

Apresentação Grupo 1

Título da Monografia: Modelo de Detecção do Uso de Máscaras Faciais Utilizando Deep Learning

(Entrega 4 de 5)

Alunos:

Daniel Campos - matrícula: A57635769 (daniel.ferraz.campos@gmail.com)

Leandro Daniel - matrícula: A57622988 (contato@leandrodaniel.com)

Ricardo Reis - matrícula: A57590919 (ricardo.l.b.reis@gmail.com)

Rodrigo Gonçalves - matrícula: A57566093 (rodrigo.goncalves@me.com)

Ygor Lima - matrícula: A57549661 (ygor_redesocial@hotmail.com)

**Apresentando nosso time
e progresso até o momento**

Chief Analytics Officer (CAO): Coordenação, Arquitetura de Software e Desenvolvimento



Leandro Daniel

- *Possui mais de 20 anos de experiência em desenvolvimento e arquitetura de software;*
- *Formação como Analista de Sistemas e atualmente como Data Scientist pela FGV;*
- *Co-fundador do primeiro grupo de arquitetura de software Microsoft: .NET Architects;*
- *Autor de diversos artigos de tecnologia no portal DevMedia, onde já foi Editor Técnico;*
- *Palestrante e coordenador de trilhas em grandes eventos de tecnologia, como o TDC (The Developers Conference), Microsoft TechNet e QCon.*



<https://github.com/ldaniel>



<https://www.linkedin.com/in/leandrodaniel/>



<https://www.kaggle.com/leandrodaniel>

Machine Learning Engineer (MLE): Modelagem, Arquitetura de Dados e Visualização



Rodrigo Gonçalves

- *Profissional multidisciplinar formado em Ciências Econômicas. Especialista em Finanças, Contabilidade e Business Analytics com grande experiência em pesquisa econômica e análise de dados.*
- *15 anos de experiência em consultorias de Finanças, Controladoria e Gerenciamento de Projetos.*
- *Atualmente Trabalha como Gerente de Tecnologia Forense e Data Analytics em projetos de investigação de fraudes.*



<https://github.com/RodriGonca>



<https://www.linkedin.com/in/rodriggonca/>



<https://www.kaggle.com/rodriggonca>

Machine Learning Engineer (MLE): Modelagem, Arquitetura de Dados e Treinamento/Teste



Daniel Campos

- *Mais de 6 anos de experiência no setor da saúde atuando como consultor de negócios, analista de inteligência de mercado e cientista de dados;*
- *Graduado em Engenharia Aeroespacial e Bacharel em Ciência e Tecnologia e atualmente como Data Scientist pela FGV;*
- *Experiência na indústria de games como cientista de dados;*
- *Co-autor de publicações relevantes para a indústria da saúde no Brasil;*
- *Criador de uma das mais relevantes publicações sobre record linkage dos dados públicos do DataSUS no Brasil.*



<https://github.com/DanielFCampos>



<https://www.linkedin.com/in/daniel-ferraz-de-campos/>



<https://www.kaggle.com/danielferrazcampos>

Research Data Scientist (RDS): Pesquisa, Testes e Engenharia de Dados



Ricardo Reis

- *Formado em ciência da computação pela Universidade Federal do Maranhão (UFMA), 2014;*
- *Pós graduado em administração de empresas pela Fundação Getúlio Vargas (FGV), 2016;*
- *Possui 7 anos de experiência com análise e desenvolvimento de sistemas nos setores público e privado ;*
- *Interesse na área acadêmica (pesquisa científica);*
- *Atualmente trabalhando e se desenvolvendo na área de data science.*



<https://github.com/ricardobreis>



<https://www.linkedin.com/in/ricardobreis/>



<https://www.kaggle.com/ricardobreis>

Research Data Scientist (RDS): Pesquisa, Testes e Engenharia de Dados



Ygor Lima

- *Formado em Administração pela Pontifícia Universidade Católica de São Paulo, 2016;*
- *Atuou nos últimos 5 anos como engenheiro de dados em um banco privado;*
- *Criador do grupo de disseminação de conhecimentos gerais de Tecnologia dentro da Diretoria em que trabalha;*
- *Consultoria em processos para encontrar oportunidades de eficiência com automação de rotinas;*
- *Trilhando conhecimento para tornar-se cientista de dados.*



<https://github.com/ygorlima1>



<https://www.linkedin.com/in/ygorlima1/>



<https://www.kaggle.com/ygorlima1>



1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico



- Onde a pesquisa será realizada (contexto)?

[1 de 5]

Vivemos tempos singulares, onde nossa sociedade enfrenta o desafio de combate e prevenção contra o crescente alastramento da COVID-19 em uma escala global. Deste desafiador contexto, emergem efeitos colaterais inéditos na era moderna, como o distanciamento social, dentre diversas mudanças de hábitos.

Com base nas evidências atuais, o vírus COVID-19 é transmitido entre pessoas através de contato próximo e gotículas. Calcula-se que uma pessoa com infecção o transmita para de duas a quatro pessoas.

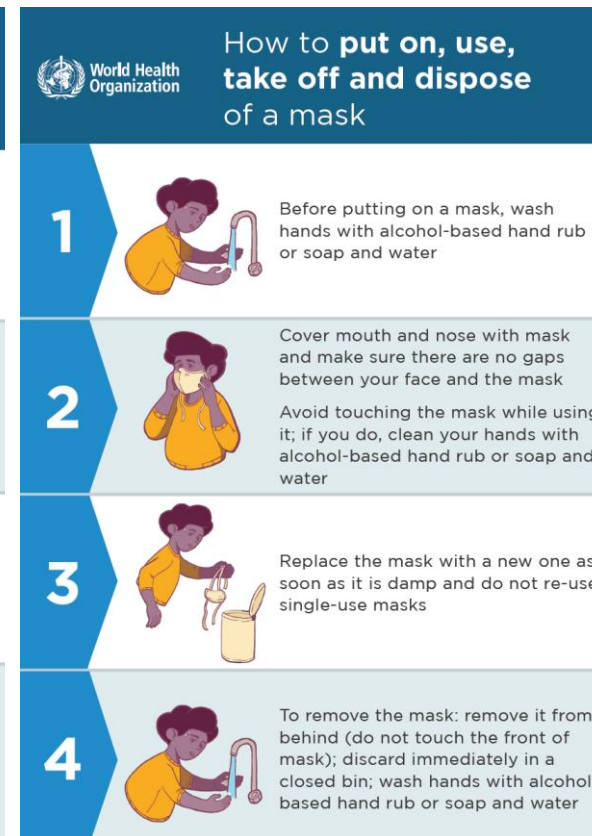
A Organização Mundial de Saúde (OMS) recomenda o uso de máscaras combinadas com a correta higiene frequente das mãos como parte da chamada EPI (Equipamento de Proteção Individual, do inglês PPE, ou Personal Protective Equipment), dentre outras orientações e recomendações.

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico

- Onde a pesquisa será realizada (contexto)?

[2 de 5]

A Organização Mundial de Saúde disponibiliza uma série de orientações e guias para o uso correto das máscaras, bem com seu descarte e combinação com procedimentos de lavagem das mãos, como uma das medidas possíveis de contenção do COVID-19.



Fonte(s) e URL(s) utilizada(s) em 24/05/2020:

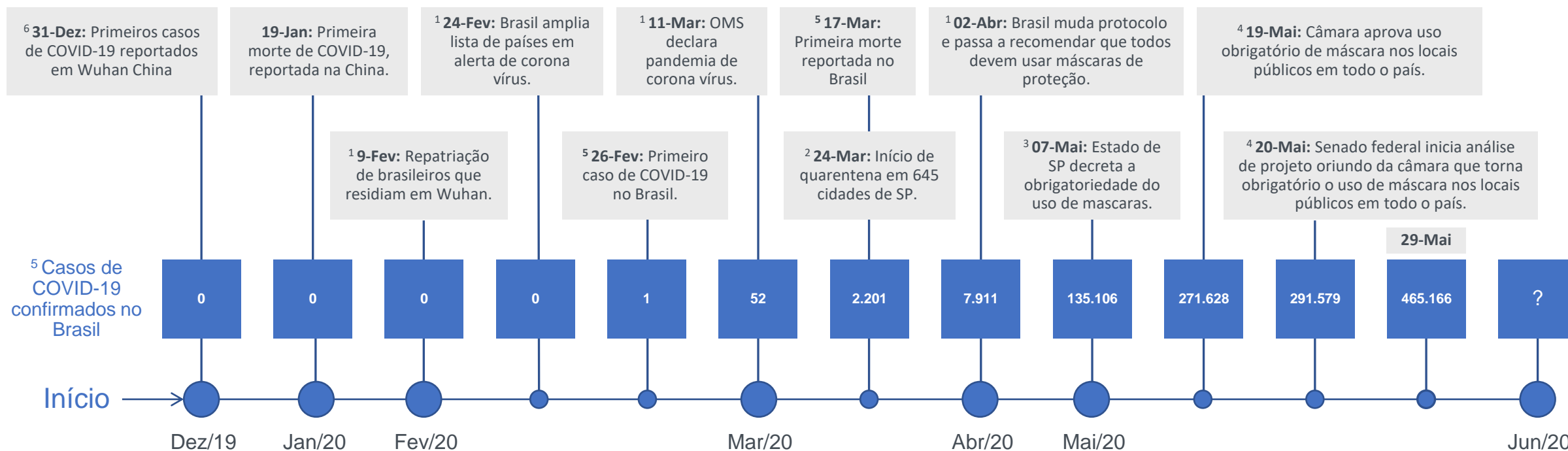
Coronavirus disease (COVID-19) advice for the public: When and how to use masks

<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks>

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico

- Onde a pesquisa será realizada (contexto)?

