REVISTA ELETRÔNICA
ACTA SAPIENTIA

DETECÇÃO DA VIOLÊNCIA DE GÊNERO E PREVENÇÃO DE FEMINICÍDIOS ATRAVÉS DA ANÁLISE DE ÁUDIOS: ESTUDO COM REDES NEURAIS CONVOLUCIONAIS (CNNs)

Rafael da Costa Brandão[1]
Universidade Federal do Ceará (UFC)
https://orcid.org/0009-0001-5762-9649

Alberto Sampaio Lima[2]
Universidade Federal do Ceará (UFC)
https://orcid.org/0000-0003-0696-5148

Wagner Bandeira Andriola[3]
Universidade Federal do Ceará (UFC)
https://orcid.org/0000-0001-6459-0992

RESUMO

O aumento da violência contra mulheres frequentemente culminando em feminicídio demanda estratégias de intervenções tecnológicas urgentes, especialmente no Brasil. Este artigo propõe o uso de Redes Neurais Convolucionais (CNNs) como estratégia para tratar informações complexas, com o fito de classificar as emoções registradas em áudio e transcritas em texto, constituindo uma poderosa ferramenta tecnológica para identificar e responder aos cenários de risco iminente, contribuindo, assim, para a prevenção de possíveis casos de feminicídio. O método CNNs fornecerá parâmetros cruciais para a tomada de decisão das autoridades, facilitando a expedição de notificações de alerta em tempo hábil para evitar possíveis casos de feminicídio. Os testes e as análises quantitativas feitas com a ferramenta confirmam a eficácia da abordagem proposta, prometendo melhorar e tornar mais eficientes as respostas das autoridades aos possíveis casos de violência contra mulheres.
Palavras-chave: Políticas Públicas; Feminicídio; Redes Neurais Convolucionais.

DETECCIÓN DE LA VIOLENCIA DE GÉNERO Y PREVENCIÓN DE FEMINICIDIOS MEDIANTE EL ANÁLISIS DE AUDIOS: INVESTIGACIÓN CON REDES NEURONALES CONVOLUCIONALES

RESUMEN

El aumento de la violencia contra las mujeres, que a menudo culmina en feminicidio, exige estrategias de intervención tecnológica urgentes, especialmente en Brasil. Este artículo propone el uso de Redes Neuronales Convolucionales como estrategia para procesar información compleja, con el fin de clasificar emociones grabadas en audio y transcritas a texto, constituyendo una poderosa herramienta tecnológica para identificar y responder a escenarios de riesgo inminente, contribuyendo así a la prevención de posibles casos de feminicidio. El método de las Redes Neurales Convolucioanles proporcionará parámetros cruciales para la toma de decisiones por parte de las autoridades, facilitando la emisión oportuna de alertas para prevenir posibles casos de feminicidio. Las pruebas y los análisis cuantitativos realizados con la herramienta confirman la eficacia del enfoque propuesto, prometiendo mejorar y hacer más eficientes las respuestas de las autoridades ante posibles casos de violencia contra las mujeres.
Palabras clave: Políticas Públicas; Femicidio; Redes Neuronales Convolucionales.

[1] Engenheiro de Telecomunicações pela Universidade Federal do Ceará (UFC), Fortaleza, Ceará, Brasil. E-mail: rafaelcosta@alu.ufc.br
[2] Professor Associado da Universidade Federal do Ceará (UFC), Doutor em Engenharia de Teleinformática (UFC) com estágio de Pós-Doutorado na Universidade Federal de Campina Grande (UFPb), Paraíba, Brasil. E-mail: albertosampaio@ufc.br
[3] Professor Titular da Universidade Federal do Ceará (UFC), Doutor em Filosofia e Ciências da Educação, *Universidad Complutense de Madrid* (UCM) com Estágio de Pós-Doutorado na *Universidad de Salamanca* (USAL), Salamanca, Espanha. E-mail: w_andriola@ufc.br

REVISTA ELETRÔNICA
ACTA SAPIENTIA

DETECTION OF GENDER-BASED VIOLENCE AND PREVENTION OF FEMINICIDES VIA AUDIO ANALYSIS: STUDY WITH CONVOLUTIONAL NEURAL NETWORKS (CNNs)

ABSTRACT

The rise in violence against women, frequently culminating in feminicide, demands urgent technological intervention strategies, especially in Brazil. This article proposes the use of Convolutional Neural Networks (CNNs) as a strategy to process complex information, aiming to classify emotions recorded in audio and transcribed into text. This constitutes a powerful technological tool for identifying and responding to imminent risk scenarios, thus contributing to the prevention of potential feminicide cases. The CNNs method will provide crucial parameters for decision-making by authorities, facilitating the timely issuance of alert notifications to prevent possible feminicide cases. Tests and quantitative analyses conducted with the tool confirm the effectiveness of the proposed approach, promising to improve and make more efficient the authorities' responses to potential cases of violence against women.
Keywords: Public Policies; Feminicide; Convolutional Neural Networks.

1. INTRODUÇÃO

A violência de gênero contra a mulher constitui um problema social e de saúde pública no Brasil, manifestando-se em múltiplas formas — física, sexual, psicológica, patrimonial e moral — e culminando, em sua expressão extrema, no feminicídio, entendido como o assassinato de mulheres por razões relacionadas ao gênero (Araújo; Ramos; Andriola, 2026). Esta violência é produto de arranjos históricos, culturais e estruturais que fortalecem a desigualdade de gênero, interseccionada com fatores socioeconômicos, raciais e regionais, exigindo abordagens intersetoriais para a prevenção, a detecção e a responsabilização (Araújo; Andriola; Ramos, 2025abc; Dias *et al.*, 2025). No Brasil, os feminicídios permanecem em patamares elevados, embora oscilem conforme a região e o período analisado, pois, segundo o Atlas da Violência e o Anuário Brasileiro de Segurança Pública, em média, 10 mulheres são assassinadas diariamente (IPEA, 2021; FBSP, 2024).

Além disso, destaca-se a subnotificação e a dificuldade de classificação correta desses crimes, o que compromete a acurácia das estatísticas, bem como a formulação de políticas públicas (IPEA, 2021). Os casos que envolvem violência contra as mulheres, especialmente o feminicídio, ocorrem predominantemente no ambiente familiar, frequentemente perpetrados por parceiros íntimos (Ribeiro; Pereira, 2022). Estes casos são contabilizados, na maioria das vezes, apenas após a consumação da agressão, criando lacunas significativas na proteção das vítimas e limitando ações preventivas de órgãos responsáveis (Rodrigues, 2016). Durante a pandemia de COVID-19, observou-se

aumento expressivo nos casos de violência contra mulheres, evidenciando a necessidade urgente de intervenções tecnológicas para enfrentar esta prática (FBSP, 2023).

