


ENSINANDO UM ROBÔ OS NÍVEIS DE ESCRITA: INTELIGÊNCIA ARTIFICIAL COMO APOIO AO DOCENTE ALFABETIZADOR

 <https://doi.org/10.56238/arev6n4-196>

Data de submissão: 12/11/2024

Data de publicação: 12/12/2024

Helaine Aparecida da Silva

Programa de Pós-graduação em Educação – Universidade Católica de Petrópolis (UCP)

Robson da Cruz Augusto

Centro de Engenharia e Computação – Universidade Católica de Petrópolis (UCP)

Karollayne Ramos dos Santos

Centro de Engenharia e Computação – Universidade Católica de Petrópolis (UCP)

Daniel de Oliveira Lima

Centro de Engenharia e Computação – Universidade Católica de Petrópolis (UCP)

Ana Carolina Carius

Programa de Pós-graduação em Educação – Universidade Católica de Petrópolis (UCP)

Centro de Engenharia e Computação – Universidade Católica de Petrópolis (UCP)

Felipe de Oliveira Baldner

Centro de Engenharia e Computação – Universidade Católica de Petrópolis (UCP)

André Maiworm Probst

Centro de Engenharia e Computação – Universidade Católica de Petrópolis (UCP)

RESUMO

O objeto de estudo desta pesquisa é a adaptação de ferramentas de visão computacional por Inteligência Artificial para identificação dos níveis de escrita dos estudantes dos Anos Iniciais do ensino fundamental, a partir de Ditados Conceituais. O objetivo da pesquisa é produzir um aplicativo que auxilie os professores alfabetizadores na identificação dos níveis de escrita dos estudantes, a partir das imagens dos manuscritos das crianças, de modo a propor atividades personalizadas. A metodologia adotada foi o Design Based Research (DBR), onde cada etapa de desenvolvimento do protótipo inclui sugestões dos próprios usuários com o objetivo de melhorar o desempenho do protótipo. As conclusões iniciais apontam para as limitações das ferramentas de visão computacional e a necessidade de aprimoramentos.

Palavras-chave: Inteligência Artificial. Alfabetização. Apoio Docente. Níveis de Escrita.

1 INTRODUÇÃO

O processo de alfabetização se constitui em uma etapa primordial nos Anos Iniciais do ensino fundamental, pois consolida a apropriação da leitura e escrita de uma criança. Porém, tal processo se configura muito mais complexo do que se pensa e vai além do domínio da codificação e decodificação das palavras [Pérez 2007].

Em grande parte das escolas públicas brasileiras, o processo de alfabetização inicial muitas vezes tem tido como resultado o insucesso e uma defasagem considerável. De modo geral, formam alunos com dificuldades em ler, escrever, interpretar e produzir pequenos textos [Pontes; Diniz & Martins-Reis 2013].

Em 12 de junho de 2023, por meio do Decreto nº 11.556 [Brasil 2023], o Governo Federal instituiu o Compromisso Nacional Criança Alfabetizada. O programa busca garantir a alfabetização de todas as crianças do país até o final do 2º ano do ensino fundamental. Além disso, busca recuperar as aprendizagens afetadas pela pandemia (SARS-CoV-2/COVID 19), de 100% das crianças matriculadas no 3º, 4º e 5º ano. O investimento será de mais de R\$ 2 bilhões ao longo de quatro anos, além de um esforço conjunto entre a União, Estados e Distrito Federal.

Diante do cenário descrito, ou seja, de recuperar as aprendizagens dos alunos no pós-pandemia, enxergar o processo de alfabetização nos Anos Iniciais do ensino fundamental, através de outra perspectiva tornou-se um desafio. Nesse sentido, os autores se debruçaram em diferentes estratégias capazes de melhorar o processo de alfabetização das crianças, incluindo ferramentas de tecnologia.

Portanto, a questão norteadora desse trabalho de pesquisa é: Como as ferramentas automatizadas a nível digital, como Inteligência Artificial para reconhecimento de imagens, poderiam ajudar o docente alfabetizador na identificação dos níveis de escrita¹ dos alunos?

A fim de responder à questão proposta intencionamos, como objetivo desse trabalho, desenvolver uma metodologia empregando algoritmos de redes neurais em *deep learning* (aprendizagem profunda) para reconhecimento e identificação dos níveis de escrita que, alinhado à classificação dos níveis de escrita já existente, seja um apoio ao trabalho do docente alfabetizador nas atividades cotidianos de sala de aula.

