

Weet wat er leeft:

Real time monitoring van plagen en bestrijders



Eindrapport TKI PPS Weet wat er leeft

Auteurs: Anne Jan Doorn, Bruce Schoelitsz, Bas Allema, Bas Boom, Janne Kool, Peter Offermans, Moniek Brederland, Jan-Kees Boerman, Ivan Valdes, Onno Calf, Jolanda Korteland, Margje Voeten, Joost Simons, Ron Saris, Thomas Verschoor, Tom Vroegop en Joop Woelke

Rapportgegevens

Auteurs: Anne Jan Doorn, Bruce Schoelitsz, Bas Allema, Bas Boom, Janne Kool, Peter Offermans, Moniek Breederland, Jan-Kees Boerman, Ivan Valdes, Onno Calf, Jolanda Korteland, Margje Voeten, Joost Simons, Ron Saris, Thomas Verschoor, Tom Vroegop en Joop Woelke

Opdrachtgever: Kennis in je Kas

Uitvoerder: Stichting Control in Food & Flowers

Contactpersoon: Anne Jan Doorn (anne.jan.doorn@scff.nl)

Datum: April 2025

Projectnummer: LWV20.166

Disclaimer

© 2025 Delfgauw, Stichting Control in Food & Flowers
Distributieweg 1, 2645 EG Delfgauw, T +31 (0)15 - 25 72 511, www.scff.nl.
Kamer van Koophandel nummer: 61916471
BTW nummer: NL 8545.47.277.B01
Stichting Control in Food & Flowers (SCFF). Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier zonder voorafgaande schriftelijke toestemming van SCFF. SCFF aanvaardt geen aansprakelijkheid voor eventuele schade voortvloeiend uit het gebruik van de resultaten van dit onderzoek of de toepassing van de adviezen.

Adresgegevens

Stichting Control in Food & Flowers
Distributieweg 1, 2645 EG Delfgauw
T +31 (0)15 - 25 72 511
www.scff.nl

Detectie van de plaaginsecten (groene perzikluisk, kaswittevlug en Californische trips) en de biologische bestrijders (*Aphidius colemani*, *Transeius montdorensis* en *Amblyseius swirskii*) met de technieken: Vision, AI, VOC en eDNA

Werkgroep Vision:

- imec OnePlanet – Bas Boom en Peter Offermans
- WUR – Bas Allema en Janne Kool
- HAS green academy – Bruce Schoelitsz en Margje Voeten
- SCFF – Joop Woelke, Onno Calf, Jolanda Korteland en Anne Jan Doorn

Werkgroep Artificieel Intelligence (AI):

- imec OnePlanet – Bas Boom en Peter Offermans
- Koppert – Thomas Verschoor en Tom Vroegop
- Letsgrow – Joost Simons en Ron Saris
- SCFF – Onno Calf en Jolanda Korteland

Werkgroep VOC's:

- EMS – Jan-Kees Boerman en Moniek Breederland
- SCFF – Onno Calf en Jolanda Korteland

Werkgroep eDNA:

- SCFF – Onno Calf en Jolanda Korteland
- Groen Agro Control – Ivan Valdes



HET NIEUWE DOEN
IN PLANTGEZONDHEID



LABORATORIUM-
ONDERZOEK
& ADVIES



Samenvatting

De Nederlandse glastuinbouw staat wereldwijd bekend als vooruitstrevend en innovatief. Hier kunnen nieuwe technieken om plaaginsecten te kunnen detecteren aan bijdragen. In dit PPS project zijn automatische beeldherkenning (Vision), environmental DNA (eDNA), Volatile Organic Compound (VOC) metingen en Artificial Intelligence (AI) detectiemethoden onderzocht, vooral om de volgende soorten te kunnen detecteren: Californische trips (*Frankliniella occidentalis*), groene perzikluis (*Myzus persicae* var. *persicae*), kaswittevlies (*Trialetrodes vaporariorum*), roofmijt *Traneius montdorensis*, roofmijt *Amblyseius swirskii* en sluipwesp *Aphidius colemani*.

Vision

Automatische beeldherkenningssystemen kunnen ondersteunen bij een vroege detectie van plaaginsecten. Er zijn verschillende manieren waarop automatische beeldherkenning kan worden toegepast. Binnen Weet wat er leeft zijn er drie toegepast: fotografie van plakplaten met de camera van een mobiele telefoon, fotografie van plakplaten met een hyperspectraal camera en fotografie van insecten en mijten afkomstig van bladmateriaal door middel van een Edapholog.

De keuze voor de mobiele telefoon is gebaseerd op de wensen van telers dat is opgehaald met behulp van een enquête. Er zijn modellen getraind van de gevleugelde doelsoorten Californische trips, groene perzikluis, kaswittevlies en de sluipwesp *A. colemani*, met data uit zowel de praktijk als onder gecontroleerde omstandigheden. Omdat van verschillende doelsoorten sterkgelijkende verwante soorten aanwezig kunnen zijn in een kas, is ervoor gekozen het bij het trainen de labels trips, wittevlies, bladluis en sluipwesp te gebruiken. Bij de toepassing van deze methode zijn verschillende factoren relevant gebleken die van invloed zijn op de prestaties van het model:

Het aantal soorten op de foto had een negatieve invloed. Dit is opgelost door de foto's op te knippen in fragmenten en de fragmenten te analyseren in plaats van de gehele foto. Bijkomend voordeel was dat hierdoor meer details zichtbaar waren van de insecten. Lichtomstandigheden in de kas zijn variabel. Er is een externe lichtbron (flits) ontwikkeld dat de lichtomstandigheden standaardiseerde. Door het model te trainen met zowel foto's met omgevingslicht als met de externe lichtbron, kon het overweg met de verschillende lichtcondities. Over het algemeen werd alleen wittevlies beter herkend met als zonder flits. De prestaties van het model door de tijd heen waren echter wel consistent als de flits gebruikt werd. Daarnaast was het mogelijk het effect van paars LED-licht te niet te doen met de flits. Dat was ook mogelijk met digitale nabewerking van de foto's.

Verschillende smartphones hebben verschillende soft- en hardware. Verschillende modellen zijn getraind voor drie verschillende telefoons (iPhone 7, iPhone 13 en Samsung A52). De prestaties van deze modellen verschilden erg tussen de verschillende telefoontypen. Door een model te trainen met foto's van alle telefoons, was het model echter in staat vergelijkbare resultaten te geven voor foto's van al deze telefoons.

Ook de vangstperiode is van invloed op de prestaties van het model. Hoe langer een plakplaat blijft hangen, hoe meer degradatie van insecten en vervuiling van de plakplaat plaatsvindt. Over het algemeen geldt dat de prestaties achteruitgaan naarmate de platen langer dan een week blijven

hangen. Dit was echter afhankelijk van de soort en het algoritme dat gebruikt is om het model te trainen. Bij het trainen van het model zijn foto's gebruikt van gele plakplaten met droge lijm. In de praktijk komen ook andere kleuren en natte lijm voor. Het model werkte redelijk met verschillende kleur-soortcombinaties bij droge lijm, maar werkte niet goed met natte lijm. Door nieuwe modellen te trainen was het mogelijk om de prestaties te verbeteren met een beperkte dataset met droge lijm, maar ging dit minder goed met natte lijm. Reflecties in de lijm werden foutief herkend als insecten en insecten zonken weg en degradeerden in de lijm.

Deze techniek is meerdere keren in de praktijk en in de HAS-kas toegepast, waar met name voor trips en wittevlieg werkbare resultaten werden behaald. De overige soorten werden weinig aangetroffen. In de laatste fase van Weet wat er leeft is de techniek in alle doelgewassen toegepast. Daaruit bleek dat de modellen nog niet goed genoeg in staat waren de insecten te herkennen en identificeren. Daarvoor zijn meerdere verklaringen: in de trainset is de achtergrond (deel van de foto achter de plakplaten) verwijderd, in de praktijk niet. In de praktijk kwamen soorten voor die niet in de trainset aanwezig waren. In de praktijklocaties kwamen veel minder doelorganismen voor dan in de trainset. Dit heeft ertoe geleid dat relatief groot aandeel van voorgestelde identificaties foutief waren. Elementen in de achtergrond werden als insecten herkend, niet eerder waargenomen niet-doelsoorten werden als doelsoorten genoemd en omdat weinig doelorganismen aanwezig waren, had een foute identificatie direct een grote impact op de prestaties van het model. Hoewel de resultaten in het laatste experiment tegenvielen, is wel aangetoond dat deze techniek potentie had bij de andere (praktijk)proeven. De foto's die in de vier kassen gemaakt zijn, zijn gelabeld, waardoor ze gebruikt kunnen worden voor verdere training van het model, zodat het model beter overweg kan met deze omstandigheden.

Hyperspectraal

Er zijn experimenten gedaan met het fotograferen van insecten met een hyperspectrale camera. Hyperspectrale camera's kunnen lichtfrequenties waarnemen die mensen niet kunnen zien, zoals infrarood. Binnen hyperspectrale camera's zijn er twee gangbare spectrumgroepen, namelijk VNIR en SWIR, waarbij we hebben aangetoond dat VNIR meer potentie heeft voor insectherkenning. Daarnaast is er gezamenlijk met de HAS een experimentele dataset gemaakt. Hierop hebben we uiteindelijk laten zien dat hyperspectrale camera's op familieniveau beter onderscheid kunnen maken dan normale camera's. De uiteindelijke vraag hierbij is echter of deze resultaten significant veel beter zijn om de hoge prijs van dit soort camera's te rechtvaardigen. Gezien deze doelsoorten denken wij dat dit waarschijnlijk nog niet interessant is voor de telers.

Edapholog

Om de nietvliegende doelsoorten van Weet wat er leeft te monitoren is de Edapholog gebruikt. De Edapholog is een commerciële geautomatiseerde versie van de klassieke Berlese val, die gebruikt wordt voor het onderzoeken van kruipend bodemleven. De Edapholog is hier toegepast voor het detecteren van de mijten en tripslarven in bladmonster. De werking is vergeleken met de nu standaard spoelmethode en Berlese val. Er is een algoritme ontwikkeld waarmee de mijten geteld kunnen worden; de roofmijten zijn niet van elkaar te onderscheiden. Wel ontwikkelde de tripseieren zich sneller tot larven doordat de opstelling warm is. Er is dus potentie om de Edapholog in te zetten om de tripsdruk van de komende dagen te voorspellen mits deze commercieel beschikbaar blijft.

eDNA

eDNA-technologieën hebben potentieel om de detectiemogelijkheden te verbeteren, met name omdat met eDNA soorten werden gevonden, die met andere detectiemethoden niet werden gevonden. Echter, tijdens dit onderzoek is ook gebleken dat zowel PCR-analyse als de shotgun methode, zoals op deze manier uitgevoerd, nog niet voldoende zijn om alle soorten goed te monitoren en zeker niet altijd betrouwbaar zijn over de aantallen van insecten. Met PCR kunnen doelsoorten aangetoond worden, maar het leverde niet consequent goede resultaten. Er werd aangetoond dat qua sampling methoden het wattenstaafje met buffer minder effectief was in vergelijking met de andere methoden zoals de vanglamp en de vangplaat. De PCR-analyse was betrouwbaarder dan de shotgunmethode. Met de shotgun methode werden doelsoorten weinig gevonden. Gezien deze techniek ook kostbaar is als het voor enkele monsters wordt gebruikt, is het niet waarschijnlijk dat de techniek andere detectiemethoden gaat vervangen, maar hoogstens als aanvulling gaat werken.

VOC

Er zijn vluchtige stoffen gemeten die als indicatie kunnen dienen dat er plaaginsecten actief zijn. Zo werd aangetoond dat het niveau van ethyleen fors hoger is als er plagen aanwezig zijn op de plant in vergelijking met planten zonder plagen. Ook heeft EMS gemeld dat er met technieken van hen waargenomen kon worden dat insecten actief waren op de plant. Deze metingen zijn echter op laboratorium niveau gedaan en nog niet toepasbaar in de praktijk. Dit heeft te maken met het feit dat in een kas veel planten en plagen zijn die samen heel veel VOC's afgeven die zo veel zijn dat het als "ruis" gezien wordt in de metingen. Bepaling wat van wat komt is daardoor niet goed mogelijk.

AI

Het doel was om de data vanuit de verschillende sensoren die getest werden tijdens dit project te integreren middels Artificiële Intelligentie (AI), waarbij de sensoren elkaar ook zouden versterken. AI kan door de data te combineren tot een uiteindelijke voorspelling komen van de plaagdruk in de kas. Tijdens het project is aangetoond dat het mogelijk is om een platform te ontwikkelen voor sensoren waarbij de sensordata met Artificiële Intelligentie wordt geanalyseerd en de data beschikbaar wordt voor de gebruiker. Het is echter niet gelukt om sensortechnologieën te combineren, doordat veel experimentele sensoren veel problemen ondervonden en niet tijdig klaar waren voor gecombineerde experimenten. We laten zien dat het belangrijk is een dataplatform te kiezen waarin de data op een gestandaardiseerde manier kan worden opgeslagen. Omdat de sensordata in het project niet altijd geleverd is, was er ook een gebrek aan data, terwijl Artificiële Intelligentie veel data nodig heeft om tot goede resultaten te komen.

Inhoud

Samenvatting	4
Algemene Introductie	9
Hoofdstuk 1: Vision	11
1.1 Inleiding	11
1.2 Foto's gemaakt met een mobiele telefoon	12
1.2.1 Inleiding	12
1.2.2 LED licht correctie	15
1.2.3 Type smartphone.....	18
1.2.4 Vangstperiode	19
1.2.5 Verschillende platen.....	20
1.2.6 Toepassing in de praktijk.....	22
1.3 Hyperspectraal imaging	25
1.3.1 Pilot experimenten	25
1.3.2 Experimentele Opzet.....	29
1.3.3 Hyper spectrale Dataset.....	30
1.3.4 Machine Learning Algoritmes	31
1.3.5 Resultaten	33
1.3.6 Conclusie en discussie	39
1.4 Edapholog	40
1.4.1 Inleiding	40
1.4.2 Technische aspecten	40
1.4.3 Terugvangst ratio's	41
1.4.4 Trainen van Edapholog voor automatische beeldherkenning.....	42
1.4.5 Vergelijking Edapholog met spoelmonsters	44
1.4.6 Vergelijking Edapholog trechter met de Berlese-Tullgren opstelling	45
1.4.7 Algemene conclusie en integratie van Edapholog in Weet wat er Leeft app.....	48
1.5 Discussie Vision	49
1.5.1 Mobiele telefoon camera.....	49
1.5.2 Hyperspectraal camera.....	51
1.5.3 Edapholog	51

Hoofdstuk 2: eDNA	52
2.1 Algemene inleiding	52
2.2 Insectendetectie met eDNA analyse	53
2.2.1 eDNA detectie met PCR-analyse.....	53
2.2.2 eDNA detectie met de shotgun methode	55
2.3 Discussie en conclusie eDNA.....	57
Hoofdstuk 3: VOC's	59
3.1 Inleiding	59
3.2 Experimenten.....	59
3.2 Resultaten	61
3.3 Conclusie VOC's.....	63
Hoofdstuk 4: Artificiële Intelligentie	64
4.1 Inleiding	64
4.2 Ontwerp.....	64
4.2.1 Inleiding	64
4.2.2 Ontwerp.....	65
4.3 Alternatieve Platformen	68
4.4 Mobiele APP	69
4.5 Integratie.....	71
4.6. Experiment met HAS	72
4.6.1 Optimalisatie van gewasscouting	72
4.7 Discussie AI	75
Aanbevelingen:	75
Hoofdstuk 5: Algemene conclusie en discussie	77
Literatuurlijst.....	80
Bijlagen.....	81

Algemene Introductie

De basis van Integrated Pest Management (IPM) is monitoren. Door tijdig schadelijke organismen waar te nemen kan het uitzetten van biologische bestrijders zorgen voor een adequate bestrijding. Voor een teler is het daarnaast belangrijk continu geïnformeerd te zijn over de populatieomvang van biologische bestrijders in zijn of haar teelt.

Het monitoren van schadelijke organismen en biologische bestrijders gebeurt veelal door het scouten van planten en vangplaten of door het verzamelen van bladmateriaal voor een spoelmonster. Wat betreft de plakplaten zijn er verschillende systemen op de markt voor een kasteler, waarmee, doormiddel van beeldherkenning, automatisch soorten worden gedetecteerd en geteld. De teler krijg hiermee in een digitaal portaal real-time informatie over de aanwezigheid van de soorten die door het systeem worden herkend en kan hierop acteren. De huidige systemen zijn echter nog niet optimaal en integreren nog geen verschillende sensoren.

Een stap verder is om monitoringssystemen te ontwikkelen en toe te voegen aan bovenstaande platform die plagen al kunnen detecteren nog voordat ze op een plakplaat worden gedetecteerd. Twee van dit soort systemen zijn detectie doormiddel van eDNA (environmental DNA) en detectie doormiddel van VOC's (Volatile Organic Compounds), of geurstoffen. Beide technologieën zijn veelbelovend wat betreft hun mogelijke toepassingen voor monitoren van organismen, maar hebben zich nog niet bewezen in een praktijk situatie (Poyntz-Wright et al., 2024; MacDougall et al., 2022). eDNA wordt vooral gebruikt voor ecologisch onderzoek en VOC's veelal voor fundamenteel onderzoek in gecontroleerde omgevingen in laboratoria. Een derde technologie is de Edapholog waarmee o.a. mijten automatisch kunnen worden gedetecteerd en geteld (Dombos et al., 2016). De Edapholog is ontwikkeld voor het analyseren van grondmonsters en zou mogelijk ook voor bladmonsters gebruikt kunnen worden.

De ambitie van het project Weet wat er Leeft is om een platform te creëren waar de verschillende datastromen uit een kas samenkomen waardoor 1) de identificatie van soorten word geoptimaliseerd; 2) de aanwezigheid van soorten over ruimte en tijd beter wordt begrepen en 3) patronen via deep-learning automatisch worden herkend.

Deze ambitie is vervolgens opgedeeld in negen specifieke projectdoelen:

1. Het precies kunnen detecteren van drie plaagorganismen: Californische trips, groene perzikluiz en kaswittevlug.
2. Het precies kunnen detecteren van de drie biologische bestrijders: *Aphidius colemani*, *Amblyseius swirskii* en *Transeius montdorensis*.
3. Inzicht in de haalbaarheid van de technologieën: vision, eDNA en VOC voor detectie van de doelorganismen.
4. Het ontwikkelen van API's voor communicatie van detectiemethodieken, omgevingssensoren (klimaat) en andere hardware met een kunstmatige intelligentie platform.
5. Het ontwikkelen van een deep learning kunstmatige Intelligentieplatform voor het optimaliseren van de (gevoeligheid van) detectie.
6. Het automatisch detecteren, identificeren en kwantificeren van verschillende soorten in een mixpopulatie door integratie van state-of-the-art technieken middels artificial intelligence;
7. Een integrale methode die geschikt is om te worden toegepast op een mixpopulatie van de drie doelplagen en biologische bestrijders;
8. Het toepassen van een Kunstmatige Intelligentie systeem met detectiemethoden en klimaatsensoren op praktijkbedrijven in de gewassen: roos, chrysant, komkommer en potorchidee.
9. Een eerste inzicht in de dynamiek van populaties in ruimte en tijd via big dataverwerking.

Dit rapport geeft een weergave van de realisatie van de projectdoelen. Hoofdstukken 1-3 gaan in op het toepassen van respectievelijk de technieken vision, eDNA en VOC voor het automatische monitoring van soorten. In Hoofdstuk 4 wordt besproken hoe de verschillende datastromen kunnen worden samengebracht in één platform en over de ontwikkeling van dat platform. Hoofdstuk 5 geeft een weergave in hoeverre de projectdoelen zijn behaald en een discussie over het project als geheel met aanbevelingen voor de toekomst.

Hoofdstuk 1: Vision

1.1 Inleiding

Van de drie technieken die in Weet wat er Leeft worden behandeld is Vision de meest gevorderde voor het automatisch monitoren van soorten in een kas. Het principe is eenvoudig. Er wordt een camera gericht om een object en dat doormiddel van automatische beeldherkenning wordt geclassificeerd als een individu van een bepaalde soort of als onbekend. De kwaliteit van de foto's en het aantal foto's waarmee het algoritme is getraind bepalen het succes van de techniek. Bij aanvang van het project waren er geen bestaande systemen die alle zes doelsoorten automatisch konden monitoren en is besloten hier zelf een systeem voor te ontwikkelen.

De vliegende doelinsecten Californische trips, kaswittevlies, groene perzikluus en de parasitaire wesp *Aphidius colemani* zijn doormiddel van gele lijmplaten gemonitord. Vervolgens is gekeken of de soorten op deze plaat konden worden herkend op afbeeldingen gemaakt met de camera van mobiele telefoons en doormiddel van hyperspectraal camera's. Een hyperspectraal camera kan veel meer verschillende kleuren licht waarnemen dan een standaard camera, waaronder voor mensen onzichtbare golflengtes.

Er is onderzocht hoe de beeldherkenningsalgoritmes presteren bij verschillende aantallen soorten doelorganismen op de plakplaten; bij het gebruik van verschillende type smartphones; bij verschillende lichtcondities in de kas; bij verschillende duur dat platen hangen voordat ze gefotografeerd worden en wat de invloed is van het type plakplaat dat wordt gebruikt op de prestatie. De trainingsfoto's voor de algoritmes moet zoveel mogelijk overeenkomen met de foto's zoals die door gebruikers wordt gemaakt. De omstandigheden binnen en tussen kassen zijn echter variabel en kunnen van grote invloed zijn op de kwaliteit van beeldherkenningsalgoritmes. Er is binnen het project daarom juist gefocust op de praktische aspecten van automatische beeldherkenning en zijn er praktische oplossingen bedacht om met deze variatie om te gaan.

De kruipende doelorganismen *Amblyseius swirskii* en *Transeius montdorensis* zijn gemonitord met behulp van de Edapholog. Met dit apparaat kunnen kleine organismen van bladmateriaal worden verzameld doormiddel van een Berlese-Tullgren opstelling en een macrolens dat deze vervolgens filmt. Een Berlese-Tullgren opstelling maakt gebruik van het principe dat arthropoden naar beneden kruipen als ze onder een felle warme lamp worden geplaatst. Vervolgens komen ze via een trechter in een reservoir waar ze worden gefilmd.

Er is onderzocht in hoeverre de Edapholog een aanvulling of vervanging kan zijn van het nemen van spoelmonsters voor het monitoren van mijten in de kas.

1.2 Foto's gemaakt met een mobiele telefoon

1.2.1 Inleiding

Uit een korte enquête onder telers van de betrokken gewascorporaties, kwam een voorkeur voor fotografie van plakplaten met een mobiele telefoon naar voren. De hoge mobiliteit en lage kosten van deze methode zijn een voordeel, maar beeldherkenning met behulp van een telefoon en plakplaten kent ook praktische uitdagingen.

Factoren zoals het aantal soorten en individuen op een plakplaat, lichtcondities in de kas en verschillende typen mobiele telefoons zorgen voor variaties in foto's waar de algoritmes mee om moeten kunnen gaan. Enerzijds kan dit door deze variaties mee te nemen in de trainingsdata, anderzijds door aanpassingen tijdens de fotografie en bewerking van de foto's. In het project zijn de effecten van deze factoren en eventuele oplossingen voor knelpunten in het herkennen van de insecten op plakplaten onderzocht.

Omdat plakplaten met name geschikt zijn voor de monitoring van vliegende insecten is de aandacht gericht op de vliegende soorten: kaswittevlies, Californische trips, gevleugelde groene perzikluiskruiper en de parasitaire wesp *Aphidius colemani*. Hierbij zijn zowel foto's gemaakt van insecten uit labkweek van SCFF en Koppert als van insecten op plakplaten in praktijksituaties.

De resultaten van de beeldherkenningsalgoritmes worden uitgedrukt met de parameters 'recall' en 'precision'. De recall is het vermogen van het algoritme om de doelinsecten op de foto's te vinden. De precision is het vermogen van het algoritme om de gevonden doelinsecten tot de juiste groep in te delen. Verder wordt soms ook de 'accuracy' en 'F-score' gegeven, die een algemener beeld geven van de prestaties van het algoritme, waarbij een hogere waarde een betere prestatie betekent.

Diversiteit insecten op platen

Om te testen of insectendiversiteit op de platen van invloed is op de prestaties van het beeldherkenningsalgoritme, zijn gele insectenplaten met droge lijm (Horiver Dry, Koppert) met verschillende combinaties van insecten, afkomstig van SCFF en Koppert, besmet.

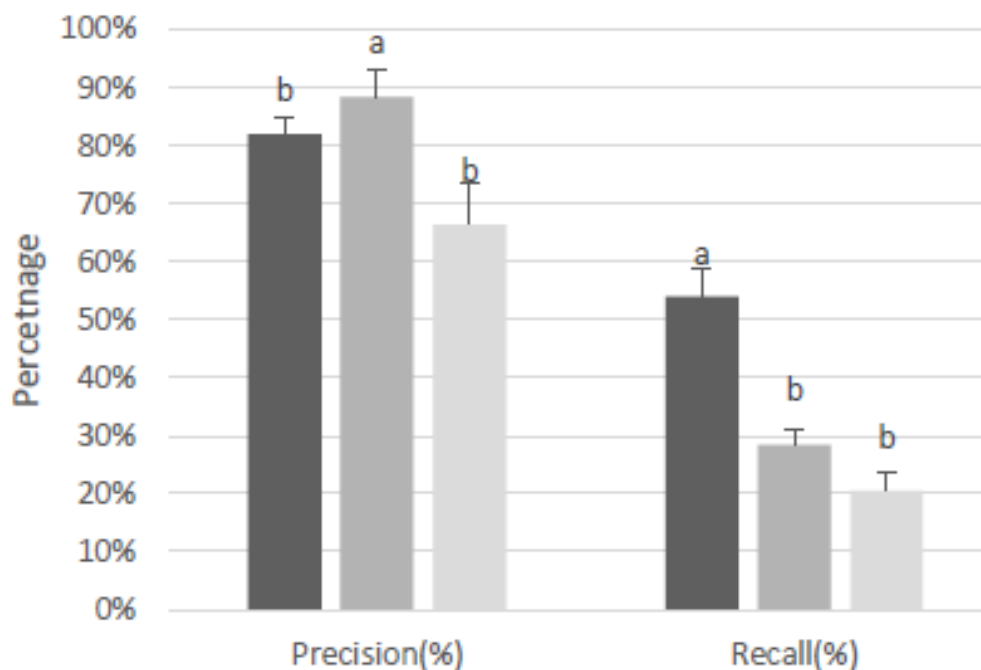
Elke plaat werd twee maal gefotografeerd met een mobiele telefoon (Oppo 52, 12 megapixels) onder kunstlicht. Hierbij zijn van één zijde de onderste en de bovenste helft gefotografeerd. De foto's zijn vervolgens geïmporteerd in Azure CustomVision en gelabeld als groene perzikluiskruiper, Californische trips, kaswittevlies en *Aphidius*. 90% van de foto's is gebruikt om een model te trainen; de overige foto's zijn gebruikt ter validatie van het model. Om het model te testen zijn plakplaten besmet met zowel individuele, als verschillende combinaties van meerdere soorten (Tabel 1.1).

Tabel 1.1: Overzicht van combinaties van soorten voor de testset.

Combinatie	Soorten	Aantal foto's
Individueel	Californische trips	7
	Groene perzikluis	9
	Kaswittevlieg	7
Mix doelsoorten	Californische trips & groene perzikluis	3
	Californische trips & kaswittevlieg	12
	Kaswittevlieg & groene perzikluis	6
	Californische trips, kaswittevlieg & groene perzikluis	11
Mix doelsoorten en niet-doelsoorten	Californische trips, groene perzikluis, kaswittevlieg & varenrouwmug	7
	Californische trips, kaswittevlieg & varenrouwmug	2

Precision en recall

Over het algemeen wordt de precision niet significant slechter als er meer soorten aanwezig zijn op een plakplaat. De recall daalt echter wel bij het toevoegen van zowel doel- als niet doelsoorten (Figuur 1.1).



Figuur 1.1: Gemiddelde precision en recall van alle doelsoorten indien deze als enkele soort op een plaat aanwezig waren (donkergrijs), in combinatie met andere doelsoorten (grijs) en in combinatie met zowel doelsoorten als niet-doelsoorten (lichtgrijs).

Onder deze omstandigheden worden de insecten met name op platen waar meerdere soorten (en meer individuen) aanwezig zijn, niet goed teruggevonden, maar wel goed gedetermineerd. Door gedetailleerdere foto's te nemen, het aantal soorten per foto en de dichtheid te verlagen, kunnen de prestaties van het model verhoogd worden (van der Spek et al., 2022; Lisi et al. 2022).

In het vervolg van het Weet wat er leeft is dit opgelost door niet hele foto's door het model te halen, maar foto's van plakplaten op te knippen in kleine fragmenten van 860x860 pixels en ieder fragment afzonderlijk te laten analyseren. Om dit te bewerkstelligen is een nieuw model getraind door Hevaci et al. (2022).

Lichtcondities

Variatie in lichtcondities vragen om een robuustere trainingsset dan wanneer foto's onder gecontroleerde, vergelijkbare omstandigheden genomen worden. Verschillen in lichtcondities binnen een kas en tussen kassen kunnen de performance van modellen beïnvloeden. Eén van de oplossingen voor dit probleem is het fotograferen van plakplaten op verschillende momenten op de dag en in het jaar in kassen. Een andere oplossing is het standaardiseren van de lichtcondities door middel van een externe lichtbron. Er zijn verschillende manieren getest om met deze variatie om te gaan. Foto's zijn gemaakt met en zonder externe lichtbron en foto's zijn door middel van nabewerking aangepast.

Flits of omgevingslicht

Om de variabele lichtcondities teniet te doen, is een externe flits ontwikkeld waar de mobiele telefoon op bevestigd werd (Figuur 1.2). De flits bestond uit een paneel van 22x12 cm met LED-lampen (12V, 70W, 7000LM, 6500K), dat was gemonteerd op een aluminium frame, voorzien van een handvat met lichtschakelaar. Door de flits onder een hoek te plaatsen werd reflectie op de plakplaten voorkomen.



Figuur 1.2: De externe flits waarop een mobiele telefoon onder een hoek bevestigd werd. De achterzijde is links, de voorzijde is rechts getoond.

Uit de resultaten van verschillende deelonderzoeken (Helvaci et al., 2022; Nijhuis et al. 2023; Verkleij et al. 2023) blijkt dat er bij bladluis, trips, wittevlies in praktijksituaties en *Aphidius* uit een kweek geen significant verschil gevonden werd in de precision, recall en F-score van de gebruikte modellen (CustomVision) bij omgevingslicht in de kas en het gebruik van een flits. Alleen bij de wittevlies uit het lab werd een verschil gevonden, waarbij de flits het beter deed (Verkleij, 2024).

