

18º Congresso Nacional de Iniciação Científica

TÍTULO: ESTUDO E IMPLEMENTAÇÃO DE UM SISTEMA DE RECONHECIMENTO DE FACES UTILIZANDO OS SERVIÇOS COGNITIVOS DA MICROSOFT

CATEGORIA: CONCLUÍDO

ÁREA: CIÊNCIAS EXATAS E DA TERRA

SUBÁREA: Computação e Informática

INSTITUIÇÃO(ÕES): FACULDADE DE TECNOLOGIA DE SÃO JOSÉ DO RIO PRETO - FATECRP

AUTOR(ES): LUIS HENRIQUE BRUNERI, LEONARDO JESUS FALCO

ORIENTADOR(ES): HENRIQUE DEZANI

1. Resumo

O reconhecimento facial é umas das subáreas da Inteligência Artificial que tem como objetivo reconhecer faces de pessoas em imagens ou vídeos. Um exemplo são os sistemas de segurança que podem utilizar esses recursos para identificar se uma pessoa está ou não presente em um ambiente. Neste contexto é importante frisar as diferenças entre as técnicas de detecção e reconhecimento facial. Enquanto a primeira somente indica se uma face está presente em uma imagem, a segunda técnica tem o objetivo de dizer de quem é a face detectada. Com o avanço das tecnologias em nuvem, surgiram também novos serviços que permitem realizar os serviços cognitivos, como da visão, que podem ser consumidos a partir de programas de computadores ou para dispositivos móveis. A Microsoft Azure é uma das plataformas que possui este serviço disponível e o qual é alvo da pesquisa aqui proposta, ou seja, objetiva-se com este projeto realizar um estudo sobre os principais conceitos inerentes aos serviços cognitivos de visão, tais como Inteligência Artificial e sua subárea, Aprendizagem de Máquina, assim como as tecnologias que permitem realizar a interação com este serviço, tais como Web Services, Web Socket e Programação para Dispositivos Móveis.

2. Introdução

Com o advento de tecnologias emergentes em Inteligência Artificial e Aprendizagem de Máquina, tem-se tornado comum a criação de *Software* como Serviço (SaaS – Software-as-a-Service) para realizar atividades até então exclusivas de filmes de ficção científica. Dentre estas atividades, destacam-se os serviços cognitivos de visão, fala e conhecimento, disponíveis na plataforma da Microsoft Azure [Microsoft, 2018].

O uso de uma câmera e o consumo de um serviço na *web* bastam para identificar uma pessoa, automaticamente, por meio do reconhecimento facial. Graficamente, nosso rosto é composto por muitos pontos e algumas medidas, como o comprimento da linha da mandíbula, tamanho do crânio, distância entre os olhos, largura do nariz, entre outros [OKABE, 2014].

Além dos serviços disponibilizados na *web*, se faz necessário um computador capaz de receber as imagens de uma câmera e processá-las a fim de atuar sobre o

ambiente. Os dispositivos móveis, além de estarem incorporados ao nosso dia a dia e facilitar a mobilidade de nossas atividades, suprem esta necessidade de maneira satisfatória, uma vez que possuem sistemas embarcados de câmera, processamento e conexão com a Internet muitas vezes com ótimo desempenho.

Segundo a Fundação Getúlio Vargas (FGV, 2018), o Brasil terá um *smartphone* em uso por habitante até o final de 2017, totalizando 208 milhões de aparelhos. Isso se deve ao fato de que o *smartphone* foi incorporado de tal forma à rotina das pessoas que é encarado como essencial.

Os usuários estão cada vez mais conectados, dadas as facilidades de compra e custos mais acessíveis de aparelhos e planos de internet 3G e 4G, e buscam por aplicativos de qualidade, preferencialmente, gratuitos. De acordo com relatório anual da eBit (2018), foi registrado no primeiro semestre de 2017 um aumento das vendas via *smartphones* de 56,2%. Isso mostra que os consumidores continuam investindo em aparelhos com mais funcionalidades e modernos, o que viabiliza esta migração do comportamento de consumo.

A indústria de aplicativos móveis tem se desenvolvido rapidamente no mundo e, só no Brasil, cresceu 83% nos últimos anos. De acordo com a ComScore (2018), o *mobile* é responsável por mais da metade de todos os minutos digitais em 13 mercados e por mais de 75% no México, Índia e Indonésia, tendo o Brasil 73% do total de minutos digitais no *mobile*, sendo que mais de 25% dos usuários globais só utilizam o *mobile*.