Embora o arcabouço legal brasileiro represente um avanço importante, o enfrentamento ao feminicídio e à violência de gênero ainda enfrenta desafios significativos, demandando esforços contínuos de pesquisa, formulação de políticas e articulação institucional (Dias; Araújo, 2024; Araújo; Andriola, 2024). Nesse cenário, o avanço de tecnologias de aprendizado de máquina, como as Redes Neurais Convolucionais (*CNNs*), permite tratar informações complexas do contexto social e utilizá-las na tomada de decisões estratégicas (Deshpande, 2016; Cho *et al*., 2014). Este trabalho propõe um modelo baseado em *CNNs* para classificar as emoções a partir de áudio, aliado à transcrição de texto, como ferramenta tecnológica para identificar e responder aos cenários de risco iminente, contribuindo para a prevenção de possíveis casos de feminicídio.

2. FUNDAMENTAÇÃO TEÓRICA E TRABALHOS RELACIONADOS

O termo "feminicídio" se refere à consumação de atos violentos contra mulheres, destacando a desigualdade de poder entre os gêneros e as violências que as mulheres enfrentam. Rodrigues (2016) destaca que o extermínio de mulheres tem raízes profundas na desigualdade de gênero, historicamente tolerada pela sociedade. Não obstante, nos últimos anos tem ocorrido maior consciência social acerca do tema, com a visibilidade de políticas públicas como a Lei Maria da Penha e legislações específicas sobre feminicídio (Ribeiro; Pereira, 2022).

No campo jurídico, dois marcos normativos se sobressaem, a saber: (i) Lei nº 11.340/2006 (Lei Maria da Penha) estabeleceu mecanismos de prevenção e enfrentamento à violência doméstica e familiar contra a mulher, criando medidas protetivas e ampliando a responsabilização do agressor (Brasil, 2006); (ii) Lei nº 13.104/2015 incluiu o feminicídio no rol dos crimes hediondos, qualificando o homicídio cometido contra mulheres em razão de sua condição (Brasil, 2015). Apesar desses avanços, pesquisas apontam lacunas na implementação efetiva, seja pela insuficiência de serviços especializados, seja pela fragmentação da rede de proteção (Pasinato, 2019; Debert; Oliveira, 2020; Araújo *et al.*, 2024).

Entre os principais fatores de risco para a ocorrência do feminicídio destacam-se a violência prévia no caso de relacionamentos íntimos, a posse de armas de fogo pelos agressores, o uso de substâncias psicoativas e a perpetuação de normas patriarcais que

reforçam a subordinação feminina (Meneghel; Portella, 2017). A literatura enfatiza ainda as desigualdades interseccionais: mulheres negras e indígenas apresentam maior vulnerabilidade à violência letal, o que reflete a interação entre racismo estrutural e desigualdade de gênero (IPEA, 2021; Waiselfisz, 2015; Araújo; Menezes; Moreira, 2024; Dias; Araújo, 2024).

No âmbito institucional, verificam-se avanços com a criação de juizados especializados, delegacias de atendimento à mulher e linhas de denúncia, como o "Ligue 180". Contudo, persistem limitações quanto à cobertura territorial, à capacitação de profissionais e à articulação intersetorial entre justiça, segurança pública, assistência social e saúde (FBSP, 2024; Pasinato, 2019). Ademais, a qualidade dos dados oficiais ainda se constitui como obstáculo, pois há inconsistências entre registros policiais e de saúde, dificultando análises longitudinais e comparativas (IPEA, 2021).

Existe, portanto, a necessidade de estratégias integradas de prevenção e enfrentamento, combinando medidas de proteção imediata às vítimas com políticas estruturais de redução das desigualdades de gênero. Ademais, o fortalecimento da produção de dados confiáveis, a ampliação de serviços de apoio e a implementação de ações educativas que promovam a igualdade de gênero e a desconstrução de estereótipos constituem ideais sociais utópicos (Meneghel; Portella, 2017; Debert; Oliveira, 2020).

O aumento alarmante nos casos de feminicídio acentua a urgência da proposição e implementação de soluções para proteger a vida das mulheres. Em resumo, o feminicídio evidencia necessidades prementes de discussões e de operacionalização de medidas eficazes visando preservar a vida das vítimas. A Tabela 1 apresenta alguns dos principais dados recentes sobre a temática, constantes no Relatório Anual Socioeconômico da Mulher (RASEAM) do Ministério das Mulheres (2025).

Tabela 1. Informações acerca da violência contra as mulheres.

Indicador	Valor / Ano	Observações
Feminicídios registrados	1.450 em 2024	12 a mais que em 2023
Homicídios dolosos + lesão seguida de morte (mulheres)	2.485 em 2024	Indicador de alta violência letal
Estupros de mulheres	71.892 em 2024 (~196/dia)	Pequena queda (-1,44%) vs 2023
Violência contra mulheres pretas/pardas	60,4% das vítimas (20-59 anos)	Possível sobreposição de gênero e raça
Local de ocorrência / gênero do agressor	71,6% no domicílio; 76,6% homens	Predominância do ambiente doméstico e agressores masculinos

Fonte: Relatório Anual Socioeconômico da Mulher (RASEAM, 2025).

Almeida e Nonato (2021) analisaram dados de violência contra a mulher em Minas Gerais e Goiás, identificando tendências espaciais e temporais relevantes. Ribeiro

e Pereira (2022) propuseram o sistema SALVE TODAS, integrando sensores e geolocalização para alertar autoridades sobre situações de risco. Mantini *et al.* (2021) apresentaram abordagem baseada em redes neurais para identificar sinais de violência em postagens de redes sociais (*Twitter*). Na sequência, apresenta-se revisão de literatura sobre processamento de sinais digitais e o uso de redes neurais.

2.1. Processamento de Sinais Digitais

O processamento de sinais digitais é fundamental para a análise de áudio em sistemas de aprendizado de máquina, permitindo extrair características como amplitude e frequência para a classificação de emoções (Livingstone; Russo, 2018). Espectrogramas oferecem uma representação visual dessas características, auxiliando o treinamento de CNNs (Xin *et al*., 2018). Na análise de emoções, os espectrogramas ajudam a identificar padrões que diferenciam medo, tristeza e nervosismo, emoções relevantes para identificar ações de violência contra as mulheres (Mantini *et al*., 2021).

Os espectrogramas são representações gráficas das características extraídas de bases de dados, que empregam paleta de cores para indicar a intensidade de energia de amostras destes dados, permitindo inferir e distinguir visualmente determinados comportamentos associados ao conjunto em questão. No entanto, nem sempre a visualização traz respostas imediatas para os seres humanos, pois há casos em que a representação de dois espectrogramas torna difícil distinguir a natureza de cada um dos gráficos, conforme exemplificado na Figura 1.