¹ Níveis de escrita se referem a uma classificação para os chamados Ditados Conceituais, desenvolvida por Emília Ferreiro e Ana Teberosky (1999). Os alunos são classificados em Pré-silábico, Silábico sem valor, Silábico com valor, Silábico Alfabético, Alfabético e Ortográfico.

2 O PROCESSO DE RECONHECIMENTO DE IMAGENS

O processo de reconhecimento de imagens não é, em si, algo muito inovador. Registra-se que os primeiros processos que envolvem reconhecimento facial, por exemplo, são da década de 1960. No entanto, a ausência de bancos de dados robustos, assim como processamento de máquina mais eficientes, impediram avanços significativos [Ambika & Suresh 2024].

Com o avanço de técnicas em *deep learning*, assim com a utilização de redes neurais convolucionais, os processos de reconhecimento de imagens se tornam mais desenvolvidos, permitindo a detecção de maiores detalhes nas imagens, a transmissão destes através das cadeias de convolução e refinando os processos de reconhecimento [Géron 2019].

A fim de realizar a proposta de reconhecimento das imagens dos Ditados Conceituais obtidos pelo grupo de pesquisadores em uma escola da rede municipal de uma cidade do estado de Minas Gerais, criou-se um código, na linguagem Python, usando de conceitos de Inteligência Artificial, tais como reconhecimento e processamento de imagens, análise de dados e conceitos estatísticos para reconhecimento de caracteres, letras, palavras ou frases escritas a mão por alunos que estão cursando do segundo ao quinto ano do ensino fundamental da referida instituição de ensino.

Os dados utilizados para essa pesquisa foram coletados de alunos e professores que estiveram fazendo as atividades de Ditado Conceitual no espaço escolar. Os dados foram convertidos em imagens e posteriormente foram pré-processadas para que pudessem ser preparadas para uso no código. Aproximadamente 1000 imagens de palavras ou frases foram usadas para essa implementação.

As imagens foram limpas usando técnicas de edição de imagens para retirar manchas, rabiscos, para separar umas das outras pois, em alguns casos, vinham com as escritas muito juntas, ou outras intempéries que pudessem atrapalhar o pré-processamento das imagens.

A ideia inicial era criar e treinar uma rede neural nova para aprender a reconhecer imagens de escrita à mão. Entretanto, como hoje já existem redes neurais treinadas para fazer tal tarefa, por exemplo as famosas *Convolutional Neural Network* (Rede Neural Convolucional), ou CNN, que são redes poderosas para processamento de imagens, usamos técnicas de visão computacional e processamento de imagens existentes para a realização dessa pesquisa.

Para tal tarefa utilizou-se as seguintes bibliotecas:

Pillow (Python Imaging Library, Fork) é uma biblioteca do Python usada para processamento de imagens. Ela nos ajuda a abrir, manipular e salvar imagens em diferentes formatos.

Além desta, utilizou-se o *Tesseract ORC*. *Tesseract* é uma ferramenta *open-source* de ORC (Reconhecimento óptico de caracteres) que extrai textos escritos ou impressos em imagens. Foi

originalmente desenvolvido por *Hewlett-Packard*, e seu desenvolvimento foi, mais tarde, assumido pela *Google*. Hoje é conhecido como *Google Tesseract OCR*.

Tesseract não é uma rede neural necessariamente dita, mas sim uma ferramenta óptica para reconhecimento de caracteres. Porém, *Tesseract* usa técnicas de *machine learning*, incluindo redes neurais, para reconhecimento e processamento de imagens.

OpenCV (Open Source Computer Vision Library) também não é uma rede neural propriamente dita, mas sim uma ampla biblioteca de funcionalidades em programação, principalmente usada em situações de visão computacional que exijam uma aproximação em tempo real. OpenCV proporciona ferramentas para processamento de imagens e vídeos, que podem ser usados em conjunto com redes neurais para realizar diversas tarefas, tais como reconhecimento de objetos, reconhecimento facial e segmentação de imagens.

OpenCV pode ser usada para pré-processamento de imagens antes de usá-las em redes neurais. E esse foi o uso dessa biblioteca em nossa implementação. OpenCV também trabalha em conjunto com redes neurais de aprendizado profundo (*deep learning*), como TensorFlow, Keras e PyTorch.

Numpy é uma biblioteca famosa no Python usada para cálculos matemáticos.

Matplotlib é uma biblioteca, também conhecida no Python, usada para plotar gráficos 2D e 3D e no nosso caso, foi usada em conjunto com a biblioteca OpenCV para plotar as imagens.

A implementação do código para testar nossa análise foi feita online através do Google Colab, no qual, tivemos todas as implementações e experimentações executadas.

