

Classificação Automática de Cenas Acústicas para Apoio à Hipersensibilidade Auditiva no TEA: Uma Abordagem Simulada

Neto, José Ricci^{*1}; Macegoza, Thiago Vinícius¹; Parreira, Wemerson Delcio²; Itturriet, Fábio Pires^{1,3}; Borges, Renata Coelho^{1,3}

1 – Universidade Tecnológica Federal do Paraná (UTFPR), Campus Curitiba, Departamento Acadêmico de Eletrotécnica (DAELT), jneto.2020@alunos.utfpr.edu.br; thiago.2000.macegoza@gmail.com

2 – Pontifícia Universidade Católica de Campinas, PUC-Campinas, wemerson.delcio@puc-campinas.edu.br

3 – Programa de Pós-Graduação em Engenharia Biomédica (PPGEB), UTFPR - Campus Curitiba, renatacoelho@utfpr.edu.br; fabioitturriet@utfpr.edu.br

* -Correspondência: Av. Sete de Setembro, 3165, Rebouças, Curitiba, Paraná, Brasil, 80230-901.

DOI 10.52050/9788579177163-20

RESUMO

A hipersensibilidade auditiva em pessoas com Transtorno do Espectro Autista (TEA) pode prejudicar a autonomia em ambientes ruidosos, afetando o bem-estar sensorial. Este estudo visa simular um sistema assistivo de reconhecimento automático de cenas acústicas para subsidiar respostas adaptativas baseadas no perfil sensorial do usuário. Utilizou-se Máquinas de Vetores de Suporte (SVM), testando diferentes funções *kernel* em base de dados pública. O modelo com *kernel*-RBF obteve alta revocação, F1-score e acurácia superior a 85%. Os achados destacam o potencial de tecnologias assistivas embarcadas para promover a inclusão e melhorar o bem-estar sensorial de pessoas neurodivergentes.

Palavras-chave: Transtorno do Espectro Autista (TEA), Máquinas de Vetores de Suporte (SVM), Tecnologias assistivas.

ABSTRACT

Auditory hypersensitivity in individuals with Autism Spectrum Disorder (ASD) can impair autonomy in noisy environments, affecting sensory well-being. This study aims to simulate an assistive system for automatic recognition of acoustic scenes to support adaptive responses based on the user's sensory profile. Support Vector Machines (SVM) were used to test different kernel functions in a public database. The model with kernel-RBF obtained high recall, F1-score and accuracy above 85%.

The findings highlight the potential of embedded assistive technologies to promote inclusion and improve the sensory well-being of neurodivergent individuals.

Keywords: Autism Spectrum Disorder (ASD), Support Vector Machines (SVM), Assistive Technologies.

1. INTRODUÇÃO

O Transtorno do Espectro Autista (TEA) é uma condição do neurodesenvolvimento que afeta aproximadamente 1 em cada 100 crianças globalmente, segundo estimativas da Organização Mundial da Saúde (OMS) (ZEIDAN et al., 2022). Caracteriza-se por dificuldades na comunicação social, padrões comportamentais restritos e alterações sensoriais, incluindo disfunções no processamento auditivo. Dentre essas alterações, a hipersensibilidade a estímulos sonoros é particularmente relevante, afetando até 70% dos indivíduos com TEA, de acordo com estudos clínicos recentes (WILLIAMS; SUZMAN; WOYNAROSKI, 2021). Essa hipersensibilidade pode se manifestar como desconforto intenso, ansiedade ou reações de sobrecarga sensorial frente a sons de alta intensidade, sons repetitivos ou ambientes imprevisivelmente ruidosos.

Ambientes urbanos modernos, como escolas, centros comerciais e transportes públicos, frequentemente apresentam paisagens sonoras dinâmicas e não estruturadas. A superposição de sons, variações rápidas de intensidade e a presença de ruídos impulsivos podem representar um desafio significativo para indivíduos com TEA, impactando sua qualidade de vida e limitando sua participação social. A exposição contínua a ambientes acusticamente hostis pode desencadear respostas de luta ou fuga, comportamentos de evitação e até crises sensoriais, dificultando a inclusão e a autonomia (POTGIETER et al., 2020). No Brasil, a Lei nº 13.146/2015 (Estatuto da Pessoa com Deficiência) estabelece o direito à acessibilidade, à autonomia e à participação plena na sociedade para pessoas com deficiência, incluindo indivíduos com transtornos do espectro autista, reforçando a importância do desenvolvimento de tecnologias assistivas que minimizem barreiras sensoriais.

Sistemas de Classificação de Cenas Acústicas (*Acoustic Scene Classification* – ASC) têm se consolidado como ferramentas eficientes para a categorização automática de ambientes sonoros, com aplicações em áreas como vigilância, dispositivos móveis e tecnologias assistivas (MACEGOZA et al., 2025). Em linhas gerais, a ASC envolve a extração de características relevantes do sinal de áudio e a aplicação de algoritmos de aprendizado de máquina para associar essas características a rótulos de cena específicos. Entre os métodos de classificação, as Máquinas de Vetores de Suporte (*Support Vector Machines* – SVM) destacam-se pela robustez em problemas de alta dimensionalidade e pela capacidade de

generalização mesmo com conjuntos de dados relativamente pequenos. O uso de SVMs em ASC tem sido explorado com sucesso em diversas aplicações, embora os desafios permaneçam no que diz respeito à variabilidade acústica e à robustez em cenários reais (DING et al., 2024).