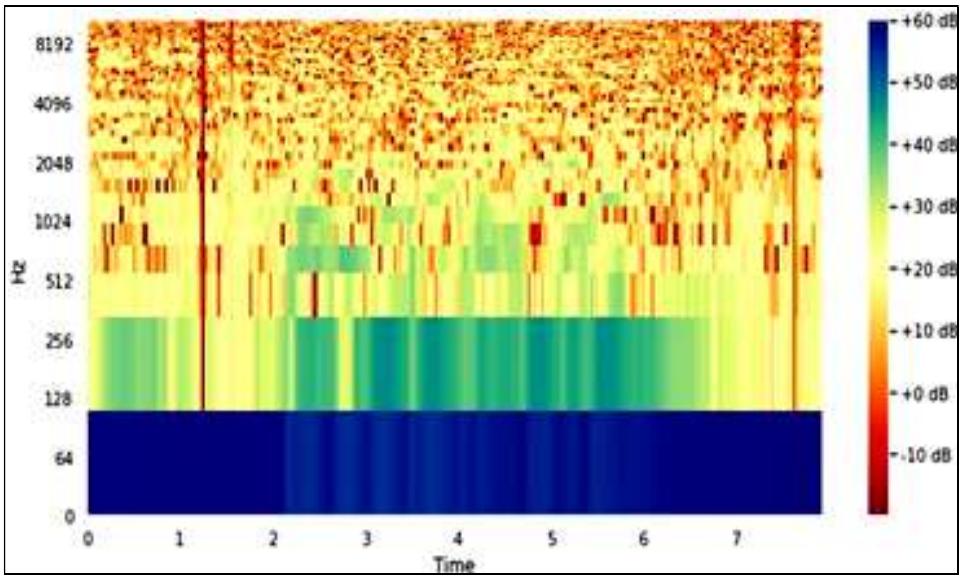

Figura 1. Espectrograma de um conjunto de 50 características extraído da base RAVDESS.

O aprendizado da rede ocorre a partir do conjunto de características extraídas de um processo, através do reconhecimento de padrões, conforme refletido no espectrograma.

2.2. Redes Neurais

Redes neurais artificiais são modelos computacionais inspirados na comunicação entre neurônios humanos, conforme ilustradas na Figura 2.



Figura 2. Regiões de seguimento de estudo e suas classes.

Nas redes neurais artificiais cada neurônio possui camadas de entrada ocultas, função de ativação e pesos, cuja otimização é possível através de algoritmos, como o Adam (Kingma; Ba, 2017). As CNNs incluem camadas de convolução, *pooling*, totalmente conectadas, otimizadas para classificação de dados complexos (Xin *et al.,* 2018). Embora as redes neurais sejam inspiradas no processo de comunicação entre os neurônios do cérebro humano, as características envolvidas na criação de um modelo de rede neural diferem das funções biológicas presentes na comunicação entre os neurônios humanos. Isso ocorre devido à natureza de sua aplicação e à necessidade de adaptação, uma vez que as redes neurais artificiais operam em um ambiente computacional para realizar os processos de aprendizagem correspondentes (Zhang, 2018).

Uma rede neural consiste em neurônios interligados responsáveis por processar informações, conforme a ilustração na Figura 3.

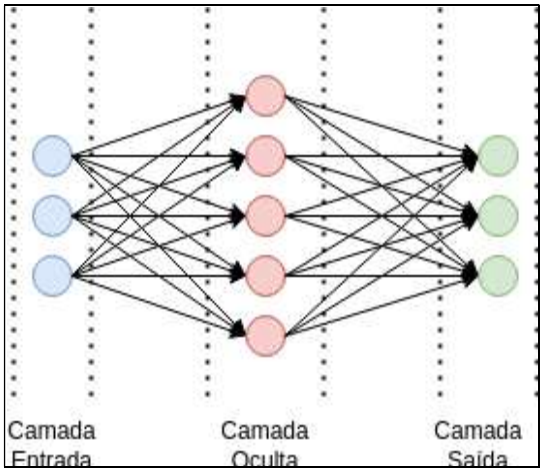

Figura 3. Rede neural e suas camadas (neurônios).

Cada neurônio artificial possui camadas de entrada, função de ativação e pesos entre camadas. Algoritmos de otimização de descida de gradiente (Ruder, 2016) são comumente empregados para otimizar redes neurais, visando minimizar o erro e encontrar os mínimos globais nos pesos da rede. Essa abordagem resulta em um aprendizado mais eficiente do modelo para problemas de classificação ou regressão, com a redução de erros e a otimização do modelo, conforme ilustrado na Figura 4.

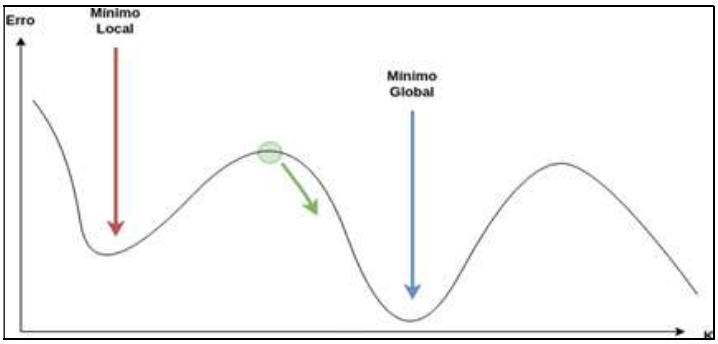

Figura 4. Minimização do erro e otimização do peso.

O otimizador Adam (Kingman; Ba, 2014; Zhang, 2018; Duchi *et. al.*, 2011) é amplamente utilizado em problemas que envolvem grandes conjuntos de dados e/ou otimização de redes neurais. É eficiente computacionalmente, requer pouca memória e é invariante ao redimensionamento diagonal de gradientes. Em comparação com a descida de gradiente, o Adam ajusta os parâmetros das redes neurais, estimando momentos de primeira e segunda ordem.

Ao se desenvolver um modelo de rede neural há de se garantir não apenas a solução final do problema, mas a qualidade do modelo e a eficácia do treinamento. A avaliação do modelo pode ser realizada por meio de uma matriz de confusão, que compara os dados previstos com as classificações reais, cujos parâmetros usados

incluem: Verdadeiro Positivo (VP): quando o valor predito corresponde à classe positiva real; Falso Negativo (FN): quando o valor predito é negativo, mas deveria ser positivo; Falso Positivo (FP): quando o valor predito é positivo, mas deveria ser negativo; Verdadeiro Negativo (VN): quando o valor predito e o real são negativos. Tais parâmetros formam uma matriz 2x2 usada para calcular a acurácia do modelo através de taxas de acerto das classes previstas em comparação com as reais, conforme a Figura 5.

