

棉花分析数据集

Akbar Ali^{*} Mahek Vyas[^] Soumyaratna Debnath^{*} Chanda Grover Kamra[†]
Jaidev Sanjay Khalane^{*} Reuben Shibu Devanesan^{*} Indra Deep Mastan[‡]
Subramanian Sankaranarayanan^{*} Pankaj Khanna^{*} Shanmuganathan Raman^{*}

^{*} Indian Institute of Technology Gandhinagar, India. [†] Ashoka University, Sonipat, India.
[^] Larsen & Toubro Technology Services Limited. [‡] Indian Institute of Technology (BHU), Varanasi, India.

ABSTRACT

本文介绍了 COT-AD，这是一套综合数据集，旨在通过计算机视觉增强棉花作物分析。COT-AD 包含了棉花生长周期内拍摄的超过 25000 张图像，其中 5000 张已标注图像，涵盖了用于田间规模检测和分割的航拍图像以及记录主要疾病的高分辨率数码单反相机图像。标注内容包括害虫和疾病识别、植被和杂草分析，填补了棉花特定农业数据集的关键空白。COT-AD 支持分类、分割、图像修复、增强、基于深度生成模型的棉花作物合成和早期疾病管理等任务，推动数据驱动的作物管理。COT-AD 数据集可以在这里找到：<https://aamaanakbar.github.io/COT-AD/>。

Index Terms— Cotton Crop Dataset, Crop Monitoring, Precision Farming

1. 介绍

棉花是一种全球重要的作物，对于纺织工业和许多国家如印度、中国和巴西的经济至关重要。它支持着包括农民和加工者在内的全球超过 2.5 亿人口。然而，棉花种植 [1] 面临着诸多挑战，如虫害侵袭、疾病和气候变化。

智能农业技术通过提高养分施用的精确度、监测植物胁迫和预测产量来帮助管理这些挑战。像土壤传感器、配备 RGB 和多光谱相机的无人机以及卫星图像这样的工具使得 [1] 能够进行有针对性的干预措施，如灌溉或农药施用，并提供更可靠的产量预测。精准农业方法可以分为侵入式（例如，手动土壤采样）和非侵入式（例如，通过无人机进行遥感），后者提供了较低的环境影响和成本。

由人工智能驱动的精细农业通过机器学习和数据分析来优化实践 [5]。这包括作物预测、害虫检测、精准灌溉，以及用于种植和收获的自主机械。然而，用于棉花图像分析的现有数据集（表 1）在范围和质量上都有限。目前的数据集主要集中在单一任务上，“数据量不足”和分辨率有限，阻碍了更广泛的应用，如疾病检测、棉花分割、物种分类、合成图像生成和文本引导增强。需要更为多样化的数据集来支持这些高级任务。

本文介绍了 COT-AD 数据集¹，用于棉花作物检

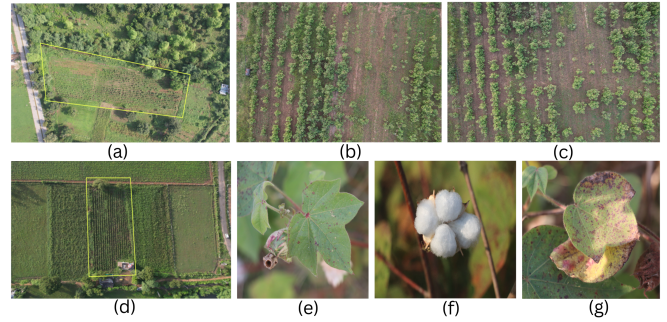


Fig. 1: 第 (a) 部分展示了站点 1 的现场视图；第 (b) 和 (c) 部分是无人机拍摄图像的例子；第 (d) 部分展示了站点 2 的现场视图；而第 (e)、(f) 和 (g) 部分是单反相机图像的例子。

测 [5]、分割、图像修复、增强和疾病分类。我们提出的数据集使机器能够设想作物的视觉和语义信息。该数据集包括在棉花作物生命周期中拍摄的棉田航拍图像（见表 3）以及用数码单反相机拍摄的各种影响棉花作物疾病的特写图像（见图 1）。主要贡献描述如下。

- 我们引入了一个新颖的大规模数据集 COT-AD，其中包括超过 25,000 张从无人机（航拍视图）和单反相机（近距离检查）在整个棉花生长周期内收集到的图像（表格 2）。
- 我们在 COT-AD 数据集中提供用于疾病分类、棉花作物检测和分割的注释（第 3 节）。
- 我们在 COT-AD 数据集上执行各种计算机视觉任务，例如作物检测、疾病分类、分割、修复、增强和合成（第 4 节）。

2. 相关工作

几个关键数据集支持疾病分类 [1, 4]，包括棉花作物植物叶片数据集 [1]、棉叶数据集 [3] 和棉花疾病数据集 [2]。前两个数据集专注于叶片图像，而第三个数据集通过包含叶片和整个植物图像提供了更广泛的视角。作物检

¹该数据集可用于非商业用途，并可在 IEEE 数据存储库 和

Kaggle 上获取。

Dataset	Images	Classes	Type	Tasks				
				Classification	Detection	Segmentation	Enhancement	Synthesis
Bishshash et al. [2]	2137	8	Hand Held	✓	✗	✗	✗	✗
Mirani et al. [3]	4000	7	Hand Held	✓	✗	✗	✗	✗
Li et al. [1]	4000	1	Aerial	✗	✗	✓	✗	✗
Karim et al. [4]	1710	4	Hand Held	✓	✗	✗	✗	✗
COT-AD(Ours)	25000+	9	Aerial + Hand Held	✓	✓	✓	✓	✓

Table 1: 各种任务的可用数据集规格，重点介绍 COT-AD 数据集，该数据集包含超过 25,000 张图像，并为分类、检测、分割、图像增强和棉花病害检测提供全面支持。

Split	Images	Classes	Type	Tasks				
				Classification	Detection	Segmentation	Enhancement	Synthesis
Drone	1800	1	Aerial	✗	✓	✓	✗	✗
DSLR Disease	3231	9	Hand Held	✓	✓	✗	✗	✗
DSLR Segment	100	12	Hand Held	✗	✗	✓	✗	✗
Drone + DSLR	14000+11000	-	Aerial + Hand Held	✗	✗	✗	✓	✓

Table 2: COT-AD 数据集划分的规格，包括无人机、DSLR 疾病、DSLR 分割和合成，总结了不同图像数量及对检测、分割和条件合成的支持，展示了数据集在不同成像类型和任务上的多样性。

