

VISUAL SEARCH

Et system til optælling af fugle og pattedyr på dronebilleder

Teknisk rapport fra DCE - Nationalt Center for Miljø og Energi

nr. 244

2022



AARHUS UNIVERSITET DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

VISUAL SEARCH

Et system til optælling af fugle og pattedyr på dronebilleder

Teknisk rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 244

2022

Thomas Eske Holm¹ Kristian Tølbøl Sørensen² Claus Lunde Pedersen¹ Jack Ord Rasmussen² Søren Rasmussen² Jakob Langdal²

¹Aarhus Universitet, Institut for Ecoscience ²Alexandra Instituttet



Datablad

Serietitel og nummer:	Teknisk rapport fra DCE - Nationalt Center for Miljø og Energi nr. 244
Kategori:	Rådgivningsrapporter
Titel: Undertitel:	Visual Search Et system til optælling af fugle og pattedyr på dronebilleder
Forfattere:	Thomas Eske Holm ¹ , Kristian Tølbøl Sørensen ² , Claus Lunde Pedersen ¹ , Jack Ord Rasmussen ² , Søren Rasmussen ² & Jakob Langdal ²
Institutioner:	¹ Aarhus Universitet, Institut for Ecoscience, ² Alexandra Instituttet
Udgiver: URL:	Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi © http://dce.au.dk
Udgivelsesår: Redaktion afsluttet:	Juni 2022 Maj 2022
Faglig kommentering: Kvalitetssikring, DCE: Finansiel støtte:	Thomas Bregnballe Jesper Fredshavn Aage V. Jensens Fonde
Bedes citeret:	Holm, T.E., Sørensen, K.T., Pedersen, C.L., Rasmussen, J.O., Rasmussen, S. & Langdal, J. 2022. Visual Search. Et system til optælling af fugle og pattedyr på dronebilleder. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 44 s Teknisk rapport nr. 244 <u>http://dce2.au.dk/pub/TR244.pdf</u>
Sammenfatning:	Aarhus Universitet har siden 2017 udviklet metoder til droneovervågning af ynglende vandfugle og hjortevildt. De nye metoder har vist sig at være tidsbesparende i felten, at give bedre data og forstyrre mindre end traditionelle optællingsmetoder. Den efterfølgende optælling fra dronebilleder har til gengæld vist sig at være tidskrævende, da dette gøres manuelt. For at afhjælpe dette, har vi i projektet "Visual Search" udviklet en prototype på et system, der ved hjælp af deep learning kan lokalisere og optælle dyr på billeder. Formålet har været at erstatte den nuværende manuelle og tidskrævende optælling på billeder og dermed nedsætte tidsforbruget med op til 90%. En sådan besparelse kan være med til at gøre droneovervågning billigere og dermed betydeligt mere udbredt end i dag. Denne tekniske rapport beskriver udviklingen og brugen af Visual Search-prototypen.
Emneord:	Deep learning, Droner, Vandfugle, Hjorte, Monitering
Layout: Foto forside:	Karin Balle Madsen Thomas Eske Holm
ISBN: ISSN (elektronisk):	978-87-7156-690-1 2244-999X
Sideantal:	44
Internetversion:	Rapporten er tilgængelig i elektronisk format (pdf) som http://dce2.au.dk/pub/TR244.pdf

Indhold

For	rord	5
Sar	mmenfatning	6
1	Indledning	7
2	Metode og udviklingsplan	11
3	Vejledning til Visual search	13
4	Analyse af resultater	19
5	Efterprocessering	29
6	Valg af tærskelværdier og filtrering af areal	33
7	Optælling af hjorte på termografiske billeder	39
8	Konklusion	42
9	Referencer	44

Forord

Aarhus Universitet (AU) har i dette projekt samarbejdet med Alexandra Instituttet (AI) og med støtte fra Aage V. Jensens Fonde udviklet Visual Search, der ved hjælp af deep learning kan finde og optælle dyr på dronebilleder.

