

4.3. Comparação com modelos de referência

Para a avaliação do custo-efetividade da adoção de estratégias de enriquecimento de assinaturas espectrais baseadas em grafos propostas neste trabalho, foram construídos dois modelos de referência, nos quais as assinaturas espectrais foram submetidas aos classificadores *Isolation Forest* e k-NN diretamente após o pré-processamento, portanto, sem enriquecimento, aqui referidos como modelos de referência IF e k-NN. A calibração destes modelos sobre o espectro pré-processado resultou em uma vizinhança de $k = 10$ para o k-NN e em uma taxa de contaminação de $c = 0,01$ para o *Isolation Forest*.

O modelo de referência IF apresentou uma AUROC de 57,33 ± 1,01%, em contraposição à AUROC de 68,52 ± 1,17% obtida pelo modelo proposto: *SVG + SDW* associada aos descritores *Graph Structural*. Desta forma, a adoção do processo de enriquecimento proposto permitiu um ganho de mais de 11 pontos percentuais, indicando que a transição para o domínio de grafos estruturou os atributos de forma muito mais favorável ao isolamento das novidades em comparação aos espectros pré-processados. Por sua vez,

esse resultado foi inferior ao do modelo de referência k-NN, que atingiu uma AUROC de 77,68 $\pm$ 1,07% e um tempo por *fold* significativamente menor (14 segundos).

A superioridade do modelo k-NN de referência (77,68%) frente às representações em grafos levanta uma discussão metodológica fundamental sobre a natureza das características acústicas extraídas. O espectro pré-processado retém a densidade de energia absoluta global do sinal, o que favorece substancialmente o cálculo de distâncias euclidianas adotado pelo k-NN. Em contrapartida, a conversão para Grafos de Visibilidade Espectral (SVG) foca na correlação harmônica e na topologia dos picos, atenuando a informação de energia global. Embora essa transição de domínio não supere a estratégia de k-NN para este conjunto de dados, ela comprova a sua eficácia ao otimizar algoritmos que falham no espectro pré-processado: o enriquecimento topológico permitiu à *Isolation Forest* um salto de mais de 11 pontos percentuais, de 57,33% para 68,52%. Isso demonstra que as abordagens em grafos são efetivas na extração de atributos relacionais complementares à energia, sendo particularmente vantajosas para algoritmos de isolamento em espaços latentes complexos.

5. Conclusões

Este trabalho investigou assinaturas espectrais enriquecidas por meio de grafos para a detecção de novidades em sonar passivo (OSR). Os resultados demonstraram que a extração de características estruturais deve ser avaliada em conjunto com a técnica de classificação de novidades adotada. Em particular, o Grafo de Visibilidade Espectral (SVG), associado à penalização espacial (SDW) e ao classificador *Isolation Forest*, atingiu 68,52% de AUROC, um ganho superior a 11 pontos percentuais frente ao uso do espectro pré-processado nesse mesmo modelo. Embora os resultados não tenham superado o classificador k-NN operando diretamente sobre o espectro pré-processado (77,68%), o estudo comprova que a transformação para o domínio de grafos extrai propriedades relacionais que melhoram fortemente a detecção por algoritmos de isolamento espacial, oferecendo uma via promissora e complementar às métricas convencionais de distância euclidiana puras. Por sua vez, constatou-se que refinamentos adicionais aos atributos avaliados neste trabalho não apresentaram ganhos significativos de desempenho e resultaram em um aumento expressivo do custo computacional das simulações.

Como trabalhos futuros, espera-se avaliar mecanismos alternativos para a definição dos atributos dos nós, a geração dos grafos, poda de arestas e enriquecimento de atributos, bem como investigar o uso de representações compactas das assinaturas derivadas por aprendizado por contraste.

6. Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Referências

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7(1):1–30.

Duda, R. O., Hart, P. E., and Stork, D. G. (2000). *Pattern classification*. John Wiley & Sons, 2 edition.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874. ROC Analysis in Pattern Recognition.

Han, J., Pei, J., and Tong, H. (2022). *Data Mining: Concepts and Techniques*. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann.

Honorato, E. S., de Oliveira e Souza Filho, J. B., and Da Silva Muniz, V. H. (2022). Clustering-and-bagging-based ensemble for novelty detection in passive sonar systems. In *2022 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, pages 1–6.

Honorato, E. S., Muniz, V. H. d. S., and Souza Filho, J. B. d. O. e. (2020). Detecção hierárquica de classes desconhecidas em sonar por "autoencoders"convolucionais. In *Anais do XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)*. Sociedade Brasileira de Telecomunicações.

Kipf, T. N. and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *CoRR*, abs/1609.02907.

Lacasa, L., Luque, B., Ballesteros, F., Luque, J., and Nuño, J. C. (2008). From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences*, 105(13):4972–4975.

Liu, F. T., Ting, K. M., and Zhou, Z.-H. (2008). Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining*, pages 413–422.

Ma, X., Wu, J., Xue, S., Yang, J., Zhou, C., Sheng, Q. Z., Xiong, H., and Akoglu, L. (2023). A comprehensive survey on graph anomaly detection with deep learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(12):12012–12038.

Montgomery, D. and Runger, G. (2010). *Applied Statistics and Probability for Engineers*. John Wiley & Sons.

Moura Junior, N. N. d. (2018). *Detecção de novidade para sistemas de sonar passivo*. Tese de doutorado, Universidade Federal do Rio de Janeiro (UFRJ/COPPE), Rio de Janeiro, Brasil.

Muniz, V. H. S., Oliveira, M. B., and SOUZA FILHO, J. B. O. (2019). Classificação hierárquica de navios em sinais de sonar passivo. In *XIV Congresso Brasileiro de Inteligência Computacional*, Belém, Brasil.

Nemenyi, P. B. (1963). *Distribution-free multiple comparisons*. PhD thesis, Princeton University.

Xu, K., Hu, W., Leskovec, J., and Jegelka, S. (2018). How powerful are graph neural networks? *CoRR*, abs/1810.00826.

Yela, D. F., Stowell, D., and Sandler, M. (2019). Spectral visibility graphs: Application to similarity of harmonic signals. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5.

Yichen, D., Xiaohong, S., Haiyan, W., and Yongsheng, Y. (2025). An open-set recognition method for ship radiated noise signal based on graph convolutional neural network prototype learning. *Digital Signal Processing*, 156:104748.