测和分割 [5] 在精准农业和植物表型分析中起着至关重要的作用。检测涉及识别和定位植物，例如作物和杂草，以促进健康监测和自动化，而分割则通过使植物与背景区别开来进行详细分析 [5]。

3. COT-AD 数据集

COT-AD 数据集如表 2 所示，相较于先前的数据集（见表 1），通过提供超过 25,000 张使用无人机和单反相机拍摄的图像，作了显著扩展。它包括在 [10 米, 15 米, 和 115 米] 高度采集的高分辨率航拍影像以及单反相机手持拍摄的详细特写图像。此外，数据集中的现场视频有助于进行动态分析，包括疾病检测、作物健康判定、分类和分割、复原甚至合成。与先前的数据集相较，表 1，COT-AD 更加丰富和灵活，可以用于棉花科学和农业技术的不同应用。这种空中与近距照片的结合以及较大的体量使其对机器学习任务和 AI 技术更加有利，使农业 AI 的研究人员、实践者和开发者更容易获取。

数据采集、标注。COT-AD 数据集是使用 DJI Mavic Air 2 无人机在六个月内两个棉花田地（有机农场（每周数据）和 Kiran 农场（每两周数据））上采集的。我们在不同的高度采集数据以评估高度效果。在有机农场，每周两次使用 Canon EOS 80D 单反相机拍摄棉花作物的特写图像，由专家验证和分类。这些数据集包括代表不同棉花作物健康状况的图像，有助于生长监测和问题检测。COT-AD 数据集的图像被分为四个主要类别，即棉叶、棉铃、棉花和昆虫（表 4 和图 2），以监测棉花健康。

一旦图像采集完成，标注过程就开始了。每张图像都被手动标注，标记出图像中的棉花作物。手动标注确保了数据的高准确性，这对于训练有效的深度学习模型至关重要。

标注数据的文件夹结构组织在检测和分割任务目录

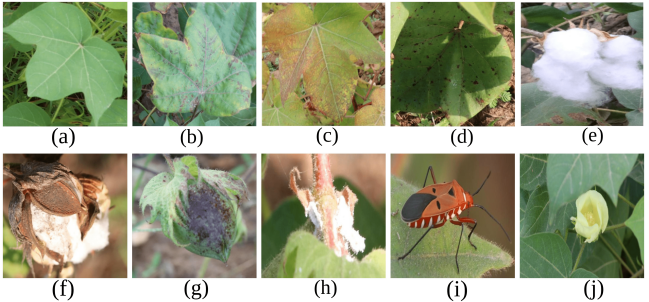


Fig. 2: 该图显示了棉花病害分类的示例图像：(a) 新鲜叶子，(b) 黄叶，(c) 叶片变红，(d) 叶斑/细菌性枯萎病，(e) 健康棉铃，(f) 棉铃腐烂，(g) 受损棉铃，(h) 粉虱，(i) 红棉虫，(j) 棉花花。

下，该目录分为四个主要部分：A 部分（前两个月的数据）、B 部分（第三个月的数据）、C 部分（第四个月的数据）和 D 部分（第五和第六个月的数据）。数据集结构的详细说明包含在补充材料中。

Age of Crop (Months)	Average Area Covered by Crops (per image)	Average Number of Crops (per image)	Total Instances of Crops
Month 1 and 2	56.95 %	49	14538
Month 3	58.01 %	58	37881
Month 4	63.24 %	73	27325
Month 5 and 6	73.01 %	44	21656

Table 3: 此表格展示了对不同年龄区间的 COT-AD 数据集航空图像的分析。

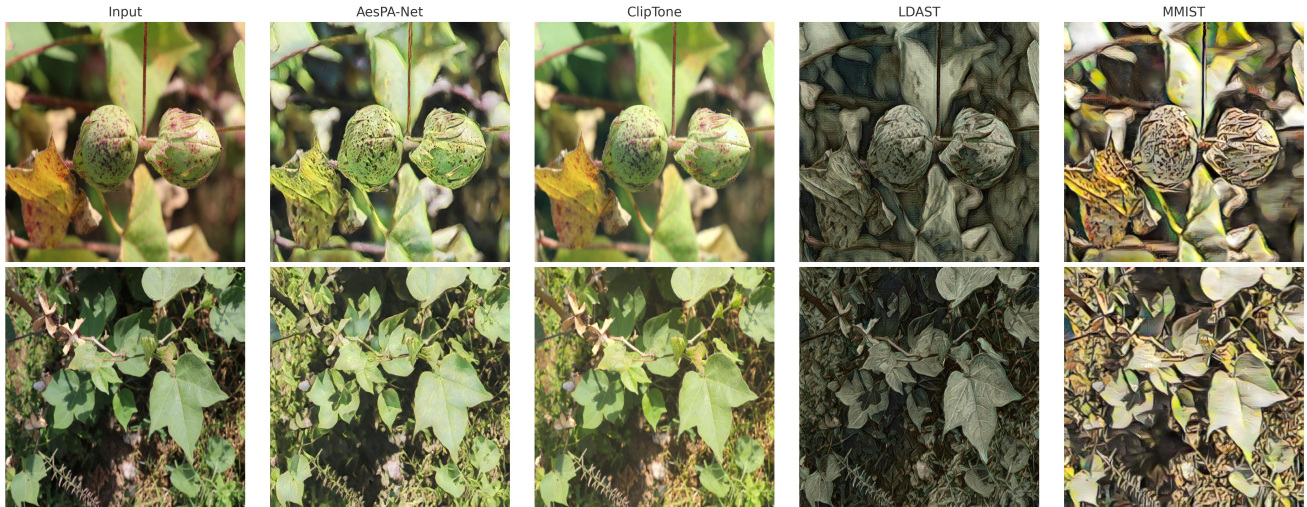


Fig. 3: 图像增强。每一行都有不同的文本提示用于增强。第 1 行：生成一张具有清晰细节、平滑纹理、均衡对比度、适当光照且无噪声或失真的高分辨率图像。第 2 行：输入图像的清晰版本，改进了可见度、对比度和细节，有效去除雾气。

Category	Types
Cotton Leaf	Yellow-ish Leaf, Bacterial Blight/Leaf Spot, Leaf Reddening, Fresh Leaf
Cotton Boll	Boll Rot, Damaged Cotton Boll, Healthy Cotton Boll
Cotton Flower	White Cotton Flower, Red Cotton Flower
Bugs	Mealy Bug, Red Cotton Bug

