Title

Automated Varroa Mite Detection Using Deep Learning and Data Augmentation

Ronald Li

Dr. Garcia-Diaz

Dr. Farhad Maleki

Dr. Farnaz Sheikhi

Daily Notes

August 29, 2025

Summary:

- Lecture on ASP expectations and goals

Notes:

Science fair:

- 9-10 ASP students go to CYSF

Important ASP skills:

- Time management
    - Track the research being done(take notes on papers, textbooks, etc.)
    - Progress
    - Daily notes(lectures, guidelines, etc.)
    - Calendar for tasks in this class(this month, future month plans)
        - Outline when I have ASP in that month, and due dates for each class, specific tasks->self accountability
- Communication
    - CC teachers with mentor communication
    - Fast response to communications, check multiple times per day, less than one day response to mentors
    - Personal emails communication to mentors, cc to teachers
    - Every meeting with mentors, send an email to the mentors afterwards with a summary of what the meeting was about.
        - What did we discuss?
        - What tasks do I have, and with what due dates?
        - When is our next meeting?
        - Did I miss anything?
    - Meet with teachers approximately weekly to discuss what is going on
        - Bring summary notes
    - Communication with schools when I will be off campus
- Preparedness
    - Proactive
- Mentors evaluating me on:
    - Showing up for meetings
    - Communication
    - Progress in research project
    - Preparedness for meetings
- Typical ASP class:
    - Mentor meetings
    - Lecture or meeting with teachers for a status update
        - Priority goes to ASP work and mentor meetings

    - Work period for ASP or catch up
    - Going to university
        - Email school ahead of time
- Note: try not to skip other school classes to visit the university
- Evaluation:
    - Schedule, communication, organization: 20% of overall mark(preparedness, on track progress, etc.)
    - Log book: end of month evaluation(10% each evaluation for 20% of overall mark)
    - Oral presentations: 30%
        - 4 in-class presentations(final presentation is 50% of 30% overall mark)
    - Written work: 30%
        - Research proposal
        - Introduction section
        - Experimental procedures
        - Results
        - Analysis
        - Combined final paper
- Homework for Wednesday September 3
    - Email summary of status update
        - Met with a mentor?
        - Discussed project topic?
        - Have a topic?
        - Started reading literature?
        - Established meeting times?
    - Reflect on what log book style works best
    - Email mentor about school starting, research proposal, schedule meeting
        - Aim to set up weekly meetings
        - Aim to have weekly communications
- Next class:
    - Log book
    - Research proposal
        - Background
        - Goals
        - Questions
        - Methods

Sept 3, 2025

Summary:

- Log book and research proposal expectations

Log book:

- Title page
    - Title, name, mentor name
- Calendar by month
    - Daily tasks
    - Monthly goals
- Daily notes(with dates)
    - Note: other work should be dated for cross referencing
    - Neutral tone
- Background research
    - Notes for each article
- Additional:
    - Data collection & results
    - Methodology
    - Conclusion
    - Discussion
    - Other

Written research proposal(October 16 deadline):

- Is the project a study, experimental project, or innovation?
    - Using variables, hypotheses, observations?
    - Innovation:
        - What is the problem?
        - What is the goal?
        - What do we need to improve?
- The goal is to find out why we want to research this?
- Will this idea get funding?
- Idea needs to be narrow and achievable
- Working title needs to be specific, yet still understandable
- Right now, the introduction section will be most highly weighted.
- Finish introduction with “therefore, this is our research question”
- Goals:
    - Short term goals(year end goals, must be achievable)
    - Long term goals(e.g. Curing cancer
- Variables if needed, not for innovation, studies
- Methodology
    - How/why does this technique help me answer my question
- Maybe around 3-5 pages?

- Background research should be extensive with a lot of literature sources
    - Know a lot about the topic and surrounding literature
- Significance
    - Why we need this project
    - Summary paragraph(s) at the end of the proposal
    - How does this contribute and is significant to the field of study

September goals:

- Find project topic
    - Narrow down scope of ideas

September 5, 2025

Daily notes:

Free block work(Right now work toward a research proposal):

- Logbook work, organization, planning
- Email mentor, meeting with mentor
- Read paper
- Outline research proposal
    - Research question
    - Intro
    - Methods
- Submit research proposal to mentor a week before October 16 due date for editing
- Create a mentor meeting plan to discuss:
    - Regular meetings(day 6 cycle, after school, weekly/monthly)
        - How are we planning this
    - Research proposal assignment
        - Topic
        - Research question
        - Goals
        - Outline intro
        - Background research
            - IEEE for engineering
            - Start with review papers
- Tuesday->setup PaperPile

September Notes

29 August 2025

29 August 2025

Summary:

- Lecture on ASP expectations and goals

Notes:

Science fair:

- 9-10 ASP students go to CYSF

Important ASP skills:

- Time management
    - Track the research being done(take notes on papers, textbooks, etc.)
    - Progress
    - Daily notes(lectures, guidelines, etc.)
    - Calendar for tasks in this class(this month, future month plans)
        - Outline when I have ASP in that month, and due dates for each class, specific tasks->self accountability
- Communication
    - CC teachers with mentor communication
    - Fast response to communications, check multiple times per day, less than one day response to mentors
    - Personal emails communication to mentors, cc to teachers
    - Every meeting with mentors, send an email to the mentors afterwards with a summary of what the meeting was about.
        - What did we discuss?
        - What tasks do I have, and with what due dates?
        - When is our next meeting?
        - Did I miss anything?
    - Meet with teachers approximately weekly to discuss what is going on
        - Bring summary notes
    - Communication with schools when I will be off campus
- Preparedness
    - Proactive
- Mentors evaluating me on:
    - Showing up for meetings
    - Communication
    - Progress in research project
    - Preparedness for meetings
- Typical ASP class:
    - Mentor meetings
    - Lecture or meeting with teachers for a status update
        - Priority goes to ASP work and mentor meetings

    - Work period for ASP or catch up
    - Going to university
        - Email school ahead of time
- Note: try not to skip other school classes to visit the university
- Evaluation:
    - Schedule, communication, organization: 20% of overall mark(preparedness, on track progress, etc.)
    - Log book: end of month evaluation(10% each evaluation for 20% of overall mark)
    - Oral presentations: 30%
        - 4 in-class presentations(final presentation is 50% of 30% overall mark)
    - Written work: 30%
        - Research proposal
        - Introduction section
        - Experimental procedures
        - Results
        - Analysis
        - Combined final paper
- Homework for Wednesday September 3
    - Email summary of status update
        - Met with a mentor?
        - Discussed project topic?
        - Have a topic?
        - Started reading literature?
        - Established meeting times?
    - Reflect on what log book style works best
    - Email mentor about school starting, research proposal, schedule meeting
        - Aim to set up weekly meetings
        - Aim to have weekly communications
- Next class:
    - Log book
    - Research proposal
        - Background
        - Goals
        - Questions
        - Methods

3 September 2025

3 September 2025

Summary:

- Log book and research proposal expectations

Log book:

- Title page
    - Title, name, mentor name
- Calendar by month
    - Daily tasks
    - Monthly goals
- Daily notes(with dates)
    - Note: other work should be dated for cross referencing
    - Neutral tone
- Background research
    - Notes for each article
- Additional:
    - Data collection & results
    - Methodology
    - Conclusion
    - Discussion
    - Other

Written research proposal(October 16 deadline):

- Is the project a study, experimental project, or innovation?
    - Using variables, hypotheses, observations?
    - Innovation:
        - What is the problem?
        - What is the goal?
        - What do we need to improve?
- The goal is to find out why we want to research this?
- Will this idea get funding?
- Idea needs to be narrow and achievable
- Working title needs to be specific, yet still understandable
- Right now, the introduction section will be most highly weighted.
- Finish introduction with “therefore, this is our research question”
- Goals:
    - Short term goals(year end goals, must be achievable)
    - Long term goals(e.g. Curing cancer
- Variables if needed, not for innovation, studies
- Methodology
    - How/why does this technique help me answer my question

- Maybe around 3-5 pages?
- Background research should be extensive with a lot of literature sources
    - Know a lot about the topic and surrounding literature
- Significance
    - Why we need this project
    - Summary paragraph(s) at the end of the proposal
    - How does this contribute and is significant to the field of study

September goals:

- Find project topic
    - Narrow down scope of ideas

5 September 2025

5 September 2025

Daily notes:

Free block work(Right now work toward a research proposal):

- Logbook work, organization, planning
- Email mentor, meeting with mentor
- Read paper
- Outline research proposal
    - Research question
    - Intro
    - Methods
- Submit research proposal to mentor a week before October 16 due date for editing
- Create a mentor meeting plan to discuss:
    - Regular meetings(day 6 cycle, after school, weekly/monthly)
        - How are we planning this
    - Research proposal assignment
        - Topic
        - Research question
        - Goals
        - Outline intro
        - Background research
            - IEEE for engineering
            - Start with review papers
- Tuesday->setup PaperPile

9 September 2025

9 September 2025

Daily notes:

- Meeting with Dr. Garcia
- Set up PaperPile
- Dr. Sheikhi said in an email that they have decided on a project topic, **Wheat head segmentation**
- I have been assigned two papers regarding this topic, both of which Dr. Maleki has worked on. I am to read them, and following that, I will schedule a meeting with Dr. Sheikhi and Dr. Maleki to discuss the next project steps.
- Yesterday I received paper 1 but not paper 2 due to technical issues, and this morning I received the second paper
- Have a detailed plan for reading and annotating the paper
    - Specific details in the paper, sections
- Make sure to not fall behind when given leeway by professors
    - The deadline for reading and annotating both papers is the end of next class
        - I will first outline the key bullet notes in each section of the paper in my first read
        - Then, I will do a second read, adding necessary details and information to my notes
        - I aim to have this done by the beginning of class on Thursday, allowing me to show Dr. Garcia to take a look at my notes before I send an email to Dr. Maleki and Dr. Sheikhi and make sure they are adequate
- Follow-up email:
    - Provide summary of notes, move forward with finding a concrete plan
    - In meeting, push for narrowed down topic, research proposal, question, goals

11 September 2025

11 September 2025

- PaperPile download and tutorial
    - Folders for organization
    - IEEE for engineering
    - PubMed for medicine
    - Google Scholar, Archive
- PaperPile can export annotations on papers
- References **need** to be peer reviewed papers
    - Google Docs has a PaperPile tool to add citations
    - Once added, we need to decide what style of citations
        - APA is general
        - AMA is for biomedical
        - **IEEE for engineering**
        - Chicago for social sciences
    - PaperPile has citation style tool
        - **Check always contain DOIs and URLs**
        - Then click format citation
    - Every sentence with info from other info needs to be sourced
        - In-text citation/parenthetical
- Add in-text citations while writing, not after
- Multiple in-text citations for APA separated by semicolon
- uses square brackets. E.g. “As shown in [1], blah blah”
- Purdue owl summarizes rules on citation styles
- PaperPile only works on the first document tab in Google Documents
- Gen AI and basic sources from Googling are fine for learning initial information, but search for peer reviewed papers when citing for research proposal and other
- Sometimes PaperPile doesn’t find all the information about the source
    - I can edit the paper information in PaperPile
    - Change information in PaperPile, not manually in the Google Document

Previous class reflection:

- My date for finishing the papers was quite ambitious. They are denser than I had expected, specifically, I am slowed down since I am unfamiliar with a lot of the terminology, requiring additional research on specific topics.

15 September 2025

15 September 2025

Pubmed
Google Scholar
Research Gate
IEEE
ArXiv

Google
Terms
Basic info

free-pmc-> otherwise ask mentors

Outline RP introduction

- Broad topics
-
-
-
- Specific topics
-
- Leading to research question

Citation style:

- Across most papers IEEE is used.
- IEEE notes:
    -

Possible research directions:

- Wheat head growth stage categorization

Citation practice: ASP Citation Practice

17 September 2025

17 September 2025

Reflection:

- It took me a lot longer than expected to be able to first understand, then take notes on to further grasp what was going on in the papers.
    - I needed to complete more background research to have a baseline understanding of the methods and topics used.
- Both papers are deeply interconnected, and understanding the first paper took a lot longer than the second.
- This topic is already very niche—specifically the use of image segmentation->not just detection, as a tool for generating datasets and reducing the need for large amounts of manually annotated data.

19 September 2025

19 September 2025

Reflection:

- Today I had a meeting with Dr. Garcia
- Right now, I need to push for moving the project forward
- Make sure to keep the project moving and keep momentum
-

23 September 2025

23 September 2025

- Paperpile does not work with Google Docs tabs
- Do not need variables section

- [ ] Prepare for meeting
    - [ ] Reaching the deadline, if it is kept as it is or moved back, is the most pressing issue

25 September 2025

25 September 2025

Meeting agenda:

- [ ] Overview of project topic area
- [ ] Questions
    - [ ] What resources are available/necessary for this project?
        - [ ] GPU->how long is training time?
        - [ ] Storage
        - [ ] Existing codebase from the two papers
- [ ] Specific project topic ideas
    - [ ] Questions
    - [ ] Goals
    - [ ] Methods
- [ ] Research proposal
- [ ] Project timeframe, upcoming deadlines
    - [ ] Deadline to submit to mentors, then to Dr. Garcia
- [ ] What are my next steps?
- [ ] Future meetings
- [ ] Future communication

29 September 2025

29 September 2025

Finding papers:

- Search on IEEE, Google, using title or DOi

Log book grading:

- Structure, organization(/5)
- Content, details(/5)
- Calendar, current and next month(/5)

Following the research proposal submission, oral presentations:

- October 22, 24, 28
- Visual component
- Grading: 34/34, 10/10. Participation=attending and participating in 4 presentations

Methodology:

- Determine how specific methodology needs to be

Research proposal:

- Does not get remarked->grade stays throughout the year
- Methodology:
    - Should outline step by step what I need to do and when
    - Explain how results will be measured and quantified
        - E.g. MaP, Dice score
    - At the end, fit together the pieces to understand how it answers the questions and reaches the goals of the project