Portanto, a oferta de programas para dispositivos móveis hoje é imensa e, por isso, os aplicativos precisam ser relevantes para se destacar. E, para isso, precisam simplificar atividades diárias, como está sendo proposto neste projeto.

3. Objetivo

Pelo exposto, objetiva-se neste projeto elaborar um sistema capaz de permitir o cadastro de alunos e a realização da chamada por meio de foto dos alunos presentes em sala de aula, agilizando o processo monótono e trazendo mais segurança no que se refere à consistência dos dados informados.

O presente projeto justifica-se pela necessidade constante de agilizar os processos manuais existentes dado o surgimento de tecnologias emergentes capazes de baratear a implantação e gerar inovação em termos de Inteligência

Artificial que nos permitam realizar tal atividade de reconhecimento de alunos de maneira não intrusiva.

4. Metodologia

O trabalho contempla uma pesquisa exploratória prática e visa entregar um produto funcional a partir do embasamento teórico obtido. O trabalho foi dividido 4 atividades, descritas a seguir.

Atividade 1: Levantamento bibliográfico. Realizou-se uma revisão bibliográfica a partir de livros, artigos, páginas *web* a fim de auxiliar a realização do trabalho e identificar o estado da arte no que condiz o uso da Inteligência Artificial aplicado ao reconhecimento de pessoas.

Atividade 2: Estudou-se a plataforma Microsoft Azure Cognitive Services. Nesta atividade, realizou-se a implementação do serviço cognitivo de visão utilizando o serviço disponibilizado pela Microsoft usando imagens coletadas dos alunos voluntários do projeto.

Atividade 3: Estudou-se e implementou-se uma aplicação para dispositivos móveis. Nesta atividade foi utilizado o *framework* Ionic [Ionic, 2018], suas APIs e recursos nativos a fim de realizar o desenvolvimento de um aplicativo capaz de consumir o serviço estudado na atividade 2.

Atividade 4: Testes. Nesta atividade realizou-se teste dos itens desenvolvidos nas etapas anteriores a fim de confirmar a hipótese de atuação de sistemas em tempo real para a realização das chamadas de aluno em sala de aula.

5. Desenvolvimento

A fim de demonstrar a aplicação dos estudos realizados, foi desenvolvida uma aplicação *mobile* para exemplificar o uso do serviço de reconhecimento facial do Microsoft Azure. A aplicação desenvolvida tem como objetivo auxiliar professores a realizarem controle de presença de seus alunos com base em uma foto geral da

turma. Obtendo a foto a partir do aplicativo desenvolvido utilizando o framework Ionic, utiliza-se do serviço de Inteligência Artificial da plataforma Microsoft Azure, a qual faz o reconhecimento das faces e identificação das faces a fim de exibir a lista de alunos presentes na sala de aula.

Na Figura 1 é apresentada a tela do aplicativo contendo a listagem das turmas cadastradas no sistema e o botão para realizar o cadastro de uma nova turma. O cadastro da turma na base de dados do Azure utiliza o método *Create* da API *PersonGroup*, e é responsável por criar um recipiente que irá conter os dados dos alunos cadastrados, incluindo as imagens faciais e recursos de reconhecimento.

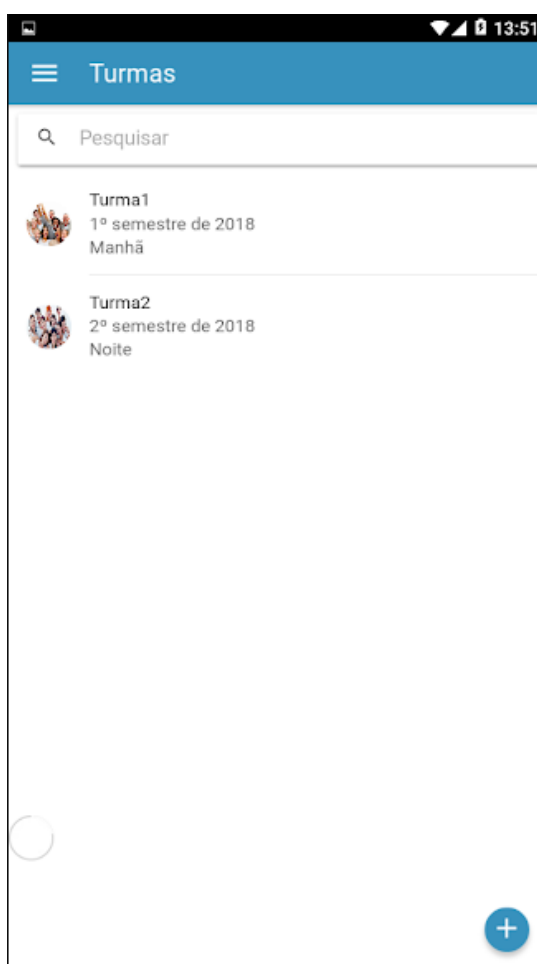


Figura 1. Lista de Turmas (GroupFace)