Echter, als plakplaten van dezelfde locaties door de tijd gevolgd werden (Helvaci et al., 2022; van den Einde et al., 2023), bleek dat het model constanter presteerde als wel een flits gebruikt werd, waardoor betrouwbaardere trends in populatiedichtheden te behalen zijn met het gebruik van een externe flits.

1.2.2 LED licht correctie

Het aanpakken van de uitdagingen die ontstaan door verschillende lichtomstandigheden in kassen, waaronder het gebruik van kunstlicht en monochromatische LED's (paarse kleur) is essentieel voor juist herkenning van insecten als deze belichting aanstaat. Deze variaties kunnen aanzienlijke verschillen veroorzaken in de RGB-beelden, wat verdere analyse kan beperken. Om dit probleem aan te pakken, is een gespecialiseerd netwerk ontwikkeld om de kleur van beelden te normaliseren tegen een witlichtstandaard, waardoor uniformiteit wordt gegarandeerd over beelden van verschillende kassen.

Belang van LAB-kleurruimte

Bij het gebruik van bichromatische LED's, die vaak een kleurcomponent elimineren (meestal groen), is de LAB-kleurruimte bijzonder nuttig. Deze ruimte heeft twee kleurkanalen die variëren van blauw tot geel en rood tot groen, wat helpt bij het omgaan met uitdagingen zoals kleurenblindheid. Bovendien verandert het luminantiekanal alleen de helderheid van pixels zonder de kleur te beïnvloeden.

Modelarchitectuur

Een aangepast neurale netwerk, de U-Net-architectuur, is gekozen voor de taak om de afbeeldingen op deze manier te analyseren. Dit neurale netwerk helpt vooral om in foto's te bepalen welke delen belangrijk zijn en is gekenmerkt door een ondiepe structuur en een verbeterde skip-verbinding die de invoer direct verbindt met de uitvoerlaag. Deze aanpassing is cruciaal voor het behouden van een hoge resolutie-uitvoer, waardoor het netwerk gedetailleerde informatie kan behouden die essentieel is voor nauwkeurige insectreconstructie.

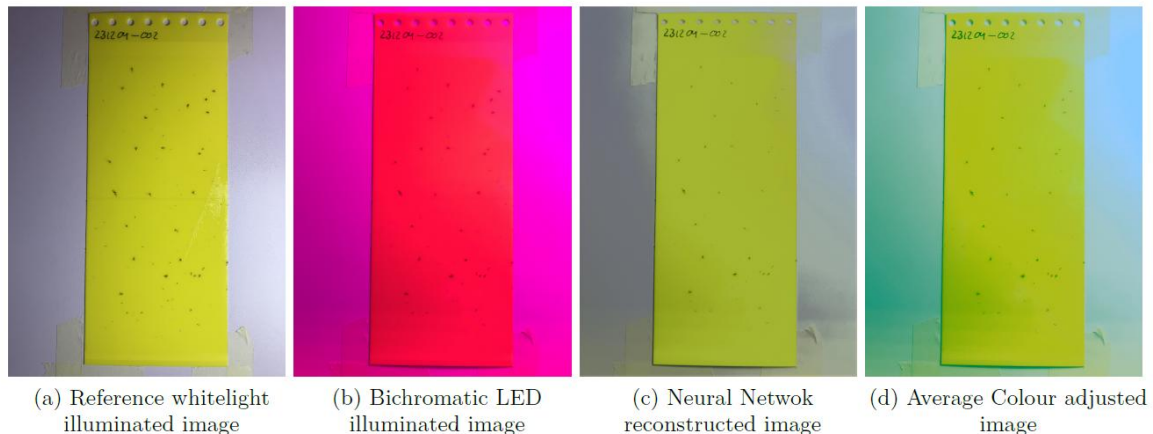
Gemiddelde Verlichtingscorrectie

Naast het gebruik van neurale netwerken is een beeldverwerkingstechniek ontwikkeld die het gemiddelde verlichtingsverschil tussen het referentiebeeld onder wit licht en het beeld verlicht door bichromatische LED's berekent. Dit proces omvat ook de conversie van beelden naar de LAB-kleurruimte. Na deze stap wordt een verlichtingscorrectie toegepast

op de beelden, wat resulteert in een betere beeldkwaliteit in vergelijking met neurale netwerken, hoewel met een compromis in kleurprecisie.

Resultaten

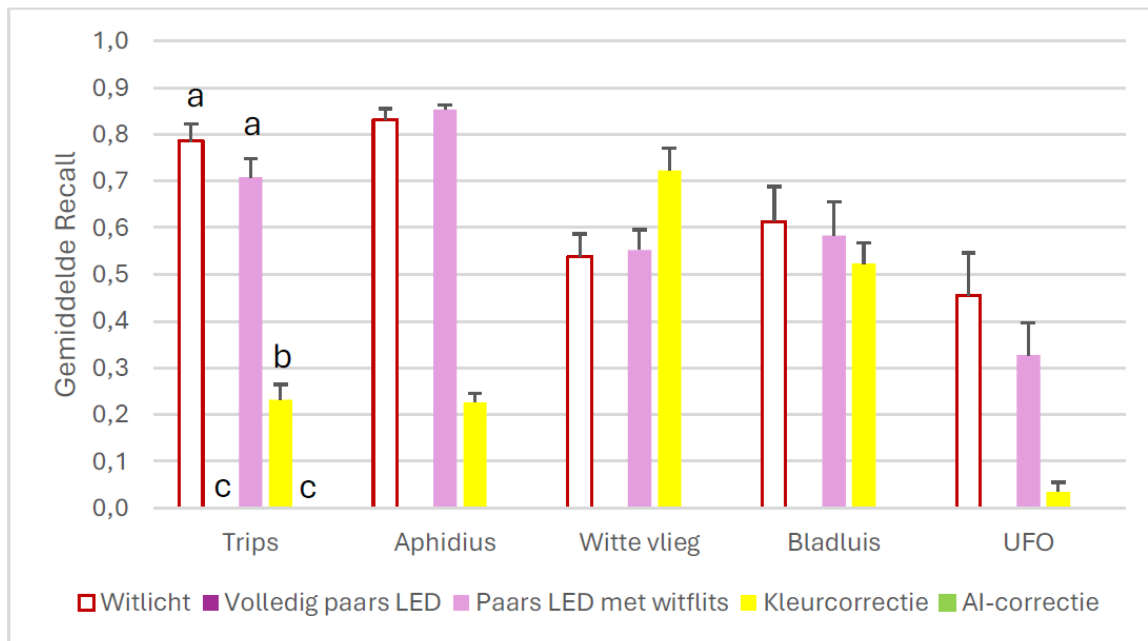
De resultaten van de verlichtingscorrectie worden getoond in Figuur 1.3, waarbij de eerste en tweede foto zijn gemaakt onder normaal licht en led licht, in de derde en vierde foto corrigeren we de LED verlichting doormiddel van een Neural Network (UNET) en met Average Colour adjustment (Gemiddelde Verlichtingscorrectie).



Figuur 1.3 Licht correcties van Bichromatic LED aan de hand van een Neural Network (UNET) en Average Colour Adjustment (Gemiddelde Verlichtingscorrectie)

Invloed van LED lichtcorrectie op herkenning

Om de invloed van de lichtcorrectie op de herkenning onderzoeken zijn de volgende stappen genomen. De besmetting van de plakplaten vond plaats bij het hoofdkantoor van 'Stichting Control in Food & Flowers'. De insecten werden gekweekt in insectententen en vervolgens gevangen in plastic bakjes. De insecten werden met een zuigexhauster over 20 gele lijmplaten verdeeld. De lijmplaten werden in vijf batches van vier platen verdeeld, elk besmet met verschillende combinaties van insectensoorten. De plakplaten werden gefotografeerd in een verduisterde ruimte met een Apple iPhone 7. Er werden drie verschillende foto's gemaakt: met witlicht, onder paarse LED verlichting, en onder paarse LED verlichting met witflits. De foto's werden ingevoerd in de AIAssist Annotatie Tool. De insecten werden handmatig geannoteerd en de annotaties werden gevalideerd met een model dat getraind is op natuurlijk licht en flitslicht. De foto's die onder volledig LED licht zijn gemaakt, werden geaugmenteerd met kleurcorrectie (Gemiddelde Verlichtingscorrectie) en AI-correctie (UNET). De prestaties van de modellen werden geëvalueerd op basis van F1 waarden.



Figuur 1.4: De resultaten van verschillende manieren in de kast om LED licht te corrigeren vergelijken met de resultaten op witlicht

Figuur 1.4 toont de gemiddelde F-waarde van vijf verschillende lichtomstandigheden per doelinsect. De F1 waarde (verder uitgelegd in Sectie 3.2) is een maat voor de algemene prestatie van het model, gebaseerd op zowel de Recall (hoeveel van de elementen goed gezien zijn en welke onterecht niet) als de Precision (hoe goed de gevonden insecten zijn gedetermineerd). De lichtomstandigheden die in figuur 1.4 worden vergeleken zijn witlicht, volledig paars LED licht, paars LED licht met witflits, kleurcorrectie en AI-correctie. De resultaten laten zien dat de F1-waarde bij volledig paars LED licht en AI-correctie 0 is, wat betekent dat er geen insecten werden herkend onder deze omstandigheden. Witlicht en paars LED licht met witflits presteerden het beste, met een F1-waarde boven de 50%. Kleurcorrectie had een lagere F1-waarde dan witlicht en paars LED licht met witflits, maar was nog steeds beter dan volledig paars LED licht en AI-correctie.

Conclusie lichtcondities fotografie

De conclusie van dit onderzoek is dat een witte flits onder een schuine hoek, een goede correctie geeft voor als de foto onder LED licht moeten worden genomen. Ook kleurcorrectie kan dit probleem zonder gebruik te maken van hardware enigszins tegengaan. De AI-correctie gaf achteraf na inzoomen op de beelden vreemde pixel patronen, waarbij we het vermoeden hebben dat die een slechte uitwerking hebben op de herkenning, meer onderzoek is nodig om dit te verbeteren omdat de kleur correctie voor het oog beter lijkt.

1.2.3 Type smartphone

Om het effect van het type smartphone op modelprestaties te testen zijn 9 modellen getraind met foto's van 3 verschillende smartphones: iPhone 7, iPhone 13 en Samsung Galaxy S52 (Tabel 1.2), met en zonder flits. Zowel plakplaten met insecten uit het lab, als plakplaten vanuit de rozenteelt zijn gebruikt bij het verzamelen van de trainings- en testdata.

Tabel 1.2: Smartphones en bijbehorende besturingssystemen en cameraspecificaties.

Smartphone	Besturingssysteem	Mexapixel	Pixelmaat (µm)	Diafragma	Brandpuntafstand (mm)	ISO
iPhone 7	IOS 15.5	12	1,25	f/1.8	4	20
iPhone 13	IOS 15.4.1	12	1,7	f/1.6	5	50
Samsung Galaxy A52	Android 12	64	0,8	f/1.8	5	25

Foto's zijn met de basis fotografie-app gemaakt, tenzij Natutec Scout is gebruikt. Omdat het aantal megapixels verschilt tussen de camera's zijn de foto's van de iPhones eerst tot 64 megapixels vergroot, voordat ze zijn opgeknipt en gelabeld in Custom Vision. Omdat vanuit de praktijklocatie andere insecten aanwezig waren dan de doelsoorten, is het label 'UFO' toegevoegd voor niet-doelsoort geleedpotigen. Overig materiaal op de plakplaten is niet gelabeld.

Vervolgens zijn modellen getraind, waarbij één model is getraind per telefoon met of zonder flits. Vervolgens is een model getraind met foto's van de drie telefoons, zowel met als zonder flits, en een model waarbij alle telefoons en lichtcondities gebruikt zijn. In totaal leverde dit 9 verschillende modellen op, die getest zijn met foto's van alle smartphones en de twee lichtcondities.

Van de modellen die op afzonderlijke telefoons zijn getraind, presteert het model van de iPhone 7 het best (Tabel 1.3). Deze herkent ook insecten op foto's van de andere smartphones goed. De iPhone 13 presteert, ondanks dat het een moderner model is, minder goed, en ook slechter dan de Samsung. De modellen die getraind zijn op foto's met flits presteren beter dan modellen zonder flits. Het model dat op alle type foto's getraind is, presteert het beste (Helvacı et al., 2022). Dat model is in een vervolgproject verbeterd, door de balans in het aantal individuen per doelsoort en praktijk- en labdata te verbeteren. Dit heeft geleid tot een model dat iets minder goed presteerde bij foto's van platen met lab insecten, maar beter presteerde onder praktijkomstandigheden (Janssen et al., 2023).

Tabel 1.3: Nauwkeurigheid van de verschillende modellen bij foto's van verschillende telefoons en lichtcondities. In de kolomnamen staat het model, in de linker kolom is aangegeven onder welke omstandigheden de foto's gemaakt zijn. De prestaties van modellen waarvan de trainings- en testdata onder dezelfde omstandigheden zijn verzameld, zijn geel gemarkeerd. Prestaties boven de 60% zijn in groen weergegeven, terwijl prestaties onder 60% rood zijn.

Foto	Testprestatie Custom Vision								
	Accuracy van het getrainde model						totaal flits	totaal omgeving	totaal flits & omgeving
	A52 flits	A52 omgeving	ip13 flits	ip13 omgeving	ip7 flits	ip7 omgeving			
A52 flits	75.5%	64.6%	32.3%	45.7%	73.9%	61.8%	76.3%	73.5%	76.0%
A52 omgeving	63.8%	72.9%	31.6%	49.4%	70.9%	59.3%	66.3%	71.0%	72.6%
ip13 flits	60.2%	39.8%	24.7%	61.7%	73.1%	47.3%	65.2%	64.1%	70.3%
ip13 omgeving	54.3%	37.0%	21.3%	51.0%	68.9%	69.0%	65.2%	77.5%	75.5%
ip7 flits	68.2%	61.7%	44.2%	76.4%	87.3%	77.3%	79.0%	75.2%	81.8%
ip7 omgeving	62.3%	59.8%	43.3%	76.0%	74.9%	80.5%	75.3%	73.2%	76.8%

1.2.4 Vangstperiode

Tijdens de onderzoeken werd duidelijk dat er verschillen waren in de frequentie van scouten en perioden tussen scoutsessies. Dit is relevant voor beeldherkenning, aangezien dit als gevolg dat niet alle insecten even lang op de platen aanwezig zijn op het moment van fotograferen. Hierdoor treedt variatie op in de data, omdat de uitwendige kenmerken van insecten kunnen veranderen door indroging en beschadiging. Om dat te testen zijn gele Horiver dry platen na verschillende vangstperioden gefotografeerd en geannoteerd. Vervolgens is bepaald wat de recall en precision was en of deze verschilden tussen de verschillende vangstperioden.

Dit is getest voor trips in de chrysantenteelt met perioden van 1, 2 en 4 weken (Van der Einde et al. 2023) en voor trips en wittevlies in de kas van de HAS met perioden van 1, 2 en 3 weken (Verkleij et al. 2023). Bladluizen en *Aphidius* werden niet gevonden op de plakplaten. In beide studies werd aangetoond dat de herkenning van trips verminderde naarmate de periode dat de insecten op de platen aanwezig waren vorderde: in de eerste studie liep de F-score af van 65,8% na 1 week tot 60,9% en 46,9% na respectievelijk 2 en 4 weken. Bij de tweede studie varieerde dat van 61,5% na 1 week tot 71,2% en 49,3% na respectievelijk 2 en 3 weken. In beide gevallen nam de precision af, maar bleef de recall gelijk. Bij wittevlies in de HAS kas werd geen verschil gevonden in modelprestaties, het model bleef vergelijkbaar presteren met een F-score tussen 96,2% na 3 weken en 98,0% na 1 week.

Hieruit kan geconcludeerd worden dat veranderingen in de modelprestaties over tijd soortafhankelijk zijn, waarbij de herkenning van trips achteruit gaat, maar van wittevlies gelijk blijft. Het advies is om minimaal wekelijks te fotograferen.

1.2.5 Verschillende platen

Bij de voorgaande onderzoeken is gebruik gemaakt van Horiver Dry gele plakvallen. In de praktijk worden echter ook platen van andere kleuren en natte lijm gebruikt. De invloed van deze factoren op het getrainde model is getest, door naast gele droge platen, ook blauwe droge, blauwe natte, gele natte en witte droge platen te besmetten met doelinsecten uit kweken. De verwachting was dat met name de blauwe achtergrond en natte lijm een grote invloed zouden hebben, omdat het contrast tussen donkere insecten en de donkerblauwe achtergrond minder groot is dan bij gele platen. De natte lijm zorgt voor veel reflectie bij het fotograferen, wat ervoor kan zorgen dat insecten niet goed herkend worden, of reflecties als insecten herkend gaan worden. Er zijn 10 platen van ieder type besmet met de verschillende doelsoorten.

Zoals verwacht heeft de kleur van de plaat een sterke invloed op de prestaties van het model, maar verschilt dit per soort. Bij trips, bladluis en *Aphidius* werden geen significante verschillen gevonden in precision, recall en F-score tussen de drie kleuren met droge lijm, hoewel foto's van gele platen wel het hoogst scoorden. Wittevlies werd op witte platen bijna niet herkend.

Het effect van natte lijm leek over het algemeen groter dan de andere kleuren. Wittevlies werd zowel op gele als blauwe platen met natte lijm vrijwel niet herkend (prestatieparameters tussen 0 en 7%). Trips werd op blauwe platen met natte lijm helemaal niet herkend en zowel bladluis als *Aphidius* werden op natte blauwe platen minder goed herkend dan op gele natte platen.

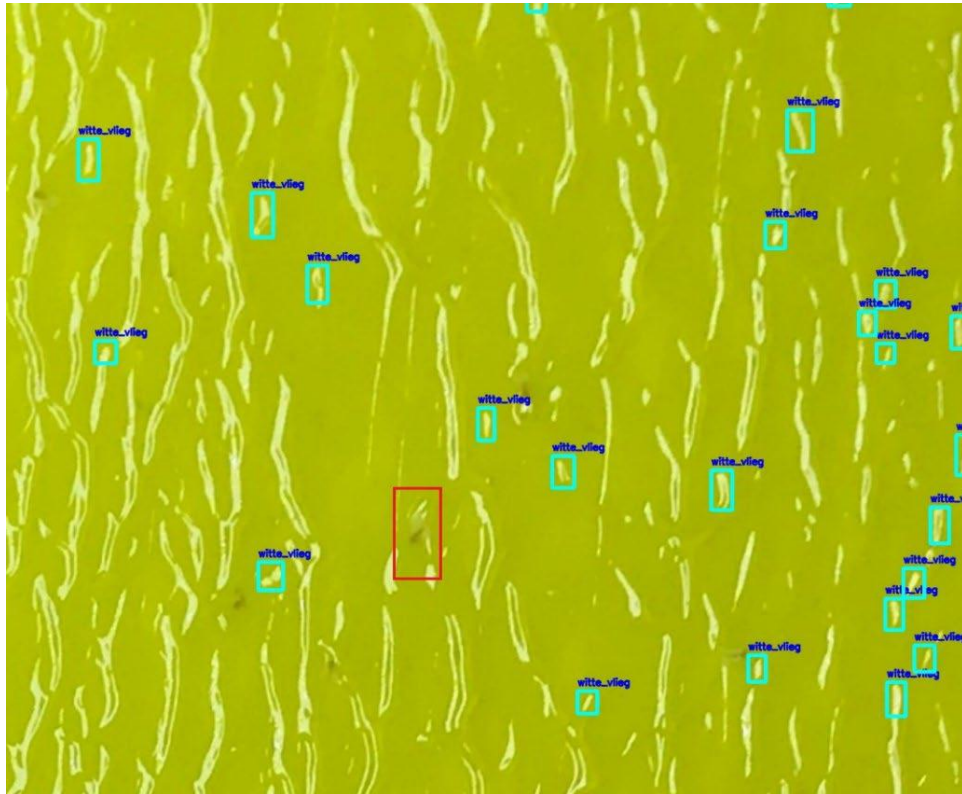
De natte lijm veroorzaakt veel reflecties, met name met de flits, die als wittevlies worden herkend (Figuur 1.5). Bovendien zakken insecten dieper in de lijm bij droge lijm, waardoor ze minder goed herkenbaar zijn. Vooral wittevlies verkleurt na verloop van tijd, waardoor ze met het blote oog al minder goed herkenbaar zijn.

Om te testen of de prestaties verbeterd kunnen worden zijn modellen specifiek op kleur en lijmtypen getraind. Hierbij is gebruik gemaakt van het YOLOv8-algoritme, dat met weinig trainingsdata getraind kan worden. 80% van de data is gebruikt voor de training en 20% voor validatie. Er zijn zowel modellen getraind per plakplaat, als een model waarbij foto's van de verschillende typen platen samen zijn gevoegd.

In tegenstelling tot bladluis en *Aphidius* werden trips en wittevlies na training op gele platen met natte lijm nog steeds niet goed gevonden en herkend. Op blauwe platen werd wittevlies wel herkend (F-measure van 69%).

De modellen die getraind waren op droge lijm lieten bij de verschillende kleuren F-scores zien van boven de 50% voor alle soorten, behalve wittevlies op witte platen, trips op blauwe platen en trips op gele platen. Door een model te trainen met alle kleuren, werden de prestaties van het model op foto's met verschillende kleuren over het algemeen licht verhoogd.

Bij droge lijm is het mogelijk om de prestaties te verbeteren door modellen met de verschillende kleuren te trainen, zelfs met een kleine trainingsdataset. Door de dataset uit te breiden is verdere verbetering mogelijk. Bij natte lijm worden trips en wittevlies op gele platen en trips op blauwe platen niet goed herkend na training. Bij deze soorten is het niet waarschijnlijk dat verdere training voor veel verbetering gaat zorgen.



Figuur 1.5: Afbeelding van een gele plakplaat met natte lijm en weerspiegeling van de flits. De blauwe vakjes zijn de boxes die door het model zijn voorgesteld, met het label in blauwe tekst. Het rode vierkant geeft een daadwerkelijke wittevlies weer.

Conclusie vangplaten

Kortom, bij de droge lijm wordt wittevlies slecht herkend op witte platen, wat te verklaren valt door de vergelijkbare kleur tussen de plaat en de insecten. Op de andere kleuren werd wittevlies beter herkend. Trips, de sluipwesp en bladluis werden op de verschillende kleuren op een vergelijkbaar niveau herkend, waarbij de prestaties het best waren bij gele plakplaten.

Natte lijm heeft een grotere negatieve impact. Wittevlies is niet herkend op blauwe en gele platen, trips alleen op geel en de sluipwesp en bladluis werden beter op geel herkend dan op blauw. Voor de soorten die wel herkend worden, moeten nog wel modellen getraind worden met uitgebreidere trainingsdata om het praktisch inzetbaar te maken.

1.2.6 Toepassing in de praktijk

De data van het herkenningsmodel is getraind op foto's van plakplaten met de doelinsecten vanuit kweken, die zowel met als zonder externe flits zijn gefotografeerd in de kas van HAS green academy en foto's van plakplaten met zowel de doelinsecten als niet doelinsecten die zijn gemaakt in een kas van de rozenteelt (Helvacı et al., 2022). Dit model is opnieuw gebalanceerd, waarbij er meer praktijkfoto's zijn gebruikt en foto's van insecten die veel aanwezig waren verwijderd zijn, terwijl foto's van insecten die minder aanwezig waren in de trainset zijn toegevoegd (Janssen et al. 2023). Foto's die onder dezelfde omstandigheden als de trainingsfoto's werden gemaakt leverden werkbare resultaten.

Bij verschillende deelonderzoeken werd echter een verschil gevonden tussen foto's onder gecontroleerde omstandigheden en de praktijk. De gegevens van drie deelprojecten in de rozen- en chrysantenteelt zijn gecombineerd en lieten zien dat de performance bij de wittevlieg, trips en bladluis significant lager was in de praktijk dan onder gecontroleerde omstandigheden (Verkleij, 2024).

Dit betekent niet dat model niet bruikbaar is: van der Eijnde et al. (2023, studentenonderzoek) toonden aan dat er een sterk significant verband bestond tussen het aantal daadwerkelijk aanwezige tripsen en de door het model aangewezen tripsen bij chrysant. Met F-scores van circa 60%, werden positieve significante correlaties gevonden met een R^2 van 85% met flits en 84% zonder. Dit betekent dat de trends in populatiedichtheden van trips goed te volgen waren met het model. Hierbij nam de R^2 af naarmate de platen langer bleven hangen, wat het eerdere advies om de platen minimaal wekelijks te fotograferen onderbouwd. In dit onderzoek is ook de Natutec Scout app gebruikt. Hierbij was het niet mogelijk de modelprestaties te bepalen, maar wel de correlatie tussen het aantal werkelijk aanwezige tripsen en het aantal door het model aangewezen aantal. Hier werden vergelijkbare waarden gevonden. Deze bleven redelijk constant, zelfs als de platen na vier weken hangen gefotografeerd werden.

In de laatste praktijkstudie werd de toepassing getest in kassen van de vier verschillende teelten (chrysant, komkommer, roos, potorchidee). Hierbij zijn nieuwe modellen getraind met exact dezelfde data als in de eerdere projecten, met YOLOv8 (HAS) en Faster R-CNN (Imec), die in de HASKas net zulke goede resultaten lieten zien als het eerdere model.

De resultaten in de praktijk waren echter zeer variabel en inconsistent. Ook correlaties tussen het werkelijk aantal aanwezige insecten op de insectenplaten en door de modellen aangegeven aantallen, waren slechts beperkt aanwezig (Houtmans et al. 2024). Hier zijn verschillende oorzaken voor aan te wijzen:

- Het aantal doelinsecten in de kassen was erg laag, waardoor een laag aantal foute detecties een grote impact heeft op de performance van de modellen.
- Daarnaast waren verschillende doelinsecten aanwezig, waar het model niet op getraind is. Dit werd het duidelijkst geïllustreerd in de potorchideeënteelt, waar veel

Atheta corriaria aanwezig was op de plakplaten. Aangezien deze niet aanwezig waren in de trainingsdata, werd deze niet goed herkend als niet-doelsoort, maar als trips. Dit gold zowel voor het YOLO-model als het model van Imec.

- In de trainingsset is de achtergrond (de ruimte achter de plakplaat die op de foto aanwezig was) verwijderd. Bij de foto's uit de praktijk is dat niet gebeurd, waarbij elementen buiten de plakplaat als insecten werden herkend.

Tijdens dit project zijn drie weken lang 12 plakplaten per teelt opgehangen en gefotografeerd, zowel met als zonder flits. Deze foto's zijn allemaal geannoteerd, en het model van Imec is daar op getraind. Dit gaf betere resultaten in herkenning, maar de invloed op de correlatie met populatieveranderingen per soort is helaas niet meer getest. Hierbij wordt geadviseerd de trainingsdata die gebruikt is voor het huidige model aan te vullen met de data uit het hierboven beschreven project en selectief subfoto's uit te zoeken waarin insecten die zeldzamer zijn meer worden meegenomen (balanceren van insectsoorten) om een nieuw model te trainen. Hierbij kunnen ook moeilijke achtergronden worden toegevoegd aan de trainingsdata.

Conclusie praktijktoepassing

Ondanks dat in meerdere deelprojecten goede resultaten behaald werden met verschillende beeldherkenningsmodellen en in de kas van HAS, bleek de performance van het model in het laatste deelproject onvoldoende te werken in de praktijk. Hiervoor zijn meerdere verklaringen. Er waren weinig doelinsecten aanwezig, waardoor een foute detectie een grote impact had op het resultaat, er waren insecten aanwezig die niet in de trainset aanwezig waren en verkeerd werden geïdentificeerd en de achtergrond was nog aanwezig op de foto, wat niet het geval was bij de trainingsset. Omdat de foto's uit de vier verschillende kassen gelabeld zijn, kunnen deze gebruikt worden voor het bijtrainen van het model. De verwachting is dat de performance daardoor significant zal verbeteren en bruikbaar is in de praktijk. Zeker aangezien bij trips is aangetoond dat het mogelijk is om trends te bepalen met een herkenning van slechts 60%.

Bij het trainen van het model is ervoor gekozen om niet te labelen met de soortnamen, maar hogere groepen (trips, bladluis, wittevlug en sluipwesp). Dit is gedaan omdat het labelen op soortniveau een valse verwachting geeft van het niveau waarop het model insecten kan onderscheiden. Omdat sterk gelijkende soorten van de doelinsecten niet in de trainingset aanwezig zijn geweest, zouden sterk gelijkende soorten een verkeerd label krijgen (dat van de doelsoorten) als wel op soortniveau was getraind. Dat betekent ook dat het met deze techniek met deze toepassing niet mogelijk is om de insecten tot op soortniveau te identificeren. Gezien het formaat van deze insecten en de zichtbare details op de foto's, is de verwachting dat de verschillende soorten trips, wittevlug en bladluis niet op soort te onderscheiden zijn, zonder aanpassingen aan de methodiek, zoals bijvoorbeeld het gebruik van een macro-lens op de camera van de telefoon. Deze techniek is daarmee vooral geschikt voor het snel tellen en monitoren van deze groepen van insecten. Voor een

nauwkeurige bepaling zou, bij verfijning van de techniek, DNA analyse of handmatige determinatie plaats kunnen vinden.

Bij de verschillende onderzoeken is aangetoond dat verschillende factoren die relevant zijn voor de praktijk van invloed zijn op de performance van de modellen, waarvoor verschillende oplossingen zijn bedacht. Over het algemeen geldt dat:

- Het aantal soorten op de plakplaat had een negatief effect. Dat is opgelost door de foto's op te knippen in fragmenten en deze te analyseren, in plaats van de hele foto. Dit had ook als voordeel dat details van insecten beter zichtbaar waren.
- De lichtcondities in de omgeving zijn van invloed op de performance. Door het gebruik van een externe lichtbron werkte het model alleen bij wittevlies beter dan zonder flits. Door de tijd gezien werkte het model echter consistent indien de lichtbron gebruikt werd. Bovendien zorgde het ervoor dat het effect van paars LED-licht teniet gedaan werd. Dat kon overigens ook bewerkstelligd worden met nabewerking van de foto's.
- Het type telefoon is van invloed op de performance. Er zijn 3 telefoons grondig getest, waarbij de performance erg verschilde tussen de telefoons. Door modellen te trainen met de verschillende telefoons werd de performance ook voor de minst presterende telefoon goed.
- De periode dat plakplaten blijven hangen een negatief effect had, doordat platen voller raken en insecten degraderen. Door wekelijks te fotograferen en platen te vervangen kan dat ondervangen worden.
- Platen met droge lijm geschikter zijn voor automatische beeldherkenning dan met natte lijm. De natte lijm reflecteert het licht van de lichtbron en omgeving, wat herkend werd als insecten. Bovendien verdwenen insecten in de lijmlaag, waardoor ze minder zichtbaar waren of degradeerden. Afhankelijk van de soort waren andere kleuren platen (blauw en wit) ook te gebruiken; met beperkte foto's zijn daar redelijke modellen voor getraind.
- De verschillende algoritmen om modellen mee te trainen gaven vergelijkbare resultaten. Het Custom Vision model, YOLOv8 model en Faster R-CNN model gaven vergelijkbare resultaten in de kas van HAS en in de praktijk. Verschillen die gevonden zijn, komen grotendeels door andere instellingen. Zij zijn alle drie met exact dezelfde trainingsdata getraind. Van Natutec Scout kon de performance niet altijd op een vergelijkbare manier bepaald worden, omdat deze gegevens niet beschikbaar waren, maar als gekeken wordt naar het aantal insecten dat gevonden werd, waren de resultaten vergelijkbaar.