AU har i projektet bidraget med annoterede datasæt med vandfugle og hjorte, test af Visual Search og har været tovholder på projektet og rapporteringen. AU har specifikt været hovedforfatter på rapportens sammenfatning, forord og indledning.

AI har udviklet og kodet Visual Search inkl. træningsalgoritmer, brugergrænseflade og serveropsætning. AI har i rapporten være ansvarlig for at beskrive den tekniske del af projektet, vejledningen i brug af Visual Search, analyse, efterprocessering og konklusion.

Vi ønsker at takke Aage V. Jensens Fonde for den økonomiske støtte til Visual Search.

Sammenfatning

Aarhus Universitet har siden 2017 udviklet metoder til droneovervågning af ynglende vandfugle og hjortevildt. De nye metoder har vist sig at være tidsbesparende i felten, at give bedre data og forstyrre mindre end traditionelle optællingsmetoder.

Den efterfølgende optælling fra dronebilleder har til gengæld vist sig at være tidskrævende, da dette gøres manuelt. For at afhjælpe dette, har vi i projektet "Visual Search" udviklet en prototype på et system, der ved hjælp af deep learning kan lokalisere og optælle dyr på billeder.

Formålet med projektet har været at erstatte den nuværende manuelle og tidskrævende optælling på billeder og dermed nedsætte tidsforbruget med op til 90%. En sådan besparelse kan være med til at gøre droneovervågning billigere og dermed betydeligt mere udbredt end i dag. Denne tekniske rapport beskriver udviklingen og brugen af Visual Search-prototypen.

1 Indledning

Traditionel overvågning af vandfugle

Overvågning af ynglende kolonirugende vandfugle udføres traditionelt med observatører i felten, som tæller fuglene på afstand ved brug af kikkert eller teleskop, og/eller går ind i ynglekolonierne for at tælle reder. Begge metoder kan være upræcise, afhængigt af koloniens størrelse og oversigtsforholdene fra observationspunkterne.

Ved optælling af kolonier på afstand gennem teleskop vil man oftest ikke kunne se alle fuglene. Et forsøg med optælling af hættemåger på øer i Egå Engsø fra land viste eksempelvis, at der i nogle tilfælde kun registreres omkring 45-60 % af de fugle, som egentlig er til stede, simpelthen fordi man ikke kan se alle fugle i kolonien. Det samme gjorde sig gældende ved Sneum Digesø, hvor hættemågekolonierne normalt udelukkende optælles fra et dige med teleskop (Holm & Bregnballe 2019).

Ved optælling inde i større kolonier, hvor fuglene ligger tæt, eksempelvis terne- eller hættemågekolonier, kan der være risiko for at optælleren har vanskeligt ved at skelne talte reder fra ikke talte reder eller overser reder. Et studie fra Langli i Vadehavet viste, at erfarne optællere i visse tilfælde kun fik registreret 1/3 af de måger, der faktisk var til stede i kolonierne på øen (Holm & Bregnballe 2019).

Når man som fugletæller bevæger sig helt ind i en koloni, går fuglene på vingerne (Figur 1.1, Holm & Bregnballe 2019). Selv om en sådan forstyrrelse oftest er relativt kortvarig, og måske kun svarer til en rovfugls passage af en koloni, kan forstyrrelsen føre til prædation af æg og unger, mens forældrefuglene ikke er på reden. Der er ligeledes risiko for, at optælleren træder i reder, der er mere eller mindre skjult i vegetationen.



Figur 1.1. En fugletæller i en splitternekoloni forstyrrer så meget, at fuglene går på vingerne. Derved øges risikoen for at rederne præderes af måger eller krager. Ved en overvågning med drone er forstyrrelsen langt mindre og fuglene bliver liggende på deres reder (se indsat foto i venstre hjørne).