Mentor meeting transcript:
I don't know how you're doing good today, so I read both of them and I sent over the notes that I wrote about them. Oh, I sent them in the email. I sent it on September 18th.
Hello. Hello. Hello. How are you guys? Hi, thank you, menu. Great, great, thanks.
Okay, perfect. How was you, Ronald? How's it going? I'm going good.
So, how have you been able to go through the paper? Yeah. Yeah, it seems that he has found through the paper and already sent us some notes, highlights up the paper. So, how did you find them?
I found them pretty interesting, basically. They're both basically trying to solve, so they're basically trying to solve the same
problem, right? Basically, using other ways to create datasets without needing a large amount of

manually annotated data, and then how we do that is through in the first paper by using the copy-a-paste method and then using that one library to do the domain adaptation, and the second
paper using the generative adversarial network to do the domain adaptation, as well.
It took me a couple of days to really understand the first paper, but after I read the first one, the second one was a lot more easier to understand because it kind of built on a lot of those down the
first one. So, I think the next step would be for you to start building a model, just to double-check if you went through the AI research. Yeah. Unfortunately, I wasn't able to go because I had
a different program that I was a part of during that time. I think it's okay. Still, you can start building some models on this. I think perhaps that some of us will be able to help you with building parts once you do this to replicate some of those models. Okay.
And ideally, when is your other competition, by the way?
Like when do I have to finish this by? Yeah. So, it is in March, and yeah, so basically this month and
next month, I need to figure out a plan for what I want to do and specifically create a research proposal for the project idea. Yeah. So, I'm thinking of how to adjust the project so you can have a good showcase. I was looking at the end of the second paper. It talks a bit about potential, like further research topics, right? Like using two networks, one for the mask and one for the actual image, or looking at using heat maps for seeing how to change the layout of the
we heads in the cut and pacing. Something, would it be something like that?
So, that could be a project, but I am also thinking of a showcase, you know, that's the model building part. Yeah. How we can showcase your project to folks. Do you really focus on that? Yeah. First, I'm not sure if we can actually run even a lightweight model in a closed environment. There are some regulation preventing and safety of regulation. My guess is that's the case.
So, your competition will be March, yeah? Yeah. The final one. Yeah. And by that time, we don't have any
closed fields outside to can see them. And this means that we need to have a digital presentation
for you. Okay. And we also want to differentiate from some of those works. Yeah.
So, you can build on top of those works. First, you need to build a model and good enough model to capture ice with heads and also separate them. You know that models, they have just segmentation. They assign a same label to all of them. While we want to actually instill instant segmentation, means that down the road, we can actually count how many we'd head we have per,
let's say, per square meter or something like that or per acre or something like that.
And have a hit map of that field. I think that would be a good thing to have focus on counting as well. That's counting is how many we'd head. Often they have a average, a reliable average of speed per we'd head. So, really, if you have from each form, you have a few four, five of those we'd head. So, I don't really select it. You get a good average of we'd head capacity and how much
speed you get from each we'd head or a spike. So, if you are able to count the number of

spikes, you somehow can calculate how much harvest you will get down the road. Yeah.
Yeah. So, I think that might be a good thing to go with. So, those are
semantics segmentation. We go further and make an instant segmentation where we actually differentiate between we'd head. And that give us ability to count and ability to count give us ability to monitor how much harvest you will get from each specific part or each specific area of the land. So, what's the use case of that? If we can have a estimate of harvest,
this give us and the question is that's why that's happening. Yeah. That give us some good information about, yeah, some parts of the land might need more fertilization, some might need less. Yeah. So, it give us opportunity for plan down the road.
And that's about the more things to know. Yeah. It's more characterized one. Okay.
So, this is more looking at like the practical application of the segmentation model rather than like,
so it's like moving away from the other papers that we've done. Yeah.
Yeah. It's a lot sort of different from segmentation. Yeah. Instance segmentation. It's like segmentation about you need to differentiate between instances as well. It's slightly more challenging.
Yeah. And the other things that we can, I'm thinking of cool projects, I want, I want winners solutions. Yeah. If this was summer, we were able to actually have a drown on the field and connect
that drown to your system and analyze things for the competition. But it is not.
If I was I'm thinking of honey bees. If we can do an instant segmentation on honey bees, we can actually count honey's
honey bees. That population is interesting to incorporate. The other things that I'm looking at this is that can we actually connect this to connect the model to a smart glass. So, we run the smart
model, the ultimate of the, you know, the sunset from a smart glass can see it. We send it to model
the character. I said, send it back. So, that's going to be a cool thing to do. Yeah.
But any chance do you have access to someone who, you know, has access to some heights?
I'm not sure. I think my dad might know some farmers. I'm not sure if they have, like,
beehives. So, I can look into that. So, one thing, I mean, the technicality, if we apply
for wheat heads, or honey bees are comparable. But for the beehives, I think we can have another
layer on it. For beehives, we run the same model, counting all of those information is relevant, even finding the queen is relevant, and also connecting it to that smart glass is very relevant.
So, it's make it very cool, and the nice thing is that you can, we can have, you can test this on a high, and you can wear your glasses, and we can actually show you as a, in the competition on
a screen, that, you know, you record that show it, and also you show what we show what we see.
You see from the, the smart glass, so it gives it a couple of layers of interesting things.
I think that would be very interesting. Also, it has fallback options. Fallback option is that the smart glasses doesn't work, or we don't have time, we still can do it with a cell phone, yeah? Because it's almost the same thing. And the nice thing is that hives are more available.

It's not, we need to wait for the end of summer. Yeah. So, January, February, it's March, yeah. We might be able to get that information, and plus, even if we,
yeah. So, that's more approachable, I would say. Okay, yeah. Take a look at that.
As one option. So, we have the, the, for wheat as, as one option is a backup option. We have data for it,
that's nice. We have data, we have labor. Honeybee is another interesting, very interesting options,
that you can actually do some models that do this incense segmentation and counting for honeybees.
The AI behind it is very similar. Yeah. I would say it's very similar.
So, however, the, the presentation of one might be way more interesting. Yeah.
So, I leave that to you. Unfortunately, I don't have access to any beekeeper here.
So, I leave it to you. Do some research. You might, you know, you might know some people. Yeah. And if that's the case, it's, if that's the case, try to collect some information.
Some, you know, ask them these beekeepers to collect some video of them, you know, bees and stuff
like that, where you, you can perhaps find a lot of them, but which are a reasonable camera and from a closed distance, because we need some data to train the model.
Yeah. For the wheat, we have the data for the honeybee we don't have.
So, there's no, like, comparable, like, data set, like, the, for honeybees.
They might be public data set as well, and we can perhaps look into that, but it's good to have the data that you go and collect, yeah. Yeah. Ideally, you go and collect, but make sure that you are not, you know, you are not, you don't have any allergy to honeybees,
stuff, and if you go, go with someone, never go yourself alone.
Your, your parents or someone should be there, and we're every safety things.
Some people might have allergy, you know.
Caring it depends not that, yeah, in general. I haven't seen anyone with significant allergy, but things happen, yeah. You need to be prepared, and you never, never, never go alone.
Oh, there is a honeybee behind this honeybee hive behind this wall. Let's go collect some data. Don't do that, yeah? Yeah. Make sure that the expert is there. Make sure that you wear enough, if you, even if you want to, you know, capture that. And when you capture that, make sure that someone captures some, some video of you, collecting those information. We might need that for the day that you present. Okay. That's, think, build those things right now.
Yeah. We'll start at the end in line. Yeah. Yeah.
So, I think the most, the next thing that I would need to do is figure out, like, specifically, what, like, questions are, like, goals. I want to, like, try to answer with the project.
Is that something we could figure out an hour? Would I need some more research to do?
Absolutely. So, for that, for the OVTA, somehow, we, we, we guess, yeah, what you can get from the
instance segmentation after we can actually have a prediction or estimate our progress.
Yeah. And that could be actually, you know, easy to explain. For honeybee, also, we want to identify the, you know, things that comes easy and early is that I, you know, identify the

colony house by seeing how many bee saved left there. Identify, we can go even further, identify the quim, that sometimes quen is the most valuable, you know, piece on your, in your hive. So, when you, when you open it and you move things and move things around, identifying your quim,
and make sure that the quim does not, you don't have the quim is the first thing you do. Sometimes it's not easy to see them. Yeah, it's a, it's not a straightforward process. A lot of bees around and quen is a little bit bigger. Yeah. And sometimes there's the dots, you know, we market with the dots, which these dots are color coded, each color represents a specific year of bears for the bee, the, the quim. So that color helps us worldwide, identify, but not always easy to see it, because there are other quim has a lot of people's supports, supported, yeah. And a
lot of these often cover the quim sometimes, and it's not, it's making it difficult to see that color. So that could be one in general, the health of it. And beside that, we can actually
identify other things on the bees. Farmers is doing a project, identifying a specific might on bees. So those might is like, it's like tick for a human. Yeah, but also, let's beg,
imagine tick for a human is like a plate. It's exactly like that, yeah. And it's attached
the bees. It's a plate size, if you scale them, it's like a plate for bees. And after a while, make them weak and also pass some wires to them. And they expand very quick. So it's not just one
to one. One comes to the hive and they become a thousand. And at the end, they kill the hive. So identifying them early is important. So in that case, you have a good
story to tell. You are saving the bees. So that's also, it's a winner solution.
So consider these two. These two could be good ideas to run with.
And on the side, based on, you know, do a lot of without research, see what we have around. If you have access to something, if not, we can go always with the
uh, uh, the weeds one and harvest addiction. Yeah. Okay. Any other question you might have? So, um, so I put it in a previous email that, um, so I have, uh, I've disamined a research proposal by October 16th. Okay. And that's feasible given the current time frame or, um, or should I?
Absolutely. We will help with your that one. Find out the answer for these and this one give us the
direction, which way to go. One way is more toward the drone imaging and the other one is more
toward a small, a small glasses. Okay. Yeah. And we'll decide. Okay. Yeah. Um, and then, if you can collect the data, that's also another valuable part of your project. Yeah. Yeah. The collection is also very. That's true. Make sure that's for the bee honey bee. We need to, we create a collect data and it is time to collect the data. Winter time is not easy. Yeah. You cannot actually disturb the hives during the winter. So if that's the case, you need to manage that. And again, this is something that's you need to, uh, you don't need to personally create the collected data. But if you, if you, uh, we have some gears and go collect those data and someone captured that, uh, you know, that the, uh, the movie that you are doing it, it's make a good impression on the presentation that you are actually doing the job and you are involved in the aspect of it. Uh, be aware of potential allergies. If you haven't been exposed to any, uh, any bee sting, if, if that's the case, uh, have a, uh, cup,

heavy pen with your saw. Okay. It's one in a, I don't know, half thousands.
And even in those cases, it is not very serious, but always be prepared. Make sure that you wear, uh, enough gears and make sure that you have perhaps a parent or someone with you as well as
a beekeeper. So you don't, uh, uh, make the hive angry. Usually that they are honeybees are very
patient. They have one life. And if they hurt you, they die. It's not like wasps. Yeah. So they are not taking that risk. They are not aggressive usually unless they feel, uh, you are hurting them or some, hurting the hive. In that case, their job is to defend the hive. Okay. Yeah.
Okay. Can't, well, you know, can't provide you enough warning. Yeah. Yeah. Of course. Um, hey, and then I think next, uh, would be could we schedule a next meeting?
Sure. Let's do that. What's the best time for you? Um, this time is hopefully in this week. Yeah. This time might be a lot of the difficult for me. Okay. If I need to pick up my daughter sometimes from daycare, it might not be, and might not be in a place that I can talk, I might be in. Okay. And them all doing the day. Uh, so during the day, uh, our school doesn't work on like a, uh, a weekday schedule. Um, and it's different, like week by week. So I could schedule that, um, but it would be a bit more like it would be, it would be, it wouldn't be like a regular. It's a biologically or it's just random. It, it's pretty much random. Yeah. Because it's a six day cycle. Cool. Yeah. That's cool. Yeah.
And that's my good with my life as that. Okay. Uh, is there any, any specific things that you know,
all of those random things they might actually, uh, a seal and keep that day to some extent? Um, I could do during a lunch time. So my lunch time is doing 1132, 1210.
Okay. And perhaps it could be a very short meeting. Now, and if we need more, uh,
a longer meeting with the schedule schedule for something separate. So this will be a, you know, a quick chat. Yeah. Uh,
uh, that's weekly updates. See what you have done. It could be really 10, 15 minutes. So you still have your lunch time. Okay. Yeah. Okay. If that's the case, uh, what Monday?
Yeah. That works. Let's go with Monday at noon. Yeah. Uh, we keep it short. Feel free to, you know, join while having your lunch. It's completely okay. So don't, don't feel
as a formal meeting. Uh, it's very casual. Have your lunch and we,
it quickly, you know, you might have four, five minutes updates. Um, you might need some information we shared with you or might say, I'm working on it. I'm collecting the data. I'm running
this and everything is good. So if they get enjoyed a lot. Yeah. Yeah. Perfect. So five minutes will send that invitation to you for a calendar invite. Yeah. Yeah. I'm just sending the calendar invitation and we're going to have it weekly. Yes. Anything weekly is good. I don't, but if we need more, if we need more, if you need more support, we'll, well,
schedule a different meeting. Yeah. Or I'll say, uh, you know, at that week, you say,
oh, this, this Thursday morning 10 to 11, I am available. So let's have a backup meeting for, to cover that specific topic. Yeah. Okay. Um, I don't have any more questions.
Perfect. So, um, let us know when you have that answer without answer reports and we'll proceed from there. Okay. Yeah. The invitation is sent now. So hopefully by the end of the week,

I can provide answers to those questions. So on Monday, we can cross it with the provost of.
Yes.
Great. Thank you very much. It was nice seeing you. Yeah. Thank you all.
Have a good rest of the day. Bye, everyone. You too. Thanks. Bye.
Okay. Let's

October Notes

1 October 2025

1 October 2025

Logbook improvements:

- More specificity needed in notes, especially in the calendar section
- Track dates for work done
- Update to do lists to demonstrate completion

RP:

- can cite multiple sources for one claim
- AI can be used to find targeted peer-reviewed sources for backing up claims(check each source)

Mentor meeting notes:

- Established regular weekly meeting schedule of Monday, at noon(extra meetings can be scheduled if more time is needed)

Project overview:

- Working on an instance segmentation model to present
- Two possible project directions are:
    - Wheat head detection(backup)
        - Count wheat heads to estimate yields to compare different areas of growth and optimize fertilization
        - Wheat out of season until the summer
    - Honeybee detection(preferred)
        - Count honey bees and detect queen bees
        - Could connect to smart glasses for real-time analysis
        - Personally can do collection of data
- Functionally, wheat head detection and honeybee detection will use similar methods
- Next steps: get a functional segmentation model working, and decide on a final project direction

3 October 2025

3 October 2025

Log book:

- Date notes
- Reference sections to link to work at a specific date
- Calendar includes current and next month specific tasks

RP:

- Intro
    - Start with statistics->what is the problem?
    - Each sentence backed up by a different source
    - If one paragraph uses one source, add the citation add the end of the first sentence, and at the end of the last sentence
    - Primary source->firsthand information
        - If citing non-primary source, must cite both primary and secondary source

7 October 2025

7 October 2025

Meeting notes:

I have gotten in contact with a beekeeper who is willing to let me do data collection of the honeybees, so I am able to go ahead with it once I start the project.

Today I shared with Dr. Maleki the previous potential project research question focused on queen bees, and discussed Varroa mites as a primary project direction that focuses on mite counting for determining intervention thresholds rather than queen bee detection. Which project direction I pursue will depend on whether or not I can collect mite data from the beekeeper, so I will get back to both of you soon on a final project topic, following which I will start writing my research proposal.

Next steps:

- Outline proposal
- Start with intro(methodology for next week)
- Start by talking about theoretical topics(computer science, machine learning, data)
- Move onto practical applications, etc.->direct toward reasons for this project
- Ask about computer power needed, uofc resources, home computer.
- How long will the project take?
- By next meeting, have outline and introduction outline, to show mentors and ask for papers I need access to
- Ethics with bees?

New research proposal deadline: October 27
Oral presentation: Early november

9 October 2025

9 October 2025

Mentor meeting follow up email includes:

- Bullet points of what happened in the meeting, next steps, and deadlines
- Ask for missed things, confirm next meeting

RP:

- Don't do abstract unless the whole project can be summarized in like 200 words
- Introduction is the biggest section->needs to be in depth for the aspects important to be understood for the project
- General data for the problem(could be a couple paragraphs) to explain what is ging on
- Then nitty gritty of the field->what specifically is going on, more narrow and more into depth. Land upon how this explains the problem that we are trying to solve, and then what we are doing.
- Research question goes at the end of introduction, to explain what the project is about. “Therefore, this study aims at answering the following question:”
- Research questions are not goals. goals/objectives are like specific steps/tasks->e.g. long term contributes to field of study
- Goals(we need short term and long term goals)->succinct, exactly what I am doing(without reading intro, should be able to understand what the project is about from the goals and question)
- Innovation and study does not have hypothesis, variables
- Innovation variables could exist, like things we are changing
- Innovation and study research proposal evaluated out of 80
- Questions and hypothesis after intro
- Goals after questions
- Variables can go within methodology

14 October 2025

14 October 2025

Biorender

Citing other people's figures:
- "Courtesy of …"
- "Taken from…"
- "Kindly provided by…"

Oral presentation November 3

RP:
- Subtitles are optional->main categories outlined by the guideline document
- Introduction needs to be very in depth into the topic

Meeting questions:
- I have reached out to more beekeepers, waiting->backup?
    - Initial beekeeper said they didn't want to disturb bees or whatever as it gets colder
    - albertabeekeepers.ca ->more commercial keepers
    - Should still be able to use sticky boards for mites as it gets colder, however depends on beekeepers willing to do the stuff and open hives
    - How many sticky board images do we need
- Ethics?
- General project timeline?
- Training resources(u of c computers, mine at home, how long?)
- Overview of research proposal outline
- What specific dataset generation techniques?
- List of papers they could access and send to me?