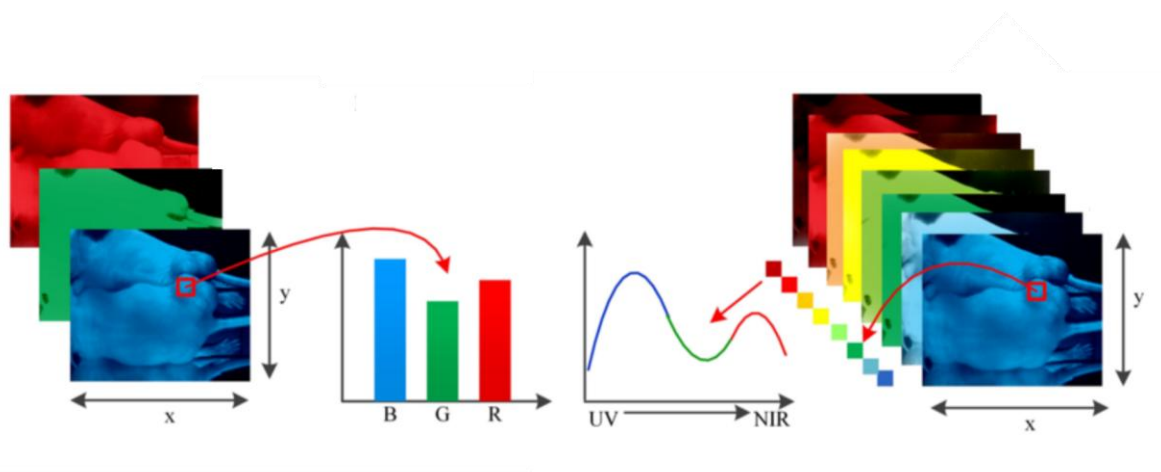
Ondanks dat de resultaten in het praktijkonderzoek in de vier kassen tegenvielen, heeft deze techniek potentie voor het snel en eenvoudig tellen en monitoren van tipsen, wittevliegen, bladluizen en sluipwespen (*Aphidius colemani*). Het model moet daarvoor verder getraind worden met de datasets die beschikbaar zijn. Daarbij wordt geadviseerd de

datasets te balanceren voordat modellen getraind worden. Bij het fotograferen wordt aangeraden dit zoveel mogelijk te standaardiseren, door het gebruik van vergelijkbare plakplaten, een externe lichtbron en minimaal wekelijks te fotograferen.

1.3 Hyperspectraal imaging

Een hyperspectrale camera is een geavanceerd type camera dat licht in veel meer details kan vastleggen dan een gewone camera. Terwijl een gewone camera alleen de kleuren rood, groen en blauw ziet, kan een hyperspectrale camera licht in honderden verschillende kleuren (of spectrale banden) zien. Dit betekent dat het licht dat door een object wordt weerkaatst, wordt opgesplitst in een reeks nauwe spectrale banden, die de camera apart detecteert en verwerkt (Figuur 1.6).

Deze technologie wordt gebruikt om materialen nauwkeurig te identificeren, omdat elk materiaal op een unieke manier met licht interageert. Bijvoorbeeld, in de landbouw kunnen hyperspectrale camera's worden gebruikt om de gezondheid van planten te controleren door te kijken naar de specifieke lichtpatronen die door de bladeren worden weerkaatst.



Figuur 1.6: Links: De kleur van elke beeldpunt wordt bepaald door een combinatie van 3 filters (RGB), Rechts: De kleur van elke beeldpunt wordt bepaald door 150 filters

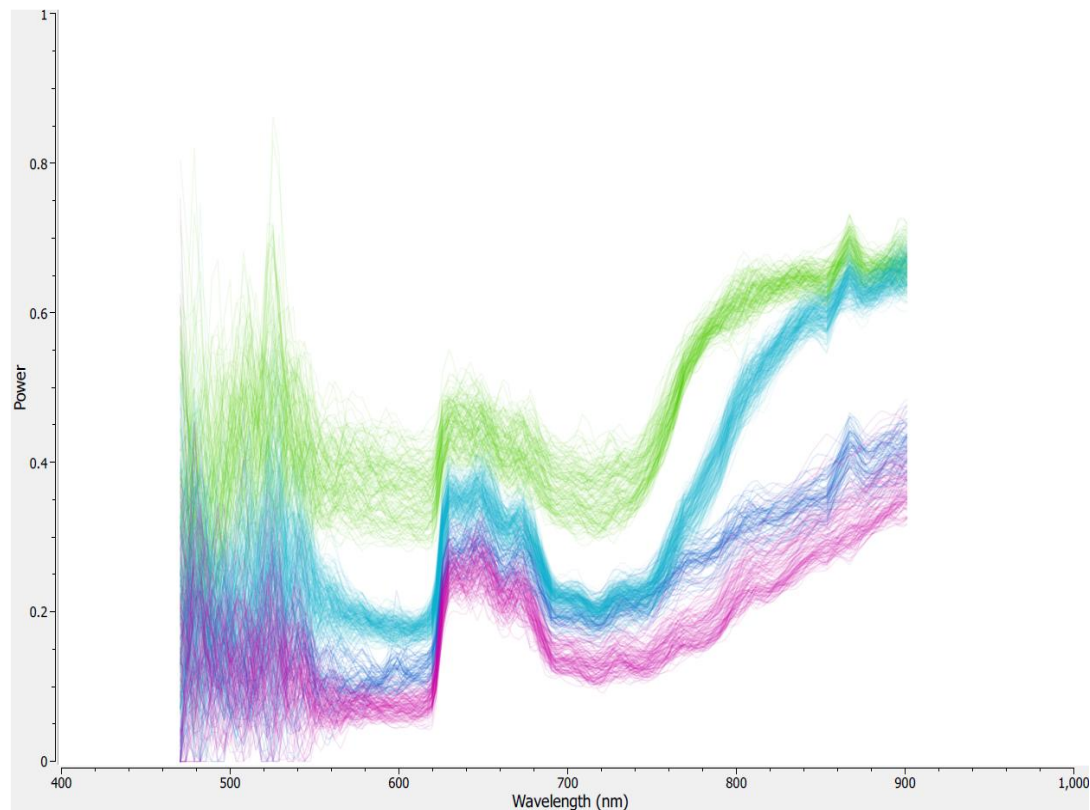
Hyperspectrale beeldvorming begon in de ruimte, waar satellieten zoals de Landsat 1 in 1972 werden gelanceerd om de geologische diversiteit van het aardoppervlak in beeld te brengen. Tegenwoordig worden hyperspectrale camera's ook gebruikt in andere gebieden zoals medische beeldvorming en voedselveiligheid.

1.3.1 Pilot experimenten

VNIR (Visible Near Infrared) en SWIR (Short Wave Infrared) zijn twee verschillende spectrale bereiken die worden gebruikt in hyperspectrale beeldvorming. Het belangrijkste verschil tussen VNIR en SWIR ligt in de golflengte van het licht dat ze detecteren.

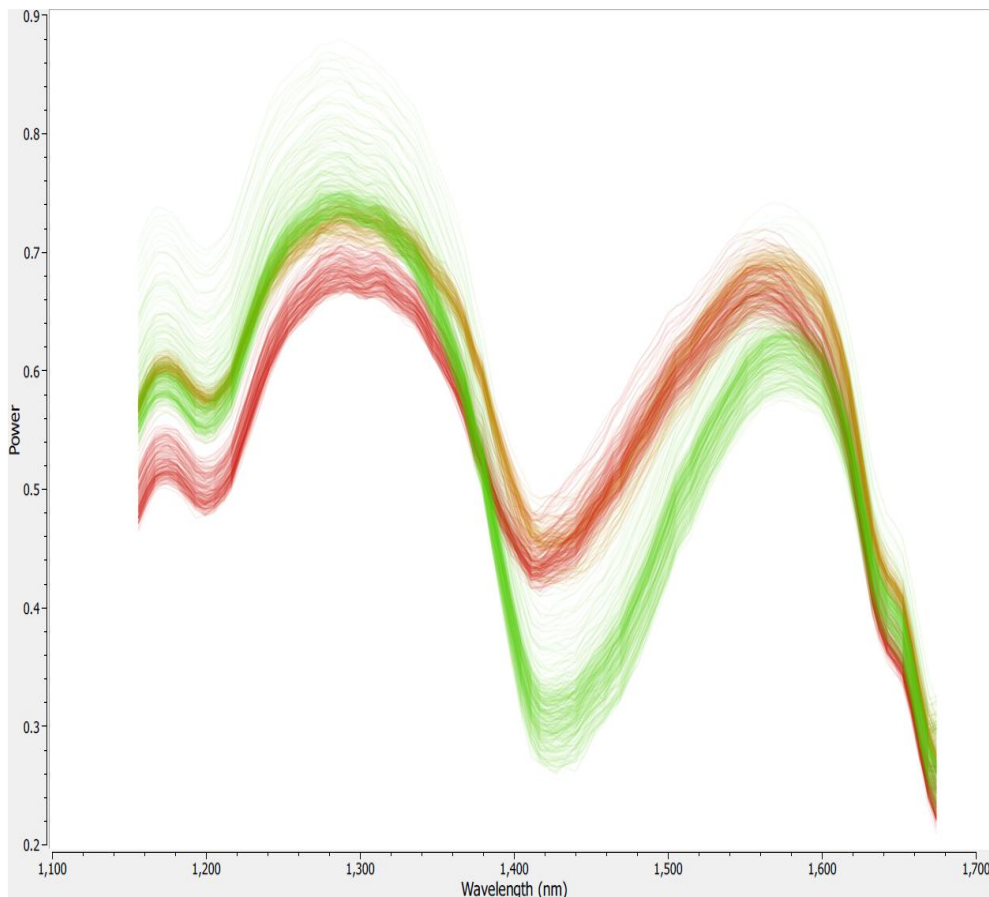
VNIR bestrijkt het golflengtebereik van ongeveer 400 tot 1000 nanometer (nm). Dit omvat zowel zichtbaar licht als nabij-infrarood licht. VNIR-camera's gebruiken meestal CMOS-sensoren, die goedkoper zijn dan de sensoren die nodig zijn voor SWIR-camera's. VNIR is

nuttig voor toepassingen waarbij zichtbaar licht en nabij-infrarood licht voldoende informatie bieden, zoals het detecteren van vegetatiegezondheid en het analyseren van oppervlakken.



Figuur 1.7: De spectraal kenmerken van verschillende insecten: wittevlies (groen), californische trips (donkerblauw), groene perzikluis nimf (lichtblauw) en groene perzikluis adult (paars).

SWIR bestrijkt het golflengtebereik van ongeveer 1000 tot 2500 nm. SWIR-camera's gebruiken vaak InGaAs-sensoren (Indium Gallium Arsenide), die duurder zijn maar noodzakelijk voor het detecteren van deze langere golflengte. SWIR is nuttig voor toepassingen waarbij het belangrijk is om materialen te onderscheiden op basis van hun chemische samenstelling, zoals het detecteren van vochtgehalte, het analyseren van voedselkwaliteit en het identificeren van verschillende materialen in recyclingprocessen.



Figuur 1.8: De spectraal kenmerken van verschillende insecten: wittevlies (oranje), californische trips (rood), groene perzikluis (groen).

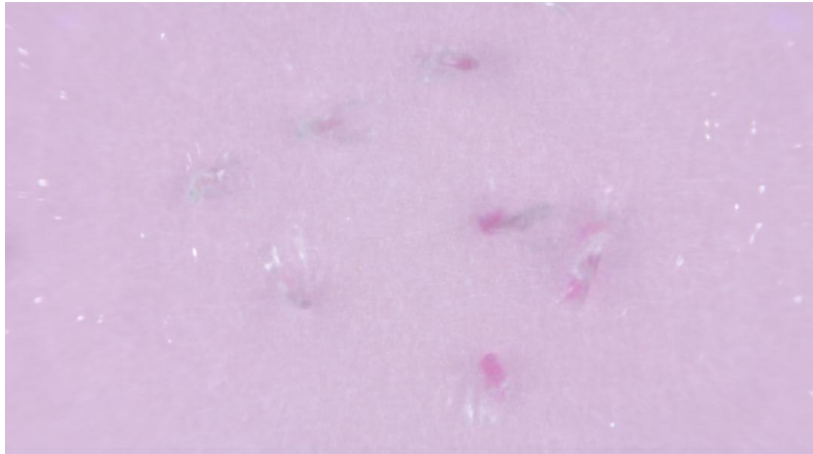
Twee aparte camera's

De reden dat je twee aparte camera's nodig hebt voor VNIR en SWIR is dat de sensoren die nodig zijn om deze verschillende golflengten te detecteren, verschillend zijn. CMOS-sensoren zijn geschikt voor VNIR, terwijl InGaAs-sensoren nodig zijn voor SWIR. Deze sensoren hebben verschillende eigenschappen en kosten, en kunnen niet eenvoudig worden gecombineerd in één enkele camera zonder aanzienlijke technische en financiële uitdagingen.

In sommige gevallen kunnen gecombineerde systemen worden ontwikkeld, maar dit is vaak complex en duur. Daarom wordt er meestal gekozen voor aparte VNIR- en SWIR-camera's om de specifieke voordelen van elk spectraal bereik optimaal te benutten

Welk Spectrale Bereik?

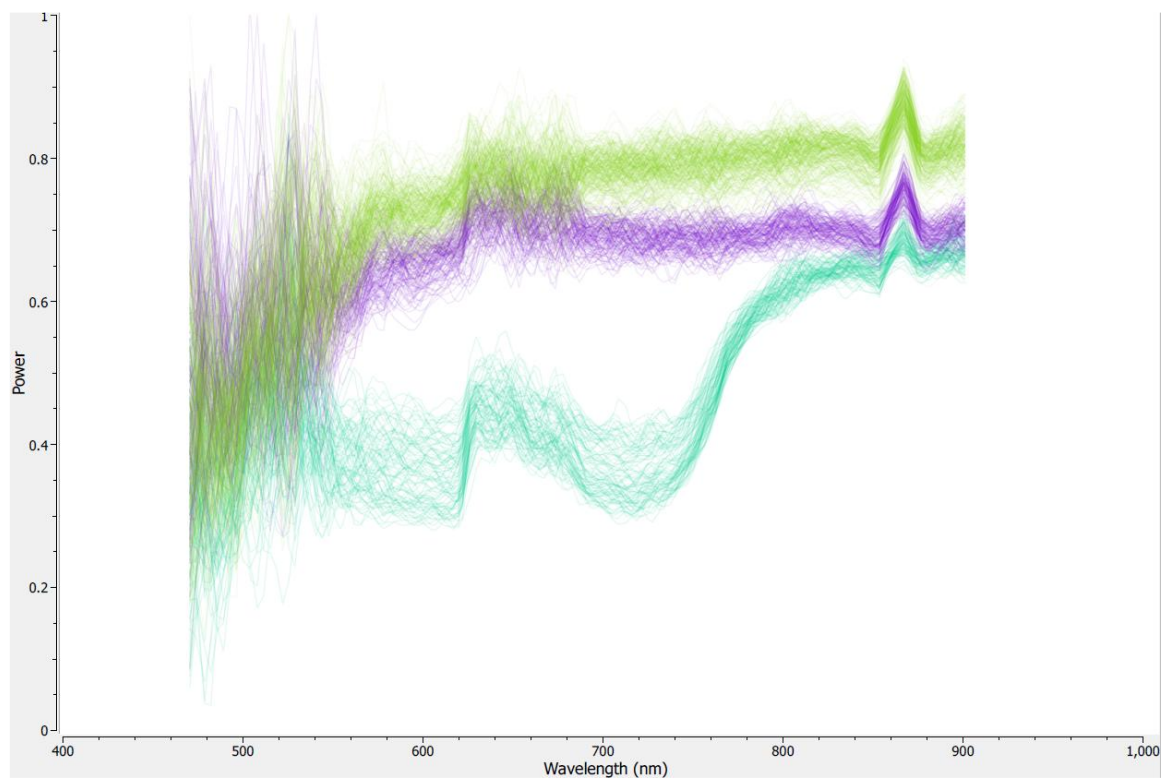
In Figuur 1.7 en 1.8 zijn voor enkele insecten op plakplaten de diverse spectra in één figuur te zien. In Figuur 1.7 zijn duidelijke verschillen in de lijnen te zien terwijl in Figuur 1.8 de lijnen op elkaar liggen voor de verschillende insecten en ze dezelfde trend volgen. Voor het onderscheiden van de verschillende doelinsecten wil je juist de variatie hebben die Figuur 1.7 laat zien, daarom is besloten met VNIR door te gaan.



Figuur 1.9: Natte lijm voorbeelden onder een hyper spectraal camera.

Type plakplaat

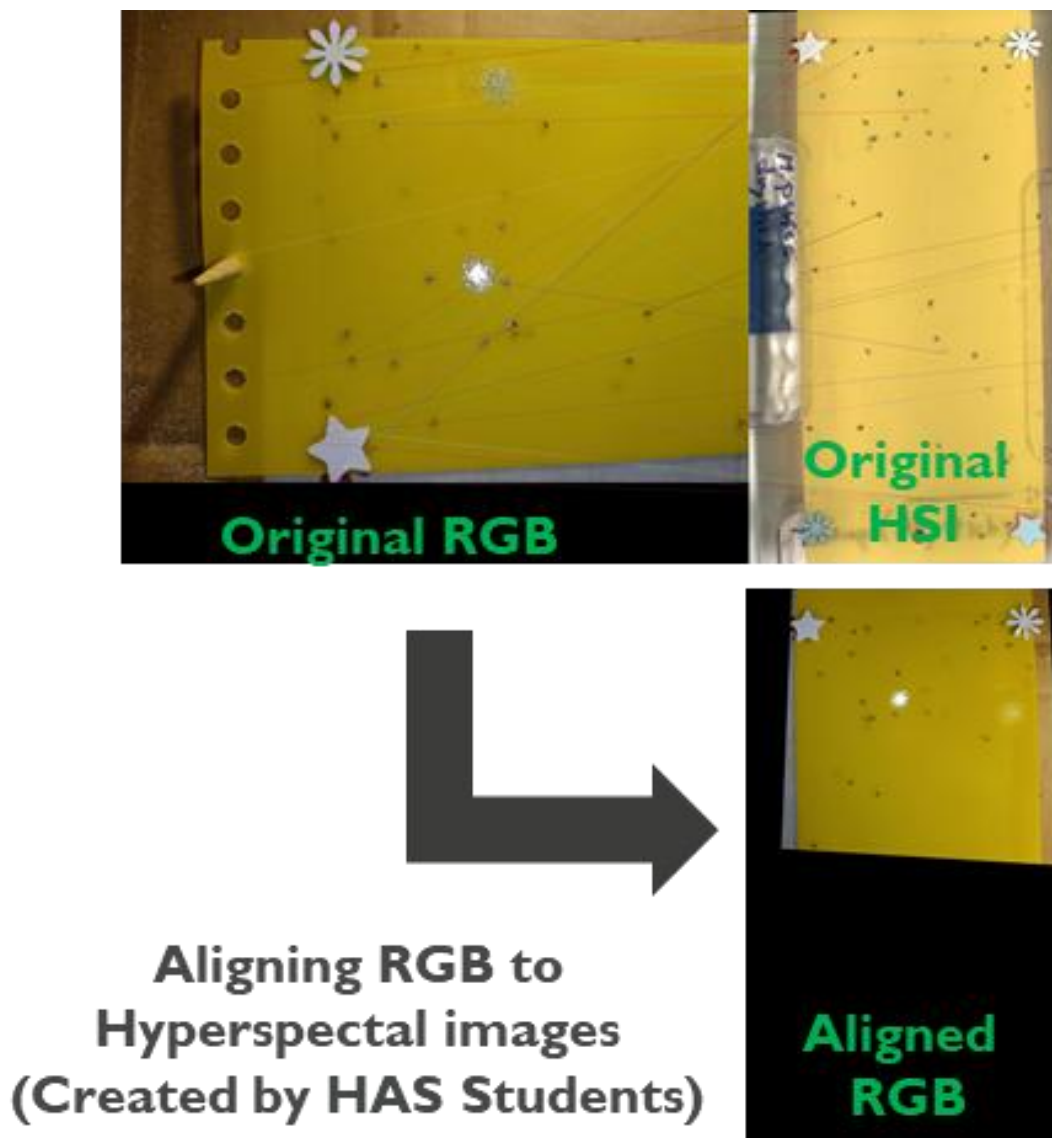
In Figuur 1.9 zijn beelden van een plakplaat met insecten met natte lijm getoond. Hierbij was na enkele dagen al niet zichtbaar welk insect er op had gezeten en ook lichaamsdelen van insecten waren moeilijk te onderscheiden. Hierdoor is in het vervolg onderzoek veelal gekozen voor droge lijm. De kleur van de plakplaten is vooral gekozen, omdat bleek dat geel het beste diverse insecten aantrok. In Figuur 1.10 is duidelijk zichtbaar dat de achtergrond kleur een effect heeft op de spectrale kenmerken en dat het beter is om 1 kleur te kiezen.



Figuur 1.10: Effect van de plakplaat kleur op de spectrale kenmerken van hetzelfde insect (wittevlies): blauwe plakplaten (mintgroen), witte plakplaten (paars) en gele plakplaten (groen).

1.3.2 Experimentele Opzet

In het experiment met de hyperspectrale camera zijn vangplaten gefotografeerd met een variatie van insecten. Van de vangplaten zijn zowel met de hyperspectrale camera foto's gemaakt als met normale camera. Omdat het visualiseren van hyperspectraal foto speciale software vereist, hebben studenten de normale foto geannoteerd en zijn deze annotatie doormiddel van beeldregistratie ook bruikbaar voor de hyperspectrale foto's (Figuur 1.11).



Figuur 1.11: Het registreren van de normale beelden (RGB) op de hyperspectral (HSI) waarbij we voor visualisatie een "false color" beeld hebben gemaakt uit het hyperspectral cube.

In het begin gaf de beeldregistratie problemen omdat de stickers (sterren/bloemen) helaas niet goed genoeg werkten. In later versies is gebruik gemaakt van tekst stickers, zichtbaar in Figuur 1.12.



Figuur 1.12: Unieke aanduiding met sticker op de plakplaat voor registratie en administratie doeleinden.

1.3.3 Hyper spectrale Dataset

Voor dit onderzoek is een unieke dataset van hyperspectrale beelden van insecten op plakplaten gemaakt, waarbij we 264 hyperspectraal foto's gebruiken. De hyperspectraal beelden zijn met de Imec Snapscan VNIR camera gemaakt (Figuur 1.13). De dataset bestaat 132 hyperspectraal beeld van 3600 x 2048 pixels en 150 spectrale banden (tussen 400 tot 1000 nanometer). De dataset bevat zo'n 4142 insecten, waarbij we ons vooral hebben gericht op de plaagsoorten en bestrijders: 305 Californische trips (*Frankliniella occidentalis*), 256 Groene perzikluiskruis (*Myzus persicae*), 278 Kaswittevlies (*Trialeurodes vaporariorum*), 3221 Sluipwesp (*Aphidius colemani*) en 82 Unidentified Flying Objects (UFO) waarbij de insecten niet bij de doelsoorten horen, dit zijn voornamelijk huisvliegen, maar kunnen ook andere insecten zijn. De locatie van het insect op de plakplaat is aangeduid met een rechthoek om het insect heen, deze rechten zijn geannoteerd in CustomVision AI (<https://www.customvision.ai/>), dit is een tool van Microsoft om grote beeld datasets te annoteren. Een probleem bij de annotatie is dat de beelden worden verkleind zodat het annoteren van de toch al kleine insecten moeilijker wordt. In later projecten is dit probleem verholpen door gebruik te maken van Labelbox (<https://labelbox.com/>).



Figuur 1.13: Imec Snapscan VNIR camera met spectral range van 470 tot 900 nm, 150 spectrale banden en 5 Mpx resolutie.

1.3.4 Machine Learning Algoritmes

FasterRCNN

De eerste methode die is gebruikt voor het herkennen van insecten op de plakplaten is automatische beeldherkenning methode, FasterRCNN (Ren et al., 2016). Hierbij kunnen we de eerst lagen in het Convolutional Neural Network van FasterRCNN vervangen zodat deze in plaats van kleuren beelden ook voor multi/hyperspectrale beelden kunnen werken. In de nieuwe methode die we hebben ontwikkelt kan de hoeveelheid spectrale banden dynamisch worden ingesteld. Hierbij moeten wij echter wel het FasterRCNN netwerk opnieuw trainen of een uitgebreide netwerk refinement runnen.

Segmentatie Netwerken

Omdat dit onderzoek nog geen significant betere resultaten uit de hyperspectraal beelden gaf, hebben we in vervolg onderzoek gebruik gemaakt van segmentatiemethoden (andere vorm van automatisch beeldherkenning waar per pixel een label wordt gegeven die daarna kunnen worden samengevoegd) in plaats van object-detection methoden (hierbij wordt niet per pixel gekeken, maar op een grover niveau een vierkant gevonden rond het insect). Automatische beeldherkenning methoden kunnen worden geleerd katten, auto, huizen, maar ook insecten te herkennen waarbij we veel over object herkenning zullen spreken i.p.v. van insect herkenning omdat we het ook vaak over de verschillende insectsoorten hebben. Hierbij hebben we de volgende stappen genomen:

1. **Onbegeleide HSI-contourextractie:** Handmatige annotatie van hyperspectrale beelden is tijdrovend en foutgevoelig. Door gebruik te maken van onbegeleide contourextractie kunnen we automatisch gedetailleerde contouren van insecten genereren zonder menselijke tussenkomst, wat de nauwkeurigheid en efficiëntie van het proces aanzienlijk verbetert. Dit omvat het gebruik van algoritmen voor het genereren van gedetailleerde insectcontouren.

2. **Dimensionaliteitsreductie op basis van golflengte:** Het identificeren van kritieke golflengtes voor effectieve plaagclassificatie door middel van een iteratieve backpropagation-gebaseerde aanpak. Hierbij is een iteratieve techniek ontwikkelt die gebruik maakt van backpropagation in neurale netwerken om de meest informatieve golflengtes te identificeren. Het model wordt getraind op een volledige dataset en vervolgens worden de belangrijkste golflengtes geselecteerd op basis van hun bijdrage aan de modelprestaties. Door de data te reduceren tot de meest informatieve golflengtes, kan het model efficiënter en nauwkeuriger leren. Het verwerken van minder golflengtes vermindert de benodigde rekenkracht en opslagruimte aanzienlijk.
3. **Geadvanceerde HSI-segmentatienetwerken:** Ontwikkeling van modellen die zich richten op ruimtelijk-spectrale kenmerken, wat resulteert in een significante verbetering van de segmentatieprestaties. Er zijn drie methoden met elkaar vergeleken:
 - a) **U-Net** (Ronneberger et al., 2015): Dit is een veelgebruikte architectuur voor semantische segmentatie. Het bestaat uit een encoderpad voor het extraheren van kenmerken en een decoderpad voor nauwkeurige lokalisatie. Deze architectuur is effectief in het vastleggen van fijne details, wat essentieel is voor het segmenteren van insecten.
 - b) **3D U-Net** (Çiçek et al., 2016): Deze variant verbetert de traditionele U-Net door volumetrische convoluties toe te voegen, waardoor zowel ruimtelijke als spectrale kenmerken tegelijkertijd kunnen worden vastgelegd. Dit is vooral nuttig voor de multidimensionale aard van hyperspectrale datasets.
 - c) **DeepLabV3+** (Chen et al., 2018): Deze architectuur maakt gebruik van atrous spatial pyramid pooling en dieptegescheiden convoluties om efficiënt multiscale kenmerken vast te leggen. Het vermindert de rekenkundige eisen terwijl het een hoge segmentatienauwkeurigheid behoudt.
 - d) **SpectralDeepLabV3+**: Dit is een nieuwe variant die speciaal is ontworpen voor hyperspectrale beeldsegmentatie. Het integreert spectrale analysecapaciteiten door een Bidirectional Long Short-Term Memory (BiLSTM) netwerk of een 1-dimensionale U-Net toe te voegen aan een afgevlakte versie van het hyperspectrale beeld (Thireou & Reczko, 2007). Deze aanpak maakt een uitgebreide analyse van de spectrale data mogelijk, wat leidt tot nauwkeurigere segmentatie.

1.3.5 Resultaten

Hoe deze resultaten worden bepaald?

De eerste stap is om de volgende nummers te berekenen, door de gevonden objecten (in ons geval de insecten) met de door de mensen geannoteerde insecten te vergelijken. We gaan er hierbij van uit dat de mensen hierbij geen fouten heeft gemaakt.

1. **True Positives (TP):** Dit zijn de objecten die correct zijn gedetecteerd door het model.
2. **False Positives (FP):** Dit zijn de objecten die ten onrechte zijn gedetecteerd door het model (vals alarm).
3. **False Negatives (FN):** Dit zijn de werkelijke objecten die het model niet heeft gedetecteerd.

Door deze waarden te berekenen, kunnen we de Precision, Recall en F1-score bepalen, wat ons inzicht geeft in de prestaties van het model in termen van nauwkeurigheid en volledigheid bij het detecteren en herkennen van objecten.

Object Recall

Object recall, ook wel gevoeligheid of true positive rate genoemd, meet het vermogen van een model om daadwerkelijk aanwezige objecten correct te identificeren. Het wordt berekend als de verhouding van het aantal correct gedetecteerde objecten (true positives) tot het totale aantal werkelijke objecten (true positives + false negatives). Een hoge recall betekent dat het model weinig werkelijke objecten mist, maar het kan ook betekenen dat het model veel valse positieven heeft.

