

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA SUL-RIO-
GRANDENSE - CÂMPUS PASSO FUNDO
CURSO DE TECNOLOGIA EM SISTEMAS PARA INTERNET**

JOEL ZATTI

**I CAN SEE: UMA PROPOSTA DE SOFTWARE DE
RECONHECIMENTO DE OBJETOS PARA AUXILIAR PESSOAS COM
DEFICIÊNCIA VISUAL**

João Mário Lopes Brezolin

**PASSO FUNDO
2018**

JOEL ZATTI

**I CAN SEE: UMA PROPOSTA DE SOFTWARE DE
RECONHECIMENTO DE OBJETOS PARA AUXILIAR PESSOAS COM
DEFICIÊNCIA VISUAL**

Monografia apresentada ao Curso de Tecnologia em Sistemas para Internet do Instituto Federal Sul-rio-grandense, Câmpus Passo Fundo, como requisito parcial para a obtenção do título de Tecnólogo em Sistemas para Internet.

Orientador (a): João Mário Lopes Brezolin

PASSO FUNDO

2018

JOEL ZATTI

**I CAN SEE: UMA PROPOSTA DE SOFTWARE DE RECONHECIMENTO DE
OBJETOS PARA AUXILIAR PESSOAS COM DEFICIÊNCIA VISUAL**

Trabalho de Conclusão de Curso aprovado em ____/____/____ como requisito
parcial para a obtenção do título de Tecnólogo em Sistemas para Internet

Banca Examinadora:

Nome do Professor(a) Orientador(a)

Nome do Professor(a) Convidado(a)

Nome do Professor(a) Convidado(a)

Coordenação do Curso

PASSO FUNDO

2018

DEDICATÓRIA

*À Deus, pela saúde, força e garra,
À minha família, pela compreensão e o estímulo
em todos os momentos de dificuldades.*

AGRADECIMENTOS

Ao professor Dr. João Mário Lopes Brezolin pela orientação e ajuda necessária para finalizar esta árdua etapa e estendo também a todos os demais professores que contribuíram de uma maneira ou outra para a minha formação.

EPÍGRAFE

“A força motriz para o desenvolvimento
de novos produtos não é tecnologia,
nem dinheiro,
mas a imaginação das pessoas.”

David Packard

RESUMO

Segundo o IBGE as pessoas com necessidades especiais representam um percentual de 6,2% da população brasileira e aplicativos dessa natureza podem melhorar a qualidade de vida dessas pessoas. Este trabalho objetiva desenvolver um aplicativo para auxiliar pessoas com deficiência visual a realizar o reconhecimento de imagens. Isso será possível através da implementação de um recurso de áudio-descrição que vai permitir ler para o usuário a descrição da imagem capturada pela câmera do *smartphone*. O aplicativo foi desenvolvido para a plataforma Android. A programação foi realizada por meio da IDE Android Studio e com auxílio dos recursos da biblioteca gráfica *OpenCV*. Futuramente pretende-se portar a aplicação para as demais arquiteturas do mercado. Para validar o aplicativo foram realizados os testes com um usuário cego. Por meio da biblioteca *OpenCV* será realizada a classificação das imagens capturadas pelo *smartphone*. Quando o usuário abre o aplicativo **I CAN SEE**, automaticamente será aberta a câmera do *smartphone* que captura a imagem e por meio do algoritmo de classificação identifica a imagem que foi capturada. Após essa classificação, o recurso de áudio – implementado através da biblioteca *Text-to-Speech* - descreve na forma de áudio a imagem capturada. Caso a imagem não esteja no repositório a mensagem vai explicitar que não foi identificada. Espera-se que esse aplicativo possa auxiliar as pessoas com necessidades visuais a ver o mundo com outros olhos. Para validar o protótipo foram realizados testes com um usuário cego. Os testes demonstraram que o aplicativo obteve êxito e realmente auxiliou o usuário na tarefa de reconhecer as imagens nas fotos, ler as suas descrições e informar o que estava ocorrendo no aplicativo.

Palavras-chave: Tecnologias assistivas, Dispositivos móveis, Android Studio. Classificação de imagens, *OpenCV*.

ABSTRACT

According to the IBGE, people with special needs represent a percentage of 6.2% of the Brazilian population and applications of this nature can improve the quality of life of these people. This work aims to develop an application to assist people with visual impairment to perform image recognition. This will be possible through the implementation of an audio-description feature that will allow the user to read the description of the image captured by the smartphone camera. The application was developed for the Android platform. Programming was done through the Android Studio IDE and with the help of the OpenCV graphic library. In the future it is intended to carry the application to the other architectures of the market. To validate the application tests were performed with a blind user. The OpenCV library will classify the images captured by the smartphone. When the user opens the I CAN SEE application, the smartphone camera that captures the image will automatically open and through the classification algorithm identifies the image that was captured. After this classification, the audio feature - implemented through the Text-to-Speech library - describes in audio form the captured image. If the image is not in the repository the message will explain that it was not identified. It is hoped that this application can help people with visual needs see the world with different eyes. To validate the prototype tests were performed with a blind user. The tests demonstrated that the application was successful and really helped the user in the task of recognizing the images in the photos, reading their descriptions and informing what was happening in the application.

Keywords: Assistive Technologies, Mobile Devices, Android Studio. Image classification, OpenCV.

LISTA DE FIGURAS

Figura 1 - Instruções para o usuário.....	21
Figura 2 - Saída do áudio <i>Add</i>	21
Figura 3 - Saída de áudio <i>Flush</i>	22
Figura 4 - Bloqueio da orientação do aplicativo.....	22
Figura 5 – Histograma - imagem capturada.....	23
Figura 6 – Histograma - imagem da base.....	23
Figura 7 - Tela inicial do aplicativo.....	24
Figura 8 - Funcionamento do Aplicativo.....	27
Figura 9 - Cálculo dos histogramas.....	29

LISTA DE ABREVIATURAS E SIGLAS

BSD - *Berkeley Software Distribution*

GPL - Licença Pública Geral

GUI - *Graphical User Interface*

IBGE - Instituto Brasileiro de Geografia e Estatística

IFSUL – Instituto Federal Sul-rio-grandense

OPENCV - *Open Source Computer Vision Library*

PC - I - Projeto de Conclusão I

PC2 – Projeto de Conclusão II

PNS – Pesquisa Nacional da Saúde

TA - Tecnologia Assistiva

SUMÁRIO

1.	INTRODUÇÃO.....	10
1.1	OBJETIVOS.....	10
1.1.1	Objetivo geral	10
1.1.2	Objetivos específicos.....	11
1.2	Organização do Trabalho de Conclusão.....	11
2	REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS	12
2.1	Tecnologias assistivas:.....	12
2.1.1	Problemas que o usuário com deficiência visual enfrenta:.....	13
2.1.2	Recursos utilizados para auxiliar as pessoas com deficiência visual:	14
2.1.3	Áudio descrição	19
3	RECURSOS UTILIZADOS PARA O DESENVOLVIMENTO DO PROTÓTIPO PROPOSTO.....	20
3.1	Sistema operacional Android	20
3.2	<i>Text-to-Speech</i>	20
3.3	<i>OpenCV</i>	24
4	DESENVOLVIMENTO DO APLICATIVO.....	27
4.1	Descrição do funcionamento do aplicativo.....	27
4.2	Processo de classificação as imagens	28
4.3	Armazenamento das imagens	29
5	VALIDAÇÃO E TESTES COM O APLICATIVO.....	30
6	CONSIDERAÇÕES FINAIS.....	31
	REFERÊNCIAS.....	32
	ANEXOS	34

