

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO RIO
GRANDE DO NORTE
CAMPUS NATAL - ZONA NORTE
CURSO TÉCNICO INTEGRADO EM ELETRÔNICA

MARIA ISABEL SOUSA FERNANDES
SARA SILVA MEIRELES

ALFA

APARELHO LÚDICO PARA FONOTERAPIA E ALFABETIZAÇÃO

NATAL/RN

2017

MARIA ISABEL SOUSA FERNANDES

SARA SILVA MEIRELES

ALFA

APARELHO LÚDICO PARA FONOTERAPIA E ALFABETIZAÇÃO

Trabalho de Conclusão de Curso apresentado ao Curso Técnico Integrado de Nível Médio em Eletrônica do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte campus Natal-Zona Norte, em cumprimento às exigências legais como requisito parcial à obtenção do título de Técnico em Eletrônica.

Orientador: Prof^o Dr. Diego Silveira Costa Nascimento.

NATAL/RN

2017

F363a Fernandes, Maria Isabel Sousa.
ALFA Aparelho lúdico para fonoterapia e alfabetização / Maria Isabel
Sousa Fernandes, Sara Silva Meireles. – 2017.
48 f .

Trabalho de Conclusão de Curso (Técnico em Eletrônica) – Instituto
Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte, Natal,
2017.

Orientador: Prof.º Dr. Diego Silveira Costa Nascimento.

1. Eletrônica. 2. ALFA. 3. Fonoaudiologia e Pedagogia. 4. Aprendizado
de máquina. I. Meireles, Sara Silva. II. Nascimento, Diego Silveira Costa.
III. Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do
Norte. IV. Título.

CDU 621.38:004.4

MARIA ISABEL SOUSA FERNANDES

SARA SILVA MEIRELES

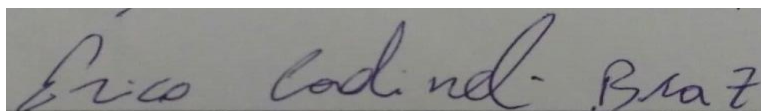
ALFA

APARELHO LÚDICO PARA FONOTERAPIA E ALFABETIZAÇÃO

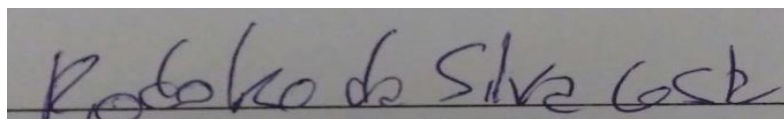
Trabalho de Conclusão de Curso apresentado ao Curso Técnico Integrado de Nível Médio em Eletrônica do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte, em cumprimento às exigências legais como requisito parcial à obtenção do título de Técnico em Eletrônica.

Trabalho de Conclusão de Curso apresentado e aprovado em 11/12/2017, pela seguinte Banca Examinadora:

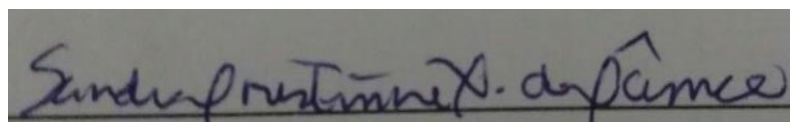
BANCA EXAMINADORA



Prof. Dr. Érico Cadineli Braz – Examinador
Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte



Prof. Esp. Rodolfo da Silva Costa – Examinador
Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte



Prof. Dra. Sandra Cristinne Xavier da Câmara
Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte

AGRADECIMENTO

Devemos reconhecer que a execução do presente trabalho, apesar de conter apenas os nomes dos autores, contou com a colaboração de terceiros, sem os quais sua realização não seria possível.

Agradecemos, especialmente, a professora Sandra Cristinne Xavier da Câmara por todo o apoio, orientação e dedicação para com o projeto. Também agradecemos a Vitor Rodrigues Greati e a Nadja Luciana pela colaboração, apoio e orientação em todo o tempo de projeto. Bem como, aos professores Jair Fernandes, Érico Cadineli Braz, Rodolfo da Silva Costa e Diego Silveira Costa Nascimento. E a toda a comunidade acadêmica do campus, em especial aos professores, técnicos de Eletrônica e à nossa turma, NPN.

Também gostaríamos de agradecer às nossas respectivas famílias e a todas as pessoas que colaboraram tanto diretamente quanto indiretamente na execução do projeto.

RESUMO

Estudos na área da Fonoaudiologia têm constatado ineficácia nos métodos utilizados no tratamento de problemas fonológicos em crianças, sendo os principais empecilhos a escassez de ferramentas e alternativas para o auxílio dos fonoterapeutas, educadores e pais envolvidos com crianças que apresentam problemas dessa natureza. A partir disso, o projeto Aparelho Lúdico para Fonoterapia e Alfabetização (ALFA), compreendido entre as áreas de Eletrônica, Informática, Fonoaudiologia e Pedagogia; tem como principal objetivo auxiliar os profissionais da área de fonoaudiologia no tratamento de pessoas, em especial crianças, que apresentam tais alterações, ou seja, dificuldade no modo de articulação e pronúncia dos sons da fala na Língua Portuguesa, como também aquelas que estão em processo de reabilitação pós implante coclear. O projeto propõe um sistema que consiste em um dispositivo de entrada de áudio cujas captações compõem um vetor de descritores que é passado para um algoritmo inteligente responsável por indicar a corretude do fonema. O ALFA consta de quatro etapas: pesquisas sobre os aspectos de inteligência do sistema; captação de dados para o treinamento, validação e testes de aprendizado de máquina; construção de um microfone de baixo custo; e, finalmente, construção da rede propriamente dita. Paralelamente, implementa-se a interface lúdica do sistema. Os resultados alcançados com a realização das três primeiras etapas foram positivos em relação à aceitação do público-alvo, os últimos testes com o algoritmo de classificação KNN apresentaram uma taxa de acerto por sílaba entre 0% e 60%, e obteve-se êxito tanto na construção do circuito de captação de áudio de baixo custo, quanto nos seus testes em laboratório. A partir dos resultados alcançados ao longo do desenvolvimento do projeto, pretende-se realizar novos testes para definir a melhor técnica de aprendizagem de máquina, assim como implementar a interface lúdica no sistema. Desse modo, será possível testar o sistema concebido com ao menos um paciente em terapia fonoaudiológica e em uma turma de alfabetização, comprovando sua eficácia.

Palavras-chave: Eletrônica. ALFA. Fonoaudiologia e Pedagogia. Aprendizado de máquina.

ABSTRACT

Studies in the area of Speech-Language Pathology have found ineffectiveness in the methods used in the treatment of phonological problems in children. The main obstacles are the shortage of tools and alternatives for the help of speech therapists, educators and parents involved with children who present problems of this nature. From this, the project Linguistic Apparatus for Speech Therapy and Literacy (ALFA), comprised between the areas of Electronics, Informatics, Speech-Language Pathology and Pedagogy; has as main objective to assist professionals in the field of speech therapy in the treatment of people, especially children, who present such changes, ie difficulty in articulation and pronunciation of speech sounds in the Portuguese language, as well as those that are in the process of rehabilitation after cochlear implantation. The project proposes a system consisting of an audio input device whose abstractions make up a vector of descriptors that is passed to an intelligent algorithm responsible for indicating the correctness of the phoneme. The ALFA consists of four steps: research on the intelligence aspects of the system; data capture for training, validation and machine learning tests; building a low-cost microphone; and finally, the construction of the network itself. At the same time, the ludic interface of the system is implemented. The results obtained with the accomplishment of the first three stages were positive in relation to the acceptance of the target public, the last tests with the KNN classification algorithm presented a syllable accuracy rate between 0% and 60%, and both success in the construction of the circuit of capture of audio of low cost, as in its tests in laboratory. From the results achieved during the project development, we intend to perform new tests to define the best machine learning technique, as well as to implement the ludic interface in the system. In this way, it will be possible to test the system designed with at least one patient in speech therapy and in a literacy class, proving its effectiveness.

Keywords: Eletronics. ALFA. Speech Therapy and Pedagogy. Machine learning.

LISTA DE ILUSTRAÇÕES

Foto 1 - Sessão de fonoterapia com o álbum fonológico de Yavas	17
Foto 2 - Detalhe do álbum fonológico	17
Figura 1 - Exemplificação do conjunto de neurônios conectados em camadas	20
Figura 2 – Representação interna do microfone de eletreto	21
Figura 3 - Diagrama do microfone de eletreto	21
Figura 4 – Comparação entre o comportamento ideal e real de um filtro PF	23
Foto 3 – Momento de recepção na ação Fale com a minha mão	26
Foto 4 – Primeira coleta de vozes do ALFA	26
Foto 5 – Segunda coleta de vozes do ALFA	26
Figura 5 - Esquemático da arquitetura de funcionamento do projeto	28
Quadro 1 – Parâmetros usados para extrair a matriz de características	29
Tabela 1 – Sílabas e quantidade se acertos em porcentagem (%)	30
Figura 6 - Simulação do circuito no software <i>Protheus</i>	31
Figura 7 - Simulação da entrada de áudio no Protheus 8 (sílabas MAR)	31
Figura 8 - Layout do circuito de captação de áudio no ARES	32
Figura 9 - Visualização 3D do circuito de captação de áudio no ARES	32
Figura 10 - Imagem frontal do circuito montado em protoboard	33
Figura 11 - Perfil do circuito montado em protoboard	33
Figura 12 - Placa de circuito impresso do microfone	34
Figura 13 – Teste com o fonema "PÉ"	35
Figura 14 - Teste do microfone no software Audacity	35
Figura 15 - Jogo do labirinto	36
Figura 16 – Esboço do jogo do trem a ser implementado na interface lúdica	37
Figura 17 – Diagrama	37
Gráfico 1 – Pesquisa sobre a utilidade do ALFA para os fonoaudiólogos	39
Gráfico 2 – Pesquisa sobre a utilidade do ALFA para os pedagogos	40

LISTA DE ABREVIATURAS E SIGLAS

ALFA Aparelho Lúdico para Fonoterapia e Alfabetização

dB Decibel

IBGE Instituto Brasileiro de Geografia e Estatística

KNN K-Nearest Neighbors

MFCC Mel-Frequency Cepstral Coefficients

MOSFET Metal Oxide Semiconductor Field Effect Transistor

RNA Rede Neural Artificial

UFRJ Universidade Federal do Rio de Janeiro

Unesp Universidade Estadual Paulista

SUMÁRIO

1	INTRODUÇÃO	10
1.2	JUSTIFICATIVA	11
1.3	HIPÓTESES	12
1.4	OBJETIVOS	13
1.4.1	Geral	13
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	ÂMBITO DA PEDAGOGIA	15
2.2	ÂMBITO DA FONOAUDIOLOGIA	16
2.3	ÂMBITO DA INFORMÁTICA	18
2.4	ÂMBITO DA ELETRÔNICA	20
2.4.1	Microfone de Eletreto	20
2.1.4	Filtro Passa-Faixa Ativo	22
3	METODOLOGIA	24
3.1	PESQUISAS	24
3.2	COLETAS DE VOZES	25
3.4	BANCO DE DADOS E EXTRAÇÃO DE DESCRITORES	29
3.5	TREINAMENTO DO ALGORITMO DE APRENDIZAGEM DE MÁQUINA	30
3.7	DESENVOLVIMENTO DA INTERFACE LÚDICA	36
4	RESULTADOS E DISCUSSÕES	39
5	CONSIDERAÇÕES FINAIS	41
	REFERÊNCIAS	42
	APÊNDICE A	46

1 INTRODUÇÃO

Atualmente, o uso de tecnologias vem sendo implementado cada vez mais no processo de aprendizado das crianças, tanto no âmbito escolar quanto no fonoaudiológico. Com isso, percebe-se que artifícios tecnológicos vêm se mostrando importantes aliados no processo da fala, ao proporcionar grandes avanços na vida de seus usuários. Como exemplo, tem-se os aplicativos educativos do mercado norte-americano desenvolvidos pela empresa *Smarty Ears*, criada pela brasileira Bárbara Fernandes, que trabalham o desenvolvimento da linguagem, da articulação e da gagueira em crianças. Contudo, segundo Fábio Henrique Pinheiro, fonoaudiólogo ligado à Unesp, no Brasil, a tecnologia na fonoaudiologia ainda é muito incipiente (GOMES, 2012), opinião corroborada durante o desenvolvimento do projeto, por meio da aplicação de um questionário online, conforme resultados apresentados em seção específica.

