



UNIVERSIDADE FEDERAL DO OESTE DO PARÁ
INSTITUTO DE ENGENHARIA E GEOCIÊNCIAS – IEG
BACHARELADO INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA

KASSYO GABRIEL OLIVEIRA ROCHA DE ANDRADE

UMA ANÁLISE DO YOLO8: DETECÇÃO DE PESSOAS COM MÁSCARAS FACIAIS

SANTARÉM-PA

2024

KASSYO GABRIEL OLIVEIRA ROCHA DE ANDRADE

UMA ANÁLISE DO YOLO8: DETECÇÃO DE PESSOAS COM MÁSCARAS FACIAIS

Trabalho de Conclusão de Curso apresentado ao Bacharelado Interdisciplinar em Ciência e Tecnologia para obtenção do grau de Bacharel em Ciência e Tecnologia na Universidade Federal do Oeste do Pará, Instituto de Engenharia e Geociências.

Orientador: Anderson Alvarenga de Moura Meneses

**SANTARÉM-PA
2024**

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema Integrado de Bibliotecas – SIBI/UFOPA

A554a Andrade, Kassyo Gabriel Oliveira Rocha de
 Uma análise do Yolov8: detecção de pessoas com máscaras faciais. / Kassyo Gabriel Oliveira Rocha de Andrade. - Santarém, 2024.
 16 p. : il.
 Inclui bibliografias.

 Orientador: Anderson Alvarenga de Moura Meneses.
 Trabalho de Conclusão de Curso (Graduação) – Universidade Federal do Oeste do Pará, Instituto de Engenharia e Geociências, Bacharelado em Ciência e Tecnologia.

 1. Desempenho. 2. Imagem. 3. Modelo. 4. Tempo. 5. Versão. I. Meneses, Anderson Alvarenga de Moura. II. Título.

CDD: 23 ed. 005.1



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO OESTE DO PARÁ
INSTITUTO DE ENGENHARIA E GEOCIÊNCIAS**



**AD REFERENDUM Nº 7 / 2024 - IEG
(11.01.09)**

Santarém-PA, 28 de outubro de 2024.

**Nº do Protocolo:
23204.014144/2024-38**

O presidente do Colegiado de Cursos do Programa de Ciência e Tecnologia* do Instituto de Engenharia e Geociências (IEG) da Universidade Federal do Oeste do Pará (Ufopa), no uso de suas atribuições legais, e considerando:

O atendimento ao princípio da continuidade e eficiência administrativa do serviço público, para que essa decisão produza efeitos de forma tempestiva na vigência prevista;

Considerando o Regimento de TCC dos cursos do Programa de Ciência e Tecnologia, em conformidade com o Projeto Pedagógico do Curso (PPC) de Bacharelado Interdisciplinar em Ciência e Tecnologia;

Considerando que o trabalho "UMA ANÁLISE DO YOLOV8: DETECÇÃO DE PESSOAS COM MÁSCARA FACIAIS" foi publicado por meio digital no XXVI Encontro Nacional de Modelagem Computacional e Encontro de Ciência e Tecnologia dos Materiais, realizado no período de 25 a 27 de outubro de 2023.

MANIFESTA-SE, ad referendum:

1) Pela aprovação do Trabalho de Conclusão de Curso "UMA ANÁLISE DO YOLOV8: DETECÇÃO DE PESSOAS COM MÁSCARA FACIAIS" do aluno KASSIO GABRIEL OLIVEIRA ROCHA DE ANDRADE - 2021001353, na modalidade Publicação em Evento Científico, com orientação do docente ANDERSON, vinculado ao período 2024.1.

2) Pela atribuição da nota 7,0 ao TCC supracitado.

Observação: A nota para TCC na modalidade de publicação em evento científico é definida pelo Regimento de TCC do Programa de Ciência e Tecnologia. Para pleitear uma nota maior o aluno deve solicitar a composição de uma banca de avaliação para a realização de uma avaliação pública. O discente não realizou a solicitação.