Object Precision

Object precision, ook wel precisie genoemd, meet het vermogen van een model om alleen de relevante objecten correct te identificeren. Het wordt berekend als de verhouding van het aantal correct gedetecteerde objecten (true positives) tot het totale aantal gedetecteerde objecten (true positives + false positives). Een hoge precision betekent dat het model weinig valse positieven heeft, maar het kan ook betekenen dat het model sommige werkelijke objecten mist.

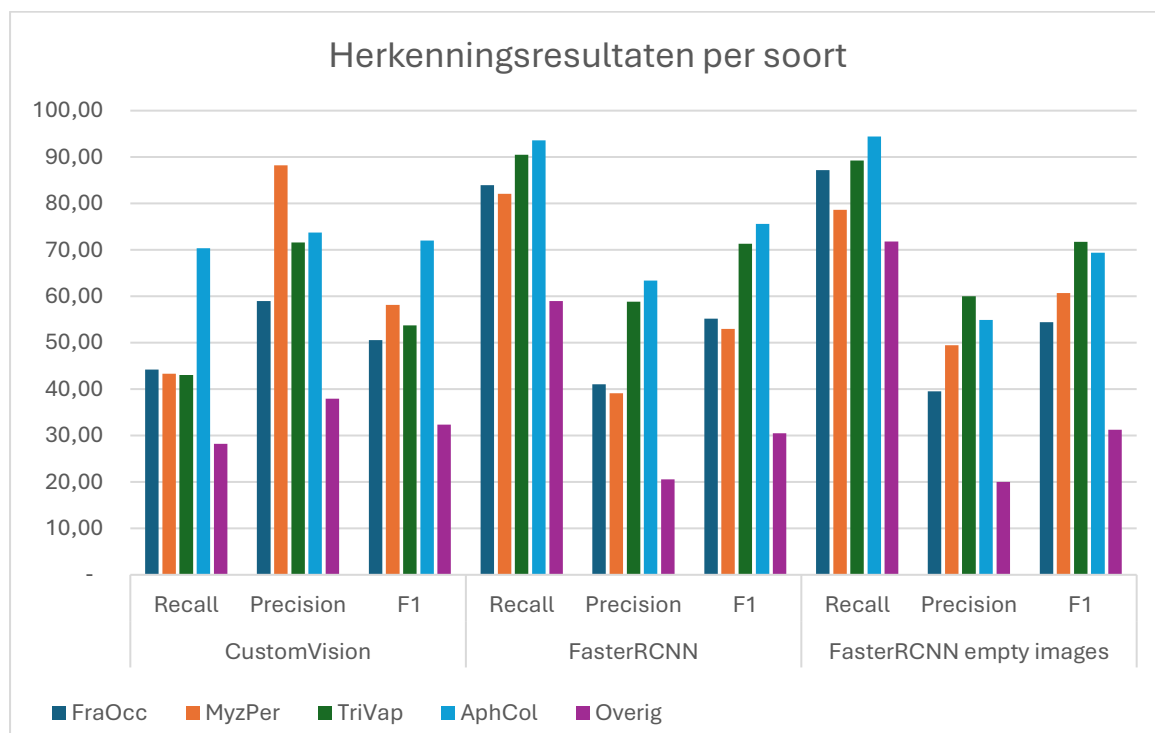
Object F1-Score

De F1-score is de harmonische gemiddelde van precision en recall. Het biedt een enkele maatstaf die zowel de precisie als de recall in overweging neemt. De F1-score is nuttig wanneer er een balans nodig is tussen precision en recall, vooral in situaties waar een van beide significant lager is dan de andere.

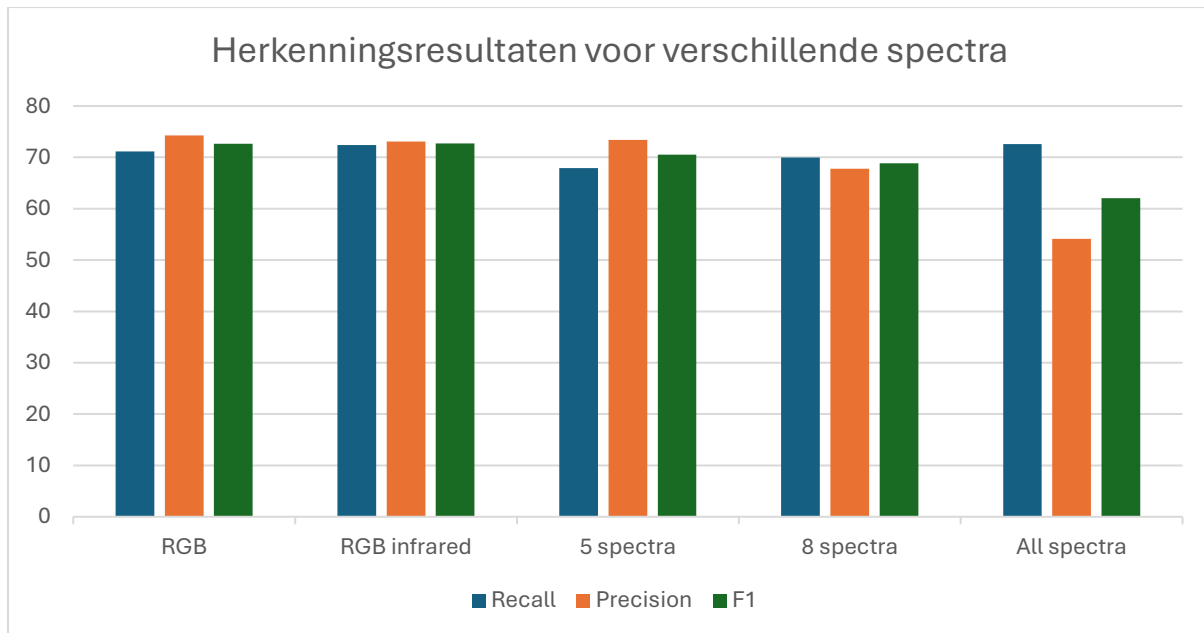
FasterRCNN resultaten

Voor de resultaten van FasterRCNN, hebben we eerst gekeken hoe FasterRCNN zich verhoudt ten opzichten van CustomVision. Resultaten hiervan zijn zichtbaar in Figuur 1.14

en laat duidelijk zien dat FasterRCNN beter is op het gebied van recall dan CustomVision, terwijl de precision van CustomVision beter is. De F1-score van beide methoden is redelijk gelijk, waarbij FasterRCNN met afbeeldingen zonder insecten (empty images) overal het beste lijkt te scoren. Een andere trend is dat als we meer annotaties van een insect hebben, we ook betere resultaten krijgen, waarbij het netwerk ook wat meer problemen heeft met de categorie “Overige”, waar ten eerste niet veel voorbeelden in de dataset zaten en ze daarnaast meerdere andere insecten bevatten. In Figuur 1.15 is de performance zichtbaar van de diverse spectra. We hebben hier gekozen voor RGB, RGB met infrarood (dit is een combinatie die veel beschikbaar is op de markt), daarna hebben we het spectrum van 400-1000 nm verdeeld in 5, 8 en 150 banden. Het trainen van 150 banden is erg duur, vooral omdat het inladen van de afbeeldingen (elk 2GB) te langzaam gaat. Het runnen van het netwerk over een hyperspectraal beeld van 150 eist vooral een computer met veel geheugen. Duidelijk is te zien dat meer banden geen betere resultaten laten zien in precision en F1-score. De 150 banden geven wel de beste resultaten in recall, waarschijnlijk omdat het specifieker kan zoeken. Doordat deze resultaten niet heel sterk waren, zijn we gaan kijken of het aan de FasterRCNN-methode lag of dat meer banden inderdaad niet helpt.



Figuur 1.14: Vergelijking tussen CustomVision and FasterRCNN in twee combinatie (waarbij empty images aangeeft dat ook deel afbeelding die geen insecten bevatten voor de training worden gebruikt, wat deels een oneerlijk vergelijk oplevert met CustomVision omdat die deze beelden negeert. De doelsoorten zijn Californische trips (Fra Occ), Groene perzikluus (Myz Per), Kaswittevlieg (TriVap), 3221 Sluipwesp (AphCol) en Overig.

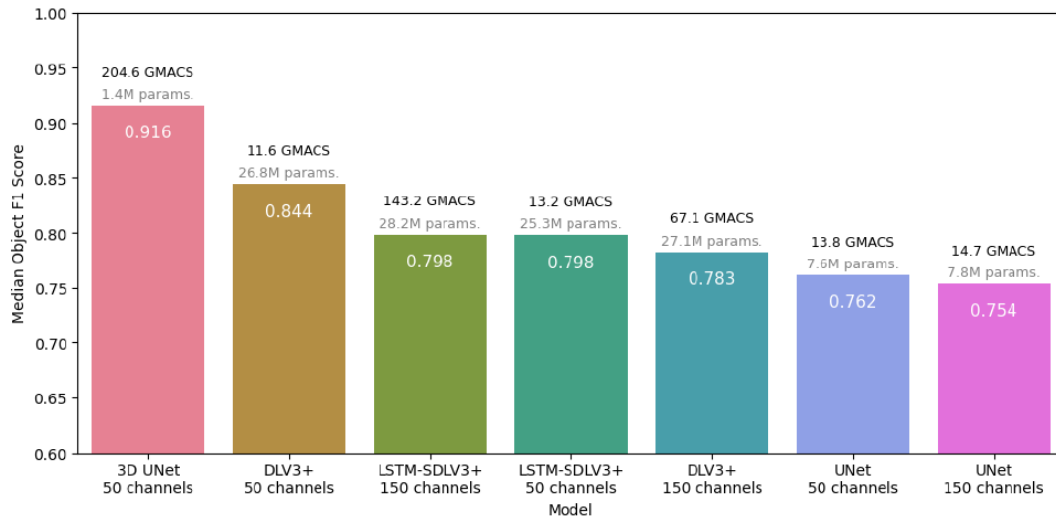


Figuur 1.15: Vergelijking tussen object detectie die gebruik maakt van meerder spectrale band en de normale kleur banden (RGB), waarbij de performance niet perse omhoog gaat als er meerder spectrale banden worden toegevoegd.

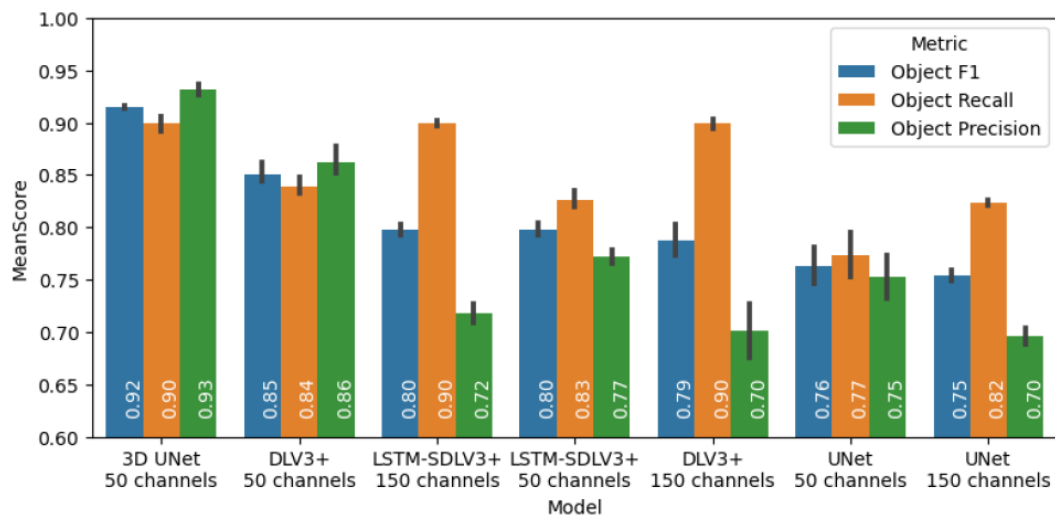
Segmentatie Networks resultaten

Segmentatienetwerken (Convolutional Neural Networks) zijn een alternatief voor object herkenning, met als grote verschil dat ze niet een rechthoek locatie waar het insect zich bevind vinden, maar per pixel kijken of die pixel tot een insect behoort of tot de achtergrond. Onze dataset bestond uit insecten die waren geannoteerd door middel van bounding boxes, door middel van onbegeleide HSI-contoureextractie hebben we pixel gebaseerde annotatie kunnen verkrijgen. Daarna hebben we HSI-segmentatienetwerken gebruikt, de resultaten van diverse methoden zijn ze te zien in Figuur 1.16, waarbij we zo wel gebruik maken van alle 150 spectrale banden en 50 gedownsamptele spectrale banden. Figuur 1.16 (b) laat zien dat 3D UNet het beste werkt met een F1 score van 0.93, maar ook in 1.16 (a) het groot netwerk in parameter en dus ook in berekening is.

DeepLabV3+ (DLV3+) werkt daarna maar heeft ook weer heel veel parameter nodig. We zien wel duidelijk dat de 150 spectrale banden niet perse betere resultaten opleveren, hierbij moet eerlijkheidshalve gezegd worden dat 3D Unet met 150 ook niet kon worden uitgevoerd gegeven onze huidige computers.



(a) Median Object F1 scores, number of parameters and GMACs



(b) Mean Objectivity scores

Figuur 1.16: Resultaten van geavanceerde HSI-segmentation model die gebruik maken van 50 and 150 spectral banden

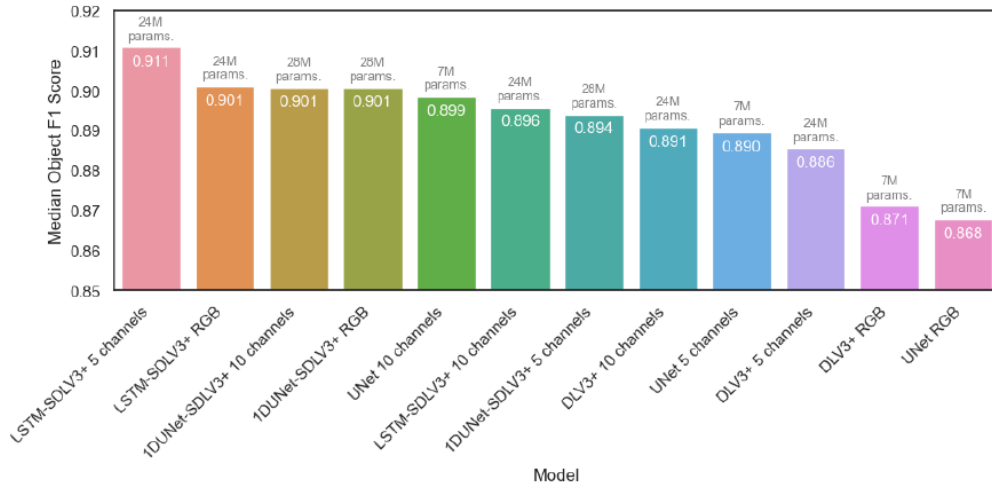
In Figuur 1.17 is duidelijk te zien dat dimensionaliteitsreductie op basis van spectrale banden niet beter presteert dan 3D Unet. De resultaten laten ook zien dat modellen die zijn getraind op datasets met minder kanalen over het algemeen beter presteren dan hun tegenhangers die zijn getraind op volledige bandbreedte datasets. Dit geldt vooral voor de LSTM-SpectralDeepLabV3+ modellen, die consistent betere prestaties laten zien in termen van object F1-scores. Dit suggereert dat het verminderen van de complexiteit van hyperspectrale beelden de prestaties van neurale netwerken voor segmentatietaken verbetert.

Een opvallende observatie is dat modellen met 10 en 5 kanalen vergelijkbare prestaties leveren, hoewel de 10-kanaals versie een lichte voorsprong heeft. Dit impliceert dat het

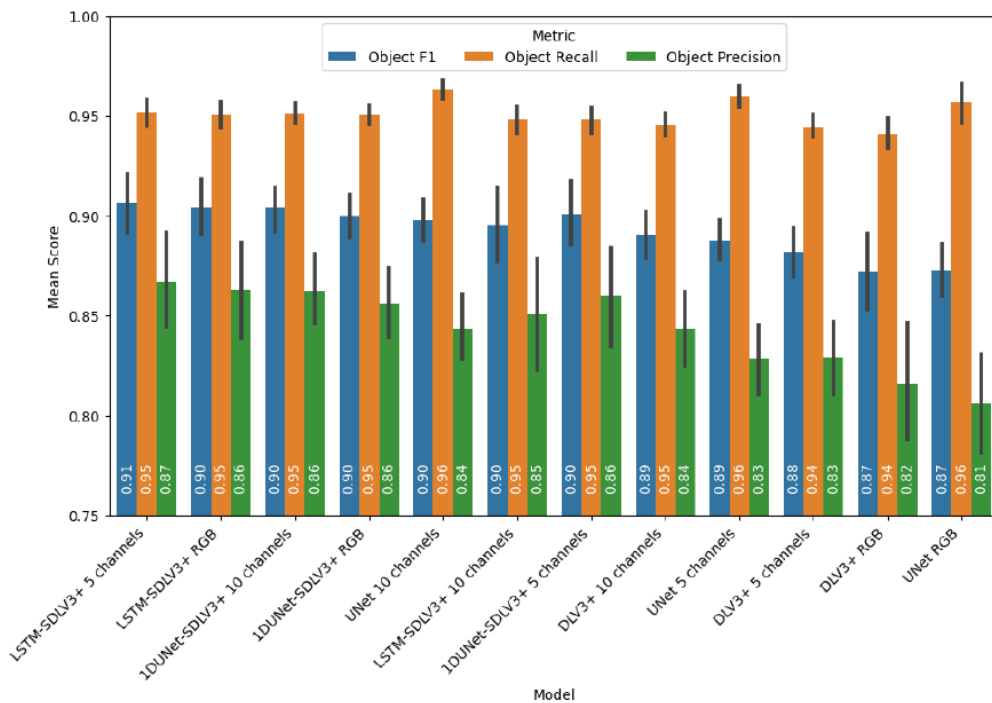
toevoegen van meer kanalen slechts een marginale verbetering biedt, wat wijst op een punt van afnemende meeropbrengst bij het verhogen van de kanaalcomplexiteit voor deze taak.

Daarnaast blijkt dat traditionele U-Net en DeepLabV3+ modellen minder goed presteren met RGB-beelden vanwege hun onvermogen om spectrale kenmerken vast te leggen. Verrassend genoeg presteren onze Spectral-Spatial modellen uitstekend met RGB-beelden, wat suggereert dat deze modellen subtiele kleurvariaties beter kunnen interpreteren, wat cruciaal is voor nauwkeurige segmentatie onder verschillende lichtomstandigheden in kassen.

Kortom, de resultaten benadrukken de voordelen van het verminderen van de complexiteit van hyperspectrale beelden en suggereren dat multispectrale en RGB-gegevens praktische alternatieven kunnen zijn voor hyperspectrale beeldvorming bij insectendetectie. Dit biedt een aantrekkelijke afweging door de operationele complexiteit te vereenvoudigen, kosten te verlagen en toch bevredigende prestaties te leveren voor insectendetectietaken.



(a) Median Object F1 scores and number of parameters



(b) Mean Objectivity scores

Figuur 1.17: Resultaten van Dimensionaliteitsreductie op basis van spectale band in combinatie met Segmentatie methods.

1.3.6 Conclusie en discussie

Spectrale beeldvorming biedt aanzienlijke voordelen voor het herkennen van insecten, vooral in complexe omgevingen zoals kassen. Door gebruik te maken van hyperspectrale beeldvorming kunnen we gedetailleerde spectrale handtekeningen van insecten vastleggen, wat leidt tot een nauwkeurige detectie en classificatie. De **segmentatienetwerken** die in dit onderzoek zijn onderzocht, zoals de geavanceerde HSI-segmentatienetwerken en de methoden voor dimensionaliteitsreductie, hebben aangetoond dat het mogelijk is om insecten 15 tot 20% effectief te identificeren en te onderscheiden van andere objecten en achtergrondruis (dit was nog wel in de experimentele fase en zou met data uit kassen mogelijk moeten worden geverifieerd omdat daar meer doelsoorten en mogelijk problemen zijn). Dit draagt bij aan een verbeterd plaagbeheer en kan de efficiëntie van landbouwtechnologieën aanzienlijk verhogen.

Toekomstig Onderzoek

1. **Integratie van Multispectrale en RGB Beeldvorming:** Hoewel hyperspectrale beeldvorming zeer gedetailleerde informatie biedt, kan het duur en complex zijn. Toekomstig onderzoek zou zich kunnen richten op het combineren van multispectrale en RGB-beeldvorming met hyperspectrale technieken om een kosteneffectieve en praktische oplossing te bieden voor insectendetectie in het veld.
2. **Uitbreiding naar Andere Spectrale Domeinen:** Onderzoek naar het gebruik van ultraviolet (UV) en infrarood (IR) beeldvorming kan nieuwe mogelijkheden openen voor het onderscheiden van insecten op basis van hun unieke spectrale kenmerken in deze golflengte bereiken.
3. **Actief Leren:** Het implementeren van actief leren kan helpen bij het efficiënt beheren van nieuwe gegevensstromen uit het veld. Door zich te concentreren op de meest informatieve monsters, kan actief leren de hoeveelheid benodigde annotaties verminderen en de prestaties van het model verbeteren.

Door deze richtingen te verkennen, kunnen we de mogelijkheden van spectrale beeldvorming verder uitbreiden en bijdragen aan de vooruitgang van technologieën voor nauwkeurige objectdetectie en -herkenning in diverse domeinen.

1.4 Edapholog

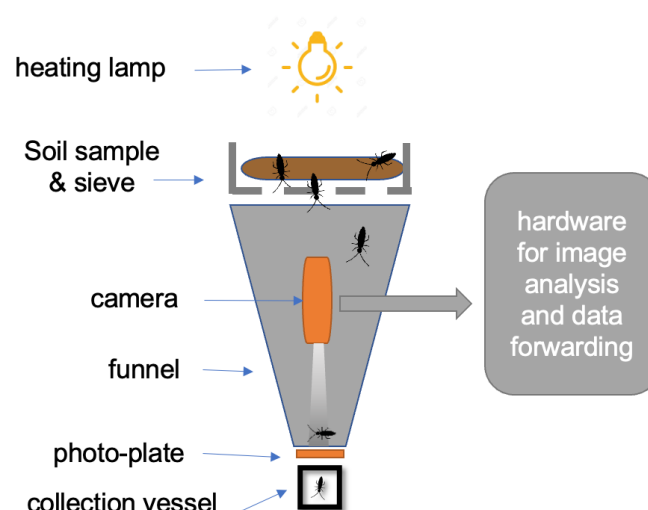
1.4.1 Inleiding

De Edapholog is een sensor die automatisch registreert welke kleine organismen in een bodemonster voorkomen zoals mijten en springstaarten. Edapholog is afgeleid van het Engelse woord Edaphology (Edafologie) wat een tak van bodemkunde is die zich bezighoudt met het bestuderen van de relatie tussen bodems en organismen. Het oorspronkelijke prototype van het apparaat wordt in een bodem geplaatst en meet tegelijkertijd ook bodemparameters (Dombos et al., 2016). De Edapholog 1.0 is ontwikkeld in een samenwerking tussen het Instituut voor Bodemonderzoek in Boedapest en Syngenta die samen de start-up Edapholog Kft hebben opgericht.

In dit onderdeel van het project wordt de Edapholog 1.0 toegepast op het detecteren van mijten en trips in bladmonsters. De Edapholog wordt vergeleken met de nu standaard spoelmethode en Berlese extractie; er wordt onderzocht of de Edapholog praktisch toepasbaar is binnen een kasteelt en het wordt getraind op de doelorganismen (mijten) van Weet wat er Leeft.

1.4.2 Technische aspecten

De Edapholog is een geautomatiseerde Berlese-Tullgren opstelling, waarbij een grond- of bladmonster op een zeef onder een warme lamp wordt geplaatst (Figuur 1.18). Door de hitte en het licht kruipen arthropoden naar beneden en komt in een trechter terecht. Onderaan de trechter komen de arthropoden op een plaat waar een camera op staat gericht. De beelden van de camera worden gebruikt voor beeldherkenning en classificatie van de objecten. Om de zoveel tijd wordt doormiddel van een overdruk de plaat schoon geblazen en komen de arthropoden in een buisje met alcohol of iso-propanol terecht.



Figuur 1.18: Schematische tekening van de Edapholog 1.0.



Figuur 1.19: Linker foto, Edapholog 1.0 met de verschillende onderdelen. 1: hittelamp, 2: bak met gaas, 3: kartonnen trechter, 4: blazertje, 5: opvangbuisje met iso-propanol. Rechterfoto, de Edapholog met afdekkap.

Om de Edapholog geschikt te maken voor bladmonsters met trips is een afdekkap boven op te bak met gaas geplaatst, zodat volwassen trips niet weg kan vliegen (Figuur 1.19). De lamp scheen door een gat in de afdekkap.

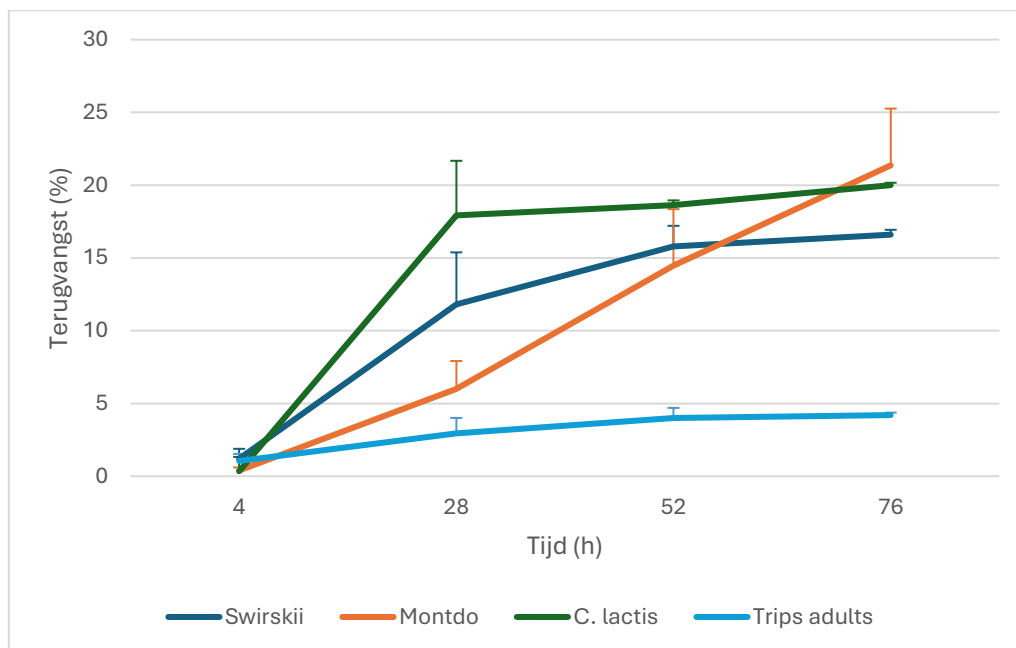
1.4.3 Terugvangst ratio's

In een experiment is bepaald hoe snel en hoeveel mijten (*Amblyseius swirskii*, *Transeius montdorensis* en *Carpoglyphus lactis*) en trips (*Frankliniella occidentalis*) worden teruggevangen nadat deze op de Edpholog zijn gelegd. Voor een gedetailleerde beschrijving van deze pilot zie het verslag van het HAS studentenonderzoek (Beekhuizen et al., 2024).

De verschillende mijten werden aangeleverd via Koppert en bij 10 °C bewaard. Omdat de roofmijten geleverd worden samen met de voermijten moesten deze eerst gescheiden worden. Pogingen om dit te volbrengen met behulp van een warmwaterbad waren niet succesvol en uiteindelijk zijn de roofmijten één voor één opgevisst met een haar onder een binoculair en overgebracht in een epje.

Het percentage individuen dat na 76 uur is teruggevangen was 21,4% voor de Montdo-mijten, 11,5% voor de Swirskii-mijten en 20.0% voor de voermijten (*C. lactis*). Van de ingezette tripslarven werd 138% teruggevangen, wat te verklaren is doordat tripseitjes in het blad uitkwamen gedurende het experiment. Van de tripsadulten is 3,6% teruggevangen. Voor de Swirskii roofmijt, de voermijt en tripsadulten was na 28 tot 52 uur de maximale

terugvangst bereikt (Figuur 1.20). Om de Montdo roofmijten terug te vangen is een langere tijd (>76 uur) nodig.



Figuur 1.20: Gemiddelde procentuele terugvangst (met standaardfout) met de Edapholog 1.0 door de tijd van *Amblyseius swirskii* (Swirskii), *Transeius montdorensis* (Montdo) en *Carpoglyphus lactis* en trips adulten (*Frankliniella occidentalis*).

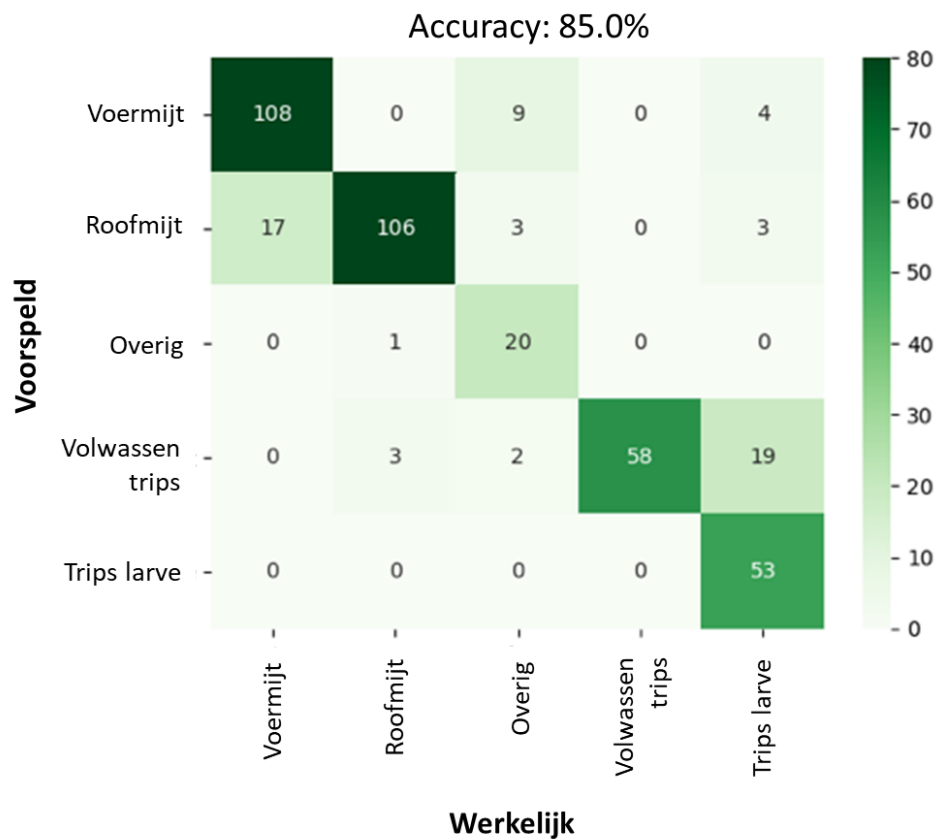
1.4.4 Trainen van Edapholog voor automatische beeldherkenning

Om de Edapholog te trainen voor automatische beeldherkenning zijn met de Edapholog opnames gemaakt van verschillende batches met *A. swirskii*, *T. montdorensis*, *C. lactis* en *F. occidentalis* adulten en larven en opnames van een combinatie van soorten. De opnames zijn door Edaphon Kft gebruikt om het algoritme te trainen en om te valideren hoe goed de beeldherkenning werkt.