3 PERCURSO METODOLÓGICO

O presente estudo se caracteriza como um estudo de caso, uma vez que o material estudado refere-se aos Dados Conceituais obtidos em uma escola da rede municipal de uma cidade de Minas Gerais, podendo ser generalizado posteriormente a partir de novos dados inseridos no banco de dados utilizado para os algoritmos de *deep learning* [Yin

De modo a melhorar os resultados obtidos com os protótipos, a cada novo ajuste, escolheu-se como metodologia de pesquisa o *Design Based Research (DBR)*, cujo objetivo é gerar vários protótipos, que são aperfeiçoados à medida que são colocados em teste pelos usuários do produto em construção. Cada etapa de validação do protótipo inclui sugestões e adaptações geradas em contato direto com os potenciais usuários da ferramenta [Matta & Boaventura 2014].

Caracteriza-se a pesquisa como quantitativa pois, até o presente, o interesse é obter um aplicativo que identifica as imagens através de visão computacional. Análise se baseia na acurácia dos resultados, de forma percentual [Creswell & Creswell 2021].

4 RESULTADOS PRELIMINARES

O código criado para analisar a tarefa foi desenvolvido em etapas, conforme a necessidade de se obter resultados coerentes, tomando-se como norteador o percentual de acurácia em cada teste realizado.

4.1 PRIMEIRA ABORDAGEM

Nesta primeira abordagem, o código foi alimentado com uma imagem aleatória e depois digitou-se a palavra na qual se esperava ser reconhecida na imagem. Com isso esse teste se objetivou reconhecer e apresentar o resultado, em percentual de acurácia, de quanto o algoritmo foi capaz de reconhecer essa imagem. Nenhuma técnica de processamento de imagem foi usada. Os resultados são apresentados nas Figuras 1 e 2.

Figura 1: Teste de semelhança entre uma palavra digitada e a imagem, sem utilização de técnicas de processamento de imagem.

A porcentagem de semelhança entre a palavra digitada e o texto na imagem é: 54.55%

Texto na imagem:

Pabsn

Robson

Figura 2: Teste de semelhança entre uma palavra digitada e a imagem, sem utilização de técnicas de processamento de imagem.

A porcentagem de semelhança entre a palavra digitada e o texto na imagem é: 13.33%

Texto na imagem:

DA nsçen

almogor

4.2 SEGUNDA ABORDAGEM

Para a segunda abordagem, sugestões foram consideradas e o código foi alimentado com uma imagem aleatória e depois digitou-se a palavra a qual se esperava ser reconhecida na imagem. Esse teste objetivou reconhecer e apresentar o resultado, em percentual de acurácia, de quanto o algoritmo foi capaz de reconhecer a imagem proposta. Técnicas de processamento de imagem foram usadas. Os resultados são apresentados nas Figuras 3 e 4.

Figura 3: Reconhecimento da palavra da imagem, aleatória, digitando-se posteriormente a qual esperava-se ser reconhecida, presente na imagem.

Digite a palavra que deseja comparar: MINECRAFT

Escolher Arquivos minecraft.jpeg

- **minecraft.jpeg**(image/jpeg) - 4620 bytes, last modified: 24/05/2024 - 100% done

Saving minecraft.jpeg to minecraft.jpeg

A porcentagem de semelhança entre a palavra digitada e o texto na imagem é: 63.16%

Texto na imagem:

M| NE CAET

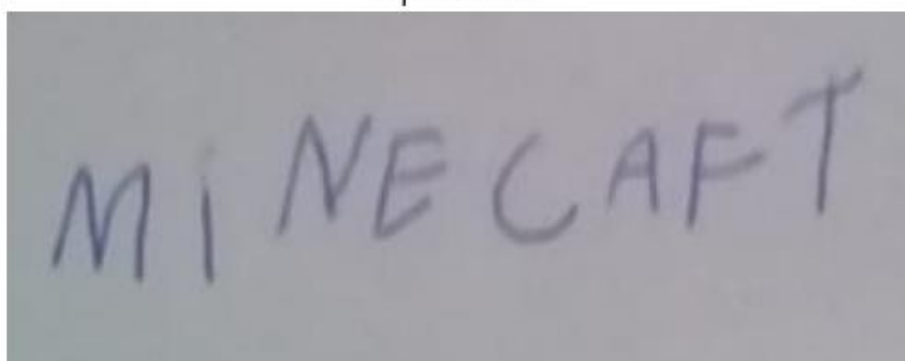


Figura 4: Reconhecimento da palavra da imagem, aleatória, digitando-se posteriormente a qual esperava-se ser reconhecida, presente na imagem.