Embora existam tecnologias voltadas para essa área, ainda há muito o que se fazer pela melhoria da terapia e sessões de apoio pedagógico para ajudar as crianças que apresentam problemas de dicção. Um dos principais empecilhos nessas áreas é a escassez de ferramentas e alternativas para o auxílio dos fonoterapeutas, educadores e pais envolvidos com crianças que apresentam problemas dessa natureza.

A utilização de sistemas informatizados em terapia, como programas computacionais, jogos, entre outros, auxiliam os profissionais de maneira mais eficiente e geram resultados mais satisfatórios em menor espaço de tempo, tendo em vista o teor lúdico dessa tecnologia proporcionar um maior interesse e melhor desempenho da criança. A Informática Educativa nos oferece uma vastidão de recursos que, se bem aproveitados, nos dão suporte para o desenvolvimento de diversas atividades com os alunos (ROCHA, 2008).

A partir disso, este trabalho buscou o desenvolvimento de um Aparelho Lúdico para Fonoterapia e Alfabetização (ALFA), um sistema de baixo custo e acessível, fundamentado em reconhecimento de padrões com interface lúdica e gamificada, para detecção automática de variações no processo fonológico, com aplicabilidade em clínicas de fonoaudiologia e no apoio pedagógico em escolas, destinado ao uso com crianças com alterações no processo fonológico.

1.1. PROBLEMÁTICA

Em nosso país, cerca de 3 a 10% da população brasileira segundo pesquisas da fonoaudióloga, doutora em Linguística e professora Adjunta da UFRJ Renata Mousinho (2008), apresentam Distúrbio Específico de Linguagem (DEL). O distúrbio fonológico apresenta grande ocorrência na população infantil (GIERUT, 1998). Um dos principais empecilhos para ajudar os fonoterapeutas, educadores e pais na educação e no tratamento das dificuldades dessas crianças é a escassez de ferramentas alternativas. Vale ressaltar que os métodos mais utilizados para o tratamento de problemas fonológicos em crianças apresentam forte caráter manual, dificultando o trabalho dos profissionais, além de serem pouco atraentes para o público infantil.

Tanto na área da educação quanto na área da fonoaudiologia, mais especificamente no âmbito fonoterapêutico, embora existam novas tecnologias, a maioria não se preocupa em aplicar técnicas lúdicas e interativas em suas respectivas sessões; além disso, não há garantia, a partir dos métodos usuais, de que o som que a criança está emitindo esteja correto, já que não existem ferramentas que detectem como a criança está se exercitando. Ademais, o processo de detecção dos distúrbios se dá através de um método extenso e enfadonho para os profissionais da área.

Com isso, mostra-se promissor o desenvolvimento de uma tecnologia que ofereça mais precisão, facilitando, desta forma, o trabalho do profissional da área específica. Para tanto, o sistema ALFA, aqui proposto, visa auxiliar tanto o meio pedagógico quanto no tratamento de fonoterapia de maneira lúdica e gamificada.

1.2 JUSTIFICATIVA

A partir da vivência de uma das componentes da equipe em uma turma de alfabetização da rede pública e de sua participação em trabalhos voluntários em ONGs de apoio à comunidade, foi observada a presença de alterações nas falas das crianças. Durante esse período, notou-se que tais alterações eram constantes em algumas delas e não apenas algo relacionado à idade de cada uma. A partir de

então, surgiu um interesse e vontade de ajudá-las — visto que já havia certo interesse por pesquisas desenvolvidas nesse âmbito, ou seja, direcionadas à melhoria (e/ou reabilitação) da dicção de crianças que haviam feito implante coclear ou que faziam uso recente de aparelho auditivo.

Desse modo, o aprofundamento das pesquisas e a formulação deste projeto demonstraram a necessidade de colaboradores de diferentes áreas, como Fonoaudiologia, Pedagogia, Informática e Eletrônica, caracterizando este estudo como multidisciplinar.

Com essa visão multidisciplinar, notou-se que a confusão na pronúncia dos fonemas mostra-se recorrente em processos de alfabetização e fonoterapia. A identificação desse problema é geralmente realizada de forma subjetiva (em processos manuais que não evidenciam resultados precisos) pelo professor ou terapeuta. Além disso, os métodos de ensino são pouco atraentes para as crianças, representantes de grande parcela do público atendido, dificultando ainda mais o processo de aprendizagem. Há, portanto, a necessidade de uma nova abordagem para o ensino e aprendizagem da fala, seja em escolas, ou em clínicas de fonoaudiologia.

É nesse conjunto de fatos que este projeto se justifica, visto que propõe um sistema para a detecção automática da alteração no processo fonológico e, a partir disso, a indicação da forma de correção para o usuário, de maneira lúdica, a fim de tornar o tratamento mais atrativo e prazeroso. Além disso, a solução proposta é de baixo custo, o que facilita sua aquisição, tanto pelos profissionais envolvidos quanto pelos pacientes, que poderão dar continuidade ao tratamento em suas próprias residências, acelerando a aprendizagem.

1.3 HIPÓTESES

Diante da problemática que foi apresentada, a ideia inicial do projeto foi discutida e aprimorada, direcionando o projeto para o objetivo de tornar o processo de identificação de alterações fonológicas mais interativo e lúdico, por meio de um software, a fim de reduzir o esforço dos profissionais de fonoaudiologia e acelerar o processo de alfabetização e tratamentos fonoaudiológicos. Ademais, o desenvolvimento de um software com uma interface lúdica e gamificada apresenta-

se como um fator motivador para o público infantil, potencializando sua participação no processo e, conseqüentemente, acelerando o aperfeiçoamento da fala e da língua.

A otimização do processo de tratamento fonoaudiológico, principalmente no que diz respeito ao tempo, legitima o ALFA como uma alternativa de solução para os especialistas envolvidos nas terapias indicadas. Vale ressaltar também o expressivo banco de dados construído por meio desta pesquisa, contendo numerosos exemplos de pronúncias e a indicação do fonema correspondente a cada gravação. Tal banco de dados poderá servir para futuras aplicações técnicas de aprendizado de máquina e/ou outros projetos na área. Finalmente, a viabilidade econômica do projeto mostra-se como fator decisivo para a aceitação do público envolvido.

1.4 OBJETIVOS

A seguir, serão descritos o objetivo geral e os específicos para o desenvolvimento e a implementação do ALFA.

1.4.1 Geral

Desenvolver o ALFA, um sistema de baixo custo e acessível, fundamentado em reconhecimento de padrões com interface lúdica e gamificada, para detecção automática de alterações no processo fonológico, com aplicabilidade em clínicas de fonoaudiologia e no apoio pedagógico em escolas, para crianças que apresentam tais alterações.

1.4.2 Específicos

- Compor um banco de dados de pronúncias de fonemas para, futuramente, viabilizar a criação de novas aplicações nesse mesmo âmbito.
- Compor um circuito eletrônico para captação de áudio utilizando componentes de baixo custo;
- Definir a melhor técnica de aprendizagem de máquina para identificar alterações no processo fonológico.

- Desenvolver uma interface lúdica e gamificada para motivar os usuários a aperfeiçoar cada vez mais a pronúncia dos fonemas.
- Aplicar o ALFA na otimização das sessões de fonoterapia, tornando o processo mais ágil e menos artesanal, e como apoio pedagógico em escolas e formular pedido de patente para o produto criado.

2 FUNDAMENTAÇÃO TEÓRICA

O ALFA configura-se como um projeto multidisciplinar, compreendido entre as áreas: da Pedagogia, no que diz respeito ao ciclo de alfabetização; da Fonoaudiologia, com estudos em específico acerca da linguagem e da fonoterapia; da Informática, abarcando a linguagem de programação Python com algoritmos de aprendizagem de máquina; e da Eletrônica com estudo em circuitos de captação de áudio. Nesse sentido, a fundamentação teórica foi organizada a partir da discussão de conceitos dessas grandes áreas necessários ao desenvolvimento desta pesquisa.

2.1 ÂMBITO DA PEDAGOGIA

O método fônico é um método de alfabetização que primeiro ensina os sons de cada letra e então constrói a mistura destes sons em conjunto para alcançar a pronúncia completa da palavra, permitindo, dessa forma, que se consiga ler toda e qualquer palavra. Capovilla (2010) afirma que esse é um método lúdico, inteligente e nada mecânico.

A implementação desse método, nas salas de aula e em sessões de apoio pedagógico, ocorre geralmente pela apresentação gradual da escrita das letras e de sua pronúncia, com os alunos sendo estimulados a repetirem e a associarem o som às suas várias representações gráficas (CAPOVILLA, 2010). Dada a natureza lúdica do método fônico, um sistema como o proposto neste trabalho almeja complementar as atividades desenvolvidas, provendo um meio eletrônico, em forma de jogo, repleto de elementos de gamificação, para tornar a experiência do aprendizado mais atrativa e eficaz.

Hoje, práticas de sentido lúdico se resumem a jogos de tabuleiro, jogos de memória, jogos com figuras e jogos de escuta — que estimulam a habilidade das crianças de prestarem atenção aos sons de forma seletiva — representando meios bastante manuais, apesar de eficazes. Ocorre que, diante do contexto tecnológico em que se vive, é possível melhorar ainda mais aproveitando-se das tecnologias de reconhecimento de fala e interfaces interativas.

Nos contextos supracitados, jogos ganham um espaço como ferramenta efetiva da aprendizagem, na medida em que propõem o estímulo ao interesse do

aluno, por desenvolver os diferentes níveis de sua experiência pessoal e social. O jogo ajuda-o a construir suas novas descobertas, desenvolve e enriquece sua personalidade e simboliza um instrumento pedagógico que leva o professor à condição de condutor, estimulador e avaliador da aprendizagem. Por meio do jogo, o aluno desenvolve e exercita sua memória, seu raciocínio, sua capacidade de percepção, a criatividade e a sua autonomia (CAMPOS, 2008).