Estudos recentes têm investigado a aplicação de ASC em contextos assistivos, incluindo a adaptação de ambientes inteligentes para pessoas com deficiências sensoriais. Entretanto, o uso específico de ASC para apoio a indivíduos com TEA ainda é limitado, com poucos trabalhos explorando classificadores leves, de baixa complexidade computacional, que possam ser futuramente implementados em sistemas embarcados ou dispositivos portáteis (MULIMANI; VENKATESH; KOOLAGUDI, 2025). A combinação de classificadores tradicionais como SVM com técnicas de extração de características acústicas eficientes surge, portanto, como uma alternativa promissora para aplicações em tempo real voltadas à mitigação de sobrecarga sensorial em indivíduos neurodivergentes.

Neste contexto, este estudo visa simular um sistema assistivo de reconhecimento automático de cenas acústicas, com o objetivo de subsidiar respostas adaptativas baseadas no perfil sensorial do usuário. A pesquisa utiliza técnicas de Classificação de Cenas Acústicas (ASC) baseadas em Máquinas de Vetores de Suporte (SVM) para identificar ambientes acústicos potencialmente desconfortáveis para indivíduos com TEA. Ao classificar diferentes ambientes sonoros, o sistema pode filtrar ou modificar sons específicos, como ruídos excessivos ou frequências perturbadoras, a fim de reduzir o desconforto sensorial. A abordagem inclui a extração de características acústicas simples, como energia e taxa de cruzamento por zero, utilizando um conjunto de dados representativo de diversas cenas. Além disso, o desempenho dos classificadores SVM, tradicionais e variantes, é avaliado com base em métricas como acurácia, precisão, sensibilidade e F1-score.

A proposta visa contribuir para o desenvolvimento futuro de sistemas embarcados assistivos personalizáveis capazes de operar em tempo real, com foco na melhoria da qualidade de vida de pessoas com hipersensibilidade auditiva.

1.1. Transtorno do Espectro Autista e Alterações no Processamento Sensorial Auditivo

Uma característica frequentemente relatada em indivíduos com TEA é a atipia no processamento sensorial, especialmente no domínio auditivo. Essas alterações podem se manifestar como hipersensibilidade a sons específicos ou, inversamente, hipossensibilidade a estímulos que normalmente despertariam atenção em indivíduos neurotípicos (RODRÍGUEZ-ARMENDARIZ;

VELA-ROMERO; GALIANA, 2024). A hipersensibilidade auditiva, em particular, pode levar a reações intensas a sons cotidianos, resultando em comportamentos de esquiva, crises de estresse ou sobrecarga sensorial.

O processamento sensorial atípico no TEA não se limita apenas à intensidade dos estímulos auditivos. Estudos indicam que existe também uma diferença na forma como as informações auditivas são integradas e interpretadas no contexto multimodal, isto é, em conjunto com estímulos visuais, táteis e proprioceptivos. Essa integração deficiente pode amplificar a sensação de imprevisibilidade e desconforto em ambientes ruidosos, dificultando a adaptação a contextos sociais dinâmicos. Ambientes considerados banais para a maioria das pessoas, como centros comerciais, escolas e eventos públicos, podem se tornar fontes de extremo desconforto para indivíduos com TEA, afetando sua qualidade de vida e participação social (THÉRIAULT; MORALES, 2022). Além disso, é importante destacar que o perfil sensorial em indivíduos com TEA não é homogêneo. Essa variabilidade torna desafiadora a criação de estratégias de suporte universais e reforça a necessidade de abordagens tecnológicas adaptativas que possam personalizar a detecção e a resposta a ambientes potencialmente desconfortáveis.

1.2. Classificação de Cenas Acústicas (ASC): Conceitos e Aplicações

A tarefa de ASC envolve a captura de sinais acústicos, a extração de características representativas desses sinais e a posterior classificação dos dados em categorias pré-definidas, como “rua movimentada”, “parque”, “metrô”, “cafeteria”, entre outros. As aplicações práticas de ASC são vastas e incluem desde a melhoria de sistemas de assistência auditiva, passando por contextos de segurança pública, até soluções de suporte para populações vulneráveis, como indivíduos com TEA.

Um aspecto crítico da ASC é a seleção adequada de características acústicas que representem de forma eficiente os padrões sonoros dos ambientes. Características como coeficientes mel-cepstrais (MFCCs), espectrogramas e parâmetros baseados na energia espectral são comumente utilizados (SURENDIRAN et al., 2024). Entretanto, pesquisas recentes exploram também características mais especializadas, como cromagrama, taxa de cruzamento por zero (*zero-crossing rate*) e contraste espectral, com o objetivo de capturar nuances mais sutis dos ambientes (DING et al., 2024). Após a extração, técnicas de redução de dimensionalidade, como Análise de Componentes Principais (PCA) ou t-SNE, são frequentemente aplicadas para otimizar o desempenho dos classificadores.

