WeedVision: Multi-Stage Growth and Classification of Weeds using DETR and RetinaNet for Precision Agriculture

Taminul Islam[a], Toqi Tahamid Sarker[a], Khaled R Ahmed[a], Cristiana Bernardi Rankrape[b], Karla Gage[b]

[a] BASE Lab, School of Computing, Southern Illinois University, Carbondale, IL 62901, United States

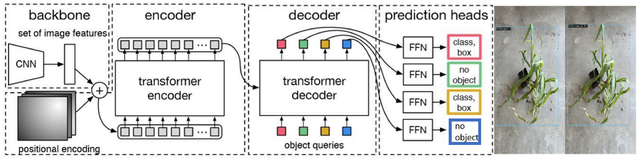
[b] School of Agricultural Sciences, Southern Illinois University, Carbondale, IL 62901, United States

Motivation

- Enhance precision in detecting and classifying diverse weed species across various growth stages using AI.
- Reduce herbicide usage and environmental impact through accurate, automated, and species-specific weed management strategies.
- Mitigate agricultural yield losses by employing advanced object detection models for precise weed identification.

Contributions

- Developed a comprehensive dataset of 203,567 images, covering 16 weed species across 11 growth stages.
- Labeling of the dataset, categorized by species and growth stage (week-wise), providing a comprehensive resource for weed identification research.
- Implemented and compared DETR and RetinaNet models for accurate weed detection and classification.



Detection Transformer (DETR) Architecture [1]. The model uses a CNN backbone, transformer encoder-decoder, and prediction heads to detect and classify objects, including weeds, as shown in the final output.

Method

- Applied data augmentation techniques, including normalization, color space conversion, scaling, and thresholding, and labeled images with bounding boxes for accurate detection.
- To ensure data accuracy, we implemented a QA process, manually refining annotations using LabelImg software [2].
- Utilized RetinaNet with a ResNeXt-101 backbone and DETR with a ResNet-50 backbone for effective weed detection and classification.
- Trained models on an 80% training, 10% validation, and 10% testing split, evaluating performance through metrics like mAP, recall, and inference speed.

Dataset

- With 16 species of weeds this research encompassed a total of 174 classes.
- The full dataset initially comprised 2,494,476 frames. After a thorough review process to remove substandard images, 203,567 images were ultimately selected for training.
- Images were captured under controlled greenhouse conditions to maintain consistency and reduce noise.
- Data includes diverse images from seedling to mature stages, ensuring robust training and evaluation, and captured with iPhone 15 Pro Max. Stopped capturing images when the weeds entered their flowering stages (week 11).

OVERVIEW OF WEED SPECIES OF ECONOMIC CONCERN, CORRESPONDING CODES, AND WEEKLY FRAME COUNTS CAPTURED FOR EACH SPECIES ACROSS 11 WEEKS IN THE GREENHOUSE

Species Code [20]	Scientific Name [21]	Common Name [22]	Family	Total Frames	Number of frames/week										
					W_1	*W_2*	*W_3*	*W_4*	*W_5*	*W_6*	*W_7*	*W_8*	*W_9*	*W_10*	*W_11*
ABUTH	*Abutilon theophrasti* Medik.	Velvetleaf	Malvaceae	14754	1084	2451	1212	1819	1414	981	677	1164	1084	1500	1368
AMAPA	*Amaranthus palmeri* S. Watson.	Palmer Amaranth	Amaranthaceae	17525	1441	1408	2110	2014	2441	1290	923	1478	1393	1667	1360
AMARE	*Amaranthus retroflexus* L.	Redroot Pigweed	Amaranthaceae	15380	1017	1363	2110	1923	1884	1150	736	1237	1082	1596	1282
AMATU	*Amaranthus tuberculatus* (Moq.) Sauer.	Water Hemp	Amaranthaceae	14852	1325	1459	1565	1664	1942	837	730	969	1638	1573	1150
AMBEL	*Ambrosia artemisiifolia* L.	Common Ragweed	Asteraceae	17427	1022	2215	1846	1739	2162	1093	1066	1432	1092	2045	1715
CHEAL	*Chenopodium album* L.	Common Lambsquarter	Chenopodiaceae	8015	1108	954	1416	661	1056	393	418	641	453	429	574
CYPES	*Cyperus esculentus* L.	Yellow Nutsedge	Cyperaceae	14275	909	1512	1032	1499	2273	978	1224	1391	1182	1170	1105
DIGSA	*Digitaria sanguinalis* (L.) Scop.	Large Crabgrass	Poaceae	16962	732	1312	2411	2596	1649	1335	1166	1261	1120	1628	1692
ECHCG	*Echinochloa crus-galli* (L.) P. Beauv.	Barnyard Grass	Poaceae	16564	1349	2067	2029	1426	2221	1240	929	1280	1371	1332	1320
ERICA	*Erigeron canadensis* L.	Horse Weed	Asteraceae	15134	930	2183	1691	1542	2715	1189	609	742	915	1217	1401
PANDI	*Panicum dichotomiflorum* Michx.	Full Panicum	Poaceae	15182	1198	1400	2143	1296	1979	952	887	1350	1425	1034	1518
SETFA	*Setaria faberi* Herrm.	Gaint Foxtail	Poaceae	14635	1614	1195	2083	1348	1944	1091	715	1466	843	1342	994
SETPU	*Setaria pumila* (Poir.) Roem.	Yellow Foxtail	Poaceae	15211	887	1390	1732	1654	2040	1093	747	1361	1325	1348	1634
SIDSP	*Sida spinosa* L.	Princkly Sida	Malvaceae	14452	1035	1782	1583	1259	2142	1373	804	1059	1186	1303	926
SORHA	*Sorghum halepense* (L.) Pers.	Johnson Grass	Poaceae	10958	0	0	1444	1268	1395	945	749	1215	1328	1116	1459
SORVU	*Sorghum bicolor* (L.) Moench.	Shatter Cane	Poaceae	9573	945	1340	1959	832	1065	525	279	748	714	592	574