Destaca-se, além disso, que o sistema se alinha a uma tendência recente na educação: a gamificação, que consiste na utilização de elementos dos jogos fora do seu contexto, com a finalidade de mobilizar os sujeitos à ação, auxiliar na solução de problemas e promover aprendizagens (FARDO, 2013 apud. KAPP, 2012). Tal fenômeno vem se inserindo cada vez mais no âmbito escolar, auxiliando na aprendizagem tanto infantil quanto adulta, por meio dos sistemas e *softwares* educativos interativos.

Os principais objetivos da gamificação nas escolas são incentivar o protagonismo, por meio da interação entre a criança e o conteúdo abordado; desenvolver habilidades, despertando talentos naturais; e promover práticas colaborativas na aprendizagem. Visto que a intenção é de promover engajamento e motivação, esse mesmo fenômeno pode se estender à área da terapia (NASCIMENTO; ALBUQUERQUE, 2015).

2.2 ÂMBITO DA FONOAUDIOLOGIA

Além do contexto da alfabetização por método fônico, as clínicas de fonoaudiologia podem se beneficiar consideravelmente das novas tecnologias digitais. Isso porque, nos atendimentos fonoaudiológicos, observam-se situações de não colaboração dos pacientes infantis. Os métodos mais utilizados para o tratamento de problemas fonológicos em crianças têm forte caráter manual, dificultando o trabalho dos profissionais, além de serem pouco atraentes para o público infantil.

Nas sessões de terapia com as crianças para avaliação da linguagem oral, é utilizado um álbum de figuras Yavas (Avaliação fonológica da criança: reeducação e terapia), como mostrado nas Fotos 1 e 2, o qual não prende completamente a

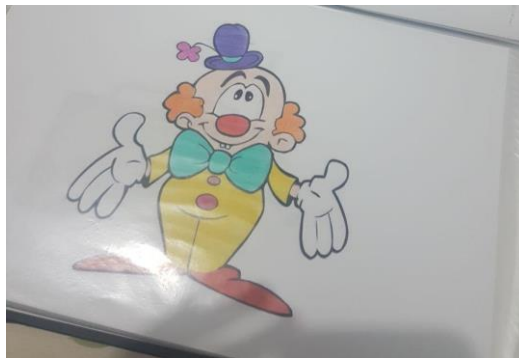
atenção das crianças, tendo em vista que “além de ser um processo manual, a cada figura o fonoaudiólogo precisa de uma pequena pausa para anotações”, como nos explica Nadja Luciana, fonoaudióloga colaboradora do projeto ALFA. Em muitos casos, não se consegue completar procedimentos de avaliação, e os procedimentos de intervenção podem ser prejudicados (NASCIMENTO et al., 2011).

Foto 1 – Sessão de fonoterapia com o álbum fonológico de Yavas.



Fonte: Acervo pessoal de Nadja Luciana

Foto 2 – Detalhe do álbum fonológico



Fonte: Acervo pessoal de Nadja Luciana

Muitas vezes, os pacientes podem necessitar de um acompanhamento contínuo, e a infraestrutura e os equipamentos de hoje não dão suporte para isso, sendo, portanto, fundamental ultrapassar o senso comum em busca de soluções baseadas em novas tecnologias para auxiliar nos tratamentos de problema de fala (CAPOVILLA, 2010). Essa falta de atratividade dos meios atuais, os quais não contribuem para a colaboração espontânea dos pacientes, bem como a falta de suporte a casos especiais, podem ser amenizadas com o uso de tecnologias lúdicas e que dinamizem o trabalho do profissional, como é o caso do sistema proposto.

Segundo a fonoaudióloga Nadja Luciana, o projeto é importante “pois otimiza e assegura nossa avaliação no quesito identificação das alterações fonológicas, e como estratégia terapêutica por ser algo visualmente atrativo para a criança”, e finaliza acrescentando que “a maior dificuldade do processo manual, que consiste em gravar a pronúncia da criança com a nomeação das imagens do álbum, fazer a transcrição dos fonemas e identificar as alterações e processos, é o tempo que isso demanda”. Principalmente nesse sentido é que considera a relevância do projeto ALFA, em desenvolvimento.

2.3 ÂMBITO DA INFORMÁTICA

A tecnologia envolvida na implementação desse sistema tem como essência técnicas de aprendizagem de máquina capazes de prover uma comparação entre uma pronúncia recebida e a pronúncia desejada, detectando quando há erros. A tarefa de reconhecimento automático de fala através de aprendizagem de máquina é alvo de muitas pesquisas recentes (DENG; LI, 2013; WITT, 2012). Mais especificamente, a detecção de problemas de pronúncia também recebe muita atenção, como no trabalho de Bang *et al.* (2014), o qual apresentou métodos e resultados positivos quanto à detecção de erros de pronúncia de palavras inglesas por coreanos.

Além dele, Qian *et al.* (2015), utilizando redes neurais profundas, conseguiu bons resultados na detecção da corretude de pronúncia, ultrapassando, inclusive, resultados da literatura com classificadores baseados em máquinas de vetores de suporte (MVS). Ademais, Wei *et al.* (2009) apresenta um trabalho utilizando MVS para propósito semelhante, obtendo bons resultados e boa aceitação.

Na classificação por aprendizado de máquinas clássico, os dados devem estar representados por conjuntos de n atributos, chamados geralmente de descritores. Sob essa perspectiva, os dados nada mais são do que pontos num espaço de características n -dimensional. A escolha dos melhores atributos é tão melhor quanto mais separadas estiverem as classes de interesse na distribuição dos pontos nesse espaço.

O método de extração de descritores mais comumente usado no reconhecimento automático de fala (RAF) é o *Mel-Frequency Cepstral Coefficients* (MFCC). Para extrair um vetor de características contendo todas as informações sobre a mensagem linguística, o MFCC imita algumas partes da produção de fala humana e da percepção logarítmica da intensidade e do tom do sistema auditivo humano e tenta eliminar as características dependentes dos falantes excluindo a frequência fundamental e seus harmônicos (LUTTER, 2014). Tendo a matriz de dados, com n colunas e m linhas, onde m corresponde à quantidade de dados disponível, é possível aplicar algoritmos de reconhecimento de máquinas para realizar a tarefa de classificação desejada. Entre esses algoritmos, os mais comuns são o *K-Nearest Neighbors* (KNN) e as Redes Neurais Artificiais (RNAs).

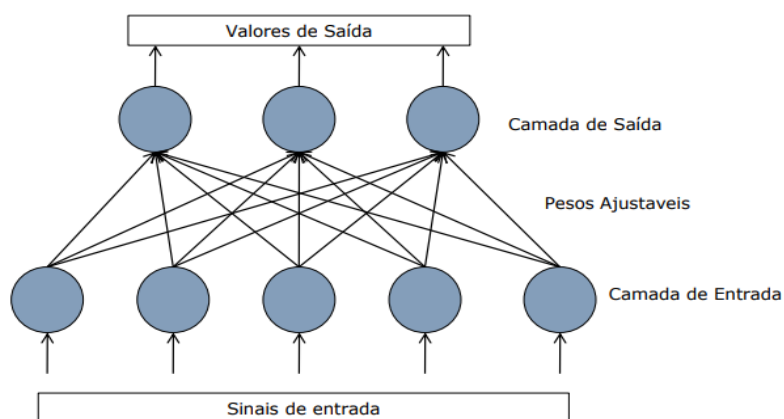
O KNN, um dos algoritmos de classificação com técnica simples e facilmente implementável, é aplicado para classificar objetos com base em exemplos de treinamento que estão mais próximos no espaço de características. Para a sua utilização é necessário um conjunto de exemplos, definir uma métrica para calcular a distância entre os exemplos e definir o valor de K (o número de vizinhos mais próximos que serão considerados pelo algoritmo) (SOARES, 2012).

As RNAs, por sua vez, iniciaram-se com o conceito simples chamado *PERCEPTRON*, um tipo de RNA (rede neural artificial) desenvolvido nos anos de 1950. Para sua execução, usa-se um conjunto de exemplos de treinamento que dão a saída desejada para uma unidade, com o objetivo de aprender pesos sinápticos de tal forma que a unidade de saída produza a resposta correta para cada exemplo. Assim, sabendo que o funcionamento é equivalente com o do cérebro humano, este processo é repetido até que um neurônio de saída seja ativado, ou seja, até chegar aos pesos corretos.

Outro tipo mais complexo comumente utilizado é o *Multilayer Perceptron* (MLP), que simula um conjunto de neurônios conectados em camadas, sendo elas a de entrada, a de saída e, possivelmente, as escondidas ou intermediárias, como ilustrado na Figura 1 a seguir, que apresenta como exemplo a camada de entrada por meio do conjunto de esferas inferiores e como camada de saída o conjunto de esferas superiores, estando ocultas as possíveis camadas intermediárias. O princípio é prover um conjunto de treino, composto de instâncias já classificadas, e reduzir o erro de classificação por meio do ajuste de parâmetros numéricos

chamados de pesos da rede neural. Uma vez treinada, a rede pode receber novos dados e classificá-los com certa garantia de acerto.

Figura 1 – Exemplificação do conjunto de neurônios conectados em camadas.



Fonte: SOARES, Edilei¹

Vê-se, com isso, que o sistema proposto para o desenvolvimento do ALFA está fundamentado em tendências recentes da tecnologia e é exequível.

2.4 ÂMBITO DA ELETRÔNICA

2.4.1 Microfone de Eletreto

Conceitua-se microfone como um transdutor eletroacústico, ou seja, um dispositivo que converte energia acústica em energia elétrica, consistindo, portanto, em um elemento de interface do homem com circuitos eletrônicos. Assim, ao longo do tempo, diversos tipos de microfones evoluíram e logo apresentaram características diferentes, dentre eles está o microfone de eletreto, que contém como elemento sensível um material denominado eletreto.

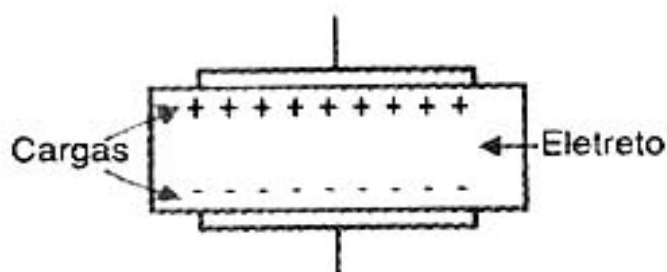
O nome eletreto vem do carregamento constante da fina folha metalizada em relação à terra da carcaça do microfone. Ao entrar, a onda mecânica (som) provoca

¹ Figura disponível no material de aula do Professor Edilei Soares

uma vibração na folha metalizada, fazendo com que as cargas elétricas permanentes do microfone de eletreto se alterem cada vez que sofrem modificações mecânicas provocadas pela perturbação do som. Com a vibração na folha metalizada, a indução elétrica do capacitor, composto pelo eletreto e por uma membrana, se altera, assim resultando num sinal, como mostrado na Figura 2. Então, o transistor MOSFET recebe o sinal elétrico e o amplifica, então o sinal é enviado pelo dreno do transistor que está ligado ao pino de saída do microfone de eletreto, assim como representado na Figura 3.