Mentor meeting notes:
- Moving forward with contacting beekeepers, I can mention involvement with the UofC for credibility. I can also highlight that all data will be anonymous, bees will experience minimal disruption from sticky boards. Given that they say no to data collection, I can offer testing of the mite detection once my project is completed in the warmer months of the year.
- We discussed the research question and goals, narrowing down goals to four main goals:
    - Generate large-scale simulated dataset from limited real sticky board images
    - Develop semantic segmentation model based on simulated data
    - Enhance model performance on real data through domain adaptation techniques
    - Assess model accuracy using real sticky board images
- In addition, long term goals focus on bee health management with mite monitoring

- For the methodology, we will use semantic segmentation for the small mites that rarely overlap. Dataset simulation will focus on cut-and-paste augmentation, and will include negative examples of debris in training data as well as mites
- I will have access to resources at the UofC for training the model
- Ethics are minimal concern since sticky board is non-invasive, but look into ethics forms for science fair

Next:

- Send email with summary, next steps(writing proposal)
- Also ask about giving me some papers

16 October 2025

16 October 2025

Citations include doi, url

All proposals are out of 100

20 October 2025

20 October 2025

Research proposal formatting:
Title page/top of first page
Name
Mentor name, dept
Research proposal
Applied science project
Date

Submit to turnitin, google classroom

Vocabulary

22 October 2025

22 October 2025

ASP papers:

- RP
- Intro(Nov - Jan)
- Methods(Feb)
- Results(Mar)
- Analysis(Apr)

Oral presentation:

- Create an outline(how many slides for each section, total 8-10 slides)
- How many slides per class or per day
- Practice
    - Make sure people understand the presentation
    - Make sure the timing fits well enough
- Note: 10 minutes is a hard cutoff
- Start the presentation with a punchline, or a brief overview of the project. What is the project about?
- Title page is a 30 second elevator pitch on what the project is about
    - Don’t give the background on the topic to start
    - Summarize the essence of the project in a few seconds

24 October 2025

24 October 2025

Little text and big fonts

More graphics(images, graphs, tables, flow diagrams)

Slides should use maximum space for bigger fonts and images

Images should be good quality

Citing images: don't use citations in oral presentations. Credit the source informally, mention where it is taken from.

- "Taken from …"
- "Courtesy of …"

Methodology is good with flow diagrams for an overview of the whole study, gives audience overview of what to do

- Then decide which components in the flow diagram needs an extra slide

Key note: start presentation with an outline:

- Title(presenting w/ punchline), 1 slides
- Background research, 3 slides
- Objectives, 1-2 slides
- Methodology, 3-4 slides
- Significance, 1 slide

10 minute hard cutoff talking

- Small grace period

5 minutes Q&A

Dr. Garcia Meeting:

- Deadlines rescheduled to November 7 for the oral presentation and the research proposal

28 October 2025

28 October 2025

Oral presentations

- Richard, Jessica

Notes: Presentation Notes

Mentor meeting:

- Methodology section needs to explain the general idea of how I will do the project

30 October 2025

30 October 2025

November goals:

- Enter info for cysf
- Submit cysf ethics
    - Especially animals
    - Enter code of mentors ethics if they have it
- Declare mentor, name of mentor and email, phone number
- Data collection
- Computer programming
- Keep up with background research

31 October 2025

31 October 2025

Dr. Garcia meeting

- Flow of the research proposal is good
- Main areas of focus for improvement
    - Make sure citations are consistent
    - Improve phrasing for clarity in certain areas
    - Add methodology section
    - Make working title more succinct
- Start on oral presentation
- Due date: November 7

November Notes

5 November 2025

5 November 2025

Oral presentation:

- Complete oral presentation
- Meeting tomorrow at 10:30 for practice

Weekly emails

- Progress report
- Summary notes of weekly meeting
- Plan for next week

Key notes:

- Review logbook, reflection on future tasks
    - 15 minutes logbook everyday
- Logbook should be able to show what I spent my time on->linking to stuff, copying and pasting stuff

Data collection:

- It is likely that mite monitoring has already ended in Alberta and most of Canada. Most bees finished by October and so there is limited opportunity for sticky board monitoring right now

7 November 2025

7 November 2025

Regular meetings and emails needed
Calendar needs to be filled with tasks

The month of November is the most important for getting the most of the project done

- Data collection
- Coding
- Analysis

December and January has limited work time
February has limited time for the project
Discuss long term view of the project
Plan for long view of what needs to get done when
Research project organizes the view of the project

Dr. Sheikhi gave feedback on the research proposal

- Focus of the project should be the mites and not necessarily the synthetic data generation
- Make sure terminology properly conveys the ultimate goals of the project

13 November 2025

13 November 2025

Marks will close next Friday
Mentor evaluations

- At some point I may need to ask mentors to provide evaluations
- Mentor evaluations will change in their value at some point
- Weekly summaries to mentors via email is important

CYSF:

Judging sheet used at school science fair and CYSF:

- Innovation project
- 3 important parts:
    - Clearly stated problem and hypothesis
        - Existing background research implemented, providing direction for the project
    - Methodology clearly mentioned in oral presentation (e.g. chi square test)
        - Mentioned verbally and not just on poster/trifold
        - # of samples
        - Log book is important->refer to logbook in presentation
            - Put link on poster/trifold
    - Analysis and conclusions
        - Proper methods used for analysis
        - Possible improvements
        - Conclusion related to project, data, and problem
- Trifold/poster needs all 3 parts
- Sources need citations (multiple independent sources)
- Problem was approached with creativity and originality
- What was the motivation for the project
- How will I take this project in the future
- What is the significance of this project
- Don't say, "my mentor said…"
    - Present the project as your thoughts
- Project needs to be clear and easy to understand

Presentation:

- Should start with a summary of the project->what the project is trying to do, its goals
    - "What my project is trying to do is…"
    - Short summary

Calendar:

- Make a long range plan with mentor from November to February. List out what are the goals of each week
    - Separate out main tasks and deadlines
        - Data collection
        - Model training on regular data
        - Data generation
        - Model training on synthetic data
        - Data analysis
        - Conclusion
    - Understand whether or not im on time
    - Given inability to finish something, have an exit plan
        - Narrow down goals
- Last 2 weeks of february should be for making poster and presentation

- Mid november to mid december is focused on methodology

17 November 2025

17 November 2025

Currently I am waiting on responses from:

- Alberta
- BC
- California
- Texas

My email communications are written here

Dr. Garcia meeting:

- Concerns about how much of the project I need to do
- Bring up in mentor meeting:
    - Tentative long term plan for project
    - Cutoff dates for data collection
    - How much data is needed
    - What can I do now before I get data?

Tentative schedule:

November	December	January	February
Best case: collect all data Worst case: continue collecting data	Test out neural network detection Stop data collection, use datasets	Synthetic data generation Neural network w/ synthetic data	Data analysis Data analysis

Mentor meeting:

- Discussed long term project plan (finish data collection by the end of November)
- Work on project over december and winter break
- Have a minimum viable product by January
- Begin learning YOLO architecture using other datasets while waiting for images

19 November 2025

19 November 2025

Continue momentum of project

Update:

- Following contact with a beekeeper, I have a beekeeper who agreed to help me: Response 2
- I received 18 images from Rushing River Apiaries. After forwarding them to Dr. Sheikhi and Dr. Maleki, I have received feedback from both:
    - Dr. Maleki says the images are too noisy
    - Dr. Sheikhi says the images I received are not sticky boards
- Dr. Maleki forwarded this video about preparing sticky boards: https://www.youtube.com/watch?v=90qgGnHLigw
- Dr. Sheikhi forwarded me images from a different research paper as reference. I sent the same images from the same paper to the beekeeper previously, so I did not forward Dr. Sheikhi and Dr. Maleki's emails to the beekeeper at Rushing River Apiaries.

Sample 2 images I received:

I found this map of honey production by state to see what places would be best for me to contact. I will look for states where the weather is warm and honey production is high to contact more beekeepers.

21 November 2025

21 November 2025

Daily notes:

- Logbook is not an agenda, it should not summarize what I've done
- Specific information and links to what work was done
- Calendars should be filled with specifics until winter break
- Implement research proposal feedback
    - Introduction section = december paper
        - Needs to already be in depth and mostly done
    - Methodology section = february/march paper
        - More in-depth later
    - Feedback is to be applied for future paper assignments
- Use empty spaces in the calendar to do miscellaneous tasks:
    - Edit research proposal
    - Work on log book
- Continue project with a long term vision of the project
- CYSF platform
    - Provide video presentation
        - Talk about difficulties throughout the project
        - Results
        - Judges preview the presentation video before the day
        - Pictures of labs, previous attempts
        - Endorsement from mentors
        - How to improve the project in the future
        - To what extent was it successful?

Testing first collected data:

- Use my images in this existing varroa mite detector to see if my images are viable. This detector is the research paper here
- Results here
    - Based on the results, I need to ask the beekeeper for some confirmation of what is a mite and what is debris->some are hard to tell by naked eye zooming in on the images
    - I believe that some of the images I received will work for the project. The background noise seems not to affect the detection too much, and the main effect is large amounts of wax debris that makes it hard for the model to pick out wax versus mites.

Goals:

- Complete data collection by the start of December
- Have a working MVP by January. I can continue testing and improving the model as I prepare for the science fair and presentation of the project.

Update:

- I have sent a follow-up email to Rushing River Apiaries asking if they could send less noisy images by the request of my mentors
- I have sent emails to beekeepers in Florida, Pennsylvania, and Philadelphia
- CYSF Ethics have been approved and can be found here

25 November 2025

25 November 2025

No class, president’s breakfast

27 November 2025

27 November 2025

- Send follow-up email to Rushing River Apiaries about the possibility of getting more sticky board images
- Make edits to logbook and submit
- Prepare for tomorrow's mentor meeting to discuss:
    - The images I received in further detail
    - Progress on data collection
    - Immediate next steps

December Notes

1 December 2025

Shortened class(assembly)

During the winter break & midterms:

- Keep in contact with mentors
- Ideally discuss project plans in person
- Make a plan for what happens during winter break and midterms
- Last 3 weeks before winter break are vital to the project
    - Identify problems and project details soon

Ethics forms:

- Adult supervisor: Dr. Garcia & Dr. Kale
- Scientific supervisor: Dr. Sheikhi & Dr. Maleki
- Phone number & email: connected to mentors

Mentor meeting:

- Getting data from the university

9 December 2025

- Plan for December to January must be clear
- Update mentors through email during the winter break and January(weekly)

11 December 2025

Notable change for logbook:

- Google Docs has a maximum of 100 tabs, so I will begin

15 December 2025

17 December 2025

19 December 2025

Absent

January Notes

20 January 2025

Absent for midterms

22 January 2026

Worked on Progress Report for meeting with mentor: Progress Report

- Gathered all previous training attempts, along with training and testing metrics, to put in report
- Added descriptions to each

26 January 2025

Methodology Paper

- Read other papers' methodology section to ensure it is good
- Content:
    - enough detail so anybody else can follow the work and recreate the work
    - Present easy to follow steps without including standard laboratory procedures
    - Used to either recreate experiments or compare methodologies
    - Specific materials and equipment should be described along with manufacturers
        - Parentheses (model, company, place), or (manufacturer, place)
    - Display deep understanding of technique(description needs to be correct, accurate)
- Style:
    - Paragraph form
    - Write in past tense: because the final paper will be completed and everything will be done
    - Materials mentioned and described in narrative text, not a list
    - For figures and tables, ask if this helps make the methodology more clear. Must be relevant. IEEE has a specific format for figures and tables. Figure accompanied by figure number and figure legend(title and description of figure). The figure must be described in narrative text before being shown. Referred to in the narrative by figure number.
- References and citation
    - Materials referenced in parenthetical along with their source
    - Describing methods or information should be cited in citations
    - Citations linked correctly to reference list
- Methodology comes after intro, after goals section
- Needs intro sentence before going into technical stuff

Even if data was not self-collected: experiment/data collection needs to be explained

- Start explaining the study and the data collected

Statistical analysis: typically near the end of the methodology section. Parentheses need to have a package version and year of statistical software used for analysis.

"This study followed the procedure described in sources [1] and [2]"

Avoid using "then". Use technical and dry language, not wordiness.

Create various subtitles->IEEE has a specific way to format subtitles and sections

Equations are well known or should be cited. Describe what variables are and do.

Animal Care and Ethics: From where, how many. Animals handled by ethical guidelinese of animal care commitee and approved under protocol number ac22-0137

Steps:

- Outline->what sections needed
    - Quick intro
    - Data collection
    - Describe general process
    - Specific details of methodology
- Fill in outline->figure out what I don't know
- Improve formatting, style, vocabulary, etc.

Methodology

- Not the longest section
- Heavy technical details

Page break before references page

Give due dates for certain sections to write

February 19 due date

28 January 2025

Meeting with Dr. Garcia

- Biggest concern right now is the balance between spending more time on the project itself and presentation of the results(methodology paper, oral presentation, poster)
- Discuss with mentors how to best present results
- Remade progress report to better standards: Progress Report 1

30 January 2025

Absent from class

February Notes

3 February 2026

Trifolds:

- (80cmx150cm)x3 panels

Posters:

- 4ftx3ft

Using Albumentations with YOLO

- https://docs.ultralytics.com/integrations/albumentations/#flip-left-right-fliplr
- Much more specific control than default Ultralytics augmentations
    - Control of each specific augmentation probability and magnitude

New YOLO model was released 2 weeks ago:
https://docs.ultralytics.com/models/yolo26/#overview

- Considering improved performance of YOLOv26 compared to YOLOv11 could help improve the project

Outlined Methodology Paper

5 February 2026

SF Judging:
Gold-100%
Silver-95%
Bronze-90%
HM-85%

Final oral presentation on last day before exam break(June), open to parents, mentors

Segmentation mask tools(manual and automated):

- https://thelinuxcode.com/how-to-use-the-segment-anything-model-sam-to-create-masks/
- https://www.mathworks.com/matlabcentral/fileexchange/38484-segmenttool-an-interactive-gui-for-segmenting-images
- https://github.com/branislavhesko/Manual-Segmentation-Tool
- https://github.com/Araf01/Simple-Segmentation-Toolkit
- https://www.brainvoyager.com/bv/doc/UsersGuide/Segmentation/ManualSegmentationTools.html

Meeting with Dr. Garcia:

- Methodology paper: don't include synthetic data
    - Focus on current work and results
- Key: poster and oral presentation
    - Practice is important for the presentation
- Methodology paper is second->but is already necessary for other results
- Continue working on project beyond school science fair->can talk about/present results in city science fair

Project:

- YOLO26 takes approximately twice as long to train as YOLO11, with no improvements on performance of my model

9 February 2026

CYSF Portal:

- Preview of project to judges
- Finish all data and analysis->understand project deeply
- Portal is lower priority than data analysis, poster, and oral presentation

Title has figure legend in paper(fig number, title, description only needed for paper)
For presentations: titles can go on top of the diagram(this vs. this is not a title)

- What is the graph actually doing

Oral presentation doesn't need formal formatting

- Base presentation on November oral presentation
- Shorten introduction of problem, focus on added methodology, results, and discussion

11 February 2026

Printing: symbols may not work, so check in with printing company

Methodology paper outline and finalizing of details

Logbook reordered to have space for enough tabs

Albumentations:

