

Detektering och klassificering av bin med hjälp av maskininlärning

Jamal Aldiwani, Filip Anjou, Gustaf Bodin, Hamid Hosseini
Ludvig Lemner, Erik Nilsson

2022-01-11

Innehåll

1	Introduktion	1
1.1	Mål	1
1.2	Syfte	1
2	Bakgrund	1
3	Metod	2
3.1	Etikettering	2
3.2	Träning	2
3.3	Utvärdering av tränade modeller	3
4	Resultat	3
5	Slutsatser	5
5.1	Förbättringar	6
5.2	Lärdomar	6
	Referenser	7

Ordlista

Augmentering - Modifiering av träningsdata för att bredda datan.

Etikettering - Manuell markering och klassificering av föremål på träningsdatan.

Google Colab - En Jupyter-notebookmiljö som körs i webbläsaren med Google Cloud. Tillåter att kombinera körbar kod och texter i ett enda dokument, tillsammans med bilder, HTML, LaTeX och mer.

Loss - Värde representerande felsvar. Beräknas genom att jämföra en bilds faktiska etiketter med de genererade av en modell.

Sample - Träningsdatans minsta beståndsdel. I vårt fall en bild med tillhörande handgjorda etiketter.

YOLOv3 - En arkitektur för neuronnät avsedd för objektdetektering.

1 Introduktion

Bin står för en väsentlig del av ekosystemet. De pollinerar blommor och grödor, vilket är en nödvändighet för oss människor och många andra arter på jorden. Bipopulationen håller drastiskt på att minska, bland annat som effekt av bekämpningsmedel, sjukdomar och parasiter. Därför är det viktigt att kunna samla in och analysera data över bisamhällen, för att förstå de underliggande problemen. Detta kan utföras med hjälp av maskininlärningsmodeller som snabbt kan identifiera och kategorisera bin, vilket till och med kan göras på relativt enkel hårdvara.

1.1 Mål

Målet är att träna en modell som kan identifiera och skilja mellan arbetar- och drönarbin från en inmatad bild. Ett delmål är att kunna använda modellen, eller en mindre variant av den, på mikrodatorn Raspberry Pi.

1.2 Syfte

Ett bis kön dikterar dess arbetsuppgifter. Möjligheten att effektivt kunna se skillnad mellan könen kan underlätta vid både underhåll och odling av bin. Det kan även bidra i undersökningar för att kunna besvara en rad frågor såsom: Hur ser könsfördelningen ut i kupan? Löper arbetar- eller drönarbin större risk för att drabbas av parasiter? Är något av könen mer känsligt för bekämpningsmedel?

2 Bakgrund

Honungsbin är en av de viktigaste pollinatörerna för kommersiella grödor och har därför en kritisk roll för matproduktionen.¹ Det här är en av fördelarna med bins överlevnad ur en antropocentrisk synvinkel. Enligt Hedenus et al. karaktäriseras denna av att värde enbart tilldelas människor medan allt annat värderas utifrån deras instrumentella värde, alltså i det här fallet vilken nytta det har för människan.² Biocentrismen beskrivs som lik antropocentrismen, men utökat så att även andra djur och till och med växter anses ha ett eget finalt värde.³ Om bin betraktas ur den här synvinkeln kan pliktetiken, så som den beskrivs av Hansson⁴, diktera att vi människor är skyldiga att bevara dem och inte endast för mänsklig fördel.

Under de senaste åren har det dykt upp artiklar i nyheter och sociala medier om bidöd, ett fenomen där arbetsbin, vilka utgör majoriteten av bina i kupan, försvinner och lämnar drottningen och larverna. Experter har försökt fastställa orsaken till detta men resultaten av dessa studier förblir blandade med många möjliga orsaker: livsmiljöförstörelse, klimatförändringar, användning av bekämpningsmedel, monokultur, patogener, parasiter, och skadedjur.^{5,6} Maskininläring kan hjälpa forskare att bättre förstå vikten av de olika faktorerna, och på så sätt assistera biodlare i att hålla bikupor friska.