Results

- RetinaNet achieved the best mean Average Precision (mAP) of 0.904 and the mean Average Recall (mAR) of 0.989, outperforming DETR's mAP of 0.840 and mAR of 0.936.
- In terms of inference speed, RetinaNet reached 7.28 FPS, significantly faster than DETR's 3.49 FPS.

PERFORMANCE COMPARISON OF DETR AND RETINANET ON TRAINING AND TEST SETS

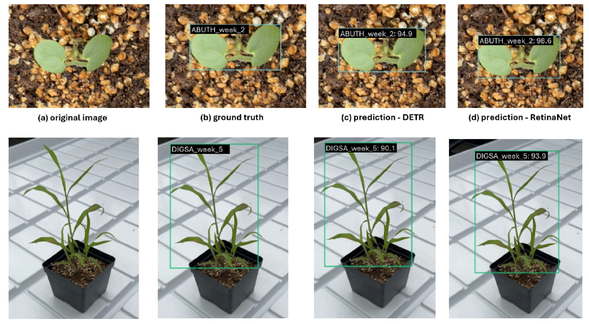
Model	mAP		mAR		FPS
	Train	*Test*	*Train*	*Test*	
DETR	0.854	0.840	0.941	0.936	3.49
RetinaNet	**0.907**	0.904	**0.997**	0.989	**7.28**

PERFORMANCE COMPARISON OF DETR AND RETINANET FOR SETFA

Class Name	DETR				RetinaNet			
	mAP	*mAP_50*	*mAP_75*	*Recall*	*mAP*	*mAP_50*	*mAP_75*	*Recall*
SETFA_Week_1	0.355	0.605	0.348	0.986	0.671	0.870	0.767	0.980
SETFA_Week_2	0.400	0.763	0.414	1.000	0.623	0.801	0.738	0.989
SETFA_Week_3	0.740	0.999	0.887	0.929	0.755	0.991	0.830	1.000
SETFA_Week_4	0.607	0.860	0.717	0.974	0.555	0.764	0.611	1.000
SETFA_Week_5	0.741	0.964	0.868	1.000	0.657	0.899	0.708	0.995
SETFA_Week_6	0.658	0.859	0.669	1.000	0.648	0.936	0.682	1.000
SETFA_Week_7	0.825	0.974	0.860	1.000	0.822	0.980	0.795	1.000
SETFA_Week_8	0.743	1.000	0.818	1.000	0.822	1.000	0.943	1.000
SETFA_Week_9	0.856	0.955	0.949	1.000	0.843	0.983	0.932	1.000
SETFA_Week_10	0.696	0.956	0.802	0.986	0.643	0.956	0.738	0.986
SETFA_Week_11	0.859	1.000	0.988	1.000	0.808	1.000	0.938	1.000

PERFORMANCE COMPARISON OF DETR AND RETINANET FOR ABUTH

Class Name	DETR				RetinaNet			
	mAP	*mAP_50*	*mAP_75*	*Recall*	*mAP*	*mAP_50*	*mAP_75*	*Recall*
ABUTH_Week_1	0.418	0.723	0.471	0.994	0.605	0.899	0.689	1.000
ABUTH_Week_2	0.576	0.988	0.530	1.000	0.829	0.990	0.952	1.000
ABUTH_Week_3	0.356	0.697	0.346	1.000	0.700	0.996	0.899	1.000
ABUTH_Week_4	0.408	0.771	0.396	0.996	0.725	0.973	0.844	0.995
ABUTH_Week_5	0.445	0.923	0.377	0.871	0.730	0.974	0.789	1.000
ABUTH_Week_6	0.850	1.000	1.000	0.886	0.924	0.970	0.970	0.972
ABUTH_Week_7	0.885	0.932	0.931	0.993	0.966	1.000	1.000	1.000
ABUTH_Week_8	0.856	1.000	0.982	1.000	0.911	1.000	1.000	1.000
ABUTH_Week_9	0.912	0.977	0.949	0.975	0.876	0.978	0.920	1.000
ABUTH_Week_10	0.860	0.967	0.923	1.000	0.868	0.971	0.893	1.000
ABUTH_Week_11	0.924	1.000	1.000	0.989	0.924	1.000	1.000	1.000



Comparison of object detection results for ABUTH and DIGSA using DETR and RetinaNet models. Row 1 displays predictions for ABUTH, and Row 2 displays predictions for DIGSA, with ground truth and model confidence scores indicated for each detection.

[1] Zhu, X., Su, W., Lu, L., Li, B., Wang, X., & Dai, J. (2020). Deformable detr: Deformable transformers for end-to-end object detection. arXiv preprint, arXiv.

[2] L. Tzutalin, "Labelimg," https://github.com/tzutalin/labelImg, 2015, accessed: Oct 17, 2024.