Todas as requisições feitas na plataforma Azure são realizadas a partir de chamadas HTTP utilizando estilo REST. A criação de um grupo, por exemplo, é realizada a partir da chamada do método HTTP PUT utilizando a URL:

<https://brazilsouth.api.cognitive.microsoft.com/face/v1.0/persongroups/{ID}>, sendo o

ID uma variável para identificar o grupo. Além do ID, no corpo da mensagem HTTP é enviada também o parâmetro *name*, que corresponde a um identificador textual do grupo.

Na Figura 2 é apresentada a tela do aplicativo referente ao cadastro de um novo aluno na turma.



Figura 2. Cadastro de um aluno na Turma

Para cadastro do aluno na base de dados do Azure realizou-se uma chamada à URL abaixo por meio de uma solicitação HTTP POST, informando o ID do grupo criando anteriormente e o nome do aluno no corpo da mensagem.

<https://brazilsouth.api.cognitive.microsoft.com/face/v1.0/persongroups/{ID}/persons>

Uma vez cadastrado o aluno, realiza-se então o cadastro das fotos deste aluno na base de dados do serviço cognitivo de face da Azure, utilizando a URL

<https://brazilsouth.api.cognitive.microsoft.com/face/v1.0/persongroups/{ID}/persons/{PID}/persistedFaces>.

Na qual é informado o ID do grupo (ID) e o Id do aluno cadastrado anteriormente (PID). A imagem é fornecida a partir de uma URL gerada pelo sistema de armazenamento de fotos utilizado, neste projeto, o Firebase Storage. A imagem de um rosto de uma pessoa em um grupo de pessoas é utilizada para uma futura identificação e verificação no reconhecimento facial.

Após o processo descrito anteriormente é necessário realizar o treinamento da do algoritmo de aprendizagem de máquina fornecido pelo serviço de reconhecimento facial da plataforma Microsoft Azure. Este algoritmo não é de conhecimento público, mas entende-se que dentro de cada grupo de pessoas criado, o algoritmo identifica padrões das faces que foram cadastradas para, somente assim, realizar o reconhecimento facial dentro dessa turma. Para esse treinamento utilizou-se a chamada à URL abaixo usando-se uma solicitação HTTP POST.

<https://brazilsouth.api.cognitive.microsoft.com/face/v1.0/persongroups/{ID}/train>

Uma vez realizado o cadastro do grupo, das pessoas no grupo, incluindo imagens de seus rostos, e o treinamento, é possível utilizar o sistema para o reconhecimento e identificação das faces cadastradas. Na Figura 3 (esquerda) é apresentada a tela do sistema que exibe a foto da turma tirada por meio do próprio aplicativo.

A imagem é novamente salva no serviço de hospedagem de arquivos do Firebase e, de maneira síncrona, seu link é enviado para o serviço de reconhecimento facial da Azure utilizando a chamada por HTTP POST da URL:

<https://brazilsouth.api.cognitive.microsoft.com/face/v1.0/identify>

No corpo da chamada é enviada um objeto JSON (JavaScript Object Notation) contendo o ID do Grupo (*personGroupId*), o número máximo de usuários reconhecido por face (*maxNumOfCandidatesReturned*), neste caso, 1, e o limiar de

confiabilidade (*confidenceThreshold*), neste caso, de 0.5, ou seja, ele retornará como encontrado um aluno na imagem apenas quando a probabilidade de ser esta pessoa for maior ou igual a 0.5.

Após realizada a identificação das faces, o método HTTP solicitado anteriormente retorna uma Array JSON com os IDs das pessoas (alunos) encontrados na imagem. A partir destes ID é possível realizar uma consulta na base de dados e indicar, numa lista, os alunos presentes na imagem, conforme pode ser visto na tela também apresentada na Figura 3 (direita).