* Bacharelado Interdisciplinar em Ciência e Tecnologia, Bacharelado em Engenharia Física e Bacharelado em Engenharia Mecânica

*(Assinado digitalmente em
28/10/2024 15:35)
UBIRAEALSON DE
LIMA RUELA
COORDENADOR DE
CURSO - TITULAR IEG
(11.01.09)
Matrícula: 2060289*

Visualize o documento original em
<https://sipac.ufopa.edu.br/public/documentos/index.jsp> informando
seu número: **7**, ano: **2024**, tipo: **AD REFERENDUM**, data de emissão:
28/10/2024 e o código de verificação: **606ceb2304**

RESUMO

A pesquisa teve como objetivo a aplicação do método Yolo (you only look once) para fazer a detecção de objetos específicos em diversas imagens, uma vez que a visão computacional avança rapidamente e com o passar dos anos são feitas melhorias de performance para que seja utilizada na área da saúde ou construção, por exemplo. Para a realização da pesquisa foi necessário ter embasamento teórico sobre a tal tecnologia e suas diversas funções e funcionamentos como: classificação, detecção, segmentação, trackeamento e pose através do github da empresa "ultralytics" a qual desenvolveu a tecnologia. Na pesquisa foi utilizado a detecção para fazer a distinção de pessoas que estão e não estão utilizando máscaras faciais. Os resultados obtidos foram considerados avaliando duas métricas: Map50 e tempo em dois tipos de GPUs: A100 e V100, nos resultados foi observado que utilizando a GPU V100 teve Map50 médio de 0.862 e tempo médio de treinamento do modelo de 56,77 minutos, enquanto a GPU A100 teve Map50 médio de 0.866 e tempo médio de treinamento do modelo de 13,06 minutos, os resultados foram realizados para 100 épocas.

Palavras-Chave: Desempenho, Imagem, Modelo, Tempo, Versão

ABSTRACT

The research aimed to apply the Yolo method (you only look once) to detect specific objects in different images, since computer vision advances rapidly and over the years performance improvements are made so that it can be used in the area of health or construction, for example. In order to carry out the research, it was necessary to have a theoretical basis on this technology and its various functions and operations, such as: classification, detection, segmentation, tracking and pose through the github of the company "ultralitics" which developed the technology. In the research, detection was used to distinguish between people who are and are not using face masks. The obtained results were considered evaluating two metrics: Map50 and time in two types of GPUs: A100 and V100. In the results it was observed that using the GPU V100 had an average Map50 of 0.862 and average training time of the model of 56.77 minutes, while the GPU A100 had an average Map50 of 0.866 and average training time of the model 13.06 minute model, both results for 100 epochs.

Keywords: Model, Version, Image, Performance, Time.

SUMÁRIO

1. INTRODUÇÃO.....	7
2. OBJETIVOS.....	8
2.1 Objetivo geral.....	8
2.2 Objetivos específicos.....	8
3. REFERENCIAIS TEÓRICOS.....	8
3.1 Inteligência artificial.....	8
3.2 Deep learning.....	8
3.3 Anchor free detection.....	9
3.4 Modelos.....	9
4. METODOLOGIA.....	9
4.1 Fluxograma de trabalho.....	9
4.2 Ambiente de desenvolvimento.....	10
4.3 Label Img.....	10
4.4 Equipamentos.....	11
4.5 Métricas de desempenho.....	11
4.6 Teste estatístico.....	11
5. RESULTADOS.....	11
6. CONCLUSÃO.....	15
Reconhecimentos.....	14
REFERÊNCIAS.....	14



UMA ANÁLISE DO YOLOV8: DETECÇÃO DE PESSOAS COM MÁSCARAS FACIAIS

Kassyo Gabriel Oliveira Rocha de Andrade¹ – kassyo.andrade@icloud.com

Yana dos Santos Pereira¹ – yana.pereira@discente.ufopa.edu.br

Carlos Henrique Pinto da Silveira² – carlos.silveira@discente.ufopa.edu.br

Anderson Alvarenga de Moura Meneses^{1, 3} – anderson.meneses@ufopa.edu.br

¹ Laboratório de Inteligência Computacional, Universidade Federal do Oeste do Pará – Santarém, PA, Brasil

² Laboratório de Modelagem Computacional, Universidade Federal do Oeste do Pará – Santarém, PA, Brasil

³ Programa de Pós-Graduação em Sociedade, Natureza e Desenvolvimento, Universidade Federal do Oeste do Pará - Santarém, PA, Brasil

