

Defect detection in textile manufacturing: Improving findings using YOLOv5 versions

Rodolfo Seidel *, Cláudio Alberto Waiandt †, Hilário Seibel Júnior‡

*†‡ *Programa de Pós-graduação em Computação Aplicada (PPComp)*

Instituto Federal do Espírito Santo (IFES) Campus Serra

rodsei@gmail.com *, waiandt26@yahoo.com.br †, hsjunior@ifes.edu.br ‡

Abstract—Currently one of the main challenges in identifying textile defects is related to the detection manual, mainly due to the complexity of irregular shapes and small objects. Finding ways to automate this detection has been studied by several authors. Thus this article expands and improves the findings and conclusions in the work, entitled 'Textile defect Detection using YOLOv5 on AITEX dataset'. This study deepened the investigation with native data augmentation, transfer learning and base rebalancing techniques, in addition to new metrics that allowed us to present results that corroborate and expand previous findings. Previous research laid the foundation for this study and provided crucial insights that guided further exploration of the issues. With the results obtained, it was concluded that the versions of YOLOv5 with the new techniques used adapted very well to the new context, enhancing their performance.

Index Terms—YOLOv5. AITEX. Detection of textile defects. Bounding boxes. IoU. Annotation of objects. Transfer learning.

I. INTRODUÇÃO

A indústria têxtil é de grande relevância no Brasil, já que representa a quinta maior indústria têxtil do mundo e a quarta no segmento de vestuário [1]. Conforme dados da Associação Brasileira da Indústria Têxtil e de Confecção [2], o Brasil é o segundo maior empregador da indústria de transformação, perdendo apenas para alimentos e bebidas, o que demonstra a relevância desse setor para a economia nacional, por demandar o uso intensivo de mão de obra, gerando emprego e renda.

O estudo realizado em [3], entretanto, destaca que a revisão de tecidos em uma indústria têxtil ainda é um processo manual, que demanda bastante tempo. E esse processo se torna ainda mais fadigante já que a capacidade humana de identificação e análise variam ao longo da jornada de trabalho [4].

Por tudo isso, se faz essencial investir no planejamento, automação, tecnologia e inovação para melhorar a produtividade deste setor. Neste contexto, soluções da indústria 4.0 proporcionam maior precisão e prometem reduzir o tempo de produção. Uma tendência que ocupa cada vez mais o espaço é a ESG (do inglês, *Environmental, Social and Governance*), que corresponde às práticas ambientais, responsabilidades sociais e de governança de uma organização. Sua ideia é a redução de desperdício em todos os processos, cuidados com a origem

da matéria-prima, valorização da mão de obra, entre outras ações sustentáveis que marcam uma era de um consumidor mais exigente e mais consciente [5].

Este artigo representa uma continuação direta da pesquisa intitulada 'Textile defect detection using YOLOv5 on AITEX Dataset' [3], na qual foi identificado o comportamento do YOLOv5 na base de defeitos têxteis AITEX¹, aplicando técnicas de anotação de objetos e *data augmentation*. A base de dados AITEX foi criada coletando imagens em uma fábrica através de um sistema real e contém 245 amostras de 7 tipos distintos de defeitos. A partir da criação da base de dados foi possível a comparação adequada entre os métodos de detecção de defeitos [7].

No estudo em que esse trabalho se baseia [3], a base chegou a 802 amostras de treino através da técnica de *data augmentation*, o suficiente para alcançar resultados significativos. A anotação de objetos com *bounding boxes* obteve os melhores resultados, mesmo sem otimizações na base original e a utilização de técnicas isoladas de *data augmentation*. O modelo produzido obteve resultados individuais superiores ao *baseline* construído.

O presente trabalho utiliza a mesma versão do YOLO em relação ao estudo original [3], além de aplicar técnicas de *transfer learning* e *data augmentation*, nativas do YOLO, novas métricas e rebalanceamento de base. Ao utilizar o mesmo modelo pré-treinado foi possível comparar os resultados obtidos nos estudos.

Este artigo contribui para as práticas ambientais e responsabilidades sociais que regem o ESG, aprofundando na investigação através de novas técnicas para aumentar a detecção de defeitos na indústria têxtil reduzindo o desperdício desde o início até o final da produção. Quanto melhor o ESG da empresa, mais chances de atrair novas oportunidades e expandir a lucratividade.

O artigo está organizado conforme detalhes a seguir. A seção 2 apresenta alguns estudos recentes que abordam os tópicos relacionados a este trabalho. Na seção 3 serão apresentados os materiais e métodos. Na seção 4 estão os experimentos e resultados e por fim a seção 5 traz as considerações finais e trabalhos futuros.