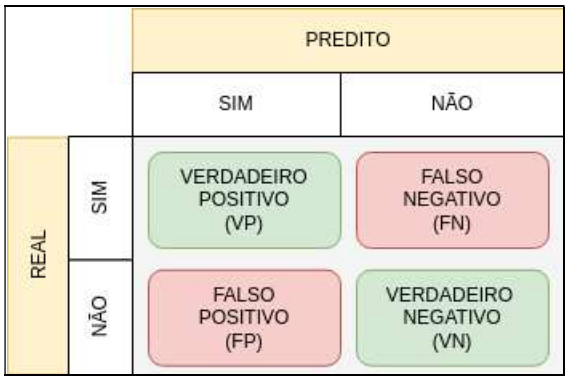

Figura 5. Esquema de uma matriz de confusão.

Uma *Rede Neural Convolucional (Convolutional Neural Network-CNN)* consiste em um *algoritmo* de *Aprendizado Profundo (Deep Learning)* com um pré-processamento muito menor em comparação com outros algoritmos de classificação (Xin *et. al.*, 2018). As CNNs são compostas por diferentes camadas com uma função específica, que incluem:

- **Camada de convolução**: fase principal que dá às CNNs o seu nome. Ela consiste em um conjunto de filtros ou kernels que são convoluídos com a entrada da imagem para produzir mapas de características. Esses filtros extraem padrões.
- **Camada de *pooling***: etapa posterior à convolução, na qual é comum se utilizar uma camada de *pooling* para reduzir a dimensionalidade dos mapas de características, diminuindo o número de parâmetros e melhorando a eficiência. Exemplos comuns de operações de *pooling são Max pooling* e *average pooling*.
- **Camada de ativação**: após cada camada convolucional e de *pooling*, é comum aplicar uma função de ativação, como a *ReLU (Rectified Linear Unit)*, que introduz não linearidade na rede e ajuda a modelar relações mais complexas nos dados.
- **Camada totalmente conectada**: após as camadas convolucionais e de *pooling*, os mapas de características são frequentemente achatados em um

vetor unidimensional e passados através de uma ou mais camadas totalmente conectadas (ou densas) para realizar a classificação final.

- **Camada de saída**: a camada final da rede, que geralmente usa uma função de ativação específica dependendo do tipo de problema a ser resolvido. Por exemplo, para problemas de classificação, é comum usar *softmax* como função de ativação na camada de saída.

O treinamento de uma CNN geralmente envolve a propagação para frente dos dados através da rede, visando fazer previsões, e em seguida, a retropropagação do erro para ajustar os pesos da rede usando um algoritmo de otimização, como gradiente descendente estocástico (SGD) ou algoritmos mais avançados, como Adam (Zhang, 2018).

3. ASPECTOS METODOLÓGICOS

De acordo com Hevner *et al.* (2004), *Design Science Research (DSR)* constitui uma abordagem de pesquisa direcionada ao desenvolvimento e à análise de artefatos inovadores que visam resolver problemas de natureza organizacional ou tecnológica relevantes, ao mesmo tempo em que promovem o avanço do conhecimento científico. A DSR prioriza a concepção de artefatos com objetivos bem definidos — como modelos, métodos, *frameworks* ou sistemas — que são avaliados tanto pela sua aplicabilidade prática quanto pela contribuição teórica que oferecem.

A natureza da presente pesquisa foi exploratória, com análise quantitativa baseada em experimentos. Na presente pesquisa foi utilizada a abordagem DSR. As etapas da pesquisa seguiram as recomendações de Peffers *et al.* (2007) tendo o processo sido estruturado em seis etapas: (1) identificação do problema e motivação, (2) definição dos objetivos, (3) design e desenvolvimento do modelo, (4) demonstração, (5) avaliação e (6) comunicação dos resultados. Utilizou-se *Deep Learning* com CNN para classificação de emoções e palavras-chave derivadas de transcrições de áudio, visando suporte à tomada de decisão. A base RAVDESS foi normalizada usando *Coeficiente Cepstral de Frequência Mel* (*MFCC - Mel Frequency Cepstral Coefficient*), e a CNN foi implementada em linguagem *Python* com uso das bibliotecas *Keras* e *TensorFlow*, incluindo duas camadas convolucionais e duas densas.

A análise exploratória visou verificar a qualidade da base de dados RAVDESS utilizada na criação do modelo da rede neural, visando utilizar *Deep Learning* no

treinamento de *rede neural convolucional* (Jeronymo, 2019), para classificação de emoções e sua relação com palavras-chave obtidas da transcrição de texto. Após a validação da base de dados, foi aplicada a *transformada de Fourier* para a extração e visualização de parâmetros dos respectivos áudios contidos na base de dados em análise, exclusivamente na visualização desse conjunto no domínio do tempo e na frequência.

Posteriormente, procedeu-se à normalização dos dados por meio da aplicação do *MFCC*. A partir da extração das características do conjunto de dados, foram inseridos os respectivos valores para o desenvolvimento do modelo de rede neural convolucional desejado. Posteriormente, um arquivo de áudio foi submetido à classificação de emoções, relacionando-o com as palavras-chave extraídas da transcrição do áudio. Dessa forma, buscaram-se os elementos necessários para a tomada de decisão relacionada às métricas.

A arquitetura da rede incluiu duas camadas convolucionais com 64 e 128 filtros, respectivamente, seguidas pela função de ativação *ReLU*. Além disso, a rede contou com duas camadas densas de 64 neurônios cada e uma camada de saída com oito neurônios, correspondentes às classes de saída. O treinamento foi realizado com o otimizador Adam (Zhang, 2018), passando pelas fases de entrada de dados, cálculo de perdas e ajuste de pesos.

4. O MODELO PROPOSTO

A base de dados utilizada para treinar a *rede neural convolucional* é conhecida como RAVDESS (*Ryerson Audio-Visual Database of Emotional Speech and Song*) (Livingstone e Russo, 2018). Esta base consiste em 1440 amostras de áudio, sendo cada amostra um arquivo de áudio único. Os áudios foram gravados por um grupo de atores profissionais, composto por 12 homens e 12 mulheres, que expressaram oito distintas categorias de emoções por meio da entonação ao pronunciar determinadas frases, a saber: calma, tristeza, nervosismo, medo, felicidade, surpresa, nojo e neutralidade.

As sete emoções referidas obtiveram o mesmo número de amostras, totalizando 192 gravadas na base RAVDESS. Apenas a emoção identificada como "neutro" contou com 96 amostras dentro do conjunto de dados. As palavras nos áudios foram pronunciadas com intensidade normal ou forte, sendo que 768 amostras de áudio tiveram intensidade de pronúncia normal e 672 tiveram intensidades fortes.

Inicialmente, foi realizado o desenvolvimento de uma cinco etapas conforme descrito no diagrama de fluxo (Figura 6) para o processamento dos sinais de áudio, os quais compreendem vários fatores que podem impactar significativamente na qualidade do modelo de rede neural convolucional treinado, como a presença de ruído nas amostras de áudio captadas. A base de dados em questão apresenta arquivos de áudio com excelente qualidade e pouca ou nenhuma presença perceptível de ruído para o ouvido humano.



Figura 6. Diagrama de fluxo.