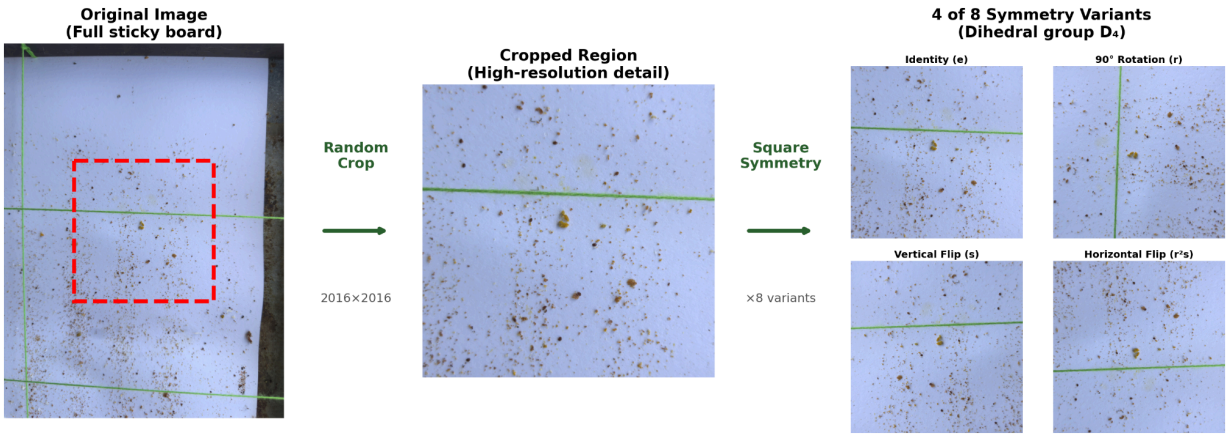
- Creating a pipeline has a 7-step approach
    - Size normalization
    - Geometric invariances
    - Dropout
    - Colour
    - Affine
    - Domain specific transforms
    - Normalization
- Some steps may not improve performance

17 February 2026

Final results of model: Progress Report 2

Results:

- Overall, the model already had a relatively high performance without augmentation despite limited data
- Two augmentations helped the most, Random Crop and Square Symmetry
    - Random Crop: zooming in on image to keep image resolution
    - Square Symmetry: flipping and rotating the images based on the axes of symmetry of the square crop
- Other augmentations worsened model performance



19 February 2026

Methodology paper will have an extended deadline

After science fair:

- 4 papers(30% of overall mark)
- Judging of science fair will be a mark
- Final oral presentation

Judging with the science fair presentation:

- Vigilant about AI usage(usage must be exactly disclosed)
- Must combat bias against mentored projects(How to present project in a light that shows personal connection and involvement in the project) ->not just mentor's project(mentor-led vs student-driven)
- Use rubric and words in the rubric during presentation
    - E.g. directly explicitly explain my originality in the presentation
- Innovation: solving a problem(how did I think about the solution and go about it)
- Avoid over-formality in oral presentation
- Stay on topic with the presentation
- Eye contact
- No cue cards

Oral presentation improvements following practice presentation with Dr. Garcia:

- Slow down when explaining mean Average Precision
- Introduce sticky boards and Varroa mites before the research question for clarity
- Present the problem of why we need to automate this process
    - Emphasize size of mites and debris
- Explain the various data augmentation transforms
- Explain dataset splits, training epochs, and other training information
- Use more visually appealing figures for results
- Emphasize next steps of the project with collected images and synthetic data

23 February 2026

Mentor meeting:

- Reviewed final oral presentation and slideshow

Oral presentation

Took notes on Andi's oral presentation: Presentation Notes

25 February 2026

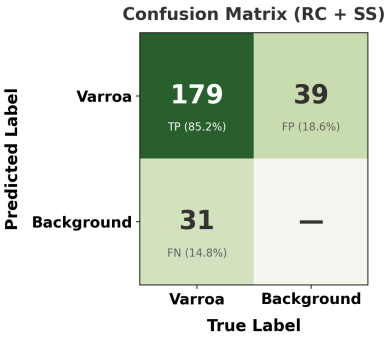
Notes on Shaayaan, Anki, and Samir's presentations: Presentation Notes

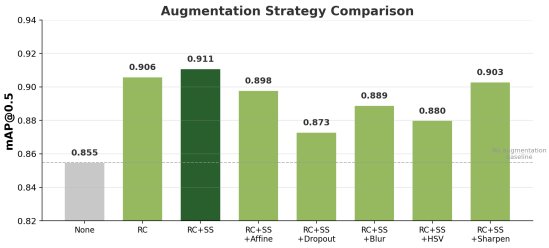
27 February 2026

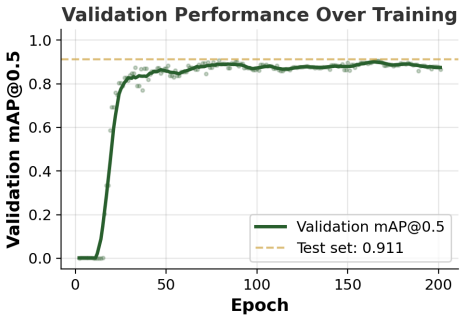
Notes on Shicheng's presentation: [Presentation Notes]

Figures generated with matplotlib library for poster (same data but with larger text and colour)

Finalize poster design to send to Staples tomorrow







Automated Varroa Mite Detection Using Deep Learning and Data Augmentation

Ronald Li[1]; Farnaz Sheikhi, PhD[2]; Farhad Maleki, PhD[2], [1]Webber Academy, [2]University of Calgary

Background

- In the 2024 season, a third of Canada's honeybee colonies died, with the largest contributor being the Varroa mite [1].
- **Varroa mites are parasites that feed on bees [2]:**
  - Decreases bee immune capacity and lifespan
  - Spreads viruses including the Deformed Wing Virus
  - Without treatment, colonies may die within a few years
- Monitoring is vital. Sticky boards track infestation severity, but manual counting is expensive due to debris and mite size [3]. We use a CNN to automate this process.

Research Question

Can we develop accurate Varroa mite detection models with limited training data for automated counting?

Objectives

- Maximize detection accuracy using only 64 labeled images
- Explore data augmentation and synthetic data strategies
- Develop a practical solution deployable in the field

Methodology

1. **Data Collection**
   64 sticky board publicly available images with annotations
2. **Data Augmentation**
   Random Crop + Square Symmetry
3. **Model Training**
   YOLOv11 object detection, 70/10/20 split, 200 epochs
4. **Evaluation**
   mAP@0.5, F1 Score, Accuracy

Methodology - Data Augmentation

Our best augmentation pipeline combines Random Crop with Square Symmetry, preserving mite detail at high resolution and training our model on all possible orientations of sticky board images.

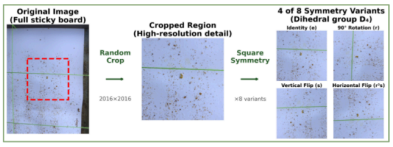


Figure 1. Augmentation pipeline: Random Crop → Square Symmetry.

Results

Augmentation	F1	Accuracy	mAP@0.5
None	0.786	0.645	0.855
Random Crop (RC)	0.852	0.701	0.906
RC + Square Symmetry	**0.836**	**0.719**	**0.911**
RC + SS + Affine	0.829	0.669	0.898
RC + SS + Dropout	0.811	0.681	0.873
RC + SS + Gaussian Blur	0.815	0.670	0.889
RC + SS + HSV	0.804	0.663	0.880
RC + SS + Sharpen	0.827	0.709	0.903

Table 1. Performance comparison across augmentation strategies.

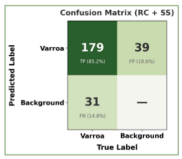


Figure 3. Confusion matrix (RC+SS).

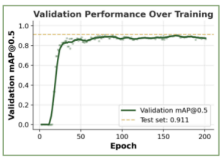


Figure 4. Training convergence (RC+SS).

Discussion

- Domain-specific augmentations (RandomCrop + SquareSymmetry) achieved 91.1% accuracy, improving by 5.6% over no augmentation.
- Generic robustness augmentations (blur, sharpen, color) hurt performance by creating a train/test mismatch.
- Small object detection benefits from high-resolution crops (2016×2016) that preserve detail without increasing training computational demand.
- The number of false positives and false negatives are similar and will balance out when calculating total mite counts.

Conclusions

- This project contributes to tools protecting honeybees
- Diversifying existing data is vital to real world applications
- Targeted augmentations improves machine learning models with scarce training data, resulting in high performance neural networks
- My application uses my trained model for automated Varroa mite counting in the field

Future Directions

- Increased data collection in Canada
  - Varied sticky board types and conditions
- Different model architectures
- Synthetic data generation

Acknowledgements

Thank you to Dr. Farnaz Sheikhi and Dr. Farhad Maleki from the University of Calgary Department of Computer Science, and Dr. Beatriz Garcia-Diaz, for the continued guidance and support throughout this project.

References

[1] G. Claing et al., "Statement on Honey Bee Wintering Losses in Canada," CAPA, 2024.

[2] K. S. Traynor et al., "Varroa destructor: A complex parasite, crippling honey bees worldwide," Trends Parasitol., vol. 36, no. 7, 2020.

[3] V. Dietemann et al., "Standard methods for varroa research," J. Apic. Res., vol. 52, no. 1, 2013.

March Notes

1 March 2026

Accidently ordered a banner instead of a poster, so then I had to buy another poster. The poster uses the Canva editor, however this messed up the formatting of my poster. So then I had to buy another poster. A few hours later, the first banner ended up being the same as the poster.

Daily Plan

September Plans

29 August 2025 Plan

29 August 2025 Plan

Intro class

3 September 2025 Plan

3 September 2025 Plan

To-do:

- [ ] Email Dr. Sheikhi for a meeting

5 September 2025 Plan

5 September 2025 Plan

To-do:

- [ ] Email Dr. Sheikhi to follow up and remind about previous email from 2 days ago
    - [ ] CC Dr. Maleki
- [ ] Set personal deadlines on calendar
    - [ ] Submit research proposal to mentor by October 9

Goals:

- Set up regular meetings with mentor
- Find project topic
- Research proposal

9 September 2025 Plan

9 September 2025 Plan

To-do:

- [ ] Set deadlines for reading and annotating both papers
- [ ] Do initial background research on wheat head segmentation(Googling, gen AI)
    - [ ] Wheat biology
    - [ ] Object detection
    - [ ] Generative adversarial network(GAN)
    - [ ] Semi-supervised learning
- [ ] Read and annotate paper 1(A Semi-self-supervised Learning Approach for Wheat Head Detection using Extremely Small Number of Labeled Samples)
    - [ ]

Goals:

- Do background research to gain a thorough understanding of wheat head segmentation
    - Break down papers into sections
- Schedule meetings with mentor to discuss future plans
- Narrow down research topic->What are possible areas in the field that can be pursued?
- Begin research proposal->tell mentors about deadlines and goals of the project

-

11 September 2025 Plan

11 September 2025 Plan

To-do:

- [ ] Add five sources to PaperPile library
- [ ] Start a Google Document and add five sources in APA style and AMA style and IEEE style
- [ ] Study citation styles for research proposal
- [ ] Finish reading and annotating paper 1
- [ ] Decide on writing style and summarize rules from Purdue Owl

Goals:

- Go in depth into the topic, especially to familiarize myself with key methods and terms
- Brainstorm questions and possible research directions for the project

15 September 2025 Plan

15 September 2025 Plan

To-do:

- [ ] Take a step back, doing necessary background information on the topic, then taking notes to refer back to when reading and annotating the papers
- [ ] Understand the theory of the papers->what are they proposing, how did they carry it out, and what were the results?

Goal:

- Get ready to discuss and propose project ideas
- Get mentors on board with the deadlines of the project
    - What do we need to do to meet the goals of the project?

17 September 2025 Plan

17 September 2025 Plan

To-do:

- [ ] Following further reading, annotations, and research on both of the papers, summarize notes on papers
- [ ] Compile summary and highlighted paper for mentors
- [ ] Write and send email to mentors
    - [ ] Update progress
    - [ ] Ask questions
    - [ ] Schedule meeting
    - [ ] Lay out research proposal and deadline for meeting discussion

Goal:

- Get ready to discuss and propose project ideas
- Get mentors on board with the deadlines of the project
    - What do we need to do to meet the goals of the project?

19 September 2025 Plan

19 September 2025 Plan

- [ ] Meeting with Dr. Garcia
- [ ] Send reminder email->plan for second reminder on monday given no response
    - [ ] Research proposal intro with guidelines, deadline information
- [ ] Explore other students' papers with Dr. Maleki's mentorship

Goals:

- Get ready to discuss and propose project ideas
- Get mentors on board with the deadlines of the project
    - What do we need to do to meet the goals of the project?

23 September 2025 Plan

23 September 2025 Plan

- [ ] Outline clear agenda for meeting
    - [ ] Questions
    - [ ] Specific explanations of paper(theory, but also coding applications)
    - [ ] Finalize project topic
    - [ ] Research proposal
        - [ ] Questions
        - [ ] Goals
        - [ ] Methods
    - [ ] Timeline->is the October 16 deadline feasible

Goals:

- Get ready to discuss and propose project ideas
- Get mentors on board with the deadlines of the project
    - What do we need to do to meet the goals of the project?

25 September 2025 Plan

25 September 2025 Plan

To-do:

- Prepare agenda for upcoming meeting
- Talk to Dr. Garcia about the research proposal deadline

Goals:

- Get ready to discuss and propose project ideas
- Get mentors on board with the deadlines of the project
    - What do we need to do to meet the goals of the project?

29 September 2025 Plan

29 September 2025 Plan

To-do:

- Complete log book for September submission
- Continue background research
    - Prepare possible:
        - Specific research topic
        - Questions
        - Goals
        - Methodology
- Confirm/change research proposal deadline

Goals:

- [ ] Push hard to get a project topic
    - [ ] Propose one specific possible idea
- [ ] Leave meeting with scheduled next meeting time
- [ ] Leave meeting with enough to do until next meeting
- [ ] Find out next necessary steps after the meeting is done

October Plans

1 October 2025 Plan

To-do:

- [ ] Research honeybee detection
    - [ ]
- [ ] Set up November calendar
- [ ] Write meeting notes

3 October 2025 Plan

- [ ] Send mentor email with meeting notes and questions
    - [ ] Proposed research topic
    - [ ] Questions about methodology and direction
- [ ] Draft working title
- [ ] Write introduction outline

November Plans

5 November 2025 Plan

- Fill in CYSF basic project info and submit ethics form
- Ask mentors about data collection
- Reflect on logbook, plan for how to improve this month

7 November 2025 Plan

- Oral presentation
- Participate in other presentations(notes)
    - Shicheng
    - Anika
- Review research proposal feedback
    - Plan for future improvement, especially for the introduction paper
    - For now, submit for marking

13 November 2025 Plan

- Participate in last oral presentation(notes)
    - Arvind
- Focus on immediate outreach
    - Email individual beekeepers outside Alberta
    - Email associations and guilds(these organizations may have more reach
- Submit CYSF ethics form for approval

17 November 2025 Plan

- Discuss long term plans with mentor
    - Data collection
    - Model training on regular data
    - Data generation
    - Model training on synthetic data
    - Data analysis
    - Conclusion

19 November 2025 Plan

- Send follow-up email to mentors about images, plus rescheduling meeting
- Set up model from another study, to test on received images

21 November 2025 Plan

- Send follow-up email to Rushing River Apiaries
- Email other beekeepers in the US

27 November 2025 Plan

- Send second follow-up email to Rushing River Apiaries
- Start figuring out how to use the public dataset for training
    - Import dataset into Python code with Pandas library

December Plans

January Plans

Background Research

Background Research

General Info

General Info

Key vocabulary:

- U-Net
    - CNN designed for image segmentation
- SAM
    - Segment anything model
    - Zero-shot generalized segmentation
- Deep learning
    - Using multiple layers in a neural network
- Convolutional neural network
    - Type of deep learning that uses convolutions, filters that detect specific features
- Levels of supervision in learning
    - Self-supervised learning
    - Supervised learning
    - Semi-self-supervised learning
- GAN
    - Generative adversarial network
- CycleGAN
    - GAN using two networks for unpaired images
- Diffusion model
    - Similar to a GAN->generative model using the diffusion process
- Adam optimizer
- YOLO
- GWHD
- Domain gap, adaptation
- Segmentation vs. detection
- Segmentation mask
- Measures of model performance(Dice, MaP, IoU)
    - Mean average precision
- Instance segmentation vs semantic segmentation

Varroa Mite:

-

Sources

Paper 1

Paper 1

A Semi-self-supervised Learning Approach for Wheat Head Detection using Extremely Small Number of Labeled Samples

Keyhan Najafian, Alireza Ghanbari, Ian Stavness, Lingling Jin, Gholam Hassan Shirdel, and Farhad Maleki

Summary Notes:

- In order to reduce the need for large amounts of manually annotated images for deep learning, a solution is finding a balance between self-supervised learning and supervised learning with manually annotated images. Using just a few annotated images, we can take a large dataset of unlabeled images and generate a labeled version of it for training.
- Using extremely limited manually segmented wheat head images, it is possible to cut and paste individual wheat heads rotated in multiple directions onto multiple background images. This generates a large dataset of wheat head images that can be used for training.
- These generated images have a domain shift problem, especially when using a limited set of annotated images, so domain adaptation is necessary to make the generated dataset usable for many real world scenarios, where the images taken will have varying levels of blur, contrast, scale, imperfections, etc..
- The models trained on the simulated data, then following domain adaptation, as well as fine tuning with pseudo-labeling and the GWHD dataset, demonstrate the potential of the method of this paper.