Table 4: 棉花植物部位的分类

4. 应用

本节介绍在 COT-AD 数据集上执行的应用。我们还在补充文件中提供了额外的结果。

4.1. 图像增强

我们使用文本引导的图像风格化方法增强了单反相机拍摄的棉花作物图像的质量。现有的方法如 mmist [6] 和 LDAST [7] 由于内容和风格图像之间的语义对应低，产生不和谐的伪影，如图 3 所示。为了解决这个问题，我们应用了美学图案感知风格迁移网络（AesPA-Net）[8]，它平衡了局部和全局风格表达，产生更好的结果。我们应用了 CLIPtone [9] 以传递图文提示的确切色调，如图 3 列 (e) 所示。正如表 ?? 所见，AesPA-Net 在视觉质量和定量分数如 Arniqa [10]、Brisque [11] 和 Clipiqa [12] 分数上优于其他方法。然而，由于其他方法采用基于文本提示的直接风格特征，Clipscores [13] 较高。

我们在单模态和多模态环境下进行语义分割。在单模态环境中，我们使用了一种强大的无监督分割方法，称为深度谱方法 [14]；在多模态环境下，我们使用了 CLIP-RC [17] 和 YOLOv11 [5]，如图 ?? 所示。在无



Fig. 4: 使用深度谱方法(DSM)进行的语义分割掩码 [14]。每种不同的颜色代表数据集的一个语义类别。

监督设置中，我们应用了深度谱方法 [14]，该方法从一个自监督预训练网络中提取特征，并利用特征相关性上的谱图理论获取特征向量。这些边缘段对应于具有明确边界的语义上有意义的区域。我们发现，与 DSM [14] 相比，如图 4 所示，CLIP-RC [17] 能够在图像中给出正确的语义类别。我们进行了零样本分类和线性探测实验，以评估 CLIP [18] 和 BioCLIP [19] 模型在棉花作物数据集上的泛化和表示能力。如表 ?? 所示，在零样本实验中，CLIP (RN50) 比 CLIP (ViT-16 或 32) 实现了更好的前 5 名准确率。这是因为我们仅使用 100 个图像点



Fig. 5: 棉花作物的疾病分类结果。图像显示了各种症状，包括红色棉桃、红色虫害和叶斑，代表了分类模型识别出的不同疾病指标。

来训练分类器，而 CLIP (ViT) 使用更多参数导致过拟合，多于 RN50。CLIP 实现了更高的前 5 名准确率，表明其具备更好的泛化能力，而 BioCLIP 在线性探测准确率上优于 CLIP (表 ?? 的第 3 列)，表明其对棉花作物数据集具有更好的表示能力。这在意料之中，因为 CLIP 是在一般图像-文本对上训练的，而 BioCLIP 是在生命树数据集上训练的。

疾病分类。该数据集分为三个主要类别：叶子疾病（新鲜叶子、黄色叶子、叶片发红、叶斑/细菌性枯萎）、虫害（粉蚧、红棉虫），以及棉花铃（损坏的棉花铃、棉花铃腐烂、健康的棉花铃）。每个类别按 70:15:15 的比例划分为训练集、验证集和测试集。VGG19 [20] 模型实现了 83.37 % 的测试准确率 (图 5)，其训练和验证损失分别为 0.3357 和 0.8805。

4.2. 图像合成

StyleGAN2-ADA(Dslr data)			StyleGAN2-ADA(Drone data)	
Metric	Value	Time (s)	Value	Time (s)
IS ($\mu \pm \sigma$)	3.214 $\pm$ 0.027	727.63	2.298 $\pm$ 0.009	726.25
FID	36.44	761.79	30.9	761.99
KID	0.0115	738.38	0.0077	734.95

Table 5: 该表比较了 StyleGAN2-ADA 在单反和无人机相机拍摄的棉花数据集上的性能指标。它报告了 Inception Score (IS)、Fréchet Inception Distance (FID) 和 Kernel Inception Distance (KID)，以及相关的计算时间 (以秒为单位)。

我们使用 StyleGAN2-ADA 进行了图像合成、数据增强和无人机及 DSLR 数据集的条件合成，如图 6 所示。我们还探索了 StyleGAN3，以利用其高级功能来改进图像生成。在 Table 5 中展示了无人机和 DSLR 数据在合成过程中使用的矩阵，显示了评估生成数据质量的关键参数和指标。

4.3. 图像修复

我们通过一种稳健的无监督恢复 [22] 框架恢复了我们提出的 COT-AD 数据集的图像，该数据集是通过无人机和 DSLR 拍摄的。为了改善在退化情景中的图像恢复，我们在无人机和 DSLR 图像数据集上使用基于 StyleGAN 的架构和双边图正则化模型 (BGRM) 执行去伪影、超



Fig. 6: 使用 StyleGAN2-ADA 创建的合成图像：前两个图像来自 DSLR 数据，而最后两个来源于无人机数据。

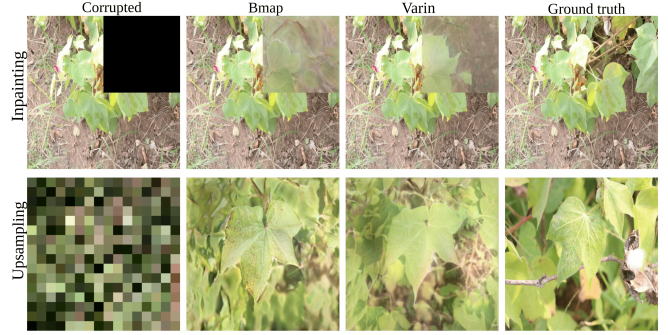


Fig. 7: 使用 BMAP 和变分修补 (VarIn) 对数码单反相机数据集进行修补与超分辨率的 BGRM [21] 图像复原。

Method	Drone			DSLR		
	LPIPS $\hat{a} \uparrow$	LPIPS- vgg $\hat{a} \uparrow$	PFID $\hat{a} \uparrow$	LPIPS $\hat{a} \uparrow$	LPIPS- vgg $\hat{a} \uparrow$	PFID $\hat{a} \uparrow$
Inpainting						
Bmap	0.214	0.198	59.67	0.214	0.212	70.78
Varin	0.203	0.194	51.57	0.210	0.208	62.07
Upsampling						
Bmap	0.678	0.643	135.13	0.656	0.669	135.32
Varin	0.674	0.649	143.03	0.655	0.661	128.46