1. INTRODUÇÃO

Segundo IGBE (2017), atualmente o Brasil possui mais de 207.983.210 de pessoas, e dentre essas, 12.894.957(6,2%) equivalem as pessoas que apresentam algum tipo de problema visual. Dentre os diferentes desafios que estas pessoas têm de enfrentar, está o reconhecimento de objetos. Nesse sentido, observou-se que o desenvolvimento de um aplicativo que realiza a captura de imagens, associado a um recurso de áudio-descrição, poderia auxiliar pessoas cegas a compreender melhor os objetos que lhe são apresentados. Assim, propõe-se a criação de um aplicativo que realiza a captura de imagens por meio de um dispositivo móvel e a sua classificação.

Esta pesquisa contempla o desenvolvimento de um aplicativo para realizar o reconhecimento e classificação de imagens (fotos) e realizar a descrição das mesmas para usuários com deficiência visual.

O desenvolvimento do aplicativo **I CAN SEE** busca contribuir para melhorar a qualidade de vida das pessoas com deficiência visual oportunizando as mesmas um meio para estabelecer o reconhecimento de fotos que lhe são apresentadas. Nesse sentido, busca-se contribuir para a inclusão do usuário com deficiência visual. Além disso esse estudo oportuniza avaliar a eficácia das tecnologias envolvidas no desenvolvimento do mesmo.

1.1 OBJETIVOS

Desenvolver um aplicativo para auxiliar pessoas com deficiência visual a reconhecer imagens representadas em fotos.

1.1.1 Objetivo geral

O Objetivo geral desta pesquisa é desenvolver um protótipo que visa auxiliar pessoas com deficiência visual a interagir com imagens que estão representadas nas fotos. O mesmo será desenvolvido inicialmente para ser utilizado na plataforma Android. O processo de classificação de imagens será realizado com auxílio dos recursos da biblioteca gráfica *OpenCV*.

1.1.2 Objetivos específicos

- Realizar a pesquisa bibliográfica com relação ao tema;
- Criar/avaliar a interface do aplicativo que será utilizada pelo usuário;
- Avaliar e implementar o uso do recurso de áudio descrição através da biblioteca *Text-To-Speech*;
- Avaliar a adequação da plataforma Android para o aplicativo proposto;
- Avaliar os recursos disponibilizados pela biblioteca gráfica *OpenCV* para realizar a extração de informações e classificação de imagens;
- Avaliar técnicas de classificação de imagens;
- Implementar algoritmo de classificação de imagens;
- Implementar recurso de áudio-descrição;
- Desenvolver o protótipo proposto;
- Realizar testes com usuários;

1.2 Organização do Trabalho de Conclusão

Este trabalho está organizado como se segue: O capítulo 2 apresenta o referencial teórico sobre tecnologias assistivas e as ferramentas que foram utilizadas para o desenvolvimento do protótipo. O capítulo 3 apresenta a arquitetura do aplicativo desenvolvido. O capítulo 4 apresenta a validação e testes com o aplicativo e por fim no capítulo 5 a conclusão.

2 REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS

As pessoas com deficiência visual necessitam de recursos que lhe permitam melhorar sua qualidade de vida. Nesse sentido, o desenvolvimento de um aplicativo pode vir a fazer diferença na vida das mesmas. O protótipo do aplicativo, denominado, **I CAN SEE**, pretende mudar o modo de como essas pessoas “enxergam o mundo”. Nessa seção serão apresentadas as características do público-alvo ao qual esse aplicativo se destina e serão descritos os recursos tecnológicos que serão utilizados na confecção do mesmo.

Atualmente o Brasil possui mais 208.200.350 milhões de pessoas e dentre essas, 12.908.421,7 milhões (6,2%) são pessoas que apresentam algum tipo de problema visual. Esse percentual representa um grande número que necessitam de recursos para melhorar sua qualidade de vida. O uso dos *smartphones* se popularizou muito no Brasil e no Mundo e tornou-se um recurso acessível para a população em geral (IBGE, 2017).

A Pesquisa Nacional de Saúde (PNS) considerou quatro tipos de deficiências: auditiva, visual, física e intelectual. O levantamento foi divulgado pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2017) e feito em parceria com o Ministério da Saúde.

Dentre os tipos de deficiência pesquisados, a visual é a mais representativa e atinge 3,6% dos brasileiros, sendo mais comum entre as pessoas com mais de 60 anos (11,5%). O grau intenso ou muito intenso da limitação impossibilita 16% dos deficientes visuais de realizarem atividades habituais como ir à escola, trabalhar e brincar. (VILELA, 2017).

Esse capítulo descreve a problemática das tecnologias assistivas e os recursos selecionados para o desenvolvimento da aplicação proposta.

2.1 Tecnologias assistivas:

Tecnologia Assistiva (TA) refere-se ao conjunto de artefatos disponibilizados às pessoas com necessidades especiais, que contribui para prover-lhes uma vida mais independente, com mais qualidade e possibilidades de inclusão social. Bersch e Tonolli (2006, p.1) identificam a TA como "todo o arsenal de Recursos e Serviços

que contribuem para proporcionar ou ampliar habilidades funcionais de pessoas com deficiência e consequentemente promover Vida Independente e Inclusão". O propósito das Tecnologias Assistivas reside em ampliar a comunicação, a mobilidade, o controle do ambiente, as possibilidades de aprendizado, trabalho e integração na vida familiar, com os amigos e na sociedade em geral" (SONZA et al, 2013, p.199).

Nesse sentido, a aplicação que está sendo proposta caracteriza-se como uma tecnologia assistiva que busca melhorar a qualidade de vida de pessoas cegas. Auxiliando-as na execução de atividades cotidianas e permitindo uma maior inclusão das mesmas.

2.1.1 Problemas que o usuário com deficiência visual enfrenta:

São muitos os problemas enfrentados pelos deficientes visuais que vão desde a locomoção até o desinteresse das autoridades em desenvolver uma política justa que os ajude com essa dificuldade.