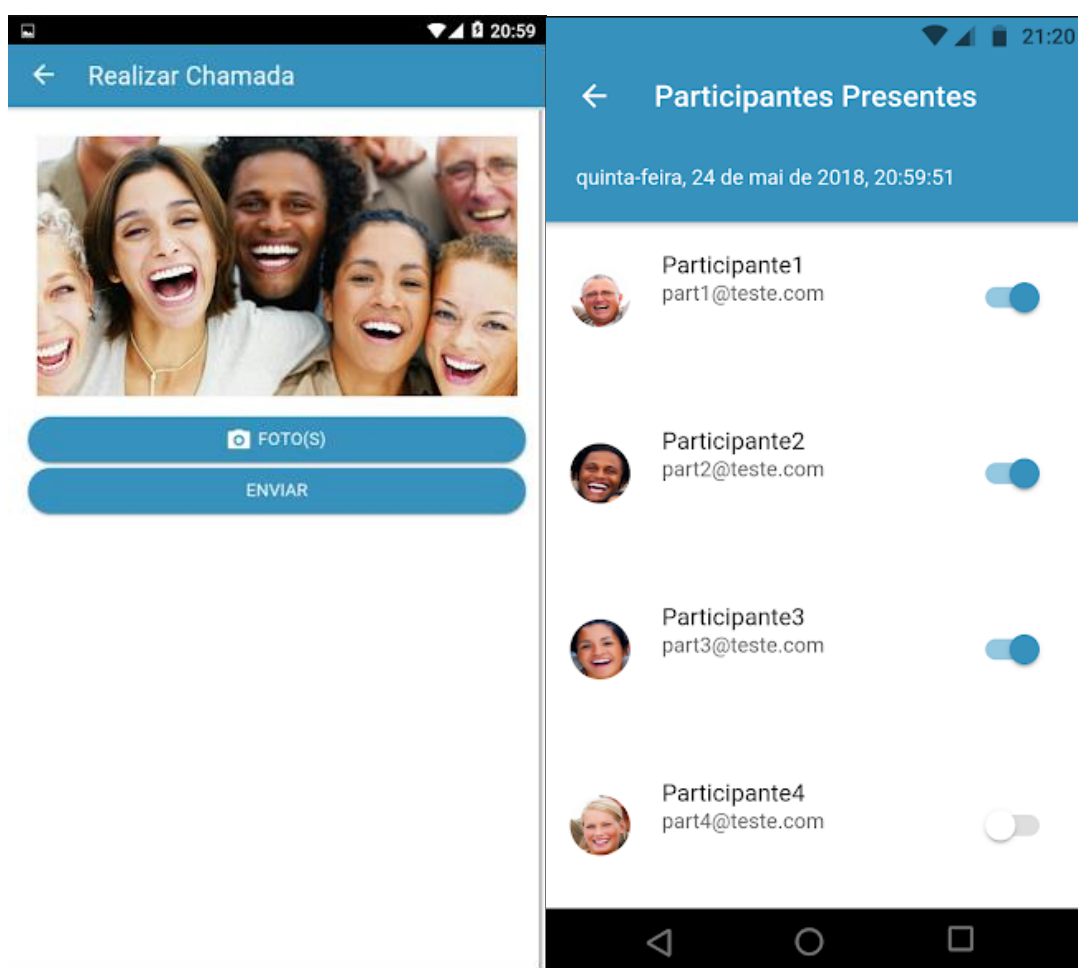


Figura 3. Telas com a foto para identificação das faces (esquerda) e a lista de alunos com as faces identificadas (direita)

6. Resultados

Após a conclusão do desenvolvimento da aplicação, esta foi testada em uma escola da cidade de São José do Rio Preto, que não será divulgada por questões de

privacidade. Os testes foram realizados nas turmas do 6º ano do ensino fundamental e no 1º ano do ensino médio. As imagens dos testes também não serão divulgadas para manter a privacidade das pessoas presentes nas imagens, portanto os resultados serão apresentados de maneira escrita e explicativa.

Antes de iniciar os testes *in loco* na instituição, foi disponibilizado pelo diretor as imagens dos alunos (carômetro) das turmas relatadas anteriormente para que fosse feito o cadastro prévio destas turmas. Vale ressaltar que havia apenas uma imagem de cada aluno, ou seja, os alunos foram cadastrados com o mínimo possível para o reconhecimento facial.

Após o cadastro das turmas, já no local de teste, o início dos testes foi feito a partir do 6º ano do fundamental, no qual foram realizadas as chamadas pelo aplicativo (reconhecimento facial) utilizando apenas uma foto de toda a turma, em seguida duas fotos da turma e por fim diversas fotos da turma.

Já no 1º ano do ensino médio, além do carômetro, também foram tiradas mais duas fotos de cada aluno para que se completasse o máximo para o reconhecimento facial e assim, como na turma anterior, foram realizadas as chamadas pelo aplicativo (reconhecimento facial) com apenas uma foto de toda a turma, outras com duas fotos da turma e outras com mais de duas fotos da turma.