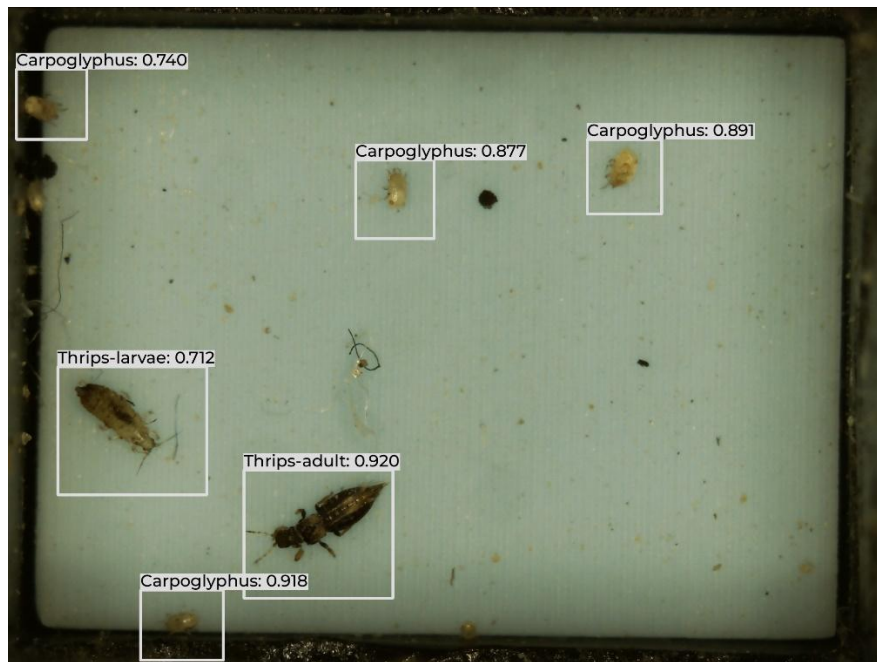
Amblyseius swirskii en *T. montdorensis* lijken heel veel op elkaar en het algoritme kon deze twee niet van elkaar onderscheiden. Het was de hoop dat deze twee soorten op basis van mogelijk ander loopgedrag van elkaar te onderscheiden zouden zijn, maar de opnames waren niet geschikt om gedragsparameters te kunnen vaststellen. Om gedrag te kunnen meten moet één enkel individu voor een bepaalde tijd goed zichtbaar zijn. In de praktijk is dit niet haalbaar, omdat er vaak meerdere individuen tegelijk zichtbaar zijn, deze elkaar kruisen en soms uit beeld verdwijnen. In de beeldherkenningsoftware van de Edapholog zijn deze twee soorten als één groep 'roofmijten' geclassificeerd.

In de matrix in Figuur 1.21 is te zien hoe goed de verschillende groepen van elkaar te onderscheiden waren. In de kolommen staan het werkelijk aantal individuen per soort(groep). In de rijen staan het aantal individuen per soort(groep) door het algoritme geclassificeerd. Er waren bijvoorbeeld 125 *C. lactis* mijten in de dataset, waarvan er 108

(86%) als *C. lactis* zijn geïdentificeerd en 17 (14%) als rootmijt. Het algoritme kon functioneert ook als er meerdere individuen van verschillende soorten zichtbaar waren op de fotografische plaat (Figuur 1.22)



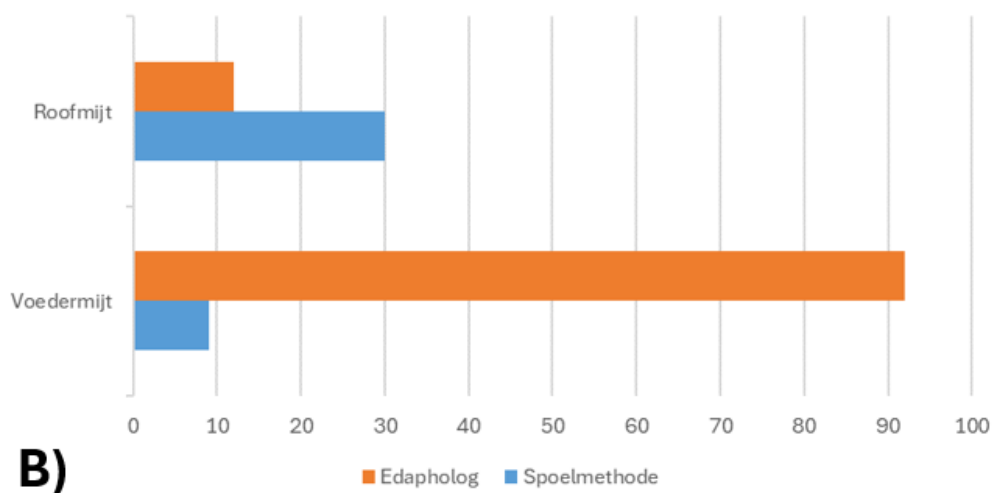
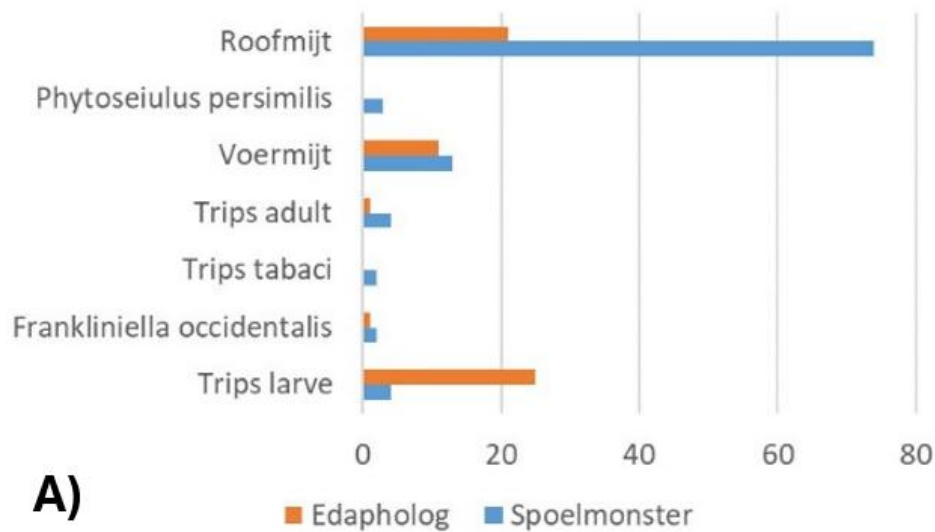
Figuur 1.21: Confusion matrix die laat zien hoe goed het algoritme beelden van vier groepen arthropoden kan onderscheiden: voermijt (*Carpoglyphus lactis*), roofmijten (*Transeius montdorensis* en *Amblyseius swirskii*), volwassen trips, trips larven en overige objecten.



Figuur 1.22: Classificatie van arthropoden met de Edapholog 1.0.

1.4.5 Vergelijking Edapholog met spoelmonsters

Om te zien hoe de Edapholog presteert ten opzichte van de huidige methode van insturen van bladmateriaal door telers voor spoelmonsters is een vergelijking gemaakt tussen deze methoden. Hiervoor zijn twee keer monsters genomen bij een komkommerteler waarvan de helft is opgestuurd naar het lab van Koppert voor analyse met de spoelmethode en de andere helft is meegenomen om op de Edapholog te leggen. Het monster van de Edapholog is vervolgens ook naar het lab van Koppert gestuurd voor identificatie. In de eerste vergelijking (linker grafiek in Figuur 1.23) zijn over het algemeen minder exemplaren per soort teruggevonden met de Edapholog dan met de spoelmethode, behalve voor het aantal tripslarven en de voermijten. Met de Edapholog hebben tripslarven de tijd om nog uit het ei te komen. In de tweede vergelijking (rechter grafiek in Figuur 1.23) zijn opvallend weinig roofmijten aangetroffen met de Edapholog en juist veel voermijten. Het grote aantal voermijten is toe te schrijven aan 2 van de 6 monsters met uitzonderlijk veel individuen (23 en 62).



Figuur 1.23: Vergelijking van tellingen van bladmonsters van een komkommerteelt tussen de Edapholog en een Spoelmonster. Figuur **A)** en **B)** zijn van verschillende telers. Voor **A)** zijn 6 emmers geanalyseerd per methode. Voor **B)** zijn 2 emmers geanalyseerd per methode. Getoond zijn het totaal aantal individuen per groep.

1.4.6 Vergelijking Edapholog trechter met de Berlese-Tullgren opstelling

De Edapholog trechter werd gebruikt voor een vergelijking met een Berlese-Tullgren opstelling (Figuur 1.24). Voor een gedetailleerde beschrijving van dit experiment zie het verslag van het HAS studentenonderzoek (Beekhuizen et al., 2024).

De vergelijking tussen beide methoden is uitgevoerd met komkommerblad uit de kas van HAS Den Bosch. Dit bladmateriaal werd willekeurig verdeeld over 4 Edapholog verzamelbakken en 4 Berlese-Tullgren opstellingen. Op gezette tijden werd het buisje met iso-propanol van beide opstellingen vervangen en werd het aantal roofmijten en trips adulten en larven geteld. Er is hierbij geen onderscheid gemaakt tussen soorten. Bij de

Edapholog was de fotografische plaat verwijderd en deze opstelling fungeerde dus precies zoals de Berlese-Tullgren opstelling.

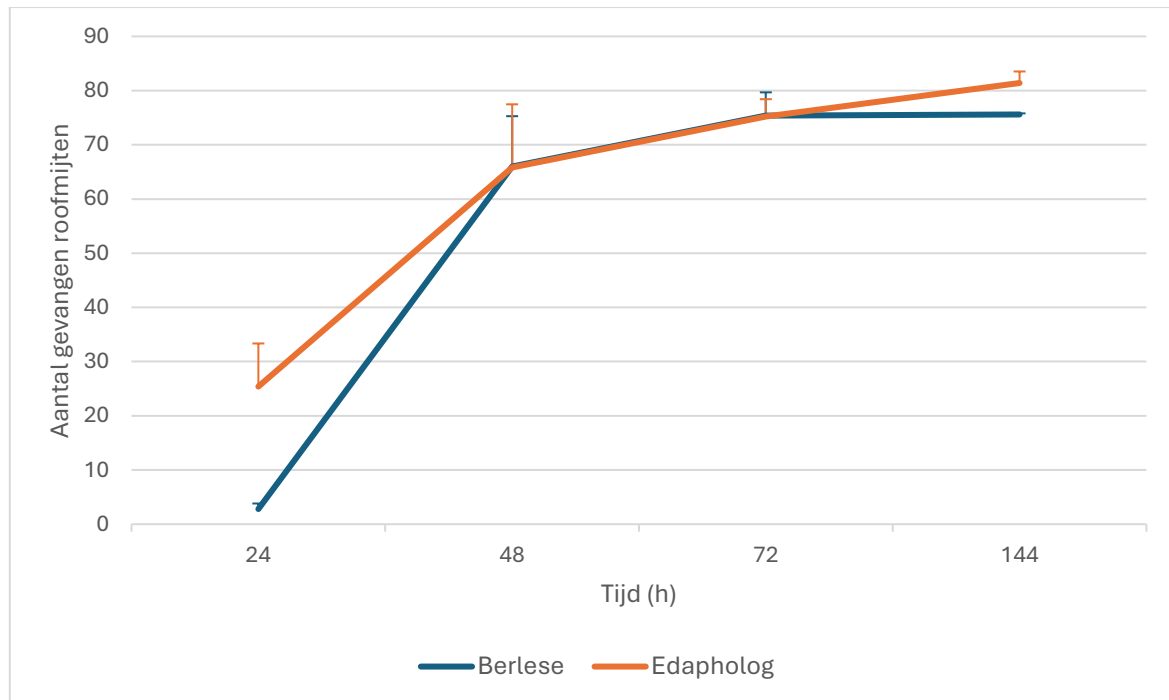
De Edapholog opstelling ving de eerste 24 uur 5 keer meer roofmijten dan de Berlese-Tullgren opstelling en ook na 144 uur had de Edapholog meer roofmijten gevangen, maar het verschil was nu veel kleiner (Figuur 1.25).

Het aantal terug gevangen tripsadulten was statistisch niet verschillend tussen beide methoden (Figuur 1.26). Dit is opmerkelijk want de Edapholog had een afdekkap en de Berlese-Tullgren opstelling niet. Het is daarom de vraag op de afdekkap nodig is om te voorkomen dat trips adulten wegvliegen.

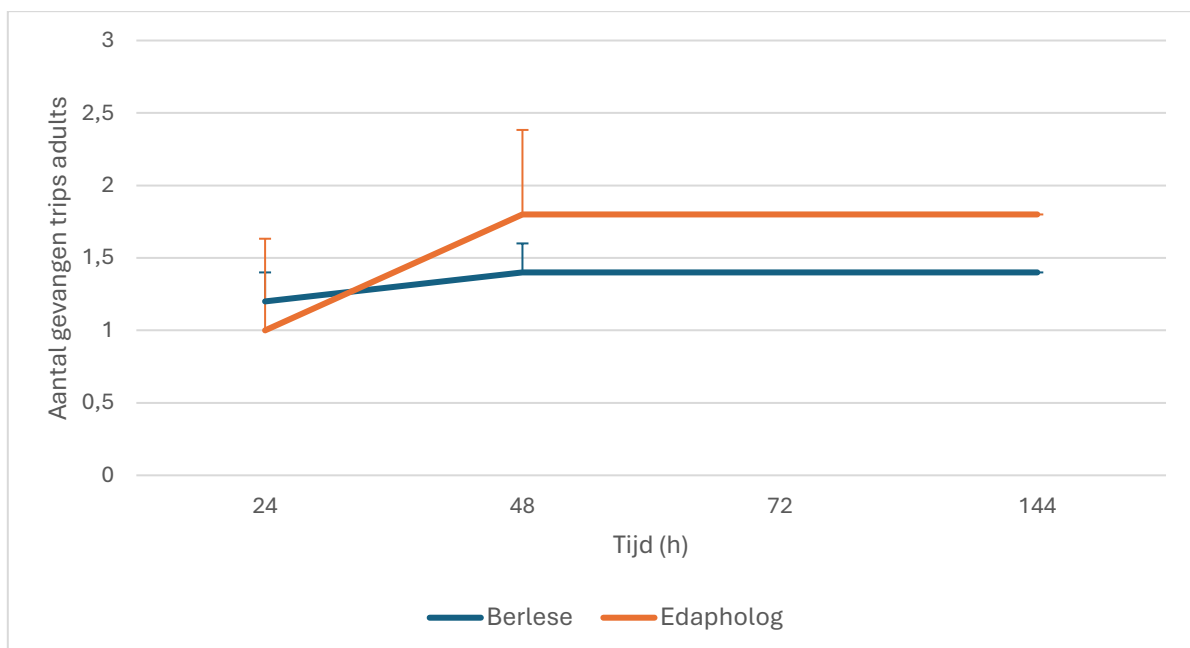
De Edaphog ving zo'n 10 keer meer tripslarven dan de Berlese-Tullgren opstelling (Figuur 1.27). Mogelijk was de lamp van de Berlese-Tullgren opstelling te heet waardoor tripslarven en eitjes dood gingen.



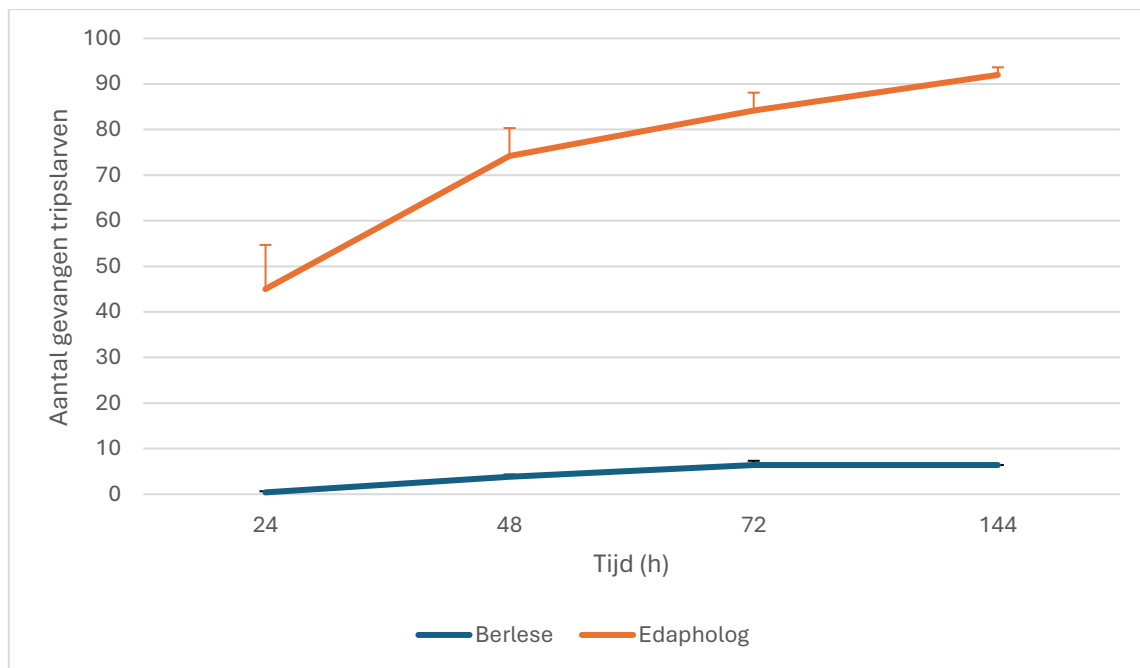
Figuur 1.24: Berlese-Tullgren opstelling bestaande uit 1: hittelamp, 2: trechter, 3: opvang potje met iso-propanol.



Figuur 1.25: Cumulatieve weergave van het gemiddelde aantal gevangen roofmijten met de Berlese-Tullgren opstelling en de Edapholog. De foutbalk geeft de standaardfout weer.



Figuur 1.26: Cumulatieve weergave van het gemiddelde aantal gevangen trips adults met de Berlese-Tullgren opstelling en de Edapholog. De foutbalk geeft de standaardfout weer.



Figuur 1.27: Cumulatieve weergave van het gemiddelde aantal gevangen trips larven met de Berlese-Tullgren opstelling en de Edapholog. De foutbalk geeft de standaardfout weer.

1.4.7 Algemene conclusie en integratie van Edapholog in Weet wat er Leeft app

Het doel van Weet wat er Leeft is om een automatisch meetprogramma te hebben voor de kas dat real-time inzicht verschaft in de populatiedichtheden van de doelorganismen. Voor twee van de doelorganismen: *Transeius montdorensis* en *Amblyseius swirskii* is onderzocht of de Edapholog 1.0 opstelling voor dit doel kan worden gebruikt.

De Edapholog was instaat om via beeldherkenning de twee roofmijten te onderscheiden van de andere veel voorkomende organismen in bladmonsters (trips en de voermijt *Carpoglyphus lactis*), maar kon de twee roofmijten niet van elkaar onderscheiden. Een opvallend resultaat is dat de Edapholog opstelling goed in staat bleek om tripslarven te extraheren. Door de opwarming van het bladmateriaal komen de eitjes uit en is in potentie te 'voorspellen' hoe de tripsdruk zich de komende tijd gaat ontwikkelen.

Een volgende vraag is hoe bruikbaar de informatie is voor een teler die met de Edapholog wordt verkregen. Om een waarneming te doen moeten bladeren van de planten worden verzameld en op een Edaphog gelegd en dit zou een teler zelf kunnen uitvoeren. Vervolgens duurt het minstens 52 uur om *A. swirskii* of 76 uur om *T. montdorensis* uit de bladmassa te laten lopen. Mogelijk kan deze terugvangstijd nog worden verkort door het protocol aan te passen (hetere lamp, blad eerst versnipperen en regelmatig keren). Hoe dan ook zal er enige tijd overheen gaan voordat een teler resultaat ziet. De Edapholog kan in die zin geen real-time oplossing bieden voor het monitoren van roofmijten.

Het aantal roofmijten dat met de Edapholog opstelling uit een bladmassa kan worden gehaald was een kwart tot een-derde lager dan dat met een spoelmonster wordt

gedetecteerd. In het terugvangexperiment lag het percentage teruggevangen roofmijten zelfs op een-vijfde. De detectiegrens van de Edapholog zal dus een stuk lager liggen dan de nu standaard meetmethode en deze grens zal in praktijk nog lager zijn, omdat de foutmarge van het automatisch classificeren nog niet is meegenomen. Er zal met de Edapholog dus meer, of vaker, bladmonsters genomen moeten worden om eenzelfde detectiegrens te krijgen voor roofmijten als die van de spoelmethode. In de praktijk zou het betekenen dat er bijvoorbeeld elke dag of twee dagen bladmonsters worden verzameld en ingezet op een Edapholog.

De laatste vraag is hoe de informatie van de Edapholog kan worden geïntegreerd in de Weet wat er Leeft app. Op dit moment komt de data van de sensor op de telefoon van de gebruiker in een Edapholog app. Het is mogelijk om de Edaphologs op termijn zo te programmeren dat deze de data doorsturen naar de Weet wat er Leeft app in een JSON of ander formaat. De vraag is eerder in hoeverre de Edapholog een aanvulling zal worden op de nu bestaande monitoring in een kas. Vanuit de telers lijkt er vooral belangstelling om de Edapholog te kunnen inzetten om een mogelijk tripsprobleem optijd waar te nemen. De Edapholog kan tripslarven detecteren, maar maakt geen onderscheid tussen soorten trips. In die zin is het de moeite waard om een paar Edaphologs een periode door een teler in de praktijk te laten gebruiken, waar hij zelf kan inschatten van welke waarde de sensor is ten opzichte van de spoelmethode. Voor eventuele verdere ontwikkeling van de Edapholog techniek in de kas is het goed te realiseren dat de Edapholog 1.0 is opgevolgd door een nieuwe versie die helemaal is afgestemd op het nemen van grondmonsters en dat de 1.0 versie niet meer wordt geproduceerd.

1.5 Discussie Vision

1.5.1 Mobiele telefoon camera

De camera op een mobiele telefoon kan een goed instrument zijn om insectenpopulaties in een kas te monitoren met behulp van plakplaten als aan een aantal voorwaarden wordt voldaan. De beste prestaties worden verkregen met een iPhone 7 camera (in vergelijking iPhone 13 en Samsung Galaxy S52), het gebruik van een externe flitsapparaat in combinatie met gele of blauwe plakplaten met droge lijm. De populatieontwikkeling kan het best worden gemonitord als er wekelijks foto's worden gemaakt van een plakplaat, waarna de plakplaat wordt vervangen. Het maken van foto's onder verschillende kleuren licht in een kas is nog wel een uitdaging en vraagt nadere uitwerking. Vooralsnog kan worden volstaan om onder gekleurd licht altijd de externe flits te gebruiken.

Bij de training van de beeldherkenningsalgoritmen is gekozen om de foto's niet te labelen op de soortnaam, maar op hogere taxonomische groepen: trips, bladluis, wittevlieg en sluipwesp. Het labelen op soortniveau wekt de indruk dat het mogelijk is om verschillende soorten binnen die groepen tot op soortniveau te kunnen onderscheiden. Dat viel met de

keuze van de doelsoorten niet binnen de scope van het onderzoek en is niet getest. De keuze voor hogere taxonomische groepen geeft een beter beeld van wat de algoritmen daadwerkelijk kunnen herkennen.

Het type mobiele telefoon had een grote invloed op de prestaties van de beeldherkenning en omdat niet iedereen dezelfde mobiele telefoon heeft is het advies om bij het trainen van beeldherkenningsalgoritmen gebruik te maken van meerdere type telefoons.

De prestaties van de beeldherkenningsalgoritmen waren onder gecontroleerde omstandigheden (met insecten uit kweken) continu redelijk tot goed. In de praktijk waren de resultaten variabler. Voor trips in een chrysantenkas was het mogelijk om betrouwbare trendlijnen te verkrijgen tussen het aantal door de mobiele telefoon waargenomen individuen op een plakplaat en het daadwerkelijke aantal. Uitdagingen bij het testen in praktijkkassen was dat er soorten op de plakplaten zaten waar het algoritme niet op was getraind en dat er op de foto's achtergrond objecten zichtbaar waren. Door het algoritme uit te breiden met foto's van andere soorten en de achtergrond weg te knippen zal de prestatie van de beeldherkenning in de praktijk verbeteren.

Het gebruik van mobiele telefoons voor fotografie heeft een aantal voor- en nadelen. Een groot voordeel is dat iedereen over een mobiele telefoon beschikt en er dus geen nieuwe apparatuur aangeschaft hoeft te worden. Het is eenvoudig in te passen in scoutprotocollen waarbij plakplaten al gebruikt worden. Met bijvoorbeeld de app van Imec (zie later hoofdstuk 4.4) of Koppert is het eenvoudig om de platen te fotograferen en te analyseren. Een nadeel van het gebruik van mobiele telefoons voor fotografie van plakplaten is de grote variatie in omstandigheden, waardoor de prestatie van de beeldherkenning verminderd. Andere, vaste opstellingen (bijvoorbeeld oplossing biobest Trap-Eye) lossen dat op met gestandaardiseerde hardware, zoals standaardcamera's en flitsapparatuur, die zich op vaste afstanden bevinden van de plakplaten. Voor een deel van de variatie kan worden gecompenseerd, door bijvoorbeeld gebruik te maken van een externe flits en het trainen van het model met verschillende type camera's. Het voordeel van het gebruik van losse plakplaten is dat het erg flexibel ingezet kan worden, zowel qua locatie als intensiteit. Het op- en afschalen van de monitoring kan gebeuren door het aantal plakplaten te verminderen of te vermeerderen.

Omdat het model is getraind op gele plakplaten met droge lijm, was het niet verwonderlijk dat het model minder presteerde bij met name natte lijm, en in mindere mate bij andere kleuren. Door het trainen van nieuwe modellen voor de specifieke plakplaten met een zeer beperkte dataset, was het mogelijk de performance te verhogen. Voor de meeste doelsoorten zou op andere platen getraind kunnen worden, waarbij met name bij de droge plakplaten een goed resultaat wordt verwacht.

De doelsoorten die zijn geselecteerd op basis van relevantie voor de teelten door de gewascoöperaties zijn klein in formaat. Bij verdere uitbreiding van het soortenaantal is de

verwachting dat het minimaal mogelijk is tot orde of familie te determineren. Met name voor de grotere soorten is, mede afhankelijk van de uitwendige verschillen tussen soorten, de verwachting dat soortidentificatie mogelijk is met het gebruik van de camera van een telefoon.

1.5.2 Hyperspectraal camera

Het onderzoek met de hyperspectraal camera heeft aangetoond dat elk van de doelsoorten een unieke 'spectrale handtekening' heeft, waardoor een zeer nauwkeurige detectie en classificatie mogelijk is. Dit biedt een significant voordeel ten opzichte van standaard camera's. De resultaten suggereren echter dat voor praktische toepassingen multispectrale of RGB-gegevens vaak voldoende zijn. Deze technieken vereisen minder complexe beeldverwerking en zijn aanzienlijk goedkoper, terwijl ze toch bevredigende prestaties leveren voor het detecteren en classificeren van de doelsoorten.

Toekomstig onderzoek op het gebied van hyperspectraal imaging met insecten zou zich daarom kunnen richten op het combineren van multispectrale (dit is goedkoper omdat minder banden worden gebruikt) en RGB-beeldvorming technieken om een kosteneffectieve en praktische oplossing te bieden voor insectendetectie in het veld. Ook kan uitbreiding naar andere multispectrale banden een optie zijn, bijvoorbeeld ultraviolet (UV) beeldvorming kan nieuwe mogelijkheden openen voor het onderscheiden van insecten op basis van hun unieke spectrale kenmerken in die golflengte.

1.5.3 Edapholog

Het onderzoek heeft aangetoond dat het met de Edapholog technisch mogelijk is om roofmijten automatisch te detecteren uit bladmonsters en te onderscheiden van voermijten. De detectiegrens voor roofmijten is met de Edapholog echter beduidend lager dan van een spoelmonster. De Edapholog zal in die zin geen vervanging zijn voor het laten uitvoeren van spoelmonsters, maar zou een hulpmiddel kunnen zijn bij het zelf scouten op mijten in een teelt.

De Edapholog bleek bijzonder effectief voor het detecteren van tripslarven, omdat het de tripseitjes die in het blad zitten versneld uitbroedt. In theorie zou het daarmee mogelijk zijn om als teler gewaarschuwd te worden voor een beginnende tripshaard.

Hoofdstuk 2: eDNA

2.1 Algemene inleiding

DNA kan op verschillende manieren gebruikt worden om de aanwezigheid van soorten aan te tonen. Daarbij wordt gebruik gemaakt van barcoding, waarbij het DNA van een individu wordt vergeleken met een databank met barcodes en een percentage aan overeenkomst wordt gegeven. Indien dat percentage hoog genoeg is, wordt aangegeven welke soort het betreft. Hierbij zijn verschillende manieren van detectie mogelijk:

- DNA barcoding: waarbij het DNA van een enkel individu wordt geanalyseerd
- DNA meta-barcoding: waarbij het DNA van een meerdere individuen tegelijkertijd wordt geanalyseerd
- eDNA meta-barcoding: waarbij DNA uit samples uit de omgeving wordt geanalyseerd. Hierbij hoeft geen biotisch materiaal verzameld te worden en, binnen de scope van dit project, geen complete geleedpotigen aanwezig te zijn in het sample.

Het DNA dat wordt geanalyseerd bij eDNA analyses, is aanwezig in onder andere huidcellen, haartjes en vervellingshuiden van insecten, maar volledige geleedpotigen kunnen, afhankelijk van samplingmethoden, wel aanwezig zijn.