[3 de 5]



Fonte(s) e URL(s) utilizada(s) em 27/05/2020:

¹ <https://www.sanarmed.com/linha-do-tempo-do-coronavirus-no-brasil>

² <https://g1.globo.com/sp/sao-paulo/noticia/2020/03/24/quarentena-passa-a-valer-nesta-terca-feira-em-sp-veja-o-que-funciona.ghtml>

³ <https://noticias.uol.com.br/saude/ultimas-noticias/redacao/2020/05/04/sao-paulo-decreta-obrigatoriedade-do-uso-de-mascaras-em-todo-o-estado.htm>

⁴ <https://g1.globo.com/politica/noticia/2020/05/19/camara-aprova-uso-obrigatorio-de-mascara-em-todo-o-pais.ghtml>

⁵ <https://covid.saude.gov.br/>

⁶ <https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/>

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico

- Onde a pesquisa será realizada (contexto)?

[4 de 5]

Recentemente o estado de São Paulo decretou o uso de máscaras de proteção facial como obrigatório, por tempo indeterminado. A medida, que entrou em vigor em 07/05/2020, estabelece o seguinte:



Uso de máscaras em:

- espaços públicos;
- estabelecimentos que executem atividades essenciais;
- repartições públicas estaduais;
- transporte por aplicativo.



Usados por:

- consumidores;
- fornecedores;
- clientes;
- empregados;
- colaboradores;
- agentes públicos;
- prestadores de serviço.

Além disso, o descumprimento da medida acarretará em infração, podendo o infrator ou o estabelecimento serem multados.

Fonte(s) e URL(s) utilizada(s) em 27/05/2020:

<https://g1.globo.com/sp/sao-paulo/noticia/2020/05/05/decreto-que-determina-uso-de-mascaras-em-sp-estipula-multa-de-ate-r-276-mil-e-4-anos-de-prisao.ghtml>

<https://www.al.sp.gov.br/repositorio/legislacao/decreto/2020/decreto-64959-04.05.2020.html>

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico

- Onde a pesquisa será realizada (contexto)?

[5 de 5]

A presente pesquisa tem como objetivo principal, desenvolver um modelo de detecção de indivíduos utilizando máscaras faciais, através de técnicas de Deep Learning, como ferramenta possível de ser implementada em sistemas de monitoramento preventivo ao alastramento da COVID-19 em locais e áreas de convívio público ou privado.

O modelo resultante do estudo, visa sua aplicação em sistemas de monitoramento por vídeo como uma ferramenta adicional de alerta em casos de pessoas que não estejam utilizando máscaras faciais. Podendo adicionalmente, ser empregado em combinação com materiais educacionais na orientação do adequado uso em estabelecimentos comerciais e artísticos, instituições de ensino ou quaisquer outros usos onde as corretas orientações científicas possam ser asseguradas de forma rápida e ampla.



"The best way forward is the blended comprehensive approach which puts containment as a major pillar."

-- Dr. Tedros Adhanom, WHO director-general.

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico



- Qual é a **PERGUNTA** do **PROBLEMA** de pesquisa?

Diante do contexto descrito, apresentam-se o questionamento a ser abordado nessa pesquisa:

É possível desenvolver, com técnicas de deep learning, um modelo de detecção de pessoas utilizando máscaras faciais para posterior implementação em sistemas de controle de entrada em espaços públicos, estabelecimentos que executem atividades essenciais, repartições públicas estaduais, transporte por aplicativo para um público alvo de consumidores, fornecedores, clientes, empregados, colaboradores, agentes públicos e prestadores de serviço?

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico



- Qual é o **OBJETIVO GERAL** e quais são os **OBJETIVOS ESPECÍFICOS**?

Objetivo geral:

Desenvolver um modelo que identifique, se as pessoas estão utilizando máscaras de proteção facial.

Objetivos específicos:

1. *Definir, e eventualmente construir, a base de imagens que será utilizada;*
2. *Realizar revisão bibliográfica de técnicas aplicadas ao tema de pesquisa;*
3. *Definir os algoritmos que serão utilizados;*
4. *Definir a arquitetura inicial dos algoritmos;*
5. *Construir os modelos e identificar o com os melhores resultados;*
6. *Iterar sobre a estrutura inicialmente proposta e realizar o tuning de hiperparâmetros, se necessário;*
7. *Analisar a performance e a estabilidade da solução.*

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico



- Quais são as **HIPÓTESES** da pesquisa (se necessário)?

Para este trabalho, dado o contexto da proposta, não se faz necessária a inclusão de hipóteses a serem validadas ou refutadas posteriormente.

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico



- Quais artigos científicos já trataram do mesmo contexto e/ou utilizaram as mesmas técnicas estatísticas e/ou de *machine learning* que serão utilizadas (**apresentar resumidamente 20 artigos no mínimo**)?

[1 de 4]

1. LI, H.; LIN, Z.; SHEN, X.; BRANDT, J.; HUA, G. **A Convolutional Neural Network Cascade for Face Detection**. 28th IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA. 2015. DOI 10.1109/CVPR.2015.7299170.
2. RANJAN, R.; SANKARANARAYANAN, S.; CASTILLO, C. D.; CHELLAPPA, R. **An All-In-One Convolutional Neural Network for Face Analysis**. IEEE 12th International Conference on Automatic Face & Gesture Recognition. Washington, USA. 2017. DOI 10.1109/FG.2017.137.
3. NAUDÉ, W. **Artificial Intelligence against COVID-19: An Early Review**. IZA – Institute of Labor Economics. Bonn, Alemanha. 2020. ISSN 2365-9793.
4. GUO, J.; LIN, C.; WU, M.; CHANG, C.; LEE, H. **Complexity Reduced Face Detection Using Probability-Based Face Mask Prefiltering and Pixel-Based Hierarchical-Feature Adaboosting**. IEEE SIGNAL PROCESSING LETTERS. VOL. 18, NO. 8, AUGUST 2011. DOI 10.1109/LSP.2011.2146772.

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico

- Quais artigos científicos já trataram do mesmo contexto e/ou utilizaram as mesmas técnicas estatísticas e/ou de *machine learning* que serão utilizadas (**apresentar resumidamente 20 artigos no mínimo**)?

[2 de 4]

5. LAWRENCE, S.; GILES, C. L.; TSOI, A. C.; BACK, A. D. **Face Recognition: A Convolutional Neural-Network Approach**. IEEE Transactions on Neural Networks, Vol. 8, No. 1. 1997. DOI 10.1109/72.554195.
6. QIN, B.; LI, D. **Identifying Facemask-wearing Condition Using Image SuperResolution with Classification Network to Prevent COVID-19**. 2020. DOI 10.21203/rs.3.rs-28668/v1.
7. WANG, Z.; et al. **Masked Face Recognition Dataset and Application**. 2020. arXiv 2003.09093v2.
8. YIN, X.; LIU, X. **Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition**. IEEE Transactions on Image Processing. 2018. DOI 10.1109/TIP.2017.2765830.
9. JIANG, M.; FAN, X. **RetinaMask: A Face Mask Detector**. 2020. arXiv 2005.03950.
10. MATSUGU, M.; MORI, K.; MITARI, Y.; KANEDA, Y. **Subject independent facial expression recognition with robust face detection using a convolutional neural network**. Neural Networks, Volume 16, Issues 5-6, Pages 555-559. 2003. DOI 10.1016/S0893-6080(03)00115-1.

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico



- Quais artigos científicos já trataram do mesmo contexto e/ou utilizaram as mesmas técnicas estatísticas e/ou de *machine learning* que serão utilizadas (**apresentar resumidamente 20 artigos no mínimo**)?

[3 de 4]

11. CHOLLET, F. **Deep Learning with Python**. Manning. ISBN 9781617294433. NY, USA. 2018.
12. GÉRON, A. **Hands-On Machine Learning with Scikit-Learn & TensorFlow: Concepts, Tools, And Techniques To Build Intelligent Systems**. O'Reilly. ISBN 978-1-491-96229-9. CA, USA. 2017.
13. JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to Statistical Learning with Applications in R**. Springer. ISBN 978-1-4614-7138-7. DOI 10.1007/978-1-4614-7138-7. NY, USA. 2015.
14. HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Springer. ISBN 978-0-387-84858-7. DOI: 10.1007/b94608. NY, USA. 2009.
15. KUHN, M.; JOHNSON, K. **Applied Predictive Modeling**. Springer. ISBN 978-1-4614-6849-3. DOI 10.1007/978-1-4614-6849-3. NY, USA. 2013.

