

APLICAÇÃO DE UMA REDE NEURAL CONVOLUCIONAL NO RECONHECIMENTO DO ALFABETO EM LIBRAS: UMA ESTRATÉGIA PARA A INCLUSÃO NA EDUCAÇÃO INFANTIL

Andreson Malaquias de Almeida

<https://orcid.org/0009-0009-7631-9687>

Gizele Melo Uchoa

<https://orcid.org/0000-0001-5541-3900>

RESUMO

Este estudo investiga a aplicação de Redes Neurais Convolucionais no reconhecimento de gestos estáticos do alfabeto da Língua Brasileira de Sinais (Libras), visando promover a inclusão educacional de crianças com deficiência auditiva. Empregou-se um modelo de aprendizagem de máquina, utilizando um conjunto de dados contendo 46.262 imagens representativas do alfabeto em Libras, como infraestrutura utilizamos a versão gratuita fornecida pelo Google Colab, obtendo assim um modelo com uma precisão de 97.58%. Os resultados demonstram o potencial das Redes Neurais Convolucionais em auxiliar a comunidade surda, indicando a necessidade de pesquisa contínua para aprimorar tecnologias assistivas neste campo.

Palavras-chave: Aprendizado de Máquina. Inclusão Educacional. Língua Brasileira de Sinais. Redes Neurais Convolucionais.

1 INTRODUÇÃO

A deficiência auditiva é uma condição que afeta aproximadamente 5% da população brasileira, correspondendo a cerca de 9,7 milhões de indivíduos (IBGE, 2010). Esta condição impõe desafios substanciais no processo de aquisição linguística para crianças surdas, uma vez que a deficiência auditiva restringe o acesso destes indivíduos à linguagem e às interações comunicativas.

De acordo com estudos realizados por Ferreira (2021), crianças surdas que não têm acesso à Língua de Sinais (Libras) apresentam um atraso no desenvolvimento linguístico, com um repertório linguístico restrito e dificuldade de comunicação. Este atraso pode levar a sentimentos de isolamento e exclusão social, bem como a desafios cognitivos e emocionais, particularmente evidentes quando estas crianças ingressam no ambiente escolar.

O Decreto nº 5.626, sob a Lei nº 10.436 de 2005, representa um marco legislativo significativo no Brasil, estabelecendo diretrizes fundamentais para os direitos educacionais de pessoas com deficiência auditiva. Este decreto assegura o acesso de alunos surdos à educação regular em território nacional, fomentando a inclusão educacional, com especial atenção à inserção da Língua Brasileira de Sinais (Libras) como componente curricular obrigatório em todos os cursos de licenciatura.

Além disso, a ascensão das tecnologias digitais tem se tornado predominante em diversas esferas do cotidiano, incluindo o setor educacional. A incorporação de tais tecnologias no processo educacional, conforme analisado por Medeiros (2011), tem sido

reconhecida como um elemento crucial na transformação dos métodos de ensino em diversos níveis acadêmicos, evidenciando o papel vital da evolução tecnológica em múltiplos setores da sociedade.

No contexto escolar, a implementação de novas tecnologias digitais de informação e comunicação (TDIC) acarreta desafios tanto para as instituições de ensino quanto para educadores e alunos. O principal desafio reside na natureza social da educação, que, historicamente, tem sido caracterizada por uma abordagem tradicional e conteudista, uma forma de educação bancária, como descrito por Paulo Freire. Este cenário aponta para uma necessidade de modernização nas estruturas educacionais, que frequentemente mantiveram uma abordagem estática na transmissão de conteúdo, conforme observado por Correia (2013).

Neste contexto, o presente trabalho propõe a aplicação de um modelo computacional utilizando a visão computacional e Redes Neurais Convolucionais (CNN) para o reconhecimento do alfabeto em Libras. Esta abordagem visa oferecer um método de ensino mais dinâmico e interativo, potencializando o processo de aprendizagem em crianças surdas e atendendo às necessidades específicas deste grupo no contexto educacional.

2 DESENVOLVIMENTO

O avanço tecnológico no campo do Aprendizado de Máquina tem aberto novas fronteiras na maneira como as máquinas interpretam e interagem com o mundo. Como elucidado por Bishop (2006), esta área é essencial para o desenvolvimento de sistemas capazes de aprender a partir de dados e realizar previsões ou tomar decisões informadas.

Intimamente ligada a este campo está a Visão Computacional, uma disciplina focada no processamento de imagens digitais para permitir a compreensão e interpretação de cenas visuais por sistemas computacionais, conforme descrito por Gonzalez e Woods (2000). Esta área tem sido fundamental no avanço da interação das máquinas com o ambiente visual, abrindo caminho para aplicações inovadoras em diversos setores.