Os autores agradecem ao Ifes, apoio da FAPES e CAPES (proc 2021-2S6CD, nº FAPES 132/2021) do PDPG (Programa de Desenvolvimento da Pós-Graduação, Parcerias Estratégica nos Estados).

¹6 AITEX Dataset: <https://www.aitex.es/afid/>

II. TRABALHOS RELACIONADOS

O referencial teórico deste estudo inclui trabalhos recentes na área, com destaque para o estudo anterior, intitulado ‘*Textile defect detection using YOLOv5 on AITEX Dataset*’ [3], que forneceu a base para as investigações tratadas aqui. Nele, o YOLOv5 adaptou-se bem a outro contexto com objetos distintos do pré-treinamento, as anotações com *bounding boxes* permitiram maior aprendizado e reconhecimento dos defeitos, mesmo com diferentes formas e tamanhos, além da combinação de *data augmentation* que potencializou seu desempenho.

O artigo citado aplica o modelo YOLOv5 na base de dados AITEX, usando a abordagem de detecção de objetos para localizar e identificar defeitos, avaliando diferentes técnicas de anotação de objetos e *data augmentation*. A base AITEX, composta por 245 imagens de 7 tipos distintos de tecidos não identificados em um tamanho de 4096×256 pixels, sendo 140 imagens sem defeitos e 105 com defeitos, estava distribuída em 12 classes.

Em uma primeira etapa, foi criado um *baseline* utilizando as imagens originais para compreender seu comportamento. Somente as imagens com defeitos foram redistribuídas entre treino (65%), validação (20%) e teste (15%). As anotações foram padronizadas com *bounding boxes* por apresentarem melhores resultados. Já com os critérios definidos, foram aplicadas as técnicas de *data augmentation* gerando novas bases de imagem e novos experimentos realizados.

A métrica para avaliar os resultados no aspecto classificação foram a *precision* e o *recall*. Para a detecção de objetos utilizou-se a métrica IoU (*intersection over union*) determinando o acerto ou erro através de um *threshold* que permitiu calcular a precisão e *recall*. Durante o treinamento, a IoU e suas variações foram utilizadas como *loss function*. Como métrica final foi utilizado o mAP (*mean average precision*).

Na Tabela I podemos observar os resultados alcançados no estudo, considerando as novas bases geradas a partir da técnica de *data augmentation* e utilizando as anotações com *bounding boxes*.

TABLE I
RESULTADOS DOS EXPERIMENTOS.

Base de Imagens	Resoluções	mAP 0.5	mAP 0.5:0.95	Precision	Recall	F1-Score
Mista Completa	1280 e 1024	84,4%	56,06%	89,66%	79,66%	84,36%
Mista	1280 e 1024	84,5%	55,00%	86,22%	78,54%	82,20%
Inversões	1280 e 1024	45,5%	21,40%	65,95%	40,15%	49,91%
Brilho	1280 e 1024	35,4%	18,45%	59,18%	33,43%	42,73%
Ruídos	1280 e 1024	20,8%	10,06%	41,26%	22,17%	28,84%
Original (Bounding Boxes)	1280 e 1024	15,9%	8,78%	34,30%	18,44%	23,99%
Simultânea	1280 e 1024	13,1%	6,54%	10,74%	20,09%	14,00%
Original (Polígonos)	1280 e 1024	7,44%	2,66%	1,68%	12,35%	2,96%
YOLOv5s (COCO)	640	56,8%	37,4%	-	-	-
YOLOv5s6 (COCO)	1280	63,7%	44,8%	-	-	-
YOLOv5x (COCO)	640	68,9%	50,7%	-	-	-
YOLOv5x6 (COCO)	1280	72,7%	55,0%	-	-	-

Na Fig. 1 podemos observar o impacto na evolução dos resultados e benefícios obtidos ao se utilizar as técnicas de *data augmentation* para tornar a base mais robusta com a presença de mais amostras.

Nos estudos de aumento de dados, de acordo com [8], o uso de *data augmentation* desempenha um papel importante na

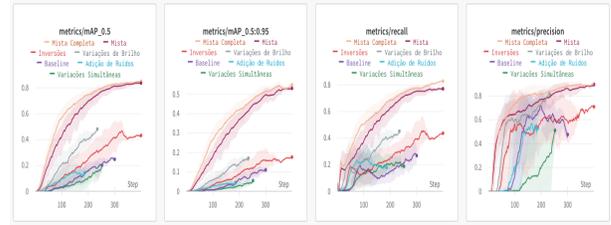


Fig. 1. Comparativo entre as técnicas isoladas de *data augmentation*.

melhoria da precisão dos modelos de aprendizagem profunda, mesmo quando os dados de treinamento são escassos. O estudo ainda destaca que o aumento de dados é uma estratégia fundamental para aprimorar a performance de modelos com limitações de dados de treinamento, transformando os dados em um formato específico. Ao aplicar o aumento de dados, tanto o volume quanto a diversidade do conjunto de dados aumentam significativamente. Essa abordagem ganhou destaque em diversas aplicações, sendo especialmente popular no campo do reconhecimento de objetos baseados em imagens, abrangendo tarefas como classificação e detecção em tempo real, entre outras.