1) Contexto da Pesquisa, Busca de Artigos Científicos, Problema, Hipóteses e Objetivos do Projeto Analítico



- Quais artigos científicos já trataram do mesmo contexto e/ou utilizaram as mesmas técnicas estatísticas e/ou de *machine learning* que serão utilizadas (**apresentar resumidamente 20 artigos no mínimo**)?

[4 de 4]

16. Keras API Reference. Disponível em <<https://keras.io/api/>>. Acesso em: 30 de Maio de 2020.
17. RUSSEL, S. J.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 2010. DOI 10.1016/j.artint.2011.01.005.
18. LUNDBERG, S.; LEE, S. **A Unified Approach to Interpreting Model Predictions**. 2017. arXiv 1705.07874.
19. World Health Organization. **Rational Use of Personal Protective Equipment for Coronavirus Disease (COVID-19) and Considerations During Severe Shortages**. 2020. WHO/2019-nCov/IPC_PPE_use/2020.3.
20. SELVARAJU, R.; et al. **Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization**. 2019. DOI 10.1007/s11263-019-01228-7.

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico

- Quais variáveis formam a **BASE DE DADOS** (mostrar a base de dados completa)?

[1 de 5]

A base de dados utilizada para este projeto é uma composição de bancos de imagens classificadas de forma binária entre faces com máscara e faces sem máscara. Ambas as fontes estão publicamente disponíveis.

A primeira delas, produzida por **Prajna Bhandary**¹, trata-se de um banco *pseudo-artificial* e possui **1.376** imagens das quais **690** imagens estão classificadas como faces com máscara e **686** para a classe complementar.



Fonte(s) e URL(s) utilizada(s) em 13/06/2020:



Disponível no GitHub de Prajna Bhandary

¹ <https://github.com/prajnasb/observations/tree/master/experiments/data>

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico

- Quais variáveis formam a **BASE DE DADOS** (mostrar a base de dados completa)?

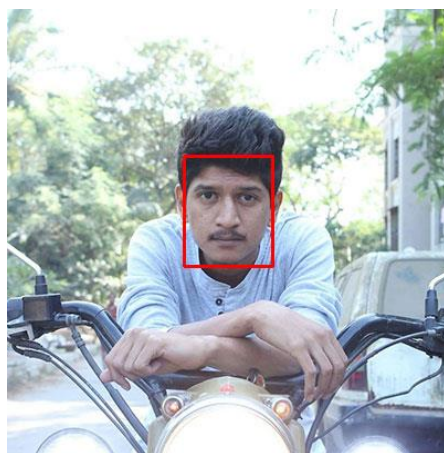
[2 de 5]

O conceito de *pseudo-artificialidade* mencionado é uma liberdade poética dos autores para expressar que, apesar de serem imagens reais de pessoas, as máscaras foram artificialmente criadas e posicionadas, via algoritmo, para o específico propósito de treinamento de modelos como o modelo que pretendemos criar.

A seguir, um exemplo de como se dá o processo de criação artificial dessas imagens:



Imagem original



Localização da face



Extração e identificação de contornos da face



Adição de uma representação de máscara



Resultado final

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico

- Quais variáveis formam a **BASE DE DADOS** (mostrar a base de dados completa)?

[3 de 5]

A segunda base de dados, por sua vez, tem como curador o *National Engineering Research Center for Multimedia Software*, do departamento de Ciências da Computação da **Universidade de Wuhan**.

Este banco de dados¹ possui 2 vertentes:

- uma artificial, similar ao processo anteriormente mencionado,
- e outra usando imagens de *mundo real*, ou seja, sem a adição artificial de máscaras de proteção.



Um exemplo de publicação usando este dataset pode ser visto no site da **Cornell University** no artigo: “*Masked Face Recognition Dataset and Application*”².

Fonte(s) e URL(s) utilizada(s) em 13/06/2020:



Disponível no GitHub

¹ <https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset>



Disponível no site da Universidade de Cornell

² <https://arxiv.org/abs/2003.09093>

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico



- Quais variáveis formam a **BASE DE DADOS (mostrar a base de dados completa)**?

[4 de 5]

Desta segunda fonte de dados, decidimos utilizar apenas as imagens de mundo real, que consiste de um total de **95.000** imagens de **525** pessoas diferentes das quais **5.000** imagens estão classificadas como faces com máscara e **90.000** para a classe complementar.

O intuito com a mescla dos *datasets* é trazer maior robustez e capacidade de generalização ao nosso modelo, bem como usar fontes de dados diferentes pode reduzir eventuais vieses dos criadores dos *datasets*.

Vale lembrar que as bases de dados foram usadas “*as-is*”, sem reclassificação das imagens, exclusão de casos dúbios ou curadoria de conteúdo. Entretanto, com uma avaliação visual superficial não-exaustiva, conseguimos observar algumas imagens mal classificadas e, apesar de entendermos que isso pode prejudicar a performance final dos modelos, seguimos para mantermos as bases originais inalteradas.

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico

- Quais variáveis formam a **BASE DE DADOS** (mostrar a base de dados completa)?

[5 de 5]

Todas as arquiteturas de modelos serão treinadas com o mesmo conjunto de treino e teste para garantirmos a justiça entre elas. A proporção definida foi de **75%** para o **treino** e **25%** para o **teste**. Para garantirmos que as classes estejam equilibradas em ambas as amostras, estratificaremos pela coluna de classes. O framework utilizado será o *train_test_split* do **scikitlearn**. Dado o desbalanceamento entre as classes também será necessária a técnica do *undersampling* (de forma aleatória) da classe majoritária.

Classe: Com Máscara



Classe: Com Máscara



Classe: Sem Máscara



Classe: Sem Máscara



Classe: Com Máscara



Classe: Com Máscara



Classe: Sem Máscara



Classe: Sem Máscara



Classe: Com Máscara



Classe: Com Máscara



Classe: Sem Máscara



Classe: Sem Máscara



2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico



- Quais variáveis foram concebidas após aplicação de *feature engineering*?

[1 de 3]

Sabemos o quão importante é termos um banco de dados robusto e grande o suficiente para habilitar a generalização do nosso modelo. Com isto em mente, uma prática comum aplicada à imagens é o processo de *data augmentation*, e este será o processo que usaremos para fazer o enriquecimento de nosso banco de imagens.

Este processo, basicamente, toma as imagens originais e aplica pequenas modificações nela tais como: **rotação, translação, cisalhamento, ampliação e espelhamento** a fim de gerar novas imagens para o treinamento dos modelos. O conceito por trás desse processo é assumirmos que uma imagem continua representando a mesma classe mesmo com estas pequenas modificações, permitindo com que tenhamos outras ópticas sobre a mesma imagem.

Esse contexto pode ser considerado como o equivalente à *feature engineering* em bases de dados tradicionais (não-imagens).

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico



- Quais variáveis foram concebidas após aplicação de *feature engineering*?

[2 de 3]

Para o *data augmentation*, a metodologia usada baseia-se no *framework* do **Keras** de *ImageDataGenerator* associado ao *flow_from_dataframe*, que cria um fluxo direto do diretório de imagens enquanto a definição de classes fica à cargo de um *dataframe* com o nome do arquivo (imagem) e sua respectiva classe. Desta forma, conseguimos trabalhar com um volume de dados maior sem onerar a capacidade computacional disponível.

Vale lembrar que, como parte do tratamento das imagens faremos também o *reshaping* das imagens para formatos quadrados (por exemplo, 256x256 pixels), que são melhores aceitos em redes neurais, especialmente as convolucionais. Faremos também a normalização dos valores dos pixels para que eles assumam valores apenas entre 0 e 1, que é uma prática padrão para o aumento da performance de modelos estatísticos de aprendizado de máquina.

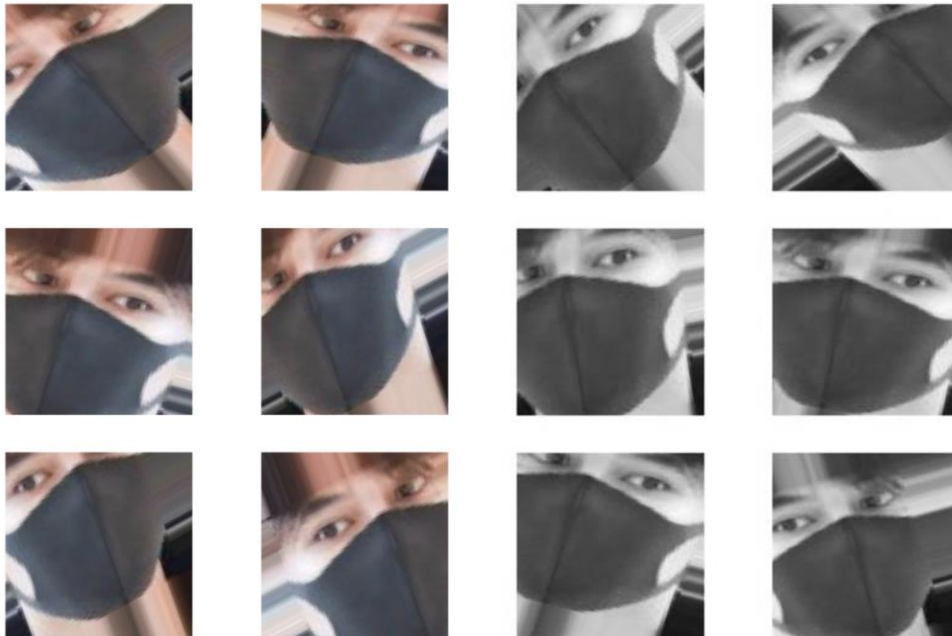
Por fim, para reduzirmos o consumo computacional e, entendendo que não existe perdas significativas ao projeto, trabalharemos com a conversão das imagens de input para escala de cinza.