Traditionel overvågning af hjortevildt

Som ved fuglene er kendskab til bestandsstørrelsen et vigtigt redskab, når større pattedyr som kron- og dåvildt skal forvaltes. Når antallet af dyr skal opgøres, tæller man ofte forårsbestanden. Typisk registrerer man antallet af individer efter jagtsæsonen, men før kalvesætning. De hidtidige bestandsopgørelser af kron- og dåvildt har været baseret på estimater foretaget på baggrund af lokale observatører og regionale indberetninger på baggrund af jagtudbyttestatistikkerne.

Fordi krondyr og dådyr bevæger sig over relativt store områder (home range), kræver en bestandsopgørelse, at man kan tælle dyrene over store arealer inden for kort tid. Den mest udbredte videnskabelige metode til bestandsopgørelse er optælling af fald (ekskrementer). Denne metode er dog primært velegnet til mindre områder og svær at benytte i stor skala. Til optælling af hjortevildt, har den bedste metode været at opgøre bestanden ved visuelt at iagttage og tælle individerne. Under forudsætning af, at man ikke tæller nogle af individerne to eller flere gange, har dette givet et minimumsestimat for bestanden, da der ikke kan være færre dyr, end der er set.

Det anslås imidlertid, at man ved en visuel tælling kun ser 60-80 % af den samlede kron- eller dåvildtbestand i et område. Især dyr der står i skove eller bevoksninger kan være svære at registrere. Såfremt man fra tælling til tælling registrerer en konstant andel af de tilstedeværende dyr, kan den visuelle tælling over tid tegne en tendens for bestandsudviklingen. Men tællingerne vil sjældent give et retvisende billede af en bestands samlede størrelse. Gode bestandsestimater er en forudsætning for en mere effektiv forvaltning, herunder en forebyggelse af fx hjortevildtets mark- og skovskader, men historisk har disse opgørelser altså været præget af en høj grad af usikkerhed.



Figur 1.2. Udsnit af øen Brasenholm ved Filsø hvor der yngler 3238 par hættemåge (Holm & Bregnballe 2020).

Overvågning af vandfugle og hjortevildt med droner

De seneste års teknologiske fremskridt har gjort det muligt med drone at affotografere selv større områder fra en højde, der ikke forstyrrer de overvågede arter (Figur 1.2, Holm m.fl. 2018, Holm m.fl. 2019, Holm m.fl. 2020). Efterfølgende kan der foretages artsbestemmelse og optælling fra billederne, hvilket ofte giver langt mere pålidelige resultater end ved de traditionelle metoder til optælling i felten. Gennemgang af fotos er i dag en manuel proces, og den er ofte tidskrævende og dermed omkostningsfuld (Holm & Bregnballe 2019). Ud over almindeligt kamera kan man benytte et termisk kamera som kan spore varmeforskelle, og herved benytte droneteknikken til at registrere større pattedyr i skovområder, hvor dyrene ikke kan spores med et almindeligt kamera og hvor dyrene heller ikke opdages ved en traditionel visuel optælling (Figur 1.3, Holm m.fl. 2020). Optælling og artsbestemmelse fra termiske billeder er ligeledes en manuel proces, der er både tidskrævende og omkostningsfuld.

Selve feltarbejdet ved brug af droner er dog oftest tidsbesparende i forhold til traditionel overvågning, dels fordi det tager kortere tid at overflyve et område frem for at gå det igennem, og dels fordi man kan flyve ud til eksempelvis yngleøer i stedet for at sejle. Flaskehalsen er den manuelle optælling og artsbestemmelse på billedmaterialet, der er meget tidskrævende. Eksempelvis foretog Aarhus Universitet en droneoptælling af yngleøerne i Filsø, Birkesø og ved Gyldensteen Strand i 2020, hvor størstedelen af de brugte timer gik til manuel optælling af billederne, og hvor feltarbejdet var den del der tog mindst tid (Holm & Bregnballe 2020). Et andet eksempel er optælling og artsbestemmelse af ynglefuglene på dronefotos fra Langli i Vadehavet som tog 75 timer (Holm & Bregnballe 2019), hvilket må vurderes at være for meget i overvågningsprogrammer, hvor økonomi er en begrænsende faktor og derfor vejer tungere end muligheden for bedre data og færre forstyrrelser.