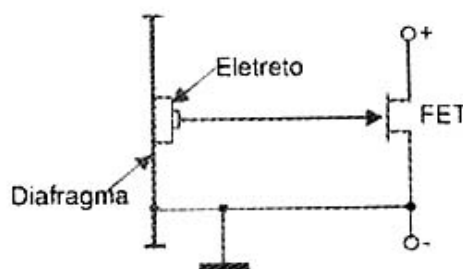
As vantagens destes microfones estão em seu baixo custo, tamanho reduzido e grande sensibilidade, o que os torna ideais em aplicações práticas contemporâneas.

Figura 2 – Representação interna do microfone de eletreto



Fonte: INSTITUTO NEWTON BRAGA²

Figura 3 – Diagrama do microfone de eletreto



Fonte: INSTITUTO NEWTON BRAGA³

² Disponível em: <<http://www.newtonbraga.com.br/index.php/como-funciona/4446-art616>>.

³ Disponível em: <<http://www.newtonbraga.com.br/index.php/como-funciona/4446-art616>>.

2.1.4 Filtro Passa-Faixa Ativo

Os filtros são circuitos eletrônicos que permitem o controle sobre quais valores de frequência irão ser efetivamente utilizados nos circuitos. Eles são constituídos, fundamentalmente, de capacitores, de indutores e de resistores, e podem também ser apresentados com uma gama de variações que permitem uma maior precisão dos valores de frequência admitidos pelo filtro. O ponto crítico, frequência de corte, em que esse circuito faz com que não haja sinal de saída relevante é reconhecível através da fórmula descrita abaixo:

$$V_{fc} = V_{in} * 0,707$$

Na qual V_{fc} representa a tensão de frequência de corte e V_{in} , a tensão do sinal de entrada. Assim, a tensão da frequência de corte sempre será 0,707 da tensão original, ou seja, há uma perda de 3dB, que significa que a potência do sinal caiu para metade. Sendo assim, podemos dizer que o sinal está “cortado”.

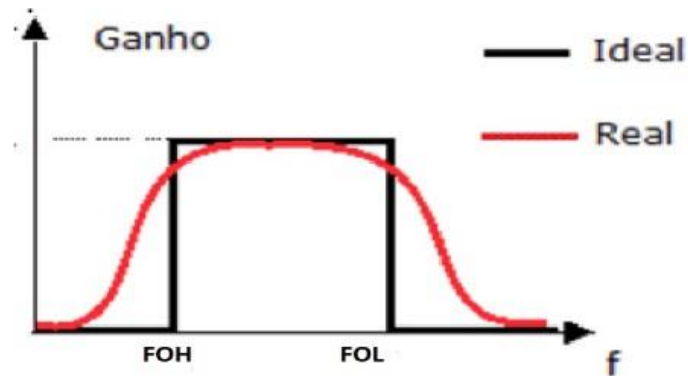
Dentre os modelos existentes de filtros, no presente relatório será abordado o filtro passa-faixa ativo, circuito o qual determina uma faixa de frequência funcional que está entre dois valores de frequência de corte, sinais acima de uma frequência de corte f_{oL} (*Frequency Out Low*) e abaixo de uma segunda frequência de corte f_{oH} (*Frequency Out High*), como mostrado na Figura 4, com um ganho de tensão constante ideal, representado por A_v , calculados através das fórmulas a seguir:

$$F_{oL} = 1 / 2\pi R_1 * C_1$$

$$F_{oH} = 1 / 2\pi R_1 * C_1$$

$$A_v = 1 + R_f / R_g$$

Figura 4 – Comparação entre o comportamento ideal e real de um filtro PF



Fonte: FERNANDES (2017)⁴

Analisando a Figura 4 acima, na qual a linha de cor preta representa o comportamento ideal de um filtro PF, enquanto que a linha de cor vermelha o comportamento real, mostrando que o circuito não consegue cortar instantaneamente a frequência de corte calculada, assim este corte sempre sofrendo variações para mais ou menos no circuito real.

⁴ Disponível no slide de aula circuitos para instrumentação do Prof. Dr. Jair Fernandes.

3 METODOLOGIA

O desenvolvimento do ALFA compreende as seguintes etapas: pesquisas em Pedagogia e Fonoaudiologia, Eletrônica, como também em Informática, sobre aspectos de inteligência do sistema; coleta de vozes; arquitetura do projeto; construção do banco de dados e extração de descritores; treinamento do algoritmo do aprendizado de máquina e, enfim, o desenvolvimento da interface lúdica.

Tendo em vista a percepção da viabilidade do ALFA junto ao público-alvo, aplicou um questionário online entre os especialistas envolvidos, a saber: fonoaudiólogos e pedagogos.

3.1 PESQUISAS

Em princípio, buscou-se um entendimento sobre as áreas que o ALFA englobava, para então se dar início às pesquisas nos núcleos de fonoaudiologia, estudando sobre o método atual de reconhecimento e identificação das alterações fonológicas, e de pedagogia, a partir de visitas a escolas da região. O objetivo foi conhecer melhor o universo em estudo e depreender quais as maiores dificuldades de dicção e alfabetização que as crianças apresentavam, por meio de levantamento de dados e depoimentos dos profissionais envolvidos.

Tendo este aporte teórico, direcionaram-se as pesquisas para o meio científico tecnológico, com estudos sobre a linguagem *Python* e algoritmos de aprendizagem de máquina específicos, como Redes Neurais Artificiais (RNAs) e *K-Nearest Neighbors* (KNN). Por último, foi preciso buscar um método de extração de descritores. Após a leitura de alguns trabalhos, chegou-se ao método *Mel-Frequency Cepstral Coefficients* (MFCC), o qual proporciona a extração de um vetor de características para representar padrões das pronúncias das sílabas.

Seguindo com as pesquisas, selecionaram-se os fonemas para gravação, com a ajuda da fonoaudióloga Nadja Luciana Costa e com base no álbum fonológico “Avaliação Fonológica da Criança”, dos autores Yavas, Hernandorena e Lamprecht, instrumento com desenhos temáticos representativos do sistema da língua portuguesa; e da tabela de aquisição de fonemas fornecida, também, pela fonoaudióloga colaboradora do projeto.

3.2 COLETAS DE VOZES

Com os fonemas selecionados, deu-se início à primeira coleta de áudios para construção do banco de dados de vozes. Para realizá-la, foi necessária antes uma conversa com os diretores de duas escolas próximas ao Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte Campus Natal-Zona Norte, a Escola Municipal professora Palmira de Souza e o Operacional Colégio e Curso, para liberação das turmas de alfabetização para a gravação de suas vozes. Essa ação do projeto foi denominada “Fale com a minha mão”, a seguir representada nas Foto 3.

Já no Instituto, a coleta procedeu com as crianças sendo direcionadas (em grupos de três) para uma sala com isolamento acústico onde uma das integrantes do projeto, com o colaborador Vitor Greati⁵, pedia para cada criança repetir uma sequência das sílabas que foram selecionadas no processo de pesquisa, como mostrado na Foto 4. Com as vozes coletadas, trabalhou-se cada gravação, separando as sílabas e armazenando-as em pastas. Logo após, com os áudios já separados, iniciou-se a construção e o teste de um código prototípico com a implementação do *k-nearest neighbors* (KNN), com o intuito de obter um vetor de características extraídas dos áudios já gravados e testar uma classificação com poucas instâncias.

No entanto, com o objetivo de inteirar o banco de dados, que já continha cerca de 2.300 áudios, foram convidados alunos e servidores do instituto para mais uma coleta de vozes, pois pretendia-se alcançar, no mínimo, quarenta pronúncias de cada sílaba até a conclusão do projeto. A segunda ação para a coleta de vozes, como apresentada na Foto 5 a seguir durou três dias e, ao final, foram obtidas cerca de sete mil e oitocentas gravações, sendo em média quarenta áudios para cada sílaba, que, quando selecionados e somados aos áudios já existentes, resultaram em cerca de dez mil e cem áudios.

⁵ Ex-aluno do IFRN-Campus Natal-Zona Norte e bacharelado do curso de Tecnologia da Informação na Universidade Federal do Rio Grande do Norte.

Foto 3 – Momento de recepção na ação fale com a minha mão



Fonte: ANDRADE (2017)⁶

Foto 4 – Primeira coleta de vozes do ALFA



Fonte: Das autoras (2017)

Foto 5 – Segunda coleta de vozes do ALFA



Fonte: IFRN em Pauta⁷

⁶ Foto disponibilizada pelo Professor Everaldo Andrade que compareceu ao Fale com a minha mão.

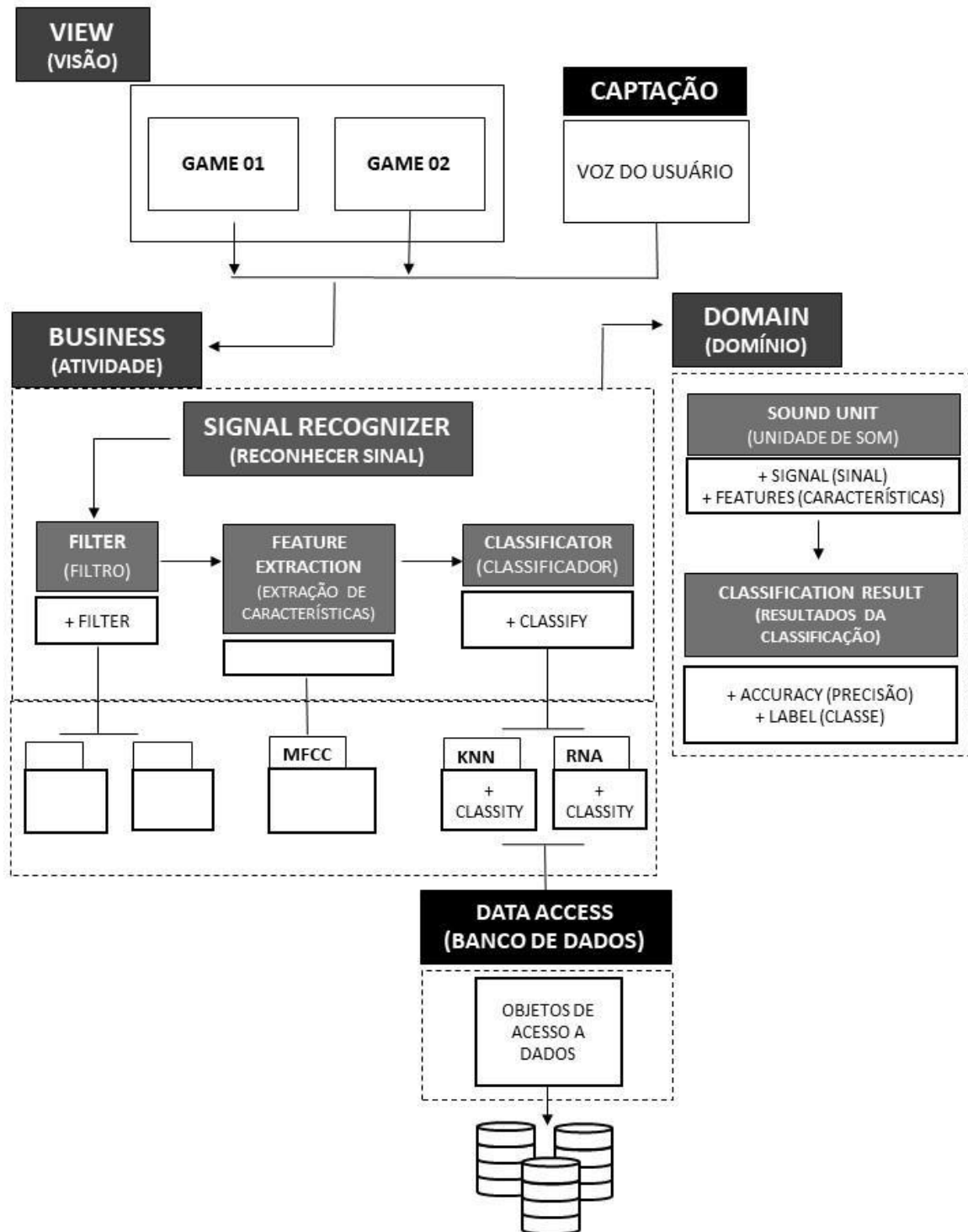
⁷ Disponível em: <<https://www.youtube.com/watch?v=Ndp9Ufy0yuk&feature=youtu.be>>.