Further Notes:

- Computers have great potential to improve food security in agriculture. Much of the success of deep learning used in the real-world depends on large datasets of manually labelled data that require humans to manually label data.
- Self-supervised learning uses the data itself to create labels, requiring no manually labelled data.
- Semi-supervised learning combines labelled and unlabeled data, creating more accurate models than self-supervised learning, and requiring fewer resources than fully supervised learning.
- The specific methods implemented in the paper for wheat head segmentation can be used for other crops, in addition to the supervised learning process of other object detection systems.

- Object detection requires object localizing and classification. YOLO and SSD are one stage detectors that do both of these steps in one stage using a single network to look at the entire image. It does this by dividing each image into a grid and determining what object is in each bounding box and what classification of object it is. R-CNN, Fast

R-CNN, and Faster R-CNN are two stage detectors that first generates location proposals, then classifies and refines the coordinates of the object.
- Efficient models such as EfficientDet balances performance and efficiency.
- The Wheat Head Dataset (GWHD) aims to help increase the size and diversity of wheat head images for wheat head detection.

- The paper creates a simulated dataset, generating a dataset of labeled wheat head images to train detection models on.
- Domain adaptation
    - Step 1: Albumentations library is used to augment the images, using rotations, colour shifts, etc..
    - Step 2: Pseudo-labelling is used to fine tune the model by making predictions on unlabeled images, taking only the highly confident predictions, and using these labels as baselines for further training and fine tuning.

- The results of the semi-self-supervised learning models demonstrate the potential of this proposed method. With domain adaptation and further fine-tuning with pseudo labeling and the GWHD dataset test set, the model achieved a mean average precision of 0.827, achieving comparable results to the baseline model trained with supervised learning on the GWHD dataset.

Paper 2

Paper 2

Efficient Wheat Head Segmentation with Minimal Annotation: A Generative Approach

Jaden Myers, Keyhan Najafian, Farhad Maleki, and Katie Ovens

Summary Notes:

- This paper explores solving a similar problem as the previous paper: reducing the need for large datasets of manually annotated images to use for training wheat head detection models.
- To improve upon methods like semi-self-supervised learning, a slightly different method is proposed, using a modified cycleGAN for domain adaptation, reducing the problem of the domain gap. This method first uses the cut-and-paste method used by Najafian et al. to generate an initial dataset of simulated wheat head images.
- In order to improve the generated dataset for training, a cycleGAN is used to transform the unpaired generated fake images to real wheat head images. This domain adaptation makes the generated images look much more real by training the network with real images, improving it for training wheat head detection models.
- This method uses a modification of the cycleGAN in order to preserve the original location of the wheat heads. Further fine tuning methods include pseudo-labeling.

Further Notes:

- Similar to Najafian et al., this paper uses a cut and paste method to create an initial dataset of generated images from videos of wheat fields and background scenes.
- GANs use two different networks competing against each other, with a generator that tries to produce data indistinguishable from real data, and a discriminator that tries to determine if some given data is fake or real. The networks help to train each other.
- CycleGANs allow us to translate between unpaired fake and real images. By using one network to make a fake image real and a second network to turn the generated real image back into a fake image, we can compare the final generated image to the original fake image.
- For an unmodified cycleGAN, the location and shape of wheat heads is not preserved, so the GAN may change them. The key modification that must be made to the cycleGAN for wheat head segmentation is preserving these details by penalizing the network for differences in the wheat head masks before and after giving the image to the GAN for domain adaptation.

- Using the GAN improved in the Dice and IoU, especially when tested on external data from the GWHD test set and the UTokyo dataset.
- Possible future research directions include utilizing two GAN generators, one for image translation between fake and real images, and one for creating masks and preserving semantic information, or implementing a pipeline to improve the layout of wheat heads in

the cut-and-paste process to better emulate real wheat fields, thereby improving the realism of images.

Paper 3

Paper 3

WheatSAM: A Two-Stage Wheat Head Automatic Segmentation Framework

Md Jaber Al Nahian, Tapotosh Ghosh, Farnaz Sheikhi, Ian McQuillan, and Farhad Maleki

- Using the Meta Segment Anything Model(a generalized model that does object segmentation for non-specific images), and adapting it for wheat head

Mites

Two papers from 2024 and 2025 in Spain

https://www.mdpi.com/2077-0472/15/9/969

- 204 sheets->816 images
    - Dataset has 357 images, all images manually annotated by experts
    - YOLO v11 nano
    - Image augmentation

https://www.mdpi.com/1424-8220/24/12/3828

- Database of 64 images, manually annotated
- Each image divided into smaller images
- Used deblurGAN to improve accuracy
- Image augmentation
- Faster R-CNN

https://zenodo.org/records/10231845 dataset

Mite Sampling

https://www.ontariobee.com/sites/ontariobee.com/files/document/Copy-of-varroa-sampling-EN.pdf

- Alcohol wash
- Ether roll
- Sticky boards

Alcohol wash ether roll both require live bee samples and kill the bees

Sticky board is relatively non-invasive

- Ethics concerns are less than other methods
- Depends on not using methods like oxalic acid for treating mites, otherwise dead mites may fall and make the infestation appear worse than it is

Sticky board can detect very small levels of mite infestations
Sticky board may not be representative of current mite levels

Damage

Non-credible source:
https://news.uoguelph.ca/2022/04/canadian-beekeepers-face-devasting-colony-losses-a-u-of-g-expert-explains-why/

Around 50% of colony losses in Canada are associated with Varroa mites

Inaccessible Sources

Inaccessible Sources

Varroa mite detection using deep learning techniques

A computer vision system to monitor the infestation level of Varroa destructor in a honeybee colony

Detection of Varroa Mites from Bee Images Using YOLO Architecture

A scoping review on the effects of Varroa mite (Varroa destructor) on global honey bee decline

Bee Disease Varroa Prediction: Utilizing Convolutional Neural Networks with Augmentation for Robust Detection and Identification of Honeybee Infection

A survey on performance metrics for object-detection algorithms

Machine learning and computer vision techniques in continuous beehive monitoring applications: A survey

Parasitic mites of honey bees (Apis spp.): A detailed review of Varroa destructor in parasitism, pathogen transmission and its management

Genetics and physiology of Varroa mites

Varroa control: A brief overview of available methods

Analysis of Varroa mite colony infestation level using new open software based on deep learning techniques

Population dynamics of Varroa mite and honeybee: Effects of parasitism with age structure and seasonality

A review of Yolo algorithm developments

Monitoring of Varroa infestation rate in beehives: A simple AI approach

Vit4v: A video classification method for the detection of Varroa destructor from honeybees

Enhancing hive monitoring and sustainability through video recognition and environmental analysis

Research Proposal

Research Proposal

Planning

Planning

Possible ideas:

- Utilizing two GANs, one for image translation between fake and real images, and one for creating masks and preserving semantic information. This differs from using one GAN and incentivising the network to maintain a minimal loss difference mask
- Implementing a pipeline to improve the layout of wheat heads in the cut-and-paste process to better emulate real wheat fields, thereby improving the realism of images
- Using a diffusion model to generate background/wheat head + masks
- Exploring the same methods with other crops
    - Maize tassel
    - Sunflower head
    - Rice->less applicable
- Testing with different amounts of data->different videos, 2, 5, 10 manually segmented images. How much does more data help? When are there diminishing returns?
- Extend the current method of segmentation to phenotyping/disease categorization->how do we use these methods practically?

- Honeybee instance segmentation for bee counting and queen bee identification
    - Instance segmentation to different between regular and queen bees
    - Question: could I use similar methods as in the paper->dataset generation for training?
    - Limited datasets of annotated bee images: The BeeImage Dataset: Annotated Honey Bee Images
    - Project goals: using limited dataset for queen bee detection
        - Do I implement the methods of cut-and-paste, cycleGAN and domain adaptation for database generation?
        - Other ways to improve training with limited data
    - Given enough time, implement into smart glasses for demonstration purposes
        - RayNeo Air 3s
- **Varroa mite counting for threshold determination**
    - **Using object detection and counting on sticky boards for threshold determination**
    - **Data collection and generated data**

Sections:

- Introduction
    - Image detection and segmentation
    - Wheat head
    - Importance of data to learning models

    - Current models, solutions
- Goals
- Research question
- Working title
- Method

Proposal Full Outline

Working Title

Sticky Board Varroa Mite Object Detection and Counting Using Detection Model Trained On a Generated Dataset Using Limited Sample Images

Introduction

Topics:

- Honeybees
    - Importance to food security for pollination
        - Honey bees, *A. Mellifera* are important pollinators that produce food and protect plant biodiversity in natural ecosystems and increase crop production in agriculture, serving as one of the most valuable insects for humans. Papa et al.
        - The Western Honey Bee, *A. Mellifera*, is the main pollinator globally, with 5% of plant species exclusively being visited by *A. Mellifera*. Hung et al.
    - Widespread bee populations globally and across Canada
        - Alberta has the largest beekeeping industry in Canada by honey production(Statistical Overview of the Canadian Honey and Bee Industry 2024)
        - Total economic contribution of $7 billion dollars per year (Statistical Overview of the Canadian Honey and Bee Industry 2024)
    - Contemporary issues
        - Despite growth in the number of hives, the past decade has seen increasing difficulty managing significant colony losses, as seen by decreasing total honey production in Canada from 2023 to 2024.
        - Current causes of issues include land management, environmental conditions, beekeeping practices
            - Anthropological pollution causes toxic substances and heavy metals to accumulate in plants, which then transfers to honeybees through pollination. Similarly, the use of chemical pesticides increases toxic exposure to bees, Hristov et al.
        - Varroa destructor is the most serious threat to *A. Mellifera*, feeding on the fat of adult bees.
        - V. destructor transmits 5 honeybee viruses as a part of its feeding behaviour. Hristov et al., Ramsey et al.
- Varroa destructor mite
    - Overview of what they do to bees
        - V. destructor impacts reproductive capability of bees and overall fitness from shortened lifespan to diminished immune function(Hristov et al., Ramsey et al.).

        - Winter mortality rate of 34.6% in Canada in 2024(Statement On Honey Bee Wintering Losses In Canada For 2024).
    - Methods and effectiveness of treatment methods
        - Organic acids and essential oils, such as formic acid and oxalic acid are effective at stabilizing Varroa mite infestations, however, concern around possible resistance to treatment methods(Kosch et al.)(Kosch et al.)(Rosenkranz et al.) means effective management of Varroa mites is vital to continued efforts for maintaining food security.
        - Synthetic chemicals(Amitraz) face similar concerns with mite resistance (Rosenkranz et al.)
        - In addition to concerns around mite resistance, these treatment methods come with the risk of contaminating bee products or affecting the bees themselves(Rosenkranz et al.).
    - Methods of monitoring infestation levels
        - Methods of monitoring Varroa mites include sticky board, sugar dusting, ether roll, and alcohol wash.(Varroa Mite Monitoring and Sampling Techniques)
        - The majority of beekeepers in Canada regularly monitor for mites, with 90% of Alberta beekeepers monitoring for mites at the end of the beekeeping season(Statement On Honey Bee Wintering Losses In Canada For 2024)
- Deep learning
    - Inspired by the brain, deep learning is a subset of machine learning using multilayered neural networks and massive amounts of data. (Alzubaidi et al.)
    - Deep learning is capable of achieving or surpassing human-level precision and accuracy for a wide range of applications(Krizhevsky et al.)
    - CNNs are multilayered neural networks that can extract low and high level features, making it good for image detection applications. One of the key factors enabling the use of CNNs is increased compute power and data. Notably, Alexnet came out of ImageNet which accumulated over 15 million manually labelled images.
    - Successes, benefits of deep learning and CNNs.
    - There is specific interest in small object detection in images. Mites are extremely small and uniform in size and shape, therefore are ideal for this use case scenario. Small object detection faces many challenges due to the poor resolution, difficulty distinguishing features, and diverse input data due to environmental reasons such as lighting. Given that real world applications of mite detection involve smartphone cameras, which although have high quality cameras in the modern era, may not be perfect for these applications. Methods of addressing this include Feature pyramid networks for multi-scale features, domain adaptation, and data generation.
- Image detection models(yolo, sam)
    - YOLO
    - R-CNN family

        - SSD
        - Feature pyramid networks
        -
        - Model evaluation methods
            - IoU
            - MAp
- Semi-supervised learning
    - For training networks, there are supervised and unsupervised learning. Supervised learning uses labelled data, and is often better for classification rather than exploration. For image detection, supervised learning allows us to directly evaluate the model accuracy with real
- Limited data
    - Manual annotation is expensive. Deep learning is dependent on lots of data that is manually labelled. This requires a high level of time and energy input.
    - In addition, these datasets are often labelled for extremely specific applications.
    - Much of current data centers around European bees, so this contributes to data diversity and Canadian bee data
- Data generation techniques
    - Cut-and-paste is a technique used for generation of synthetic training data for image detection models.
    - Domain adaptation
        - Standard techniques
        - pseudo-labelling
        - We must highlight that this step is especially important for us since we are using limited data for training
        - In addition, we want to make sure that our results are applicable and robust for real-world non-generated data
- Related works
    - Yániz et al.
    - Divasón et al.
    - Najafian et al.

Research Question

Can we apply dataset generation from limited sample data to practical applications of object detection for Varroa mite detection?