Figur 1.3. Dronefotografering af krondyr med termisk og almindeligt kamera. Med det termiske kamera kan man registrere dyr der på de almindelige billeder er skjult under træerne.

Udvikling af en automatisk optælling - Visual Search

En automatisering af optællingsproceduren vil gøre optællingen af billeder langt hurtigere. En sådan automatisering udvikler man med Computer Vision og Deep Learning, hvor der er sket enorme landvindinger de seneste år, og som Alexandra Instituttet har omfattende erfaring med. Alexandra Instituttet er et af Danmarks Godkendt Teknologisk Service (GTS-) institutter. Automatiseret billedanalyse har bl.a. været anvendt til at sammensætte luftfotos over hele Danmark for Styrelsen for Dataforsyning og Effektivisering, og Deep Learning har været brugt til automatisk lokalisering af markgrænser og træformationer for Landbrugsstyrelsen. Flere delelementer i den foreslåede løsning er desuden blevet udviklet til Proof of Concept stadiet, i samarbejde med Det Kongelige Bibliotek, hvor teknikken har været brugt til at genkende fotograf-stempler, mhp. bestemmelse af ophavsret. Forarbejdet med at afklare, om teknologien egner sig til at automatisk genkendelse af visuelt lignende objekter (såsom fugle), og træning af neurale netværk til genkendelse af objekterne i nye datasæt, er dermed allerede foretaget. Det næste naturlige skridt har derfor været at forbedre systemet og lave en første protype hen imod et brugervenligt system, der kan anvendes i den fremtidige droneovervågning af fugle og pattedyr, til gavn for forvaltningen af Danmarks natur.

2 Metode og udviklingsplan

Konkret er der i projektet udviklet en brugergrænseflade, hvor brugeren kan uploade dronebilleder, hvorefter systemet foretager en fuldautomatisk præprocessering og analyse af billedmaterialet. Dernæst kan brugeren klikke på de objekter der skal tælles, og angive mærkater. Et mærkat kan eksempelvis være et artsnavn, et køn, en adfærd eller andet. Hvert objekt kan have flere mærkater, så man fx kan optælle sølvmåger der ligger ned og har synligt redemateriale omkring sig, altså mærkater med artsnavn, adfærd og omgivelser. Systemet træner dernæst automatisk et dybt neuralt netværk til at kunne genkende de relevante objekter. Efterfølgende vil det neurale netværk kunne bruges direkte på tilsvarende billeder, hvor der så ikke kræves manuel udvælgelse og angivelse af mærkater, fordi systemet nu er trænet til at kunne genkende objekterne.

Den nødvendige kodebase til Visual Search er udviklet af Alexandra Instituttet i tæt samarbejde med Aarhus Universitet, Institut for Ecoscience (AU), som har stillet datasæt til rådighed til udvikling og test af systemet. Som eksempler på datasæt kan nævnes 1) et sammensat drone-billede af øen Langli ud for Esbjerg, hvor der er angivet art, positur og andre parametre på 22.000 ynglefugle, 2) billede af rastende troldænder ved Eskildsø, 3) billede af ynglende hættemåger ved Egå Engsø og 4) termografiske dronebilleder af krondyr og dådyr fra bl.a. Oustrup Hede. Flere af datasættene er beskrevet i Holm & Bregnballe (2019) og Holm m.fl. (2020).

Udviklingsplanen har været formuleret i faser, som er gået fra eksperimentering, afklaring og Proof of Concept, og gradvist gået mod automatiserede løsninger, som AU og andre slutbruger vil kunne anvende.