Table 6: 通过生成模型的贝叶斯重建 (BRGM) 在无人机和单反棉花作物数据集上的图像修补和放大表现。

分辨率和图像修复任务 (见图 7)。BGRM 的矩阵表示在表 6 中给出。

5. 结论

在本文中，我们提出了一个 COT-AD 数据集，它对农业数据资源作出了重要贡献，特别针对棉花作物分析而定制。该数据集通过包含不同生长阶段的航拍和特写图像，能够广泛应用于分类、分割和恢复。实验评估表明，COT-AD 支持准确的病害识别、高效的作物检测和图像恢复，证明对于模型训练和现实农业用例都是有价值的。在未来的工作中，我们建议利用该数据集解决棉花球数数的问题。

6. 致谢

我们要感谢 Jibaben Patel AI 讲席、L & T Technology Services Limited 和 Visvesvaraya 博士计划对这项工作

的支持。

7. REFERENCES

- [1] Fei Li, Jingya Bai, Mengyun Zhang, and Ruoyu Zhang, “Yield estimation of high-density cotton fields using low-altitude uav imaging and deep learning,” *Plant Methods*, vol. 18, no. 1, pp. 55, 2022.
- [2] Prayma Bishshash, Asraful Sharker Nirob, Habibur Shikder, Afjal Hossan Sarower, Touhid Bhuiyan, and Sheak Rashed Haider Noori, “A comprehensive cotton leaf disease dataset for enhanced detection and classification,” *Data in Brief*, vol. 57, pp. 110913, 2024.
- [3] Azeem Ayaz Mirani, Shahzaman Nizamani, Mehwish Leghari, Qazi Ejaz, Mushtaque Ahmed Rahu, et al., “A dataset of cotton leaf images for disease classification using deep learning,” Qazi and Rahu, Mushtaque Ahmed, A Dataset of Cotton Leaf Images for Disease Classification Using Deep Learning.
- [4] Serosh Karim, “Cotton leaf disease dataset,” 2025, <https://www.kaggle.com/datasets/seroshkarim/cotton-leaf-disease-dataset>.
- [5] Fengying Dang, Dong Chen, Yuzhen Lu, and Zhaojian Li, “Yoloweeds: A novel benchmark of yolo object detectors for multi-class weed detection in cotton production systems,” *Computers and Electronics in Agriculture*, vol. 205, pp. 107655, 2023.
- [6] Hanyu Wang, Pengxiang Wu, Kevin Dela Rosa, Chen Wang, and Abhinav Shrivastava, “Multimodality-guided image style transfer using cross-modal gan inversion,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024, pp. 4976–4985.
- [7] Tsu-Jui Fu, Xin Eric Wang, and William Yang Wang, “Language-driven artistic style transfer,” in *Computer Vision –ECCV 2022: 17th European Conference*, Tel Aviv, Israel, Berlin, Heidelberg, 2022, Springer-Verlag.
- [8] Kibeom Hong, Seogkyu Jeon, Junsoo Lee, Namhyuk Ahn, Kunhee Kim, Pilhyeon Lee, Daesik Kim, Youngjung Uh, and Hyeran Byun, “Aespa-net: Aesthetic pattern-aware style transfer networks,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 22758–22767.
- [9] Hyeonmin Lee, Kyoungkook Kang, Jungseul Ok, and Sunghyun Cho, “Cliptone: Unsupervised learning for text-based image tone adjustment,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2024, pp. 2942–2951.
- [10] Lorenzo Agnolucci, Leonardo Galteri, Marco Bertini, and Alberto Del Bimbo, “Arniqa: Learning distortion manifold for image quality assessment,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024, pp. 189–198.
- [11] Anish Mittal, Anush K Moorthy, and Alan C Bovik, “Blind/referenceless image spatial quality evaluator,” in 2011 conference record of the forty fifth asilomar conference on signals, systems and computers (ASILOMAR). IEEE, 2011.
- [12] Jianyi Wang, Kelvin CK Chan, and Chen Change Loy, “Exploring clip for assessing the look and feel of images,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.
- [13] Jack Hessel, Ari Holtzman, Maxwell Forbes, Ronan Le Bras, and Yejin Choi, “Clipscore: A reference-free evaluation metric for image captioning,” in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2021.
- [14] Luke Melas-Kyriazi, Christian Rupprecht, Iro Laina, and Andrea Vedaldi, “Deep spectral methods: A surprisingly strong baseline for unsupervised semantic segmentation and localization,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 8364–8375.
- [15] Hossein Talebi and Peyman Milanfar, “Nima: Neural image assessment,” *IEEE transactions on image processing*, vol. 27, no. 8, pp. 3998–4011, 2018.
- [16] Gihyun Kwon and Jong Chul Ye, “Clipstyler: Image style transfer with a single text condition,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 18062–18071.
- [17] Yi Zhang, Meng-Hao Guo, Miao Wang, and Shi-Min Hu, “Exploring regional clues in clip for zero-shot semantic segmentation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2024, pp. 3270–3280.
- [18] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al., “Learning transferable visual models from natural language supervision,” in *International conference on machine learning*. PMLR, 2021.

- [19] Samuel Stevens, Jiaman Wu, Matthew J Thompson, Elizabeth G Campolongo, Chan Hee Song, David Edward Carlyn, Li Dong, Wasila M Dahdul, Charles Stewart, Tanya Berger-Wolf, et al., “Bio-clip: A vision foundation model for the tree of life,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 19412–19424.
- [20] Sheldon Mascarenhas and Mukul Agarwal, “A comparison between vgg16, vgg19 and resnet50 architecture frameworks for image classification,” in 2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON). IEEE, 2021.
- [21] Razvan V Marinescu, Daniel Moyer, and Polina Golland, “Bayesian image reconstruction using deep generative models,” arXiv preprint arXiv:2012.04567, 2020.
- [22] Yohan Poirier-Ginter and Jean-François Lalonde, “Robust unsupervised stylegan image restoration,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 22292–22301.