- **Locomoção:** Para as pessoas que tem deficiência fica difícil a travessia e/ou locomoção por ruas e avenidas devido ao grande fluxo de veículos e pedestres. É certo que uma pessoa com problema visual, seja ele cego ou com alguma deficiência que o impossibilite de enxergar normalmente, sofre com muitos problemas e dentre eles estão as calçadas sem o piso tátil, buracos, irregularidades nos passeios, carrinhos de compras com cabo (esse a bengala não acusa, somente a roda do carrinho e por isso muitas das vezes a pessoa com deficiência visual acaba por esbarrar e fere a boca, nariz, rosto, etc.) e na maioria das cidades não possui o meio-fio da calçada que "mostra" para o deficiente visual onde termina a calçada.
- **Indiferença, desprezo e preconceito:** Pessoas com deficiências visuais são desprezadas frequentemente por outras pessoas que não dialogam com elas quando as mesmas estão acompanhadas por outra (guia), as pessoas anulam o deficiente visual e conversam apenas com o acompanhante, sendo que as perguntas que deveriam ser feitas para a pessoa com deficiência são feitas para o acompanhante.

- Um dos grandes problemas que o deficiente visual enfrenta diariamente é o auxílio nas tarefas do dia-a-dia, onde podemos citar como por exemplo: a identificação dos utensílios, localização dos móveis dentro da sua casa, objetos de pequeno porte como pratos, talheres, copos e local onde estão as suas roupas.

Nesse sentido, observa-se que o desenvolvimento de um aplicativo como o proposto neste trabalho pode promover a inclusão da pessoa cega e promover a sua autonomia.

2.1.2 Recursos utilizados para auxiliar as pessoas com deficiência visual:

Até hoje foram desenvolvidos diversos recursos que ajudam as pessoas com deficiências visuais. Entretanto algumas ferramentas são gratuitas e necessitam de auxílio de terceiros para que seja alimentada como por exemplo o **Be my eyes**. No entanto temos outras que são pagas tornando a sua aquisição impossível para pessoas de baixa renda. Abaixo citamos ferramentas que ajudam as pessoas com deficiências visuais e dentre os softwares desenvolvidos especificamente para auxiliar no reconhecimento de objetos pode-se destacar alguns exemplos que estão no mercado atualmente.

- **BlindTool:**

Criado pelo cientista da computação *Joseph Cohen*, pesquisador da Universidade de Massachusetts, o aplicativo reconhece objetos. Funciona da seguinte maneira: o usuário deve apontar o celular para seu entorno até senti-lo vibrar. Isso significa que o aplicativo detectou um objeto reconhecível e pode verbalizar qual é. Essa leitura de objetos tridimensionais é feita por uma rede neural artificial capaz de relacionar o que está diante da câmera do aparelho com imagens armazenadas em um banco de dados, buscando semelhanças. O sistema, claro, está sujeito a erros, mas é programado para descrever o objeto apenas se há possibilidade de ao menos 30% de acerto. O *BlindTool* é gratuito e está disponível para sistema Android no Google Play (OLIVEIRA, 2017).

- ***Be my eyes:***

Esse aplicativo é especialmente interessante, pois permite que pessoas que enxergam ajudem cegos a resolver problemas pontuais, como ler uma etiqueta, um rótulo, uma conta etc. Ao se cadastrar no sistema, o usuário pode atuar como voluntário ou como alguém que precisa de auxílio. Este envia imagens em vídeo do que precisa ver; a outra pessoa responde por escrito e o aplicativo verbaliza. *Be my eyes* pode ser baixado gratuitamente para iPhone no iTunes (BE MY EYES, 2017).

- ***Eye-D - Assistente para cegos:***

EyeD (acrônimo para o dispositivo do olho) é uma aplicação conceituada como solução autônoma para melhorar a qualidade de vida dos deficientes Visuais. É uma solução completa que auxilia os deficientes em sua vida diária. Este aplicativo ajuda os deficientes visuais “mostrando” a sua localização, navegarem para pontos de sua escolha próximos, avaliar o ambiente com a câmera de seu *smartphone* e ler texto um impresso (GOOGLE PLAY, 2017).

Também podemos citar outros softwares que permitem que as tarefas mais rotineiras dos deficientes visuais se tornem menos complexas e mais ágeis. Seguem abaixo alguns aplicativos desenvolvidos para resolver esse problema:

- ***Color ID:***

Este aplicativo está disponível para iPhone e Android, e o mesmo é capaz de reconhecer os mais variados tons de cores e verbalizar (em inglês) para o usuário. Visando ajudar as pessoas com baixa visão a descobrir, por exemplo, a cor da roupa que pretende usar, ou se uma fruta ainda não está madura (OLIVEIRA, 2017).

- ***IBraille Notes:***

Permite digitar anotações na tela do iPad ou iPhone e compartilhá-las diretamente em braile. Quando o usuário posiciona os dedos sobre a tela, teclas dinâmicas aparecem e com isso melhora o conforto do usuário (OLIVEIRA, 2017).

- **Ariadne GPS:**

O GPS foi especialmente desenvolvido para usuário com deficiência visual e ajuda a saber onde ele está e a seguir rotas. Quando o usuário passa o dedo sobre o mapa o aplicativo fala onde ele está e oferece as coordenadas para chegar ao destino. O celular vibra caso seja preciso atravessar um cruzamento e também sinaliza as paradas em ônibus em movimento (OLIVEIRA, 2017).

- **Ubook:**

O *Ubook* é uma audioteca com um acervo que conta com vários gêneros literários e mais de mil títulos. A ideia é parecida com a de serviços de streaming populares, e pode ser uma ótima saída para quem tem dificuldades para ler e os usuários podem baixar o aplicativo *Ubook* pela Internet, iOS ou Android, e salvar os livros que mais interessam dentro do catálogo. Um grande diferencial do software é a possibilidade que ele tem de compartilhar trechos de livros nas redes sociais para seus amigos (OLIVEIRA, 2017).

- **CPqD Alcance:**

O CPqD Alcance é um projeto da instituição brasileira disponível gratuitamente na loja do Google. O sistema é um guia completo para deficientes visuais, com narração automática da tela e com auxílio para quase todas as funções básicas e avançadas do celular. O sistema é de simples navegação e tem configurações que mantêm a privacidade permitindo que o usuário escreva sozinho. Após a instalação, o aplicativo se torna a interface padrão do celular sem a necessidade de cadastro do usuário (OLIVEIRA, 2017).

- **Liane TTS:**

O LianeTTS é um aplicativo (compilador) que analisa texto e o traduz em texto compilado no formato de difones (.pho) para processamento e síntese de voz pelo sistema mbrola (sintetizador de voz baseado na concatenação de difones). O

LianeTTS é uma aplicação software livre, que se comunica com o usuário através de síntese de voz. Pode ser acoplado a diversos programas para produzir a síntese de voz, seja através de um uso direto de suas rotinas, ou preferencialmente pelo uso do sistema *Speech Dispatcher*, que é uma camada de dispositivo independente para a síntese de voz que fornece uma interface de uso comum e fácil para ambas as aplicações clientes (programas que disponibilizam texto para a conversão) e softwares sintetizadores (convertem texto para fala) (ACESSIBILIDADE INCLUSIVA, 2017).