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico

- Quais variáveis foram concebidas após aplicação de *feature engineering*?

[3 de 3]

A seguir exibimos alguns exemplos desse processo e valores que testamos. Nesta visualização, aplicamos o mesmo procedimento que será usado para treinamento dos modelos tal qual uso de **escalas de cinza ou esquema de cores** e **redimensionamento de imagens** para dimensões quadradas.



- **rotation_range (40)** é um valor em graus (0-180), um intervalo no qual rotacionaremos aleatoriamente as imagens.
- **width_shift (0.2)** e **height_shift (0.2)** são intervalos (como uma fração entre 0 e 1 da largura ou altura total) dentro dos quais é possível transladar aleatoriamente imagens na vertical ou na horizontal.
- **rescale** é um valor pelo qual multiplicaremos os dados antes de qualquer outro processamento. Nossas imagens originais consistem em coeficientes RGB no intervalo de 0 a 255, mas é recomendável a normalização dos dados para o processamento dos nossos modelos (dada uma taxa de aprendizado típica); portanto, transformaremos em valores entre 0 e 1 pela multiplicação do fator $1/255$. O mesmo vale para o caso de usarmos imagens BW.
- **shear_range (0.2)** é o intervalo o qual aplicaremos, aleatoriamente, transformações de cisalhamento.
- **zoom_range (0.2)** é o intervalo destinado à transformações aleatórias de ampliação.
- **horizontal_flip (Sim)** argumento do tipo booleano indicando que transformações aleatórias de espelhamento horizontal são permitidas - relevante quando não há suposições de assimetria horizontal (e.g. imagens reais).
- **fill_mode (Mais próximo/Nearest)** é a estratégia usada para preencher pixels recém-criados, que podem aparecer após alguma das transformações previamente mencionadas.

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico

- Quais técnicas estatísticas e/ou de *machine learning* serão utilizadas (dentre as aprendidas no curso)?

[1 de 2]

Para o projeto em questão, pretendemos abordar duas das arquiteturas mais usadas no campo de inteligência artificial:

- **ANNs** (**Artificial Neural Networks**) e
- **CNNs** (**Convolutional Artificial Neural Networks**).

As **ANNs** ou sistemas conectados são sistemas de computação inspirados nas redes neurais biológicas que constituem cérebros de animais. Esses sistemas são uma coleção de unidades ou nós conectados chamados de neurônios artificiais. Cada conexão, como as sinapses no cérebro biológico, pode transmitir um sinal para outros neurônios. Nas ANNs podemos ter várias camadas de neurônios, aumentando a complexidade da rede.

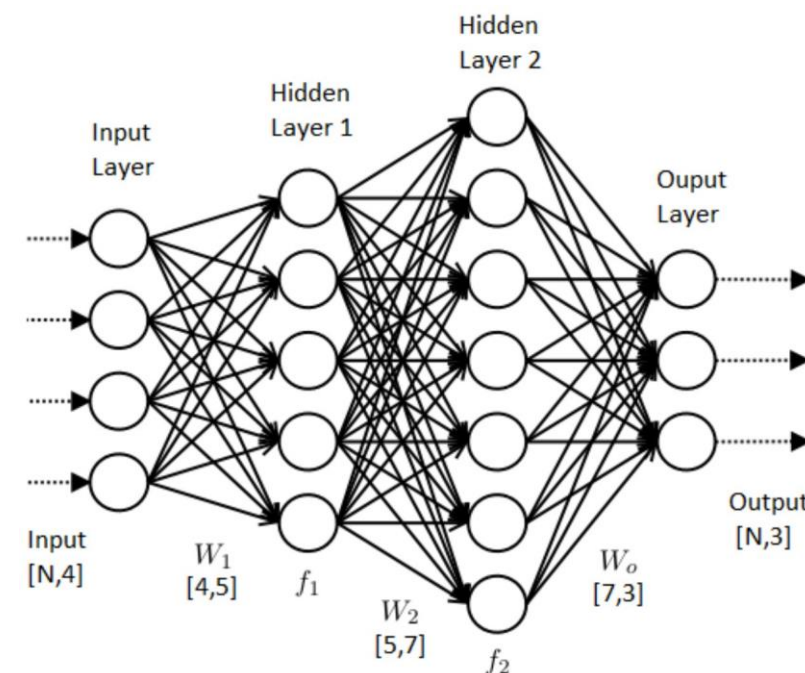


Imagem ilustrativa

2) Base de Dados e Técnicas Estatísticas e/ou de *Machine Learning* do Projeto Analítico

- Quais técnicas estatísticas e/ou de *machine learning* serão utilizadas (dentre as aprendidas no curso)?

[2 de 2]

As **CNNs** (ou *ConvNet*), por sua vez, são uma classe de redes neurais profundas, mais comumente aplicada à análise de imagens (nosso caso). Assim como as ANNs, elas são formadas por camadas conectadas por operações matriciais.

Entretanto, apesar de também serem acopladas à uma ANN em seu output, elas possuem camadas ocultas convolucionais, bem como outras camadas de operações multidimensionais.

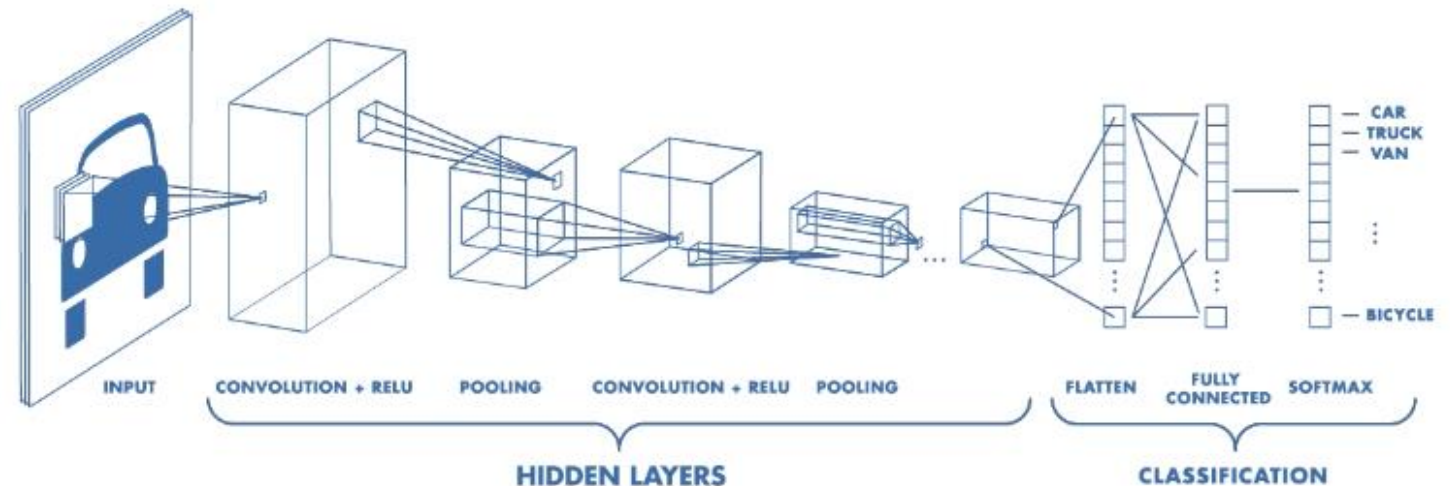


Imagem ilustrativa

3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (**ler para a turma**)?

[1 de 12]

Conforme apontado por Naudé Wim (2020), a inteligência artificial (IA) é uma ferramenta potencialmente poderosa na luta contra a pandemia do COVID-19. São discutidas seis áreas em que a IA pode contribuir para a luta contra o COVID-19:

- i) alertas e alertas precoces;
- ii) rastreamento e previsão;
- iii) painéis de dados;
- iv) diagnóstico e prognóstico;
- v) tratamentos e curas;
- vi) controle social.**



Este trabalho concentrou-se especialmente em medidas de **controle social**, notadamente, no uso de máscaras faciais para controle do alastramento da COVID-19.