9. 数据集文档

描述

该数据集由高质量的航拍和 DSLR 图像组成，专为高级农业应用设计，特别是针对棉花作物。航拍图像提供了广泛的覆盖范围，适合于检测和分割任务，辅助分析作物生长和土地使用效率。DSLR 图像则专门策划用于详细分类作物病害，促进早期检测和管理。

数据集链接

<https://ieee-dataport.org/documents/cot-adcotton-analysis-dataset>

<https://www.kaggle.com/datasets/aamaanakbar/cot-ad1>

数据主体

检测、分割、疾病检测、修复、合成

数据类型

数据集由两个农场整个生命周期内的棉花作物的航拍图像和数码单反相机图像，以及视频构成。

体积约为 310 GB。

图像总数

该数据集包含超过 25,000 张图像。

视频总数

数据集包含超过 140 个视频。

图像格式

这些图像是 JPG、JPEG 和 mp4 格式的。

数据收集时间框架

在两座农场的整个棉花作物收获周期期间收集。

数据来源

通过无人机以 [10,15,115] 米的高度拍摄的航拍图像。

使用变焦模式拍摄的单反相机图像。

航空影像标签详情

分别包含图像的文件夹；分割掩码，即用于作物分割的二值掩码，以 JPG 格式；用于作物分割的 YOLO 标准.txt 格式的分割标签；以及用于作物检测的 YOLO 标准.txt 格式的检测标签。

DSLR 图像的标签细节：每种疾病类型都整理在包含相应 JPG 图像的不同文件夹中。

注释细节

每张图像都经过细致的人工注释，并经过多轮严格的质量检查，以确保高精度和可靠性。

数据集所有者及其隶属机构

印度古吉拉特邦甘地纳格尔理工学院

版本详情

当前版本是 1.0

维护状态

限定维护 - 数据将不会更新，但任何技术问题都会得到解决。

图 1 和图 2 主论文中提供的示例数据。

数据集组织方式

数据集的组织方式在第 15 节中提供

许可

署名-非商业性 4.0 国际

DOI

10.21227/bpqn-9a12(IEEE Dataport)和 10.34740/kaggle/dsv/10674370 (Kaggle)

作者声明

如果发生权利等的侵犯，我们承担所有责任。

农业和食品管理是未来世界将面临的关键挑战。在作物管理和可持续实践方面的进步对于确保充足的食物供应至关重要。利用航空器和相机结合先进的计算机视觉技术可以解决作物管理的各个方面，如病害检测、杂草密度分析和收获信息，从而使农民能够做出及时和明智的决定。本文提出了两个专注于棉花作物的综合数据集：一个航拍数据集，记录了作物整个生命周期，以及一个详细描绘不同阶段棉花作物所受疾病影响的近距离单反相机图像数据集。航拍数据集经过精心标注，用于棉花作物的检测和分割，提供了特定农场区域内植被和杂草竞争的重要见解。单反相机数据集包括叶病、棉铃腐烂和红棉虫的图像，以促进棉花病害的分类和早期管理。

我们的数据集在现有农业数据库中填补了关键空白，提供了针对棉花作物特别优化的细粒度分割数据，使其在植物级别和田间级别的检测和分割任务中更具实用性。以每周整个收获周期拍摄的超过 5000 张手动注释的图像为基础，这是最广泛的数据集之一。这些数据集可用作创建创新技术解决方案的基准，以提升农业实践，为更高效和可持续的农业打开大门。食品管理和农业是世界在不久的将来将面临的重要问题。作物管理和可持续实践的进步将是确保充足食品供应的关键。利用无人机和相机结合先进的计算机视觉技术可以解决作物管理的各个方面，如疾病检测、杂草密度分析和收获信息，使农民能够就作物管理做出及时且明智的决策。本文介绍了整个作物生命周期中拍摄的棉田航空数据集，以及有关棉花作物在不同生命周期阶段受到各种疾病影响的近距离 DSLR 图像数据集。

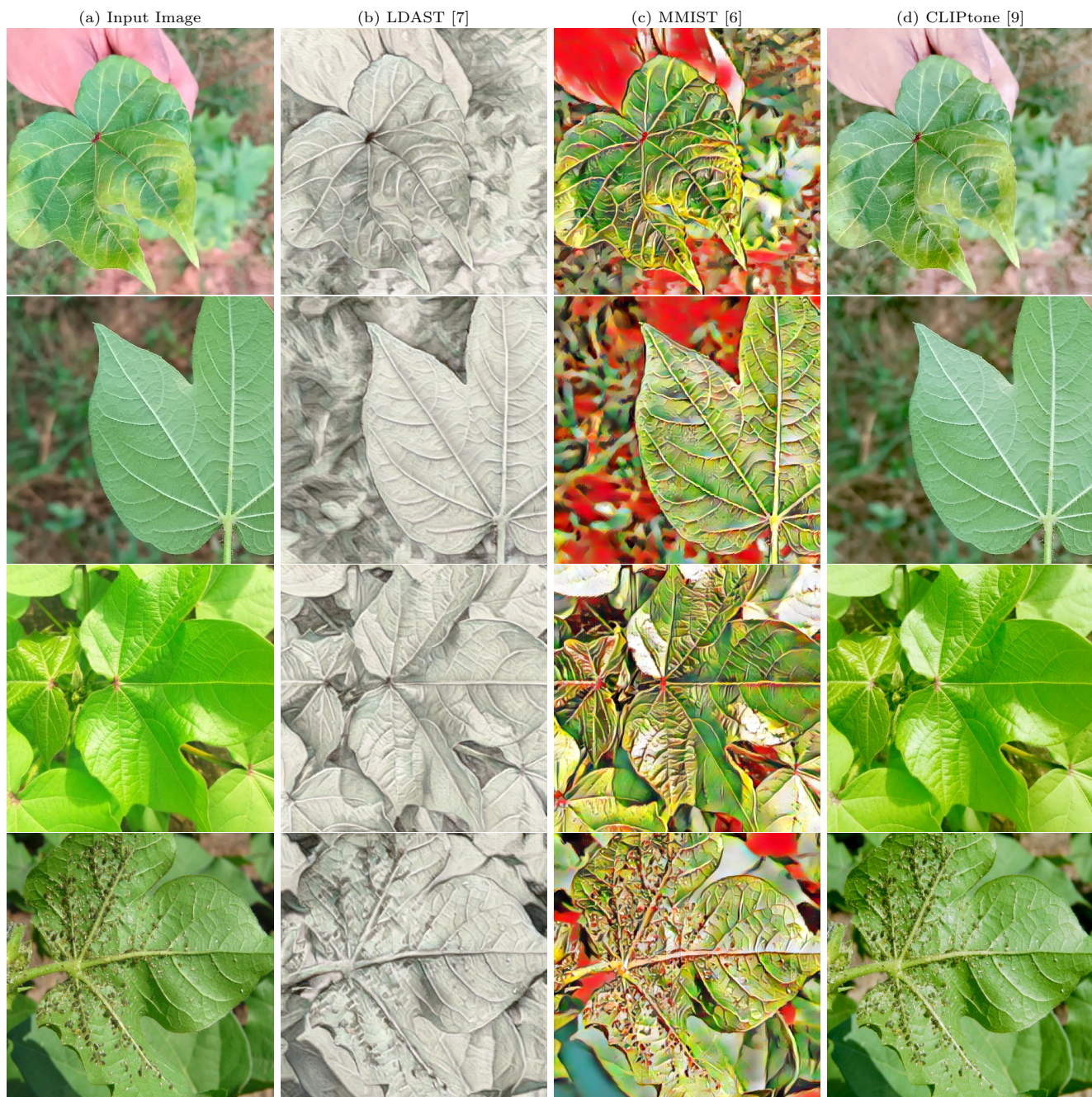


Fig. 9: 图像增强结果。每行显示一个不同的数据集（参见主论文表 1）（从上到下：增强（棉花叶片病害检测 [2] ）、原始（棉花叶片病害检测）数据集 [2]、棉花叶片病害数据集 [4]，以及用于病害检测和分类的棉花叶片图像数据集 [3]）的图像和四种方法的增强结果：LDAST [7]、MMIST [6] 和 CLIPtone [9]。提示：一张干净的、高质量的图像，具有平衡的对比度和亮度，没有噪点或任何失真。

