JOGO SOCIOEDUCATIVO BASEADO EM ESTRATÉGIAS DE VISÃO COMPUTACIONAL E COOPERAÇÃO DE SURDOS E OUVINTES PARA O APRENDIZADO DE LIBRAS

William Marques[1], Anderson Pereira das Neves[1]

[1]Instituto Federal de Mato Grosso do Sul

junior.stuart34@gmail.com, anderson.neves@ifms.edu.br

Resumo

(O resumo deve apresentar de 100 a 150 palavras)

Libras é a língua de sinais usada pelos brasileiros, sendo reconhecida oficialmente em 2002. Apesar de receber mais atenção nos últimos anos, ainda existem diversas barreiras no seu ensino, que é realizado por técnicas tradicionais, podendo ocorrer o desinteresse tanto pelos surdos como de ouvintes. À vista disso, a tecnologia pode ser aplicada em um papel de inclusão, através de diferentes áreas do conhecimento, podendo aproximar e permitir uma melhor interação entre surdos e ouvintes, além de facilitar e principalmente motivar o aprendizado de Libras. Sendo assim, com o propósito de auxiliar na promoção de interatividade, cooperação e aprendizado em Libras, neste trabalho é proposto o desenvolvimento de um jogo digital socioeducativo utilizando técnicas de visão computacional (VC) para o reconhecimento de gestos, possibilitando a interação surdo-ouvinte e consequentemente o ensino de Libras.

Palavras-chave: Libras, Inclusão, Jogo Socioeducativo.

Introdução

Segundo o último censo realizado em 2010 pelo Instituto Brasileiro de Geografia e Estatística (IBGE), aproximadamente 9,7 milhões de pessoas possuem alguma deficiência auditiva, o que ultrapassa 5% da população do país. Desse número mais de 2 milhões possuem deficiência severa. (BRASIL, 2016).

Conforme os estudos de Carvalho (2007), Libras é a língua de sinais usada pelos surdos brasileiros e foi criada junto com o Instituto Nacional de Educação de Surdos (INES), a partir de uma mistura entre a Língua Francesa de Sinais e de gestos já utilizados no Brasil.

Segundo o Estatuto da Pessoa com Deficiência, a acessibilidade se refere a possibilidade e condição de alcance para utilização, com segurança e autonomia, de espaços, mobiliários, equipamentos urbanos, edificações, transportes, informação e comunicação, inclusive seus sistemas e tecnologias, bem como de outros serviços e instalações abertos ao público, de uso público ou privados de uso coletivo, tanto na zona urbana como na rural, por pessoa com deficiência ou com mobilidade reduzida". (BRASIL, 2015).

Um importante ramo de pesquisa que engloba Libras e tecnologias é o reconhecimento de gestos por meio de técnicas de Visão Computacional (VC). Essa tecnologia tem o objetivo de extrair informações úteis das imagens para tomar decisões úteis sobre o ambiente e cenas por meio de imagens e vídeos (PRINCE, 2012; SHAPIRO, STOCKMAN, 2001).

Um jogo pode ser definido como uma atividade lúdica, sendo também, um ato voluntário concretizado como evasão da vida real, limitado pelo tempo e espaço, criando a ordem através de uma perfeição temporária (HUIZINGA, 2003).

Sendo assim, um jogo digital (eletrônico) é uma atividade lúdica formada por ações e decisões que resultam numa condição final. Essas ações e decisões são limitadas por um conjunto de regras e por um universo, que no contexto dos jogos digitais, são regidos por um programa de computador.

O universo contextualiza as ações e decisões do jogador, fornecendo a ambientação adequada à narrativa do jogo, enquanto as regras definem o que pode e o que não pode ser realizado, bem como as consequências das ações e decisões do jogador. Além disso, as regras fornecem desafios a fim de dificultar ou impedir o jogador de alcançar os objetivos estabelecidos (SCHUYTEMA, 2008).

O presente trabalho faz parte de um conjunto de obras já desenvolvidos que com o auxílio de um algoritmo de extração de características já projetado, busca por meio de técnicas de visão computacional, desenvolver um jogo digital que promova o interesse na aprendizagem de Libras.

Metodologia

Inicialmente foi realizado o estudo da linguagem Python e suas bibliotecas que foram usadas no decorrer do projeto para possibilitar o reconhecimento de gestos, como a biblioteca OpenCV. Também foi realizado um estudo sobre Deep Learning e como aplicar no projeto. Começando pela rede neural mais simples, o perceptron, para compreender a mecânica de funcionamento das redes neurais.

O perceptron é um modelo de rede neural, desenvolvido na década de 1950 e 1960, pelo cientista da computação Frank Rosenblatt. Basicamente é formado por três camadas:

entrada, oculta e saída. Na camada de entrada recebemos dados binários e atribuímos um peso a cada um, na camada oculta ocorre o processamento dos dados, realizada por alguma função matemática, por exemplo, a função degrau. Por último na camada de saída deste modelo temos uma única saída binária. O aprendizado deste modelo é baseado no ajuste dos pesos. (MICHAEL NIELSEN, 2015)

Após os estudos sobre o perceptron, foi iniciado um estudo sobre uma rede neural mais complexa, a MLP ou Multi Layer Perceptron (Perceptron de múltiplas camadas). Esta rede que possibilita resolver problemas não lineares, suportando funções mais complexas e principalmente fazendo uso da Deep Learning.