O campo de ASC também vem expandindo suas aplicações para sistemas embarcados e *context-aware computing*. Em robótica assistiva, por exemplo, ASC é utilizado para ajudar robôs sociais a interpretar o ambiente e ajustar seu comportamento (FURLETOV; WILLERT; ADAMY, 2021). No caso de tecnologias assistivas para pessoas com TEA, a capacidade de detecção automática de ambientes permite a implementação de sistemas de alerta, adaptação sensorial (como redução automática de volume de aparelhos auditivos) ou mesmo a sugestão de estratégias de enfrentamento, promovendo maior autonomia e segurança para o usuário.

1.3. Máquinas de Vetores de Suporte (SVM) na ASC

Máquinas de Vetores de Suporte são algoritmos amplamente utilizados em problemas de classificação e regressão, conhecidos por sua robustez em contextos de alta dimensionalidade e pela capacidade de encontrar hiperplanos ótimos de separação entre classes. Em ASC, SVMs têm se destacado pela capacidade de lidar bem com conjuntos de dados relativamente pequenos, onde a generalização é mais crítica do que a capacidade de aprender representações extremamente complexas, como em redes neurais profundas (DENG, TIAN e ZHANG, 2012).

Estratégias como *One-vs-One* e *One-vs-Rest* possibilitam a expansão da SVM, originalmente projetada para problemas binários, para cenários com múltiplas classes de ambientes. Adicionalmente, SVMs probabilísticas, que estimam a probabilidade de um dado pertencer a cada classe, têm sido desenvolvidas, aumentando a interpretabilidade dos resultados, especialmente em sistemas assistivos onde decisões baseadas em níveis de confiança são desejáveis.

Em termos de eficiência computacional, algoritmos como o *Sequential Minimal Optimization* (SMO) foram desenvolvidos para acelerar o treinamento de SVMs em conjuntos grandes de dados. Outras iniciativas têm focado na adaptação de SVMs para implementação eficiente em dispositivos de baixo consumo, como microcontroladores ARM Cortex-M e FPGAs, viabilizando aplicações embarcadas para ASC em dispositivos *wearable*. Essas estratégias são essenciais para sistemas assistivos que precisam operar continuamente e de forma energeticamente eficiente, características fundamentais para o suporte em tempo real a indivíduos com TEA.

1.4. Trabalhos Relacionados

Diversos estudos recentes têm buscado soluções para mitigar os efeitos da hipersensibilidade auditiva em indivíduos com Transtorno do Espectro Autista

(TEA). Um desses estudos desenvolveu um sistema de cancelamento ativo de ruído (ANC) adaptado, baseado em avaliações de percepção auditiva e testes de eletroencefalografia, especificamente voltado para crianças com TEA e hiper-reatividade auditiva. Os resultados mostraram que determinadas frequências, como 250 Hz e 8 kHz, foram consistentemente classificadas como mais desconfortáveis, levando ao desenvolvimento de fones de ouvido com funções de controle de ruído individualizado, que melhoraram significativamente o conforto auditivo dos participantes (KWONG; MUNG; CHOY, 2024).

Outro estudo utilizou modelos de redes neurais profundas para prever as características de percepção sonora e sugerir modificações auditivas destinadas a reduzir o estresse em indivíduos com transtornos do neurodesenvolvimento, incluindo TEA. Os resultados demonstraram que tais modelos podem auxiliar na construção de sistemas personalizados que adaptam estímulos sonoros de acordo com as reações subjetivas dos usuários, proporcionando reduções significativas nos níveis de desconforto (ICHIKAWA et al., 2025).

A literatura sobre tecnologias de controle de ruído também traz contribuições relevantes para a área de assistência a pessoas com TEA, mesmo quando o foco inicial não é diretamente em transtornos do neurodesenvolvimento. Laska, Goubran e Wallace (2024) avaliaram a qualidade do som residual e a preservação das pistas binaurais em fones de ouvido ANC, aspectos fundamentais para a percepção espacial auditiva. Considerando que indivíduos com TEA frequentemente apresentam dificuldades na integração sensorial auditiva e na localização de sons, a preservação adequada de pistas espaciais torna-se crítica para evitar desorientação ou aumento de estresse em ambientes ruidosos.

Outro avanço importante é a proposta de uma topologia de circuito ANC de latência ultra baixa com características ajustáveis (KWONG; CHOY; CHAN; MUNG, 2024). A capacidade de adaptar dinamicamente o cancelamento de ruído a diferentes ambientes é especialmente relevante para o TEA, dada a variabilidade na sensibilidade auditiva entre indivíduos e a necessidade de soluções altamente personalizadas para conforto e segurança auditiva.

Recentemente, um novo estudo reforçou a importância do controle individualizado da percepção sonora em crianças com TEA (KWONG et al., 2025). Através da combinação de testes de percepção auditiva e respostas eletrofisiológicas, foi desenvolvido um *headset* com sistema híbrido de controle ativo de ruído e função de ajuste de percepção. Os resultados mostraram redução significativa na magnitude do potencial evocado auditivo e melhoria na experiência sonora em ambientes comuns, como o tráfego intenso ou durante tempestades, indicando potencial para a diminuição de comportamentos relacionados à hiper-reatividade auditiva.