于获取关于农场上存在的疾病的信息，例如它们的名称、成因、采取的措施以及减少其影响的杀虫剂信息。

我们已经在 Kaggle 和 IEEE DataPort 上共享了 COT-AD 数据集，该数据集大约包含 308 GB 的数据。对访问该数据集感兴趣的研究人员和从业人员可以访问我们的项目页面 ([https://aamaanakbar.github.io/COT-](https://aamaanakbar.github.io/COT-AD/)

AD/), 或者我们在数据集文档中也提供了数据集的链接。该数据集仅供非商业用途。

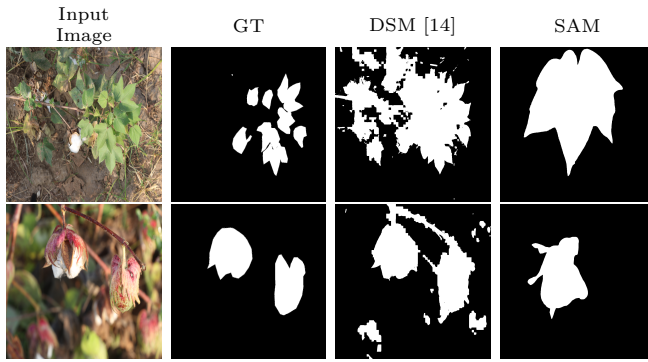


Fig. 10: 使用监督 (SAM) 和无监督 [14] 方法的物体分割。



Fig. 11: 数据采集设备: (a) DJI MAVIC Air 2s (b) 单反相机 Canon 80D (w) 图像来源: Canon 亚洲)

15. 数据集组织

用于检测和分割的航拍图像: 用于检测和分割任务的数据组织在一个名为“用于检测和分割的航拍图像”的目录下, 该目录分为四个主要部分:

- 第一部分: 前两个月的数据。
- 第二部分: 第三个月的数据。
- 第三部分: 第四个月的数据。
- 第 D 部分: 来自第五和第六个月的数据。

每个部分包含四个子文件夹:

疾病分类数据集: 在目录“高分辨率 DSLR 棉花作物信息图片”下组织, 图像的分类基于三个类别——叶子疾病、棉铃和虫害。

- Leaf 文件夹包含四个子目录——黄叶、叶斑病、细菌性枯萎病、叶片变红和新鲜叶。
- 棉铃文件夹包含三个子目录: Boll Rot、Damaged Cotton Boll 和 Healthy Cotton Boll。
- “Bugs” 文件夹包含两个子目录——“Mealy Bug”和“Red Cotton Bug”。

该组织提高了数据处理的访问性和效率, 适用于与检测、分割和疾病分类任务相关的各种模型训练目的。

在主论文的表 (1) 中展示了不同数据上的

16. 结果

图像分类, 我们对这些数据进行了分类和分割任务。请参见表格 8 中的结果。我们使用多个数据集和多种方法进行了图像增强。这些增强实验的定性结果在图 9 中进行了说明。

对于图像预测, 我们采用了 YOLOv11 模型, 并在图 (14, 17) 中使用混淆矩阵展示了 DSLR 和无人机数据集的结果

该研究和数据收集得到了印度瓦多达拉的 L & T 技术服务有限公司的支持, 与印度甘地讷格尔的印度理工学院和 Jibaben Patel AI 主席合作。我们感谢 L & T 技术服务的研究团队, 特别是 Ashokkumar Jain 先生、Pushkar Shaktawat 先生和 Nikhil Dev 先生。此外, 我们还感谢位于甘地讷格尔的印度理工学院的 Prakram Singh Rathore 先生、Medhansh Singh 先生、Pavidhar Jain 先生、Yash Khandelwal 先生和 Harshvardhan Vala 先生为航拍图像的标注所作的努力。

Method	DSLR Cotton Crop Data			Drone Cotton Crop Data		
	LPIPS $\hat{a}^\dagger$	LPIPS-	PFID $\hat{a}^\dagger$	LPIPS $\hat{a}^\dagger$	LPIPS-	PFID $\hat{a}^\dagger$
	“	vgg $\hat{a}^\dagger$	“	“	vgg $\hat{a}^\dagger$	“
Deartifacting (JPEG Compression)						
L	0.584	0.619	121.419	0.598	0.588	120.462
M	0.576	0.613	115.922	0.583	0.580	112.291
S	0.573	0.611	114.228	0.579	0.576	113.928
XL	0.6	0.633	118.723	0.615	0.601	157.034
XS	0.571	0.609	115.213	0.577	0.575	113.201
Denoising (Clamped Poisson and Bernoulli Mixture)						
L	0.58	0.628	116.625	0.576	0.583	130.393
M	0.573	0.62	113.264	0.570	0.579	115.636
S	0.567	0.613	109.726	0.568	0.575	114.738
XL	0.592	0.64	115.037	0.591	0.591	146.684
XS	0.56	0.607	113.883	0.562	0.571	107.531
Inpainting (Random Strokes)						
L	0.553	0.596	126.934	0.527	0.553	112.856
M	0.548	0.59	121.841	0.528	0.552	106.999
S	0.542	0.585	114.523	0.524	0.549	104.970
XL	0.558	0.602	133.558	0.533	0.557	109.472
XS	0.551	0.59	114.447	0.543	0.556	96.843
Upsampling (Bilinear, Bicubic or Lanczos)						
L	0.619	0.642	129.074	0.604	0.599	154.395
M	0.598	0.632	126.482	0.606	0.598	140.140
S	0.581	0.619	119.499	0.592	0.587	103.984
XL	0.628	0.651	130	0.620	0.609	161.933
XS	0.566	0.605	116.894	0.578	0.577	102.341
2 Degradations						
AP	0.57	0.61	121.229	0.542	0.569	112.517
NA	0.584	0.627	114.7	0.584	0.589	136.752
NP	0.57	0.615	122.015	0.538	0.566	118.027
UA	0.639	0.673	127.588	0.636	0.617	163.628
UN	0.623	0.66	116.295	0.604	0.605	140.503
UP	0.588	0.626	120.72	0.573	0.583	139.341
3 Degradations						
NAP	0.578	0.621	122.239	0.542	0.569	123.345
UAP	0.617	0.66	124.197	0.592	0.596	119.654
UNA	0.644	0.68	119.279	0.614	0.612	138.181
UNP	0.624	0.654	124.385	0.557	0.578	108.822
4 Degradations						
UNAP	0.638	0.67	120.707	0.598	0.591	107.681