Drönarbin, dvs. hanar, kännetecknas av större, rundare kroppar med en mörk bakdel och stora ihopsittande ögon. Arbetare, eller honor, är till formen mer lika getingar. Enbart arbetare kan bära pollen vilket är tacksamt eftersom det därför även går att avgöra könet på suddiga bin om de syns bära pollen. I Figur 1 nedan ges exempel på bin av de två könen. Alla bilder innehåller Buckfastbin.



Figur 1: Drönare/hanar (vänster) och arbetare/honor (höger)

3 Metod

Det fanns flera tänkbara tillvägagångssätt för att nå det slutgiltiga målet. Tidigare kunskap gjorde att det inom kort stod mellan ett fåtal alternativ och arbetet kunde inledas. Det första steget i arbetet var att dela ut bilderna mellan medlemmarna, där datan bestod sammanlagt av ungefär 3500 bilder från ett antal bikupor. Målet var att markera alla bin i varje bild och, där det gick, även avgöra kön. Dessa bilder tillsammans med deras etiketter användes sedan för att träna en neuronätsmodell, vars uppgift var att placera ut markeringar med tillhörande kategori kring bin i en inmatad bild. Till sist utvärderades den färdigtränade modellen för att identifiera eventuella brister. Försök att åtgärda dessa gjordes och träning upprepades. Samma procedur applicerades för en mindre modell mer anpassad för långsammare och mer energisnål hårdvara.

3.1 Etikettering

Den tilldelade datan *etiketterades* manuellt med hjälp av programmet LabelImg. Detta utfördes i mindre delgrupper om två medlemmar som parallellt markerade alla bin på bilderna och placerade etiketter efter kön; totalt etiketterades 1252 bilder. Ett tidigt problem som uppstod var att det ofta var svårt att med säkerhet identifiera somliga bins kön när det var dålig skärpa eller ljus. Därför infördes en tredje kategori *okänd/bi*, utöver *arbetare* och *drönare*, i etiketteringen.

Bilder innehållande drönare prioriterades eftersom de var många gånger färre än antalet bilder enbart innehållande arbetare. Det ansågs först givet att försöka balansera uppsättningen bilder innehållande arbetare och drönare i den mån det gick för att ge modellerna en sportslig chans att lära sig se skillnad mellan dem. Detta innebar att en del bilder av arbetare exkluderades. Vidare behövde bilderna från en viss kupa uteslutas eftersom datan ansågs vara för lågkvalitativ för etikettering. Tillsammans ledde detta till en markant minskning i antalet bilder att träna på. För att maximera mängden träningsdata tillades därför fler bilder från de andra kuporna och träning utfördes på alla bilder utan någon balansering.

3.2 Träning

Merparten av arbetet med att träna modeller utfördes i molntjänsten *Google Colab*. Först antogs att Tensorflow skulle vara ett lämpligt verktyg som modeller skulle kunna tas fram med hjälp av. Efter tekniska hinder med Tensorflow valdes istället att skapa modellerna med hjälp av kodbiblioteket Darknet. Biblioteket har ett medföljande konsolprogram med funktioner för träning och användning av tränade nätverk. Dessutom medföljer konfigurationsfiler för vanliga arkitekturer på modeller, vilka kan anpassas efter egna ändamål. Utöver Darknet valdes arkitekturen *YOLOv3*, det vill säga version tre av *You Only Look Once*. Denna är en typ av neuronät som skiljer sig från tidigare metoder genom att den inmatade bilden endast behandlas av nätverket en gång, därav namnet. Dess utdata är en matris av grupper av rektangulära markeringar, alla med tillhörande sannolikhetsfördelning över de möjliga klasserna på objekt, och ett värde representerande modellens säkerhet att markeringen faktiskt täcker ett objekt. Den mindre modellen använde en snarlik arkitektur med färre lager, som var anpassad för hastighet. Syftet var att ha en modell med hög träffsäkerhet, och en annan med kortare slutledningstid på bekostnad av träffsäkerhet.