- **DosVox (gratuito):**

O sistema operacional DOSVOX permite que pessoas com deficiência visual utilizem um microcomputador comum (PC) para desempenhar uma série de tarefas, adquirindo assim um nível alto de independência no estudo e no trabalho (ACESSIBILIDADE INCLUSIVA, 2017).

- **Virtual Vision:**

O *Virtual Vision* foi desenvolvido em 1997 a partir de pesquisas da *MicroPower* com modelos de processamento de linguagem natural. É hoje o único software de leitura de telas desenvolvido nacionalmente capaz de funcionar sobre os aplicativos mais comuns utilizados na maior parte dos computadores (utiliza sistema operacional do Windows e reconhece Word, Excel, Internet Explorer, Outlook, MSN, Skype, entre outros) (ACESSIBILIDADE INCLUSIVA, 2017).

- **JAWS:**

Desenvolvido para usuários de computador, cuja perda de visão impede de ver o conteúdo da tela, o JAWS lê em voz alta o que está na tela do PC, tornando possível assim a compreensão do conteúdo que ele deseja acessar na máquina (ACESSIBILIDADE INCLUSIVA, 2017).

- **MAGIC:**

O *Magic* é um software que amplia a tela do computador para pessoas com baixa visão. Além de ampliar a tela em até 16 vezes, permite escolher entre diversas configurações visuais e formas de exibição para conseguir a melhor condição de visualização possível para a condição visual do usuário. Uma versão com voz permite ainda vocalizar textos da tela ao mesmo tempo em que esta é ampliada. Com isso o usuário consegue visualizar o conteúdo da tela da máquina (ACESSIBILIDADE INCLUSIVA, 2017).

- **NVDA (gratuito):**

O “*Non Visual Desktop Access*” ou NVDA, foi registrado sob a licença GPL (sigla em Inglês para “LICENÇA PÚBLICA GERAL GNU”), de autoria da Fundação para o Software Livre e adotada pelos sistemas GNU/Linux e outros. A tecla INS juntamente com a tecla N acessarão ao menu para configurar o funcionamento do NVDA (ACESSIBILIDADE INCLUSIVA, 2017).

- **MACDAISY:**

Baseado no padrão internacional Daisy – *Digital Accessible Information System* - a ferramenta brasileira traz sintetizador de voz (narração) e instruções de uso em português brasileiro. O software permite converter qualquer texto em formato *Daisy* e, após a conversão, é possível manusear o texto sonoro de maneira semelhante ao texto escrito. “O *Macdaisy* permite que o usuário folheie, consulte o índice, pesquise, faça comentários”, enumerou o analista de sistemas da UFRJ, João Sérgio Assis, que participou da equipe de desenvolvimento da ferramenta (ACESSIBILIDADE INCLUSIVA, 2017).

- **Braille Translator:**

Braille Translator é uma ferramenta online grátis para traduzir textos em Braille de até mil caracteres. Muito útil para escrever notas pequenas no sistema de leitura para cegos. O *Braille Translator* é bem fácil de usar. Basta escrever num

campo de texto e clicar no botão “Translate! ” que o programa traduz o conteúdo em uma imagem e em alfabeto Braille. Você pode salvar a imagem para imprimir em impressoras normais (depois basta furar os pontos), ou imprimir diretamente em uma impressora especial para Braille. *Braille Translator* ainda oferece a opção de copiar em formato ASCII para colar em um editor de textos com uma fonte Braille. Mesmo assim, ele pode ser muito útil para escrever pequenas mensagens e ajudar quem está aprendendo o sistema (ACESSIBILIDADE INCLUSIVA, 2017).

Como pode-se observar, diversos aplicativos e programas foram criados, entretanto a maioria precisa de algum periférico adicional tornando difícil para o deficiente visual o seu uso e esse foi o motivo pelo qual foi escolhido desenvolver um aplicativo *mobile*. Essa solução que está sendo proposta é acessível e a sua implementação não vai necessitar aquisição de nenhum recurso adicional além do *smartphone* para o usuário.

2.1.3 Áudio descrição

A áudio-descrição, uma forma de tradução visual semiótica, que traduz imagens em palavras, é hoje, compreendida como uma das mais importantes ajudas técnicas (tecnologia assistiva) para dar à pessoa cega ou com baixa visão, a acessibilidade comunicacional aos eventos visuais (LER PARA VER, 2017).

A Áudio descrição é fundamental para que o cego consiga compreender o ambiente e os objetos que se encontram no mesmo ambiente em que ele está. Bem sabemos que por não enxergar, a audição de uma pessoa com deficiência visual é muito aguçada.

No aplicativo proposto, será adicionado o recurso de áudio descrição visando auxiliar o usuário deficiente visual a compreender os objetos que lhes estão sendo apresentados.

3 RECURSOS UTILIZADOS PARA O DESENVOLVIMENTO DO PROTÓTIPO PROPOSTO

Nesse capítulo serão apresentados os recursos selecionados para o desenvolvimento do protótipo do aplicativo de reconhecimento de imagens em fotos ***I CAN SEE***.

3.1 Sistema operacional Android

Inicialmente pensou-se em desenvolver a aplicação para o sistema operacional Android. Android é o nome do sistema operacional baseado em Linux que opera em celulares (*smartphones*), *netbooks* e *tablets*. É desenvolvido pela *Open Handset Alliance*, uma aliança entre várias empresas, dentre elas a Google.

O funcionamento do Android é idêntico a outros sistemas operacionais (como Windows, Mac OS, Ubuntu, entre outros), cuja função é gerenciar todos os processos dos aplicativos e do hardware de um computador. A diferença é que o Android foi desenvolvido para ser utilizado em dispositivos móveis. Android é o sistema operacional mobile mais utilizado do planeta segundo (StatCounter, 2017), bem sabemos que há uma grande fatia no mercado de celulares e afins que usam essa tecnologia.

Os *smartphones* estão a cada dia que passa mais baratos, tecnológicos e são de fácil aquisição por pessoas de diversas classe sociais e é justo que o aplicativo seja desenvolvido para essa plataforma. O Sistema Operacional Android disponibiliza uma grande quantidade de bibliotecas que permitem o desenvolvimento de aplicativos através da linguagem Java. Por meio de uma extensa quantidade de bibliotecas, a linguagem permite interagir com os recursos dos *smartphones* e com outros *frameworks* como o *OpenCV* que será apresentado a seguir.