Table 7: RUSIR [22] 在单一降解类型（去工件化、去噪、修补和上采样）和多种降解类型中，在五个级别（S, M, L, XL, XS）上对来自单反和无人机相机的棉花作物数据进行图像修复的性能指标。RUSIR 方法在所有级别上使用一致的超参数集，而基准线在每个单独的级别上都针对准确性进行优化。报告的指标包括准确性（LPIPS）、保真度（LPIPS-vgg）和真实感（PFID），值越低表示性能越好。

Model	Zero Shot (Top-5)	Linear Probing (Accuracy)
Augmented (Cotton Leaf Disease Detection) Dataset [2]		
CLIP (ViT-B/32) [18]	93.67	91.54
CLIP (ViT-B/16) [18]	96.72	95.86
CLIP (RN50) [18]	98.88	88.74
BioCLIP [19]	67.41	94.54
BioCLIP (ViT-B/16, iNat-only) [19]	69.98	93.21
Original (Cotton Leaf Disease Detection) Dataset [2]		
CLIP (ViT-B/32) [18]	91.76	90.74
CLIP (ViT-B/16) [18]	92.85	93.38
CLIP (RN50) [18]	95.73	84.37
BioCLIP [19]	61.98	91.87
BioCLIP (ViT-B/16, iNat-only) [19]	64.93	91.36
cotton leaf disease dataset [4]		
CLIP (ViT-B/32) [18]	92.40	87.85
CLIP (ViT-B/16) [18]	94.95	90.37
CLIP (RN50) [18]	96.83	84.96
BioCLIP [19]	64.81	88.36
BioCLIP (ViT-B/16, iNat-only) [19]	66.32	89.76
A Dataset of Cotton Leaf Images for Disease Detection and Classification [3]		
CLIP (ViT-B/32) [18]	94.00	86.69
CLIP (ViT-B/16) [18]	95.88	89.43
CLIP (RN50) [18]	97.62	84.86
BioCLIP [19]	65.27	87.46
BioCLIP (ViT-B/16, iNat-only) [19]	67.66	90.96