Fase 1: Minimum viable product

Målet i denne fase har været at afklare, hvilke Deep Learning-teknikker der har egnet sig bedst til at løse problemet, og eksperimentere med de neurale netværk, så den bedst mulige genkendelsesprocent har kunne opnås. I denne fase foregik udviklingen på lokale arbejdsstationer og servere. Herefter er de nødvendige brugergrænseflader opbygget, for at slutbrugerne på egen hånd kan anvende systemet.

Det er i denne fase udviklet en metode til automatisk opdeling af billeder og beregning af visuelt indhold. Det er ikke usædvanligt, at sammensatte dronebilleder fylder over 10 gigapixels (100.000 x 100.000 pixels), så de er nødt til først at opdeles i mindre felter, før de kan behandles. Billedet opdeles i overlappende felter af varierende størrelse, sådan at hvert område er afbilledet i adskillige udsnit. Via et neuralt netværk beregnes en slags "fingeraftryk" for hvert udsnit, som siger noget om dets indhold, og gør det muligt at finde billeder der minder om hinanden, selvom de pixel-for-pixel kan have meget lidt til fælles.

Der er herefter udviklet en metode til automatisk træning af et neuralt netværk til datakompression for effektiv lagring og søgning. For et typisk dronebillede genereres fx ca. 8 millioner udklip, som hvert får et unikt "fingeraftryk", men lagringen og efterfølgende søgning i sådanne "fingeraftryk" bliver meget tung. Der trænes derfor et neuralt netværk, der kan lære at komprimere fingeraftrykkene til ca. 1 % af deres oprindelige datamængde, med minimalt informations-tab.

Efterfølgende er der udviklet en metode til interaktiv udvælgelse af søgeområde, samt søgning i millioner af billedudklip på under 1 sekund. Der er udviklet en enkel måde at udvælge, hvilke objekter der skal søges efter. Med et for-beregnet komprimeret fingeraftryk for hvert udsnit, kan de udvalgte områder sammenlignes, og hvis man fx markerer en sølvmåge, kan man i eksemplet forvente, at systemet finder tusindvis af andre steder, hvor der ses sølvmåger.

Store datamængder kan være svære at håndtere. Derfor er der også i Fase 1 udviklet en metode til visualisering, så billederne kan ses direkte i en browser, selvom de fylder mange gigabyte.

Der er også udviklet en metode til iterativ forfining af udvalgte udsnit. Man skal have mulighed for at justere på, hvor mange matches der vises, så der findes så mange objekter som muligt, men uden at der findes for mange fejl. Man skal desuden kunne fravælge eventuelle falsk-positiver, altså felter, hvor det valgte objekt faktisk ikke var.

Endelig er der i Fase 1 udviklet en metode til automatisk træning af neuralt netværk til genkendelse af de fundne objekter. Med søgeresultatet fra forrige trin trænes nu et neuralt netværk, som netop kan genkende de valgte motiver. Dette neurale netværk kan lokalisere en stor andel af de objekter, som ikke var indeholdt i den indledningsvise søgning.

Fase 2: Udrulning

I denne fase blev algoritmerne fra Fase 1 udrullet og gjort tilgængelige for slutbrugere. Målet har været, at de der benytter systemet selv skal kunne foretage de nødvendige undersøgelser, fx via en browser.

Der er i fase 2 blevet tilføjet en mulighed for at gemme en Deep Learning model, som efterfølgende kan genbruges på nye datasæt i samme domæne (fx optælling af fugle fra et andet område). Der er også lavet en brugergrænseflade til valg af neuralt netværk, afhængig af data og anvendelse.

Der er lavet en simpel hjemmeside, hvor man kan uploade/vælge datasæt, træne/anvende neurale netværk, og få præsenteret resultater.

Fase 3: Generalisering

I fase 3 er løsningerne generaliseret og der er indbygget ny funktionalitet, så også fremtidige problemer vil kunne løses med systemet. Til dette formål er det termografiske vildtdatasæt benyttet som et eksempel på, hvordan ganske anderledes problemer kan løses med væsentligt anderledes data, men med samme system.