Innan träning inleddes utfördes två förberedande steg: först *augmenterades* alla bilder för att öka mängden *samples* i syfte att underlätta för modellen att generalisera. Augmentering innebar att bilderna vändes horisontellt och/eller vertikalt för att omvandla en bild till fyra. Sedan valdes 80 % av alla samples ut slumpmässigt för att träna på. Resterande 20 % användes för utvärdering av tränade modeller. En sådan uppdelning är viktig för att kunna avgöra modellens verkliga prestanda. Träningsprocessen avslutades då modellens träffsäkerhet på utvärderingsdatan slutade stiga. Träning upprepades ett flertal gånger med olika mängder data, med och utan augmentering, med olika etikettering på existerande bilder, olika fördelningar mellan tränings- och utvärderingsdata, olika arkitekturer på modellen, osv.

Den större modellen kördes även på en Raspberry Pi 4 för att fastställa att den var applicerbar i verkligheten. En Raspberry Pi är billig och lätt att använda, och lämpar sig därför för biodlare världen över att själva nyttja modellen. Den mindre modellens prestanda kunde härledas från resultatet av att köra den större modellen, tillsammans med de två tiderna från att köra de båda i Colab.

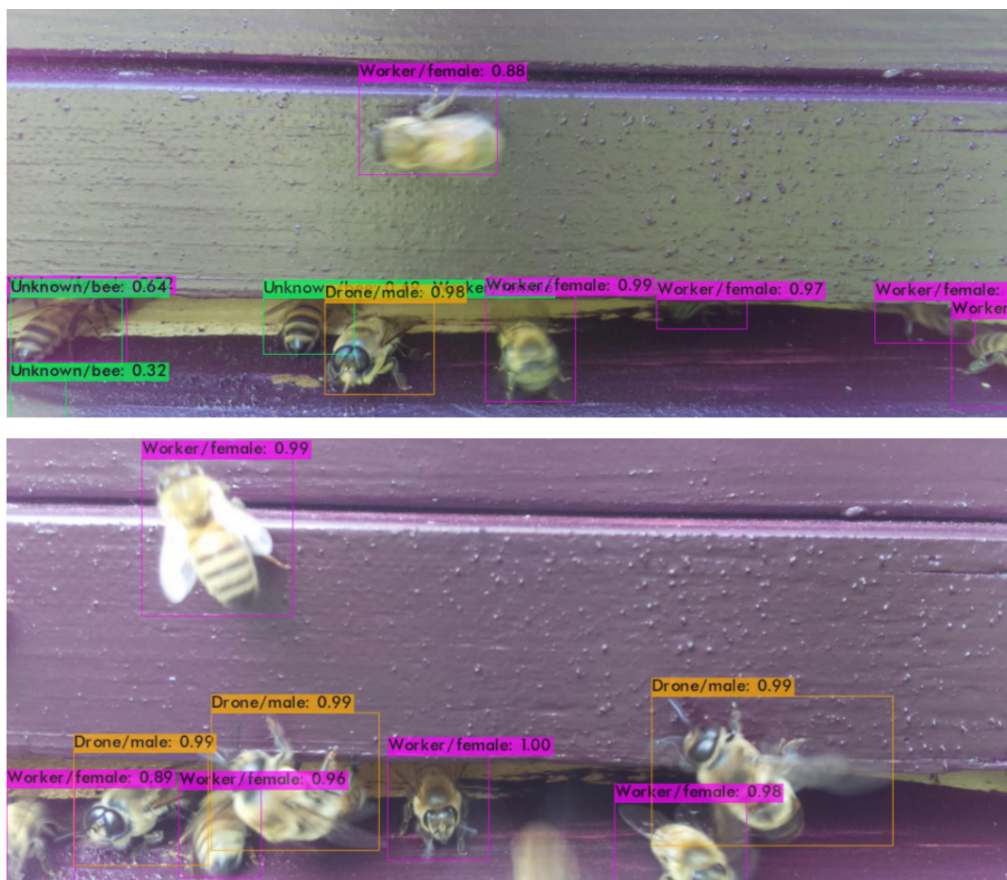
3.3 Utvärdering av tränade modeller

För testning användes den andel data som valts ut för utvärdering i föregående steg. Från denna data valdes några bilder ut slumpmässigt och matades in i modellen. Resultatet kunde sedan utvärderas manuellt, dvs. genom inspektion, eller genom att jämföra med existerande etiketter. Utöver detta beräknades ett slutgiltigt mAP-värde som tog hänsyn till all utvärderingsdata. mAP står för *mean average precision* (ungefär *genomsnittlig precision*) och motsvarar en träffsäkerhet på en procentuell skala. Beräkning av ett sådant värde just på utvärderingsdatan är ett bra mått på hur väl en modell kommer prestera i verkligheten eftersom den även där kommer exponeras för helt nya bilder.