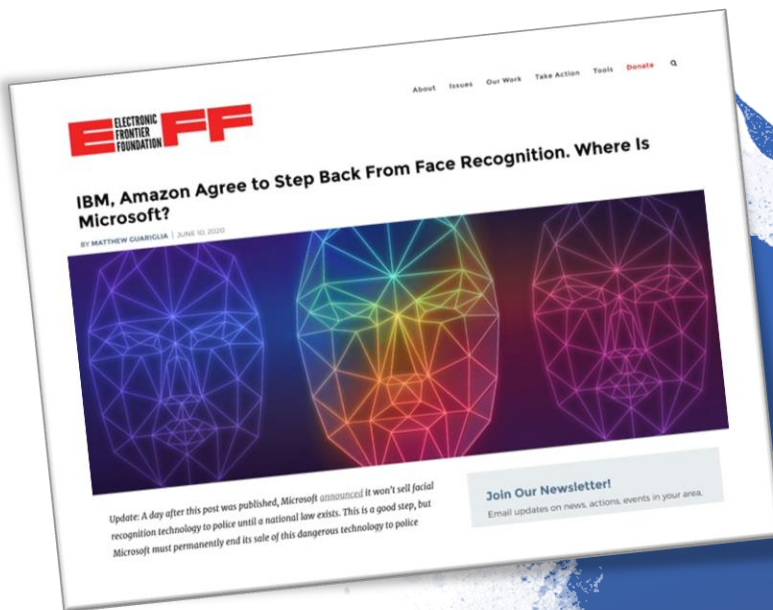
3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[2 de 12]

Naudé Wim (2020), enfatiza que a aplicação de técnicas de IA, especialmente no contexto do combate a COVID-19 é dificultada pela falta de dados. Ademais, a superação dessas restrições exige um equilíbrio cuidadoso entre privacidade de dados e saúde pública e interação rigorosa entre humanos e IA. É improvável que essas questões sejam tratadas a tempo de ser de grande ajuda durante a presente pandemia. Enquanto isso, a coleta extensiva de dados de controle social será benéfica para salvar vidas, treinar a IA e limitar os danos econômicos. A IA tem sido, e pode ser usada ainda mais, para gerenciar a pandemia usando imagens térmicas para varrer espaços públicos em busca de pessoas potencialmente infectadas e aplicando medidas de distanciamento e bloqueio sociais (RIVAS, 2020).

Por exemplo, conforme descrito por Chun (2020) no ***South China Morning Post***: "*Nos aeroportos e estações de trem na China, câmeras infravermelhas são usadas para escanear multidões em busca de altas temperaturas. Às vezes, eles são usados com um sistema de reconhecimento facial, que pode identificar o indivíduo com uma temperatura alta e se ele está usando uma máscara cirúrgica.*"



Privacidade

Abusos por
parte do Estado

Sistemas de
Vigilância

Prisões
indevidas

Direitos
Humanos

Erros de
identificação

I.A.

Fonte(s) e URL(s) utilizada(s) em 24/06/2020:

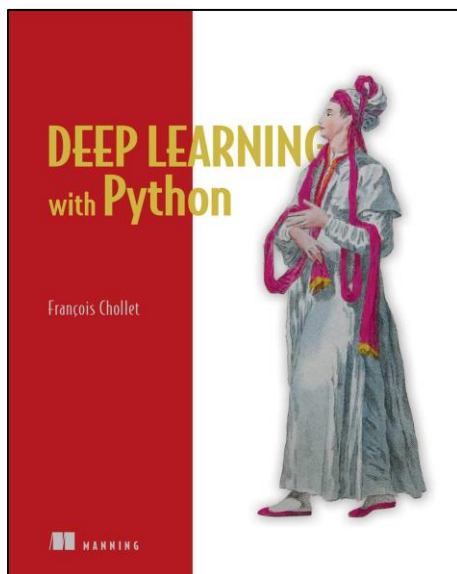
<https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/>
<https://www.eff.org/deeplinks/2020/06/ibm-amazon-end-and-pause-fr-programs-where-microsoft>
<https://www.fastcompany.com/90516450/ibm-microsoft-and-amazons-face-recognition-bans-dont-go-far-enough>

3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[3 de 12]

Na sequência descrevemos o referencial teórico das técnicas e métodos utilizados neste trabalho de pesquisa baseados nos trabalhos de **CHOLLET, F. Deep Learning with Python**¹. e a documentação oficial do Keras TensorFlow 2.0, disponível em seu site oficial ².




François Chollet é o autor do Keras, um framework de *Deep Learning* para Python. Pesquisador de IA e aprendizagem profunda no Google e autor de trabalhos aceitos em grandes conferências como CVPR, NIPS, ICLR, dentre outras.



Fonte(s) e URL(s) utilizada(s) em 24/06/2020:

 Livro disponível no site da Amazon
¹ <https://amzn.to/3fYOfVU>

 Site oficial do Keras TensorFlow
² <https://keras.io/about/>

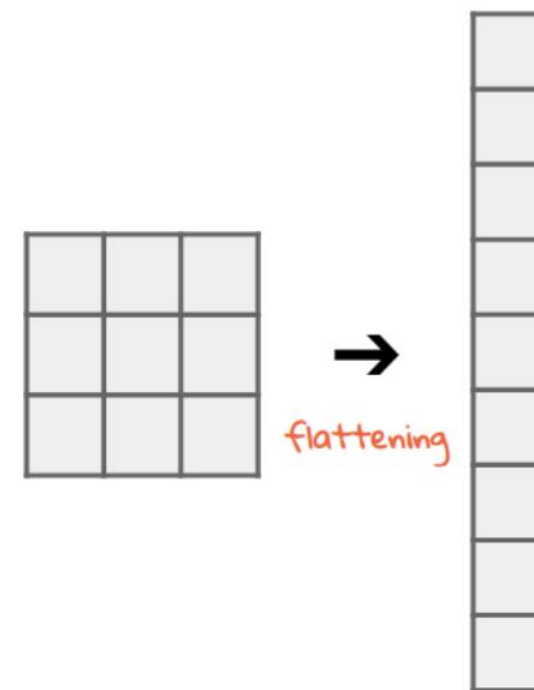
3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[4 de 12]

Para a comparação de diferentes arquiteturas, construímos diferentes estruturas baseadas nas técnicas de **Flatten**, **Dense**, **Dropout**, **Conv2D**, **BatchNormalization** and **MaxPooling2D**. Cada técnica é aplicada a uma camada diferente das redes neurais.

Flatten: a técnica de achatamento (*flatten*) converte os dados de uma matriz multidimensional em uma matriz unidimensional (vetor) para os dados serem usados como input em uma próxima camada. É comum achatarmos a saída das camadas convolucionais ou a entrada multidimensional (e.g. imagens) de redes densas para criarmos um único vetor longo de *features*. Esta camada, portanto, precede uma camada densa conectada (*fully connected dense layer*, explicada a seguir) costumeiramente responsável pela classificação final do modelo. De forma mais simples, colocamos todos os dados de pixels (no caso de imagens) em linha para usarmos como *input* em uma rede neural.

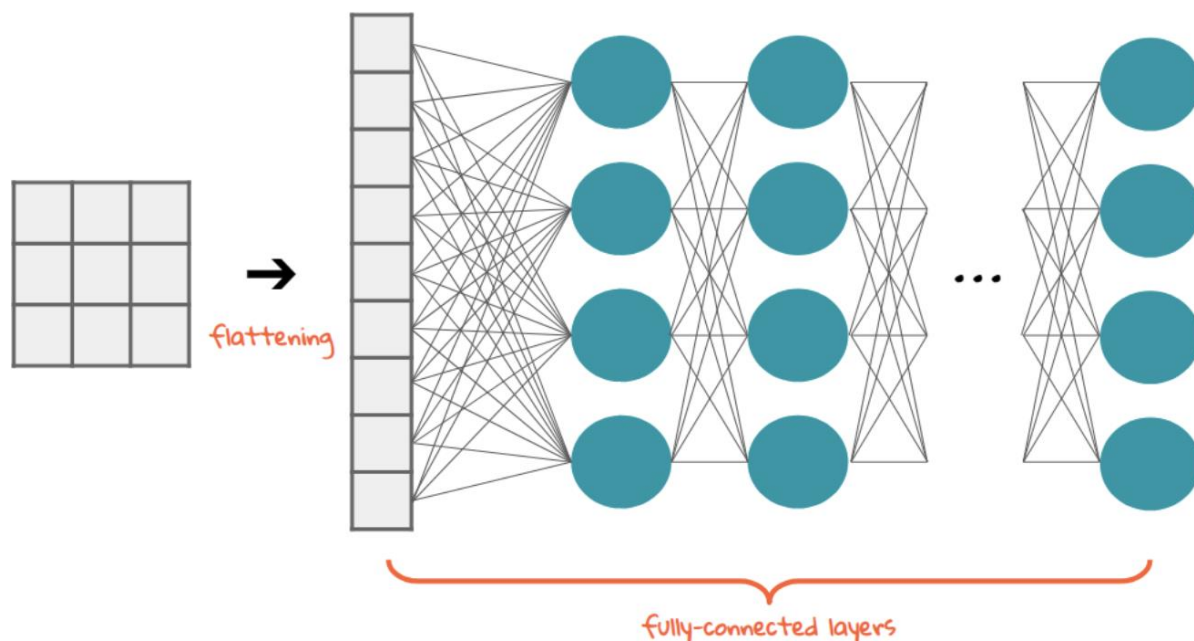


3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[5 de 12]

Dense: uma camada densa (*dense layer*) é apenas uma camada regular de neurônios em uma rede neural. Cada neurônio recebe a entrada de todos os neurônios da camada anterior estando, portanto, densamente conectados. Eles também, possuem uma função de ativação que é responsável por normalizar a saída dos neurônios (e.g. ReLU).