Através da condução dos seus estudos, [9] utilizou a arquitetura YOLOv5 para treinar um modelo personalizado através da técnica de aprendizado por transferência. Esta abordagem usa modelos pré-existentes como base para tarefas específicas, economizando tempo e recursos em comparação com a criação de redes neurais do zero. Os resultados revelaram que o YOLOv5 com aprendizado por transferência é altamente eficaz para detecção e reconhecimento, alcançando uma precisão de 96% e uma pontuação mAP de 0,93, com uma pontuação F1 estável e robusta em níveis de confiança de até 90%. Isso ressalta a utilidade e eficácia dessa técnica na área de visão computacional. Na Fig. 2, podemos visualizar a arquitetura da técnica de aprendizado por transferência.

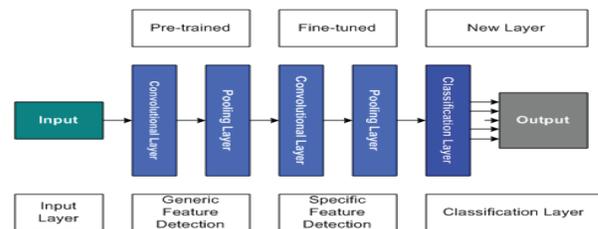


Fig. 2. *Transfer learning* arquitetura.

Já na Fig. 3, podemos observar como o treinamento de um modelo é impactado pela presença ou ausência dessa técnica. A análise do gráfico revela uma melhoria substancial no treinamento quando a aprendizagem por transferência é aplicada desde o início.

O YOLOv5 [10] (acrônimo para *you only look once*) foi apresentado inicialmente por [11] como uma família de arquiteturas e modelos de detecção de objetos, utilizando o framework *PyTorch*. Os modelos disponibilizados foram pré-treinados com a base de dados COCO, criada por [12] e que

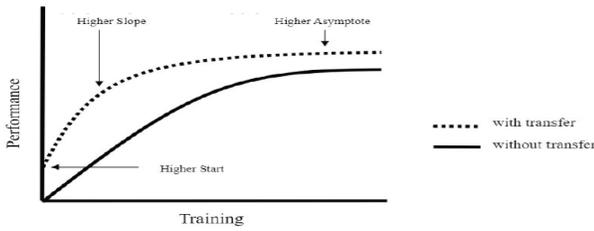


Fig. 3. Treinamento com e sem aprendizagem por transferência.

reúne imagens com cenas complexas do cotidiano contendo objetos em seu contexto natural. O YOLOv5 fornece alguns modelos pré-treinados. A proposta padrão e recomendada para bases de dados pequenas e médias é realizar o *transfer learning* com os pesos dos modelos pré-treinados.

Avaliando o tempo total de execução dos modelos e os resultados obtidos, pode-se perceber que enquanto o tempo de execução entre o YOLOv5s e o YOLOv5m é maior, o resultado máximo obtido não chega ao dobro do inicial no modelo YOLOv5x, que tem os melhores resultados, porém apresenta um tempo de execução muito acima dos demais como podemos visualizar na Fig. 4.

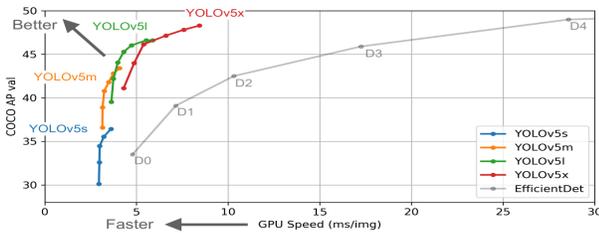


Fig. 4. Eficiência versus tempo de execução.

III. MATERIAIS E MÉTODOS

A. Rebalanceamento da base de dados

A base de dados AITEX, foi utilizada para realizar todos os testes desta pesquisa, com o objetivo de melhorar os resultados encontrados na pesquisa original [3]. Foram realizados alguns ajustes na base original de imagens da AITEX. Considerando que as imagens sem defeitos comprometeram o balanceamento da base, elas foram removidas e as demais imagens restantes foram redistribuídas. Todas as classes de defeitos continuaram com a distribuição original de amostras, passando apenas por uma calibragem para manter uma razão aproximada de 65%, 20% e 15% nas etapas de treino, validação e teste, respectivamente.