Goals

Generate large-scale simulated dataset from limited real sticky board images
Develop semantic segmentation model based on simulated data
Enhance model performance on real data through domain adaptation techniques
Assess model accuracy using real sticky board images
Develop integrated application for image capture and counting

Methodology

Project:

- Image collection
    - Public datasets are available, but we can provide Canadian data with self collection
    - Sticky boards
        - Instructions
        - Impact on bees->ethics
- Manual annotation of limited sample images
- Cut and paste method
    - Works well for uniform mite shapes
- Augmentation
- General learning architecture
    - Different examples
- Validation on test images
    - IoU
    - mAP
    - mAR

Significance

- Application of mite detection practically
- Contribution to field of machine learning, specifically deep learning

Proposal

Proposal

Research Proposal

Calendar

Calendar

September

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
	1	2	3 Class Email mentor	4	5 Class Email mentor	6
7	8	9 Class Dr. Garcia Meeting Background research	10	11 Class Finish reading and annotating paper, show Dr. Garcia to check for details and format	12	13
14	15 Class Log Book update, continue background research	16	17 Class Log book update Summarize notes and annotations, send email to mentors to schedule a meeting	18	19 Class Dr. Garcia Meeting Log book update Continue background research, filling in knowledge gaps on vocab and adjacent research topics	20
21	22	23 Class Log book update Work on research proposal topic finalizing. Prepare 2 in depth possible research topics.	24	25 Class Log book update Prepare agenda for mentor meeting, continue brainstorming 2 more possible research topics	26	27
28	29 Class Log book update Send reminder email Dr. Garcia meeting Mentor meeting	30 Send meeting summary email to mentor Do final research and planning for my project topic				

Month Goals:

- ☐ Rough draft/outline of research proposal
- ☐ Background research
  - ☐ What topics and in what order to be in the intro, methodology, etc.
  - ☐ Outline the introduction(which topics do I need to talk about?)
- ☐ Research question
- ☐ Goals
- ☐ Variables
- ☐ Methods

October

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
			1 Class Complete the RP outline, filling out details once topic and information is finalized Write meeting notes.	2 Complete working title & Introduction->sections split into	3 Class Complete objectives	4 Complete questions
5 Complete methodology	6 Complete Significance Mentor meeting	7 Class Check with Dr. Garcia for errors, citations, and submit RP full draft to Mentor	8	9 Class Make edits to RP, waiting for mentor feedback	10	11
12	13 No school Mentor meeting	14 Class Show outline to mentor, ask further questions about project details. Start writing research proposal.	15	16 Class	17	18
19	20 Class Submit RP draft to mentor	21	22 Class	23	24 Class Get feedback for RP from Dr. Garcia	25
26	27 Mentor meeting	28 Class Research Proposal Due	29	30 Class Mentor meeting Begin project: using current resources, set up coding environment->download existing codebase, libraries	31 Run initial code tests, run model using existing database	

Month goals:

- ☐ Finalize, edit, and submit RP
  - ☐ Working title
  - ☐ Introduction
  - ☐ Objectives
  - ☐ Questions
  - ☐ Methodology
  - ☐ Significance
- ☐ Start project

November

November

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
						1
2	3 Class Oral Presentation Show draft of research proposal to mentor Work on slideshows for oral presentation	4	5 Class Practice oral presentation with Dr. Garcia Finalize research proposal, oral presentation	6	7 Class Submit research proposal and do oral presentation	8
9	10 Confirm next project steps Get data from mentors	11 Class Start creating a detection model Choose and architecture to try	12	13 Class Test out YOLO model on other dataset(non-bee related)	14	15
16	17 Class Discuss long term project timeline with mentors, when to stop collection data	18	19 Class Continue outreach outside Canada to state beekeeper associations	20	21 Class Find public dataset from previous studies on sticky board mites. Test my images with a detector	22
23	24	25 President's breakfast	26	27 Class Download software for manual segmentation, test out on dataset	28 Mentor meeting, discuss project ethics form approval and whether or not to continue data collection	29
30						

Month goals:

- [x] ~~Submit RP~~
- [x] ~~Do oral presentation~~
- [x] ~~Fill CYSF portal~~
  - [x] ~~Basic project info~~
  - [x] ~~Ethics form~~
- [ ] Begin project
  - [ ] Complete data collection
  - [ ] Build model
    - [ ] Use public datasets or collected real images
    - [ ] Use only real data for a baseline model
    - [ ] Generate data, train the final model

December

December

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
	1 Class Start working with a mentor on manual segmentation->how do I do it?	2	3 Class Meeting with Dr. Garcia, discuss progress Finish model with public dataset and regular YOLO model, no augmentation or synthetic data	4	5 Class Evaluate the model on multiple metrics and visualizations for a baseline model	6
7	8 Mentor meeting	9 Class Test the model on more diverse data, including my own data, and then implement augmentation to increase generalizability	10	11 Class Pick 1-2 images to manually segment. Start doing manual cut-and-paste generation with few variations(position and number of mites)	12	13
14	15 Class Learn how to keep the bounding boxes(the generated images need labels of where mites are) in order to train the model	16	17 Class Test out model training with simple generated data combined with existing data. Evaluate impact on model performance	18	19 Class Last day before break, rest of December is break Expand the data generation by augmenting size, orientation, and contrast	20
21	22	23	24	25	26	27
28	29	30	31			

Month goals:

- ☐ Create a functional mite detection model using publicly available datasets
  - ☐ Baseline control model
  - ☐ YOLO model
  - ☐ Use image augmentation to generalize the model, and test it on diverse data
- ☐ Learn how to manually segment, and then generate synthetic images for my dataset
  - ☐ Determine what software works best for segmentation
  - ☐ Learn how to preserve bounding boxes on generated images
  - ☐ Determine the number of generation images needed
- ☐ Continue outreach if I have extra time to get as many images as possible

January

January

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
				1	**2**	3
4	5 End of break	6 Class Study for midterms	7	8 Class Study for midterms	9 Exam Break	10
11	12	13	14	15	16	17
18	19 Mentor Meeting	20 Last midterm exam	21	22 Class Work on progress report	23	24
25	26 Class Mentor Meeting	27	28 Class Implement mentor feedback to retrain model on new dataset split	29	30 Class Logbook Due	31

Month goals:

- ☐ Continue refining baseline model
  - ☐ Apply data augmentation to achieve best performance
- ☐ Apply mentor feedback to improve model

February

February

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	2 Mentor Meeting	3 Class Update Logbook Set up Albumentations for data augmentation Outline Methodology paper	4	5 Class Write Methodology Paper data collection Gather all software, equipment used for project	**6** Train model with new augmentation pipeline	7
8	9 Mentor Meeting Class Write Methodology Paper introduction (adapt introduction paper)	10	11 Class Write Methodology Paper detailed section on procedure Submit paper to Dr. Garcia and mentors for review	12	13 No school Work on poster (decide upon layout and themes)	14
15	16 Mentor Meeting No school	17 Class Create methodology flowcharts with Python	18	19 Class Methodology Paper Due,	20	21
22	23 Mentor Meeting Class Oral Presentation Take notes on oral presentations	24	25 Class Take notes on oral presentations, edit poster based on feedback	26	27 Class Finalize poster, creating final figures for results	28 Print poster via Staples

Month goals:

- ☐ Methodology paper
  - ☐ Outline paper
  - ☐ Individual sections
- ☐ Oral presentation
  - ☐ Slideshow Presentation
  - ☐ Practice with Dr. Garcia
- ☐ CYSF Portal
  - ☐ Problem
  - ☐ Method
  - ☐ Analysis
  - ☐ Conclusion
  - ☐ Citations
  - ☐ Acknowledgement
- ☐ Final poster
  - ☐ Design
    - ☐ Decide themes/colours
    - ☐ Create figures
  - ☐ Printing at Staples

March

March

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	2 School science fair Complete CYSF Portal, including video	3 Away for robotics tournament	4 Away for robotics tournament	5 Away for robotics tournament	**6** Away for robotics tournament	7
8	9 Mentor Meeting Class Submit final Methodology Paper	10	11 Class Research domain adaption techniques to improve synthetic data usability	12	13 Class Outline Discussion Paper	14
15	16 Mentor Meeting	17 Class Finalize results of synthetic data and compare metrics	18	19 Class Work on discussion paper	20	21
22	23 Mentor Meeting Spring break	24 Spring break	25 Spring break	26 Spring break	27 Spring break	28
29	30 Spring break	31 Spring break				

Month goals:

- ☐ Methodology paper
  - ☐ Outline paper
  - ☐ Individual sections
- ☐ CYSF Portal
  - ☐ Problem
  - ☐ Method
  - ☐ Analysis
  - ☐ Conclusion
  - ☐ Citations
  - ☐ Acknowledgement
- ☐ Discussion & Conclusions paper
- ☐ Finalize project work
  - ☐ Synthetic data
  - ☐ Outreach to beekeepers

Oral Presentation

Oral Presentation

Planning Notes

Planning

10 minutes presentation, 5 minutes questions

10% of marks is from participation in other people's presentations. Take notes in log book.

Technically sound and simple enough(concise)

Maybe 10 slides for a powerpoint

Methodology->flow diagram(for overview at least, to show how i answer research question)

Don't use distracting templates

Biorender(software for creating biology diagrams)

Sections:

- Title
- Intro(2-3 slides)
    -
- RQ and Goals(1 slides)
- Methodology)2-3 slides)
- Significance(1 slides)

Presentation Notes

Presentation Notes

Richard:

Evaluating the implementation of copper-molybdenum alloys as heat exchangers in nuclear reactors

- Transmit heat more efficiently in nuclear reactors

Nuclear energy is safe and clean->climate change
Nuclear development has stagnated due to lack of investment

- High construction cost
- Long payout time
- Low operational lifespan
- Previous failures and public opinion->germany canceled reactors due to politics

Radiation makes for bad public perception

Material degradation makes reactor lifespan short

- Wearing down of materials to have less strength, more brittle, more fractures
- Reactor make fail and be inoperable

How a reactor works:

- Reactor core undergoes fission, generates heat with radiation as a byproducts
    - Turn heat into steam, water is good at retaining and transferring heat
    - Steam spins turbine
- Two main areas of heat exchange
    - Materials right now very resistant to radiation, bad at conducting heat

Use copper->very conductive, but not strong and melting point too low

Create alloy to replace current materials

Increase energy efficiency
Lifespan
Increase investment through reliability and capability

What kind of radiation, a lot of reactors use what materials? Uranium, other, differences?
What are cost of molybdenum
Public perception->is the larger issue public perception or like lifespan and efficiency

Jessica

Investigating Changes in Grey Matter Associated with Cognitive Difficulties in Multiple Sclerosis(MS): The Influence of Gender, Age, and Disease Duration

- Changes in thinking and reasoning in MS, and effect of other

Nerve cells

- Myelin sheets insulate axons->improve sending of electrical signals

White matter:

- Myelinated axons
- Daily cognitive activity

Grey matter:

- Cognitive, sensory processing, and motor control

Multiple Sclerosis:

- Destroys myelin sheet and grey matter, disrupting nerve signals

Preventing MS:

- Exercise
- Regulate emotional shift
- Muscle injections

No known cause of MS, but certain living habits may make one more susceptible

What changes in grey matter explain difficulties with cognition

Unique treatment based on personal patients

Shicheng

Evaluating the Effectiveness of 2-Phage and 5-Phage cocktails against multi drug resistant e-coli

MDR are resistant to typical chemical antibiotics

Prolonged exposure to antibiotics results in MDR bacterial infections

MDR bacteria are typically found in hospitals

Proposed solution is to use phages instead of antibiotics

- Phages are viruses that kill bacteria
- Phages don't harm other cells
- Antibiotics target all cells, causing side effects
- MDR are mostly non-resistant to phages
- Phages cannot become resistant to both phages and antibiotics

Methodology;

- 5-phage vs 2-phage cocktail vs control

Phages are taken from phage banks

Anika

Astrocyte Reactivity in the Orbitofrontal Cortex During Diet-Induced Obesity

How your brain is affected by and causes obesity

Astrocytes:

- Cells supporting neurons
- Help transmission
- Maintain health
- Non-neuronal cells
- Too reactive: releases chemicals
- Become reactive if they release too many cytokines

Obesity affects 60-70% of people in North America

Obesity is largely habit based

OFC is a brain region involved with decision making and reward

- Should stop overeating->food is less desirable when full
- Astrocyte activity is believed to affect proper OFC function

Do astrocytes become reactive in the OFC when diet-induced obesity happens

Methodology:

- High vs low fat diet
- Using mice
- Using brain imaging and scanning to look at reactivity of astrocytes

DREADD is a receptor injected into the brain. It is let to develop. A drug is then given to react with the receptor and stimulates reactive astrocytes

DCZ needs DREADD to make astrocytes reactive

We can “turn on” astrocytes in experiment 2 using DCZ and DREADD

- Only astrocytes are affected
- Astrocytes are sensitive to this drug

In experiment 2:

- All mice are controlled, they have the same diet

- This doesn’t specifically connect to obesity, just measuring something that is believed to impact obesity

Astrocytes surround neurons

Understanding astrocytes
How high fat diets cause obesity
Understanding why people become obese

Arvind

Examining the relationship between stretch induced syncope and something

Syncope
- Faint due to no blood in the brain
    - Cardiovascular issue
- Common, most people experience this in their life

Stretch-induced syncope
- Rare, poorly understood
- Triggered by stretching
- Limited documentation

Theories:
- Vertebral artery compression? Little evidence
- Valsalva-esque strain? Little evidence
- Carotid baroreceptor activation? No evidence

Most recent work on SIS links it to muscle mechanoreceptor activation in upper-extremities
We observe how

Hemodynamics
- Flow of blood throughout the body
- Eg blood pressure, heart rate
- Cardiac output

Rq
Does upper extremity stretch induce a transient decrease in BP? Hypothesis is yes
Is hypotension induced in stretch due to decreased venous return and SV(heart not pumping enough), or due to arterial vasodilation(pressure of arteries on blood)? Hypothesis is both
Is orthostatic stress necessary for induced hypotension?
- Blood pooling in lower body

Testing standing vs laying down

Variables:
- Body position
- SIS prevalence
- Dependent are hemodynamic outcomes

- Controlled are medications, eating, stretch procedure
- Confounding: sex & age due to low numbers of patients

Testing protocol:
- Stretching upper extremities

Determine significance

SIS has very little research
- 40 documented cases, more prevalent
- Inform how to avoid SIS onset for patients
- Provide future framework for research

February presentations

Andi

Sarcasm detection

Using ML to differentiate between sarcasm and sincerity->voice systems

Neural networks are expensive->features not based on researched features

No system has utilized the sarcasm types

Applications:
- Voice assistants
- Application for people who can't distinguish sarcasm(weaker hearing, autism spectrum)

Research question:
- Simple ML using linguistic research features to distinguish sarcasm
- Logistic regression model with local features

Methodology:
- Sitcom audio with labels about sarcasm/not sarcasm
- Manual labelling

Legendre coefficients:
- No significance in any models

Pitch and intensity aren't distinct enough
Speech rate is a weak predictor

Model performance similar to neural network performance

Dataset:
- English speakers only, american english only

Shaayaan

Aluminum silicon carbide metal matrix composites

- Combining the metals to create smth stronger
- Lighter weight plus more strength
- Stainless
- Wear resistant

Pellets

Research question:

- Difference percentages of silicon carbide
- Hypothesis: amt of SiC is proportional to hardness

Methodology:

- Mixture by weight percentage
- Powders used(uniform mixture of powders achieved with mortal and pestle)
- Cold press uses pressure + vacuum to form pellets
- Atomic force microscopy(3d map of surface) -> arrangement of particles
- Nano indentation

Results:

- Promising results
- More SiC is larger gaps
- Increase in hardness of 87.6%

E.g.: jet engine blade

Anika

Astrocyte reactivity in the orbitofrontal cortex during diet-induced obesity

Obesity is largely a problem of habits
Obesity:

- Multiple issues

Astrocytes:

- Brain cells that surround and protect neurons
- Fats in the bloodstream are sensed by astrocytes->hypertrophic and reactive
- OFC is responsible for decision making and reward evaluation
- Reactivity->more projections, more inflamed, more chemicals released

Do astrocytes become reactive when mice are fed a high-fat diet

Analysis:

- 20x confocal microscope
- SOX shows cell body
- S100 shows projections

Results:

- Cell numbers increased with high fat diet
- Cell intensity not significant(what does this mean practically)

Samir

Histological analysis of kidney tissue structural differences between healthy and type 2 diabetic female goto-kakizaki rats

Diabetes increasing rapidly

Diabetes:

- Kidneys have decreased function

Lack of female data(different hormonal ratios)

Artificially induced diabetes via evolution
Sample size of 1(however multiple samples of kidney tissue)

Research question: how does kidney tissue differ between non-diabetic and diabetic rats

Methodology:

- Kidney isolated
- Glomerulus isolation

Shicheng

Most prokaryotic bacteria can be treated w/ antibiotics
Developing antimicrobial resistance

Research question:

- Using bioinformatic processing, attempting to analyze AMR.

What amr genes present in each isolate?
Which antiviral defense systems were produced?

Results and conclusion

- AMR genes aad and aph detected

Outreach

Outreach

Alberta
Sticky Board Image Collection For a Computer Science Research Project

To Whom It May Concern,

Good day! My name is Ronald Li and I am a grade 12 student at Webber Academy working on a science fair research project with Dr. Sheikhi and Dr. Maleki at the University of Calgary. Our goal is to develop an AI-powered tool to combat Varroa mites by helping beekeepers better monitor Varroa mite infestations.

I am aware that many beekeepers in Alberta have problems with Varroa mites that have caused a large number of damage and colony losses in the past few years. The project I am working on is using smartphone images of sticky boards to create a noninvasive automatic Varroa mite detection model to quickly count and monitor mite levels for beekeepers.

