

Reconhecimento de imagens alimentar

Ithalo Leonel

Faculdade Impacta Tecnologia (FIT)

São Paulo, São Paulo – Brasil

Abstract

This work presents a machine learning application with the objective of interpreting images and identifying the selected foods, so that it can allow the target audience (which are people with visual impairment and with some type of food allergy), to be able to make sure that there really is no allergic food on your plate and so you can also have more independence. Since the beginning of 2022, 28,000 images have been collected for the database, which were divided into parts: training, testing and also validation for each chosen ingredient (sausage and shrimp), these images were collected from different sources such as: Google, Bing and Pinterest. The presented algorithm was trained with a convolutional neural network with the focus on identifying whether there is shrimp or sausage on a plate of food, which were the selected models.

Resumo

Este trabalho apresenta uma aplicação de machine learning com objetivo de interpretar imagens e identificar os alimentos selecionados, para que assim possa permitir que o público-alvo (que são as pessoas com deficiência visual e com algum tipo alergia alimentar), possa se certificar que realmente em seu prato não há o alimento alérgico e assim também podendo ter mais independência. Desde o início de 2022 foram coletadas para o banco de dados 28.000 imagens, que foram divididas em partes sendo elas: treino, teste e validação igualmente para cada ingrediente escolhido (Salsicha e Camarão), essas imagens foram coletadas de diferentes fontes como: Google, Bing e Pinterest. O algoritmo apresentado foi treinado com uma rede neural convolucional com o foco de identificar se em um prato de comida há camarão ou salsicha, que foram os modelos selecionados.

1. Introdução

Em janeiro de 2022 foram registrados 7,8 bilhões de habitantes em nosso planeta e segundo pesquisas a alergia alimentar atinge aproximadamente 3 a 8% da população mundial [1]. A identificação de

ingredientes nos alimentos que consumimos em restaurantes, festas ou afins nem sempre é clara, o que pode acarretar problemas alérgicos, esta dificuldade é maior ainda a pessoas que possuem alguma deficiência visual. A UMC (União Mundial dos Cegos) é uma organização global que representa um total estimado de 253 milhões de pessoas cegas ou com deficiência visual em todo mundo [2]. Com o intuito de auxiliar estas pessoas criamos o Luke.

Luke é uma plataforma tendo como principal um modelo que realiza o reconhecimento de Imagem, para identificar ingredientes que possam ser prejudiciais à saúde do usuário. Sendo assim Luke possui uma Rede Neural Convulacional (CNN), é utilizado uma arquitetura especial que é adequada para a classificação de imagens.

Foram coletadas mais de 28.000 imagens de diversos sites de pesquisa além de que estas imagens foram tratadas e padronizadas para uma melhor classificação por parte do algoritmo.

O artigo foi dividido em Seções sendo a Seção 02 apresentando a Caracterização do Problema; Seção 03 apresentando a Descrição da Base de Dados; Seção 04 apresentando a Aplicação das Técnicas; Seção 05 os Resultados e pôr fim a Seção 06 com as Referências.

2. Caracterização do Problema

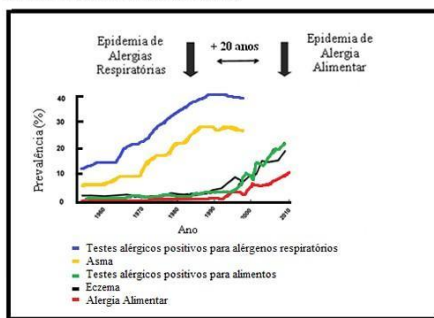
Uma única refeição pode acabar em sérios danos à saúde inclusive podendo levar a óbito, quando se trata de uma pessoa alérgica que se expôs a um alimento que precisava evitar.

Os principais alimentos que causam alergia são frutos do mar, embutidos, leite, ovo, canteios e aditivos alimentares [4].

Os sintomas das alergias alimentares podem ser desde espirros; coceiras; nariz entupido; inchaço; diarreia; erupções cutâneas e dores de estômago podendo chegar a dificuldades de respirar; arritmia cardíaca; tonturas e desmaios e até a morte [5].

Abaixo segue o estudo realizado pela Universidade Federal de Goiás demonstrando o aumento das alergias alimentares ao longo dos anos.

Figura 1 - Mudanças nos padrões das doenças alérgicas ao longo do tempo, em países desenvolvidos (Austrália, Nova Zelândia, Hong Kong, Cingapura)



A epidemia de sensibilização e alergia respiratória ocorrida nas décadas de 80 e 90 parece estar se deslocando para uma epidemia de sensibilização e alergia alimentares a partir dos anos 2000. Fonte: Adaptado de (PRESCOTT; ALLEN, 2011).