3.2 *Text-to-Speech*

O recurso *TextToSpeech* é utilizado para transformar o texto em áudio e o Android permite converter texto em voz em diversos idiomas. O Android fornece a classe *TextToSpeech* para essa finalidade. Sintetiza a fala do texto para reprodução

imediate ou para criar um arquivo de som. Uma instância *TextToSpeech* só pode ser usada para sintetizar o texto depois de concluir sua inicialização (ANDROID DEVELOPER, 2018). A Figura 1 apresenta um trecho de código que demonstra o uso da biblioteca no aplicativo *I CAN SEE*. No evento *OnInitListener* é chamado o recurso de áudio onde é dado primeiramente uma mensagem de boas vinda ao usuário e a partir desta mensagem segue as instruções de como proceder para utilizar da forma correta o aplicativo.

Figura 1 - Instruções para o usuário

```
text1 = new TextToSpeech(getApplicationContext(), new TextToSpeech.OnInitListener() {
    @Override
    public void onInit(int status) {
        text1.speak(text: "Bem vindo ao aplicativo Ai can see", TextToSpeech.QUEUE_ADD, params: null);
        text1.speak(text: "Para iniciar a câmera do celular pressione na tela", TextToSpeech.QUEUE_ADD, params: null);
        text1.speak(text: "Aponte para o objeto para capturar a imagem", TextToSpeech.QUEUE_ADD, params: null);
        text1.speak(text: "Mantenha o aparelho a uma distância de mais ou menos 10 centímetros da imagem", TextToSpeech.QUEUE_ADD, params: null);
        if (status != TextToSpeech.ERROR)
            text1.setLanguage(Locale.getDefault());
    }
});
```

Fonte: autor

Já na figura 2 é enviada uma mensagem para o usuário para que aguarde por alguns instantes informando que o aplicativo está processando a classificação das imagens. Nesse método as mensagens são enfileiradas.

Figura 2 - Saída do áudio Add

```
text1.speak(text: "Processo de classificação de imagens iniciado... " +
    "aguarde alguns instantes e finalizaremos o processo! ",
    TextToSpeech.QUEUE_ADD, params: null); //aqui que sai a voz
```

Fonte: autor

Já na figura 3 é enviada uma mensagem para o usuário para que aguarde por alguns instantes informando que o aplicativo está processando a classificação das imagens. Nesse método as mensagens não são enfileiradas.

Figura 3 - Saída de áudio *Flush*

```
text1.speak( text: "Processo de classificação de imagens iniciado... " +  
              "aguarde alguns instantes e finalizaremos o processo! ",  
            TextToSpeech.QUEUE_FLUSH, params: null);
```

Fonte: autor

A sintaxe do código é composta pelo texto que será reproduzido na forma de áudio e o método de gerenciamento das mensagens que são recebidas pelo dispositivo móvel que podem ser:

- **QUEUE_ADD** - Modo de fila em que a nova entrada é adicionada no final da fila de reprodução.
- **QUEUE_FLUSH** - Modo de fila onde todas as entradas na fila de reprodução (mídia a ser reproduzida e texto a ser sintetizado) são descartadas e substituídas pela nova entrada.

Os dois métodos que foram utilizados:

O método **QUEUE_FLUSH** foi utilizado para dar mensagens mais rápidas para o usuário como é o caso das mensagens de que indicam o resultado do processo de classificação das imagens e o método **QUEUE_ADD** foi utilizado nas demais situações.

Na figura 4 está sendo bloqueada a orientação do celular na posição RETRATO, para que não ocorra de o usuário sem querer girar a tela e causar falhas no aplicativo. Essa medida de precaução foi tomada para evitar problemas na captura de imagens e saída de áudio pois quando a tela era rotacionada o aplicativo apresentava a saída de voz repetindo a mesma mensagem de boas-vindas.

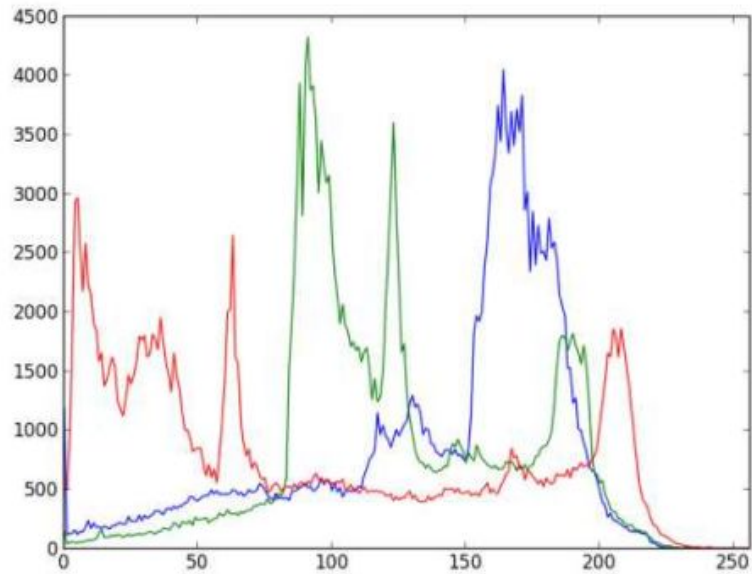
Figura 4 - Bloqueio da orientação do aplicativo

```
super.onCreate(savedInstanceState);  
setContentView(R.layout.activity_main);  
setRequestedOrientation(ActivityInfo.SCREEN_ORIENTATION_PORTRAIT);
```

Fonte: autor

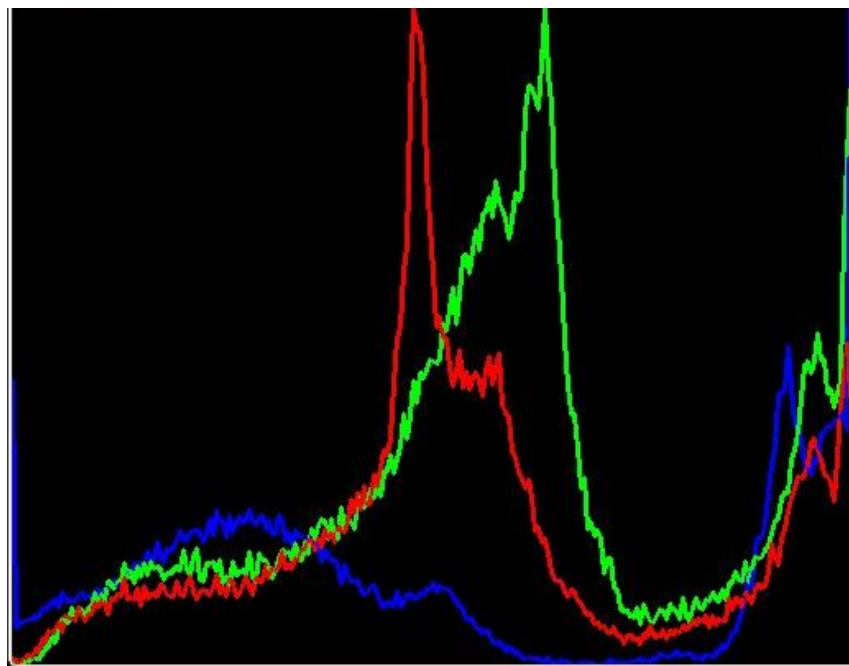
Na figura 5 estão os histogramas que são gerados pelo aplicativo **I CAN SEE** a partir das imagens capturadas pela câmera do *smartphone* e os histogramas das imagens que estão armazenadas na base de dados dentro do dispositivo.