4 Resultat

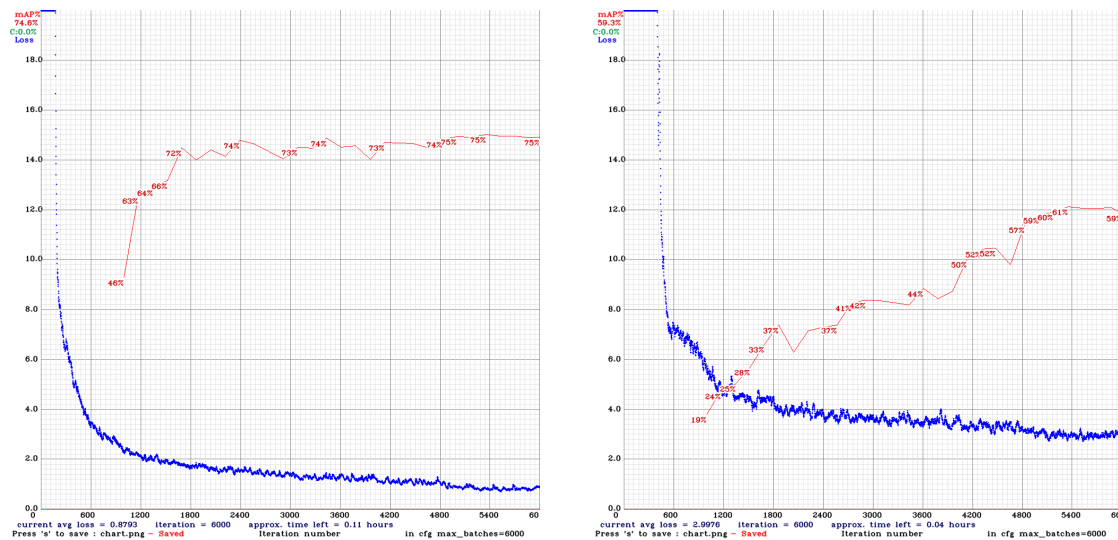
Den primära modellens slutgiltiga mAP-värde var omkring 74.6 %, vilket indikerar att den kommer att etikettera korrekt större delen av tiden. Den mindre modellen fick träna i fler omgångar på grund av dess låga precision, men tränades ändå inte lika länge i tid som den stora modellen eftersom den mindre storleken gjorde träningsprocessen snabbare. Den mindre modellens slutgiltiga mAP-värde var 65.9 %. På Colabs GPU tar det ungefär 18 ms att evaluera en bild med den primära modellen, men under 3 ms med den mindre. Den stora modellen tar dock nästan en minut för en enda slutledning på vår Raspberry Pi 4. Den mindre modellen skulle kräva färre än 10 sekunder förutsatt samma hastighetsökning som på Colabs GPU.

Nedan i Figur 2 visas resultatet av modellen från två slumpvalda bilder från utvärderingsdatan. I den övre bilden dubbelmarkerar modellen ett arbetsbi som både arbetare och okänd, men lyckas identifiera suddiga och skymda arbetsbin. I den nedre markerar den korrekt ut de fem arbetsbina och de tre drönarbina, men missar ett suddigt bi. Markeringarna omger till större del hela bin, men missar ibland lite av vingarna eller benen.