As Redes Neurais Artificiais (RNAs) e, particularmente, as Redes Neurais Convolucionais (RNCs) são componentes centrais do aprendizado de máquina. Estes sistemas, baseado em neurônios biológicos, são notáveis pela sua habilidade em reconhecer padrões e resolver problemas complexos, especialmente no que diz respeito ao processamento de imagens. LeCun *et al.* (2015) destacam a relevância das RNCs na captura de características hierárquicas em dados visuais, tornando-as ideais para tarefas de visão computacional.

No contexto das tecnologias assistivas, a combinação de Aprendizado de Máquina e Visão Computacional tem se mostrado uma ferramenta poderosa. Para pessoas com deficiência auditiva, essas tecnologias oferecem meios avançados para facilitar a comunicação e o acesso à informação.

A inclusão de crianças surdas no ambiente escolar, reforçada pela Lei Brasileira de Inclusão da Pessoa com Deficiência (2015), é uma questão de grande relevância nas políticas educacionais atuais. Neste sentido, a implementação de tecnologias assistivas baseadas em Aprendizado de Máquina e Visão Computacional representa um avanço significativo. Essas tecnologias têm o potencial de transformar o ambiente educacional, promovendo a inclusão efetiva e a participação plena de alunos surdos.

2.1 Metodologia Implementada

Na implementação do algoritmo de reconhecimento de sinais estáticos do alfabeto de Libras, empregou-se um ambiente computacional disponibilizado pelo Google Colab. A infraestrutura utilizada incluiu um processador Intel Xeon de 2GHz, acompanhado de 16Gb de RAM e uma placa gráfica Tesla T4 de 16Gb. O sistema operacional Linux 64 bits foi escolhido, juntamente com a linguagem de programação Python, reconhecida pela sua vasta aplicabilidade em tarefas relacionadas ao aprendizado de máquina e visão computacional.

A seleção do conjunto de dados representou uma etapa crítica do estudo, culminando na escolha do dataset Libras, referenciado por Oliveira (2022). Este conjunto abrange 46.262 imagens das letras do alfabeto em Libras. Algumas letras, de natureza dinâmica, foram excluídas da análise. As imagens, padronizadas para uma dimensão de 64x64 pixels e em formato RGB, foram divididas em 34.714 para treino e 11.548 para teste, abrangendo 21 classes correspondentes às letras do alfabeto, com exceção de H, J, K, X e Z.

O pré-processamento dos dados incluiu uma fase crucial de normalização, escalando os valores dos *pixels* para um intervalo específico. Esta técnica assegurou a consistência e estabilidade dos dados, dada a variabilidade das fontes e das condições de iluminação. A biblioteca OpenCV2 foi utilizada para a leitura das imagens, mantendo o canal de três cores (RGB).

A arquitetura da CNN adotada consistiu em camadas sequenciais. Inicialmente, foram implementadas três camadas convolucionais, seguidas por camadas de pooling para redução da dimensionalidade. Os recursos foram então achatados e inseridos em uma camada densa, utilizando a ativação ReLU. O método de *dropout* foi aplicado para regularização da rede, minimizando o *overfitting*.

2.2 Análise dos Resultados

A eficácia do modelo foi demonstrada pela acurácia de 97.58% e pela baixa taxa de perda de 0.1128 após 50 épocas de treinamento. Esses resultados indicam que o modelo conseguiu aprender efetivamente as características distintivas das letras estáticas em Libras. As técnicas de normalização e a utilização da OpenCV2 contribuíram significativamente para a uniformidade e consistência dos resultados.

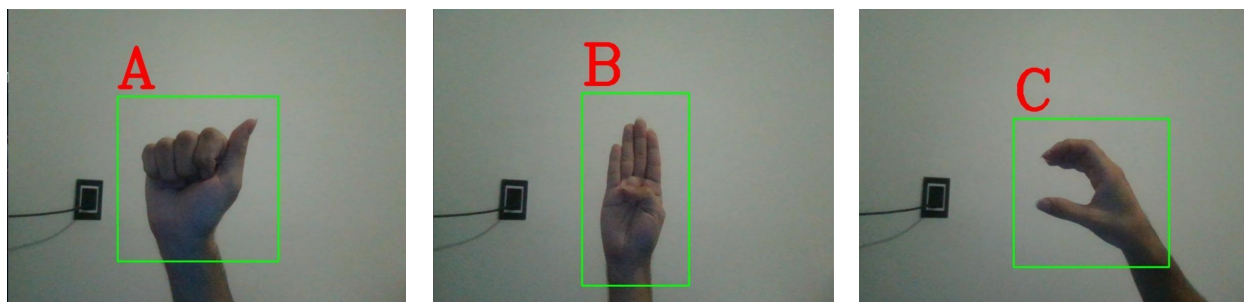
No entanto, o modelo apresentou limitações no reconhecimento de letras com gestos dinâmicos, como H, J, K, X e Z. Esta restrição ressalta a necessidade de pesquisa adicional para integrar a captura de movimento ou vídeo, visando o reconhecimento de sinais dinâmicos.