A nova organização da base apenas com as imagens defeituosas está apresentada na Tabela II. Ao fim dos ajustes, restaram 9 classes em um total de 103 imagens.

Com a aplicação das técnicas de *data augmentation* diretamente na base, ela se tornou mais robusta com a presença de mais amostras de treino, principalmente para as classes de defeitos com menor representação. Para isolar possíveis efeitos negativos e compreender os respectivos impactos, foram

TABLE II
CLASSE DE DEFEITOS APÓS REDISTRIBUIÇÕES GERAIS.

#	Classe	Total Amostras	Distribuição	Treino	Validação	Teste
0	002 - Broken End	9	8,74%	6	1	2
1	006 - Broken Yarn	8	7,77%	5	1	2
2	010 - Broken Pick	10	9,71%	5	3	2
3	016 - Weft Curling	3	2,91%	2	0	1
4	019 - Fuzzyball	39	37,86%	27	7	5
5	022 - Cut Selvage	9	8,74%	6	2	1
6	023 - Crease	5	4,85%	3	1	1
7	025 - Warp Ball	6	5,83%	4	1	1
8	030 - Nep	14	13,59%	8	3	3
	Totais	103	100%	66	19	18

criadas outras duas bases com um volume maior de amostras, dessa vez agrupando as imagens de todas as variantes de cada técnica. Para maior compreensão dos possíveis impactos, uma delas contou apenas com as variantes isoladas, ou seja, uma junção das três primeiras bases e a outra incluiu na junção a quarta base com as variações simultâneas. Elas foram nomeadas bases "Mista" e "Mista Completa", respectivamente.

O formato final das bases pode ser observado na Tabela III. As classes mantiveram a proporção de amostras da AITEX original e contaram com um aumento significativo de amostras.

TABLE III
CLASSES DE DEFEITOS E SUAS RESPECTIVAS DISTRIBUIÇÕES DE AMOSTRAS APÓS APLICAR DATA AUGMENTATION NAS IMAGENS DE TREINO.

#	Classe	Baseline	Variação de Brilho	Inversão horizontal e vertical	Adição de Ruídos	Variações Simultâneas	Mista	Mista Completa
0	002 - Broken End	6	18	12	18	18	54	72
1	006 - Broken Yarn	5	15	9	15	15	44	59
2	010 - Broken Pick	5	15	11	15	15	46	61
3	016 - Weft Curling	2	5	4	6	6	17	23
4	019 - Fuzzyball	27	80	61	81	81	249	330
5	022 - Cut Selvage	6	17	15	18	18	56	74
6	023 - Crease	3	9	6	9	9	27	36
7	025 - Warp Ball	4	12	10	12	12	38	50
8	030 - Nep	8	24	17	24	24	73	97
	Totais	66	195	145	198	198	604	802
	% Aumento Base	-	295%	220%	300%	300%	915%	1215%

A partir dessa etapa os experimentos foram realizados apenas com as opções de treinamento mais promissoras, sendo elas 300, 350 e 400 épocas nas resoluções de imagens de 1024px e 1280px.

B. Aplicações e impacto do transfer learning

Para avaliar como esses formatos de treinamento influenciam no desempenho do modelo, foram utilizadas três configurações possíveis de acordo com a combinação dos parâmetros de treinamento *weights*, *cfg* e *freeze*. Os modelos escolhidos foram o yolov5s e yolov5x e suas variações foram definidas como: yolov5s_pt sendo o treinamento padrão do yolov5s com os pesos pré-treinados da base COCO. A segunda, nomeada yolov5s_scratch foi o treinamento do yolov5s com as configurações padrão da rede neural, porém sem os pesos do pré-treinamento, o que exige que o modelo realize todo o aprendizado desde o início. E por último, a terceira variação nomeada yolov5s_freeze10 foi o treinamento com os pesos pré-treinados, porém com o *backbone* congelado, ou seja, o *backbone* não é retreinado, reutilizando uma parte maior do aprendizado obtido anteriormente.

C. Native data augmentation

Também estão presentes nas versões do YOLOv5 os parâmetros nativos relacionados a utilização de técnicas de *data augmentation*. Essas técnicas Tabela IV, são aplicadas durante o carregamento de imagens pelo modelo para treinamento.

TABLE IV
PARÂMETROS DE DATA AUGMENTATION NATIVOS NO MODELO YOLO.