3.3 ARQUITETURA DO PROJETO

Paralelamente às pesquisas, iniciou-se a organização do sistema, desenvolvido a partir do paradigma de orientação a objetos, cuja finalidade é representar objetos do mundo real e seus comportamentos por meio de classes, com métodos e atributos. Um conjunto de objetos com os mesmos métodos e atributos é uma classe. Por exemplo, cada sílaba que se quer representar foi denominada de unidade de som, como “ba”, “be”, “bi”. O “ba”, por exemplo, possui os atributos sinal e label (classe), os quais definem exatamente que o “ba” é um “ba”, e assim por diante. O mesmo ocorre para os outros. Assim, embora cada som possua valores diferentes, os atributos definem que eles fazem parte da mesma classe.

A fim de modularizar a arquitetura do projeto, utilizou-se um modelo em camadas. Inicialmente, tem-se a camada de visão, com a qual os usuários finais irão interagir, ou seja, o ambiente lúdico com os jogos. Cada jogo acessa a camada de aprendizagem de máquina, dentro da qual encontram-se classes abstratas responsáveis por cada etapa do reconhecimento. A fachada do processo é a classe de reconhecimento, que, quando recebe um sinal, direciona-o para uma classe de filtro, a partir do qual o sinal processado segue para o método de extração de características. Estas partem para o método de classificação, o qual deve ser implementado como algum classificador. O banco de dados participa externo a esse fluxo, provendo os dados de treinamento, de testes e de validação. Ao final desse processo, tem-se a sílaba como resultado de interesse. Na Figura 5 segue a exemplificação através do diagrama.

Figura 5 – Esquemático da arquitetura de funcionamento do projeto.



Fonte: Elaborado pelas autoras (2017).

Vale salientar que a arquitetura é um conjunto de classes abstratas e interfaces, as quais devem ser estendidas por subclasses para se constituir uma

implementação concreta do sistema. Graças ao modelo em camadas, é possível substituir, por exemplo, o método de extração de descritores sem afetar a classificação, e vice-versa, facilitando o processo de desenvolvimento e a adequação a cenários particulares.

3.4 BANCO DE DADOS E EXTRAÇÃO DE DESCRITORES

Com a quantidade de áudios desejada, deu-se início à construção do banco de dados. A priori foi necessário extrair as características de cada áudio e organizá-las em vetores. Para isso, fez-se uso do método de extração de descritores MFCC (Mel-Frequency Cepstral Coefficients) através da biblioteca `python_speech_features`.

Para tanto, fez-se uso dos módulos `os` e `numpy`, os quais possibilitam abrir diretórios e arquivos como também trabalhar com matrizes multidimensionais de alto desempenho. Posteriormente, seguiu-se com a extração da matriz de descritores por MFCC usando os parâmetros listados no Quadro 1.

Quadro 1 – Parâmetros usados para extrair a matriz de características

Parâmetros	Descrição
Sig	o sinal de áudio para enquadrar
Rate	velocidade da resposta
Nfft	amostragem do comprimento da FFT (Transformada Rápida de Fourier)

Fonte: Elaborado pelas autoras (2017)

O método utilizado para a extração via MFCC retornava uma matriz de descritores. Como o interesse era de se obter um vetor, calculou-se a média e o desvio padrão de cada linha da matriz do MFCC, sendo esses valores agrupados em um vetor.

Os vetores obtidos foram também automaticamente direcionados para uma planilha no Excel, com a implementação do módulo `Xlsxwriter` do Python.

3.5 TREINAMENTO DO ALGORITMO DE APRENDIZAGEM DE MÁQUINA

Com os dados já organizados, foi usado o KNN com 7, 8 e 10 vizinhos, que correspondem à quantidade de pontos mais próximos que o algoritmo deve considerar. Para essa etapa, 30 gravações de cada sílaba foram utilizadas. Para testar a acurácia, 10 outras gravações de cada sílaba foram aplicadas.

Por causa dos descritores utilizados ou do próprio KNN, que talvez seja simples para o objetivo do projeto, ou por causa da falta de um filtro, os resultados, até então, não foram tão satisfatórios.

O algoritmo apresentou acurácia entre 0% e 60% para a maioria das sílabas consideradas. A Tabela 1 abaixo exhibe as porcentagens de classificação correta para algumas sílabas:

Tabela 1 – Sílabas e quantidade de acertos em porcentagem (%)

SÍLABAS	ACERTOS (%)
RA	20
RROR	30
VI	20
RO	0
SÃO	10
SÕES	50
CHE	40
CHA	10
CHI	40
CHU	40
PRU	10
SUL	40
TA	0
TROM	10
PI	60
SEM	30
PU	20

Fonte: Elaborado pelas autoras (2017)

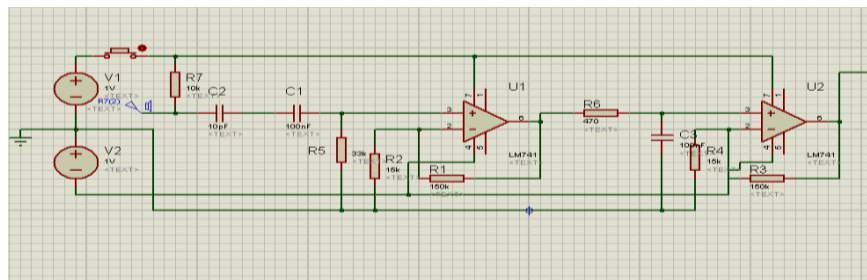
3.6 CIRCUITO DE CAPTAÇÃO DE ÁUDIO

Para a captação de áudio do ALFA surgiu a necessidade de se confeccionar além de um circuito de captação de voz também um circuito de filtragem, visto que os limites inferiores e superiores do espectro de frequência da voz humana variam

entre 50 Hz e 3400 Hz. O circuito pensado constitui-se de um microfone de eletreto acoplado a um filtro passa-alta ativo com frequência de corte de 50Hz e um filtro passa-baixa ativo com frequência de corte de 3400 Hz, em cascata, assim formando um filtro passa-faixa ativo, a fim de reduzir a presença de ruídos.

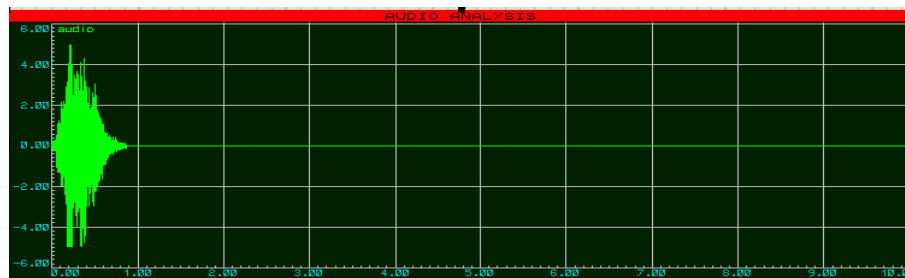
E para a simulação desse circuito, descrito no parágrafo acima como mostra a Figura 6, usou-se como ferramenta o módulo ISIS do software Protheus 8 com o qual obteve-se bons resultados tanto na simulação (Figura 7) quanto na prática referente aos filtros.

Figura 6 – Simulação do circuito no software Protheus.



Fonte: Das autoras (2017)

Figura 7 – Simulação da entrada de áudio no Protheus 8 (sílabas MAR)



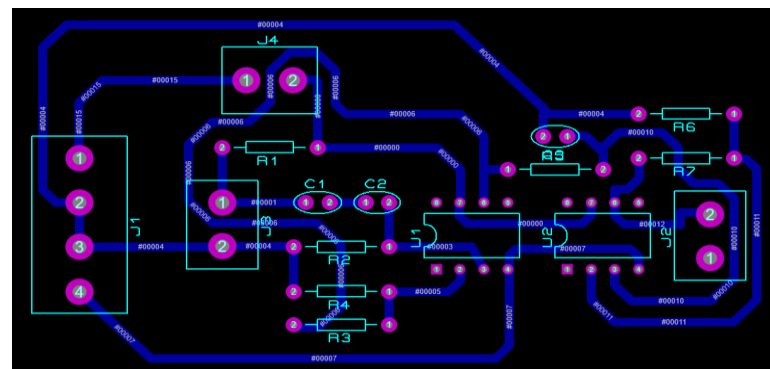
Fonte: Das autoras (2017)

Inicialmente fora disposto sobre a tela de edição uma fonte simétrica de 12V, para a alimentação dos amplificadores operacionais LM741, dois resistores de 15KΩ e 150KΩ que foram usados para a obtenção do ganho de cada filtro desejado no circuito igual a 10, que quando somados daria um ganho ideal igual a 20. Com configuração do circuito que foi proposto, o objetivo era de que ao se obter uma onda mecânica no microfone de eletreto, o sinal elétrico seguiria para o filtro passa-alta, no qual qualquer frequência detectada maior que 50 Hz seguiria para o segundo filtro, passa-baixa, com um ganho de tensão igual a 10. E já no segundo

filtro caso a frequência fosse menor que 3400 Hz, o sinal elétrico sairia ao final do circuito total com um ganho de 20, assim seria totalmente viável fazer a captação da fala com o mínimo de ruído possível.

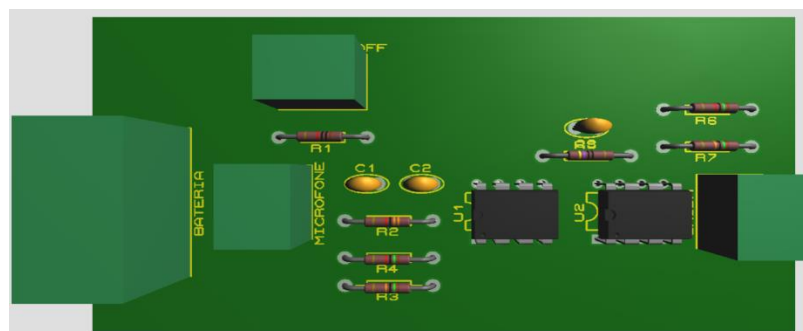
Após a simulação do esquemático do circuito de captação de áudio e de filtragem juntos, formulou-se então o layout do mesmo no módulo ARES do software Protheus, a Figura 8 mostra o layout feito depois dos ajustes de trilhas e ilhas. E com este mesmo módulo foi possível a visualização 3D da placa do circuito, como apresentado na Figura 9.