Figura 5 – Histograma - imagem capturada



Fonte: autor

Figura 6 – Histograma - imagem da base



Fonte: autor

Na figura 5 está a tela principal do Aplicativo *I CAN SEE*, pois é nesta que o usuário vai pressionar o botão para abrir a câmera e iniciar a captura da imagem.

Figura 7 - Tela inicial do aplicativo



Fonte: autor

Após esse procedimento de abertura da câmera, o aplicativo vai iniciar a comparação das imagens e em seguida a mensagem de áudio informando se a foto foi encontrada e a descrição da mesma, senão a mensagem será - Imagem não localizada.

3.3 *OpenCV*

O *OpenCV* é uma biblioteca concebida especialmente para o processamento e análise de imagens em tempo real e implementa algoritmos bem estabelecidos na literatura de Processamento de Imagens, Visão Computacional e Inteligência

Artificial. A biblioteca, originalmente desenvolvida pela Intel em 2000, tem uma implementação multiplataforma, totalmente livre para o uso acadêmico e comercial. Ela possui módulos de Processamento de Imagens e Vídeo, estrutura de dados, álgebra linear, Interface Gráfica (GUI) com sistema de janelas independentes, controle de mouse e teclado, além de mais de 350 algoritmos de Visão Computacional (Processamento e Análise) como: filtros de imagem, calibração de câmera, reconhecimento de objetos, análise estrutural e outros. A biblioteca possui versões disponíveis em Python, C++, C, Java e Android. Através desta biblioteca é possível desenvolver aplicações de processamento de imagens em tempo real (RODCOSTA, 2017).

A biblioteca disponibiliza recursos para determinar a similaridade entre imagens o que permite classificá-las com base no comparativo das características encontradas nas mesmas. Entre as técnicas utilizadas para realizar esse comparativo está o uso de histogramas que permite comparar as características dos pixels que compõe a imagem.

Os recursos da biblioteca gráfica *OpenCV* oportunizam o desenvolvimento de uma aplicação dessa natureza diretamente no *smartphone* ou através da interação do mesmo com um servidor de imagens. A presente pesquisa foi desenvolvida seguindo essa segunda alternativa pois é possível criar um banco mais robusto de imagens enquanto garante um melhor desempenho do seu processo de classificação.

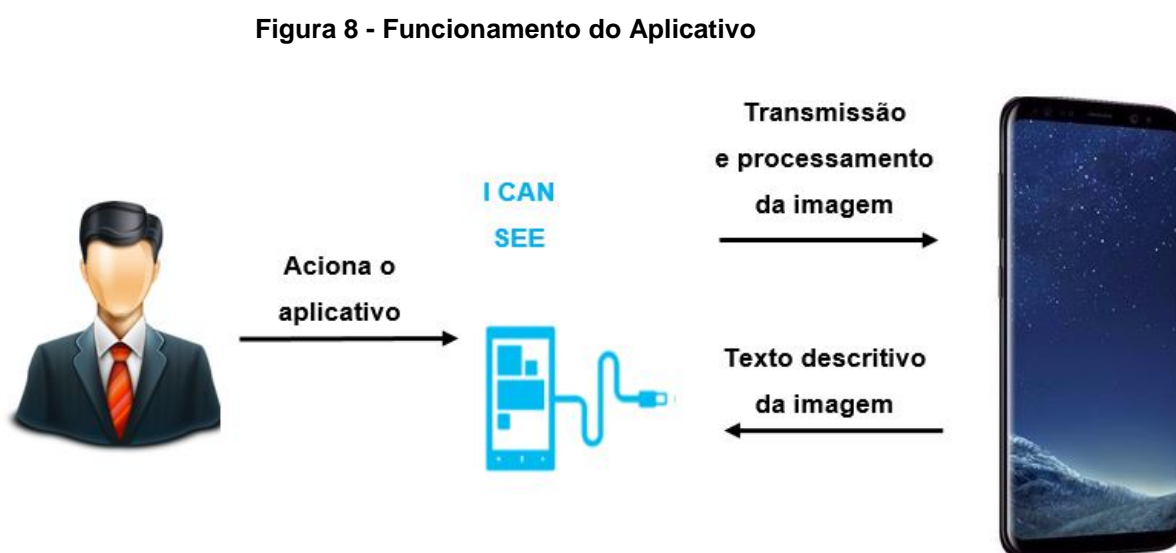
A biblioteca *OpenCV* possui mais de 500 funções. Foi idealizada com o objetivo de tornar a visão computacional acessível a utilizadores e programadores em áreas tais como a interação humano-computador em tempo real e a robótica. A biblioteca está disponível com o código fonte e os executáveis (binários) otimizados para os processadores Intel. Um programa *OpenCV*, ao ser executado, invoca automaticamente uma DLL (*Dynamic Linked Library*) que detecta o tipo de processador e carrega, por sua vez, a DLL otimizada para este. Juntamente com o pacote *OpenCV* é oferecida a biblioteca IPL (*Image Processing Library*), da qual a *OpenCV* depende parcialmente, além de documentação e um conjunto de códigos exemplos. A biblioteca está dividida em cinco grupos de funções: Processamento de imagens; Análise estrutural; Análise de movimento e rastreamento de objetos; Reconhecimento de padrões e Calibração de câmara e reconstrução 3D.

Dentre muitos recursos que a biblioteca *OpenCV* oferece, foi utilizado o recurso de **HISTOGRAMAS** para o desenvolvimento do protótipo do aplicativo **I CAN SEE**.

4 DESENVOLVIMENTO DO APLICATIVO

O aplicativo **I CAN SEE** foi desenvolvido utilizando a linguagem Java com Android, a biblioteca gráfica *OpenCV* e o recurso de áudio-descrição *Text-to-speech*. A arquitetura do aplicativo será descrita na seção a seguir e a figura 6 ilustra essa informação.

4.1 Descrição do funcionamento do aplicativo



Fonte: autor

O aplicativo foi criado no intuito de facilitar ao máximo a vida do usuário, para acessar o aplicativo com um click (pressiona na tela) o usuário já inicia o processo de captura de imagens. Através da biblioteca *Text-to-Speech* foi desenvolvido um recurso de áudio que auxilia o usuário em todo o uso do aplicativo.