Uma camada densa representa nada mais que uma multiplicação matricial. A camada tem uma matriz de peso e um vetor de vieses aplicados às ativações da camada anterior. Os valores que modificam as ativações da camada anterior são parâmetros treináveis que são atualizados durante o *backpropagation* (treinamento da rede).

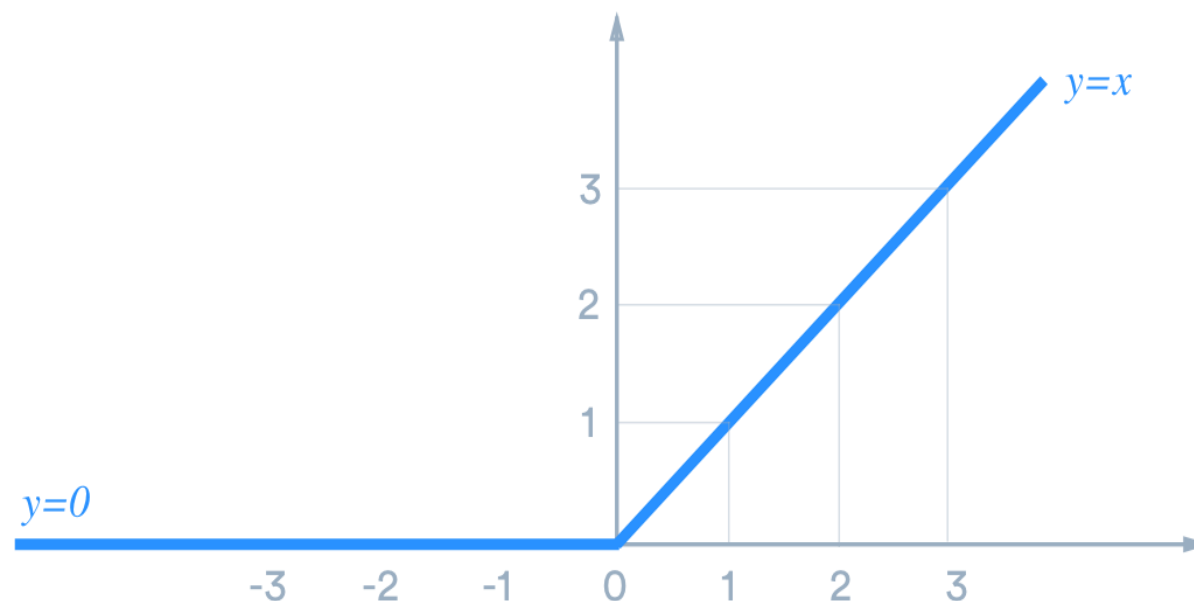
3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[6 de 12]

ReLU: em uma rede neural, conforme mencionado previamente, a função de ativação é responsável por normalizar a saída dos neurônios.

A função de ativação linear retificada (ReLU) é uma função linear que produzirá a entrada 1 se for positiva; caso contrário, produzirá 0. Tornou-se a função de ativação padrão para muitos tipos de redes neurais, porque um modelo que a utiliza é mais fácil de treinar por ser de menor consumo computacional, geralmente obtendo melhor desempenho.



3) Referencial Teórico e Metodologia do Projeto Analítico

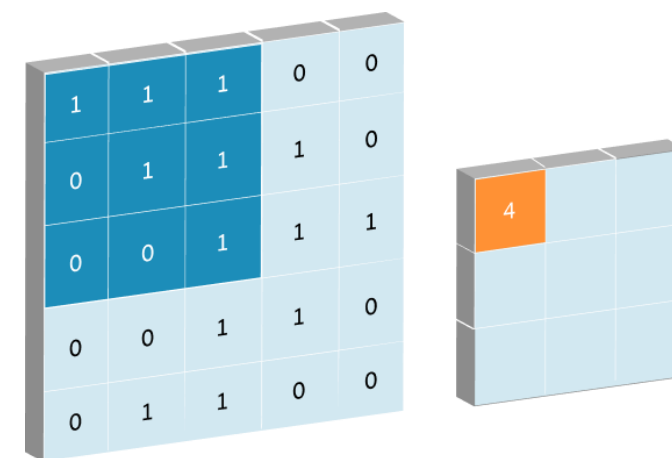
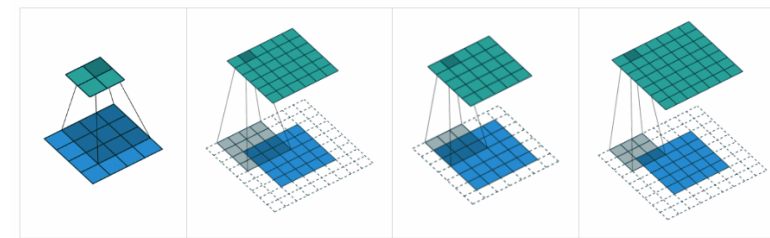
- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[7 de 12]

Conv2D: as camadas convolucionais são os principais componentes usados nas redes neurais convolucionais. Em uma convolução, você movimenta seu kernel sobre a entrada, calcula as multiplicações elemento a elemento e as soma.

O passo desse movimento é chamado de *stride*. A adição de uma borda para que nenhuma informação seja perdida é chamado de *padding*.

De maneira simplificada, uma convolução é a aplicação de um filtro a uma entrada que resulta em uma ativação. A aplicação repetida do mesmo filtro a uma entrada resulta em um mapa de ativações chamado mapa de características, indicando os locais e a força de uma característica detectada em uma entrada, como uma imagem.



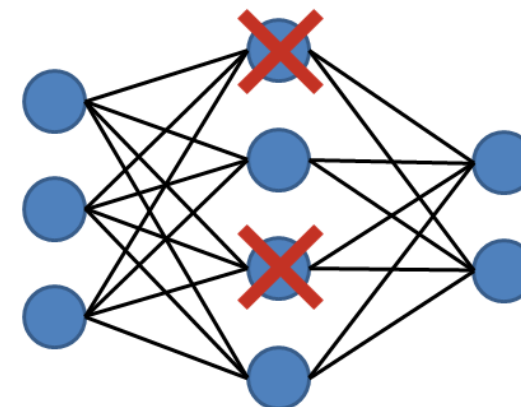
3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (**ler para a turma**)?

[8 de 12]

BatchNormalization: o processo, em si, consiste na normalização da camada ajustando os valores e dimensionando as ativações. Por exemplo, quando temos recursos de 0 a 1 e alguns de 1 a 1000, devemos normalizá-los para uma mesma escala. A normalização de lotes (*batch normalization*) é uma técnica para o treinamento de redes neurais muito profundas que padroniza as entradas em uma camada para cada mini-lote (*batch*). O efeito da técnica é a estabilização do processo de aprendizado e redução do número de épocas necessárias para treinar redes profundas.

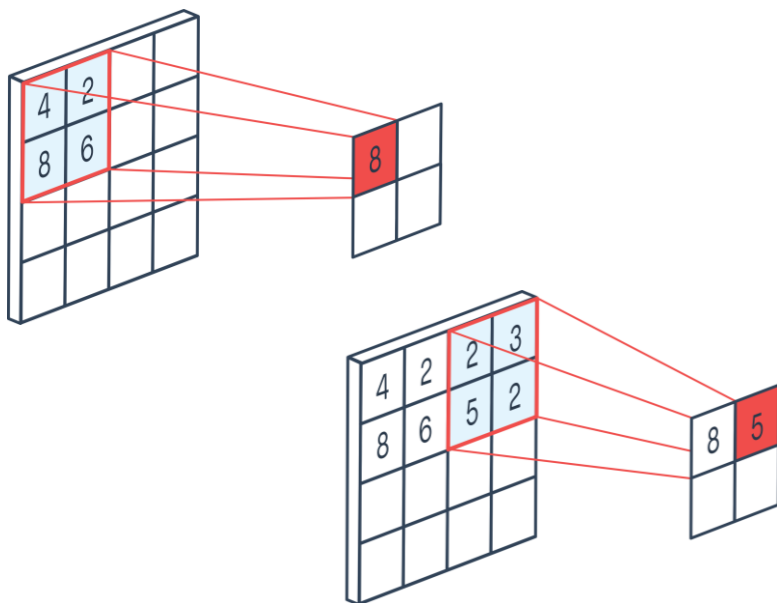
Dropout: é uma técnica usada para combater o *overfitting* (o estado de *overfit* de um modelo faz com que ele não seja generalizável para fora do universo de possibilidades onde foi treinado). O método consiste no descarte aleatório de neurônios de uma camada densa. A técnica é, portanto, usada para regularização.