Figura 8 – Layout do circuito de captação de áudio no ARES



Fonte: Das autoras (2017)

Figura 9 – Visualização 3D do circuito de captação de áudio no ARES

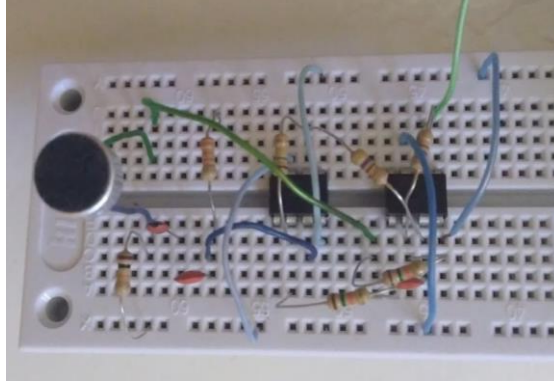


Fonte: Das autoras (2017)

Dando continuação a etapa, partiu-se para a confecção do circuito como um todo em protoboard, Figuras 10 e 11. Primeiramente montou-se o circuito de captação de áudio, composto por um microfone de eletreto, um resistor de 10 K Ω e um capacitor de 10pF, em seguida foi confeccionado o circuito de filtragem e

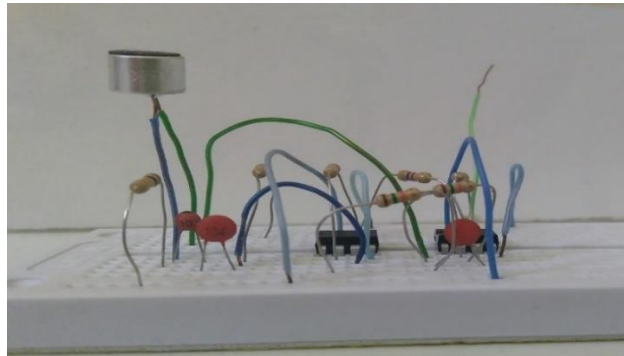
amplificação, composto por um filtro passa-alta ativo e passa-baixa ativo, em cascata, de ganho igual a 20.

Figura 10 – Imagem frontal do circuito montado em protoboard



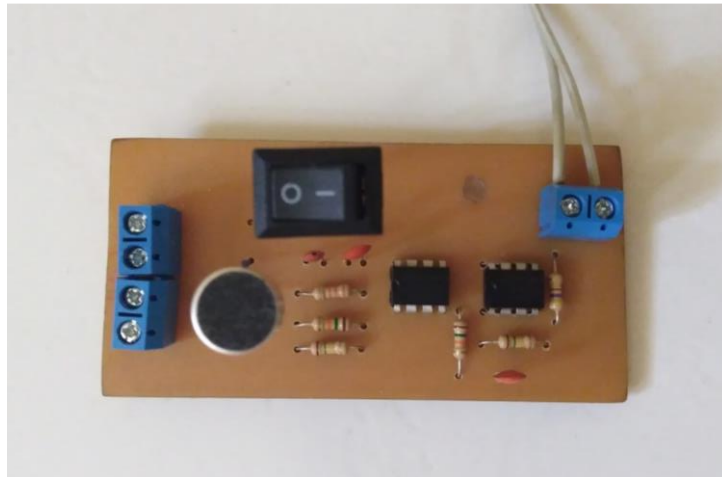
Fonte: Das autoras (2017)

Figura 11 – Perfil do circuito montado em protoboard



Fonte: Das autoras (2017)

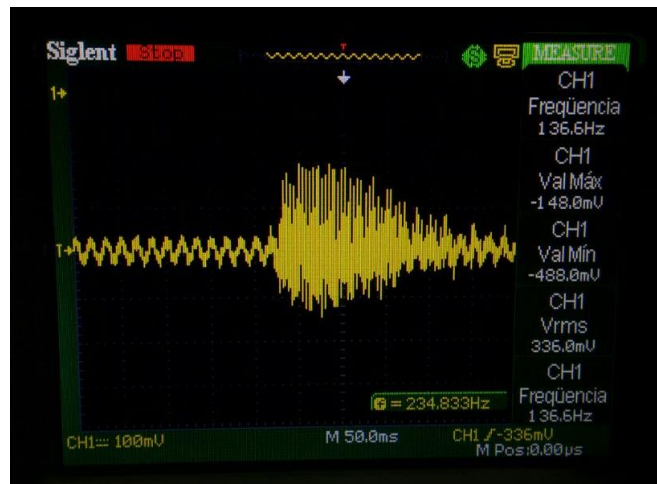
Figura 12 – Placa de circuito impresso do microfone



Fonte: Das autoras (2017)

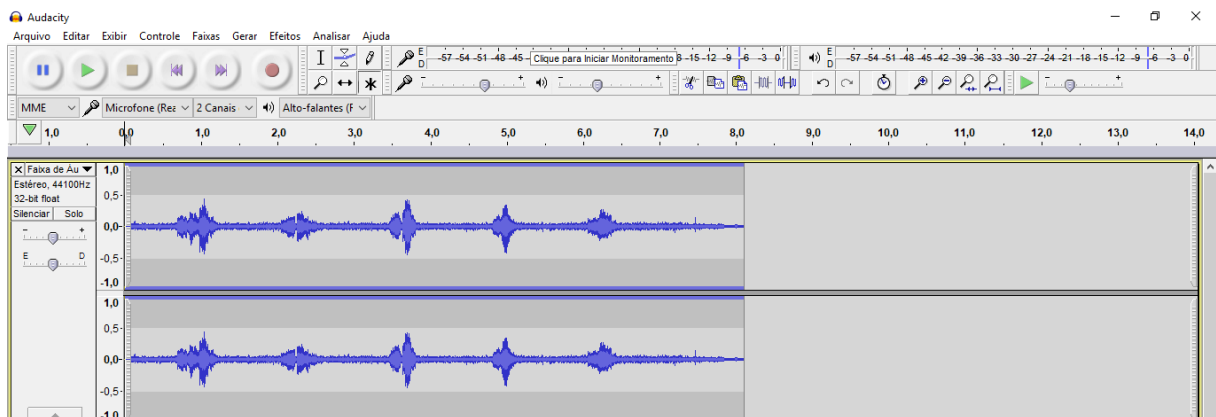
Com a confecção do circuito em protoboard, o próximo passo se caracterizou pela a fabricação da placa de circuito impresso do microfone (Figura 12). Para a realização dessa etapa, foi preciso a aquisição dos componentes necessários para a construção do circuito, como o microfone de eletreto, os resistores, capacitores, amplificadores operacionais, botão ON/OFF e os conectores. Com os componentes em mãos, seguiu-se para a transferência do layout impresso em folha de papel fotográfico para a placa de cobre por meio da técnica de choque térmico, e posteriormente partiu-se para a corrosão do cobre na solução de perclorato de ferro. Ao final, deu-se continuidade com a limpeza, perfuração da placa e soldagem dos componentes na mesma.

Figura 13 – Teste com o fonema "PÉ"



Fonte: Dos autores (2017)

Figura 14 – Teste do microfone no software Audacity



Fonte: Das autoras (2017)

Para os testes do microfone, a princípio buscou-se testar o circuito em laboratório com o auxílio de um osciloscópio para a verificação da tensão e frequência de saída, fora testado o microfone com a pronúncia de fonemas, sendo entre eles, por exemplo o fonema PÉ como mostrado na Figura 13. Posteriormente usou-se o software Audacity (Figura 14), o qual foi usado também para todas as gravações de fonemas ao longo das coletas de vozes realizadas. O áudio obtido como resultado do circuito, mostrou-se de boa qualidade e sem a presença de grandes ruídos, ainda de baixa amplitude, mas isso pode ser melhorado ao longo do trabalho. Ao final, obteve-se um microfone de custo menor que 40 reais.

3.7 DESENVOLVIMENTO DA INTERFACE LÚDICA

Com a etapa do treinamento em andamento, deu-se início às pesquisas referentes ao desenvolvimento da interface lúdica do ALFA. Depois de algumas pesquisas e reuniões para decidir se a interface seria online ou não, decidiu-se que, até então, não seria. Essa decisão foi fundamentada tendo em vista que nem sempre o usuário teria acesso à internet, dificultando a utilização do ALFA.

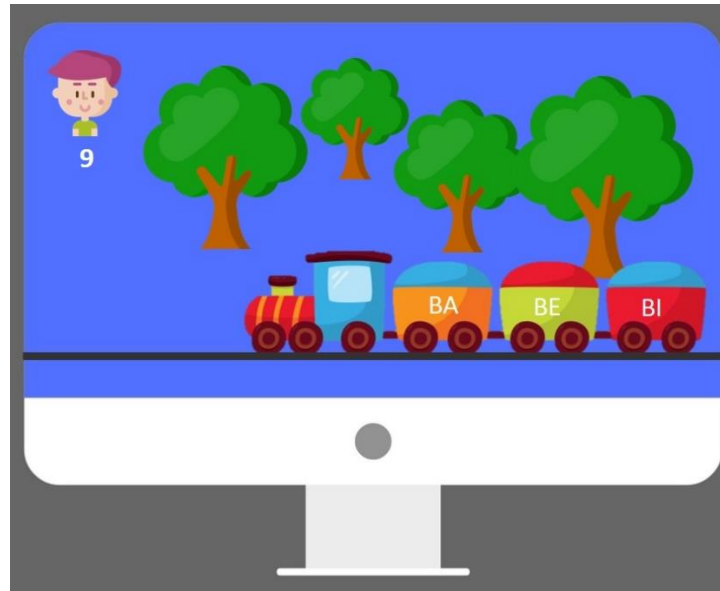
Em seguida, foram esboçados alguns jogos como, por exemplo: a ideia de um labirinto no qual os movimentos seriam realizados a partir das vogais faladas pelo usuário, como mostrado no filme *Le Pays des sourds* (1992) — *O País dos surdos*, em português —, de Nicolas Philibert, representado na Figura 15; um trem andando quando o usuário dissesse algum fonema, demonstrado na Figura 16; e uma corda bamba, onde haveria um boneco que somente andaria quando o usuário falasse corretamente cada palavra.