Resumo

A pesquisa teve como objetivo a aplicação do método Yolo (you only look once) para fazer a detecção de objetos específicos em diversas imagens, uma vez que a visão computacional avança rapidamente e com o passar dos anos são feitas melhorias de performance para que seja utilizada na área da saúde ou construção, por exemplo. Para a realização da pesquisa foi necessário ter embasamento teórico sobre a tal tecnologia e suas diversas funções e funcionamentos como: classificação, detecção, segmentação, trackeamento e pose através do github da empresa "ultralytics" a qual desenvolveu a tecnologia. Na pesquisa foi utilizado a detecção para fazer a distinção de pessoas que estão e não estão utilizando máscaras faciais. Os resultados obtidos foram considerados avaliando duas métricas: Map50 e tempo em dois tipos de GPUs: A100 e V100, nos resultados foi observado que utilizando a GPU V100 teve Map50 médio de 0.862 e tempo médio de treinamento do modelo de 56,77 minutos, enquanto a GPU A100 teve Map50 médio de 0.866 e tempo médio de treinamento do modelo de 13,06 minutos, os resultados foram realizados para 100 épocas.

Palavras-chave: Modelo, Versão, Imagem, Desempenho, Tempo.

1. INTRODUÇÃO

Em 2020 aconteceu a pandemia do covid-19, pessoas perderam entes queridos e com as perdas foi necessário redobrar a atenção em relação ao uso de máscaras, álcool em gel e outros itens de segurança. Com isso o avanço da aprendizagem de máquina ao passar dos anos foi essencial, visto que diversos mecanismos de auxílio à proteção da população não precisavam mais ser feitos manualmente por pessoas, e sim por máquinas.

Com o lockdown e o número de circulação de pessoas reduzidos em locais fechados como: salas de aula, reuniões, hospitais etc, foi necessário que existisse um software que fizesse a detecção de pessoas usando os equipamentos necessários para frequentar tais ambientes, com isso acontecendo a yolo teve grande importância para desenvolver tecnologia para monitoramento em tempo real de pessoas que estão e não estavam utilizando equipamentos adequados para frequentar determinados ambientes e com alta performance, havendo uma taxa de acerto alta (Zhenggon et al., 2022). Então, pode-se dizer que o Yolo teve um grande papel durante a pandemia.

A cada ano foram feitas melhorias para que tivesse inúmeras utilidades, como por exemplo no uso de detecção de animais, detecção de rachaduras em uma construção civil, detecção de armas de fogo, e principalmente na área da saúde como detecção de câncer e detecção de pessoas utilizando máscaras a qual é o foco da pesquisa.

2. OBJETIVOS

2.1 Objetivo geral

Os objetivos gerais da pesquisa consiste em utilizar o método Yolo para fazer a detecção em mais de 100 imagens de teste de pessoas que estão utilizando máscaras faciais e as que não estão utilizando máscaras faciais, assim como avaliar a qualidade do modelo treinado para fazer tais detecções.

2.2 Objetivos específicos

- Implementar o estudo de caso
- Comparar os modelos: small, medium e large quanto a fluidez, velocidade e performance
- Verificar se houve mudança de performance entre as duas GPUs utilizadas
- Realizar testes estatísticos entre os modelos

3. REFERENCIAIS TEÓRICOS

Esta etapa da pesquisa aborda conceitos importantes para conseguir utilizar o método da melhor maneira possível . As principais são: IA, Deep learning, Bounding box, anchor free detection e modelos.

3.1 Inteligência artificial

A inteligência artificial (IA) é uma subárea da ciência da computação responsável por pesquisar e propor a elaboração de dispositivos computacionais capazes de simular aspectos do intelecto humano, ao modo da capacidade de raciocinar, perceber, tomar decisões e resolver problemas (Silva, 2013).

3.2 Deep learning

Deep learning é uma subárea de aprendizagem de máquina que investiga técnicas para simular o comportamento do cérebro humano em tarefas como reconhecimento visual, reconhecimento de fala e processamento de linguagem natural (Bezerra, 2016). Algumas aplicações são: Detecção de fraudes, diagnósticos médicos etc.

3.3 Anchor free detection

Significa que o método YoloV8 prevê diretamente o centro de um objeto ao invés do deslocamento de uma anchor box já conhecida, ou seja, faz com que o processo de detecção seja mais rápido e eficiente que as versões anteriores.

3.4 Modelos

Ao todo possuem cinco modelos de detecção, segmentação, classificação e pose, que são: nano, small, medium, large, extra large. A versão nano é a que possui maior velocidade e requer uma quantidade menor de recursos gráficos, enquanto que a versão extra large é a mais precisa, porém, é lenta e requer uma grande quantidade de recursos gráficos (Ultralytics, 2023).