O pré-processamento de dados é fundamental para obter resultados diversos, conforme mostrado na Figura 6. Este processo envolve: o encaminhamento do dado de entrada para extração de informações; descrição e verificação de relacionamentos entre os dados para interpretação direta e clara das métricas; identificação e minimização de ruído para aumentar a legibilidade dos dados; normalização dos dados em escalas compreensíveis; retorno do áudio tratado com qualidade superior.

A CNN automatizou a compreensão das características dos dados de entrada e permitiu a extração e mapeamento das características do conjunto de dados RAVDESS, contribuindo para a sumarização dos dados. O conjunto de dados utilizado possui oito emoções, algumas das quais possuem características semelhantes. O treinamento da CNN para classificação foi realizado com 80% dos dados para treinamento e 20% para teste. A linguagem de programação utilizada foi Python, executada em um ambiente com processador Intel Core i7-6500U, 8 GB de RAM e sistema operacional Ubuntu 22.04.1 LTS.

A otimização da CNN foi realizada com o otimizador Adam, que adapta as estimativas durante a descida do gradiente. Após o treinamento, amostras foram testadas e verificou-se que as predições do modelo foram consistentes com os rótulos originais. Para avaliação do modelo, foram utilizadas métricas como F1-Score e acurácia, detalhadas para cada emoção. A transcrição de áudio foi feita com a *Google Web Speech*

API, devido à facilidade de manipulação e resultados coesos. Essa transcrição é utilizada juntamente com a saída da CNN para tomada de decisão.

5. EXPERIMENTO E ANÁLISE DE RESULTADOS

Visando a validação do modelo proposto e implementado nesta pesquisa, foi definido um experimento da *CNN,* com o objetivo de classificar as emoções, em conjunto com a transcrição de áudio em texto, com o fito de estabelecer um mecanismo para suporte à tomada de decisão. Os resultados obtidos foram analisados e são apresentados a seguir.

5.1. Treinamento

Inicialmente, torna-se necessário plotar os dados ao longo da duração (em segundos), visando compreender o quão diferentes são as amostras de sinal de áudio. Dessa forma, torna-se perceptível que cada emoção apresenta componentes distintos ao longo do tempo, como evidenciado na Figura 7.

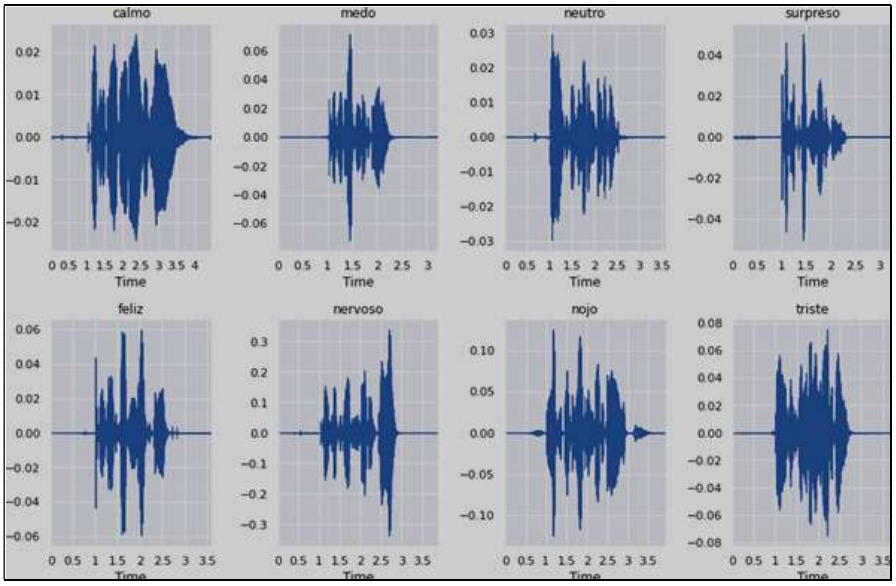

Figura 7. Sinal de áudio correspondente a cada emoção.

A *rede neural convolucional* (CNN) usa as características extraídas do conjunto de dados para gerar os respectivos espectrogramas correspondentes a cada uma das emoções. Foram extraídas 50 características do sinal de áudio de cada amostra da base de dados, e posteriormente gerado o gráfico correspondente, conforme a Figura 8.

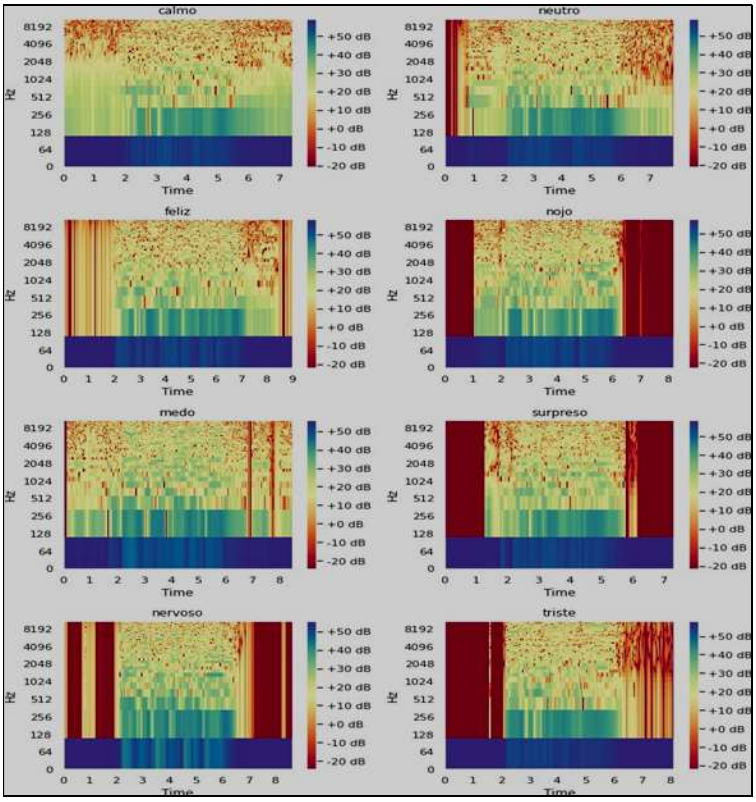

Figura 8. Espectrograma correspondente a cada emoção.

As informações dos espectrogramas contêm dados usados no treinamento do modelo de rede neural convolucional. Após a extração de características das amostras, procedeu-se ao desenvolvimento da *CNN* para lidar com a classificação das emoções. O treinamento consistiu no uso de 1152 amostras, cada uma contendo 50 características extraídas, cujo modelo resultante de 35 épocas obteve acurácia de 90,1% no treinamento e 71,18% no teste, conforme a Figura 9.