Até o momento não há grandes estudos nacionais ou regionais que tenham avaliado com acurácia a prevalência de alergia alimentar. Inquérito epidemiológico multicêntrico realizado em consultórios de gastroenterologistas de 20 cidades brasileiras apontou uma prevalência de sintomas suspeitos de alergia alimentar em 7,4% [6]. Pessoas com deficiência visual correm um risco ainda maior ao frequentarem restaurantes, festas ou jantares com amigos, necessitando sempre do um auxílio de alguém próximo.

3. Descrição da Base de Dados

Para realizar a identificação dos ingredientes Camarão e Salsicha que se enquadram em dois das maiores categorias de alimentos com grande concentração de alérgicos: Frutos do Mar e Embutidos respectivamente, foram coletadas 28.550 imagens de diversas fontes de pesquisa como Google, Bing e Pinterest. Coletamos estas imagens através da Extensão Fatkun Batch [7]. As imagens extraídas estavam no formatos: .gif, .jpg, .jiff, .html, .png e .webp. As imagens foram tratadas, imagens desnecessárias, poluídas ou que não atendessem os parâmetros de qualidade como por exemplo imagens em preto e branco, imagens de banners, animações, menus ou desenhos foram excluídas. As imagens foram convertidas para o formato .jpg após esta análise restaram 25.500 imagens sendo 8.500 imagens de prato com o ingrediente Camarão em sua composição e 4.250 imagens de pratos com o ingrediente Salsicha em sua composição e 12.750 imagens de outros alimentos que não possuem os ingredientes Camarão e Salsicha. Para realizar o treinamento do algoritmo as imagens foram subdividas em pastas para

serem alocadas futuramente aos datasets de treino, teste e validação.

Para o ingrediente Camarão foram alocadas 11.000 imagens para treino, 2.000 imagens para teste e 4.000 imagens para validação. Para o ingrediente Salsicha foram alocadas 6.500 imagens para treino, 700 imagens para teste e 1.300 imagens para validação.

As imagens estavam em diversos tamanhos para padronizar as imagens foi utilizado a ferramenta Power Toys [8] tornando todas as imagens em tamanho médio (1.366 x 768)

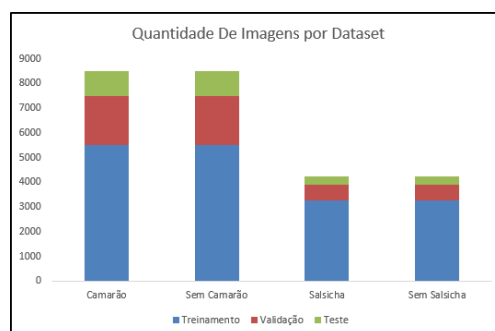


Gráfico 01: Comparação da Quantidade de Imagens em cada pasta de treino.

4. Aplicação das Técnicas

Com o auxílio da biblioteca ImageDataGenerator[9], foram criados três datasets contendo as regras de pré-processamento das imagens; dataset treino, dataset teste e dataset validação.

Ao preparar os datasets iniciamos o processamento das imagens de treino, validação e teste.

O Luke é uma Rede Neural Convolucional (CNN) que combina as bibliotecas TensorFlow[10] e Keras[11].

As Redes Neurais Convolucionais se utilizam de uma arquitetura especial que é particularmente bem adequada para classificar imagens. O uso dessa arquitetura torna as redes convolucionais rápidas de treinar, o que é vantajoso para trabalhar com redes profundas.

Para o seu funcionamento foram criadas duas camadas de convulação, onde foram ativadas com Relu.

O objetivo do Relu é a aplicação de uma função não linear na saída da camada não convolucional, assim os dados não são

modulados por uma combinação linear e o algoritmo e a classificação é facilitada.

Foram adicionadas duas camadas de pooling também.

Uma camada de pooling serve para simplificar a informação da camada anterior.

A saída da segunda camada passamos o flatten para transformar a matriz da imagem para uma array para identificações de padrões.

Ao final da rede é colocada uma camada Full Connected, onde sua entrada é a saída da camada anterior e sua saída são N neurônios, com N sendo a quantidade de classes do seu modelo para finalizar a classificação.

Com a conclusão do treinamento das imagens, o treinamento do algoritmo começa, para determinar a quantidade de steps por épocas dividimos o total de imagens alocadas dentro do dataset treino dividido pelo batch 64, foi utilizada uma época[12] para analisarmos o comportamento do algoritmo e iniciarmos o balanceamento dos hiperparâmetros sendo eles a quantidade apropriada de épocas; tamanho do batch; tamanho das imagens; quantidade necessária de imagens em cada dataset.

Utilizamos a Matriz de Confusão como Avaliador de Desempenho principal. A Matriz de Confusão é uma tabela que mostra as frequências de classificação para cada classe do modelo. O Indicador Falso Negativo é o que necessita de maior atenção, devido ao fato de o algoritmo não conseguir identificar o ingrediente e considerar o alimento como seguro. Foi utilizado um score de probabilidade para validar se o prato possui o ingrediente ou não.