Terminada essa fase de captura das imagens necessárias para teste, as mesmas foram submetidas à API através da aplicação, e o resultado foi satisfatório mas com alguns pontos a serem ressaltados:

Para faces totalmente visíveis se obteve 100% de assertividade no reconhecimento dos alunos presentes na imagem; o ângulo das faces presentes na imagem não influenciaram no reconhecimento, a não ser que realmente não de para identificar uma face inteira nesse ângulo; a distância das faces em relação a onde a imagem foi capturada não tem tanta influência, a não ser que seja impossível de enxergá-las; a luminosidade das faces não influencia no reconhecimento, a não ser que haja grande falta de luz na face que a torne não visível; o reconhecimento facial é realizado com sucesso mesmo com feições diferentes dos alunos presentes na imagem; não foi possível estimar ao certo o percentual de assertividade no reconhecimento de faces parcialmente visíveis, pois foi constatado que algumas faces foram possíveis ser reconhecidas corretamente apesar de não estarem 100% visíveis e outras não; faces divididas ao meio em eixo vertical não são possíveis serem reconhecidas.

Outro ponto interessante a ressaltar é que no 1º ano do ensino médio havia algumas gêmeas e mesmo diante dessa situação o serviço de reconhecimento facial conseguiu realizar a identificação com sucesso fazendo com que as duas estivessem presentes na lista montada pela aplicação, o que comprova a grande assertividade e confiabilidade do serviço.

Quanto a utilização do serviço, não houveram dificuldades aparentes para relatar, pois seu uso é relativamente simples, assim como a integração entre a aplicação e API utilizando o plugin HTTPClient do framework Angular dentro da aplicação desenvolvida em Ionic. Porém, para armazenar todos os dados da aplicação utilizou-se o Firebase Firestore (Serviço de armazenamento de dados NoSQL em nuvem disponibilizado pela Google) e o maior desafio do projeto foi manter os dados persistidos corretamente entre a base de dados no Firestore e a base de dados no Microsoft Azure.

7. Considerações Finais

Após todo o estudo e implementação do projeto utilizando uma aplicação para consumir os recursos do reconhecimento facial do Microsoft Azure é possível concluir que é viável a utilização desse serviço oferecido pela Microsoft, pois ele apresenta um resultado satisfatório com ótimo desempenho, assertividade, confiabilidade e um custo relativamente baixo levando-se em consideração o benefício e a facilidade da utilização. Durante o percurso do projeto utilizou-se uma licença do Azure gratuita para estudantes, porém a Microsoft oferece 30 dias de teste para novas contas, o que torna possível a implementação e teste para esse serviço. Após esse período, o serviço passa a ser cobrado, mas os valores são acessíveis considerando o custo/benefício.

Diante dos resultados obtidos é claramente visível o sucesso da Microsoft em investir em serviços cognitivos e disponibilizá-los facilitando o acesso àqueles que necessitam de serviços de Inteligência Artificial, os quais tem crescido a cada dia e se tornando mais difundidos, ou seja, esse é o futuro da tecnologia e vale a pena se aliar ao Microsoft Azure para criar soluções melhores e mais inteligentes para agregar valores e auxiliar a humanidade.

8. Fontes Consultadas

ComScore. **Estudo Global do Mobile**. Disponível em: http://icabrasil.org/2016/files/557-corporateTwo/downloads/The+Global+Mobile+Report_FINAL_PT.pdf. Acesso em: 14 de março de 2018.

eBit - **WEBSHOPPERS 36ª Edição 2017**. Disponível em: https://iabbrasil.com.br/estudo-ebit-webshoppers-36a-edicao-2017/?gclid=EAlaIQobChMI3PW0-zr2QIVFQaRCh38NQ-qEAAYASAAEgl_7fD_Bw. Acesso em: 14 de março de 2018.

FGV. **Pesquisa Anual do Uso de TI**. Disponível em: <http://eaesp.fgv.br/ensinoeconhecimento/centros/cia/pesquisa>. Acesso em: 14 de março de 2018.

Ionic. **Ionic Framework**. Disponível em: <https://ionicframework.com>. Acesso em 08 de abril de 2018.

MICROSOFT. **Serviços Cognitivos**. Disponível em: <https://azure.microsoft.com/pt-br/services/cognitive-services>. Acesso em 08 de abril de 2018.

OKABE, R. K. **Reconhecimento facial em imagens capturadas por câmeras digitais de rede**. 2014. 14 p. Estudo de Caso - Artigo (Informática)- UNOESTE, Universidade do Oeste Paulista, Presidente Prudente, 2014.

OpenCV. **Open Source Computer Vision Library**. Disponível em: <https://opencv.org>. Acesso em: 03 de abril de 2018.