

Grupo de Pesquisas em Sistemas Inteligentes
Laboratório de Sistemas Inteligentes
Escola Superior de Tecnologia
Universidade do Estado do Amazonas



Um Estudo de Caso da Detecção do Uso de Máscaras Faciais com Redes Neurais Convolucionais Regionais

Diego Lucena de Medeiros, Elloá B. Guedes, Carlos Maurício S. Figueiredo

{dldm.eng18, ebgcosta, cfigueiredo}@uea.edu.br

III Workshop Brasileiro de Cidades Inteligentes

CSBC 2022 – Niterói – Rio de Janeiro

Introdução

- Contexto da Pandemia do COVID-19
 - Fardo avassalador para a saúde em todo o planeta
 - Uso de máscaras faciais mitiga o contágio (Prather *et al.* 2020)
 - Máscaram são comprovadamente efetivas ainda quando o distanciamento social não podia ser assegurado (Kwon *et al.* 2021)

- Organização das Nações Unidas (ONU 2018)
 - 57 % da população mundial vive em áreas urbanas
 - Até 2050 essa proporção será de 68 %
 - Novas doenças altamente contagiosas podem surgir em áreas densamente povoadas

Introdução

- Gestão da Informação pode ser tão vital quanto qualquer outro elemento dos sistemas de Saúde
 - Dados fornecidos pelas cidades podem ser cruciais (Costa & Peixoto 2020)

Introdução

- Gestão da Informação pode ser tão vital quanto qualquer outro elemento dos sistemas de Saúde
 - Dados fornecidos pelas cidades podem ser cruciais (Costa & Peixoto 2020)

Cidades Inteligentes (CIs)

Área urbana que usa a informação coletada por diferentes tipos de sensores para monitorar e administrar os recursos disponíveis de forma eficiente, com capacidade de aprender e se adaptar continuamente, melhorando as condições de conforto e segurança para as pessoas que nela habitam (Du *et al.* 2019)

Introdução

- Espera-se que o desenvolvimento sustentável das CIs forneça **inteligência situacional** e uma **resposta direcionada automatizada** para garantir a segurança da saúde pública para as pandemias atuais e futuras (Shorfuzzaman *et al.* 2021)

Introdução

- Espera-se que o desenvolvimento sustentável das CIs forneça **inteligência situacional** e uma **resposta direcionada automatizada** para garantir a segurança da saúde pública para as pandemias atuais e futuras (Shorfuzzaman *et al.* 2021)

Solução Proposta

- Solução de monitoramento para CIs que ajuda a mitigar a disseminação de doenças infecciosas cujas estratégias de prevenção contemplem o uso de máscaras faciais
- Tarefa de Detecção de Objetos com R-CNNs da Família YOLO
- Validação e avaliação perante bases de dados realísticas e diversas

Trabalhos Relacionados

- Tarefa de Visão Computacional para detecção de objetos
 - Localização da face
 - Classificação binária (com máscara vs sem máscara)

- Grande parte das soluções baseadas em *Deep Learning* com o uso de Redes Neurais Convolucionais Regionais (R-CNNs)
 - YOLO: Família de arquiteturas de R-CNNs *single-shot* (Redmond *et al.* 2016)
 - Previsão simultânea das caixas delimitadoras e da classificação das regiões
 - Detecção em tempo real

Trabalhos Relacionados

- Loey *et al.* 2021
 - YOLOv2 com ResNet-50
- Singh *et al.* 2021
 - Comparação da YOLOv3 com detectores de dois estágios
 - Duas bases de dados abertas e rotuladas pelos próprios autores
- Mahukar & Gadge 2021: YOLOv4
- Yang *et al.* 2020: YOLOv5

Trabalhos Relacionados

- Loey *et al.* 2021
 - YOLOv2 com ResNet-50
- Singh *et al.* 2021
 - Comparação da YOLOv3 com detectores de dois estágios
 - Duas bases de dados abertas e rotuladas pelos próprios autores
- Mahukar & Gadge 2021: YOLOv4
- Yang *et al.* 2020: YOLOv5

Survey de Nowrin *et al.* 2021

- Grande variedade de bases de dados e soluções para o detecção de máscaras faciais
- **Problema:** A ausência de um *benchmark* dificulta uma comparação objetiva entre as soluções propostas

Materiais e Métodos

- Análise comparativa entre diferentes modelos da família YOLO sob diferentes bases de dados publicamente disponíveis
- **Dados Experimentais:**
 - **Mask Dataset (Mask)**. 848 imagens rotuladas com sujeitos de diferentes idades, em condições não controladas e com tipos diferentes de máscaras ([MakeML 2020](#))
 - **Face Mask Dataset (Face Mask)**. 920 exemplos coletados e rotulados a partir da consulta de imagens em buscadores. Mais de um sujeito por imagem ([Purohit 2020](#))
 - **MaskDetection at YOLO format (Mask Detection)**. Base de dados com 1226 imagens anotadas e previamente particionadas ([Lorenzo 2020](#))
 - **Base Unificada**: União das bases anteriores

Materiais e Métodos: Dados Experimentais



(a) Mask

(b) Face Mask

(c) Mask Detection

Figura 1: Exemplos de diferentes imagens oriundas das três bases de dados.