5. Resultados

Todos os resultados aqui apresentados foram realizados sobre o conjunto de testes do algoritmo identificador do ingrediente Camarão. 2000 imagens com independência estatística em relação aos dados de treinamento e validação.

Como observado na Tabela 1 ao aumentarmos a quantidade de épocas a acurácia teve uma melhora.

Batch	Épocas	Qtd de Img Treino	Qtd de Img Validação	Qtd de Img Teste	Acuracia
64	1	11000	4000	2000	64,23%
64	10	11000	4000	2000	74,45%
64	25	11000	4000	2000	77,30%
64	100	11000	4000	2000	81,29%

Tabela 1: Resultados obtidos através da alteração de épocas

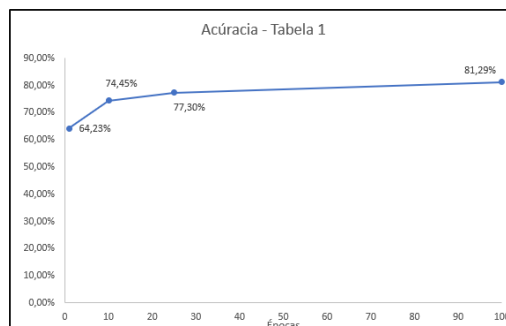


Gráfico 02: Comparação de Acurácia entre épocas levando em consideração os dados da Tabela_1

Ao utilizar o batch com 96 submetendo a mesma quantidade de épocas obtivemos uma melhora na acurácia.

Batch	Épocas	Qtd de Img Treino	Qtd de Img Validação	Qtd de Img Teste	Acuracia
96	1	11000	4000	2000	59,96%
96	10	11000	4000	2000	73,49%
96	25	11000	4000	2000	75,36%
96	100	11000	4000	2000	82,13%

Tabela 2: Resultados obtidos através da alteração de épocas e alteração do batch de 64 para 96

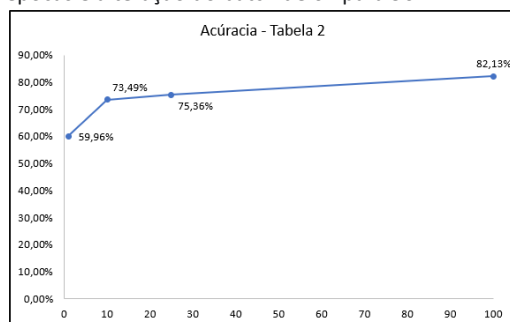


Gráfico 03: Comparação de Acurácia entre épocas levando em consideração os dados da Tabela_2

6. Referências

[1] *PEBMED - Alergia alimentar atinge 10% da população* - <https://pebmed.com.br/alergia-alimentar-atinge-10-da-populacao/>

[2] *Ministério da Educação - Data reafirma os direitos das pessoas com deficiência visual* - <http://portal.mec.gov.br/component/tags/tag/deficiencia-visual#:~:text=A%20UMC%20C3%A9%20uma%20organiza%C3%A7%C3%A3o,com%20defici%C3%AAn%20cia%20visual%20no%20mundo.>

[3] A Matriz de Confusão é uma tabela que mostra as frequências de classificação para cada classe do modelo.

[4] HCOR - Nutróloga do HCOR aponta os alimentos que mais causam alergia alimentar -
<https://www.hcor.com.br/imprensa/noticias/nutrologa-do-hcor-aponta-os-alimentos-que-mais-causam-alergia-alimentar/#:~:text=para%20combat%C3%AA%2Dlas,-,Leite%20de%20vaca%2C%20soja%2C%20amendoim%2C%20ovo%2C%20castanhas%2C,do%20alimento%20e%20seus%20derivados.>

[5] Hermes Pardini – Tudo o que você precisa saber sobre alergia alimentar. -
<https://www.hermespardini.com.br/blog/?p=427>

[6] Universidade Federal de Goiás programa de Pós- Graduação em medicina tropical e saúde pública - GERMANA PIMENTEL STEFANI - chrome-extension://efaidnbmnnnibpcajpcgclefindmkaj/https://files.cercomp.ufg.br/weby/up/59/o/GermanaStefani_2015_Vers%C3%A3oFinal.pdf

[7] Fatkun Batch Download da imagem – chrome web store -
<https://chrome.google.com/webstore/detail/fatkun-batch-download-ima/nnjjahliabnchcpehckpkdeckfgnohf?hl=pt-BR>

[8] Power Toys Microsoft Ignite -
<https://learn.microsoft.com/pt-br/windows/powertoys/image-resizer>

[9] ImageDataGenerator:
<https://keras.io/api/preprocessing/image/>

[10] TensorFlow: <https://www.tensorflow.org/>

[11] Keras: <https://keras.io/>

[12] Épocas são as repetições que o algoritmo irá realizar para aprender determinada tarefa.