Parâmetro	Valor	Descrição
<i>hsv_h</i>	0.015	ajuste parcial de Matiz (<i>Hue</i>)
<i>hsv_s</i>	0.7	ajuste parcial de saturação (<i>Saturation</i>)
<i>hsv_v</i>	0.4	ajuste parcial de Valor (brilho) (<i>Value</i>)
<i>translate</i>	0.1	porção de <i>pixels</i> de deslocamento (+/-)
<i>scale</i>	0.5	porção de escala (+/-)
<i>flipr</i>	0.5	probabilidade de inversão (Horizontal)
<i>mosaic</i>	1.0	probabilidade de criação de mosaicos (4 imagens)
<i>degrees</i>	0.0	grau de rotação (+/-)
<i>shear</i>	0.0	grau de cisalhamento (+/-)
<i>perspective</i>	0.0	grau de angulação de perspectiva (+/-)
<i>flipud</i>	0.0	probabilidade de inversão (Vertical)
<i>mixup</i>	0.0	probabilidade de mesclar duas imagens sobrepostas
<i>copy_paste</i>	0.0	probabilidade de inclusão de partes de imagens em outras

D. Métricas de Avaliação

As métricas de avaliação usadas neste artigo levaram em consideração as métricas aplicadas em [13]. Para avaliar os resultados no aspecto classificação, foram utilizados o *precision* e o *recall*, calculados a partir das detecções corretas de objetos existentes na imagem. Além das métricas citadas, também foi utilizado a F1-Score, calculado a partir do *precision* e *recall* conforme equação 1.

$$F1 = \frac{2 \times Precision \times Recall}{(Precision + Recall)} \quad (1)$$

A principal métrica utilizada foi o mAP (*mean average precision*), que é representada pelas médias das áreas abaixo da curva *precision-recall* (AP) de todas as classes e o mAP@0,5:0,95, combinando os resultados de vários *thresholds* da IoU, de 0,5 até 0,95, com intervalos de 0,05. Para o desafio COCO 2017 por exemplo, que foi utilizado como base para a avaliação YOLOv5, foram consideradas como métricas o AP com limiar de 0,5 e o mAP de 10 limiares de IoU [14].

O limiar do IoU determina a acurácia da localização do objeto, sendo calculado entre a predição do *bounding box* e o alvo. Normalmente definido em 0,5, representa o mAP@0,5, e caso o valor esteja acima, é considerada uma localização correta, caso contrário, um erro. A nova medida realiza um cálculo da média obtida considerando vários limiares simultaneamente, entre 0,5 e 0,95, caracterizando o mAP@0,5:0,95 [15]. Uma nova abordagem chamada SIOU, proposta por [16], redefine as penalidades da função considerando a direção do erro a partir do ângulo vetorial entre a predição e o alvo. Essa solução obteve melhores tempos de treinamento e acurácia nos resultados do estudo. A fórmula para o cálculo da métrica pode ser observada na equação 2.

$$mAP@0,5:0,95 = \frac{1}{10} \times \sum_{IoU=0,5; i=0,05}^{0,95} \left(\sum_{c=1}^9 Precision(Recall_c)_{@IoU} \right) \quad (2)$$

IV. EXPERIMENTOS E RESULTADOS

Os experimentos foram realizados diretamente nas novas bases disponibilizadas no estudo [3], contendo bases de imagens produzidas com as técnicas de *data augmentation* no tamanho 1024px e 1280px e treinamento entre 350 e 400 épocas. Com o aprofundamento dos estudos no modelo, foram realizados novos experimentos para explorar as características e possibilidades.

É importante relembrar que o modelo YOLO entrega uma configuração pré-definida para aplicação de *data augmentation*. Parâmetros como: *mosaic*, *mixup* e *copy_paste*. Conforme os gráficos apresentados na Fig. 5, observa-se que o novo conjunto de amostras utilizando os parâmetros nativos apenas potencializou o aprendizado do modelo yolov5s.



Fig. 5. Comparativo das combinações de treinamento do yolov5s com as técnicas nativas de *data augmentation* ligadas e desligadas.

Observa-se também, conforme resultados apresentados na Fig. 6, que entre a utilização dos modelos pré-treinados, o yolov5s_pt obteve os melhores resultados dentre as três configurações. Apesar de ser um contexto diferente, os pesos do pré-treinamento contribuem para o modelo aprender como localizar objetos.

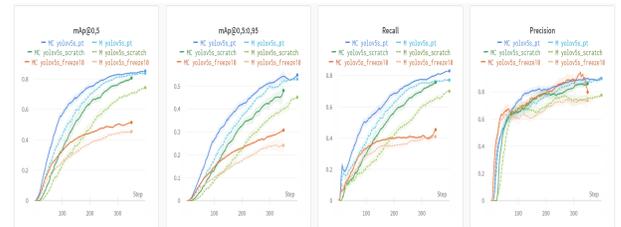


Fig. 6. Comparativo das combinações de treinamento do yolov5s nas bases Mista Completa (MC) e Mista (M).