Model	Size (pixels)	mAP val	Speed CPU (ms)	Speed A100 (ms)	Params (M)	FLOPs (B)
YOLOv8n	640	37,3	80,40	0,99	3,2	8,7
YOLOv8s	640	44,9	128,4	1,20	11,2	28,6
YOLOv8m	640	50,2	234,7	1,83	25,9	78,9
YOLOv8l	640	52,9	375,2	2,39	43,7	165,2
YOLOv8x	640	53,9	479,1	3,53	68,2	257,8

Figura 1 - Modelos disponíveis e suas métricas. Fonte: Ultralytics 2023.

4. METODOLOGIA

Para realização do treinamento da YoloV8 para o estudo de caso foi necessário verificar no GitHub da empresa que criou (ultralytics) as funções que podem ser utilizadas com esse método.

4.1 Fluxograma de trabalho



Figura 2 - Fluxograma de trabalho. Fonte: Autor

Os objetivos destinados para a pesquisa foram conseguir fazer a detecção de pessoas utilizando ou não máscaras faciais, já para o estudo de caso foram utilizados o modo de detecção com os modelos "medium" e "large" pois ambos possuem bons parâmetros, em seguida o dataset utilizado foi obtido no site "kaggle" que possui 3 tipos de conjunto de dados: Conjunto de treinamento, conjunto de validação e conjunto de teste. O conjunto de treinamento possui 700 imagens, enquanto o de validação 100 imagens e teste 200 imagens. No treinamento do modelo foram utilizados dois tipos de GPU: Tesla A100 e Tesla V100. Com cada GPU foram treinados os modelos "medium" e "large" por cem épocas cada. Os resultados do treinamento ficaram salvos no google drive para que seja de fácil acesso. Por último foi aplicado o teste estatístico de de Mann Whitney para comparar os dados obtidos de tempo e performance.

4.2 Ambiente de desenvolvimento

Foi utilizado o Google Colaboratory pelo motivo de que a utilização da ferramenta é simplificada, pois não necessita configuração para utilização, conta com bibliotecas pré instaladas e o compartilhamento do código é facilitado.

4.3 Label Img

Para realizar o treinamento é necessário utilizar um programa para criar manualmente as bounding boxes, imagem por imagem. O passo a passo para instalação da ferramenta está disponível no site "github.com/HumanSignal/labelimg". É necessário rotular nas imagens os objetos que são importantes para o estudo, ou seja, é necessário criar um bounding box para n objetos presentes em cada imagem de treinamento e validação. Ao final do processo são criados n arquivos txt com as posições dos objetos correspondentes que vão variar de acordo com a quantidade de imagens que o dataset possui, os arquivos são necessários no momento em que for realizado o treinamento para que o processo seja facilitado.



Figura 3 - Interface da ferramenta label img. Fonte: Autor

4.4 Equipamentos

Os equipamentos necessários para treinamento do conjunto de dados foi utilizado pelo Google Colaboratory pro que disponibiliza dois tipos de GPUs: Tesla A100 e Tesla V100. As bibliotecas necessárias para otimizar a pesquisa foram: OS, matplotlib, time, cv2 e random.

4.5 Métricas de desempenho

Na pesquisa foi utilizado como métrica de desempenho o map50, baseada no cálculo entre a precisão e a recall de um modelo. As equações podem ser observadas abaixo:

$$\text{Map50} = \frac{\text{precisão}}{\text{recall}} \quad (1)$$

$$\text{Precisão} = \frac{\text{Verdadeiro positivo}}{\text{Verdadeiro positivo} + \text{Falso positivo}} \quad (2)$$

$$\text{Recall} = \frac{\text{Verdadeiro positivo}}{\text{Verdadeiro positivo} + \text{falso negativo}} \quad (3)$$

Utilizada como métrica nos principais detectores de objetos, o map50 é fundamental para comprovar o quão eficaz o modelo foi treinado e consegue realizar as predições determinadas. Os valores de avaliação da métrica map50 variam entre 0 e 1, quanto mais próximo de 1, melhor o modelo é. Baseada nas submétricas: Matriz de confusão, intersecção sobre união (IoU), recall e precisão, em que a matriz de confusão é baseada em 4 atributos: Verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo, já a intersecção corresponde a área entre a bounding box prevista e real entre os objetos em estudo (Shah, 2022).