Esses trabalhos demonstram que a combinação de técnicas de classificação de cenas acústicas (ASC), tecnologias de cancelamento ativo de ruído e personalização sensorial pode representar um caminho promissor para o desenvolvimento de tecnologias assistivas destinadas a melhorar a qualidade de vida de pessoas com TEA.

2. MATERIAIS E MÉTODOS

2.1. Base de dados

Para o desenvolvimento e avaliação do sistema proposto, foi utilizada a base de dados TUT Urban Acoustic Scenes 2018 (HEITTOLA; MESAROS; VIRTANEN, 2018), composta por gravações reais de ambientes urbanos. Foram selecionadas 10 classes distintas de ambientes: *airport* (classe 1), *bus* (classe 2), *metro* (classe 3), *metro station* (classe 4), *park* (classe 5), *public square* (classe 6), *shopping mall* (classe 7), *street pedestrian* (classe 8), *street traffic* (classe 9) e *tram* (classe 10). Como a base já é pré-processada para tarefas de classificação, técnicas adicionais de remoção de ruído não foram aplicadas nesta etapa. A escolha desta base se justifica pela relevância dos sons urbanos no contexto de hipersensibilidade auditiva em indivíduos com TEA, considerando que ambientes externos frequentemente apresentam sons imprevisíveis e potencialmente estressantes.

2.2. Extração de características

Foram extraídas 11 características acústicas de cada arquivo de áudio, selecionadas por sua relevância na caracterização de sinais sonoros ambientais. As *features* computadas incluíram: energia, ZCR (*zero crossing rate*), duração, coeficientes mel-cepstrais (MFCC - média dos coeficientes), valor RMS (*root mean square*), centróide espectral, largura de banda espectral (*spectral bandwidth*), *roll-off* espectral, além das estatísticas de média e desvio padrão da ZCR, e *chroma mean*. Estas características oferecem uma representação multidimensional do conteúdo espectro-temporal de cada gravação, capturando tanto aspectos de intensidade e periodicidade quanto propriedades espectrais relacionadas à percepção auditiva.

2.4. Classificação

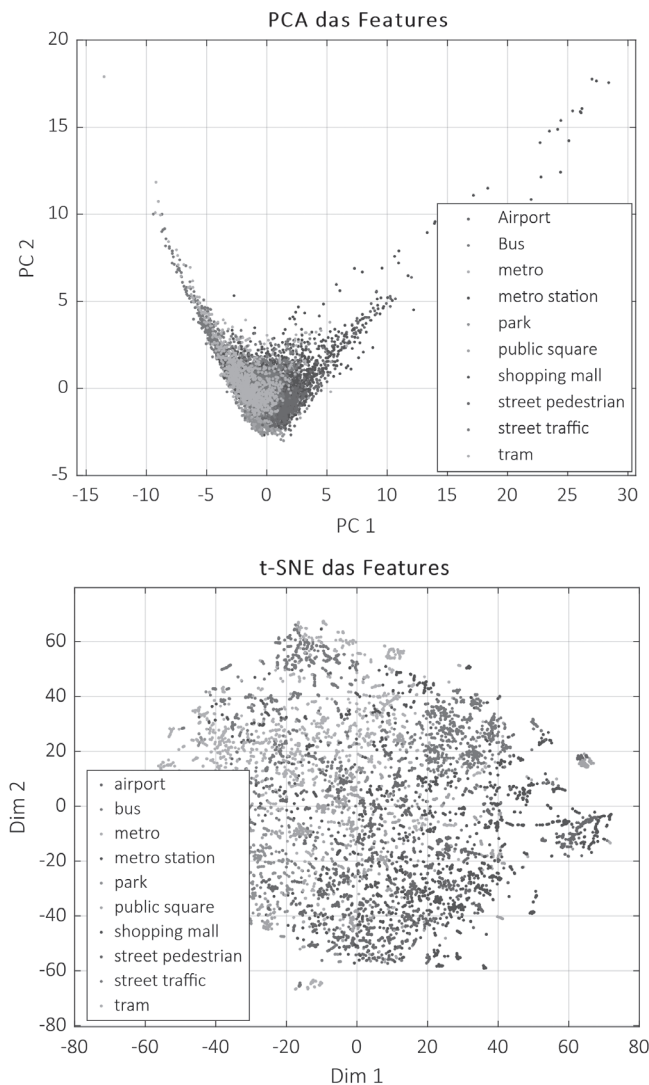
A classificação dos sons foi realizada utilizando SVM, explorando diferentes funções de kernel para investigar a melhor configuração para os dados extraídos. Foram testados os kernels linear (separação linear entre as classes), polinomial (decisão baseada em polinômios de graus superiores) e *radial basis function* (RBF

- modelagem não linear eficiente para dados com distribuição mais complexa). A validação do desempenho dos modelos foi conduzida por meio de validação cruzada estratificada com 5 *folds*, assegurando que todas as classes estivessem proporcionalmente representadas em cada partição de treinamento e teste. Esse procedimento não apenas preserva o balanceamento das classes ao longo das divisões, mas também permite uma avaliação mais robusta e generalizável do modelo ao dividir os dados em múltiplos subconjuntos de treinamento e teste. A validação cruzada reduz a dependência de uma única divisão aleatória dos dados, e com isso mitiga o risco de viés ou variância decorrentes de amostragens particulares, assim proporcionando uma estimativa mais confiável do desempenho do classificador.