Digite a palavra que deseja comparar: CINZENTO

Escolher Arquivos CINZENTO.JPEG

- **CINZENTO.JPEG**(image/jpeg) - 19842 bytes, last modified: 24/05/2024 - 100% done

Saving CINZENTO.JPEG to CINZENTO.JPEG

A porcentagem de semelhança entre a palavra digitada e o texto na imagem é: 47.06%

Texto na imagem:

Si sSENTO



4.3 TERCEIRA ABORDAGEM

Por fim, para a terceira abordagem, novas sugestões foram acolhidas e o código foi alimentado com duas imagens: uma sendo da palavra esperada e a outra com a do Ditado Conceitual do banco de imagens, o qual se pretendia reconhecer. Esse teste objetivou reconhecer e apresentar o resultado, em percentual de acurácia, de quanto o algoritmo foi capaz de reconhecer a semelhança entre as duas imagens. Técnicas de processamento de imagem foram usadas. Os resultados são apresentados nas Figuras 5, 6 e 7.

Figura 5: Abordagem com a palavra que se esperava reconhecer e com a imagem proveniente do Ditado Conceitual.

Por favor, faça o upload da imagem com a palavra esperada.

Escolher Arquivos MINECRAFT...perado.PNG

• **MINECRAFT_esperado.PNG**(image/png) - 2396 bytes, last modified: 24/05/2024 - 100% done

Saving MINECRAFT_esperado.PNG to MINECRAFT_esperado.PNG

Por favor, faça o upload da imagem que você quer comparar.

Escolher Arquivos minecraft.jpeg

• **minecraft.jpeg**(image/jpeg) - 4620 bytes, last modified: 24/05/2024 - 100% done

Saving minecraft.jpeg to minecraft (1).jpeg

A porcentagem de semelhança entre o texto esperado e o texto na imagem é: 63.16%

Texto esperado:
MINECRAFT

MINECRAFT

Texto na imagem:
M| NE CAET

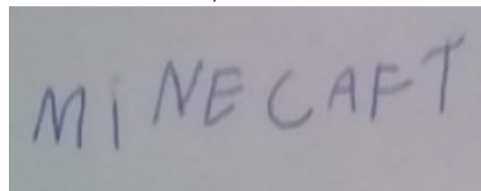


Figura 6: Abordagem com a palavra que se esperava reconhecer e com a imagem proveniente do Ditado Conceitual.

Por favor, faça o upload da imagem com a palavra esperada.

Escolher Arquivos aprendizagem...perado.PNG

• **aprendizagem_esperado.PNG**(image/png) - 4153 bytes, last modified: 24/05/2024 - 100% done

Saving aprendizagem_esperado.PNG to aprendizagem_esperado.PNG

Por favor, faça o upload da imagem que você quer comparar.

Escolher Arquivos aprendizagem.jpeg

• **aprendizagem.jpeg**(image/jpeg) - 19916 bytes, last modified: 24/05/2024 - 100% done

Saving aprendizagem.jpeg to aprendizagem.jpeg

A porcentagem de semelhança entre o texto esperado e o texto na imagem é: 37.04%

Texto esperado:
aprendizagem

aprendizagem

Texto na imagem:
Opremas 39 G Mm

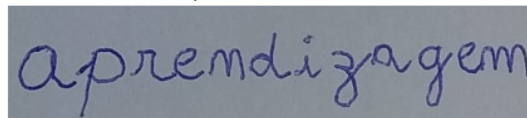
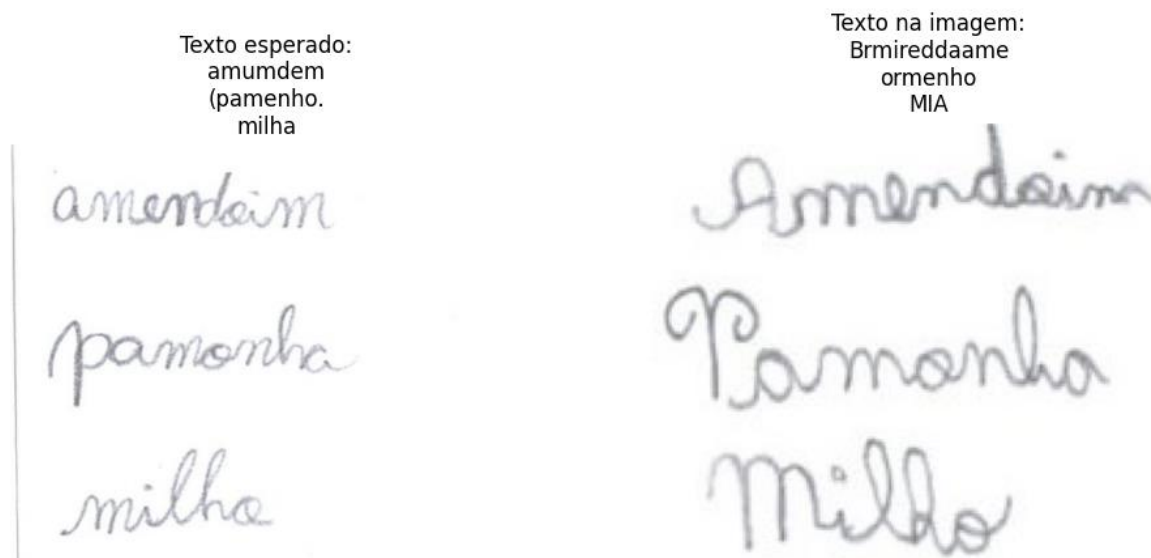