4.6 Teste estatístico

O teste estatístico utilizado foi o de Mann Whitney (wilcoxon rank-sum test) que é um teste não paramétrico aplicado para duas amostras independentes, existem duas hipóteses para o teste de Mann Whitney.

- Sob hipótese nula: não existe diferença significativa entre os grupos
- Sob hipótese alternativa: há diferença significativa entre os grupos

Para escolher sobre qual hipótese o modelo se enquadra é necessário avaliar o valor de U que testa a igualdade entre a mediana das amostras. Os valores de U calculados pelo teste avaliam o grau em que as amostras se entrelaçam, quanto mais distantes estão os valores, maior a distinção entre as amostras. Para o teste foram comparados tempo de treinamento entre os modelos "medium" e "large" e map50 entre os mesmos modelos citados.

5. RESULTADOS

A tabela 2 apresenta os resultados obtidos a partir do treinamento do modelo "medium". Cada treinamento foi realizado por cem épocas utilizando a GPU Tesla A100 disponibilizado pelo Google Colaboratory pro (figura 3).

Tabela 1 - Resultados do modelo "medium".

medium	Map50	Tempo(min)
1	0,837	12.76
2	0,836	13.33
3	0,862	14.31
4	0,848	14.33
5	0,861	14.35
Média	0,8488	13,816
Mediana	0,848	14,31
Desvio Padrão	0,011	0,65
Máximo	0,836	14,35
Mínimo	0,862	12,76

Fonte: Autor

A figura 4 mostra a confiabilidade do modelo treinado, relacionando a precisão e a recall do modelo em um gráfico, enquanto as linhas estiverem mais próximas do ponto (1,1) mais confiável é o modelo.

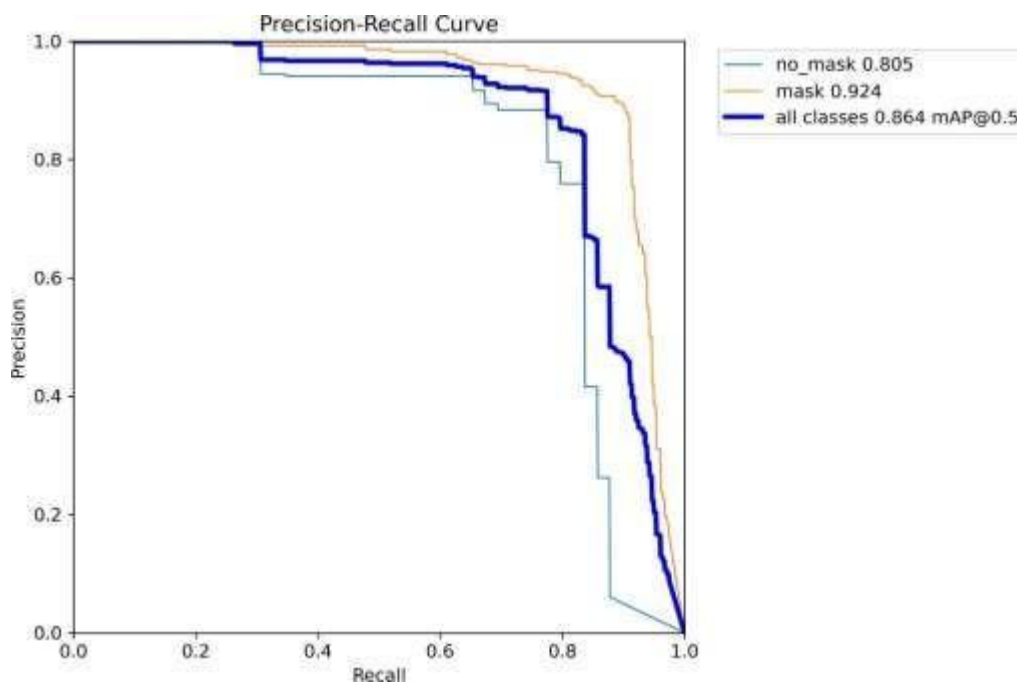


Figura 4 - Relação entre a precisão e a recall para as imagens de pessoas usando ou não usando máscaras faciais. Fonte: Autor.

A tabela 2 apresenta os resultados obtidos a partir do treinamento do modelo "large". Cada treinamento foi realizado por cem épocas utilizando a GPU A100 disponibilizada pelo Google Colaboratory pro (figura 6).