Durante o treinamento, enfrentaram-se desafios como o estouro de memória na máquina do Google Colab, que foram superados mediante ajustes nos hiperparâmetros. Esta situação enfatiza a importância de um ambiente de treinamento adequado e a necessidade de uma otimização contínua do modelo.

As imagens abaixo exemplificam alguns dos resultados alcançados pelo modelo. Estas ilustrações demonstram a capacidade do modelo de identificar corretamente as

letras do alfabeto em Libras.

Figura 1 - Letras do alfabeto em Libras



Fonte: Autor (2023).

Este estudo, portanto, destaca o potencial das Redes Neurais Convolucionais no desenvolvimento de tecnologias assistivas para a inclusão educacional de crianças surdas, abrindo caminho para futuras pesquisas e aplicações práticas neste campo.

3 CONCLUSÃO

Este estudo sobre o uso de Redes Neurais Convolucionais (RNCs) para o reconhecimento de sinais estáticos do alfabeto em Libras representa um passo significativo na direção da inclusão educacional de indivíduos surdos. Os resultados alcançados, com uma precisão de 97.58% e uma taxa de perda de 0.1128, demonstram a eficácia das RNCs no processamento e classificação de imagens relativas à linguagem de sinais.

No entanto, o estudo identificou limitações no que tange ao reconhecimento de letras do alfabeto em Libras que envolvem movimentos dinâmicos. Esta constatação sugere a necessidade de futuras pesquisas que incluam métodos para capturar e interpretar gestos dinâmicos, expandindo assim a capacidade do modelo para abranger todas as letras do alfabeto em Libras.

Os desafios enfrentados durante o treinamento, particularmente relacionados ao estouro da memória, destacam a importância de um ambiente de treinamento otimizado e da constante revisão dos hiperparâmetros. Essas questões são essenciais para melhorar a performance e a aplicabilidade de modelos de aprendizado de máquina em situações reais.

Em relação à inclusão educacional, este estudo sublinha a importância das tecnologias assistivas no suporte à aprendizagem de alunos surdos. A aplicação de RNCs em ferramentas educacionais para a interpretação de Libras pode contribuir para um ambiente de aprendizagem mais inclusivo e acessível, promovendo a igualdade de oportunidades para todos os alunos.

Em suma, a pesquisa demonstra o potencial significativo das Redes Neurais Convolucionais na criação de ferramentas assistivas para a comunidade surda. Embora haja espaço para melhorias e desenvolvimentos futuros, os resultados obtidos apontam para o impacto positivo que tais tecnologias podem ter na inclusão educacional e na superação de barreiras comunicativas enfrentadas por indivíduos surdos.

REFERÊNCIAS

BISHOP, C. M.; NASRABADI, N. M. **Pattern recognition and machine learning**. New York: springer, 2006.

BRASIL. **Decreto nº 5.626, de 22 de dezembro de 2005**. Regulamenta a Lei nº 10.436, de 24 de abril de 2002, que dispõe sobre a Língua Brasileira de Sinais – Libras, e o art. 18 da Lei nº 10.098, de 19 de dezembro de 2000. Diário Oficial da União. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/Decreto/D5626.htm. Acesso em: set. 2023.

BRASIL. **Lei nº 13.146, de 6 de julho de 2015**. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Brasília, DF: Presidência da República, 2015. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: set. 2023.

CORREIA, W. F.; GALLO, S. O que é conservadorismo em educação. **Conjectura: Filos. Educ., Caxias do Sul**. v. 18, n. 2, p. 78-90, 2013.

FERREIRA, L. da C. **Educação Inclusiva a Importância das Libras no Ensino Infantil para Crianças Surdas e Ouvintes**. 2021. Dissertação de Mestrado.

GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais**. Editora Blucher, 2000.

IBGE - Instituto Brasileiro de Geografia e Estatística . **Censo Brasileiro de 2010**. Rio de Janeiro: IBGE, 2012. Disponível em: <https://www.ibge.gov.br/apps/snig/v1/?loc=0&ind=4643&cat=-1,-2,-3,128>. Acesso em: set. 2023

LECUN, Y. *et al.* **Gradient-based learning applied to document recognition**. **Proceedings of the IEEE**. v. 86, n. 11, p. 2278-2324, 1998.

MEDEIROS, C. B. **Computação: o terceiro pilar**. Revista USP, n. 89, p. 134-147, 2011.

OLIVEIRA, W. **Libras**. 2022. Disponível em: <https://www.kaggle.com/williansoliveira/datasets>. Acesso em: set. 2023.