A avaliação do desempenho dos modelos considerou a acurácia média e o desvio padrão obtidos nos *folds* de validação. Além disso, foram geradas matrizes de confusão para cada kernel, permitindo a análise da distribuição dos acertos e erros de classificação entre as classes. A partir das matrizes de confusão, também foram calculadas métricas adicionais, incluindo precisão, revocação e F1-score, possibilitando uma avaliação mais detalhada da capacidade de generalização dos classificadores e da robustez dos resultados frente a possíveis desbalanceamentos de classes.

Além das etapas quantitativas de avaliação dos classificadores, foram utilizadas técnicas de redução de dimensionalidade para inspecionar visualmente a separabilidade entre as classes no espaço das características extraídas, conforme apresentado na Figura 1. Os métodos PCA (*Principal Component Analysis*) e t-SNE (*t-distributed Stochastic Neighbor Embedding*) foram aplicados às amostras da base para projeção bidimensional dos dados. O gráfico gerado por PCA, Figura 1a, evidenciou uma distribuição relativamente contínua entre as classes, com alguma sobreposição, sugerindo que as variáveis extraídas contêm variação discriminativa, mas não necessariamente linearmente separável. Por sua vez, a visualização t-SNE, Figura 2b, revelou agrupamentos mais nítidos e estruturas locais mais bem definidas, indicando que a complexidade não linear dos dados pode ser capturada por modelos como o SVM com kernel RBF. Essas visualizações apoiam empiricamente a escolha de kernels não lineares para a tarefa de classificação.

Figura 01 - Projeções bidimensionais das amostras no espaço de características.



(a) PCA: Principal Component Analysis (b) t-SNE: *t-distributed Stochastic Neighbor Embedding*

3. RESULTADOS

Os modelos SVM avaliados com diferentes funções kernel apresentaram desempenhos significativamente distintos na tarefa de classificação de cenas acústicas da base TUT-8, conforme apresentado na Tabela 1.

Tabela 01 - Desempenho obtido nos diferentes modelos de SVM

Kernel	Acurácia (%)	Desvio Padrão	Melhor Classe (F1)	Pior Classe (F1)
Linear	64,39	±1,36	<i>park</i> (88,50)	<i>public square</i> (49,30)
Polinomial	82,44	±1,24	<i>bus</i> (88,90)	<i>metro station</i> (70,81)
RBF	85,89	±0,98	<i>park</i> (94,15)	<i>metro station</i> (73,33)

O modelo com kernel RBF obteve o melhor desempenho global, com acurácia média de 85,89% ($\pm 0,92\%$), revocação média de 85,74% ($\pm 1,09\%$) e F1-score médio de 85,69% ($\pm 0,98\%$). Esse modelo mostrou-se particularmente robusto na classificação de cenas como *bus*, *park*, *shopping mall* e *street traffic*, cujas taxas de revocação e precisão ultrapassaram 90%. A consistência dos resultados também se reflete nos baixos desvios padrão, sugerindo estabilidade entre os *folds*. A matriz de confusão associada a esse classificador, Figura 2b, mostrou uma clara superioridade na separação entre as classes, com acertos próximos a 90% em ambientes como *street traffic*, *bus* e *shopping mall*, com destaque para a classe *park* que obteve 93,2%. Ainda assim, algumas confusões pontuais persistiram, principalmente entre *metro station* (17,6%), *street pedestrian* (18,5%) e *public square* (17,9%), evidenciando que ambientes semiabertos e dinâmicos continuam desafiadores mesmo para modelos altamente não lineares. Também foi observado um pequeno grau de confusão entre *airport* e *shopping mall*, o que pode ser atribuído à presença de sons de fundo semelhantes, como ruídos de movimento, anúncios e vozes humanas.

Os valores da matriz confusão podem diferenciar-se entre falsos positivos (FP) — o sistema identifica incorretamente algo como positivo, quando na verdade é negativo — ou falsos negativos (FN) — o sistema identifica incorretamente algo como negativo, quando na verdade é positivo. Para o contexto de usuários com hipersensibilidade auditiva, deve ser levado em conta em grande parte os falsos positivos, pois, neste contexto, esses erros são especialmente críticos, podendo levar a ajustes inadequados. Isso pode envolver, por exemplo, a ativação de filtros redutores de ruído desnecessariamente, resultando em uma supressão abrupta de sons que pode causar desorientação sensorial ou ansiedade, agravando o desconforto.