To train the detection model on, we need sticky board images from Alberta colonies, which requires beekeepers that are currently in the following weeks or months monitoring for mites using sticky boards. My project focuses on using special techniques to work with limited data, so even 5-10 images from your operation would be greatly appreciated.

I am reaching out to ask if you would be willing to help me with my project in either taking photos of your sticky boards when you check them, or letting me visit your operation to photograph the boards myself as is convenient for you.

The data will be completely anonymized, with all images I collect just being identified as honey bee colonies in Alberta.

If you do not have sticky boards, the University of Calgary and I are open to the possibility of supplying sticky board supplies or funds as a courtesy for helping with this project.

I understand that it is getting colder in Alberta and that the timing is not ideal, but any images from sticky board monitoring in the next few weeks would be incredibly helpful.

Otherwise, if you are not able to participate, I would still greatly appreciate the opportunity to have you test our detection model once my project is completed in the coming spring or summer.

Please feel free to contact me at this email or at my phone number 403-988-5435. You can also contact my school mentor Dr. Beatriz Garcia-Diaz or my mentors at the UofC, Dr. Farhad Maleki and Dr. Farnaz Sheikhi, with the following emails:

bgarcia-diaz@webberacademy.ca
farhad.maleki1@ucalgary.ca
farnaz.sheikhi@ucalgary.ca

Thank you again for considering this request.

Yours sincerely,
Ronald Li

External outreach

Canada-wide
Sticky Board Image Collection For a Computer Science Research Project

To Whom It May Concern,

Good day! My name is Ronald Li and I am a grade 12 student at Webber Academy in Calgary, Alberta, working on a science fair research project with Dr. Sheikhi and Dr. Maleki at the University of Calgary. Our goal is to develop an AI-powered tool to combat Varroa mites by helping beekeepers better monitor Varroa mite infestations.

I am aware that many beekeepers in Canada have problems with Varroa mites that have caused a large number of damage and colony losses in the past few years. The project I am working on is using smartphone images of sticky boards to create a noninvasive automatic Varroa mite detection model to quickly count and monitor mite levels for beekeepers.

To train the detection model on, we need sticky board images from diverse geographic locations and colony conditions. Most existing research datasets come from Europe, so including data from BC is helpful for developing a reliable counting tool across different beekeeping environments.

My project focuses on using special techniques to work with limited data, so even 5-10 images from your operation would be greatly appreciated.

I am reaching out to ask if your colony is using sticky boards for mite monitoring in the upcoming few weeks or months, and if you could help me with my project by taking photos of your sticky boards when you check them.

The data will be completely anonymized, with all images I collect just being identified as honey bee colonies.

If you do not have sticky boards, the University of Calgary and I are open to the possibility of supplying sticky board funds as a courtesy for helping with this project.

I understand that it is getting colder and that the timing is not ideal, but any images from sticky board monitoring in the next few weeks would be incredibly helpful.

Otherwise, if you are not able to participate, I would still greatly appreciate the opportunity to have you test our detection model once my project is completed in the coming spring or summer.

Please feel free to contact me at this email or at my phone number 403-988-5435. You can also contact my school mentor Dr. Beatriz Garcia-Diaz or my mentors at the UofC, Dr. Farhad Maleki and Dr. Farnaz Sheikhi, with the following emails:

bgarcia-diaz@webberacademy.ca
farhad.maleki1@ucalgary.ca
farnaz.sheikhi@ucalgary.ca

Thank you again for considering this request.

Yours sincerely,
Ronald Li

California
To Whom It May Concern,

Good day! My name is Ronald Li and I am a grade 12 student at Webber Academy in Calgary, Alberta, working on a science fair research project with Dr. Sheikhi and Dr. Maleki at the University of Calgary. Our goal is to develop an AI-powered tool to combat Varroa mites by helping beekeepers better monitor Varroa mite infestations, one of CSBA's top priorities in saving the bees.

I recently read your November 2025 White Paper identifying Varroa destructor as the leading cause of colony losses in California, and I was particularly interested in your discussion of AI as a potential solution to this challenge. In Alberta, Canada's largest honey producer, Varroa mites are similarly the most critical threat to colony survival, which is why my university supervisors have focused their previous research on this problem.

Our project uses smartphone images of sticky boards to create an automated mite detection model that can count mites faster than the 15-30 minutes required for manual counting. By making monitoring faster and more accessible, we aim to encourage the frequent checking necessary for effective integrated pest management.

To train the detection model on, we need sticky board images from diverse geographic locations and colony conditions. Most existing research in sticky board detection use datasets from Europe, so including data from your region would help develop a reliable counting tool across different beekeeping environments.

Would CSBA be willing to share this opportunity with your member clubs and associations? I'm seeking out beekeepers who would be willing to:

- Photograph their sticky boards with smartphones during their normal monitoring routine
- Share 5-10 images (even from a single monitoring session would be valuable)
- All data will be completely anonymized and identified only as "California honey bee colonies"

Given that most Alberta beekeepers have already stopped monitoring for winter, any images collected from Californian colonies in the coming weeks would be incredibly helpful for training our model.

Beekeepers who cannot participate now are also welcome to test our detection model once completed in spring/summer 2026.

Please feel free to contact me at this email or at my phone number 403-988-5435. You can also contact my school mentor Dr. Beatriz Garcia-Diaz or my mentors at the UofC, Dr. Farhad Maleki and Dr. Farnaz Sheikhi, with the following emails:

bgarcia-diaz@webberacademy.ca
farhad.maleki1@ucalgary.ca
farnaz.sheikhi@ucalgary.ca

Thank you for considering my research request. CSBA's leadership in addressing Varroa challenges makes your members' participation particularly valuable to developing tools that can help reduce California's colony loss rates.

I am happy to provide additional information or answer any questions about the project.

Yours sincerely,
Ronald Li

BC
Dear BC Honey Producers Association Leadership,

My name is Ronald Li, a Grade 12 student at Webber Academy in Calgary, working on a computer science research project with Dr. Farnaz Sheikhi and Dr. Farhad Maleki at the University of Calgary.

I recently read about BCHPA's support for Dr. Erika Plettner's work on novel compounds to control Varroa at Simon Fraser University, your previous commitment to innovative solutions for Varroa management. Our project takes a complementary approach, developing an AI-powered tool that automates Varroa mite counting from sticky board images to make monitoring more accessible for BC beekeepers.

Our project uses smartphone images of sticky boards to create an automated mite detection model that can count mites faster than the 15-30 minutes required for manual counting. By making monitoring faster and more accessible, we aim to encourage the frequent checking necessary for effective integrated pest management.

To train the detection model on, we need sticky board images from diverse geographic locations and colony conditions. Most existing research in sticky board detection use datasets from Europe, so including data from your region would help develop a reliable counting tool for Canadian beekeeping environments.

Given that most Alberta beekeepers have already stopped monitoring for winter, BC's milder climate, particularly in the Fraser Valley, makes your region uniquely valuable for our November data collection needs.

Would BCHPA be willing to share this research opportunity with your member clubs and regional associations? I'm particularly hoping to connect with Fraser Valley beekeepers who may still be actively monitoring. I’m seeking out beekeepers who would be willing to:

- Photograph their sticky boards with smartphones during their normal monitoring routine
- Share 5-10 images (even from a single monitoring session would be valuable)
- All data will be completely anonymized and identified only as "BC honey bee colonies"

Any images collected from BC colonies in the coming weeks would be incredibly helpful for training our model.

Beekeepers who cannot participate now are also welcome to test our detection model once completed in spring/summer 2026.

Please feel free to contact me at this email or at my phone number 403-988-5435. You can also contact my school mentor Dr. Beatriz Garcia-Diaz or my mentors at the UofC, Dr. Farhad Maleki and Dr. Farnaz Sheikhi, with the following emails:

bgarcia-diaz@webberacademy.ca
farhad.maleki1@ucalgary.ca
farnaz.sheikhi@ucalgary.ca

Thank you for considering my research request. BCPHA's previous support of research in addressing Varroa challenges makes your members' participation particularly valuable to developing tools that can help reduce colony loss rates.

I am happy to provide additional information or answer any questions about the project.

Yours sincerely,
Ronald Li

Florida
To Whom It May Concern,

Good day! My name is Ronald Li and I am a grade 12 student at Webber Academy in Calgary, Alberta, working on a science fair research project with Dr. Sheikhi and Dr. Maleki at the University of Calgary. Our goal is to develop an AI-powered tool to combat Varroa mites by helping beekeepers better monitor Varroa mite infestations.

I recently learned about your Save the bees plates, and think it's a great way to fundraise and promote awareness for bees and the FSBA. In Alberta, Canada's largest honey producer, Varroa mites are similarly the most critical threat to colony survival, which is why my university supervisors have focused their previous research on this problem.

Our project uses smartphone images of sticky boards to create an automated mite detection model that can count mites faster than the 15-30 minutes required for manual counting. By making monitoring faster and more accessible, we aim to encourage the frequent checking necessary for effective integrated pest management.

To train the detection model on, we need sticky board images from diverse geographic locations and colony conditions. Most existing research in sticky board detection use datasets from Europe, so including data from Florida would help develop a reliable counting tool across different beekeeping environments.

Would FSBA be willing to share this opportunity with your member clubs and associations? I'm seeking out beekeepers who would be willing to:

- Photograph their sticky boards with smartphones during their normal monitoring routine
- Share 5-10 images (even from a single monitoring session would be valuable)
- All data will be completely anonymized and identified only as "Florida honey bee colonies"

Given that most Alberta beekeepers have already stopped monitoring for winter, any images collected from Floridian colonies in the coming weeks would be incredibly helpful for training our model.

Beekeepers who cannot participate now are also welcome to test our detection model once completed in spring/summer 2026.

Please feel free to contact me at this email or at my phone number 403-988-5435. You can also contact my school mentor Dr. Beatriz Garcia-Diaz or my mentors at the UofC, Dr. Farhad Maleki and Dr. Farnaz Sheikhi, with the following emails:

bgarcia-diaz@webberacademy.ca
farhad.maleki1@ucalgary.ca
farnaz.sheikhi@ucalgary.ca

Thank you for considering my research request. Your members' participation is particularly valuable to developing tools that can help reduce Florda's colony loss rates.

I am happy to provide additional information or answer any questions about the project.

Yours sincerely,
Ronald Li

Pennsylvania

To Whom It May Concern,

Good day! My name is Ronald Li and I am a grade 12 student at Webber Academy in Calgary, Alberta, working on a science fair research project with Dr. Sheikhi and Dr. Maleki at the University of Calgary. Our goal is to develop an AI-powered tool to combat Varroa mites by helping beekeepers better monitor Varroa mite infestations.

I read the Winter Colony Survival Analysis report done on data from your surveys, and appreciated your contribution to that research. The findings correlating mite treatment, and mite monitoring, to colony survival, is one of the main reasons why I chose my project idea. In Alberta, Canada's largest honey producer, Varroa mites are the most critical threat to colony survival, which is why my university supervisors have focused their previous research on this problem.

Our project uses smartphone images of sticky boards to create an automated mite detection model that can count mites faster than the 15-30 minutes required for manual counting. By making monitoring faster and more accessible, we aim to encourage the frequent checking necessary for effective integrated pest management.

To train the detection model on, we need sticky board images from diverse geographic locations and colony conditions. Most existing research in sticky board detection use datasets from Europe, so including data from Pennsylvania would help develop a reliable counting tool across different beekeeping environments.

Would PSBA be willing to share this opportunity with your member clubs and associations? I'm seeking out beekeepers who would be willing to:

- Photograph their sticky boards with smartphones during their normal monitoring routine
- Share 5-10 images (even from a single monitoring session would be valuable)
- All data will be completely anonymized and identified only as "Pennsylvania honey bee colonies"

Given that most Alberta beekeepers have already stopped monitoring for winter, any images collected from Pennsylvanian colonies in the coming weeks would be incredibly helpful for training our model.

Beekeepers who cannot participate now are also welcome to test our detection model once completed in spring/summer 2026.

Please feel free to contact me at this email or at my phone number 403-988-5435. You can also contact my school mentor Dr. Beatriz Garcia-Diaz or my mentors at the UofC, Dr. Farhad Maleki and Dr. Farnaz Sheikhi, with the following emails:

bgarcia-diaz@webberacademy.ca
farhad.maleki1@ucalgary.ca
farnaz.sheikhi@ucalgary.ca

Thank you for considering my research request. Your members' participation is particularly valuable to developing tools that can help reduce Pennsylvania’s colony loss rates.

I am happy to provide additional information or answer any questions about the project.

Yours sincerely,
Ronald Li

Philadelphia
To Whom It May Concern,

Good day! My name is Ronald Li and I am a grade 12 student at Webber Academy in Calgary, Alberta, working on a science fair research project with Dr. Sheikhi and Dr. Maleki at the University of Calgary. Our goal is to develop an AI-powered tool to combat Varroa mites by helping beekeepers better monitor Varroa mite infestations.

I read the page on your website about your March monthly meeting speaker Dr. Phil Lester who is working on a biochemical approach to Varroa mite treatment, and contributing to tackling this issue is one of the main reasons why I chose my project idea. In Alberta, Canada's largest honey producer, Varroa mites are the most critical threat to colony survival, which is why my university supervisors have focused their previous research on this problem.

Our project uses smartphone images of sticky boards to create an automated mite detection model that can count mites faster than the 15-30 minutes required for manual counting. By making monitoring faster and more accessible, we aim to encourage the frequent checking necessary for effective integrated pest management.

To train the detection model on, we need sticky board images from diverse geographic locations and colony conditions. Most existing research in sticky board detection use datasets from Europe, so including data from Philadelphia would help develop a reliable counting tool across different beekeeping environments.

Would the Philadelphia Beekeepers Guild be willing to share this opportunity with your member beekeepers? I'm seeking out beekeepers who would be willing to:

- Photograph their sticky boards with smartphones during their normal monitoring routine
- Share 5-10 images (even from a single monitoring session would be valuable)
- All data will be completely anonymized and identified only as "Pennsylvania honey bee colonies"

Given that most Alberta beekeepers have already stopped monitoring for winter, any images collected from Pennsylvanian colonies in the coming weeks would be incredibly helpful for training our model.

Beekeepers who cannot participate now are also welcome to test our detection model once completed in spring/summer 2026.

Please feel free to contact me at this email or at my phone number 403-988-5435. You can also contact my school mentor Dr. Beatriz Garcia-Diaz or my mentors at the UofC, Dr. Farhad Maleki and Dr. Farnaz Sheikhi, with the following emails:

bgarcia-diaz@webberacademy.ca
farhad.maleki1@ucalgary.ca
farnaz.sheikhi@ucalgary.ca

Thank you for considering my research request. Your members' participation is particularly valuable to developing management tools that can help reduce Pennsylvania's colony loss rates.

I am happy to provide additional information or answer any questions about the project.

Yours sincerely,
Ronald Li

Responses

Nov 4, 2025, 9:16 PM (11 hours ago)

Terry Greidanus

to **gstrom@me.com**, me

Hi Ronald:

Thanks for reaching out to me with your research project, coming up with an AI app to help us beekeepers count varroa mites. Unfortunately, our bee season is now over, and we will not be looking into our beehives until spring. I also have not used sticky boards for a number of years. We seem to prefer using an alcohol wash to decide whether or not we have a mite problem. I may have pictures of used sticky boards somewhere, but I will need to go back a number of years to find them. I can forward you these at a later time if I run across them. I will cc my sister Grace Strom in this email, as she has also used Sticky boards in the past.

For your information most beekeepers are now finished with their beehives for the season and will not be doing any inspections again until spring or sometime in March.

Thanks

Terry Greidanus
Mountainview Honey
Box 126
Cayley AB
T0L 0P0
403-809-0122

From: Ronald <li.ronald2008@gmail.com>
Sent: Monday, November 3, 2025 12:07 PM
To: Terry Greidanus <terry@mountainviewhoney.com>; Yetta Greidanus <yetta@mountainviewhoney.com>
Subject: Sticky Board Image Collection For a Computer Science Research Project

9:11 AM (0 minutes ago)

Rona
ld

to
Terry

Dear Mr. Greidanus,

Thank you for your email and advice on my project. I plan on asking other beekeepers in Alberta, but I understand it is difficult because beekeepers are wintering bees now. Do you happen to know about any other beekeepers in Alberta who might still be monitoring for mites right now using sticky boards? Otherwise, do you think it might be possible to approach beekeepers in BC or another province? I appreciate all the help.