Figura 7: Abordagem com a palavra que se esperava reconhecer e com a imagem proveniente do Ditado Conceitual.
A porcentagem de semelhança entre o texto esperado e o texto na imagem é: 42.55%



5 DISCUSSÕES E PERSPECTIVAS FUTURAS

O presente trabalho se propôs a implementar, através de técnicas envolvendo Inteligência Artificial, na forma de *machine learning* (*deep learning*) o reconhecimento de imagens no campo da visão computacional, com o objetivo de classificar os Ditados Conceituais de estudantes dos Anos Iniciais do ensino fundamental em seus diferentes níveis de escrita. A proposta é inovadora, uma vez que não há, até o presente, ferramentas de Inteligência Artificial que consigam identificar o nível de escrita em estudantes no processo de alfabetização no Brasil.

Uma vez que o trabalho se encontra em andamento, os resultados apresentados até agora restringem-se à testagem de diferentes abordagens que visam o reconhecimento de imagens dos Ditados Conceituais, de modo a avaliar a acurácia das propostas. Ainda não se verificou uma acurácia satisfatória. Diversos fatores contribuem para a ausência de resultados melhores: ampla variabilidade da grafia dos estudantes, poucas imagens que compõem o banco de imagens para treinamento das ferramentas de processamento de imagens e testagem de outras abordagens, como comparação entre grafias dos Ditados Conceituais de alunos no nível ortográfico com alunos no nível silábico alfabético, por exemplo, os quais serão avaliados na continuidade desta pesquisa.

Os próximos passos referem-se a, a partir da detecção da imagem, classificá-las no nível de escrita adequado para, posteriormente, ser implementado em um aplicativo para utilização nas escolas.

REFERÊNCIAS

- Ambika, G. N. & Suresh, Y. (2024). Mathematics for 2D face recognition from real time image data set using deep learning techniques. *Bulletin of Electrical Engineering and Informatics*, 13(2), 1228-1237.
- Brasil (2023). Decreto nº 11556, de 12 de junho de 2023. Institui o Compromisso Nacional com a Criança Alfabetizada. Secretaria Especial para Assuntos Jurídicos. Casa Civil. Presidência da República.
- Creswell J.W. & Creswell, J. D. (2021) Projeto de pesquisa: métodos qualitativo, quantitativo e misto. Tradução: Sandra Maria Mallmann da Rosa. 5ª edição. Porto Alegre: Penso.
- Géron, A. (2019). Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow. Alta Books.
- Matta, A. E. R., Silva, F. D. P. S. D., & Boaventura, E. M. (2014). Design-based research ou pesquisa de desenvolvimento: metodologia para pesquisa aplicada de inovação em educação do século XXI. *Revista da FAEEBA: educação e contemporaneidade*, 23(42), 23-36.
- Pérez, C. L. V. (2007). Alfabetização para além do método: uma sintaxe freiriana . *Acolhendo a Alfabetização Nos Países De Língua Portuguesa*, 1(2), 46-55. <https://doi.org/10.11606/issn.1980-7686.v1i2p46-55>
- Pontes, V. L., Diniz, N. L. F., & Martins-Reis, V. D. O. (2013). Parâmetros e estratégias de leitura e escrita utilizados por crianças de escolas pública e privada. *Revista Cefac*, 15, 827-836.
- Yin, R. K. (2015) Estudo de caso: planejamento e métodos. Tradução: Cristhian Matheus Herrera. 5ª edição. Porto Alegre: Bookman.
- Teberosky, A., & Ferreiro, E. (1999). *Psicogênese da língua escrita*. Porto Alegre: Artmed.