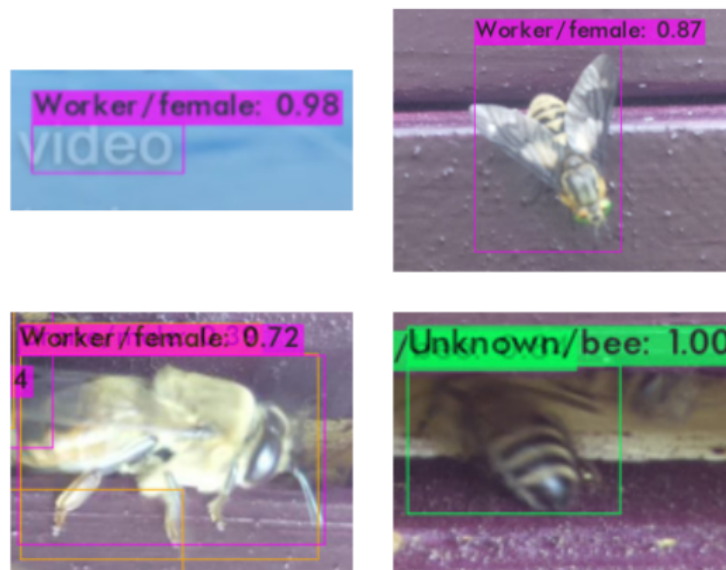
Figur 2: Slumpvalda bilder från utvärderingsdatan med etiketter från den primära modellen. Rosa markering: arbetare, orange markering: drönare, grön markering: okänd

Figur 3 visar två grafer över modellernas prestanda under träning. De röda kurvorna representerar träffsäkerheten i procent på utvärderingsdatan under träningsgången, och de blå visar *loss* på träningsdatan. Det bör poängteras att figuren för den mindre modellen endast visar den första av två omgångar träning.



Figur 3: Grafer från träning av den primära modellen (vänster) och den mindre modellen (höger)

Modellerna placerar ut "okänd/bi" på både suddiga bin och bin som inte helt var i bild. Trots det höga mAP-värdet betecknade den stora modellen en fluga som ett arbetsbi och har även problem med vattenstämpeltexten som den markerar fel. Detta visas i Figur 4 nedan.



Figur 4: Felmarkeringar av den stora modellen på utvärderingsdatan

5 Slutsatser

Den framtagna modellen har uppnått målet att från en inmatad bild markera ut och skilja mellan arbetsbin och drönarbin. Om den inte med större säkerhet kan avgöra könet på ett bi, markeras det som ”okänd/bi”. Om den dessutom inte är säker nog att nå tröskeln för någon av de tre klasserna så låter den bli att etikettera. För majoriteten av de testade bilderna lyckades modellen markera bin korrekt, men som påpekat i Kapitel 4 förekommer det att modellen gissar fel. Felen som visas där beror delvis på att bilderna innehåller objekt som modellen aldrig bearbetat under träningsförloppet, delvis på att den tredje kategorin ofta valdes vid osäkerhet under den manuella etiketteringen. Eftersom modellen grundar slutledningar på manuellt etiketterad träningsdata så finns alltid risken att den skulle kunna följa felaktig etikettering. Om forskning blint förlitar sig på modellens utdata hade en felaktig slutledning kunnat stå till grund för en falsk slutsats. Detta gäller även vid hög rapporterad träffsäkerhet, vilket kan vara vilseledande, eftersom denna enbart grundas i modellens förmåga att efterlikna de manuella etiketterna.

Vi uppskattar att vår användning av Google Colabs GPU:er sammanlagt har motsvarat 4,4 kg koldioxidutsläpp.⁷ Enligt Google själva är alla Googles tjänster, Colab inkluderat, koldioxidneutrala då de klimatkompenserar.⁸ Följer man definitionen av ekologisk hållbarhet av Hedenus et al.⁹, så kan träningen då anses vara ekologiskt hållbar. Det hade givetvis varit bättre om utsläppen var noll, vilket är något Google påstår sig sträva efter.

Träning av neuronnät använder mycket energi, men vi sparade både tid och datorkraft genom att utgå från redan tränade lager. Vidare behöver man bara utföra träningen en gång och eventuella utsläpp från användningen av modellen baseras på energiförbrukningen av hårdvaran som kör den. Produkten kan betraktas vara ekonomiskt hållbar förutsatt en låg och framför allt effektiv energiförbrukning av Raspberry Pi-datorn.

Med avseende på rättvis fördelning av resurser skulle osämja kunna uppstå mellan de som har tillgång till en sådan här modell gentemot de som inte har det. Men eftersom hårdvaran är billig och mjukvaran kan göras öppen anser vi tröskeln vara låg att som biodlare själv ta del av teknologin. Produkten kan därför anses vara socialt hållbar.