Materiais e Métodos: Dados Experimentais

Tabela 1: Descrição estatística dos *pixels* das *bounding boxes* nas bases de dados selecionadas.

	Comprimento			Largura		
	Média e Desvio Padrão	Máx	Mín	Média e Desvio Padrão	Máx	Mín
Mask	34.95 ± 32.64	340	2	31.07 ± 27.91	317	1
Face Mask	95.34 ± 108.93	2016	8	83.09 ± 89.24	1612	7
Mask Detection	163.29 ± 140.42	847	11	165.04 ± 139.48	1337	10

Materiais e Métodos: Arquiteturas, Parâmetros e Hiperparâmetros

- **YOLOv3**: Melhoria continuada do modelo original (Redmond & Farhadi 2018)
 - YOLOv3 Small
 - YOLOv3
- **YOLOv5**: Estado da arte dessa família (Jocher *et al.* 2020)
 - YOLOv5 Nano 6
 - YOLOv5 Small 6
 - YOLOv5 Medium 6
- Máximo de 1000 épocas de treinamento com *early stopping*
- Sem uso de transferência de aprendizado

Materiais e Métodos: Avaliação de Desempenho

- Validação cruzada do tipo *holdout*
- Métricas de avaliação:
 - Precisão
 - Revocação
 - F_1 -Score
 - *Average Precision* (AP)
 - *Mean Average Precision* mAP@0.5

Materiais e Métodos: Avaliação de Desempenho

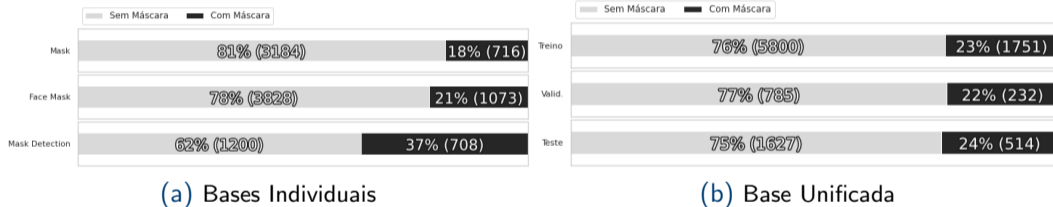


Figura 2: Propoção de exemplos por classe nas bases de dados.

Resultados e Discussão

Tabela 2: Avaliação de desempenho dos modelos perante as bases de dados consideradas.

	Base de Dados	Precisão	Revocação	F ₁ -Score	AP ₀	AP ₁	mAP	FPS
YoloV3 Small	Mask	88,6	72,9	80,0	72,1	89,2	80,7	37,3
	Face Mask	90,1	81,9	85,8	78,4	94,5	86,4	37,0
	Mask Detection	96,3	96,7	96,5	99,2	98,8	99,0	83,3
	Base Unificada	91,9	84,9	88,3	84,3	94,2	89,3	107,5
YoloV3	Mask	95,1	79,6	86,7	84,1	93,2	88,7	32,8
	Face Mask	93,6	87,5	90,4	87,6	97,2	92,4	38,3
	Mask Detection	99,8	99,3	99,5	99,5	99,5	99,5	35,5
	Base Unificada	94,2	90,3	92,2	89,5	95,8	92,7	40,2

Resultados e Discussão

Tabela 3: Avaliação de desempenho dos modelos perante as bases de dados consideradas.

	Base de Dados	Precisão	Revocação	F ₁ -Score	AP ₀	AP ₁	mAP	FPS
YoloV5 Nano 6	Mask	93,6	75,2	83,4	77,1	90,5	83,8	38,0
	Face Mask	87,5	83,3	85,3	78,1	95,2	86,7	37,3
	Mask Detection	97,2	97,2	97,2	99,4	98,4	98,9	87,0
	Base Unificada	91,6	87,5	89,5	88,0	94,7	91,3	123,5
YoloV5 Small 6	Mask	93,3	77,5	84,7	81,2	93,6	87,4	37,3
	Face Mask	90,4	89,2	89,8	85,3	97,5	91,4	35,8
	Mask Detection	99,2	98,5	98,8	99,5	99,1	99,3	77,5
	Base Unificada	94,0	90,1	92,0	89,8	95,9	92,8	101,0
YoloV5 Medium 6	Mask	91,2	79,9	85,2	80,5	92,5	86,5	35,6
	Face Mask	90,2	85,2	87,6	83,8	93,3	88,6	32,7
	Mask Detection	99,3	98,5	98,9	99,5	99,4	99,4	56,0
	Base Unificada	92,2	90,5	91,3	89,0	95,5	92,3	78,7