Figura 15 – Jogo do labirinto



Fonte: Filme *Le Pays des sourds* (1992)

Figura 16 – Esboço do jogo do trem a ser implementado na interface lúdica



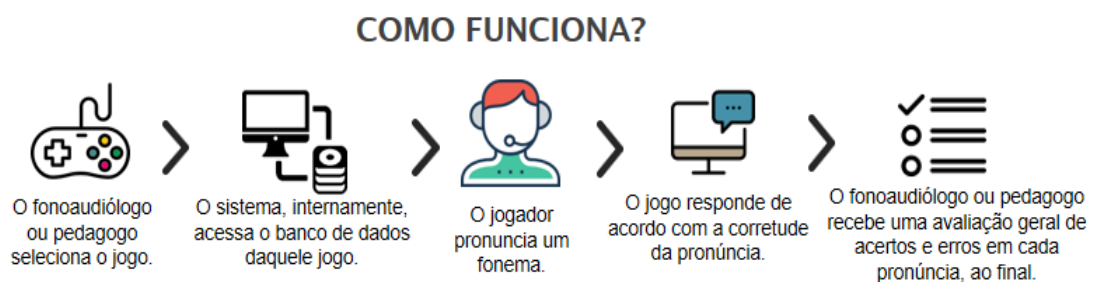
Fonte: Elaborado pelas autoras (2017)

A fim de melhor integrar o sistema de reconhecimento com a visão, preferiu-se adotar também a linguagem Python para o desenvolvimento dos jogos, através do módulo PyGame, de ampla utilização nesse âmbito, com ferramentas facilitadoras para a programação.

3.8 SISTEMA FINAL

Com base no exposto, a Figura 17 abaixo resume o funcionamento do sistema proposto neste projeto:

Figura 17 – Diagrama



Fonte: Elaborada pelas autoras (2017)

Como mostrado na Figura 17 acima, o fonoaudiólogo ou pedagogo seleciona um jogo de acordo com a necessidade do paciente ou aluno. Em seguida, ao selecionar o jogo, o sistema, internamente, acessa o banco de dados daquele jogo buscando os fonemas que serão “trabalhados”. Após a seleção, é solicitado que o jogador pronuncie os fonemas exercitados no jogo escolhido, assim o jogo segue respondendo de acordo com a corretude de cada fonema. Ao final, o fonoaudiólogo ou pedagogo recebe uma avaliação geral de acertos e erros em cada pronúncia; vale ressaltar que o jogo só responde com uma ação positiva quando o usuário falar corretamente o fonema.

4 RESULTADOS E DISCUSSÕES

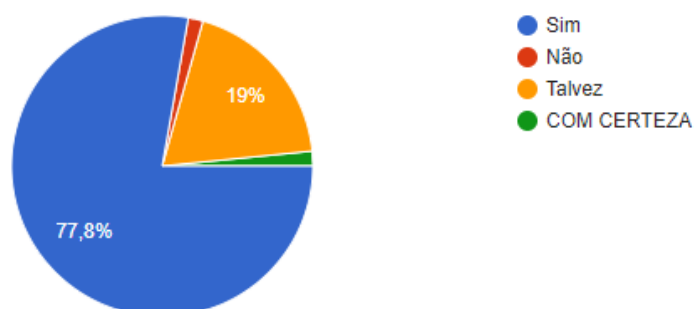
O processo de coleta de vozes culminou em um banco de dados com dez mil e cem gravações, com aproximadamente 240 áudios de cada uma das vinte e seis famílias de sílabas de interesse.

Os testes preliminares com o algoritmo de classificação KNN não apresentaram resultados satisfatórios, com uma taxa de acerto por sílaba abaixo de 30%. Isso se deve, provavelmente, à simplicidade da técnica dos vizinhos mais próximos ou à escolha do vetor de descritores. Espera-se que classificadores neurais apresentem melhores resultados.

Quanto ao circuito de captação de áudio, os testes com o filtro passa-faixa ativo foram satisfatórios, visto que foi possível a captação, filtragem e amplificação dos áudios no circuito. Os áudios obtidos por meio do software Audacity mostraram-se de boa qualidade, não apresentando tantos ruídos e nem alterações nas vozes captadas. Além de obtermos bons resultados, também conseguimos construir um circuito com componentes de baixo custo, o que possibilitou chegar ainda mais perto do desenvolvimento de um sistema final sem maiores custos.

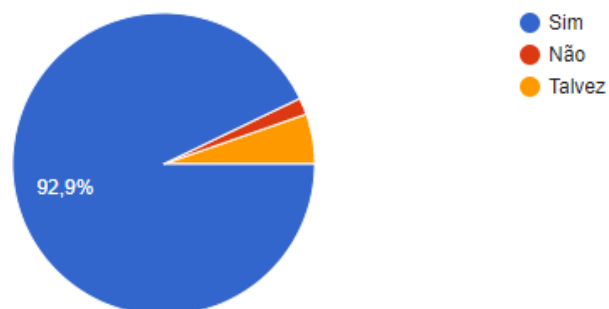
Os resultados das pesquisas sobre a aceitação do projeto por parte do público-alvo foram positivos, conforme os gráficos abaixo, correspondentes, respectivamente, às opiniões dos fonoaudiólogos sobre se o ALFA é realmente útil para sessões de fonoterapia, e dos pedagogos, sobre se é útil para sessões de apoio pedagógico.

Gráfico 1 – Pesquisa sobre a utilidade do ALFA para os fonoaudiólogos



Fonte: Das autoras (2017)

Gráfico 2 – Pesquisa sobre a utilidade do ALFA para os pedagogos



Fonte: Das autoras (2017)

Ainda na mesma pesquisa, 93,5% dos fonoaudiólogos ou estudantes de fonoaudiologia concordam que novas tecnologias devem ser empregadas no âmbito da fonoaudiologia, 79,4% acreditam que o ALFA pode aumentar a eficácia da identificação de alterações fonológicas. Ademais, 77,2% dos pedagogos acredita que o ALFA pode ser aplicado tanto em salas de aula, quanto em sessões individuais, 78,9% concordam que a interface lúdica do sistema estimulará o aluno no processo de correção da dicção.

5 CONSIDERAÇÕES FINAIS

Este trabalho propõe um sistema fundamentado em reconhecimento de padrões para detecção automática de alterações no processo fonológico de pessoas, especialmente de crianças.

O desenvolvimento deste projeto, até o momento, culminou na composição de um banco de dados amplo de pronúncias e em testes preliminares com a técnica K-Nearest Neighbors, além de um circuito de captação de áudio ainda a ser aperfeiçoado. Embora, do ponto de vista técnico, muito ainda haja para ser feito, o potencial e a factibilidade do projeto foram demonstrados por pesquisas e por trabalhos do estado da arte, o que está sendo muito motivador para todos os colaboradores do projeto.

Tem-se a expectativa de que o ALFA, em um ambiente real, proporcione ao usuário e ao profissional (fonoaudiólogo e/ou pedagogo) uma experiência mais interessante, lúdica e interativa. As próximas etapas de desenvolvimento do sistema proposto serão voltadas para o reconhecimento dos fonemas pronunciados pelo usuário diretamente do microfone para o algoritmo reconhecedor, e para a conexão da interface lúdica ao programa. Paralelamente, serão realizados testes de outros métodos de aprendizado de máquina como RNAs, a fim de se comparar qual a melhor alternativa para o sistema proposto.

Além disso, outra etapa prevista no desenvolvimento do trabalho é o teste do ALFA com, ao menos, um paciente em fonoterapia e em uma turma de alfabetização. Por fim, pretende-se formular um pedido de patente do sistema desenvolvido.

Como extensão do projeto, sugere-se ainda que sejam ampliadas as versões de jogos, para que estes atendam às necessidades dos mais diversos tipos de tratamento fonoterapêutico e auxiliem o maior número de pessoas possível, seja na alfabetização, seja na fonoterapia.

REFERÊNCIAS

- BANG, Jeosoo et al. Pronunciation variants prediction method to detect mispronunciations by korean learners of english. **Acm Transactions On Asian Language Information Processing** (talip), v. 13, n. 4, p.1-21, dez. 2014.
- BARBOSA, Mirna Rossi; BARBOSA, Luiza Augusta Rosa Rossi; SAMPAIO, Cristina. A fonoaudiologia no curso de pedagogia: percepções dos estudantes. **Unimontes científica**, v.12, n1/2, Montes Claros, 2010.
- OLIVEIRA, João Batista Araujo. **Por que Joãozinho não sabe ler?**, 2010 Disponível em: <<http://www.alfaabeto.org.br/por-que-joaozinho-nao-sabe-ler/>>. Acesso em: 06 abr. 2017.
- _____. **Alfabetização**: como ensinar a ler e a escrever com método fônico. Disponível em: <<http://www.alfaabeto.org.br/blog/alfabetizacao-com-metodo-fonico/>> Acesso em: 11/04/2017.
- PORTAL BRASIL. **Apesar de avanços, surdos ainda enfrentam barreiras de acessibilidade**. [Brasília], 2016. Disponível em: <<http://www.brasil.gov.br/cidadania-e-justica/2016/09/apesar-de-avancos-surdos-ainda-enfrentam-barreiras-de-acessibilidade>> Acesso em: 12 abr. 2017.
- CAPOVILLA, Fernando C.; CAPOVILLA, Alessandra G. S. Problemas de aquisição de leitura e escrita: efeitos de déficit de discriminação fonológica, velocidade de processamento e memória fonológica. **Estudos e Pesquisas em Psicologia**, São Paulo, v. 2, n. 1, p. 26-50, 2002. Disponível em: <<http://www.e-publicacoes.uerj.br/index.php/revispsi/article/view/7703>>. Acesso em: 15 mai. 2017.
- CAPOVILLA, Alessandra G. S.; CAPOVILLA, Fernando C. **Alfabetização: Método fônico**. 2007. Curso de Psicologia, Instituto de Psicologia, Universidade de São Paulo e Instituto de Psicopedagogia, Universidade de Santo Amaro, São Paulo, 2007.
- CAPOVILLA, Fernando; SEABRA, Alessandra G. **Alfabetização: método fônico**. 5. ed. São Paulo: Memmon, 2010.
- CROVATO, César David Paredes. **Classificação de sinais de voz utilizando a transformada Wavelet Packet e redes neurais artificiais**. Porto Alegre: [s.l.], 2004.
- DENG, Li; LI, Xiao. Machine learning paradigms for speech recognition: an overview. **IEEE transactions on audio, speech, and language processing**, v. 21, n. 5, p.1060-1089, maio 2013.
- PORTAL EDUCAÇÃO. **Linguagem que se aprende**. Campo Grande, 2012. Disponível em: <<https://www.portaleducacao.com.br/conteudo/artigos/pedagogia/linguagem-que-se-aprende/23680>> Acesso em: 09 abr. 2017.
- FARDO, Marcelo Luis. A gamificação aplicada em ambientes de aprendizagem.

Novas Tecnologias na Educação, Caxias do Sul, v. 11, p.1-9, jul. 2013. Disponível em: <<http://www.seer.ufrgs.br/index.php/renote/article/view/41629/26409>>. Acesso em: 22 dez. 2017.

FEATURE Extraction. Disponível em: <<http://recognize-speech.com/feature-extraction>>. Acesso em: 17 jul. 17.

FERNANDES, Anita Maria da Rocha. **Inteligência artificial: noções gerais**. Florianópolis: VisualBooks, 2005.

NOCETI FILHO, Sidnei. **Filtros seletores de sinais**. 3 ed. Florianópolis: UFSC, 2010.