Tabela 2 - Resultados do modelo "large".

large	Map50	Tempo(min)
1	0,877	16,25
2	0,859	16,63
3	0,856	17,78
4	0,855	16,41
5	0,848	16,4
Média	0,859	16,694
Mediana	0,856	16,41
Desvio Padrao	0,0096	0,59
Máximo	0,877	17,78
Mínimo	0,848	16,25

Fonte: Autor.

A figura 5 mostra a confiabilidade do modelo treinado relacionando a precisão e a recall do modelo em um gráfico, enquanto as linhas estiverem mais próximas do ponto (1,1) mais confiável é o modelo.

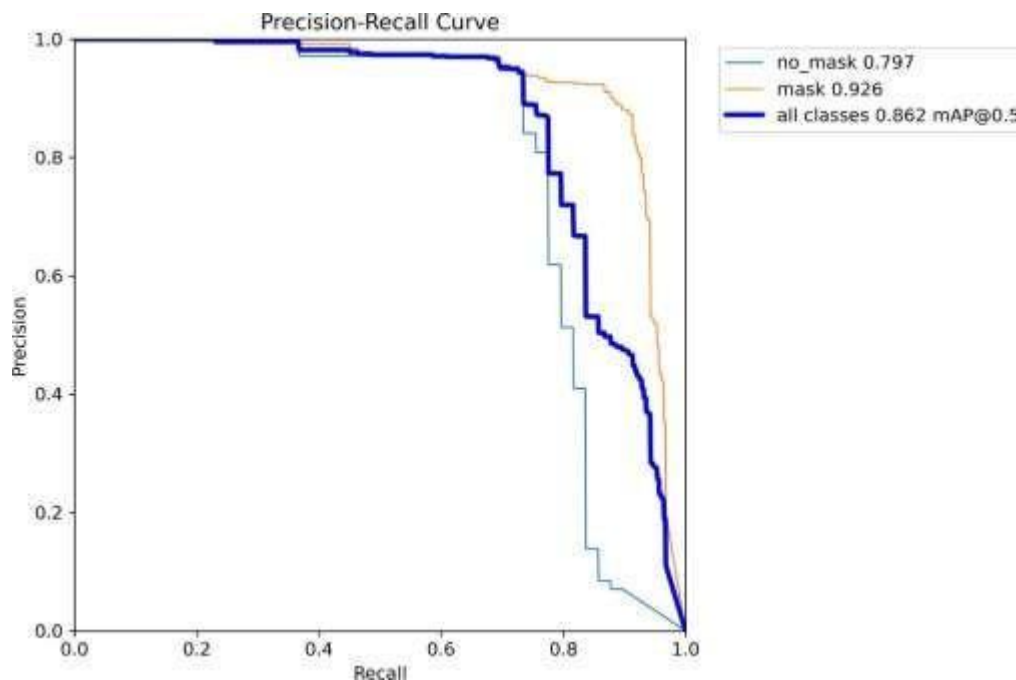


Figura 5 - Relação entre a precisão e a recall para as imagens de pessoas usando ou não máscaras faciais. Fonte: Autor.

Intuitivamente é normal pensar que o modelo "large" é superior em vários quesitos em relação a versão "medium", porém, fazendo repetidos testes é notório que não há uma diferença significativa entre os modelos por conta do peso da rede neural que pode acabar não tendo um bom rendimento com o dataset em questão. O dataset trabalhado contém setecentas imagens para treinamento e cem imagens para validação, então não necessariamente será preciso uma rede pesada para fazer a detecção de "poucas" imagens, por isso não houve uma diferença significativa entre os modelos, supondo que o dataset em questão tivesse uma quantidade de imagens três vezes maior por exemplo, haveria a necessidade de usar um modelo mais bruto (Cimirro,2022). Foi realizado o teste estatístico de Mann Whitney para verificar se havia diferença significativa entre os tempos que foram necessários para o treinamento ser realizado dos modelos "medium" e "large" e outro teste para verificar se havia diferença significativa entre o map50 dos modelos "medium" e "large". Os resultados mostraram que apenas o tempo entre

os modelos teve diferença	$\alpha = 0,$	O resultado de "p"
significativa, assumindo	05.	para o

tempo foi de 0,0079, mostrando que H_0 foi rejeitado e consequentemente temos evidências de que há diferença significativa. O resultado de "p" para o map50 foi de 0,46, então, não H_0 não deve ser rejeitada, mostrando que não há evidência suficiente para comprovar que existe diferença significativa.