Trots att modellen kan brukas för att hjälpa bin är den inte i sig en komplett lösning. Enligt Hedenus et al. faller lösningar till hållbarhetsproblem in på ett spektrum från tekniska lösningar till livsstilsförändringar¹⁰ där denna modell faller närmare en teknisk lösning och inte i sig kommer rädda bina. Det behövs fortfarande göras viktiga ändringar inom biodlingsmetoder och jordbrukspraxis, särskilt när det kommer till användningen av bekämpningsmedel och monokulturer. Tvång på mer restriktioner inom jordbruk kan dock slå hårdare mot fattigare bönder med mindre odlingar då exempelvis en parasitinfestation skulle kunna förstöra en hel skörd om den inte skyddas av bekämpningsmedel.

Modellen kan dock visa sig vara negativ för andra arter än de bin den appliceras på. Om produkten realiserar kan den stå till grund för forskning som leder till metoder som hjälper öka populationen av domesticerade bin. Det skulle kunna öka konkurrensen med andra pollinatörer, såsom vilda bin, om de begränsade resurserna i närområdet. En sådan konkurrens har visat sig vara destruktiv för de andra arterna¹¹, och belyser därför en nackdel med vår modell. Bör vi då prioritera vår matproduktion, eller överlevnaden av andra arter än honungsbin? Ska de andra arternas överlevnad ges ett finalt värde? Ur en antropocentrisk synvinkel behöver vi kanske inte prioritera de andra arterna, om de inte bidrar med något till oss människor som inte kan utföras av honungsbin. Men från biocentrismen bör allt levande tilldelas finalt värde³, och därmed skulle det vara negativt att blint öka mängden honungsbin.

5.1 Förbättringar

Till att börja med hade träningsdatan kunnat behandlas mer och med fler metoder. Skulle modellen användas för forskning eller i kommersiellt syfte behöver den ytterligare diversifierad träningsdata för att undvika feletikettering och kunna appliceras på data olik den använd under träning. Man hade kunnat blanda in bilder innehållande annat än bin för att modellen skulle bli mer robust och orsaka mindre feletikettering med okända föremål. Som exempel hade man kunnat inkludera bilder innehållandeflugor och andra insekter som liknar bin i utseende. Det hade sannolikt löst problemet med flugan i Figur 4. På liknande vis kan träningsdata med varierande upplösning och skärpa förbättra modellens möjlighet att prestera väl på bilder från olika källor. I ett annat fall gissade modellen att text över en arkivbild tagen från internet var ett arbetsbi, antagligen då träningsdatan saknade bilder som innehöll text helt och hållet. En förbättring hade då varit att träna modellen på data innehållande mer varierande innehåll.

Det uppstår ett dilemma ur att etikettering är lättare att utföra på skarpa och högupplösta bilder, men där man hellre tränar på bilder av varierande kvalitet för att göra modellen mer robust. En möjlig lösning på detta hade varit att etikettera bilder av hög kvalitet, för att sedan artificiellt försämra dem genom att lägga på oskarpa, smuts, distortion och liknande genom augmentering.

Den mindre modellen är inte lika träffsäker som den större. Den hade kunnat tränas längre för en potentiell förbättring, men grafen i Figur 3 verkar tyda på att det inte hade lett till någon större förbättring då träffsäkerheten börjar plana ut.

Om projektet gjordes om på nytt skulle en senare version av YOLO-algoritmen användas, det vill säga version fyra eller fem. Valet av version tre gjordes som hastigast innan det var beslutat att Darknet skulle användas och innan vi visste att det överhuvudtaget skulle fungera. En uppgradering skulle sannolikt leda till både högre träffsäkerhet och kortare slutledningstid.¹²

5.2 Lärdomar

Att arbeta kring en fråga utan att bli tilldelade konkreta lösningssteg ledde till en mer experimentell arbetsmetod. Flera tillvägagångssätt övervägdes och provades innan den slutgiltiga lösningen var utformad. Kollektivt hade vi tillräckligt med erfarenhet för att angripa problemet och arbetet flöt på ganska bra. Dessutom har Discord och planeringsverktyget Notion varit en gagn för kommunikationen i gruppen som fungerade väl. Fördelningen av arbetsuppgifterna gjordes i två dimensioner: vissa uppgifter, såsom etikettering, fördelades jämnt bland medlemmarna, medan andra, såsom testning på Raspberry Pi, tilldelades individuella gruppmedlemmar.