Er zijn meerdere moleculaire technieken waarmee DNA aangetoond kan worden. Een veelgebruikte techniek is het gebruik van PCR, waarmee door middel van primers heel specifieke delen van het DNA (cytochroom oxidase I; COI) worden vermeerderd en meetbaar worden gemaakt door middel van sequencing. Een andere methode, is de shotgun methode, waarbij het volledige DNA in stukken wordt geknipt en gesequenced. Hierbij hoeven geen primers gebruikt te worden die specifiek zijn voor bepaalde soortgroep, en wordt al het DNA geanalyseerd.

De voordelen van eDNA-analyses is dat de sampling gestandaardiseerd uitgevoerd kan worden en zeer specifieke uitslagen geeft op een breed spectrum, zonder dat taxonomische vaardigheden nodig zijn. Het zou daarmee een goede aanvulling kunnen vormen op handmatig scouten en herkenning van insecten op basis van visionetechnieken.

Het nadeel van deze techniek is dat het voornamelijk een kwalitatieve analysemethode is en weinig zegt over aantallen. In dit onderzoek wordt de praktische toepassing van eDNA-analyse onderzocht voor de detectie van soorten geleedpotigen in kasteelten (roos, chrysant, potorchidee en komkommer), met in het bijzonder aandacht voor de kaswittevlieg, Californische trips, groene perzikluis, de roofmijten *Transeius montdorensis* en *Amblyseius swirskii*, en de parasitaire wesp *Aphidius colemani*.

2.2 Insectendetectie met eDNA analyse

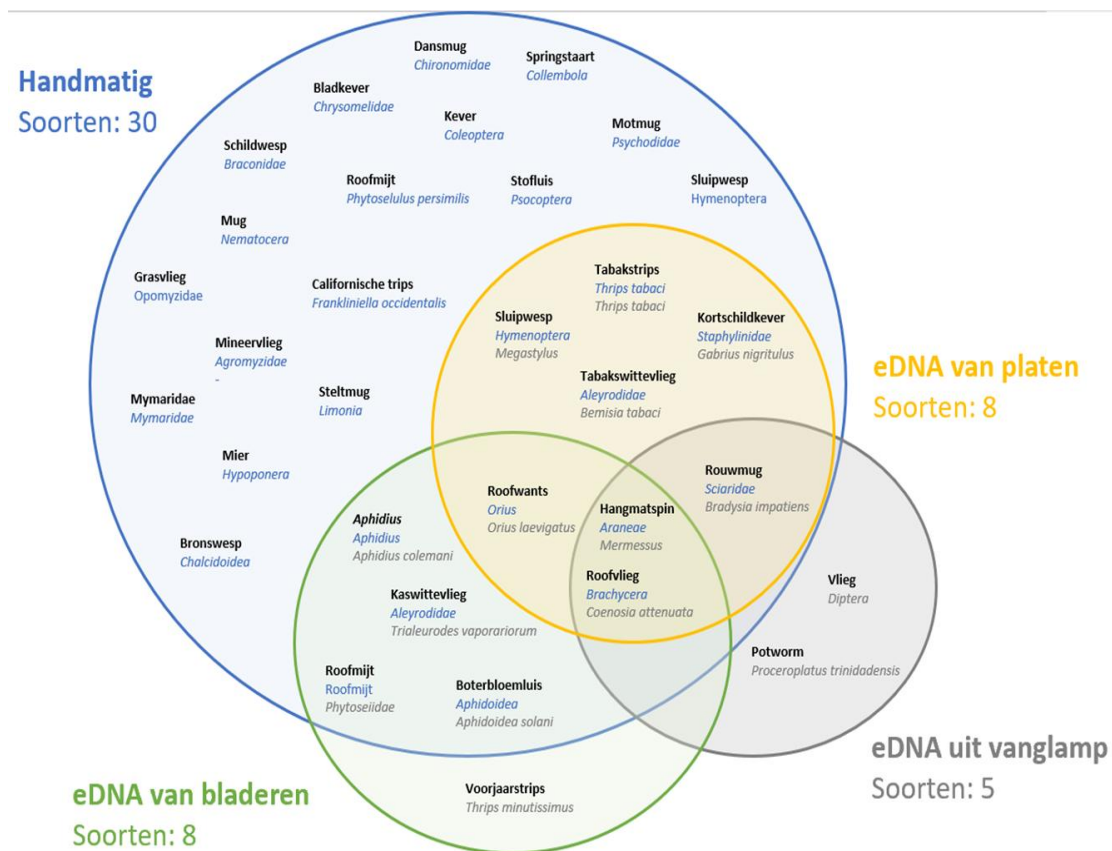
2.2.1 eDNA detectie met PCR-analyse

Om te testen of het mogelijk was insecten op plakplaten te herkennen met eDNA analyses en of de eDNA reads (het aantal keer dat een specifiek stukje DNA vermeerderd was) gekwantificeerd konden worden, zijn 45 Horiver Dry stick plakplaten (geel, droge lijm, 10 x 25 cm, Koppert Biological Systems) besmet met verschillende samenstellingen van de doelinsecten en geanalyseerd. Insecten waren afkomstig van de kweek van SCFF (trips, wittevlug, bladluis) en Koppert (*Aphidius colemani*). Deze platen werden vervolgens gewassen in buffer en het lysaat is geanalyseerd door het COI-gen te sequencen.

41 van de 45 plakplaten gaven een uitslag, waarbij de aanwezigheid van *Aphidius colemani*, de groene perzikluis en kaswittevlug aangetoond werden. De Californische trips werd niet aangetoond. In totaal was de eDNA-analyse in 87,1% van de gevallen correct in het aantonen van de aanwezige soorten. Daarnaast zijn ook 7 soorten gevonden die niet actief op de platen geplaatst waren, maar waarvan DNA in de omgeving aanwezig was. Van de gevonden doelsoorten zijn het aantal eDNA reads gecorreleerd aan het daadwerkelijk aantal aanwezige insecten en het aantal insecten dat door middel van vision-technieken met een model in Custom Vision is geteld. Hierbij werd een zwakke maar significante correlatie gevonden bij de groene perzikluis. Bij de kaswittevlug werd alleen een zwakke correlatie gevonden bij de handmatige tellingen. Tussen het relatief aantal individuen en het relatief aantal reads werd geen correlatie gevonden (Benschop et al. 2023).

Na dit onderzoek is de methode van eDNA extractie verbeterd (intern protocol SCFF/Groen Agro Control) en is een gaas over de plakplaten geplaatst, dat met een roller plat is gedrukt, om meer DNA vrij te maken. Bovenstaande onderzoek is herhaald, met de tabakstrips in plaats van de Californische trips, omdat laatstgenoemde niet gedetecteerd werd als het in combinatie met andere insectensoorten op de platen aanwezig waren. Hier werden naast significante correlaties tussen aantallen *Aphidius colemani* ($R^2 = 0,57$) en kaswittevlug ($R^2 = 0,64$), ook significante correlaties gevonden tussen relatieve aantallen insecten en het relatief aantal reads, voor *Aphidius colemani* ($R^2 = 0,90$), de tabakstrips ($R^2 = 0,69$) en kaswittevlug ($R^2 = 0,94$). Dit werd in de praktijk (chrysantenkas) echter niet gevonden. Daarin werden met eDNA 8 soorten gevonden op 15 platen, terwijl bij determinatie met het oog 19 taxonomische groepen werden geïdentificeerd. Bovendien was er geen correlatie tussen het aantal individuen en DNA reads: het DNA werd gedomineerd door de vliegsoort *Coenosia attenuata* met een aandeel van 99,8%, terwijl het aantal individuen op de platen gemiddeld 50,0% bedroeg. Dit is te verklaren door het formaat van de vliegen en diens vliegspieren in vergelijking met andere soorten op de platen. Het DNA dat gebruikt wordt bij de analyses is afkomstig van mitochondrieën, dat in hoge mate aanwezig is in vliegspieren die bij deze vliegen relatief groot zijn. Naast het wassen van plakplaten is tevens DNA verzameld door middel van het swabben van bladmateriaal in de kas en de vangsten met een Ferris UVA vanglamp van twee weken te analyseren. Beide methoden

gaven aanvullende vondsten ten opzichte van handmatig determineren van plakplaten en de DNA analyses van plakplaten (Figuur 2.1, van den Einde et al. 2023).



Figuur 2.1: Venn diagram voor de verschillende soorten gevonden in de kas voor de methodes: handmatig scouten van vangplaten (n=56 vangplaten, blauw), eDNA verzameld van bladeren doormiddel van swabs (n=3 swabs, groen), eDNA verzameld van vangplaten (n=15 vangplaten, geel) en eDNA verzameld van een Ferra vanglamp (n=1, grijs). De zwarte naam geeft de algemene naam weer, de blauwe naam is de handmatig geïdentificeerde soort op een zo klein mogelijk taxonomisch niveau, de grijze naam geeft de naam gevonden door eDNA de analyses weer.

Aanvullende experimenten

Omdat de swabs aanvullende uitslagen lieten zien ten opzichte van het DNA van de plakplaten, is besloten een vervolgonderzoek uit te voeren, waarbij drie experimenten zijn uitgevoerd. In het eerste experiment zijn bladeren van chrysant in een petrischaal geplaatst en is een combinatie van verschillende dichtheden van schadelijke soorten (tabakstrips, kaswittevlieg en groene perzikluis) of biologische bestrijders (*Aphidius colemani*, *Amblyseius swirskii* en *Traneius montdorensis*) toegevoegd. Een week later zijn deze bladeren geswabd met katoenen swabs en zijn deze geanalyseerd.

Dit leverde echter geen positief signaal voor DNA op. In het tweede experiment werden chrysanten met 12 tot 15 bladeren onder glazen stolpen met een waterfilter besmet met hoge (15) en lage dichtheden (3) van dezelfde combinaties van soorten. Na een week

werden drie of vijf bladeren verwijderd en op dezelfde manier geswabd als in het eerste experiment. De bladeren uit de stolpen met natuurlijke vijanden leverden geen betrouwbaar PCR-product op. Van de schadelijke soorten werd wittevlieg bij alle behandelingen teruggevonden. Tabakstrips werd in één behandeling teruggevonden (3 bladeren met lage dichtheid). Daarnaast werd DNA gevonden van *Aphidius colemani* in de samples waar ze niet waren toegevoegd en werden andere mijten (Tarsonemidae) aangetoond. In het derde experiment werden vier Horiver dry plakplaten (Koppert Biological Systems, droge lijm, geel) opgehangen. Twee daarvan leverde geen betrouwbaar PCR-product, de overige twee geen betrouwbare DNA uitslagen (van den Berg & Groot, 2024).

PCR (nog) niet consistent genoeg

Uit deze onderzoeken kan geconcludeerd worden dat PCR DNA analyse van de plakplaten niet tot consistente resultaten leidt. Indien er betrouwbare data uitkomt, geeft het aanvullende informatie, waaronder de aanwezigheid van DNA van soorten die met het oog niet gevonden zijn, maar mogelijk wel in de omgeving aanwezig zijn. Vergeleken met handmatige determinaties leverde het minder soorten op, maar was het significant minder bewerkelijk om data te verkrijgen. Het kwantificeren van de DNA uitslagen was onder gecontroleerde omstandigheden mogelijk. In de praktijk echter niet.

2.2.2 eDNA detectie met de shotgun methode

Waar met PCR een specifiek deel van het DNA wordt vermeerderd, wordt bij de shotgun-methode het gehele DNA in kleine stukken opgebroken. De sequenties worden vervolgens vertaald, waarna het genoom in beeld wordt gebracht door overlappende stukken aan elkaar te verbinden. Vervolgens wordt dit genoom vergeleken met bekende genomen in een database om soorten op naam te brengen. Deze methode beperkt zich niet tot dierlijk DNA, ook virussen, schimmels, bacteriën en archaea worden geïdentificeerd. Daarnaast zijn naast plakplaten, UV-vallen en swabs nog andere methoden om DNA van insecten en andere geleedpotigen te verzamelen. Dit zijn onder meer het spoelen van blad met demiwater, het samplen van drain water en het samplen van water van vanglampen. Deze methoden zijn getest in de HAS kas, voordat besloten is welke methoden in de praktijk toegepast en geanalyseerd zijn (Tabel 2.1).

Tabel 2.1: multiple criteria analysis voor de zes eDNA-methoden die getest zijn bij de pilot in de HASKas, met een schaal van 1/rood (slecht) tot 5/groen (goed).

Methoden:	Arbeidsefficiëntie	Prijs	Toepasbaarheid voor alle teelten	Aantal relevante soorten	Totaal aantal soorten	Gemiddelde
1. Wattenstaafje zonder buffer	4	5	5	1	1	3,2
2. Wattenstaafje met buffer	4	3	5	3	3	3,6
3. Afspoelen met water	3	5	5	2	2	3,4
4. Drainwater	5	5	2	1	1	2,8
5. Vanglamp	3	4	5	3	3	3,6
6. Vangplaat	1	2	5	5	5	3,6

Beste sampling methodes

Uit deze analyse bleken de swab (wattenstaafje) met buffer, het samplen van het water van de vanglamp en de vangplaten het meest efficiënt en effectief. Deze methoden zijn toegepast in de praktijk bij de komkommer-, rozen-, chrysanten- en potorchideeënteelt (m.u.v. vanglampen bij orchidee) en samples zijn gesequences middels de shotgunmethode en intern protocol van Groen Agro Control. De resultaten van de DNA analyses zijn vergeleken met traditionelere monitoringsmethoden, zoals spoelsamples (uitgevoerd door Koppert Biological Systems), handmatig scouten en automatische beeldherkenning (Tabel 2.2).

Tabel 2.2: Aangetroffen relevante taxonomische groepen per detectiemethode. De afkortingen bij locaties staan voor: HASKas (H), Komkommer (K), Potorchidee (P), Roos (R), Chrysant (C).

Groep	Wetenschappelijke naam	Uitgevoerde locaties ->	H,K,P,R,C	H,K,R,C	H,K,P,R,C	H	H	K,P,R,C	H,K,P,R,C	H,K,P,R,C
		Vangplaat	Vanglamp	Swabs	Water van blad	Drainwater	Spoelsample	Handmatig	Beeldherkenning	
Bladluizen	<i>Aphidoidea</i>	H, P, R		H			C	H		
Roofmijten	<i>Mesostigmata</i>	H	K, C	H, C	H		K, P, R, C			
Voermijten	<i>Astigmata</i>		C				K, C			
Overige mijten	<i>Prostigmata & Oribatida</i>		H, K, R, C							
Motten/vlinders	<i>Lepidoptera</i>	H, R	H, K, R, C	H						
Rouwmuggen	<i>Sciaridae</i>	H, K, R	H, K, R					K		
Sluipwespen	<i>Ichneumonidae</i>	H							H, K, C	
Tripsen	<i>Thysanoptera</i>	H	H	H	H	H	C	H, P	H, K	
Wantsen	<i>Heteroptera</i>	H	R	H						
Witte vliegen	<i>Aleyrodidae</i>	H	H	H	H	H		H, K, R	H, K	
Wolluizen	<i>Pseudococcidae</i>						P			

Wat bij de plakplaten opvalt, is dat trips, wittevlies en sluipwespen met DNA analyse niet terug gevonden worden in de praktijklocaties, terwijl die wel zijn aangetoond in de HAS kas. Bovendien zijn zij wel met handmatige scouting, spoelen en automatische beeldherkenning gevonden. Bladluizen zijn wel aangetroffen bij potorchidee en roos. Bij de vanglampen worden mijten aangetroffen, wat niet geheel te verwachten was, aangezien vanglampen voornamelijk vliegende insecten aantreffen. De mijten zijn waarschijnlijk door luchtverplaatsing in het water van de val terecht gekomen. Ook bij deze methode worden de doelsoorten niet aangetroffen. Met swabs zijn bij de praktijklocaties alleen roofmijten aangetoond bij chrysant. De overige doelsoorten zijn niet gedetecteerd.

Detectie op soortniveau

Op soortniveau zijn de groene perzikluiz, kaswittevlies en Californische trips alleen in de HAS kas aangetroffen, niet in de praktijklocaties, terwijl Californische trips wel in spoelsamples van chrysant en *montdo*-roofmijten in spoelsamples bij zowel chrysant en komkommer aangetroffen zijn (Tabel 2.3). In de HAS kas wordt niet chemisch en vrijwel niet biologisch bestreden, waardoor hoge dichtheden van deze insectensoorten aanwezig kunnen zijn.

Tabel 2.3: Aangetroffen relevante soorten per detectiemethode. De afkortingen bij locaties staan voor: HASKas (H), Komkommer (K), Potorchidee (P), Roos (R), Chrysant (C).

	Uitgevoerde locaties ->	H,K,P,R,C	H,K,R,C	H,K,P,R,C	H	H	K,P,R,C
Doelsoorten	Wetenschappelijke naam	Vangplaat	Vanglamp	Swabs	Water van blad	Drainwater	Spoelsample
Groene perzikluiskruis	<i>Myzus persicae</i>	H					
Kaswittevlug	<i>Trialeurodes vaporariorum</i>	H	H	H	H	H	
Californische trips	<i>Frankliniella occidentalis</i>		H	H	H	H	C
Sluipwespen	<i>Aphidius colemani</i>						
Roofmijt	<i>Amblyseius swirskii</i>						
Roofmijt	<i>Transeius montdorensis</i>						K, C

Dit duidt erop dat het met deze sample- en analysemethoden niet mogelijk is de aanwezigheid van de doelorganismen binnen Weet wat er leeft betrouwbaar aan te kunnen tonen in praktijksituaties. De DNA-analyses leveren vooral aanvullende informatie over de diversiteit van geleedpotigen binnen kassystemen, waarbij de vanglamp en vangplaat de hoogste diversiteit aan groepen laten zien.

2.3 Discussie en conclusie eDNA

eDNA-technologieën hebben potentieel om de detectiemogelijkheden te verbeteren, met name omdat met eDNA soorten werden gevonden, die met andere detectiemethoden niet werden gevonden. Echter, tijdens dit onderzoek is ook gebleken dat zowel PCR-analyse als de shotgun methode zoals nu uitgevoerd nog niet voldoende zijn om alle soorten goed te monitoren. Hoewel in de literatuur eDNA wel als veelbelovende techniek voor insectenmonitoring wordt gezien, blijkt ook uit ander onderzoek dat de praktijktoepassing nog een uitdagende tweede stap is. Het bleek bijvoorbeeld in een meta-analyse uit 2024 dat monitoring met eDNA, van onder andere insecten in het water, ongeacht de gebruikte genetische marker, gemiddeld genomen minder taxa detecteerde dan morfologische bemonstering (Poyntz-Wright, 2024).

Met PCR kunnen doelsoorten aangetoond worden, maar het leverde niet consequent goede resultaten. Er werd aangetoond dat qua sampling methoden het wattenstaafje met buffer minder effectief was in vergelijking met de andere methoden zoals de vanglamp en de vangplaat. De vanglamp en vangplaat leverden consistente resultaten op in verschillende gewassen, waarbij de vanglamp bijzonder effectief was in het detecteren van mijten, terwijl de vangplaat succesvol bladluizen kon aantonen.

De PCR-analyse was betrouwbaarder dan de shotgunmethode. Met de shotgun methode werden doelsoorten weinig gevonden, terwijl ze wel aanwezig waren. Vergeleken met de andere detectiemethoden geldt wel dat het langer duurt om de data te verkrijgen en dat er relatief veel werk zit in het uitvoeren van de DNA-determinatie en analyse. Het is ook duidelijk geworden dat aantallen per soort bepalen met eDNA niet altijd goed werkte met de huidige methode. Dit kwam voornamelijk omdat het formaat van bepaalde insecten ertoe leidde dat geen andere insecten meer werden gemeten. Het is aan te bevelen hiervoor te corrigeren, oftewel door meer specifieke primers te gebruiken, of door in de na-analyse deze grote soorten uit te filteren.

Toekomstig onderzoek

Om de nauwkeurigheid en betrouwbaarheid van eDNA-methoden te verbeteren, is het essentieel om een gestandaardiseerde database van de genomen van interessante organismen te implementeren. Deze database zou helpen om de detectie van niet-doelsoorten te vermijden en vals-positieve resultaten te verminderen. Door een uitgebreide referentie te hebben kunnen de eDNA-analyse nauwkeuriger worden. Bovendien kan een gestandaardiseerde genoomdatabase helpen om de kosten per monster en per analyse te verlagen door het identificatieproces te stroomlijnen en de noodzaak van herhaalde tests te minimaliseren.

Om het volledige potentieel van eDNA-technologieën te gebruiken is vooral verbetering van de betrouwbaarheid belangrijk: dit kan bijvoorbeeld door de ontwikkeling van gestandaardiseerde protocollen en contaminatiebeheersingsmaatregelen. Dit omvat het vaststellen van minimale drempelwaarden voor eDNA-lezingen en het verbeteren van de nauwkeurigheid van soortidentificatie.

Hoofdstuk 3: VOC's

3.1 Inleiding

EMS heeft een techniek ontwikkelt waardoor in potentie Vluchtige Organische Stoffen/Compounds (VOC) gebruikt kunnen worden om soorten te identificeren en kwantificeren. Het onderzoek in het kader van dit project richtte er zich op of insecten die aanwezig zijn in het gewas geïdentificeerd kunnen worden door middel van VOC metingen? Belangrijk hierbij is of er onderscheid gemaakt kan worden tussen plagen en niet schadelijke insecten en VOC's van planten. Er is gekozen voor de doelsoorten: Californische trips (*Frankliniella occidentalis*) met *Transeius montdorensis* (roofmijt).

Groene perzikluisk (Myzus persicae var. persicae) met *Aphidoletes aphidimyza* (galmug) of *Aphidius colemani* (sluipwesp). En tot slot de combinatie kaswittevlug (*Trialeurodes vaporariorum*) met *Amblyseius swirskii* (roofmijt).

3.2 Experimenten

Er zijn 3 experimenten uitgevoerd. Bij SCFF zijn ethyleen metingen gedaan, met meetapparatuur van EMS. Er is een setup gecreëerd voor detectie van ethyleen bij plaagdruk (Figuur 3.1). Vervolgens zijn op jonge planten met trips, wittevlug en controle de niveaus van ethyleen gemeten.



Figuur 3.1: testopstelling voor de detectie van ethyleen in verschillende behandelde planten.

In het lab van EMS zijn proeven gedaan met de uitstoot van VOCs met losse insecten zonder planten. In dit stadium is onderzocht of we ons moeten focussen op enkel de uitstoot van het insect of dat we vooral moeten kijken naar de signalering van de plant veroorzaakt door de aanwezigheid van het insect op de plant. Het resultaat van deze fase was duidelijk dat de signalering primair veroorzaakt wordt door de plantreactie (Dicke, 2015). De proef met enkel insecten was een belangrijk startpunt om te weten hoe het mechanisme werkt.

In het lab van EMS is er getest met de uitstoot van VOC's met insecten in paprika en komkommer.

Paprika

Om in het lab te kunnen meten is er getest met planten die in de testopstelling van EMS goed ondergebracht konden worden. Een ideale plant om groene perzikluis op te testen is paprika. De afmetingen zijn beheersbaar, kleine plant met toch nog 1-3 vruchten per plant en relatief veel blad (in vergelijking met andere gewassen). EMS heeft een proefopstelling opgebouwd waarin 2 paprikaplanten / komkommerplanten als test- en referentie plant konden worden geplaatst in volledige geconditioneerde omgeving (licht, temperatuur, RV en voeding; Figuur 3.2). Het was technisch uitdagend om een gesimuleerde omgeving te realiseren waarin aan alle voorwaarden voldaan kon worden om betrouwbaar VOC's te meten. Zo zijn er bijvoorbeeld stress signalen afkomstig van licht, voeding, droogtestress en RV variaties die dwars door de eigenlijke proeven signalen kunnen geven.



Figuur 3.2: Testopstelling met referentie- en testplanten, hier met komkommer.

De focus in de proeven lag voornamelijk op het herkennen van unieke VOC's geproduceerd na infestatie met plagen of predatoren in de test-plant. De referentie-plant bleef vrij van plagen. Er zijn variaties op dit thema uitgevoerd door predatoren in de referentie wel bewust te plaatsen en variaties op dit thema waarbij in de laatste fase van een run alsnog in de referentie plagen zijn geplaatst om het principe van VOC productie uit de plant dubbel te toetsen met de test-plant.

De volgende proeven zijn uiteindelijk uitgevoerd (zie Tabel 3.1):

3 runs luizen (2 paprika, 1x komkommer)

1 run met trips

2 runs met wittevlies

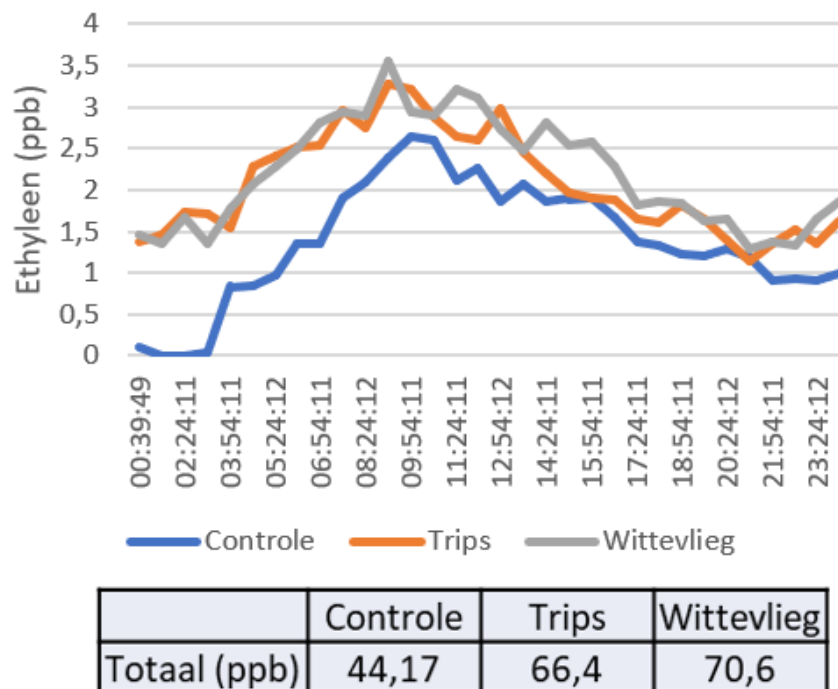
1 run praktijkproef in de kas met luizen

3.2 Resultaten

Ethyleen metingen

Uit de ethyleenmetingen blijkt dat een hogere ethyleen uitstoot stress indiceert (Figuur 3.3).

De ethyleenniveaus zijn hoger in planten waar trips of wittevlies op zitten.



Figuur 3.3: Ethyleenniveaus in controle planten, planten met trips en planten met wittevlies.

Run 1 & 2:

Er zijn zowel wat betreft de plaag alsook de predator goede resultaten geboekt in de detectie van de VOC's. Bij het plaatsen van circa 25 luizen zijn er een relatief grote hoeveelheid plant VOC's onmiddellijk zichtbaar. Dit komt met name door de explosieve reproductie. Het toevoegen van de sluipwesp in run 2 laat zien dat de VOC's afnemen door een afstervende luizen populatie. De stress signalen van de luizen kort na het inbrengen van de predatoren zijn ook waargenomen in deze run.

Tabel 3.1: samenvatting van de experimenten die zijn uitgevoerd in welke plant en met welke plaag. Ook is het aantal VOC's dat gedetecteerd is weergegeven.

Run nr.	Plant	Plaag	Predator	Aantal VOC's plaag	Aantal VOC's predator	Locatie
1	paprika	groene perzikluis	-	++	n.v.t.	lab
2	paprika	groene perzikluis	sluipwesp	++	+	lab
3	paprika	californische trips	-	+	n.v.t.	lab
4	paprika	kas wittevlies	-	-	n.v.t.	lab
5	paprika	kas wittevlies	swirskii	-	-	lab
6	komkommer	groene perzikluis	-	--	n.v.t.	lab
7	paprika	groene perzikluis	-	+-	n.v.t.	kas

Run 3:

Deze proef is slechts 1x uitgevoerd in verband met het hoge risico op kruisbesmetting van trips in de voorraad planten. In run 3 hebben we voldoende VOC's waargenomen, maar de beoordeling blijft + en geen ++ omdat er een 2e proef moet worden uitgevoerd om alle VOC's bevestigd te krijgen. Een aantal VOC's waren onmiskenbaar duidelijk aanwezig. Zonder extra proeven is al aangetoond dat trips goed waar te nemen is met VOC's.

Run 4 & 5:

In run 4 is er wittevlies geplaatst rechtstreeks na levering van Koppert. Deze wittevliegen zijn niet goed aangeslagen op de paprikaplanten, en er was weinig signaal zichtbaar. Hierna heeft de wittevlies enige maanden op kweek gestaan op paprikaplanten, om zo een generatie van wittevliegen te kweken die een voorkeur hebben voor paprika. Deze zijn nogmaals in de proefopstelling geplaatst, en hebben daar langere tijd ingezet om zo de volledige reproductiecyclus van de wittevlies op paprika te bekijken. Ondanks deze methode zijn er helaas alsnog weinig VOC's in deze proef geobserveerd.

Run 6:

De combinatie in run 6 (groene perzikluis in komkommer) heeft de slechtste resultaten die totaal niet bruikbaar zijn. Er zijn geen VOC's meetbaar. De vermoedelijk reden is dat de komkommer haren op het blad heeft en dat de groene perzikluis daar niet in gedijt. De reproductie komt niet op gang en de groene perzikluizen gaan dood. De groene perzikluis is geen echt probleem in de teelt van komkommer. De plaag die in komkommer meer een probleem is, is de katoenluis. Deze katoenluis had beter getest kunnen worden, en de verwachting is dat dit wel VOC-signalen op gaat leveren.

Run 7:

In deze run is in de praktijk in een kas gemeten. Dit is geweest omstreeks de periode van oktober 2024. Het probleem van VOC's meten in deze omstandigheden is dat er ontzettend veel plagen al in de kas aanwezig zijn. Al deze planten en plagen geven VOC's af die zo veel zijn dat het als "ruis" gezien wordt in de metingen. Bepaling wat van wat komt is niet meer mogelijk. Dit staat ook haaks op het concept van "early warning". Om toch een bruikbaar resultaat in de praktijk te realiseren zijn 2 planten schoon gemaakt van luizen, ingepakt als test-plant en referentie-plant en is na een stabilisatie periode is de test-plant voorzien van groene perzikluizen. Het resultaat is dat een aantal VOC's die ook in het lab waargenomen zijn (run 1 & 2) ook in de praktijk run 7 zijn waargenomen.

3.3 Conclusie VOC's

Ethyleen zou in theorie behulpzaam kunnen zijn in de detectie van plagen, omdat aangetoond is dat er hogere ethyleenniveaus zijn waargenomen als er plagen actief zijn op de plant. Echter, deze waarnemingen zijn gedaan in zeer gecontroleerde omstandigheden. In de praktijk zou dit alleen waarneembaar zijn als er een grote aantasting is door plaaginsecten, maar die zou waarschijnlijk al eerder met het oog zijn waargenomen.

