

## AVALIAÇÃO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA NO RECONHECIMENTO DE GESTOS ESTÁTICOS DE LIBRAS UTILIZANDO O SENSOR LEAP MOTION

THAINAN B. REMBOSKI<sup>1</sup>; WILLIAM D. DE SOUZA<sup>1</sup>; MARILTON S. DE AGUIAR<sup>1</sup>

<sup>1</sup>Universidade Federal de Pelotas – {tbremboski,wddsouza,marilton}@inf.ufpel.edu.br

### 1. INTRODUÇÃO

No Brasil, segundo dados do censo realizado em 2010 pelo Instituto Brasileiro de Geografia e Estatística (IBGE), cerca de 9,7 milhões de brasileiros possuem deficiência auditiva (DA), representando 5,1% da população brasileira. Dentre eles, 1 milhão de deficientes auditivos são crianças e jovens de até 19 anos (IBGE, 2013). Estes, por sua vez, são introduzidos à língua brasileira de sinais (LIBRAS), que se torna o principal meio de comunicação entre as pessoas com deficiência auditiva e a sociedade.

Assim como as demais línguas, LIBRAS possui sua própria estrutura gramatical, que diferentemente da opinião pública, não se assemelha ao português. Ela apresenta diversos níveis linguísticos, como fonologia, morfologia, sintaxe e semântica, não sendo composta somente de sinais, tornando-se uma linguagem bastante complexa, compreendida de gestos com as mãos, expressões faciais, orientação do corpo, entre outros (TORTELLI; RUTZ; AGUIAR, 2015). Em razão dessa complexidade, mecanismos capazes de auxiliar no aprendizado da LIBRAS, que utilizem as tecnologias atuais, podem se mostrar bastante eficientes no aprendizado de jovens e crianças, reduzindo consideravelmente a complexidade da língua.

Com o surgimento de novos dispositivos e sensores 3D, como o Leap Motion (SPIEGELMOCK, 2013) tornou-se possível o desenvolvimento de novos aplicativos, voltados a diferentes áreas e tecnologias. Dentre essas diferentes áreas, destacam-se as Tecnologias Assistivas, que são recursos e serviços que visam facilitar o desenvolvimento de atividades diárias de pessoas com alguma deficiência (MELO; COSTA; SOARES, 2016).

O Leap Motion é um pequeno sensor 3D cujo foco é dado no mapeamento das mãos do usuário. Ele é composto por duas câmeras e três LEDs infravermelhos que geram um campo de interação de 150 graus em um formato de pirâmide invertida, logo acima do dispositivo. O alcance de detecção das mãos é de aproximadamente 60 centímetros acima do dispositivo, em razão da propagação da luz infravermelha, que se torna imprecisa após uma certa distância (LEAP MOTION, 2014).

Os dados gerados pelo Leap Motion são agrupados em quadros (do inglês, *frames*) que apresentam diversas informações de rastreamento, tais como: posição espacial das mãos, dedos, punho e braço; vetores de direção das mãos e dedos; tamanho da mão; vetor normal da palma da mão; entre outros. Com a disponibilidade dessas informações, o Leap Motion mostrou-se um dispositivo que apresenta um alto potencial para a captura e posterior reconhecimento de gestos de LIBRAS.

O presente trabalho visa desenvolver um aplicativo utilizando técnicas de aprendizado de máquina e o sensor Leap Motion para realizar o reconhecimento de gestos da LIBRAS, convertendo-os para texto e/ou áudio, a fim de auxiliar o aprendizado de crianças com deficiências auditivas.

## 2. METODOLOGIA

Os gestos em LIBRAS são divididos em duas categorias, gestos estáticos e gestos dinâmicos. O trabalho aqui apresentado tem como foco a identificação dos gestos estáticos, que são aqueles gestos que não levam em consideração a variação do tempo de que o gesto leva para ser realizado.

Para a captura dos gestos, cada participante realizou cinco repetições dos gestos correspondentes aos dígitos de 0 a 9, totalizando uma quantia de 300 amostras na base de dados. Por se tratar de gestos estáticos, cada gesto corresponde a um *frame* capturado pelo dispositivo. Para cada *frame*, o Leap Motion disponibiliza cerca de 183 dados, inferidos através da leitura por infravermelho da mão do usuário. Esses dados correspondem a informações como centro da palma da mão, direção que os dedos apontam, localização das articulações, entre outras, quando identificada somente uma mão acima do sensor.

Durante o procedimento de captura, todas as informações foram então armazenadas em formato JSON em um banco de dados MongoDB (CHODOROW, 2013). Isso possibilitou a criação de uma base de dados consistente para a realização dos testes deste trabalho.

Após a fase de captura, todos os gestos foram transladados para a origem e discretizados e, então, foi realizada uma análise de quais características disponibilizadas pelo Leap Motion são consideradas mais importantes para a previsão dos gestos

Esta análise foi executada através do uso da ferramenta Visual Studio da Microsoft e permitiu constatar que os dados mais relevantes para a classificação foram os correspondentes à posição da união das falanges médias com as falanges distais e proximais, junto com informações da direção que cada dedo aponta e raio do círculo gerado indicando o quão aberta a mão se encontra.

Assim, estas informações compõem o conjunto de dados utilizados para o treinamento dos algoritmos de aprendizado de máquina.

## 3. RESULTADOS E DISCUSSÃO

Para a realização dos testes, foi utilizado a API *scikit-learn*<sup>1</sup> para a linguagem de programação Python, que conta com diversos algoritmos de aprendizagem de máquina.