A. YOLOv5X

O yolov5s é um dos dois menores modelos da família, então para mensurar o potencial máximo que o YOLOv5 poderia alcançar com a base AITEX nas condições apresentadas foi realizados os experimentos no yolov5x com as bases Mista (M), Mista Completa (MC) e o Baseline (B).

A Fig. 7 e a Tabela V apresentam os resultados obtidos comparando os dois modelos pré-treinados. De fato, o yolov5x superou o yolov5s conforme esperado. A base “Mista Completa” alcançou um $mAP@0,5:0,95$ de 60,12%, *precision* de 91,81% e *recall* de 83,82%. O yolov5s apresenta um custo-benefício com relação ao yolov5s, porém sendo neste mesmo cenário, o yolov5x superou as métricas do yolov5s em 2,96% a mais do $mAP@0,5:0,95$, 2,39% a mais de *precision* e 5,22% a mais de *recall*.

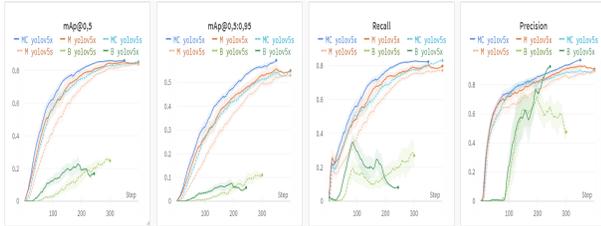


Fig. 7. Comparativo de resultados entre os modelos yolov5x e yolov5s.

TABLE V
COMPARATIVO DE RESULTADOS ENTRE OS MODELOS YOLOV5X E YOLOV5S.

Modelo	Base de Imagens	Resoluções	mAP 0.5	mAP 0.5:0.95	Precision	Recall	F1-Score
yolov5x	Mista Completa (MC)	1280 e 1024	86,9%	60,12%	91,81%	83,82%	87,63%
	Mista (M)	1280 e 1024	86,5%	57,53%	92,02%	81,23%	86,29%
	Baseline (B)	1280 e 1024	22,7%	11,63%	57,20%	20,60%	30,29%
yolov5s	Mista Completa (MC)	1280 e 1024	84,4%	56,06%	89,66%	79,66%	84,36%
	Mista (M)	1280 e 1024	84,5%	55,00%	86,22%	78,54%	82,20%
	Baseline (B)	1280 e 1024	15,9%	8,78%	34,30%	18,44%	23,99%

Durante a realização dos experimentos foram selecionados 3 tipos de GPUs, listadas a seguir com suas respectivas capacidades de alocação de memória: Tesla T4 (15Gb), Tesla P100-PCIE (16Gb) e NVIDIA A100-SXM4 (40Gb). A Fig. 8 apresenta os tempos de treinamento de cada modelo com os equipamentos utilizados.

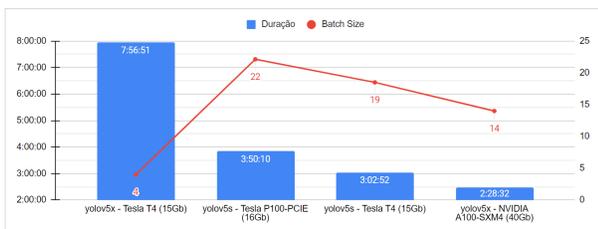


Fig. 8. Comparativo dos tempos de execução com diferentes GPUs (yolov5x e yolov5s).

Outro ponto importante que pode ser observado é o total de treinamento paralelo suportado, onde o yolov5x, por realizar um volume muito maior de operações (GFlops), suporta um *batch_size* muito inferior aos demais ambientes. Mesmo utilizando uma GPU mais robusta, com 40Gb de memória disponível ainda precisa ser executado com *batch_sizes* menores. Com esses comparativos, entende-se que

apesar dos resultados obtidos, o modelo yolov5x se torna viável de ser utilizado apenas em cenários que o tempo de treinamento não seja uma restrição e que se tenha um equipamento adequado à disposição.

Ao final dos experimentos, o yolov5x superou os resultados alcançados pelo yolov5s assim como esperado com base nos resultados dos modelos pré-treinados na base COCO. A base “Mista Completa” alcançou um $mAP@0,5:0,95$ de 60,12%, *precision* de 91,81% e *recall* de 83,82%. Esses resultados demonstraram a disparidade dos modelos entre o consumo de recursos e tempo de treinamento em relação aos resultados obtidos. O yolov5x alcançou na principal métrica, $mAP@0,5:0,95$, um resultado aproximadamente 7% superior ao yolov5s consumindo cerca de 8h de treinamento contra 3h do yolov5s utilizando a mesma placa gráfica (Tesla T4).