FREITAS, Patrícia Gomes de. **Um olhar sobre o método fônico**. 2011. Disponível em: <<http://www.uel.br/ceca/pedagogia/pages/arquivos/PATRICIA%20GOMES%20DE%20FREITAS.pdf>>. Acesso em: 07 abr. 2017.

GOMES, Patrícia. **Tecnologia e fono se unem para melhorar fala**. 2012. Disponível em: <<http://porvir.org/tecnologia-fono-se-unem-para-melhorar-fala/>>. Acesso em: 03 nov. 2017.

GONÇALVES, Leila Laís; GIACOMAZZO, Graziela Fátima; Rodrigues, Flávia; MACAIA, César Bráulio Sumbo. Gamificação na Educação: um modelo conceitual de apoio ao planejamento em uma proposta pedagógica. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 27., 2016. **Anais...** 2016.

HORTA, Leila; TOMITA, Shiro. **Um método de investigação dos distúrbios da fala e voz: a espectrografia vocal**. 2001. Disponível em: <<http://fonoaudiologia.com/artigo/um-metodo-de-investigacao-dos-disturbios-da-fala-e-voz-a-espectrografia-vocal.html>> Acesso em: 12 abr. 2017.

JÚNIOR, Antonio Pertence. **Amplificadores operacionais e filtros ativos**. 6 ed. Porto Alegre: Bookman, 2003.

KRÜGER, Stefan. **Introduction to Artificial Neural Networks**, 2014. Disponível em: <[http://recognize-speech.com/basics/introduction-to-artificial-neural-networks#\[object HTMLHeadingElement\]](http://recognize-speech.com/basics/introduction-to-artificial-neural-networks#[object HTMLHeadingElement])> Acesso em: 22 dez. 2017.

MACEDO, Thiago; RABELO, Camila. **Aplicativo com exercícios de fonoaudiologia para criança com deficiência auditiva**. 2014. Disponível em: <http://www.puc-rio.br/pibic/relatorio_resumo2014/relatorios_pdf/ctch/ART/ART-Thiago%20Macedo%20e%20Camila%20Rabelo.pdf>. Acesso em: 15 abr. 2017.

MOREIRA, Luciano. **Atraso da fala: sinais de alerta**. 2015. Disponível em: <<http://portalotorrino.com.br/atraso-na-fala-ate-quando-e-normal/>>. Acesso em: 15 abr. 17.

MOUSINHO, Renata. **Aquisição e desenvolvimento da linguagem: dificuldades que podem surgir neste percurso**. 2008. Disponível em:

<http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0103-84862008000300012>. Acesso em: 16 out. 2017.

NASCIMENTO, Camila Lima et al. **Colaboração e participação de crianças no atendimento fonoaudiológico**. 2010.

NASCIMENTO, Francisco de Assis do. **Como ocorrem os distúrbios da linguagem oral e da comunicação na criança**. 2007. Disponível em: <http://www.psicologia.pt/artigos/ver_artigo_licenciatura.php?codigo=TL0086>. Acesso em: 15 nov. 2017.

NASCIMENTO, Matheus Batista; ALBUQUERQUE, Eduardo Simões. **Uso a gamification para melhorar a adesão ao tratamento**. 2015. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/sbsi/2015/052.pdf>>. Acesso em: 15/04/2017.

QIAN, Xiaojun; SOONG, Frank; MENG, Helen. **Discriminatively Trained Acoustic Model for Improving Mispronunciation Detection and Diagnosis in Computer Aided Pronunciation Training (CAPT)**. 2010. Disponível em: <http://www1.se.cuhk.edu.hk/~hccl/publications/pub/xiaojun_interspeech2010.pdf>. Acesso em: 16 abr. 2017.

ROCHA, Sinara Socorro Duarte. O uso do computador na educação: a informática Educativa. **Espaço Acadêmico**, Maringá, v. 85, p.1-6, jun. 2008. Mensal. Disponível em: <https://3c726bc3-a-62cb3a1a-sites.googlegroups.com/site/prntextos2/educacao-matematica/TC-InfomáticaEducativa.pdf?attachauth=ANoY7cpHYvF80R_5HTmPrwEY5PWGZps7og1rOdFh3Rkhkp2uZ_mXZuFT1Li0d0f2L511DoQzUnxfCePvUpqinUmRDIUcJEf3ZcvwogjO8y0zvttL1FtFDRHR2cLgyrWz0Y0f1OQ9cNbgFRNhUGf83Wb8sCTi9eTfQMmuAtC8PgUx1BObBsl-3ghut2qyRW8MO_KJrANK4YjY6UtAPMmk4B-LVDoyyzRyyA2pbKf_yUhl6ZZ2jqRaUzsXKhQJGBzM5MNPnZeU1IWGO&attredirects=0>. Acesso em: 22 dez. 2017.

SANTOS, Karoline Weber dos Santos et al. **Utilização de softwares em pesquisas científicas de fonoaudiologia**. 2012. Disponível em : <<http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/185/116>> Acesso em: 07 nov. 2017.

SITTA, Erica. **Qual é a diferença entre fonética e fonologia?**. 2012. Disponível em: <<https://ericasitta.wordpress.com/2012/07/07/qual-e-a-diferenca-entre-fonetica-e-fonologia/>> Acesso em: 09 abr. 2017.

SIEVES, Cristiano. **Gamificação na escola: 3 exemplos para professores**. Disponível em: <http://playtable.com.br/blog/gamificacao-nas-escolas-3-exemplos-para-professores/> Acesso em: 15/04/2017.

SEMIS, Laís. **Avaliação Nacional de Alfabetização (ANA) 2016 acontece entre 14 e 25 de novembro**. Disponível em: <<https://gestaoescolar.org.br/conteudo/1671/avaliacao-nacional-de-alfabetizacao-ana-2016-acontece-entre-14-e-25-de-novembro>>. Acesso em: 07/04/2017.

THE Speech Recognition. Disponível em: <<http://recognize-speech.com/>>. Acesso em: 17 jul. 2017.

VARELLA, Drauzio. **Problemas da fala na criança**. [S.l. : s. n.], 2012. Disponível em: <<https://drauziovarella.com.br/crianca-2/problemas-da-fala-na-crianca/>>. Acesso em: 16 abr. 2017.

VICARIA, Luciana. **Testes em crianças identificam problemas de linguagem que podem prejudicar a alfabetização**. 2012. Disponível em: <<http://revistaepoca.globo.com/Sociedade/noticia/2012/11/testes-em-criancas-identificam-problemas-de-linguagem-que-podem-prejudicar-alfabetizacao.html>>. Acesso em: 06 abr. 2017.

YAVAS, Mehmet; HERNANDORENA, Carmen L. Matzenauer; LAMPRECHT, Regina Ritter. Avaliação fonológica da criança. Porto Alegre: Artes Médicas, 1991.

ZIEGLER, Maria Fernanda. **Crianças que nasceram surdas passam a escutar após cirurgia**. 2013. Disponível em: <<http://saude.ig.com.br/minhasaude/2013-11-18/criancas-que-nasceram-surdas-voltam-a-escutar-apos-passarem-por-cirurgia.html>>. Acesso em: 06 abr. 2017.

APÊNDICE A – Código prototípico de extração de características e reconhecimento

```

from python_speech_features import mfcc           # extrair descritores
from python_speech_features import delta         # extrair descritores
from python_speech_features import logfbank      # extrair descritores
from sklearn.neighbors import NearestNeighbors   # para usarmos o KNN
import scipy.io.wavfile as wav                  # carregar o áudio
import xlswriter
import numpy as np
import os    # para navegarmos nos diretórios e lermos todos os arquivos

audio_dir = '../audio'

print(os.listdir(audio_dir))

# Vamos contar quantos áudios nós temos e quantas features por áudio

sylfile = "map_labels.txt"
fo = open(sylfile, "w")

features_mfcc_qtd = -1
cod_sy = 0
map_syl_to_number = dict()
map_number_to_syl = dict()
audios_qtd=0
for d in os.listdir(audio_dir):
    family_dir = audio_dir + '/' + d
    for s in os.listdir(family_dir):
        sy_dir = family_dir+ '/' + s
        map_syl_to_number[s] = cod_sy
        map_number_to_syl[cod_sy] = s
        fo.write(str(cod_sy) + "," + str(s) + "\n")
        cod_sy = cod_sy + 1
        for a in os.listdir(sy_dir):
            print(a)
            audio_path = sy_dir + '/' + a
            (rate, sig) = wav.read(audio_path);
            # Extrair a matriz de features por MFCC (113x13).
            mfcc_feat = mfcc(sig, rate, nfft=1105)
            features_mfcc_qtd = max(features_mfcc_qtd, mfcc_feat.shape[0])
            audios_qtd = audios_qtd + 1

fo.close()
# Agora, vamos extrair os descritores arquivo a arquivo

features = np.zeros(shape=(audios_qtd,features_mfcc_qtd*2)) # Nossa
matriz dos descritores de todos os áudios
labels = np.zeros(shape=(1, audios_qtd)) # Nosso
vetor de classes com uma coluna para cada áudio (cada áudio tem uma
classe, que é a sílaba que ele representa)

r = 0 # Esta variável aponta para a linha de matriz de descritores que
estamos computando agora
for d in os.listdir(audio_dir):
    family_dir = audio_dir + '/' + d
    for s in os.listdir(family_dir):
        sy_dir = family_dir+ '/' + s
        for a in os.listdir(sy_dir):

```

```

print(a)
# Definimos a label desse audio (a sílaba dele)
labels[0][r] = map_syl_to_number[s]
audio_path = sy_dir + '/' + a
# Carregar o áudio (cada sílaba pronunciada)
(rate, sig) = wav.read(audio_path);
# Extrair a matriz de features por MFCC.
mfcc_feat = mfcc(sig, rate, nfft=1105)
# Como queremos um vetor, vamos fazer assim: calculamos a
média e o desvio padrão de cada linha da matriz
# do MFCC e vamos agrupar em um vetor. Ele será nosso vetor
de características.
j = 0
for i in mfcc_feat:
    features[r][j] = np.mean(i)
    features[r][j+1] = np.std(i)
    j = j + 2
r = r + 1

np.savetxt('features.txt', features, delimiter=',')
np.savetxt('labels.txt', labels, delimiter=',')
#workbook = xlswriter.Workbook('features.xlsx')
#worksheet = workbook.add_worksheet()
#row = 0
#for i in range(len(features)):
#    for j in range(len(features[i])):
#        worksheet.write_column(i, j, features[i][j])

#for col, data in enumerate(features):
#    worksheet.write_column(row, col, data)
#workbook.close()

#knn = NearestNeighbors(n_neighbors=8,
algorithm='ball_tree').fit(features)

#(rate, sig) = wav.read("../PI-guilherme.wav");
#mfcc_feat = mfcc(sig, rate, nfft=1105)
#new_sample = np.zeros(shape=(1, features_mfcc_qtd*2)) # Nossa matriz
dos descritores de todos os áudios
#j = 0
#for i in mfcc_feat:
#    new_sample[0][j] = np.mean(i)
#    new_sample[0][j+1] = np.std(i)
#    j = j + 2

#distances, indices = knn.kneighbors(new_sample)

#print(map_number_to_syl[labels[0][indices[0][0]]])

```