- Qual foi o **REFERENCIAL TEÓRICO** concebido (**ler para a turma**)?

[9 de 12]

MaxPooling2D: a técnica consiste em um processo de discretização baseado em amostra. O objetivo é fazer uma amostragem reduzida de uma representação de entrada (imagem, matriz de saída de uma camada oculta, etc.), reduzindo sua dimensionalidade e permitindo suposições sobre os recursos contidos nas sub-regiões classificadas.



Isso é feito em parte para ajudar no *overfitting*, fornecendo uma forma simplificada da representação. Além disso, reduz o custo computacional, reduzindo o número de parâmetros a serem aprendidos. O maxpooling é feito aplicando um filtro (máximo) a sub-regiões (normalmente) não sobrepostas da representação inicial.

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[10 de 12]

Accuracy: a acurácia é uma métrica para avaliar modelos de classificação. Informalmente, precisão é a fração de previsões que nosso modelo acertou. Formalmente, a acurácia tem a seguinte definição:

$$\text{Acurácia} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

Para a classificação binária, a acurácia também pode ser calculada em termos de positivos e negativos da seguinte maneira:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN}$$

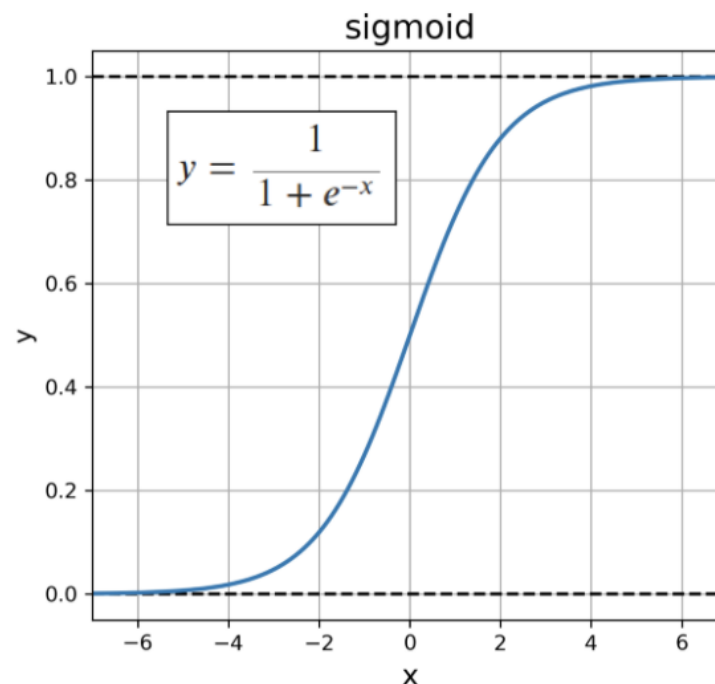
Onde VP = verdadeiros positivos, VN = verdadeiros negativos, FP = falsos positivos e FN = falsos negativos.

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[11 de 12]

BinaryCrossEntropy: é uma função de perda usada para medir os resultados de um modelo binário (nosso caso, com ou sem máscara). É também chamada de **Sigmoid Cross-Entropy loss** (função de perda de entropia cruzada sigmoide) ou **Binary Cross-Entropy loss** (função de perda de entropia binária cruzada).

Sigmoid Activation: a função de ativação sigmoide, também chamada de função logística, é tradicionalmente uma função de ativação muito popular para redes neurais. O resultado dessa ativação é a transformação de uma entrada em um valor entre 0 e 1. Esta função é especialmente usada em casos de classificação binária e é, atualmente, a única função de ativação compatível com a função de perda de entropia cruzada binária.



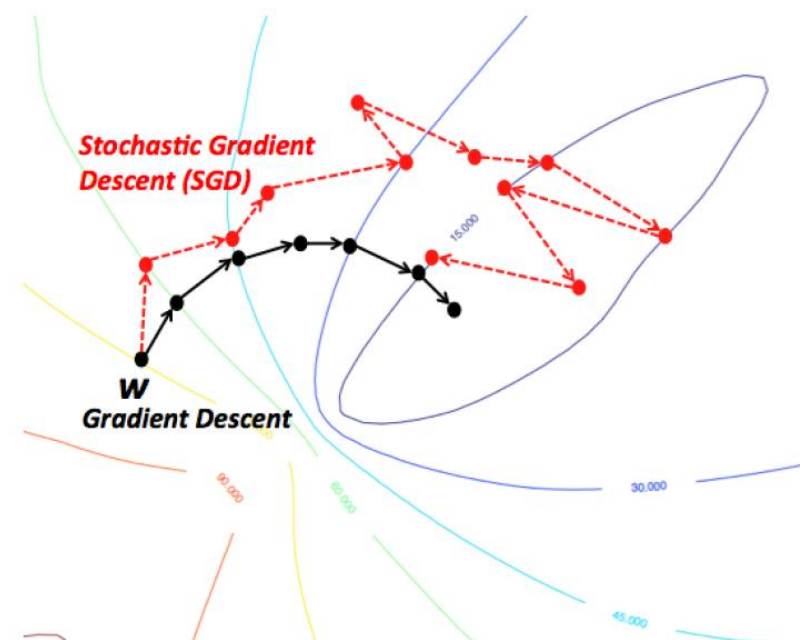
3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual foi o **REFERENCIAL TEÓRICO** concebido (ler para a turma)?

[12 de 12]

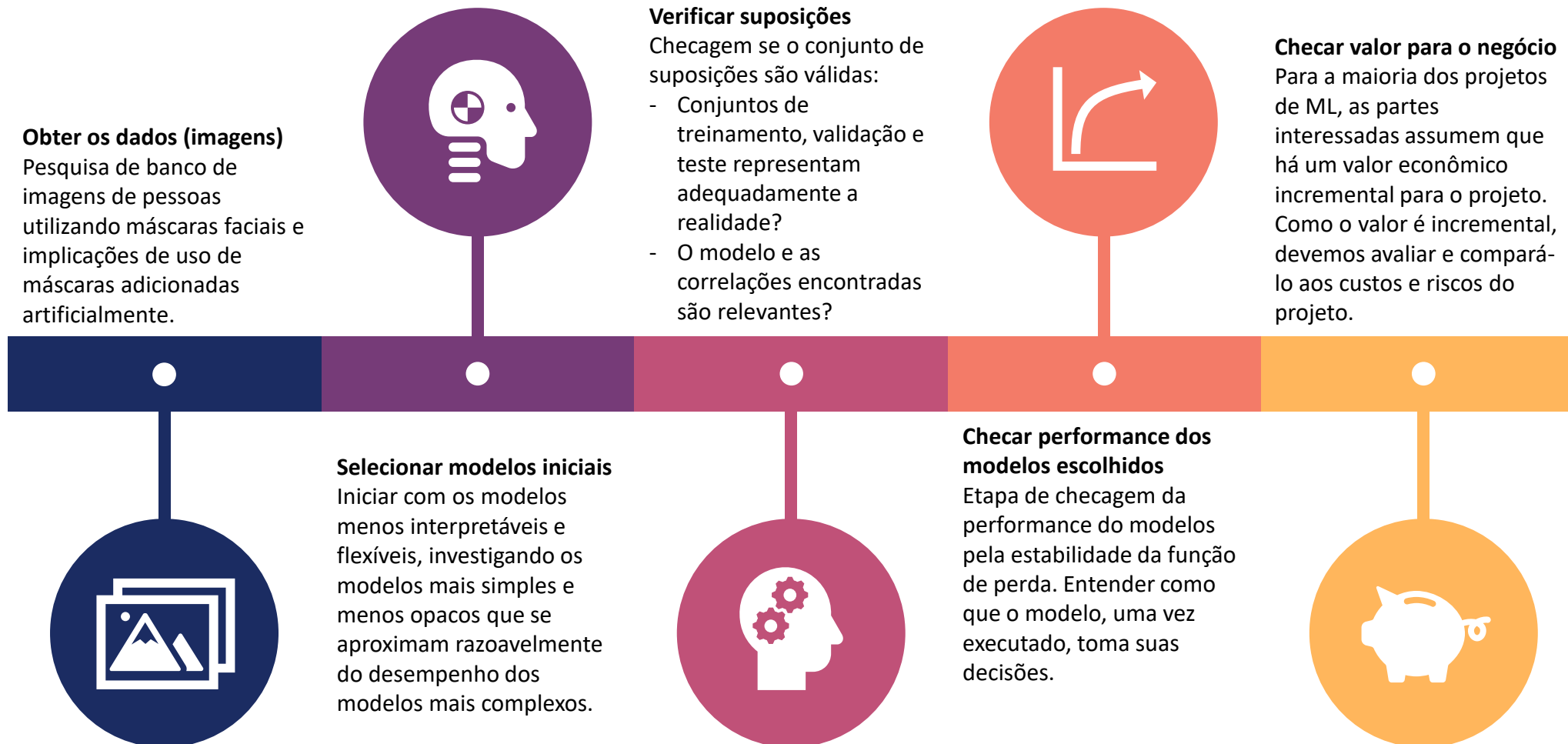
SGD Optimization: a sigla é uma referência à **Stochastic Gradient Descent** (gradiente descendente estocástico) e esta é a metodologia usada para a otimização dos parâmetros do modelo. A palavra "estocástico" significa um sistema ou processo que está vinculado a uma probabilidade aleatória. Portanto, no gradiente descendente estocástico, algumas amostras são selecionadas aleatoriamente em vez de todo o conjunto de dados para cada iteração.