Na Tabela VI podemos observar um comparativo entre os resultados desta pesquisa e dos trabalhos correlatos apresentando as métricas disponíveis. Não tivemos acesso aos códigos e *datasets* para unificar as medidas de comparação dos trabalhos correlatos. Sendo assim demonstra-se as métricas informadas em cada estudo destacando os melhores resultados alcançados. O yolov5x alcançou o maior F1-Score de 87,63%, superando [17] (82,58%). Alcançou também a maior *precision*, com 91,81%, superando [18] (85,33%) e o maior $mAP@0,5$ em 86,9%, superando [19] (71,7%). Apesar dos demais resultados e da otimização do *recall*, com um resultado de 83,82%, foi possível superar apenas [17] (82,58%), permanecendo abaixo dos demais trabalhos que utilizaram outros métodos não relacionados ao YOLO na base de dados AITEX. No topo, [20] alcançou um *recall* de 92,70%. Mesmo não superando o melhor resultado, o resultado obtido ainda está muito acima dos demais trabalhos que utilizaram o YOLOv5, como [18](48,95%).

TABLE VI
COMPARATIVO COM OS RESULTADOS OBTIDOS PELOS TRABALHOS CORRELATOS.

Autores	Base de Dados	Modelo	Resoluções	Precision	Recall	F1-Score	Accuracy	mAP 0.5	mAP 0.5:0.95
Silvestre-Bianes et al. (2019b)	AITEX	Gabor Filter	256	-	86.67%	-	87.76%	-	-
Rong-jiang et al. (2021)	AITEX	U-Net Aprimorada	-	-	92.70%	-	98.30%	-	-
Zhou et al. (2021)	AITEX	VAE + GMM	256	73.30%	90.70%	81.00%	-	-	-
Seçkin and Seçkin (2022)	AITEX	IFV	7	82.68%	82.59%	82.58%	82.59%	-	-
Weng et al. (2021)	Foshan Dataset	YOLOv5s Aprimorado	608	-	-	-	-	71.70%	-
Zheng et al. (2021)	Tianchi Dataset	SE-YOLOv5	-	85.33%	48.95%	62.21%	95.52%	-	-
Zheng et al. (2021)	Tianchi Dataset	YOLOv5	-	79.72%	48.62%	60.40%	94.22%	-	-
Jim and Niu (2021)	Xuelang Tianchi	Teacher Network	256	-	-	-	-	44.70%	-
Jim and Niu (2021)	TILDA	Teacher Network	256	-	-	-	-	45.10%	-
Este Trabalho	AITEX Mista Completa	yolov5s	1280 e 1024	89,66%	79,66%	84,36%	-	84,40%	56,06%
Este Trabalho	AITEX Mista Completa	yolov5x	1280 e 1024	91,81%	83,82%	87,63%	-	86,90%	60,12%

V. CONCLUSÃO E TRABALHOS FUTUROS

Esse estudo foi uma contribuição direta ao trabalho [3], onde foi aprofundado a compreensão dos conceitos e melhorias significativas foram obtidas em relação aos resultados da pesquisa inicial.

Ao construir o estudo sobre as bases estabelecidas anteriormente também foram avaliados os modelos mais robustos da família, como o YOLOv5x, o SIOU como nova função de regressão de localização, diferentes configurações de *transfer learning*, treinando desde o início sem repasse de aprendizado

e também congelando a camada inicial do *backbone* e desligando as configurações nativas de *data augmentation* dos modelos.

Com a base no formato final os modelos pré-treinados em seu estado da arte obtiveram os resultados mais relevantes na detecção de defeitos têxteis localizando e classificando os defeitos identificados de forma eficiente, atingindo um mAP@0,5:0,95 de 60,12%.

A partir dos resultados em seu estado da arte superiores às demais variações avaliadas, entende-se que o modelo se adaptou muito bem ao contexto de detecção de defeitos, mesmo com o desafio de detectar objetos pequenos.

Trabalhos futuros poderão utilizar os modelos deste estudo, treinados na base AITEX, como fonte de origem para aplicação de *transfer learning* em outras bases de detecção de defeitos em tecidos ou outros cenários.

Também é possível realizar uma avaliação dos limites de aprendizado do modelo, avaliando volumes maiores de épocas para identificar até que ponto o modelo continuaria aprendendo e se em determinado momento os treinamentos sem transferência de conhecimento ultrapassariam os modelos pré-treinados. Adicionalmente, a calibração de hiper parâmetros para o novo contexto poderia ser realizada para refinar o desempenho dos experimentos. Por fim, a versão final da base AITEX pode ser utilizada e validada nas novas versões do YOLO como 6, 7 e 8 e em outras alternativamente como por exemplo DETR e Deformable DETR.