Plagen zijn meetbaar in een testopstelling op planten d.m.v. VOC's. Het moeten wel plagen zijn die affiniteit met de plant hebben, zoals probleem-plagen die in de praktijk kunnen voorkomen.

De run 7 in de praktijk kwam in grote lijnen overeen met de run 1 & 2 in het lab.

Het is belangrijk te vermelden dat de tijd, kosten en energie die gestoken moet worden in een onderzoek waarbij het laboratorium aansluit op de praktijk enorm is geweest. In veel bestaande literatuur wordt totaal geen rekening gehouden met praktijkomstandigheden. Real-time metingen (kwalitatief en kwantitatief) zoals hier zijn uitgevoerd in de kas zijn eigenlijk niet bekend. Het maken van de transitie is een niet ontgonnen terrein dat niet eenvoudig is.

Zover we weten zijn dit de allereerste proeven waar real-time kwalitatief en kwantitatief in gemeten is aan plagen. Door de vele 'ruis' in kassen, vanwege veel planten en plagen die aanwezig zijn en samen heel veel VOC's afgeven, kon dit nog niet toegepast worden in de kas en niet ingebouwd worden in het AI model. Er is meer onderzoek in de praktijk nodig waar EMS ook op aanstuurt. Het opbouwen en opschalen van lab-situatie (zoals ook in veel onderzoeken staat beschreven) naar de praktijk kost veel tijd, geld en middelen maar lijkt veelbelovend.

Hoofdstuk 4: Artificiële Intelligentie

4.1 Inleiding

De ontwikkelingen in de tuinbouwsector hebben de laatste jaren een enorme versnelling doorgemaakt doordat er nieuwere sensoren gecombineerd met Artificiële Intelligentie (AI) zijn gekomen. In het Weet Wat Er Leeft-project is er gekeken of deze ontwikkeling ook kan worden toegepast op insectendetectie. Door het samenbrengen van verschillende expertises in sensortechnologie wilde dit project nieuwe vormen van insectenmonitoring mogelijk maken. Doordat deze data soms moeilijk te begrijpen is, kunnen Artificiële Intelligentie (AI) en Data Science deze technologieën beter samenbrengen en presenteren aan gebruikers. Het idee was om de verschillende sensoren die hiervoor besproken zijn samen te voegen in een dashboard waar de tuinder alle informatie die uit die sensoren kwam tegelijk zou kunnen zien.

Dit rapport biedt een uitgebreide inkijk in het proces tijdens het project. In alinea 2 worden de technische ontwerpen besproken. Alinea 3 gaat in op de keuze voor een dataplatform, mede doordat 30MHz failliet ging. Alinea 4 laat de mobiele app zien die op dit moment kan worden gebruikt om foto's van plakplaten te analyseren, maar later mogelijk ook voor andere data kan worden uitgebreid. Alinea 5 gaat in op de huidige softwarearchitectuur en in grote lijnen hoe het systeem achter de app werkt. Alinea 6 bespreekt de twee veldtesten die zijn gedaan met de app. In alinea 7 volgt de discussie en mogelijke aanbevelingen.

4.2 Ontwerp

4.2.1 Inleiding

Het doel van het AI-werkpakket was om diverse sensormodaliteiten samen te voegen, waarbij de sensoren elkaar zouden kunnen versterken. Het gaat hierbij vrijwel niet om lopende insecten, omdat er gewerkt wordt met vangplaten. Lopende insecten kunnen mogelijk wel door een andere sensor, zoals eDNA of de Edapholog, worden gezien. Daarnaast kan eDNA op soortniveau kijken naar de aanwezigheid van insectsoorten, terwijl het met computer vision vaak maar mogelijk is op familieniveau insecten te onderscheiden. In het AI-werkpakket zitten echter geen specialisten voor elke sensor, en ondanks dat er overleg is geweest met de diverse andere werkpakketten, is er gezamenlijk besloten om naar een markerdatarepresentatie te gaan, wat neerkomt op een histogramrepresentatie per insectensoort, waarin relatieve distributies van de gedetecteerde insecten die gemeten zijn, worden gedeeld met de Artificiële Intelligentie. De Artificiële Intelligentie kan deze data dan met elkaar combineren en tot een uiteindelijke voorspelling komen van de plaagdruk in de kas, waarbij rekening wordt gehouden met de locatie en tijd waarop sensormetingen worden gedaan.

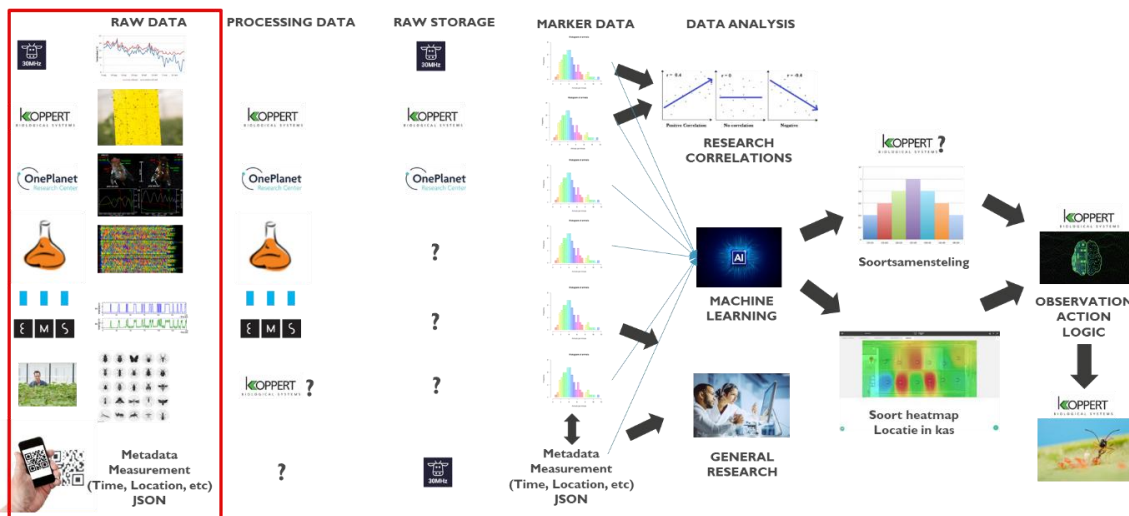
4.2.2 Ontwerp

In het eerste ontwerp (Figuur 4.1), is er gekeken naar de inbreng die de diverse partners kunnen leveren. In de linker twee kolommen zijn de initiële data van de diverse startende partijen zichtbaar: de initiële data, 30MHz voor de klimaatdata, Koppert heeft hun plakplaats analyse app, OnePlanet de hyperspectrale analyses, Normec Groen Agro Control voor de eDNA, EMS voor de VOC-analyse en daarnaast kunnen ook nog telersobservaties worden meegenomen. Door middel van een telefoonapp zou de locatie-informatie en de tijd gelinkt kunnen worden aan een meting (zie onderste rij). Alle sensoren en mogelijk combinaties van sensoren zouden markerdata genereren; dit komt neer op een gestandaardiseerde histogram in JSON-format waarbij per insectensoort of -familie een telling wordt bijgehouden. De machine learning om die telling te leveren wordt (na overleg met de projectpartners) door de diverse partners zelf aangeleverd, omdat daar vaak de kennis van de specifieke sensorgebieden heel belangrijk is. Dit was ook de kennis die de diverse partners (zoals EMS, Normec Groen Agro Control) zelf wilden ontwikkelen. Alleen in de vision is er op het gebied van machine learning nauw samengewerkt tussen de diverse partners: Koppert, HAS, WUR en OnePlanet. Hierbij hebben de diverse partijen niet hun IP afgestaan, maar zijn de diverse oplossingen wel door de HAS vergeleken. De HAS heeft meerdere projecten gedaan waarbij het verzamelen van data (door de HAS) essentiële informatie heeft opgeleverd, zowel voor het verbeteren als het verder ontwikkelen van de sensoren. De markerdata die de hoeveelheid plaagsoorten geeft, kan op verschillende manieren worden gebruikt:

1. Onderzoek naar Correlaties (Research Correlations): Het onderzoeken van correlatie tussen sensoren geeft ons inzicht of de sensor dezelfde trends in plaag/bestrijder populatie kunnen monitoren. Diverse onderzoeken met combinatie van sensoren zijn uitgevoerd in kooien vaak samen met de HAS.
2. Machine Learning: Aan de hand van de output van verschillende sensors kan doormiddel van Machine Learning worden gekeken naar een gecombineerde telling die zwakte van bepaalde sensoren kan meenemen.
3. General Research: Door middel van het verzamelen van data kan er in de verder toekomst ook worden gewerkt aan verder correlaties, bijvoorbeeld de invloed van het weer op bepaalde plaagsoorten

Het idee was dat alle onderzoek uiteindelijk zou leiden tot verschillende visualisaties, met een focus op de soortsaamenstelling in de kas, en de andere kijk naar de plaagdruk in de diverse compartimenten. Als laatste zou Koppert aan de hand van deze resultaten zijn logica en bestrijdingsmiddel inzetten voor de uiteindelijke plaagbestrijding

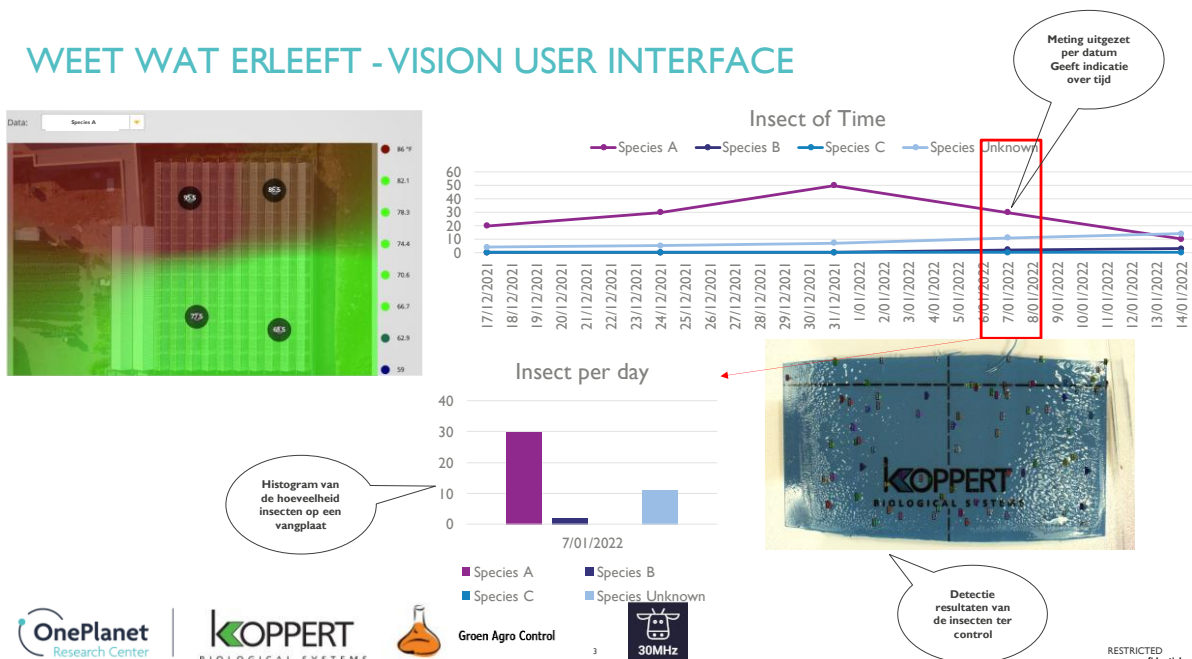
DATASTROMEN DIAGRAM “WEET WAT ER LEEFT”



Figuur 4.1: Het eerste globale ontwerp voor het project waarbij AI de diverse sensoren aan elkaar kan koppelen.

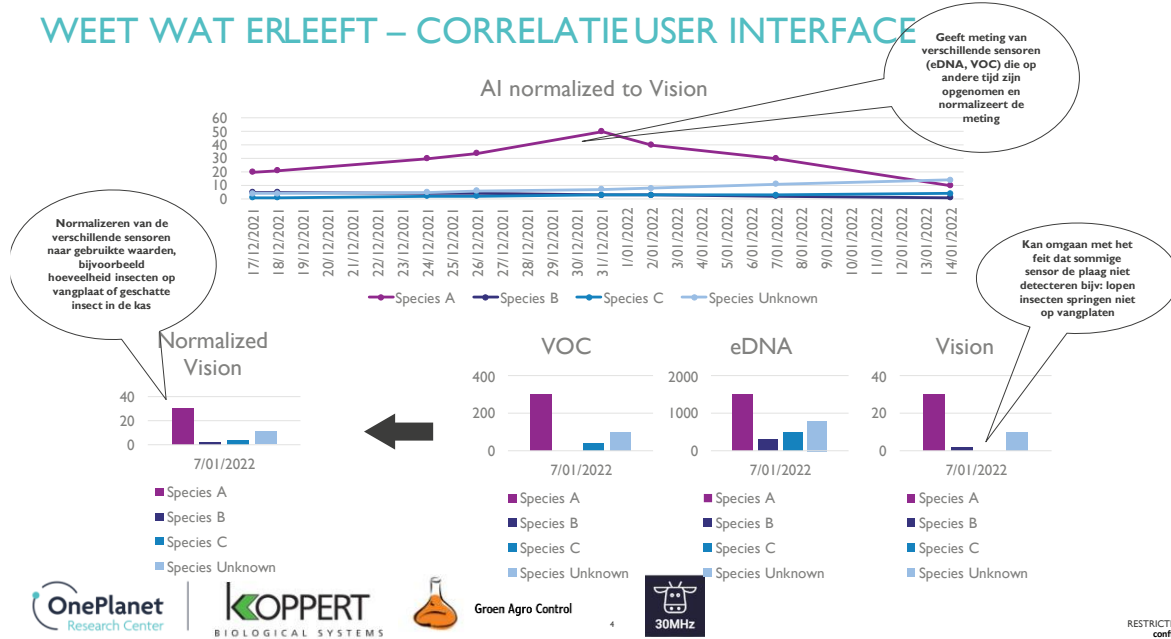
Samen met de partners in het AI werkpakket zijn 3 mockup schermen gemaakt zodat 30MHz een application kon bouwen voor plaagmonitoring in hun platform. De drie mockup schermen zijn te zien in Figuur 4.2-4.4. Figuur 4.2 geeft per sensor de plaagdruk weer, waarbij je in de interface zowel de positie als de tijd van de sensormetingen kan zien en terug kan gaan naar de origineel meting als je deze niet mocht vertrouwen Figuur 4.3 geeft de gecombineerde informatie van alle sensoren weer en hoe deze informatie is opgebouwd en Figuur 4.4 laat de voorspelling van de plaagdruk zien aan de hand van modellen (van Koppert), weer en geschiedenis.

WEET WAT ERLEEFT - VISION USER INTERFACE



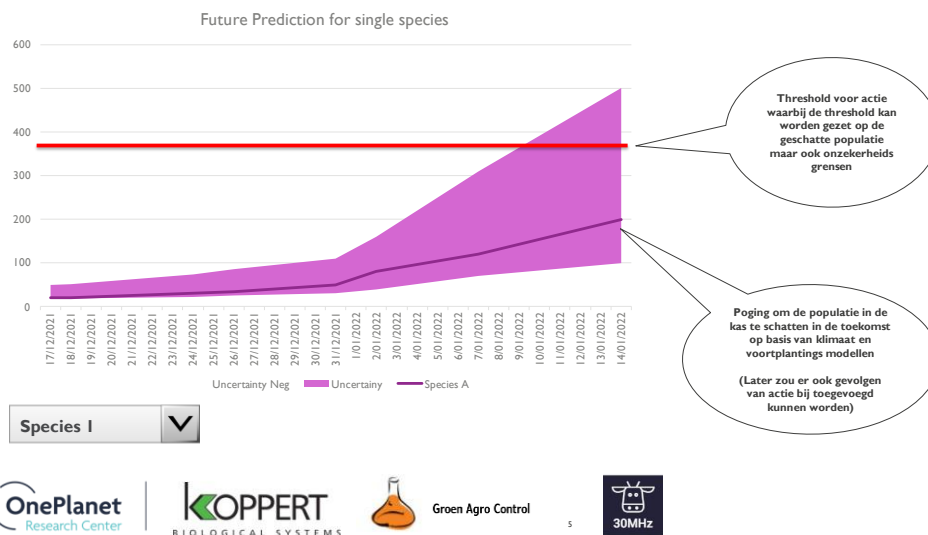
Figuur 4.2: Weet Wat Er Leeft Sensor User Interface Mockup, links een kaart van de kast waarbij de cirkel de sensor posities zijn, door te klikken op een sensor positie kan je de insectplaagdruk krijgen in de figuur recht boven, door op een datum te klikken wordt een histogram met de hoeveelheid insecten op de plakplaat getoond, samen met bijvoorbeeld de foto van de plakplaats als je de automatisch tellen wil controleren.

WEET WAT ERLEEFT – CORRELATIE USER INTERFACE



Figuur 4.3: Weet Wat Er Leeft Correlatie User Interface Mockup, waarbij de bovenste grafiek de plaagdruk per insect vastgesteld door de gecombineerde data van de diverse sensoren. Door te klikken op een dag zou je per sensor verder kunnen achterhalen hoe de data is gecombineerd en welke individuele sensor zijn gebruikt voor die conclusie.

WEET WAT ERLEEFT – VOORSPELLING USER INTERFACE



Figuur 4.4: Weet Wat Er Leeft Voorspelling User Interface Mockup, het idee van deze interface was om aan de hand van de huidige sensor data en de weersvoorspelling ook de plaagdruk verder te gaan voorspellen binnen een onzekerheid marge.

Niet alle interfaces gebouwd

Van de huidige mockup interfaces zijn lang niet alle interfaces gebouwd. De twee voornaamste redenen zijn dat voor de correlatie van diverse sensoren zowel de VOC als

eDNA niet ver genoeg ontwikkeld zijn binnen dit project. Er zijn wel wat positieve resultaten (goede correlatie) gevonden tussen Vision en eDNA, maar deze experimenten konden niet altijd herhaald worden, en tussen Vision en VOC is zelfs nooit onderzoek gedaan. Daarnaast is 30MHz tijdens dit project failliet gegaan en heeft het ons helaas bijna een jaar gekost om Letsgrow.com onder andere voorwaarden bij het project te betrekken. We hebben een werkend systeem op het gebied van Vision kunnen laten zien dat makkelijk te vertalen valt naar andere sensoren. Voor het werk op het gebied van voorspellingsmodellen hebben we echter weinig kunnen doen omdat de beschikbaarheid van data op klimaat in de kas helaas door het faillissement laat beschikbaar kwam.

4.3 Alternatieve Platformen

Door het wegvallen van 30MHz is er gekeken naar alternatieve data platforms.

Data Platform	Voordelen	Nadelen
 GBIF Global Biodiversity Information Facility	+ Officiële standaard voor soorten tellen + Open Standaard	- Informatie is door buitenwereld zichtbaar - Geen klimaat data ondersteuning - Moeilijk op te zetten
 farmmaps	+ GeoPlatform voor alle soorten data verzamelen + Beveiliging per gebruiker + Ondersteuning van App	- Importeren van klimaat data - Vendor Lock in - Zelf bouwen van interfaces
 Open Geospatial Consortium	+ Open Standaard voor alle sensor data + Makkelijke API voor data importeren + Beveiliging per gebruiker wordt gebouwd	- Importeren van klimaat data - Zelf bouwen van interfaces
 LETS GROW .COM	+ Klimaatdata is beschikbaar + API voor data importeren + Beveiliging per gebruiker + Standaard interfaces	- Vendor lock in - Geen open standard (wel open API)

Figuur 4.5: De voor- en nadelen van de diverse dataplatformen voor insecten op een rij gezet

In Figuur 4.5 zijn voor- en nadelen op een rijtje gezet van de volgende platformen:

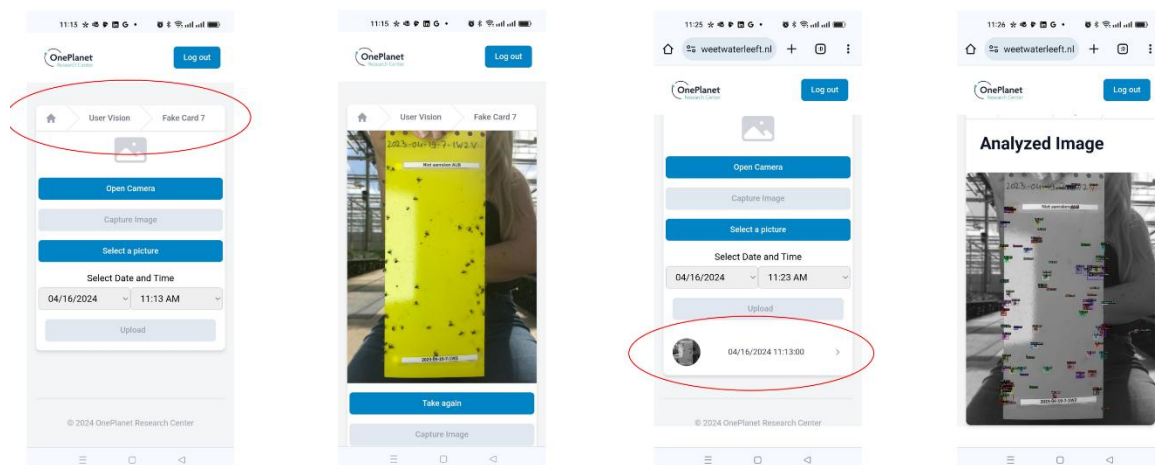
- GBIF (Global Biodiversity Information Facility), niet commercieel: Wetenschappelijke data opslag infrastructuur voor biodiversiteits data. In Nederland wordt deze beheert door Naturalis, de data is gestandaardiseerd en dus open voor iedereen, het eigendom recht blijft bij de data verstrekker.
- Farmmaps, commercieel: Het platform verzamelt, analyseert en distribueert bedrijfsgegevens, biedt gewasmanagementadvies en voert scenario studies en voorspellingen uit. Het verbindt app-ontwikkelaars met gebruikers zoals telers en wetenschappers. De data blijft eigendom van de telers, en gebruikers moeten toestemming krijgen om toegang te krijgen tot deze gegevens.
- OGC Sensorthing API, niet commercieel: Open protocol voor alle sensors (geen vendor lock), die beschikbaar kan worden gemaakt door middel van HTTP API en binnen OnePlanet wordt het ook in andere project gebruikt. Data is niet openbaar, beveiliging per gebruiker is afhankelijk van de implementatie, binnen OnePlanet werd daar al aan gewerkt.

- Letsgrow(.com), commercieel: een dataplatform voor de glastuinbouw waarin klimaatdata vaak al beschikbaar is, zonder dat export nodig is. Ze hebben een API voor ontwikkelaars en mogelijkheden voor nieuwe modules, zoals plaagbestrijding. Het platform heeft bestaande interface-opties, maar kan leiden tot vendor lock-in bij het bouwen van oplossingen.

In overleg met het consortium is uiteindelijk voor Letsgrow.com gekozen, mede omdat ze mogelijk deel van het consortium wilden worden. Er is ook hierover met Farmmaps gesproken, maar dit bedrijf staat verder weg van de glastuinbouw sector.

4.4 Mobiele APP

Binnen dit project hebben we een mobiele app in de vorm van een website ontwikkeld om de sensordata op te sturen, in eerste instantie afbeeldingen van plakplaten, maar later ook andere data zoals beelden, hyperspectrale data, video van Edapholog, DNA-sequenties, VOC-data, etc. De website of app zorgt ervoor dat de sensordata wordt gekoppeld aan een locatie en tijd. De sensordata wordt in de cloud geanalyseerd, waarna de markerdata wordt gegenereerd, wat in de basis een histogram is met de telling van de diverse insecten. Voor vision is dit makkelijk voor te stellen, waarbij het de hoeveelheid insecten op de plakplaat is; voor de Edapholog is het de hoeveelheid insecten die beneden bij de camera op de video worden gezien. Voor eDNA is het de hoeveelheid sequenties die bij een bepaald insect horen.



Figuur 4.6: User Interface van de WeetWatErLeeft App: 1) Het selecteren van een locatie en datum, 2) Het uploaden van de sensor data in dit geval een afbeelding, 3) Het terug ontvangen van de analyse per locatie, 4) Het visualiseren van de analyse door middel van een afbeelding met de gevonden insecten in het geval van een plakplaat of een ander visualisatie voor andere sensoren.

Figuur 4.6 toont de User Interfaces die zijn ontwikkeld binnen dit project. Hierbij hebben wij het login-scherm niet laten zien omdat het redelijk standaard is; hierbij kan je door middel van een e-mail en wachtwoord inloggen. De beveiliging is op dit moment geregeld door OnePlanet, ons initiële idee was om gebruik te maken van Letsgrow-authenticatie, maar dit was nog niet mogelijk gezien hun huidige architectuur. In Figuur 4.6 zie je de verschillende interfaces die zijn gemaakt om:

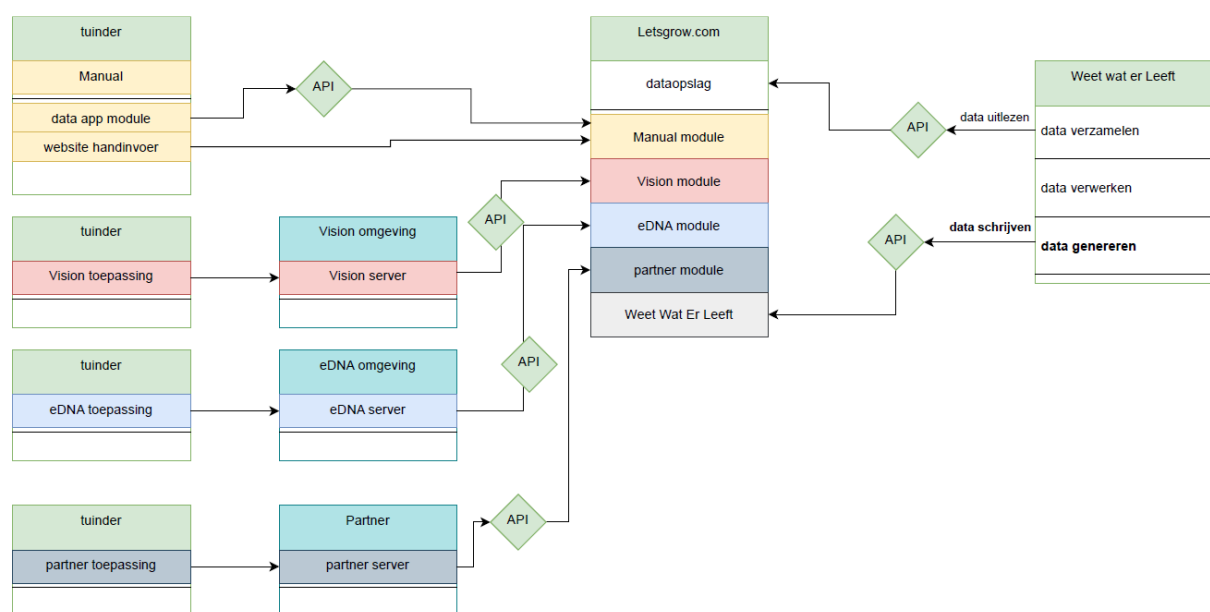
1. Data te voorzien van context rond tijd en plaats van de meting
2. De data zelf naar de cloud te uploaden om het daar te verwerken
3. Het terugontvangen van de analyse als samenvatting
4. Visualisatie voor de gebruiker om te checken dat de automatische analyse juist is gebeurd (waar mogelijk)



Figure 4.7: Letsgrow user interface waarbij per insect of voor alle insecten de verschillende metingen worden weergegeven.

Figuur 4.7 laat de interface van Letsgrow.com zien waar de data naartoe wordt verzonden. Per locatie kan je dan de hoeveelheid van een specifiek insect op de plakplaat volgen, in dit geval wittevlieg. Andere visualisaties zijn in Letsgrow ook mogelijk, die bijvoorbeeld een heatmap geven van de kas met daarop de meest recente plaagdruk.

In de laatste versie van de softwarearchitectuur zijn de volgende keuzes gemaakt (Figuur 4.8). Dit is ontwikkeld door OnePlanet in samenspraak met Letsgrow.com.

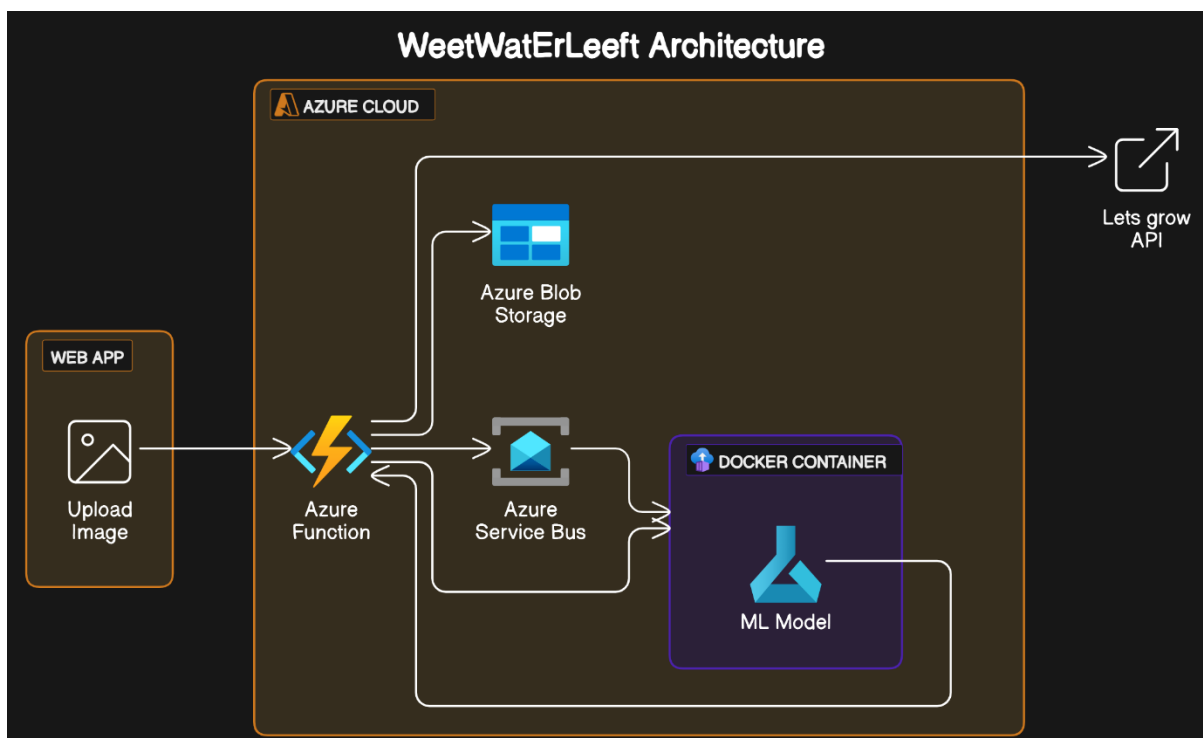


Figuur 4.8: Letsgrow architectuur om verschillende partijen op hun API aan te sluiten

4.5 Integratie

Binnen OnePlanet wordt er gebruik gemaakt van de Microsoft Azure cloud, een andere cloud heeft echter vergelijkbare componenten die kunnen worden gebruikt. In figuur 4.9 is een schematisch weergave van de verschillende componenten te zien. Het begint met de webapp die de data naar de cloud verstuurt en op de Azure Blob Storage zet door middel van een Azure functie. Naast de afbeelding wordt ook de meta-informatie meegegeven over de plaats en tijd waarop de data is gemaakt. De Azure functie communiceert over de Azure Service Bus dat er data klaarstaat, waardoor de machine learning (ML-model) in ons geval de afbeelding analyseert en de insecten detecteert en telt. De analyse wordt teruggezonden op de Azure Blob Storage en door de Azure functie ook naar de Letsgrow API gecommuniceerd. De machine learning (ML-model) draait in Docker, wat ervoor zorgt dat het makkelijk is om deze oplossing te updaten met betere modellen. Daarnaast kunnen andere partners, als ze analysesoftware hebben gemaakt voor hun sensoren, deze oplossing ook makkelijk gebruiken.