Resultados e Discussão

- Bom desempenho da YOLO na detecção
 - mAP mínimo: 80,7 %
 - mAP máximo: 99,5 %
 - Características de Haar em [Viola & Jones \(2001\)](#)
- Mask Detection: mAP médio de $99,2\% \pm 0.23$
 - Pouca representatividade para a complexidade real do problema
- Melhor FPS observado na Base Unificada

Resultados e Discussão

- Bom desempenho da YOLO na detecção
 - mAP mínimo: 80,7 %
 - mAP máximo: 99,5 %
 - Características de Haar em [Viola & Jones \(2001\)](#)
- Mask Detection: mAP médio de $99,2\% \pm 0.23$
 - Pouca representatividade para a complexidade real do problema
- Melhor FPS observado na Base Unificada

Solução de Referência

YOLOv5 Small 6: mAP de 92,8 % na Base Unificada

Resultados e Discussão

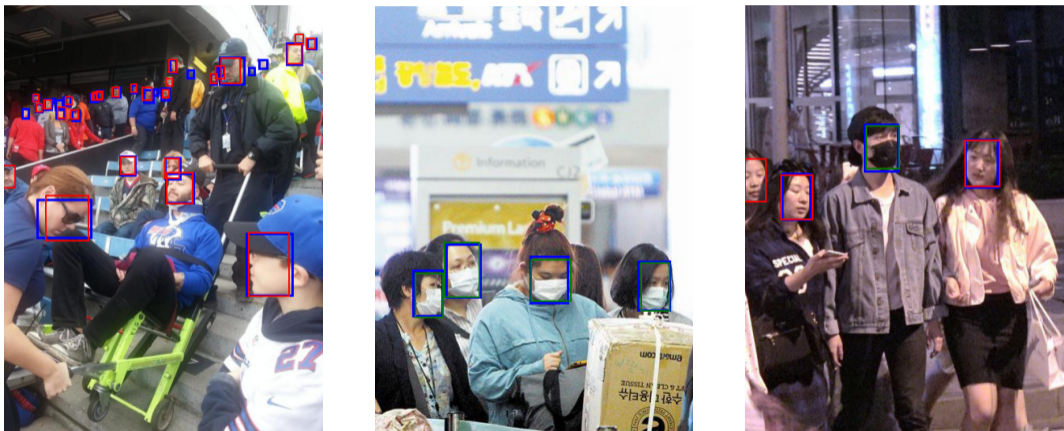


Figura 3: Exemplos de detecções com a YOLOv5 *Small 6* no conjunto de testes.

Resultados e Discussão

Tabela 4: Comparativo de desempenho perante a partição de testes da AIZOO.

Modelo	AP₀	AP₁	mAP
RetinaFace (Deng <i>et al.</i> 2020)	92,8	93,1	93,0
RetinaFaceMask-M (Fan & Jiang 2021)	93,6	90,4	92,0
SL-FMDet (Fan <i>et al.</i> 2021)	93,6	94,0	93,8
Solução Proposta	89,4	97,0	93,2

Considerações Finais

- Avaliação comparativa de 5 arquiteturas da R-CNN YOLO para detecção de máscaras faciais
- Solução de referência: YOLOv5 *Small* 6 treinada e testada com a Base Unificada
 - F_1 -Score igual a 92 %, mAP igual a 92,8 % e capacidade de detecção em tempo real a 101 FPS
 - É menos de 1 % inferior ao estado da arte na detecção de máscaras faciais, mas é mais eficiente para detectar faces com máscara
 - Solução competitiva
- **Trabalhos Futuros:** ajustes finos na solução de referência e validar a solução no contexto da detecção em tempo real em vídeos de câmeras de monitoramento

Agradecimentos

Os autores agradecem o apoio financeiro provido pela Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) por meio do Programa PAIC nas Edições 2020/2021 e 2021/2022.



SECRETARIA DE ESTADO DE
PLANEJAMENTO,
DESENVOLVIMENTO, CIÊNCIA,
TECNOLOGIA E INOVAÇÃO



GOVERNO DO ESTADO DO
AMAZONAS

Grupo de Pesquisas em Sistemas Inteligentes
Laboratório de Sistemas Inteligentes
Escola Superior de Tecnologia
Universidade do Estado do Amazonas



Um Estudo de Caso da Detecção do Uso de Máscaras Faciais com Redes Neurais Convolucionais Regionais

Diego Lucena de Medeiros, Elloá B. Guedes, Carlos Maurício S. Figueiredo

{dldm.eng18, ebgcosta, cfigueiredo}@uea.edu.br

III Workshop Brasileiro de Cidades Inteligentes

CSBC 2022 – Niterói – Rio de Janeiro