Existe, também, um parâmetro (lote), que indica o número total de amostras de um conjunto de dados usado para calcular o gradiente para cada iteração. Na otimização típica de gradiente descendente, o lote considerado é o conjunto de dados inteiro. Embora o uso de todo o conjunto de dados seja realmente útil para obter os mínimos de uma função de maneira menos ruidosa e menos aleatória, o problema surge quando os conjuntos de dados ficam grandes demais.



3) Referencial Teórico e Metodologia do Projeto Analítico

- Qual é a **METODOLOGIA** (procedimento passo a passo) da pesquisa?



4) Análise dos Dados e Resultados Encontrados

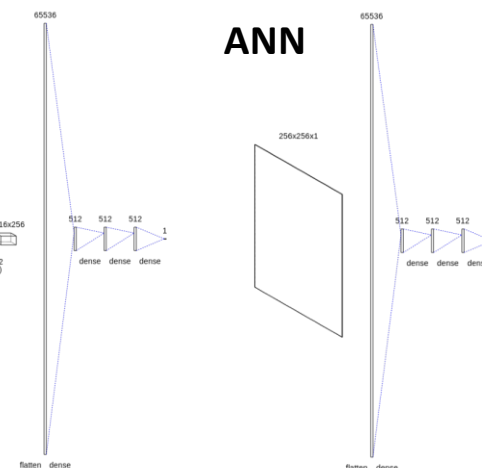
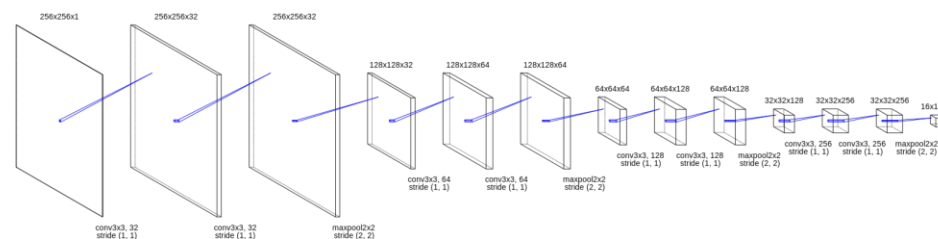
- Quais **RESULTADOS** foram encontrados na análise nos dados?

[1 de 4]

Usando o referencial teórico mencionado anteriormente, treinamos dois tipos de redes, uma rede neural artificial densa tradicional (ANN) e uma rede neural convolucional (CNN). No processo de seleção das nossas finalistas, construímos diferentes profundidades de redes até encontrarmos arquiteturas satisfatórias. Este processo foi não-exaustivo. Ambas tiveram ativação final sigmoide, com função de perda Binary Cross-entropy e otimizador SGD. Para a ANN final usamos camadas Flatten, Dense (ReLU) e Dropout e obtivemos um total de 34.080.769 parâmetros treináveis.

Já para a CNN final usamos camadas Conv2D, Dropout, BatchNormalization, MaxPooling2D, Flatten e Dense (ReLU) e obtivemos um total de 35.254.369 parâmetros treináveis.

Arquiteturas finais: CNN

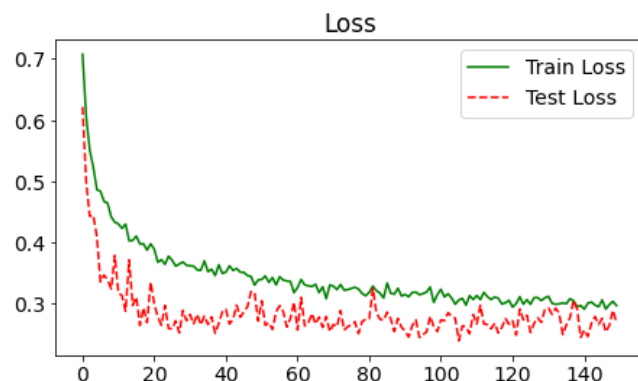


4) Análise dos Dados e Resultados Encontrados

- Quais **RESULTADOS** foram encontrados na análise nos dados?

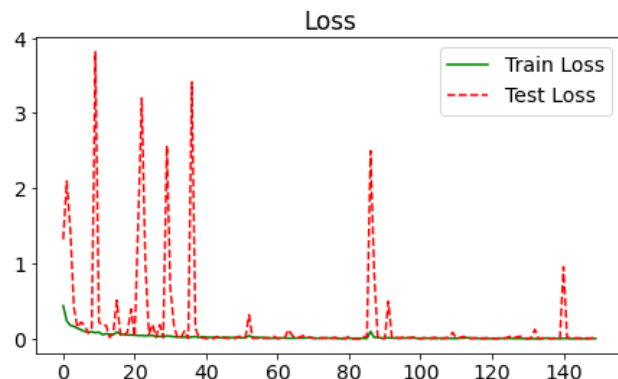
[2 de 4]

ANN



A métrica de performance usada foi acurácia. Ela foi usada tanto para a seleção da arquitetura desejada quanto para a escolha do modelo final. Todos os treinamentos foram feitos com um total de 150 épocas.

CNN



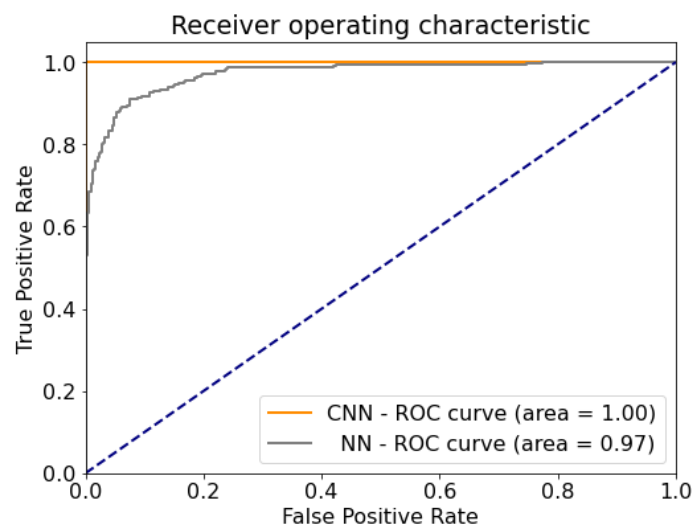
Para as arquiteturas finais, observamos que a CNN converge mais rápido que a ANN a resultados ótimos. Entretanto, observamos que a ANN converge com menor variância da função perda.

4) Análise dos Dados e Resultados Encontrados

- Quais **RESULTADOS** foram encontrados na análise nos dados?

[3 de 4]

Os resultados pós-treino nos mostra que a CNN performou muito melhor que a ANN (avaliado pela curva ROC). Apesar de os resultados da ANN terem sido bastante satisfatórios a rede convolucional, como descrito em literatura (RANJAN, R. et al. e LAWRENCE, S. et al.) costumam ter resultados extraordinários com processamento de imagem.



Fato de não termos tido erros da CNN na base de teste e uma AUROC = 1 nos levou a querer entender melhor como o modelo estava, de fato, decidindo sua classificação. Apesar de termos tomado todos os cuidados necessários para evitarmos o overfit, gostaríamos de nos certificar que o modelo funciona como esperado.

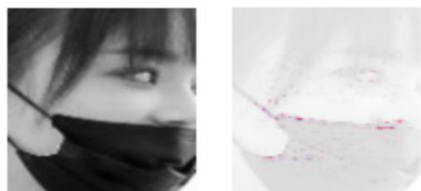
Neural Network		
true label	without_mask	with_mask
	420	21
true label	without_mask	with_mask
	60	381

Convolutional Neural Network		
true label	without_mask	with_mask
	441	0
true label	without_mask	with_mask
	0	441

4) Análise dos Dados e Resultados Encontrados

- Quais **RESULTADOS** foram encontrados na análise nos dados?

[4 de 4]



Para isto, usamos as técnicas SHAP (LUNDBERG, S.; LEE, S.) e Grad-CAM (SELVARAJU, R.; et al.) que visam desmistificar a não interpretabilidade de modelos tidos como "black-box".

Os resultados, à esquerda, mostram que a decisão de uso de máscaras vem exatamente do contorno das máscaras pelo método SHAP (SHapley Additive exPlanations). Já os resultados à direita, usando o método Grad-CAM (Gradient-weighted Class Activation Mapping), nos mostram que, além do contorno das máscaras, o nariz e boca também são relevantes para a identificação do uso delas.

Observamos, portanto, que o comportamento do modelo final está adequado conforme esperado.