Voor het combineren van sensoren was bedacht dat alle sensorinformatie in de Letsgrow API verzameld werd. Wanneer er data bijgevoegd zou worden, zou de Letsgrow API ook kunnen worden gebruikt om de data uit Letsgrow te halen door middel van een Azure functie en meerdere sensoren met elkaar te wegen.



Figuur 4.9: Software Architecture op een hoog niveau hoe de web app en achterliggende systemen werken om uiteindelijk de data in het Letsgrow portal te visualiseren.

4.6. Experiment met HAS

4.6.1 Optimalisatie van gewasscouting

Er is door student van de HAS gekeken naar de optimalisatie van gewasscouting in komkommer, rozen, chrysanten en potorchideeën door middel van geïntegreerde beeldherkenning (waarbij gebruik is gemaakt van de app) en eDNA. Hier is een samenvatting van de belangrijkste punten met betrekking tot automatische beeldherkenning, de opzet van het experiment en de herkenning in de kas, evenals de problemen met machine learning:

Automatische Beeldherkenning

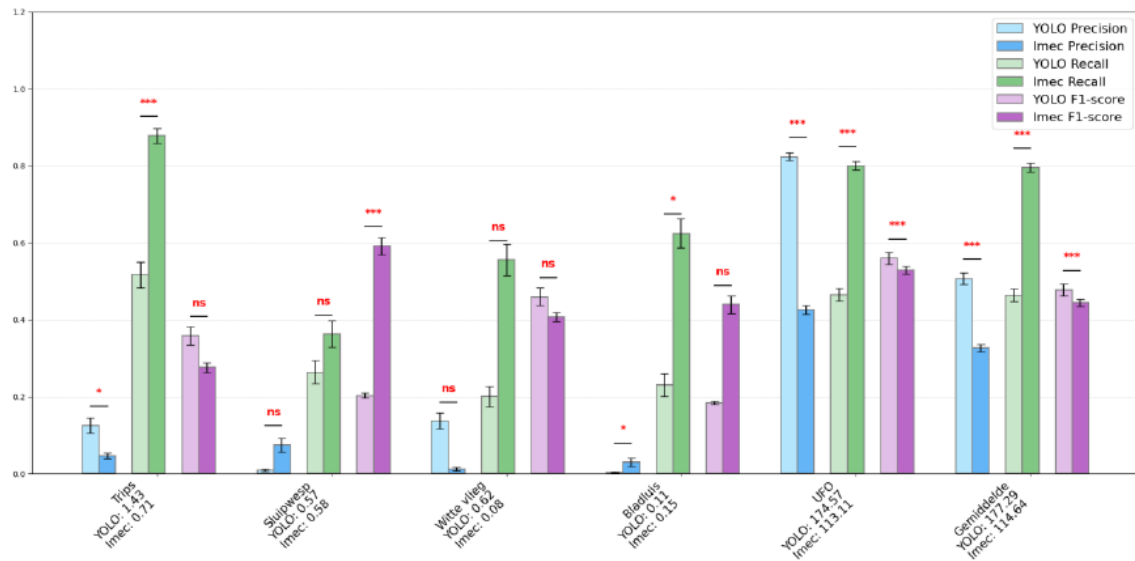
Het deelonderzoek beeldherkenning richtte zich op de recall, precision en F1-score van YOLOv8 model (ontwikkeld door HAS) en vergeleek deze met de waarden van een door Imec getraind model. De prestaties van deze modellen werden verder vergeleken met het commercieel beschikbare beeldherkenningsmodel Natutec van Koppert. De invloed van een externe flits op de prestaties van het model werd ook getest (Zie 1.2).

Opzet van het Experiment

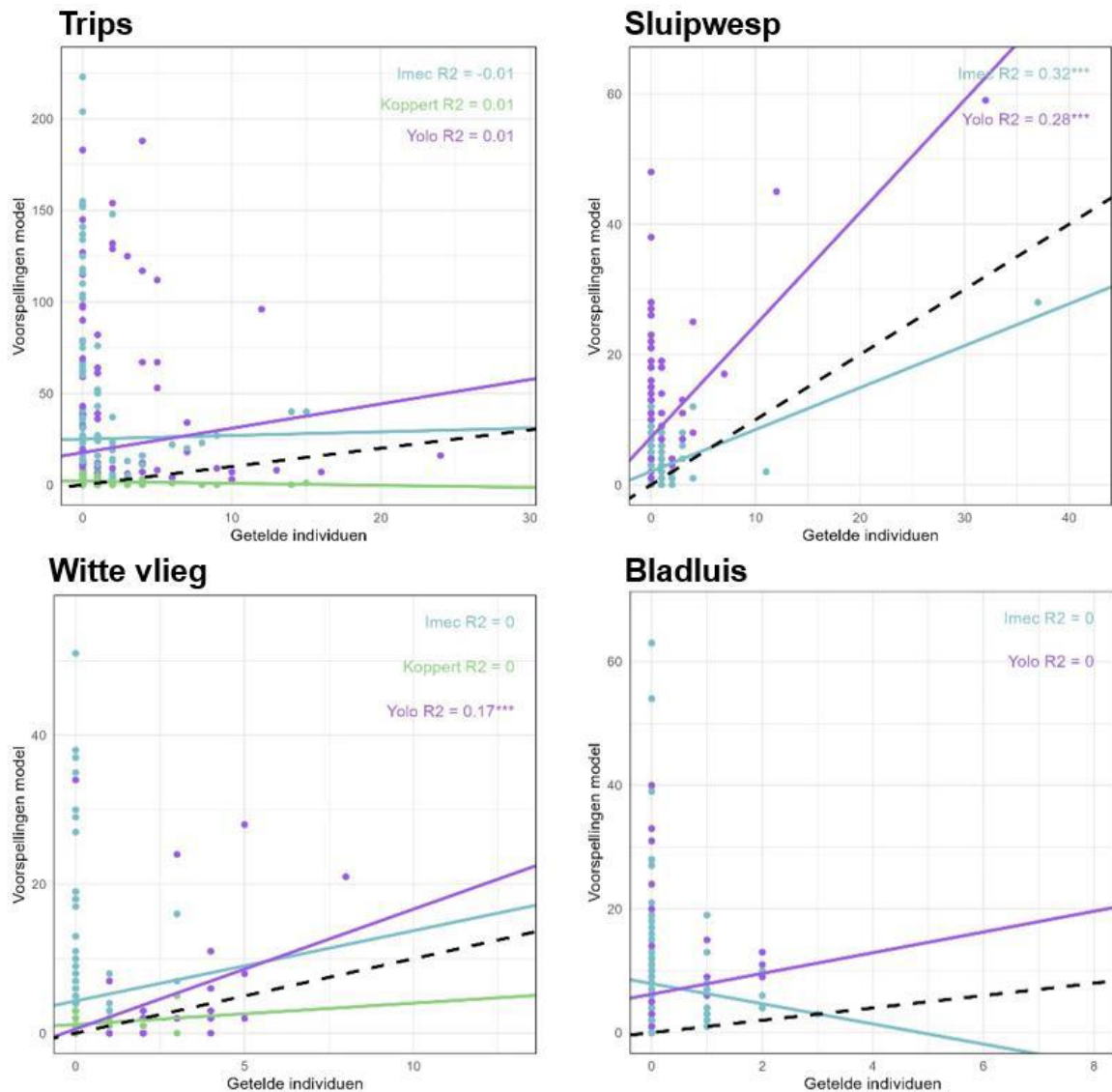
Het onderzoek bestond uit twee deelonderzoeken: automatische beeldherkenning en de samplingmethoden van eDNA. De experimenten werden uitgevoerd in vier teelten: komkommer, rozen, chrysanten en potorchideeën. Voor de beeldherkenning werden gele droge vangplaten gebruikt die vijf centimeter boven het gewas werden opgehangen. Na zeven dagen werden deze vangplaten gefotografeerd met een iPhone 7, zowel met als zonder externe flits. De foto's werden vervolgens geanalyseerd door verschillende beeldherkenningsmodellen.

Herkenning in de Kas en Problemen met Machine Learning

De resultaten toonden aan dat de betrouwbaarheid van zowel eDNA als beeldherkenning sterk afhankelijk is van de aantallen aanwezige geleedpotigen. Het automatische beeldherkenningsmodel YOLOv8 en het model van Imec presteerden variabel, met inconsistenties in precision, recall en F1-score, voornamelijk door lage aantallen doelsoorten op de vangplaten (Figuur 4.10). Ook lieten de drie modellen zeer beperkt correlatie zien (Figuur 4.11). Waarbij bij sommige teeltsystemen de correlatie hoopgevend zijn, zijn er bij andere toch nog veel problemen (Figuur 4.12).



Figuur 4.10: Gemiddelde precision, recall en F1-score met flits van YOLOv8 en het model van Imec in de vier gewassen: komkommer, rozen, chrysanten en potorchidee. Het gemiddelde van alle soorten is berekend aan de hand van het totaal aantal individuen per plaat. De getallen onder de labels op de x-as geven het gemiddelde aantal soorten per plaat weer. De foutbalk geeft de standaardfout weer. Significantie is weergegeven met '*', '**' of '***' voor een respectievelijke p-waarde kleiner dan 0.05, 0.01 en 0.001.



Figuur 4.11: De lineaire regressie voor YOLOv8 (Paarse lijn en R2), het model van Imec (blauwe lijn en R2) en Natutec (groene lijn en R2) op basis van de individu aantallen met flits. De getallen op de x-as geven daarbij aantal getelde individuen per plaat weer. Op de y-as staat het aantal voorspelde individuen per plaat van de beeldherkenning modellen. De zwarte gestreepte lijnen in de figuren geven de ideale lijn weer waarbij de formule $y=x$ is. De R-squared (R2) staat per model weergegeven met daarachter de bijbehorende significantieniveaus '*', '**' of '***' voor een respectievelijke p-waarde kleiner dan 0.05, 0.01 en 0.001.

	K : TR	K : WV	R : WV	C : TR	C : SL	C : WV	T : SL	T : WV
Natutec	0,02	0,04	0,00	-0,02	n/a	n/a	n/a	0,00
Yolo	0.67***	0.15**	0.38***	0,07	0.64***	0.26***	0.28***	0.17***
Imec	0.68***	-0,02	0,01	0.36***	0.50***	n/a	0.32***	0,00

Figuur 4.12: De R2 voor Natutec, YOLOv8 en het model van Imec per combinatie tussen gewas en doelsoort. Hierbij staat de eerste letter voor het gewas, 'K' voor komkommer, 'R' voor roos, 'C' voor chrysant en 'T' voor een totaal van de vier geteste gewassen uit dit onderzoek. De twee letters achter het gewas staan voor de doelsoort, 'TR' voor trips, 'SL' voor sluipwesp en 'WV' voor wittevlieg. Achter de R2 is het significantieniveau aangegeven met '*', '**' of '***' voor een respectievelijke p-waarde kleiner dan 0.05, 0.01 en 0.001. Voor de combinaties waar niet genoeg data beschikbaar is voor een regressieanalyse, is 'n/a' genoteerd.

Hoewel de innovatieve technieken potentie tonen, zijn ze nog niet ver genoeg ontwikkeld om conventionele methoden volledig te vervangen. Bij deze experimenten, zijn modellen gebruikt met trainingsvoorbeelden die vooral in het lab zijn gemaakt en niet in de kas. Dit leverde veel problemen op waarbij verdere training van beeldherkenningsmodellen met diverse vangplaat geleedpotigen samenstellingen nodig zijn om de betrouwbaarheid te verbeteren.

4.7 Discussie AI

Binnen het Weet Wat Er Leeft project hebben we laten zien dat het mogelijk is om een platform te ontwikkelen voor sensoren waarbij de sensordata met Artificiële Intelligentie wordt geanalyseerd en de data beschikbaar wordt voor de gebruiker. Het is echter niet gelukt om sensortechnologieën te combineren, doordat veel experimentele sensoren veel problemen ondervonden en niet tijdig klaar waren voor gecombineerde experimenten. Wel beschrijven we de visie over hoe de sensoren mogelijk gecombineerd kunnen worden in ons ontwerp. We laten zien dat het belangrijk is een dataplatform te kiezen waarin de data op een gestandaardiseerde manier kan worden opgeslagen. Hiervoor geven we de keuzeoverwegingen, waarna mede door de samenwerking met Letsgrow.com, er voor hun platform is gekozen. Op het gebied van vision (automatische plakplaatherkenning) laten we zien dat onze visie gerealiseerd kan worden door het maken van een app waarbij de data kan worden verwerkt in de cloud. We geven hiervoor een softwarearchitectuur die ook hergebruikt kan worden voor andere data. Daarnaast laten we experimenten zien die gedaan zijn met de app. Hierbij is echter wel de conclusie dat er meer werk moet worden gestoken in de dataverzameling. Dit komt deels omdat eerdere dataverzameling in labs is gedaan en dit een van de eerste keren is dat de data in kassen grootschalig verzameld is. Ten tweede zijn er andere omstandigheden in de kassen waar geen rekening mee is gehouden (voorbeelden zijn vangplaten vol huisvliegen, veel meer achtergrond bij de foto's, etc).

Aanbevelingen:

- Dit project heeft laten zien dat het maken van een softwareoplossing mogelijk is, er is echter ook een gebruiksdrempel en een prijs waartegen deze oplossingen kunnen worden gerealiseerd. De oplossing die wij hebben gemaakt op het gebied van vision is misschien niet uniek, maar we zien ook dat de business case niet erg makkelijk is voor een softwarebedrijf om dit te onderhouden. Er moet vanuit de sector worden nagedacht of het de moeite waard is om hierin als sector gezamenlijk in te investeren of dit aan bedrijven over te laten.
- Artificiële Intelligentie en Data Science hebben meestal veel data nodig om tot goede resultaten te komen. In dit project was voor meerdere sensoren deze data niet aanwezig zodat er binnen deze werkgroep moeilijk kon worden begonnen met dit deel van het project. Het zou een overweging zijn om projecten op het gebied van Data Science strikter te beoordelen op de haalbaarheid en de aanwezigheid van

data. Als de data bij sommige sensoren echter wel tijdig was gekomen, is het voordeel dat er in een eerdere fase is nagedacht over het dataontwerp.

- Glastuinbouw Nederland heeft door de app die is ontwikkeld een unieke kans om de app uit te breiden en systematisch plakplaten te gaan fotograferen. Dit zou ook kunnen door van oude smartphones een soort automatisch monitoring device te maken die constant naar een plakplaat kijkt. Zo kunnen telers zelf een unieke database van insecten opbouwen en informatie uitwisselen op het gebied van (early) detectie van plagen. Hierbij moet echter nog meer worden geïnvesteerd in deze infrastructuur (zie aanbeveling 1).

Het standaardiseren van de data op het gebied van plaagdata kan leiden tot platforminteroperabiliteit. Dit betekent dat je niet meer aparte oplossingen hoeft te maken per dataplatform dat je gebruikt. Als teler kan je makkelijk naar een ander platform switchen zonder dat je de data verliest.

Hoofdstuk 5: Algemene conclusie en discussie

Het project *Weet wat er leeft* is erin geslaagd een platform te creëren waarin verschillende datastromen uit een kas samenkomen. Echter, het is niet gelukt om deze data effectief te benutten voor een geoptimaliseerde soortidentificatie, een beter inzicht in de verspreiding van soorten over ruimte en tijd, of het automatisch herkennen van patronen met deep learning. Behalve de vision-datastroom bevonden de andere datastromen, VOC en eDNA, zich nog niet in een voldoende vergevorderd technologisch stadium om succesvol in het platform te worden geïntegreerd. Desondanks heeft het project significant bijgedragen aan de ontwikkeling van beide technologieën waarmee een toekomstige toepassing dichterbij is gebracht. Er is een demonstrator gemaakt met het data platform van Letsgrow.com. Daarnaast is er een mobiele app gemaakt door OnePlanet research center waarbij het mogelijk is trips, wittevlieg, Aphidius en bladluizen op plakplaten te fotograferen en met behulp van AI die foto's te analyseren en de resultaten in Letsgrow.com te tonen.

Hieronder wordt per detectiemethoden beschreven wat er is geleerd en wat het project aan de ontwikkeling van deze techniek heeft bijgedragen.

Vision

Het project heeft een methodologie ontwikkeld waarmee classificatie van insecten op plakplaten in een kas mogelijk is met de camera van een mobiele telefoon. Met deze methode kan een teler zonder veel bijkomende kosten op een eenvoudige manier het scouten van plakplaten automatiseren. Met name voor trips was de techniek in staat om populatietrends in een kas te registreren. Het gebruik van een externe flitser en gele of blauwe vangplaten met droge lijm leverde de beste resultaten op. Voor de uitdaging van het gebruik van led-kleurverlichting heeft het project passende oplossingen gevonden. Er is een Technology Readiness Levels (TRL) bereikt van 7 (Demonstratie prototype in operationele omgeving). Om de technologie verder te ontwikkelen zijn meer foto's uit de praktijk nodig om het beeldherkenningsalgoritme mee te trainen.

Het project heeft laten zien dat nauwkeurige en detectie en classificatie van insecten mogelijk is met een hyperspectraal camera en heeft laten zien hoe de complexiteit en kosten van deze technologie kunnen worden verminderd door gebruik te maken van multispectrale en RGB-gegevens. De Technology Readiness Level van dit onderdeel is 5 omdat er validatie met echte plakplaten is geweest maar niet in praktijk bedrijven.

Wat betreft het monitoren van roofmijten op het blad is er verkennend onderzoek gedaan naar het gebruik van een sensor die ontwikkeld is voor het automatisch herkennen van mijten in bodemonsters (Edapholog). Voor het monitoren van roofmijten presteerde deze sensor beduidend minder goed dan de nu gangbare spoelmonsters. Deze sensor was wel heel geschikt om tripslarven te monitoren, door het uitbroeden van tripseitjes in het blad. Het bereikte TRL level wat dit onderdeel betreft is 4 (Implementatie en test prototype).

eDNA

Het project heeft aangetoond dat met behulp van eDNA detectie met PCR het onder gecontroleerde omstandigheden mogelijk is een betrouwbare schatting te krijgen van de soorten op een plakplaat en aantal individuen per soort. Echter, in een praktijksituatie gaf deze methode geen betrouwbaar resultaat meer. Eén van de redenen hiervoor is dat in een praktijksituatie DNA van een enkele soort de plakplaat kan domineren, waardoor DNA van de andere soorten niet meer wordt geamplificeerd. Het bereikte TRL van deze techniek is 5 (Validatie prototype).

Het project heeft ook aangetoond dat eDNA detectie met de shotgun methode niet geschikt is voor de doelsoorten van het project. Geen van de soorten werd consistent in praktijksituaties teruggevonden terwijl deze wel in de kas aanwezig waren. Het bereikte TRL van deze techniek is 5 (Validatie prototype).

VOC

Meerdere plagen zijn meetbaar met ethyleenmetingen en in een testopstelling op paprika en komkommer planten in gecontroleerde omstandigheden. Het opschalen van lab-situatie naar de praktijk kost veel tijd en is een grotere stap dan ingeschat op basis van literatuur. Dit heeft te maken met het feit dat in een kas veel planten en plagen zijn die samen heel veel VOC's afgeven die zo veel zijn dat het als "ruis" gezien wordt in de metingen. Bepaling wat van wat komt is daardoor niet goed mogelijk. Het bereikte TRL level wat dit onderdeel betreft is daarom 4 (Implementatie en test prototype).

Oorspronkelijke doelen en de mate waarin deze zijn gerealiseerd

1. Het precies kunnen detecteren van drie plaagorganismen: Californische trips, groene perzikluis en kaswittevlieg.
 - Met Vision is het mogelijk deze drie soorten op groepsniveau te detecteren (trips, bladluis en wittevlieg). Met DNA methoden was het mogelijk de groene perzikluis en kaswittevlieg op soort te determineren. Dit werkte niet voor de Californische trips. Met VOC was het mogelijk trips en bladluizen te detecteren, wittevlieg niet.
2. Het precies kunnen detecteren van de drie biologische bestrijders: *Aphidius colemani*, *Amblyseius swirskii* en *Transeius montdorensis*.
 - Met Vision kunnen *A. colemani* en *A. swirskii* en *T. montdorensis* op groepsniveau worden gedetecteerd (sluipwesp, roofmijt). Met de huidige eDNA methoden was het niet mogelijk om de bestrijders betrouwbaar te identificeren. Met VOC's was het mogelijk indirect sluipwespen te detecteren, door stress bij de plaag.
3. Inzicht in de haalbaarheid van de technologieën: vision, eDNA en VOC voor detectie van de doelorganismen.
 - De resultaten van dit onderdeel zijn terug te vinden in het huidige rapport.

4. Het ontwikkelen van API's voor communicatie van detectiemethodieken, omgevingssensoren (klimaat) en andere hardware met een kunstmatige intelligentie platform.
 - Voor de Vision techniek is er een API ontwikkelt. Voor de andere technieken niet.
5. Het ontwikkelen van een deep learning kunstmatige Intelligentieplatform voor het optimaliseren van de (gevoeligheid van) detectie.
 - Voor Vision wel, andere sensoren niet. De gevoeligheid kan alleen toenemen als er meerdere sensoren worden gekoppeld.
6. Het automatisch detecteren, identificeren en kwantificeren van verschillende soorten in een mixpopulatie door integratie van state-of-the-art technieken middels artificial intelligence.
 - Voor Vision en eDNA is er een vergelijking gemaakt. eDNA blijkt echter niet zo goed te presteren in een praktijk omgeving als verwacht. Met VOC is geen koppeling gemaakt.
7. Een geïntegreerde methode die geschikt is om te worden toegepast op een mixpopulatie van de drie doelpLAGen en biologische bestrijders;
 - In principe kan Vision werken voor een mixpopulatie, maar integratie met andere technieken is zeker nog niet compleet, maar in theorie wel mogelijk.
8. Het toepassen van het Kunstmatige Intelligentie systeem met detectiemethoden en klimaatsensoren op praktijkbedrijven in de gewassen: roos, chrysant, komkommer en potorchidee.
 - Voor vision is dit eenmaal gelukt, voor de andere sensoren niet.
9. Een eerste inzicht in de dynamiek van populaties in ruimte en tijd via big dataverwerking.
 - Niet gerealiseerd.

Eindconclusie

Concluderend heeft het project "Weet wat er Leeft" een basis gelegd voor toekomstige vooruitgang in het detecteren en monitoren van plagen in kassen. Het heeft het potentieel van het combineren van verschillende detectiemethoden met AI benadrukt, maar ook de uitdagingen die moeten worden aangepakt voordat deze technologieën in de praktijk kunnen worden toegepast. De bevindingen en aanbevelingen van het project bieden waardevolle richtlijnen voor toekomstig onderzoek en ontwikkeling op dit gebied.

Literatuurlijst

- Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19* (pp. 424-432). Springer International Publishing.
- Dicke, M. Herbivore-induced plant volatiles as a rich source of information for arthropod predators: Fundamental and applied aspects. *J. Indian Inst. Sci.* 2015, 95, 35–42.
- Dombos, M., Kosztolányi A., Szilávecz, K., Gedeon, C., Flórián, N., Groó, Z., Dudás, P., Bánszegi, O. (2016). EDAPHOLOG monitoring system: automatic, real-time detection of soil microarthropods. *Methods in Ecology and Evolution* 8(3), 313-321.
- MacDougall, S., Bayansal, F., & Ahmadi, A. (2022). Emerging methods of monitoring volatile organic compounds for detection of plant pests and disease. *Biosensors*, 12(4), 239.
- Poyntz-Wright, I. P., Harrison, X. A., Pedersen, S., & Tyler, C. R. (2024). Effectiveness of eDNA for monitoring riverine macroinvertebrates. *Science of The Total Environment*, 173621.
- Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6), 1137-1149.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer International Publishing.
- Thireou, T., & Reczko, M. (2007). Bidirectional long short-term memory networks for predicting the subcellular localization of eukaryotic proteins. *IEEE/ACM transactions on computational biology and bioinformatics*, 4(3), 441-446.

Bijlagen

Alle studentrapporten zijn te vinden op de website van Glastuinbouw Nederland. Het betreft de volgende studentenrapporten:

Studentenrapporten

Vision:

- Beekhuizen, C., Boer, J., Nissen, K. & Op 't Veld, W. (2024). Onderzoek naar het vangen van mijten, trips en andere geleedpotigen met behulp van de Edapholog. Studentenverslag HAS green academy.
- Benschop, J., Heijmen, H. & van der Kuijl, C. (2023). EDNA reinforced photo-scouting. Exploring insect scouting methods using eDNA and machine vision. Studentenverslag HAS green academy.
- van Boxtel, S., Botman, S. & Smit, L. (2024). Automatische beeldherkenning van plaaginsecten bij verschillende lichtomstandigheden. Het effect van witlicht, paars LED licht, paars LED licht met witflits, kleurcorrectie en AI-correctie op automatische beeldherkenning van doelinsecten op gele plakplaten. Studentenverslag HAS green academy.
- Dezutter, V., van Eekelen, W., Fletcher, E., Mutsers, K. & Roovers, J. (2022). Automatische beeldherkenning van (plaag)insecten in de glastuinbouw. Onderzoek naar de verschillende automatische beeldherkenningssystemen voor plaaginsecten, de meningen/wensen van telers omtrent deze systemen en het bepalen van de correcte resolutie en afstand van een mobiele telefoon ten opzichte van vangplaten voor het trainen van Azure Custom Vision. Studentenverslag HAS green academy.
- van den Einde, T., ter Heerdt, A. & Leenaars, L. (2023). Automatische beeldherkenning en eDNA als monitoringstechnieken voor plagen en biologische bestrijders in de kas. Studentenverslag HAS green academy.
- Helvaci, B., Heijmen, H. & Seelenman R.-J. (2022). Fotoscouting: heden of toekomst? Een uitgebreide studie naar de methoden en validiteit van automatische beeldherkenning van plaaginsecten en biologische bestrijders in kassen. Studentenverslag HAS green academy.
- Houtmans, A., Geurts, I., Veenema, J.P., Haaring, N., Barten, S. & Stals, T. (2024). Optimalisatie van gewasscouting in komkommer, rozen, chrysanten en potorchideeën door geïntegreerde beeldherkenning en eDNA. Studentenverslag HAS green academy.
- Janssen, G., Kramer, S. & Barten, S. (2023). Insectherkenning. Onderzoek naar het verbeteren van een insectherkenningsmodel in Custom Vision. Studentenverslag HAS green academy.
- Lisi, M., Ruizendaal, E. & de Waard, M. (2022). Herkenning van natuurlijke bestrijders door middel van artificial intelligence. Studentenverslag HAS green academy.

- Nijhuis, J., Kapteijns, K. & van der Heide, P. (2023). Herkenning van plaaginsecten op diverse plakplaten met artificial intelligence. Studentenverslag HAS green academy.
- van der Spek, C., Walravens, J. & Mol, M. (2022). Herkenning van plaaginsecten door middel van machine learning. Studentenverslag HAS green academy.
- Verkleij, R., de Kloe, L. & Knijf, N. (2023). Is automatische beeldherkenning de oplossing voor plaaginsecten? Efficiënte monitoring en bestrijding van plaaginsecten in de glastuinbouw: het potentieel van automatische beeldherkenning. Studentenverslag HAS green academy.
- Verkleij, R. (2024). Van pixel tot innovatieve monitoring: factoren van invloed op de performance van automatische beeldherkenningssystemen. Een literatuurstudie en meta-analyse over de factoren die invloed hebben op de performance van automatische beeldherkenning van (plaag)insecten in de glastuinbouw. Studentenverslag HAS green academy.

eDNA:

- Benschop, J., Heijmen, H. & van der Kuijl, C. (2023). EDNA reinforced photo-scouting. Exploring insect scouting methods using eDNA and machine vision. Studentenverslag HAS green academy.
- van den Berg, R. & Groot, N. (2023) eDNA swabbing as an early detection method in greenhouses. Using eDNA swabbing to monitor common greenhouse pests and natural enemies. Studentenverslag HAS green academy.
- van den Einde, T., ter Heerdt, A. & Leenaars, L. (2023). Automatische beeldherkenning en eDNA als monitoringstechnieken voor plagen en biologische bestrijders in de kas. Studentenverslag HAS green academy.
- Houtmans, A., Geurts, I., Veenema, J.P., Haaring, N., Barten, S. & Stals, T. (2024). Optimalisatie van gewasscouting in komkommer, rozen, chrysanten en potorchideeën door geïntegreerde beeldherkenning en eDNA. Studentenverslag HAS green academy.



De Stichting Control in Food & Flowers voert onderzoek uit op het gebied van agrarische productie, voeding en hieraan gerelateerde biotechnologie. De Stichting heeft als doel het bevorderen van innovatieve technologische kennis op het gebied van productie en kwaliteit van levensmiddelen en agrarische producten in de sector. Dit vindt plaats door het uitvoeren van onderzoek en ontwikkeling, samenwerken met andere organisaties, bevorderen van technologische kennis, kennisoverdracht, voorlichting en wetenschappelijke publicaties.

Stichting Control in Food & Flowers

Distributieweg 1

2645 EG Delfgauw

T: +31(0)15-2858124

E: info@scff.nl

www.scff.nl

KvK: 61916471