Como a aplicação funciona: Inicialmente foi criada uma base de dados composta por um conjunto de imagens, essas imagens são de fotos que representam situações reais vivenciadas pelo usuário cego. Para iniciar o uso da aplicação o usuário deve iniciar a aplicação no *smartphone*. Uma mensagem de áudio será reproduzida para indicar que aplicação foi iniciada. O usuário então é informado pelo aplicativo que deve posicionar a câmara do *smartphone* para a foto

que deseja capturar. Após a captura da imagem da câmera o processo de classificação é iniciado. A imagem capturada é comparada com as imagens da base de dados previamente cadastrada. O aplicativo tem uma base de dados e a câmera captura a imagem, realiza a comparação da imagem capturada com todas as imagens da base de dados até que seja encontrada a imagem que mais se aproxima da capturada. O tempo de processamento é de aproximadamente alguns milissegundos, pois neste protótipo as imagens cadastradas são poucas.

4.2 Processo de classificação as imagens

Quando o usuário executa o aplicativo a câmera do *smartphone* é aberta e a partir desse momento ao pressionar na tela é capturada uma imagem. Neste momento a imagem capturada é convertida para o formato Bitmap e armazenada em uma variável temporária. Após realiza-se o cálculo dos canais RGB (*Red Green Blue*) gerando um histograma dos valores das cores presentes na imagem (Figura 7). Esse histograma é então comparado com os histogramas das imagens presentes em uma base dos dados pré-cadastrada que contém imagens dos objetos que poderiam ser encontrados. Nesse caso, a base de imagens é formada por um conjunto de fotos que representam diferentes momentos da vida do usuário cego. Estabelece-se na sequência o comparativo do histograma da imagem capturada e os histogramas das imagens da base de dados. O comparativo resulta em valores que representam a distância entre os pontos dos canais RGB das duas imagens, gerando um valor final que resulta no somatório de todas as distâncias [OpenCV, 2018]. A partir desse cálculo seleciona-se a imagem da base de dados cuja a distância é a menor em relação a imagem capturada. Para minimizar a quantidade de erros foi estabelecido um ponto de corte (5000). Esse ponto de corte foi determinado após efetuados diversos testes com aproximadamente 50 imagens e esse foi o valor que mais reconheceu as fotos armazenadas no repositório. Portanto esse valor é o que determina se a imagem capturada pelo aplicativo foi encontrada ou não.

4.3 Armazenamento das imagens

Neste protótipo, a base de dados será localmente em uma pasta dentro do *smartphone*, as fotos serão no formato jpg e a base de dados será alimentada por uma pessoa normal que vai cadastrar previamente as imagens através da captura pela câmera do *smartphone*.

Figura 9 - Cálculo dos histogramas

```
//Cria os histogramas para comparar
Mat hist1 = new Mat();
Mat hist3 = new Mat();
MatOfInt histSize = new MatOfInt( ...a: 180);
MatOfInt channels = new MatOfInt( ...a: 0);
ArrayList<Mat> bgr_planes1= new ArrayList<Mat>();
ArrayList<Mat> bgr_planes3= new ArrayList<Mat>();
Core.split(img1, bgr_planes1);
Core.split(img3, bgr_planes3);
MatOfFloat histRanges = new MatOfFloat ( ...a: 0f, 180f);
boolean accumulate = false;

//Calcula o histograma da imagem capturada da câmera
Imgproc.calcHist(bgr_planes1, channels, new Mat(), hist1, histSize, histRanges, accumulate);
Core.normalize(hist1, hist1, alpha: 0, hist1.rows(), Core.NORM_MINMAX, dtype: -1, new Mat());

//Calcula o histograma da imagem lida do diretório
Imgproc.calcHist(bgr_planes3, channels, new Mat(), hist3, histSize, histRanges, accumulate);
Core.normalize(hist3, hist3, alpha: 0, hist3.rows(), Core.NORM_MINMAX, dtype: -1, new Mat());
img1.convertTo(img1, CvType.CV_32F);
img3.convertTo(img3, CvType.CV_32F);
hist1.convertTo(hist1, CvType.CV_32F);
hist3.convertTo(hist3, CvType.CV_32F);

//Compara os 2 histogramas gerados e armazena o resultado na variável calibrar
double calibrar = Imgproc.compareHist(hist1, hist3, Imgproc.CV_COMP_CHISQR);
Log.d( tag: "Compara Imagem", msg: "calibrar: "+calibrar);
if(calibrar>0 && calibrar<5000) { // Essa variável é para calibrar os valores gerados.
```

Fonte: autor

Nos testes efetuados a abordagem escolhida provou-se adequada para os objetivos propostos que era de fazer com o que o usuário cego conseguisse utilizar o aplicativo e o mesmo reconhecesse a imagem ou não.

5 VALIDAÇÃO E TESTES COM O APLICATIVO

Para realizar a validação do aplicativo criado foi realizado um teste com um usuário cego. Inicialmente foi lido o termo livre e esclarecido (Anexo 1) para deixar o usuário ciente dos propósitos da pesquisa. Após realizou-se um questionário pré-teste (Anexo 2) para delinear o perfil do usuário. Segundo o **A1** encontra-se faixa-etária de **46 a 50 anos**, possui cegueira adquirida a mais de **25 anos** e possui experiência limitada com *smartphones*. Sua expectativa com relação ao uso do aplicativo era muito importante pois como ele não nasceu com essa deficiência ele poderia imaginar o que o aplicativo estava lhe informando através da áudio-descrição

O teste com o usuário teve alguns problemas no início, pois o mesmo nunca havia utilizado um *smartphone* até o momento do teste. Mas após algumas explicações ele conseguiu operar o *smartphone* normalmente sem ajuda. No começo do teste ele foi auxiliado por sua companheira como devia proceder e depois de várias tentativas obteve sucesso no uso do aplicativo. Visto que se familiarizou com o mesmo, gostou do aplicativo e achou de fácil utilização. Essa informação pode ser observada na página de anexos.

O reconhecimento das fotos pelo aplicativo operado pelo usuário cego foi difícil no início e precisou de orientações, mas depois destas ele conseguiu gradativamente a cada utilização, operar melhor até que nas últimas ele conseguiu fazer com que o aplicativo obtivesse sucesso reconhecendo as fotos, pois, o foco da câmera sobre a imagem estava correto em relação as primeiras tentativas.

Após foi aplicado um questionário pós-Teste (Anexo 2) no qual o usuário respondeu as questões que lhe foram propostas. O usuário considerou que o aplicativo desenvolvido lhe auxiliou na tarefa que foi solicitada.

6 CONSIDERAÇÕES FINAIS

Podemos concluir que o desenvolvimento do presente estudo possibilitou o desenvolvimento de um sistema capaz de reconhecer objetos em fotos e uma análise de como um aplicativo pode melhorar muito a vida de uma pessoa com deficiência visual. Além disso, foi possível realizar testes com um usuário real e com isso foi obtido dados mais consistentes sobre a usabilidade, eficiência e eficácia do aplicativo.