Deep Learning ou traduzindo do inglês, aprendizagem profunda é um método da área de aprendizagem de máquina, que tem como objetivo extrair padrões de maneira automática a partir de dados, como imagens e vídeo. Basicamente na aprendizagem profunda recebemos um entrada X e o usa para prever uma saída Y. O Deep Learning atualmente é muito utilizado na classificação de imagens, reconhecimento de fala, detecção de objetos e descrição de conteúdos.

Após concluir as tarefas anteriores, ocorreu o Enict (Encontro de Iniciação Científica e Tecnológica) no IFMS, onde estes estudos preliminares foram apresentados, para fins de experiência e sugestões de melhorias.

Concluindo esta etapa focamos em procurar algoritmos em python de reconhecimento da mão e realizamos testes para avaliarmos o desempenho e o que precisaria ser feito para conseguir extrair características somente da mão.

Posteriormente foi dada a tarefa de cortar e delimitar a mão no vídeo para possibilitar a evolução do projeto. A função escolhida para realizar tal tarefa foi o crop, que faz parte do PIL (biblioteca de imagens python).

O crop funciona passando quatro parâmetros de coordenadas, esquerda, direita, superior e inferior de pixels, e retorna uma imagem com uma região de acordo com os parâmetros na imagem.

Durante o tempo estudando e implementando o crop, obtivemos uma outra perspectiva do projeto, e concluímos que não seria viável fazer o reconhecimento da mão, consequentemente da classificação do gesto em Libras em tempo real, pois exigiria alto poder de processamento e tempo.

Sendo assim, mudamos a dinâmica do jogo, na qual a mão será capturada pela câmera, onde o crop é aplicado em um curto período de tempo, salvando esta imagem e o algoritmo de extração de características faria seu papel, possibilitando mais tarde a classificação do gesto da mão, com base numa base de dados com imagens estáticas de gestos de Libras.

Em síntese, trabalhamos com python e suas bibliotecas para adaptar um algoritmo de reconhecimento da mão já existente. Buscando aplicar a classificação dos gestos mais tarde, usamos o método crop para delimitar a mão na imagem. Com o crop funcionando e salvando imagens somente da mão de forma contínua , construímos um fluxograma do funcionamento do jogo.

Resultados e Discussão

Dessa forma conseguimos implementar o método de crop no algoritmo, permitindo delimitar a mão no vídeo capturado e salvar a imagem da mão nos arquivos do projeto de maneira periódica.

Figura 1. Imagem extraída com aplicação do crop.

Isto posto, construímos um esquema geral do jogo, possibilitando termos uma ideia do funcionamento dos processos, desde a captura da mão até o resultado do usuário.

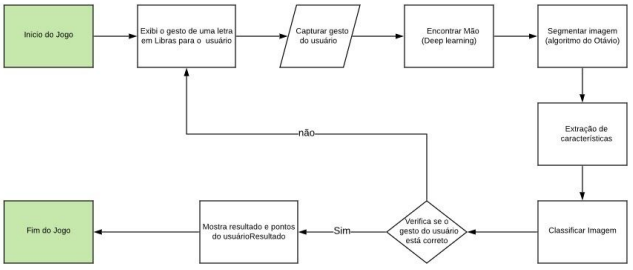

Figura 2. Esquema básico do jogo.

Considerações Finais

Em resumo, atingimos êxito na aplicação do crop no algoritmo de reconhecimento da mão, assim como também construímos um fluxograma para indicar o funcionamento do jogo. Dessa forma preparamos o projeto para invocar o algoritmo de extração de características e mais tarde, realizar a classificação dos gestos de Libras. Ademais, fica a sugestão para trabalhos futuros, ao decorrer do projeto foi visto que a Deep Learning pode ser uma tecnologia um pouco pesada e inviável pensando no público em geral.

Agradecimentos

À Instituição pelas oportunidades que proporciona.

Ao professor Diego Saqui, pelo conhecimento compartilhado, confiança e excepcional humor.

Ao meu orientador, pela paciente revisão deste trabalho e suporte.

Ao meu amigo Otávio Nathanael, pelos conselhos, horas de atenção e apoio.

Referências

NIELSEN, MICHAEL. “Neural Networks and Deep Learning”, Determination Press, 2015. Disponível em: http://neuralnetworksanddeeplearning.com/chap1.html

HUIZINGA, J. Homo ludens: o jogo como elemento da cultura. 5o. ed. [S.l.]: Perspectiva, 2003. p. 256.

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).

PRINCE, S. Computer Vision: Models Learning and Inference. [S.l.]: Cambridge University Press, 2012.

SHAPIRO, L.; STOCKMAN, G. Computer Vision. Prentice Hall, 2001. ISBN 9780130307965.

SCHUYTEMA, Paul. Design de games: uma abordagem prática. Cengage Learning, 2008.

BRASIL, Governo do Brasil, Cidadania e Justiça, Apesar de avanços, surdos ainda enfrentam barreiras de acessibilidade, 2016. Disponível em: <http://www.brasil.gov.br/cidadania-e-justica/2016/09/apesar-de-avancos-surdos-ainda-enfrentam-barreiras-de-acessibilidade>, acesso em Julho 2020.

TITLE IN ENGLISH

Abstract: (*Write the English version with the same structure using italic characters*)

Keywords: (*Write the same words in English using italic characters)*