Table 8: 图像分类：在四个数据集上对比模型的零样本（Top-5）和线性探测（Top-1）准确率。

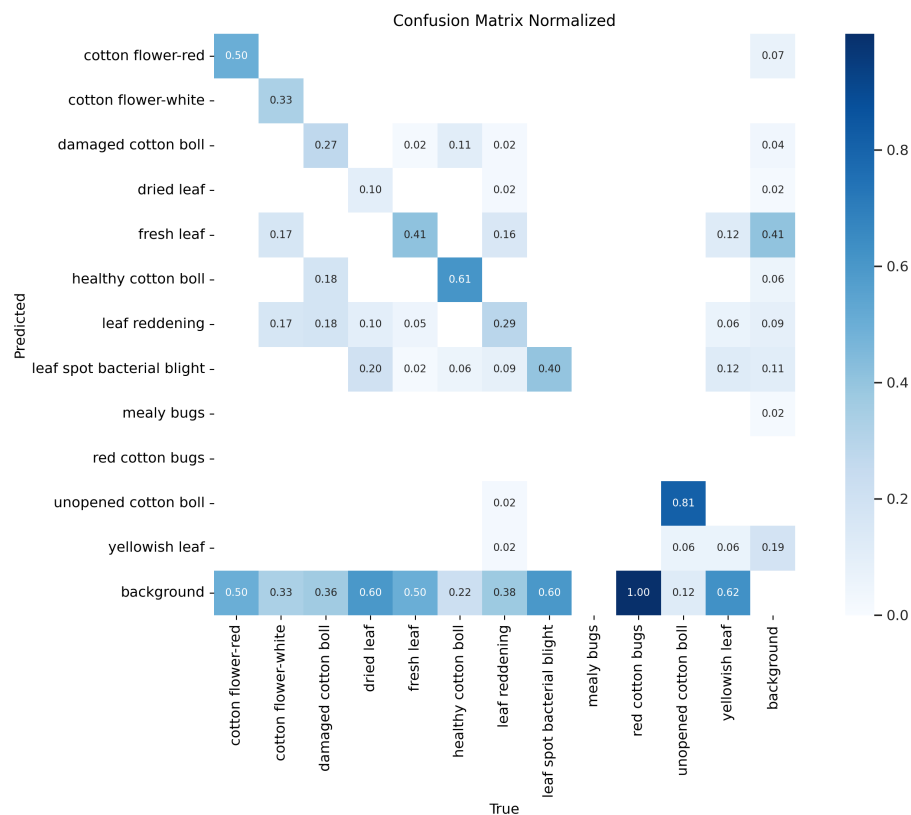


Fig. 12:
(a) 大—normalised confusion matrix

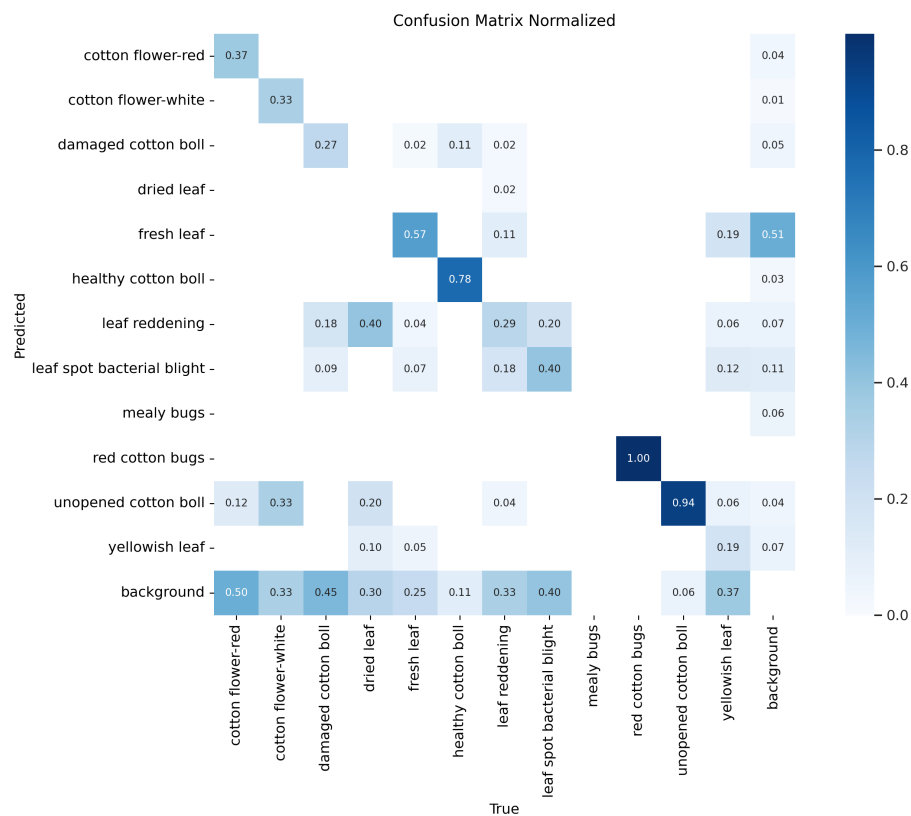


Fig. 13:
(c) Xtra —Normalised Confusion Matrix

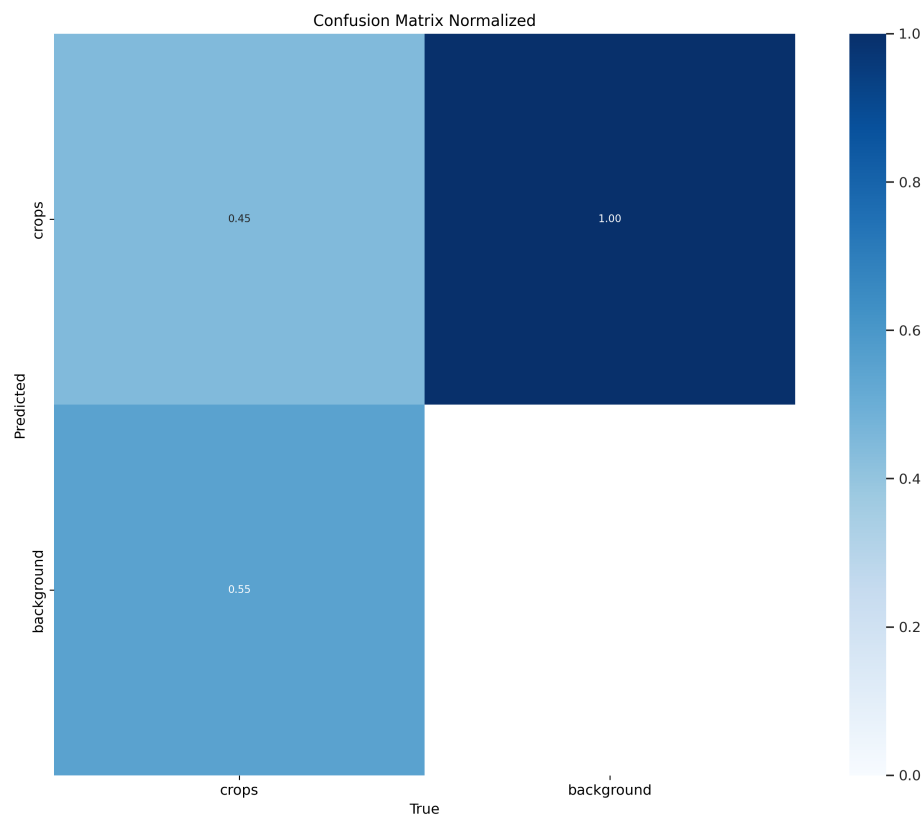


Fig. 15:
(a) 大—normalised confusion matrix

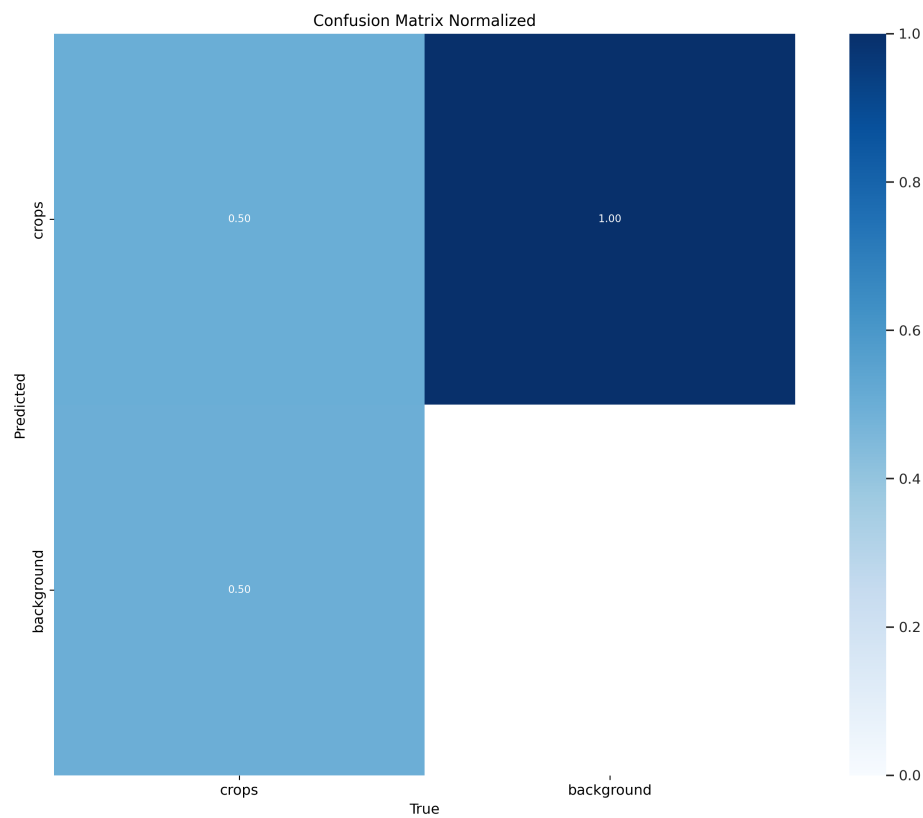


Fig. 16:
(c) Xtra â€” Normalised Confusion Matrix