A seguir estão algumas figuras de teste utilizando os resultados modelo "large".



Figura 6 - Imagem teste usando o método Yolo treinado. Fonte: Autor.



Figura 7 - Imagem teste usando o método Yolo treinado. Fonte: Autor.

Na Figura 6 é possível observar que é feita a detecção da pessoa com 90% de probabilidade e ao lado da placa verde o Yolo consegue fazer a detecção de uma pessoa em um cartaz e classifica que a pessoa não está utilizando máscara. Na Figura 7 é possível observar que o Yolo consegue fazer a detecção de maneira precisa mesmo tendo diversas pessoas em uma imagem, variando a probabilidade entre 35% a 79% de acerto. E também consegue fazer a detecção de uma única pessoa ao lado direito da imagem que não está utilizando máscara.

6. CONCLUSÃO

A partir dos resultados obtidos pode-se concluir que o método yolo conseguiu obter resultados satisfatórios visto que na figura 9 por exemplo foi feita a detecção e classificação de todas as pessoas que estão e não estão usando máscaras faciais e com grande confiança, ademais, como dito anteriormente é possível obter resultados ainda melhores, porém, é necessário obter um dataset com um número maior de imagens de treinamento.

Reconhecimentos

CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo apoio.

REFERÊNCIAS

- Bezerra, E. (2016). "Introdução à aprendizagem profunda". Em: OGASAWARA, E; VIEIRA, V. (Eds.). Tópicos em Gerenciamento de Dados e Informação. Porto Alegre: Sociedade Brasileira de Computação, 2016. p. 57.
- Carmo, V. (2012). Teste de Mann Whitney. Disponível em: www.inf.ufsc.br/~vera.carmo/Teste_de_Hipoteses/Testes_nao_parametricos_Mann-Whitney.pdf. Acesso em: 30 de mai. de 2023.
- Cimirro, J. (2022). "Reconhecimento de imagens: Uso do método yolo no reconhecimento de placas de trânsito". Tese de conclusão de curso, UNIPAMPA, Bagé.
- Luvsandorj, Z. (2021). Precision recall curve explained. Disponível em: levelup.gitconnected.com/precision-recall-curve-explained-fabfe58fb52e. Acesso em: 2 de jun. de 2023
- Métricas de avaliação em machine learning: Classificação. (2020). Disponível em: medium.com/kunumi/métricas-de-avaliação-em-machine-learning-classificação-49340dcd198. Acesso em: 31 de jun de 2023
- Silva, R. (2013). Inteligência Artificial. Enciclopédia da conscienciologia. Disponível em: repositorios.org/handle/123456789/3737. Acesso em: 5 de jul. de 2023
- Zhengong, H et al. (2022). "SMD-YOLO: An efficient and lightweight detection method for mask wearing status during the COVID-19 pandemic". Computer methods and Programs in Biomedicine, vol 221.
- Shah, D. (2022). Mean Average Precision explained: Everything you need to know. Disponível em: [https://www.v7labs.com/blog/mean-average-precision#:~:text=let%27s%20dive%20in!-,What%20is%20Mean%20Average%20Precision%20\(mAP\)%3F,values%20from%200%20to%201](https://www.v7labs.com/blog/mean-average-precision#:~:text=let%27s%20dive%20in!-,What%20is%20Mean%20Average%20Precision%20(mAP)%3F,values%20from%200%20to%201). Acesso em: 21 de jul. de 2023.

AN ANALYSIS OF YOLOV8: DETECTION OF PEOPLE WITH FACIAL MASKS

Abstract.

The research aimed to apply the Yolo method (you only look once) to detect specific objects in different images, since computer vision advances rapidly and over the years performance improvements are made so that it can be used in the area of health or construction, for example. In order to carry out the research, it was necessary to have a theoretical basis on this technology and its various functions and operations, such as: classification, detection, segmentation, tracking and pose through the github of the company "ultralytics" which developed the technology. In the research, detection was used to distinguish between people who are and are not using face masks. The obtained results were considered evaluating two metrics: Map50 and time in two types of GPUs: A100 and V100. In the results it was observed that using the GPU V100 had an average Map50 of 0.862 and average training time of the model of 56.77 minutes, while the GPU A100 had an average Map50 of 0.866 and average training time of the model 13.06 minute model, both results for 100 epochs.

Keywords: Model, Version, Image, Performance, Time.