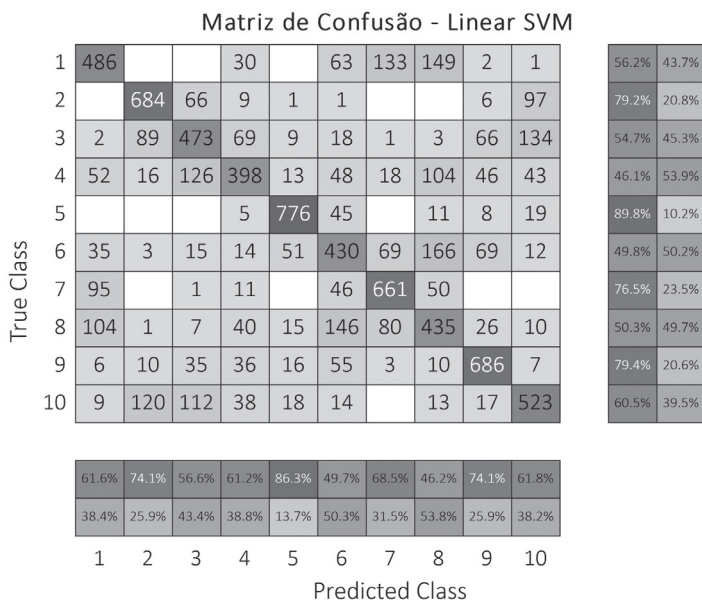
Em contraste, o classificador com kernel linear apresentou o pior desempenho, com acurácia média de 64,39% ($\pm 1,36\%$), revocação média de 64,48% ($\pm 1,44\%$) e F1-score médio de 64,18% ($\pm 1,45\%$). A matriz de confusão desse modelo, Figura 2a, revelou altos índices de confusão entre ambientes

sonoramente semelhantes, como *metro* e *metro station*, ou *street pedestrian* e *public square*. Esses resultados indicam a limitação do kernel linear para capturar a não linearidade presente na distribuição dos dados no espaço de características.

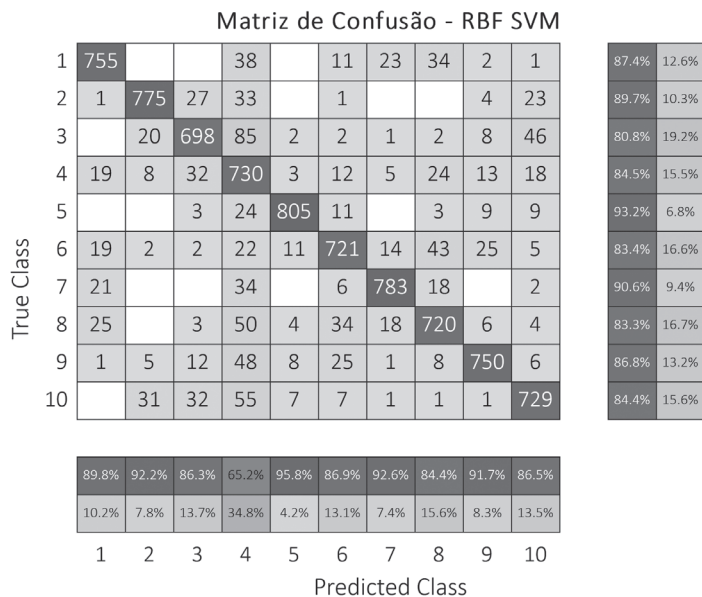
O modelo SVM com kernel polinomial (grau 3) apresentou desempenho intermediário, com acurácia média de 71,72% ($\pm 1,73\%$), revocação média de 71,72% ($\pm 1,60\%$) e F1-score de 71,64% ($\pm 1,66\%$). Embora tenha superado o kernel linear em quase todas as classes, sua complexidade não foi suficiente para alcançar a performance do modelo RBF. A matriz de confusão desse modelo, Figura 2c, evidenciou redução das confusões nas classes mais estruturadas, como *airport*, *street pedestrian* e *public square*, que passaram a ser corretamente identificadas com maior frequência, obtendo uma melhora de 27.1, 25.6 e 29.1 pontos percentuais em relação ao modelo linear. No entanto, ainda que as classes *public square* e *street pedestrian* obtiveram uma melhora expressiva em relação ao modelo linear, ainda foram observadas trocas frequentes entre elas, o que sugere que, embora o kernel polinomial seja mais expressivo, ele ainda apresenta limitações para modelar relações mais sutis entre classes com maior sobreposição acústica. Ainda vale destacar que, quando voltamos a observar os falsos positivos, o modelo polinomial teve o melhor desempenho na classe *metro station*, ainda que no geral tenha ficado abaixo do RBF, o que pode indicar uma especificidade acústica inerente a este tipo de ambiente que é melhor reconhecida pelo modelo polinomial.

A análise comparativa entre os três modelos mostra uma clara progressão no desempenho à medida que a complexidade do kernel aumenta, como esperado. O modelo linear falha em capturar a estrutura do espaço de características, enquanto o polinomial oferece ganhos, mas ainda enfrenta dificuldades em fronteiras ambíguas. Por outro lado, o kernel RBF se mostrou mais eficaz em generalizar a classificação, mesmo frente à variabilidade intraclasse e interclasse presente nos dados.

Figura 02 - Matrizes de confusão para três funções kernel distintas
 – linear, polinomial (grau 3) e *radial basis function* (RBF).