Findando esse trabalho, fica implícito que os objetivos propostos - que era trabalhar com o reconhecimento de imagens em foto - foram alcançados. A metodologia aplicada foi eficaz com relação aos testes com o usuário, que se adequou perfeitamente ao uso do aplicativo no final do teste. As dificuldades foram muitas, mas com esforço e dedicação todas foram superadas. Citando uma das maiores dificuldades que foi fazer com que o algoritmo de comparação funcionasse perfeitamente obtendo o valor necessário para a comparação das imagens (calibração). A expectativa era que o *smartphone* reconhecesse a imagem capturada pela câmera, transmitisse em forma de áudio o resultado para o usuário pois era o objetivo almejado e foi atingido. No entanto, adicionar um recurso para repetir as mensagens para que o usuário fique orientado como proceder, pois, como ele não enxerga fica difícil saber o que está ocorrendo na tela do *smartphone* e com isso facilitar o seu aprendizado na utilização do aplicativo.

Fica para um projeto futuro, um estudo mais detalhado sobre outros métodos de classificação, melhoras na interface do aplicativo, desenvolvimento para outras plataformas do mercado, como por exemplo, o IOS e também a implementação do recurso de reconhecimento do comando de voz, um recurso extremamente importante, de grande facilidade e de fácil utilização pelo usuário com deficiência visual.

REFERÊNCIAS

Acessibilidade. Acessibilidade Inclusiva. Disponível em:< <http://www.acessibilidadeinclusiva.com.br/programas-para-computadores/> >. Acesso em: 02 nov. 2017.

Android Studio. O que é Android. Disponível em:< <https://www.significados.com.br/android/> >. Acesso em: 02 nov. 2017.

ANDROID DEVELOPER. Disponível em:< <https://developer.android.com/> >. Acesso em: 19 set. 2017.

Áudio-descrição: Opinião, Crítica e Comentários. Técnicas para Áudio-descrição: Acessibilidade e Usabilidade na Penn State. Disponível em:< <http://www.lerparaver.com/lpv/tecnicas-audio-descricao-acessibilidade-usabilidade-penn> >. Acesso em: 08 out. 2017.

BERSCH, R.; TONOLLI, J. C. Tecnologia Assistiva. 2006. Disponível em: < <http://www.assistiva.com.br/> >. Acesso em: 03 set. 2017.

BY MY EYES. Disponível em:< <http://bemyeyes.com> >. Acesso em: 10 dez. 2017.

BORTOLINI, Sirlei; FERNANDES, Woqiton Lima. Acessibilidade e Tecnologia Assistiva: pensando a inclusão sociodigital de pessoas com necessidades especiais. Bento Gonçalves - RS, 2013.

GOOGLE PLAY. Assistente para cegos. Disponível em:< https://play.google.com/store/apps/details?id=in.gingermind.eyed&hl=pt_BR >. Acesso em: 10 nov. 2017.

IBGE. Projeção da população do Brasil e das Unidades da Federação. Disponível em:< <http://www.ibge.gov.br/apps/populacao/projecao/>>. Acesso em: 11 set. 2017.

OpenCV. Docs master. Disponível em:< <http://opencv.org/>>. Acesso em: 17 set. 2017.

OpenCV. Biblioteca OpenCV. Disponível em:< <http://rodcosta.eadti.com.br/tutoriais/opencv> >. Acesso em: 02 nov. 2017.

OLIVEIRA, Hamilton. Tecnologias inovadoras para deficientes visuais. Disponível em:< <http://www.casadaptada.com.br/2016/11/7-tecnologias-inovadoras-para-deficientes-visuais/> >. Acesso em: 25 set. 2017.

ROS. ROS (Robot Operating System). Disponível em:< http://wiki.ros.org/pt_BR >. Acesso em: 02 nov. 2017.

SONZA, Andréa Poletto; KADE, Adrovane; FAÇANHA, Agebson; REZENDE, André Luiz Andrade; NASCIMENTO, Gleison Samuel do; ROSITO, Maurício ovolan; BOROTOLINI, Sirlei; FERNANDES, Woqiton Lima. Acessibilidade e Tecnologia

Assistiva: pensando a inclusão sociodigital de pessoas com necessidades especiais. Bento Gonçalves - RS, 2013.

TextToSpeech. Android - texto para fala. Disponível em:<https://www.tutorialspoint.com/android/android_text_to_speech.htm >. Acesso em: 02 Jun. 2018.

VILELA, Flavia. IBGE: 6,2% da população têm algum tipo de deficiência. Disponível em: <<http://www.ebc.com.br/noticias/2015/08/ibge-62-da-populacao-tem-algum-tipo-de-deficiencia>>. Acesso em 05 set. 2017

ANEXOS

Questionário aplicado para usuário cego:

Antes do teste:

- 1) Faixa-etária:
 - a) 18 a 25 anos
 - b) 26 a 35 anos
 - c) 36 a 45 anos
 - d) 46 a 50 anos
 - e) mais de 50 anos
- 2) No que se refere à sua deficiência visual, você possui cegueira ou baixa visão?
- 3) Sua deficiência é congênita ou adquirida? Se for adquirida, há quanto tempo você se considera deficiente visual?
- 4) Você faz uso de tablet, *smartphone*, ou dispositivos similares? Em caso afirmativo, qual o sistema operacional (Android, iOS, Windows phone, etc.)?
- 5) Que tipo de informações você espera de um aplicativo que possa lhe dar no reconhecimento das fotos?

Após o teste:

- 1) Você considerou o aplicativo de fácil uso? Se você teve dificuldades para usar o aplicativo?
- 2) Você considera que é fácil de aprender a utilizar o aplicativo?
- 3) Teria sugestões para melhorar o aplicativo?

Termo de Consentimento Livre e Esclarecido

- 1) Você está sendo convidado para participar da pesquisa intitulada ***I CAN SEE: UMA PROPOSTA DE SOFTWARE DE RECONHECIMENTO DE OBJETOS PARA AUXILIAR PESSOAS COM DEFICIÊNCIA VISUAL*** realizada pelo aluno **Joel Zatti** do curso **TSPI** sob orientação do professor **João Mário Lopes Brezolin**. Essa pesquisa tem por objetivo avaliar a **eficiência e eficácia do aplicativo *I CAN SEE***.

- 2) Sua participação nesta pesquisa consistirá em responder a um questionário com questões relacionadas à sua experiência no uso **do aplicativo *I CAN SEE***.

- 3) Os dados não serão divulgados de forma a possibilitar sua identificação.

Declaro que tenho mais de dezoito (18) anos e entendi os objetivos, riscos e benefícios de minha participação na pesquisa e concordo em participar.

Assinale a concordância (ou não) com relação ao termo de consentimento:

Concordo Discordo

Nome: _____

Assinatura: _____