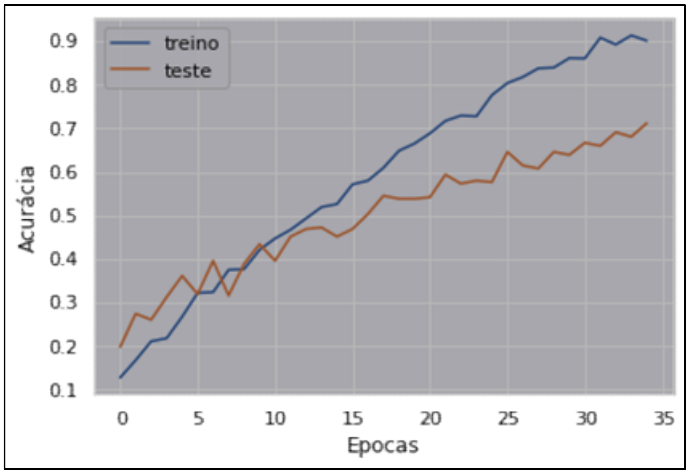

Figura 9. Acurácia do Modelo.

5.2. Análise da Matriz de Confusão

A avaliação do desempenho da *CNN* para classificação pode ser obtida por meio da análise da matriz de confusão, que oferece uma representação clara e concisa do número de previsões corretas e incorretas dos dados. Como mostrado na Figura 10, algumas emoções foram erroneamente classificadas como pertencentes a outras classes emocionais. Por exemplo, a emoção "triste" teve parte de suas amostras classificadas como "calmo", "feliz", "medo" e "neutro", respectivamente.

As emoções de interesse para a identificação de violência contra mulheres são especificamente a tristeza, o nervosismo e o medo. Quando combinadas com a transcrição de texto, ambas as métricas fornecem suporte para definir parâmetros para a tomada de decisão. Deve-se realizar uma análise minuciosa do desempenho do modelo, em relação à precisão com que o modelo classifica as emoções de interesse. O uso de dados diversificados possibilita evitar qualquer viés no modelo e garantir que ele possa lidar com diferentes classes de emoções. O modelo demonstra uma capacidade maior na classificação de um conjunto com 8 classes distintas de emoções (Figura 10).

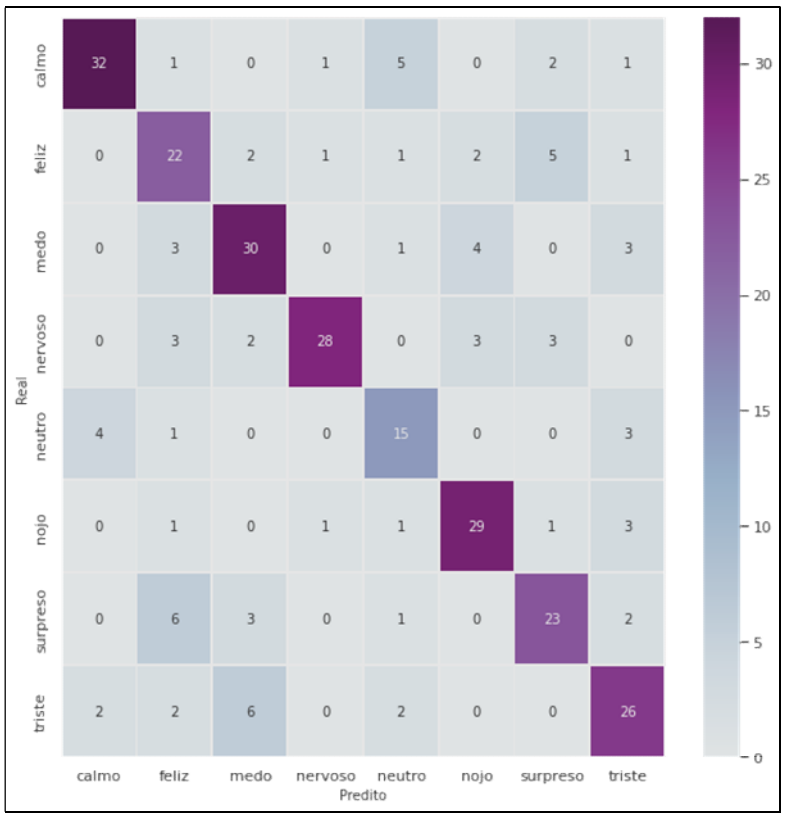

Figura 10. Confiabilidade do Modelo.

5.3. Apresentação e Discussão dos Resultados

Conforme a Tabela 2, pode-se verificar algumas métricas pertinentes ao modelo proposto, cujos resultados indicam que o modelo apresenta boa precisão para classificar

as emoções “calma”, “medo”, “nervosismo” e “nojo”. Em contrapartida, as emoções “felicidade”, “neutralidade” e “tristeza” estão com baixa precisão baixa.

Tabela 2. Métricas do Modelo.

Emoções	Precisão	F1-Score	Amostras
Calma	0,84	0,80	42
Felicidade	0,56	0,60	34
Medo	0,70	0,71	41
Nervosismo	0,90	0,80	39
Neutralidade	0,58	0,61	23
Nojo	0,76	0,78	36
Surpresa	0,68	0,67	35
Triste	0,57	0,68	38
Acurácia (média)	0,71		

Fonte: os Autores.

Pode-se inferir a partir das informações apresentadas na Tabela 2, que os dados e suas respectivas relações indicam uma boa precisão. Após submeter várias amostras de áudio à API *Google Web Speech*, foi possível obter transcrições precisas do áudio para texto. O resultado foi satisfatório, uma vez que o texto extraído reflete com fidelidade o conteúdo original. Dessa forma, os resultados indicaram que é viável extrair palavras-chave por meio da transcrição do áudio para texto. Em um contexto de violência contra a mulher, palavras como "socorro", "me ajuda pelo amor de Deus", "alguém me ajuda", entre outras, podem ser reconhecidas, contribuindo para a extração de palavras-chave.

5.7. Contribuições e Aplicabilidade

Em termos de contribuições científicas, tem-se que a presente pesquisa incorpora conhecimento na aplicação de redes neurais convolucionais (CNNs) para análise de áudio em um contexto socialmente relevante, o que ainda é pouco explorado. Além disso, a pesquisa demonstra que técnicas de classificação de emoções e extração de padrões acústicos podem ser aplicadas à detecção de situações de risco de violência de gênero. E ainda aborda a interseção entre ciência da computação e ciências sociais, ao enquadrar o estudo em teorias de sistemas de informação (*social network theory e social information processing theory*). Já no que se refere às contribuições tecnológicas, houve a proposta de um modelo computacional inovador para identificar emoções e termos relevantes em transcrições de áudio, servindo como base para sistemas inteligentes de prevenção. O trabalho destaca o potencial de integração com plataformas digitais, assistentes virtuais e aplicativos móveis, ampliando as possibilidades de monitoramento em tempo real. Dessa

forma, existe uma abertura para o desenvolvimento de sistemas de apoio à decisão, voltados a órgãos de segurança e redes de proteção às mulheres. Em termos de contribuições sociais, foi apresentada uma abordagem que pode antecipar situações de risco, permitindo respostas mais rápidas e eficazes por parte das autoridades competentes. O trabalho ainda contribui para políticas públicas e estratégias de prevenção do feminicídio e da violência de gênero, apoiando agendas nacionais e internacionais de direitos humanos. Além disso, incentiva a produção e uso de bases de dados em português, fundamentais para ampliar a aplicabilidade do modelo no Brasil.