Os algoritmos utilizados neste trabalho foram: *Naive Bayes* (duas diferentes abordagens oferecidas pela API que utilizam os princípios estatísticos da distribuição de Bernoulli e da distribuição Gaussiana), *Decision Tree*, *K-Nearest-Neighbors* e *Support Vector Machines*, por serem considerados pela literatura os principais algoritmos de classificação com aprendizado supervisionado. Além destes, foram testados outros dois algoritmos fornecidos também pela API, *One-vs-One* e *One-vs-The-Rest*, que também suportam classificação com mais de duas classes de saída e possuem como base o mesmo algoritmo utilizado pelas SVMs.

Os dados foram divididos em dois conjuntos, um conjunto de treinamento e um conjunto de teste, como normalmente é feito em testes de algoritmos de aprendizado de máquina. A base foi dividida em 50% das amostras para

1 <http://scikit-learn.org/stable/>

treinamento e 50% para teste. A seleção de qual conjunto de dados seria destinada a amostra foi feita aleatoriamente, automatizada por uma ferramenta fornecida pela API. Os resultados obtidos para cada algoritmo podem ser observados na Tabela 1.

**Tabela 1:** Comparação entre os métodos de aprendizagem de máquina

Algoritmo	Média		
	Precisão	Recall	F <sub>1</sub>
<i>Support Vector Machine</i>	0.88	0.82	0.82
<i>Bernoulli Naive Bayes</i>	0.77	0.74	0.73
<i>Desicion Tree</i>	0.74	0.68	0.66
<i>Gaussian Naive Bayes</i>	0.79	0.66	0.61
<i>K-Nearest-Neighbours</i>	0.80	0.72	0.71
<i>One-vs-One</i>	0.87	0.82	0.82
<i>One-vs-The-Rest</i>	0.92	0.90	0.90

As métricas utilizadas por este trabalho são as informações de precisão e *recall*. Precisão indica, neste trabalho, a razão entre o número de gestos reconhecidos corretamente pelo classificador para uma dada classe X e o número total de gestos da mesma classe reconhecidos. *Recall* indica, aqui, a razão entre o número de gestos corretamente classificados da classe X e o número total de gestos que pertencem à classe X encontrados na base de dados. Por último, a métrica F<sub>1</sub> busca balancear as duas métricas anteriores utilizando uma média harmônica ponderada entre elas. Estas são métricas comumente utilizadas para avaliar algoritmos de Aprendizado de Máquina.

A partir da Tabela 1, é possível notar que o algoritmo *One-vs-The-Rest* obteve um resultado de 90% de acertos com base na métrica F<sub>1</sub>, sendo considerado o melhor resultado entre os algoritmos testados por este trabalho.

#### 4. CONCLUSÕES

Este trabalho mostrou que os dados disponibilizados pelo dispositivo Leap Motion são suficientes para a geração de um classificador e mostrou que o melhor algoritmo, entre os aqui apresentados, para o desenvolvimento deste classificador é o algoritmo *One-vs-The-Rest*, disponibilizado pela API *scikit-learn*.

Entretanto, durante o percorrer deste trabalho, o Leap Motion mostrou-se bastante instável na captura dos gestos. Assim, foi necessário realizar a captura em um ambiente controlado, onde se tinha a garantia de que o gesto observado pelo aparelho era igual ou semelhante ao gesto realizado pela pessoa ali presente. Portanto, é esperado obter resultados menos satisfatórios caso o dispositivo seja empregado em uma situação real.

Como trabalhos futuros, pretende-se avaliar os métodos escolhidos com gestos dinâmicos, começando pelo estudo dos gestos mais utilizados em LIBRAS que levem em consideração a variação do tempo na sua representação. Pretende-se também repetir a metodologia de captura proposta neste trabalho aos gestos dinâmicos feitos por voluntários.

Por fim, os autores agradecem ao CNPq (CNPq-SETEC/MEC No 17/2014, Proc. 468487/2014-0) pelo auxílio na forma de bolsa.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

CHODOROW K. **MongoDB: The Definitive Guide Second Edition**. Estados Unidos, O'Reilly Media, 2013.

IBGE. **Deficiência auditiva atinge 9,7 milhões de Brasileiros**. ADAP, 24 mai. 2013. Acessado em 9 ago. 2016. Online. Disponível em: <http://www.adap.org.br/site/index.php/artigos/20-deficiencia-auditiva-atinge-9-7-milhoes-de-brasileiros>

LEAP MOTION. **How does the Leap Motion controller work?** Leap Motion blog, São Francisco, 9 ago. 2014. Acessado em 9 ago. 2016. Online. Disponível em: <http://blog.leapmotion.com/hardware-to-software-how-does-the-leap-motion-controller-work>

MELO A.M.; COSTA J.B.; SOARES S.C.M. Tecnologias Assistivas. **Assessibilidade: discurso e prática no cotidiano das bibliotecas**. Cap. 8, pág. 62-70, 2006.

SPIEGELMOCK M. **Leap Motion Development Essentials**. Birmingham: Packt Publishing, 2013.

TORTELLI, L.M.; RUTZ, S.D.; AGUIAR, M.S. Aplicação de redes neurais convolucionais em dados capturados pelo Kinect: o estudo de caso do reconhecimento de sinais de LIBRAS. In: **XXIV CONGRESSO DE INICIAÇÃO CIENTÍFICA**, Pelotas, 2015. **Anais...** Pelotas: UFPel, 2015. p 1-4.