En av de största utmaningarna som stöttes på var etiketteringen; dels var det en mycket tidskrävande uppgift, dels var det inte alltid möjligt att säkerställa könet på alla bin. Vi stod inför avvägningar såsom att utesluta bilder av lägre kvalitet eller behålla dem för att modellen inte ska ha för få bilder att träna på. En värdefull insikt var då vikten av att augmentera träningsdatan, där vi kunde väga upp förlusten genom att fyrdubbla de kvarvarande bilderna i träningsdatan.

Något som tydliggjorts under projektet är vilken stor roll bin spelar för ekosystemet och att det finns flera synvinklar att resonera ifrån för att motivera deras stora betydelse. Ansvar kan anses ligga på oss ingenjörer, som bärare av teknisk kunskap, att med försiktighet arbeta för samhällets, miljöns och mänskighetens gynning,¹³ där även bin är inkluderade.

Referenser

- [1] U.S. Food and Drug Administration. "Helping Agriculture's Helpful Honey Bees." Tillgänglig: <https://www.fda.gov/animal-veterinary/animal-health-literacy/helping-agricultures-helpful-honey-bees>. (2018). (Hämtad: 2022-01-07).
- [2] F. Hedenus, M. Persson och F. Sprei, *Hållbar utveckling - nyanser och tolkningar*. Studentlitteratur, 2018, s. 70–71, 81–82, (Hämtad: 2021-12-21).
- [3] F. Hedenus, M. Persson och F. Sprei, *Hållbar utveckling - nyanser och tolkningar*. Studentlitteratur, 2018, s. 72–74, (Hämtad: 2021-12-31).
- [4] S. O. Hansson, *Teknik och etik*. Kungliga Tekniska Högskolan, Stockholm, 2009, s. 44, (Hämtad: 2021-12-05).
- [5] U.S. Environmental Protection Agency. "Colony Collapse Disorder." Tillgänglig: <https://www.epa.gov/pollinator-protection/colony-collapse-disorder>. (2021). (Hämtad: 2021-12-15).
- [6] M. Donaldson-Matasci. "Honeybees and Monoculture: Nothing to Dance About." Tillgänglig: <https://blogs.scientificamerican.com/guest-blog/honey-bees-and-monoculture-nothing-to-dance-about/>. (2013). (Hämtad: 2022-01-07).
- [7] V. Schmidt, A. Luccioni, A. Lacoste och T. Dandres. "ML CO2 Impact." Tillgänglig: <https://mlco2.github.io/impact/#co2eq>. (2021). (Hämtad: 2022-01-07).
- [8] Google LLC. "Cloud sustainability." Tillgänglig: <https://cloud.google.com/sustainability/>. (2021). (Hämtad: 2022-01-06).
- [9] F. Hedenus, M. Persson och F. Sprei, *Hållbar utveckling - nyanser och tolkningar*. Studentlitteratur, 2018, s. 32–38, (Hämtad: 2021-12-21).
- [10] F. Hedenus, M. Persson och F. Sprei, *Hållbar utveckling - nyanser och tolkningar*. Studentlitteratur, 2018, s. 85–86, (Hämtad: 2022-01-07).
- [11] J. Oldham. "Will Putting Honey Bees on Public Lands Threaten Native Bees?" Tillgänglig: <https://e360.yale.edu/features/will-putting-honey-bees-on-public-lands-threaten-native-bees>. (2020). (Hämtad: 2022-01-07).
- [12] A. Bochkovskiy, C.-Y. Wang och H.-Y. M. Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection." Tillgänglig: <https://arxiv.org/pdf/2004.10934v1.pdf>. (2020). (Hämtad: 2022-01-08).
- [13] F. Hedenus, M. Persson och F. Sprei, *Hållbar utveckling - nyanser och tolkningar*. Studentlitteratur, 2018, s. 127, (Hämtad: 2022-01-07).