A solução desenvolvida na pesquisa deve ser implementada em dispositivos capazes de captar áudio. A Figura 11 ilustra o ambiente onde ocorrem agressões contra mulheres, muitas resultando em feminicídio. Com a presença crescente de assistentes virtuais domésticos, o uso de redes neurais em conjunto com a transcrição de áudio para texto emerge como ferramenta capaz de garantir a segurança das vítimas.

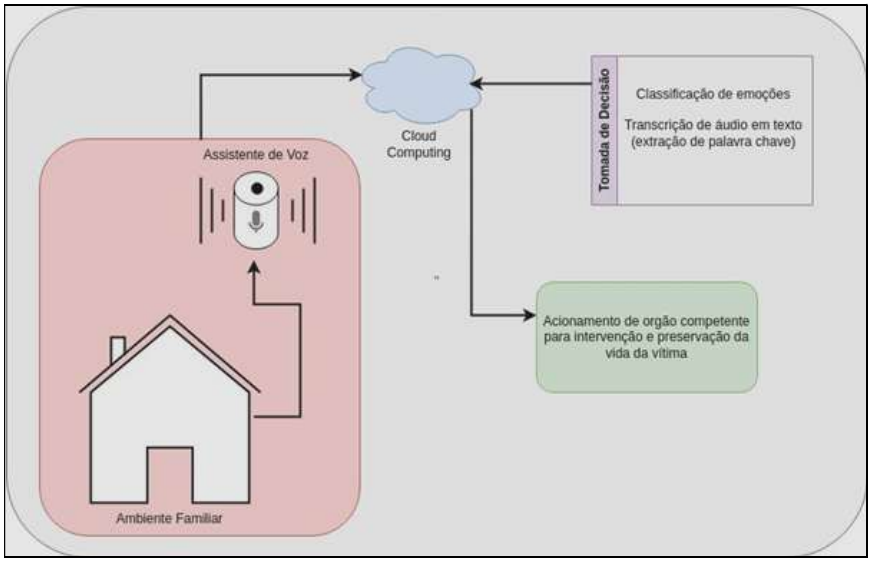

Figura 11. Esquematização da problemática e respectiva integração.

O áudio captado pelo assistente virtual é processado, havendo a extração de palavras-chave e classificação de emoções para desencadear um processo de acionamento para intervenção (Figura 11). O processo de tomada de decisão utiliza a classificação das emoções combinada com as palavras-chave extraídas, acionando um gatilho quando emoções e palavras-chave específicas são identificadas, com o envio de uma notificação de alerta (Figura 12).

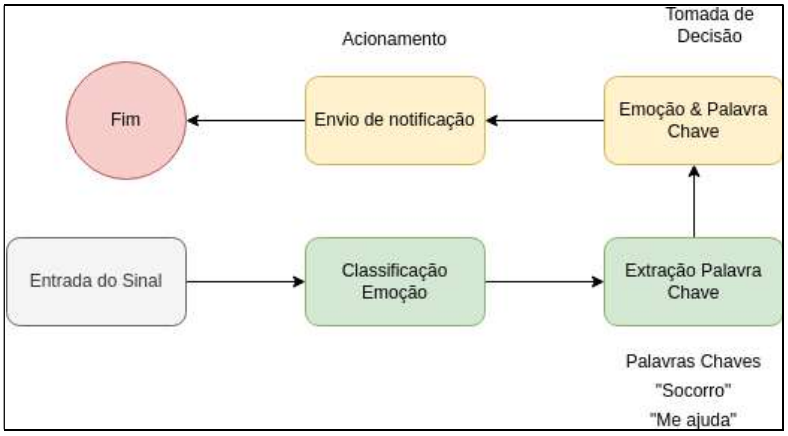

Figura 12. Fluxo para tomada de decisão.

6. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Esta pesquisa apresentou um modelo baseado em redes neurais convolucionais para a prevenção de feminicídios a partir da análise de áudio. A proposta contempla a identificação de emoções e a extração de termos relevantes de transcrições, possibilitando apoiar processos de tomada de decisão por meio de diferentes aplicações e protocolos de resposta.

O modelo foi treinado com o conjunto de dados *RAVDESS*, demonstrando capacidade de classificar emoções de forma consistente. Os resultados obtidos mostraram-se promissores e validam a viabilidade do uso de *CNNs* na detecção de padrões emocionais relacionados à violência de gênero. Como limitação e ameaça à validade da pesquisa, destaca-se a escassez de bases de dados em língua portuguesa, o que restringe a generalização e a aplicabilidade direta dos modelos no contexto brasileiro.

Para concluir, em termos de trabalhos futuros, propõe-se: (i) a ampliação das bases de dados com foco em cenários nacionais; (ii) a integração do modelo em uma plataforma inteligente de prevenção, que incorpore múltiplas fontes de dados (áudio, imagem e texto) e recursos de inteligência artificial; (iii) o desenvolvimento de aplicações *web*, assistentes virtuais e soluções móveis voltadas ao monitoramento em tempo real; e (iv) a efetiva integração com os sistemas de segurança pública, de forma a potencializar o impacto social da abordagem proposta.

REFERÊNCIAS

AGNES, A.; SAGAYARAJ, F. ADAGRAD **An Optimizer for Stochastic Gradient Descent**. Department of Computer Science and Engineering. Pondicherry Engineering College, 2019.

ALMEIDA, D. S.; NONATO, L. G. **Visualização de dados de violência contra a mulher em Minas Gerais e Goiás usando o Orange**. In: Anais do XI Encontro Nacional de Computação (ENACOMP), 2021.

ARAÚJO, A. C.; RAMOS, E. M.; ANDRIOLA, W. B. The Problem of Underrepresentation for Black Women in Public Universities. **Multidisciplinary Journal of Gender Studies**, v. 15, p. 156-171, 2026.

ARAÚJO, A. C.; ANDRIOLA, W. B.; RAMOS, E. M. **Interseccionalidade de gênero e raça nas universidades brasileiras: reflexões necessárias para superação de vulnerabilidades e fortalecimento da justiça social**. In: 58 International Congress of Americanists, 2025. Proceedings of 58 International Congress of Americanists. Novi Sad: Ministerio de Ciencia, Desarrollo Tecnológico e Innovación de la República de Serbia, 2025a.