REFERENCES

- [1] A. M. Cavalcanti and G. F. dos Santos, "A indústria têxtil no brasil: uma análise da importância da competitividade frente ao contexto mundial," *Revista Exacta*, vol. 20, 2022. [Online]. Available: <https://periodicos.uninove.br/exacta/article/view/17784>
- [2] ABIT, "Perfil do setor," 2022. [Online]. Available: <https://www.abit.org.br/cont/perfil-do-setor>
- [3] R. Seidel, H. S. Júnior, and K. Komati, "Textile defect detection using yolov5 on aitex dataset," in *Anais do XIX Encontro Nacional de Inteligência Artificial e Computacional*. Porto Alegre, RS, Brasil: SBC, 2022, pp. 763–774. [Online]. Available: <https://sol.sbc.org.br/index.php/eniac/article/view/22830>
- [4] A. M. G. Conceição, "Critérios de classificação de tecidos quanto à qualidade," 1998. [Online]. Available: <http://www.textilia.net/materias/ler/textil/maquina-textil>
- [5] Sebrae, "Moda, o que esperar da indústria têxtil em 2023," 2023. [Online]. Available: <https://sebrae.com.br/sites/PortalSebrae/artigos/moda-o-que-esperar-da-industria-textil-em-2023,3b84c93a0b8a6810VgnVCM1000001b00320aRCRD>
- [6] J. Silvestre-Blanes, T. Albero Albero, I. Miralles, R. Pérez-Llorens, and J. Moreno, "Aitex fabric image database," 2019. [Online]. Available: <https://www.aitex.es/afid/>
- [7] J. S. Blanes, T. Albero, I. Miralles, R. Pérez-Llorens, and J. Moreno, "A public fabric database for defect detection methods and results," *Autex Research Journal*, vol. 19, no. 4, pp. 363–374, 2019. [Online]. Available: <https://doi.org/10.2478/aut-2019-0035>
- [8] I. Zulkarnain, R. R. Nurmalasari, and F. N. Azizah, "Table information extraction using data augmentation on deep learning and image processing," in *2022 16th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, 2022, pp. 1–6.
- [9] U. Kulkarni, M. S. M. A. Shanbhag, A. A. Ayachit, P. Kulkarni, S. Halemani, and S. S. Kadappanavar, "Signboard detection using yolov5 with transfer learning," in *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*, 2023, pp. 1–7.
- [10] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, NanoCode012, Y. Kwon, TaoXie, J. Fang, Imyhxy, K. Michael, Lorna, A. V. D. Montes, J. Nadar, Laughing, Tkianai, YxNONG, P. Skalski, Z. Wang, A. Hogan, C. Fati, L. Mammana, AlexWang1900, D. Patel, D. Yiwei, F. You, J. Hajek, L. Diaconu, and M. Thanh Minh, "ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference," Zenodo, Feb. 2022.
- [11] G. Jocher, L. Changyu, A. Hogan, L. Y. , changyu98, P. Rai, and T. Sullivan, "ultralytics/yolov5: Initial Release," Zenodo, Jun. 2020.
- [12] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.
- [13] T. Liu and S. Chen, "Yolov4-dcn-based fabric defect detection algorithm," in *2022 37th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2022, pp. 710–715.
- [14] S. Yohanandan, "map (mean average precision) might confuse you!" 2020. [Online]. Available: <https://towardsdatascience.com/map-mean-average-precision-might-confuse-you-5956f1bfa9e2>
- [15] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," *Proceedings of the IEEE*, 2023.
- [16] Z. Gevorgyan, "Siou loss: More powerful learning for bounding box regression," *arXiv preprint arXiv:2205.12740*, 2022.
- [17] A. Ç. Seçkin and M. Seçkin, "Detection of fabric defects with intertwined frame vector feature extraction," *Alexandria Engineering Journal*, vol. 61, no. 4, pp. 2887–2898, 2022.
- [18] L. Zheng, X. Wang, Q. Wang, S. Wang, and X. Liu, "A fabric defect detection method based on improved yolov5," in *2021 7th International Conference on Computer and Communications (ICCC)*. IEEE, 2021, pp. 620–624.
- [19] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "You only learn one representation: Unified network for multiple tasks," *arXiv preprint arXiv:2105.04206*, 2021.
- [20] L. Rong-qiang, L. Ming-hui, S. Jia-chen, and L. Yi-bin, "Fabric defect detection method based on improved u-net," in *Journal of Physics: Conference Series*, vol. 1948, no. 1. IOP Publishing, 2021, p. 012160.