(a) SVM Linear

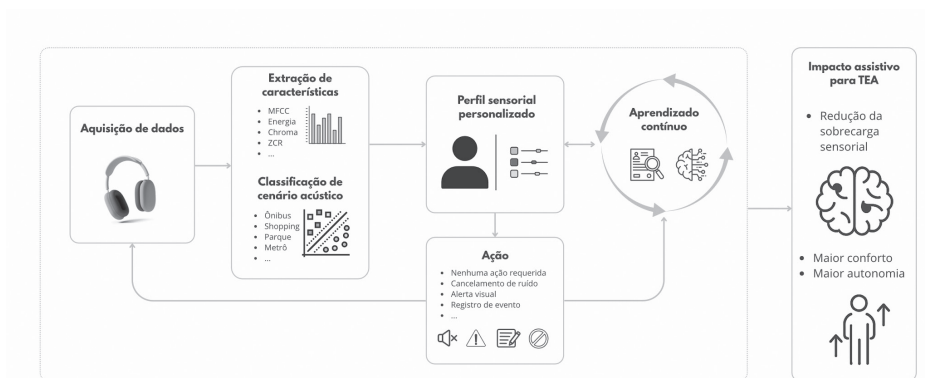


(b) SVM RBF

4.1. Aplicação e Potencial Assistivo

A aplicação prática do sistema de classificação de cenas acústicas (ASC) proposto neste trabalho visa apoiar indivíduos com Transtorno do Espectro Autista (TEA) na gestão de ambientes auditivos potencialmente estressantes. Sua integração em dispositivos vestíveis, como *hearables*, ou em aplicativos móveis, representa uma solução viável para o monitoramento contínuo do ambiente sonoro e a identificação, em tempo real, de contextos acústicos de alto risco sensorial. A proposta é que o sistema atue como uma camada sensorial inteligente, capaz de antecipar situações de sobrecarga auditiva e fornecer respostas adaptativas que minimizem o impacto desses estímulos, conforme apresentado na Figura 3.

Figura 03 - Arquitetura conceitual de um sistema assistivo adaptativo baseado em classificação de cenas acústicas (ASC) para suporte à hipersensibilidade auditiva no Transtorno do Espectro Autista (TEA).



Em uma possível implementação embarcada, sensores de áudio integrados a fones inteligentes capturariam o som ambiente, aplicariam o modelo de ASC desenvolvido neste estudo, e, mediante a detecção de cenas classificadas como de alto risco sensorial para aquele perfil, desencadeariam ações automáticas. Para que essa resposta seja eficaz, é fundamental que o sistema opere de forma personalizada, respeitando a variabilidade individual da percepção auditiva entre pessoas com TEA. Em vez de adotar reações padronizadas a todos os ambientes urbanos, o sistema pode aprender com o histórico de interações do usuário, identificando padrões de desconforto associados a determinadas cenas. Por exemplo, se um indivíduo demonstrar reações negativas frequentes em *public square*, mas não em *shopping mall*, o sistema pode ajustar automaticamente parâmetros como a sensibilidade de detecção, o nível de filtragem sonora ou o tipo de intervenção adotada. Essa personalização pode ser implementada a partir

de perfis sensoriais configuráveis, como ajustes manuais, *feedback* do usuário ou aprendizado incremental. Com isso, o modelo de ASC deixa de ser apenas uma ferramenta de reconhecimento ambiental e se torna a base de um sistema assistivo centrado no indivíduo, com potencial para realizar ações específicas como filtragem de frequências incômodas, ativação de cancelamento de ruído, emissão de alertas discretos entre outros.

Além do suporte direto ao usuário, a integração do sistema com plataformas móveis permitiria o registro de exposições sonoras e intervenções adotadas, possibilitando a criação de relatórios individualizados ao longo do tempo. Esses dados poderiam ser compartilhados com cuidadores, terapeutas e profissionais de saúde, apoiando o planejamento de intervenções mais eficazes e fundamentadas em evidências do cotidiano do usuário.

Entretanto, algumas limitações importantes devem ser consideradas. A principal delas diz respeito à ausência de bases de dados acústicos específicas para indivíduos com TEA, o que limita a representatividade dos modelos de classificação frente às particularidades da percepção sensorial dessa população. Além disso, fatores contextuais como o estado emocional, o nível de familiaridade com o ambiente e a presença de estímulos simultâneos não sonoros podem influenciar significativamente a resposta auditiva, dificultando a generalização das decisões do sistema.

5. CONCLUSÕES

O presente estudo demonstrou a viabilidade da utilização de máquinas de vetor de suporte (SVM) na classificação de cenas acústicas urbanas a partir de um conjunto de características acústicas extraídas da base TUT-8. Foram testadas diferentes configurações de kernel, com destaque para o kernel RBF, que apresentou desempenho superior em termos de acurácia, revocação e F1-score. A abordagem multiclasse adotada permitiu a identificação de 10 tipos distintos de ambientes urbanos, com potencial aplicação em dispositivos assistivos voltados à mitigação de sobrecarga sensorial auditiva em indivíduos com TEA.