ARAÚJO, A. C.; ANDRIOLA, W. B.; RAMOS, E. M. **Vivências e violências: reflexões sobre as experiências acadêmicas de mulheres negras nas universidades públicas do Brasil**. In: Congreso Internacional de Antropología - Narrativa de resistencia: comprendiendo las violencias en la actualidad, 2025, Valencia. Resúmenes del Congreso Internacional de Antropología - Narrativa de resistencia: comprehendiendo las violencias en la actualidad. Valencia, Spain: Editora de la Universitat de Valencia, 2025b.

ARAÚJO, A. C.; ANDRIOLA, W. B.; RAMOS, E. M. **Black Women and the problem of representation in Brazilian Public Higher Education**. In: 18th annual International Conference of Education, Research and Innovation, 2025. Seville: ICERI2025 Editor, 2025c.

ARAÚJO, A. C.; ANDRIOLA, W. B. **Mulheres negras na educação superior do Brasil: novos desafios para as políticas afirmativas**. In: XXI Congreso Internacional de Investigación Educativa, 2024. Resúmenes del XXI Congreso Internacional de Investigación Educativa. Barcelona: Editora de la Universidad de Barcelona, 2024.

ARAÚJO, A. C.; MENEZES, N. V. S.; MOREIRA, T. P.; PRESTES, T. K. A. Interseccionalidade nas universidades públicas do Ceará: práticas pedagógicas sob o olhar de mulheres negras universitárias. **Revista de Instrumentos, Modelos e Políticas em Avaliação Educacional**, v. 5, p. e024011, 2024.

ARAÚJO, A. C.; MENEZES, N. V. S.; MOREIRA, T. P. Os meus olhos coloridos me fazem refletir: Vivências acadêmicas de mulheres negras em uma universidade pública do Estado do Ceará. **Revista Cocar**, v. 21, p. 1-20, 2024.

BRASIL. **Lei nº 11.340, de 7 de agosto de 2006**. Cria mecanismos para coibir a violência doméstica e familiar contra a mulher. Diário Oficial da União, Brasília, DF, 8 ago.

BRASIL. **Lei nº 13.104, de 9 de março de 2015**. Altera o art. 121 do Decreto-Lei nº 2.848, de 7 de dezembro de 1940 - Código Penal, para prever o feminicídio como circunstância qualificadora do crime de homicídio. Diário Oficial da União, Brasília, DF, 10 mar. 2015.

BUENO, S. **Violência contra mulheres em 2021**. Disponível em: https://forumseguranca.org.br/wp-content/uploads/2022/03/violencia-contra-mulher-2021-v5.pdf. Acesso em: 02 mai. 2026.

CHO, K.; VAN MERRIËNBOER, B.; BAHDANAU, D.; BENGIO, Y. **On the Properties of Neural Machine Translation: Encoder–Decoder Approaches**. In Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation, p. 103–111. Doha, Qatar. Association for Computational Linguistics, 2014.

DIAS, L. S.; ARAÚJO, A. C. O legado do racismo estrutural na Educação Brasileira. **Revista Eletrônica Acta Sapientia**, v. 11, p. 1-26, 2024.

DIAS, L. S.; ARAÚJO, A. C.; NUNES, A. O.; ANDRIOLA, W. B.; RAMOS, E. M. Vivências de mulheres negras em cursos de graduação: estudo de caso na Universidade Vale do Acaraú (UVA). **Caderno Pedagógico**, v. 22, p. e17990-21, 2025.

DUCHI, J.; HAZAN, E.; SINGER, Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. **Journal of Machine Learning Research,** v. 12, p. 2121-2159, 2011.

FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA (FBSP). **Anuário Brasileiro de Segurança Pública 2024**. São Paulo: FBSP, 2024. Disponível em: https://forumseguranca.org.br. Acesso em março de 2025.

GRIN, D. G.; OLIVEIRA, M. B. O feminicídio e as formas de administração da violência contra as mulheres. **Revista Brasileira de Ciências Criminais**, v. 28, n. 165, p. 17-40, 2020.

HEVNER, A. R.; MARCH, S. T.; PARK, J.; RAM, S. Design Science in Information Systems Research. **MIS Quarterly**, v. 28, n. 1, p. 75–105, 2004.

IPEA – Instituto de Pesquisa Econômica Aplicada. **Atlas da Violência 2021**. Brasília: IPEA/Fórum Brasileiro de Segurança Pública, 2021.

JERONYMO, V. **Implementando métodos de otimização para treinamento de redes neurais com Pytorch**, 2019. Disponível em: https://www.ime.unicamp.br/~mac/db/2019-2S-157490.pdf.

KINGMA, D. P.; BA, J. **Adam: A method for stochastic optimization**. Proceedings of International Conference on Learning Representations, 2014.

LIVINGSTONE, S. R.; RUSSO, F. A. The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. **PLoS ONE**, v. 13, n. 5, e0196391, 2018.

MANTINI, L. A.; DA SILVA, Y. P. A.; MEIRELES, M. R. G. Reconhecimento Automático de Narrativas Violentas contra a Mulher Utilizando Classificadores Neurais. **Pista: Periódico Interdisciplinar: Sociedade, Tecnologia e Ambiente**, v. 3, n. 2, p. 141-163, 2021.

MENEGHEL, S. N.; PORTELLA, A. P. Feminicídios: conceitos, tipos e cenários. **Ciência & Saúde Coletiva**, v. 22, n. 9, p. 3077-3086, 2017.

MINISTÉRIO DAS MULHERES. **Relatório Anual Socioeconômico da Mulher (RASEAM)**. Brasília: Ministério das Mulheres, 2025.

PASINATO, W. Feminicídios e violência contra a mulher no Brasil: contribuições para o debate. **Revista Estudos Feministas**, v. 27, n. 1, p. 1-15, 2019.

PEFFERS, K.; TUUNANEN, T.; ROTHENBERGER, M. A.; CHATTERJEE, S. A. Design Science Research Methodology for Information Systems Research. **Journal of Management Information Systems**, v. 24, n. 3, p. 45–77, 2007.

RIBEIRO, M. F. C.; PEREIRA, M. M. **SALVE TODAS: um Sistema Inteligente para Auxiliar na Segurança de Mulheres**. Anais do XVI Women in Information Technology, 2022. DOI: https://doi.org/10.5753/wit.2022.222940.

RODRIGUES, A. S. C. **Feminicídio no Brasil: Uma reflexão sobre o direito penal como instrumento de combate a violência de gênero**. Universidade Federal Fluminense, 2016.

RUDER, S. **An overview of gradient descent optimization algorithms**, 2016.

WAISELFISZ, J. J. **Mapa da violência 2015: homicídio de mulheres no Brasil**. Brasília: Flacso Brasil, 2015.

XIN, R.; ZHANG, J.; SHAO, Y. Complex network classification with convolutional neural network. **Tsinghua Science and Technology**, v. 25, n. 4, p. 447–457, 2020.

ZANG, Z. **Improved Adam Optimizer for Deep Neural Networks**, IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), Banff, AB, Canada, 2018, pp. 1-2.