Thanks again,
Ronald Li

Hi Ronald

I don't believe that anyone in the province is doing anything with their beehives now. Everyone is or should be finished getting their beehives wrapped for winter. The Finish line date is considered to be November one.. We finished October 15. The normal time for this is in late august through mid-September. To get a feel for the varroa problem. If there is a varroa problem now beehive is already dead even though the population looks healthy even if all the Varroa were instantly killed. This monitoring and control window is now closed for us. and in my opinion for all of Canada. The reason being the beehive needs to create healthy winter bees, and this can only be done with low varroa counts before the fall season.

There may be the odd late beekeeper in the Frazer valley in BC that is still doing varroa control now, you would need to talk to the provincial apiculturist there to find out what he knows

Terry

Bottom board photos

Inbox

3:41 PM

Rushing River <rushingriverapiaries@gmail.com>

to me

Hi Ronald,
I received your email from the BCHPA and I would love to provide some bottom board photos for your research. Does it matter if we have done a treatment recently or when we last cleaned them? Can you tell me if there is any specific way to get the photos you need?
Christine McDonald

Does it matter if we have done a treatment recently or when we last cleaned them? Can you tell me if there is any specific way to get the photos you need?

Dear Christine,

Thanks so much for reaching out and offering to help, I really appreciate it.

To answer your questions, I'm interested in sticky boards at all stages, whether you have done a treatment recently or not, and actually collecting data with a variety of mite levels will help make the model more robust. Similarly, don't worry about when they were last cleaned, sticky boards with typical levels of debris like wax are great. I'm looking to use sticky boards from real-world conditions without any special preparation.

To take the photos, just use a smartphone camera, taking the photos from directly above the board (around 1-2 feet away). If you can't fit the full board clearly in one photo, feel free to take multiple photos covering different sections of the board, overlapping them slightly. Good lighting helps, and don't worry about perfect quality, since we hope to train the model for real-world conditions.

I've attached two sample images from published research in Spain to help give an idea of what photos I'm looking for. Please keep in mind that your boards may look different from these examples. For example, you might notice that they divide each board into sections using green string, but you don't need to do that.

The beekeepers I contacted here in Alberta stopped monitoring in early/mid October, so I wasn't sure if beekeepers in your area are still actively monitoring or if you've already winterized for the season.

If you have boards from monitoring, those would be incredibly helpful. Images of anywhere from 2-10 different boards would be valuable.

If you're still actively monitoring, any boards you check in the coming weeks would be amazing. Even just sending photos whenever you do your regular checks (no special monitoring needed!) would help build a more diverse dataset.

Either way, whatever works with your monitoring schedule is perfect, since I'm grateful for any images you're able to share!

You can email the photos directly to this email address, or use another file transfer software if it works better for you.

Thanks again for supporting my project, it means a lot.

Best regards,
Ronald Li

Dear Christine,

Hi, I shared the images with my project mentors at the UofC, and they have let me know that the images may not work since they are too noisy. I must apologize, since I underestimated how clean the sticky boards must be. Would you be willing to clean your boards and send me images of some of your sticky boards after a few days?

In addition, could I ask for some details about your sticky boards for the project, such as the dimensions, how many days of mite fall, and the sticky material used?

Thanks a lot,
Ronald Li

Dear Christine,

Good morning, it's been a few weeks, and I wanted to email you again to ask if you are able to provide the details about your sticky boards. I understand if you are busy, but I wanted to ask again in case you missed my last email. Thanks.

Sincerely,
Ronald Li

Hi Ronald,
I agree with Eric and would add that mite testing season is pretty much over here as well.

However, I personally am interested in what you're doing and would be willing to start using sticky boards and sending you images.

My understanding of sticky boards is that they are not effective for evaluating mite infestation levels because the number of mites on any given board are dependent of the inspection interval and population size of the colony. It seems that the preparation of the board could add another confounding factor.

If your approach somehow addresses these concerns I'm interested. Most of us here use alcohol wash to do mite counts. Pulling a sticky board and taking a picture seems a lot easier if it can work.

So I'm open to helping you in general and I think other people might be as well. A detailed description of what you'd like us to do would be helpful - what kind of sticky board, how to prepare it, inspection interval, other data needed, etc. A step by step approach would let us decide if we can comply with it and evaluate the effort involved. If it seems feasible we can do a presentation at our monthly meeting or a post on our web site.

Yours

Dave Harrod
Past President
Philadelphia Beekeepers Guild
http://www.phillybeekeepers.org

> On Nov 19, 2025, at 6:01 PM, eric schoefer <ericschoefer@gmail.com> wrote:
>
> Hi Ronald,
>
> Sounds like an interesting project. I believe that most beekeeper members of the Philadelphia Beekeepers Guild do not use sticky boards to monitor varroa. Most members use alcohol washes, sugar rolls, or uncapping of drone brood. I imagine we could make your request to members to take images of sticky boards, but I don't predict you will get many beekeepers who are already monitoring or using sticky boards. It may be more likely to get some beekeepers who would be willing to set up sticky boards next season and take images as you request. It sounds like this would be too late for your project, though.
>
> This is just my opinion. If you would like, we can pass your request on to our members and see if there are any takers.

Peace,
eric

Sent from my iPhone

CYSF Ethics

Ethics Form

APPROVED

Description of experiment or research study

The first step in the project is data collection. With the help of beekeepers, I will use the sticky board methodology to collect data on mite infestations. A beehive with a screened bottom board will have a wide board with a sticky material placed underneath it such as petroleum jelly. The board is left there for a few days, and mites that fall off the bees stick to the board. The board is taken out, and I will take images of the boards using a smartphone camera. I chose this methodology since it is noninvasive and does not harm the bees. After I collect images of the sticky boards, the rest of my project is done digitally. I will generate synthetic data using the sticky board images, then train a mite detection neural network.

Evaluating Test Images

Using VarroDetector on one sample image I received(split into four quadrants)

Progress Report

Progress Report 1
Progress Report 2

Methodology Paper

Methodology Paper

Outline

Quick intro

Data collection(referencing the paper that collected the data)

Describe general process (data cleaning, training and model configuration, augmentation)

- Separate baseline model, augmentations
    - Model parameters

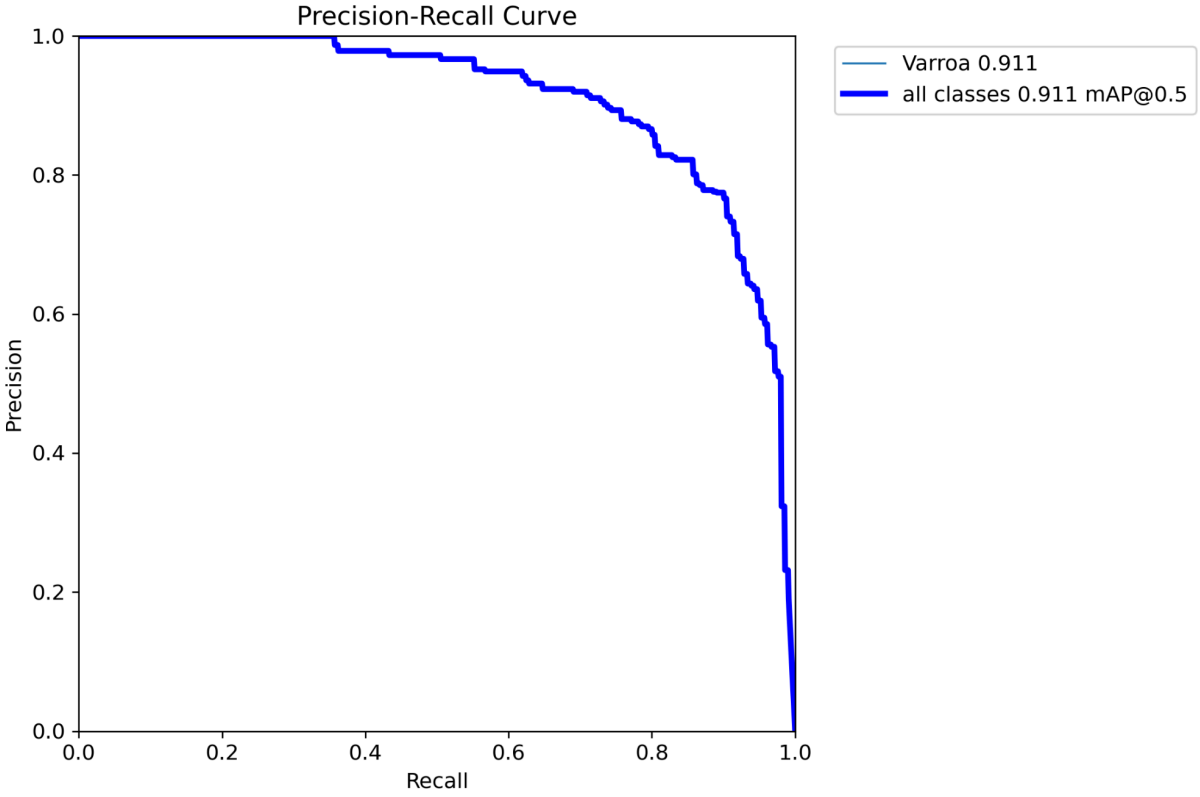
Computing:

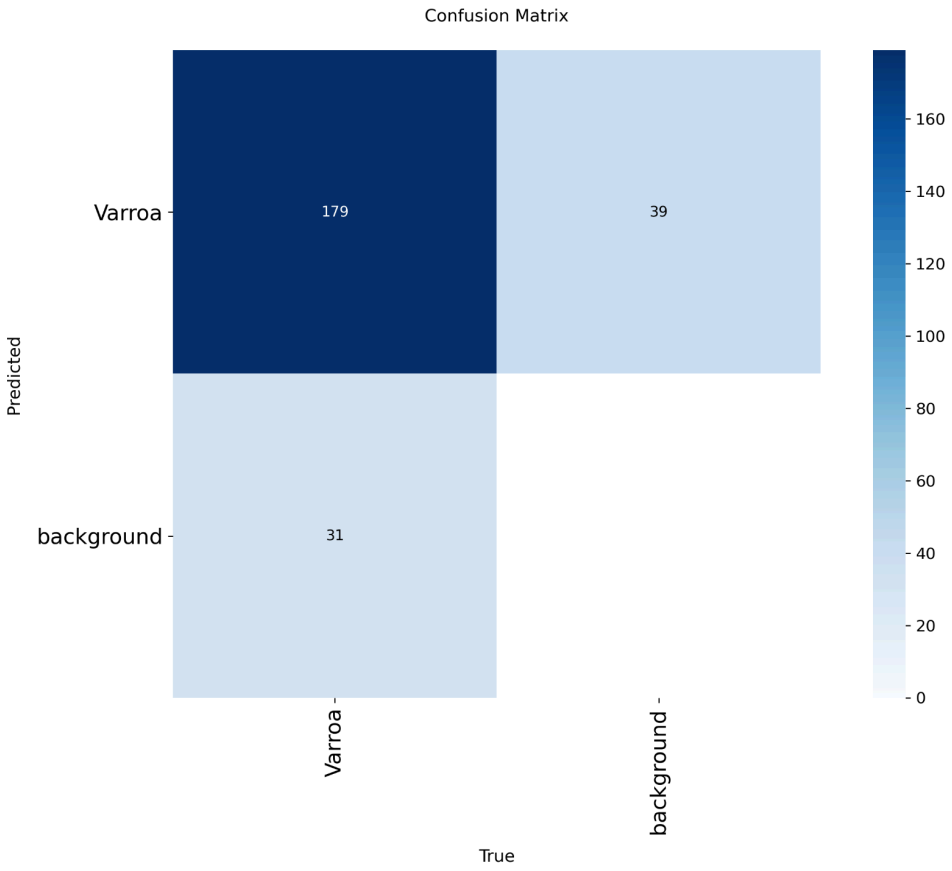
- Computer with:
    - Radeon 6700XT(12GB VRAM), with ROCm driver
    - 32GB RAM
    - AMD Ryzen 9 7900 CPU
    - Arch linux distribution
    - Pyenv, 3.10
    - Unsupported gpu->specific commands needed
    - yolov11(sept 2024, ultralytics)
    - Yolov11 small model
        - 9.4M parameters
- Python libraries:
    - Ultralytics version()
    - Albumentations version
    - Pytorch 2.10.0 ROCm 7.1
    - yolosplitter(august 2024)

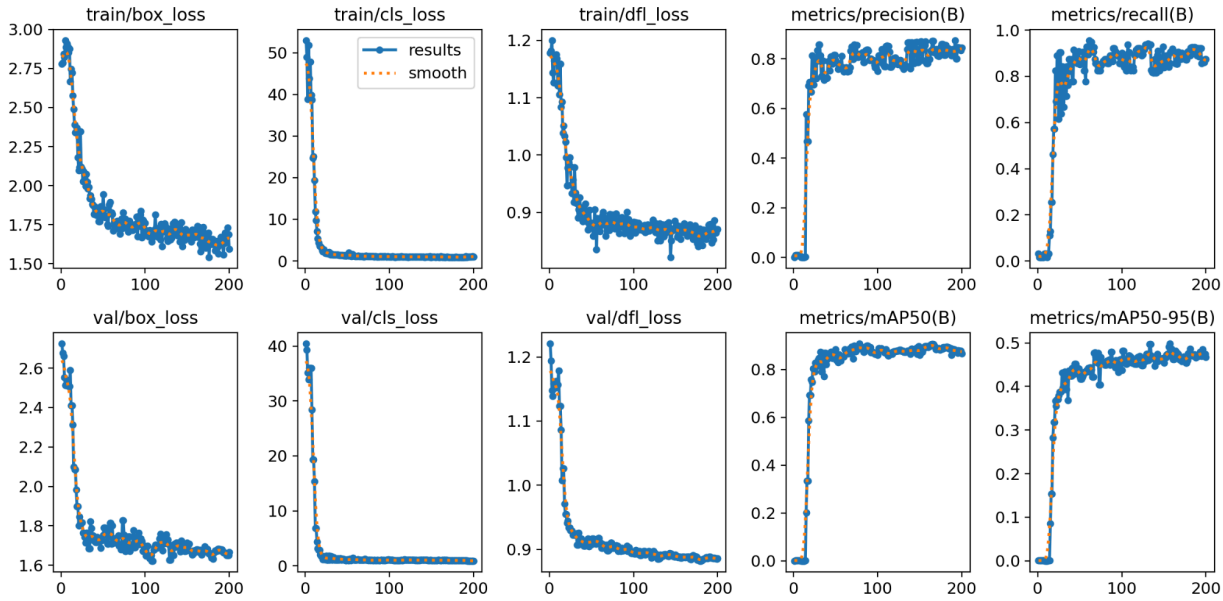
Data:

- 64 images of 8064x6048 resolution smartphone images. Images manually annotated
- Randomly split into 70 10 20 split for training, validation, and testing. (45, 6, 13)
- Metrics: mAP, IoU, R & P, accuracy
- Rather than improving image quality, we focus on data augmentation
- Albumentations
    - Size normalization->effectively increasing resolution w/o sacrificing model size or speed
    - Geometric invariances
    - dropout->minimal for small object
    - Affine
    - Robustness augmentations for real-world applications
- Augmentations applied to images and then passed to yolo->applies training hyperparameters

Final Run Curves


Precision-Recall Curve
Varroa 0.911
all classes 0.911 mAP@0.5
Precision
Recall
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


Confusion Matrix
Varroa
background
Predicted
179
39
31
Varroa
background
True
160
140
120
100
80
60
40
20
0


train/box_loss
train/cls_loss
results
smooth
train/dfl_loss
metrics/precision(B)
metrics/recall(B)
val/box_loss
val/cls_loss
val/dfl_loss
metrics/mAP50(B)
metrics/mAP50-95(B)