A principal contribuição desta pesquisa está no direcionamento do ASC para uma perspectiva sensorial voltada ao TEA, explorando uma aplicação assistiva que visa reduzir episódios de desconforto auditivo e promover maior qualidade de vida. Ao considerar as peculiaridades da sensibilidade auditiva no espectro autista, o sistema proposto representa um avanço no desenvolvimento de tecnologias assistivas personalizadas.

Para futuras etapas, destaca-se a necessidade de coleta de dados auditivos de usuários reais de diferentes faixas etárias e níveis de severidade

do TEA, permitindo o treinamento de modelos mais robustos e adaptativos. A implementação dos algoritmos em plataformas embarcadas, como *hearables*, será fundamental para validar sua aplicabilidade prática. Estudos clínicos também serão indispensáveis para avaliar a eficácia do sistema em ambientes reais e ajustar suas funcionalidades de acordo com as necessidades específicas dos usuários.

REFERÊNCIAS

DENG, N.; TIAN, Y.; ZHANG, C.. *Support vector machines: optimization based theory, algorithms, and extensions*. CRC Press, 2012

DING, B.; et al. Acoustic scene classification: a comprehensive survey. *Expert Systems with Applications*, v. 238, p. 121902, 2024.

FURLETOV, Y.; WILLERT, V.; ADAMY, J. Auditory scene understanding for autonomous driving. In: *IEEE INTELLIGENT VEHICLES SYMPOSIUM (IV)*, 2021,

HEITTOLA, T.; MESAROS, A.; VIRTANEN, T. TUT Urban Acoustic Scenes 2018, development dataset [Data set]. *Zenodo*, 2018. Disponível em: <https://zenodo.org/record/1234567>. Acesso em: 05 fev 2025.

ICHIKAWA, Itsuki; et al. Machine learning model for reproducing subjective sensations and alleviating sound-induced stress in individuals with developmental disorders. *Frontiers in Psychiatry*, v. 16, p. 1412019, 2025.

KWONG, T. C.; CHOY, Y. S.; CHAN, C. C.; MUNG, S. W. Tunable active noise control circuit topology for multiple-feature applications. *Scientific Reports*, v. 14, n. 1, p. 18629, 11 ago. 2024.

KWONG, T. C.; et al. Intervention technology of aural perception controllable headset for children with autism spectrum disorder. *Scientific Reports*, v. 15, n. 1, p. 5356, 2025.

KWONG, T. C.; MUNG, S. W.; CHOY, Y. S. Development of hybrid active noise cancellation system for the children with autism spectrum disorder. In: *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, v. 270, n. 7, p. 4168-4178, 2024. Institute of Noise Control Engineering.

LASKA, B.; GOUBRAN, R.; WALLACE, B. Binaural characterization of active noise cancelling headphones. In: *2024 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 2024.

MACEGOZA, T. V.; RICCI NETO, J.; COMPANHONI, L. M. A.; PARREIRA, W. D.; ITTURRIET, F. P.; BORGES, R. C. Classificação de cenários acústicos em dispositivos auditivos: uma revisão sistemática da literatura sobre desafios, avanços e perspectivas. In: *XVI COMPUTER ON THE BEACH*, 2025, Itajaí. *Anais do XVI Computer on the Beach*, 2025.

MULIMANI, M.; VENKATESH, S.; KOOLAGUDI, S. G. Acoustic event and scene classification: a review. *SN Computer Science*, v. 6, p. 54, 2025.

POTGIETER, I.; FACKRELL, K.; KENNEDY, V.; CRUNKHORN, R.; HOARE, D. J. Hyperacusis in children: a scoping review. *BMC Pediatrics*, v. 20, n. 1, p. 319, 29 jun. 2020.

RODRÍGUEZ-ARMENDARIZ, E.; VELA-ROMERO, M.; GALIANA, A. Sensory processing challenges in children with neurodevelopmental disorders and genetic conditions: an observational study. *NeuroSci*, v. 5, n. 3, p. 339-353, 2024.

SURENDIRAN, J.; et al. A systemic review on automatic acoustic scene classification. In: *INTERNATIONAL CONFERENCE ON POWER, ENERGY, CONTROL AND TRANSMISSION SYSTEMS (ICPECTS)*, 2024, [local]. IEEE, 2024.

THÉRIAULT, W.; MORALES, E. Accessibility and social participation in urban settings for people with autism spectrum disorder (ASD) or an intellectual disability (ID). *Journal of Accessibility and Design for All*, v. 12, n. 1, p. 155-179, 2022.

ZEIDAN, J.; FOMBONNE, E.; SCORAH, J.; IBRAHIM, A.; DURKIN, M. S.; SAXENA, S.; YUSUF, A.; SHIH, A.; ELSABBAGH, M. Global prevalence of autism: a systematic review update. *Autism Research*, v. 15, n. 5, p. 778-790, maio 2022.

WILLIAMS, Z. J.; SUZMAN, E.; WOYNAROSKI, T. G. Prevalence of decreased sound tolerance (hyperacusis) in individuals with autism spectrum disorder: a meta-analysis. *Ear and Hearing*, v. 42, n. 5, p. 1137-1150, set./out. 2021.