3 Vejledning til Visual search

Den endelige prototype af Visual Search kan findes på hjemmesiden <u>https://visualsearch.projects.alexandra.dk</u>. Hvis man ikke er logget på, vil siden vise beskeden "There are currently no users signed in!"

Login eller opret bruger

På forsiden klikker man først på "Login", hvorefter man præsenteres for en login-formular. Hvis man ikke allerede har oprettet en bruger, vælger man "Sign up now". Der indtastes e-mailadresse, og klikkes "Send verification code". Efter indtastning af den kode som modtages på mail, kan man vælge password, og klikke "Create".

Når man har en brugerprofil, indtastes e-mail og adgangskode, og der klikkes "Sign in". Bemærk dog, at da det er relativt omkostningstungt at træne neurale netværk, kræves der en forhåndsgodkendelse fra systemets administrator, før login rent faktisk giver adgang til systemet.

Opret projekt

Under "Create New Project" indtastes et projektnavn, og der klikkes på "Create".

Upload foto

Når projektet er oprettet og valgt, klikkes "Upload image(s)", og man vælger den georefererede tif-fil, der ønskes analyseret (Figur 3.1).



Figur 3.1. Screenshot af Visual search. Når projektet er valgt klikkes der på "Upload image(s)".

Indeksér foto

Når billedet er uploadet, skal systemet analysere billedet. Dette gøres ved at trykke på den blå "play"-knap under "indexed" til højre for filnavnet (Figur 3.2).

Activities 🕴 Firefox Web	Browser 🔻	May 3 15:09 🕱		4	💎 9(i)	ŧ,	•
Visual Search	+						8
$\leftrightarrow \rightarrow C$	O A https://naturfondenvisualsearchdev.azurewebsites.net/projects/fdfc8e64-7989-44ff-91	32-973cccc26eed	☆	. ⊘	Ł #	-	=
LOG OUT PROJECTS							
Projects							Т
Flojecis							
Select project							
Eskildsoe 👻							
Create new project							
Project name	CREATE						
Eskildsoe							
Id: fdfc8e64-7989-44ff-9132	373cccc26eed						
Images							
Name Uploaded	Size Id Indexed						
fd1.tif 5/3/2022 3:09:05 P	457652267 aac3a01c-71a8-47c1-9805-a4be3d225297						
UPLOAD IMAGE(S)							
JODS No jobs have been started							
Labels	+						
Config							
settings will be used for next jo							
{ "general": {							
"tile_size_px": 56, "train_size_px":							
56, "tile_overlap":							
0.5							

Figur 3.2. For at indeksere billedet, trykkes på den blå play-knap under "Indexed".

Indekseringen kan tage tid, dog måske kun nogle minutter for små billeder.

Mens processen er i gang, er der en gul cirkel under "Jobs" og "Indexing". Når indekserings-jobbet er afsluttet vises en grøn cirkel og billedet er nu forberedt til at der kan udføres en visuel søgning (Figur 3.3).

Activities 👌 Firefox We	b Browser ▼ May 3 15:38		- 1 9	👻 191	1	1 -
Visual Search	< +					8
$\leftarrow \ \rightarrow \ \mathbf{G}$	O A https://naturfondenvisualsearchdev.azurewebsites.net/projects/fdfc8e64-7989-44ff-9132-973cccc26eed	☆	♥ .	Ł #		=
LOG OUT PROJECTS						
Projecto						
Flojecis						
Loading Select project						
Eskildsoe 👻						
Create new project						
Project name	CREATE					
Fakildaga						11
Eskildsoe						
ld: ldtc8e64-7989-4411-913	-973cccc256ed					
Images						
Name Uploaded	Size id indexed					
fd1.tif 5/3/2022 3:09:05 F	M 457652267 aac3a01c-71a8-47c1-9805-a4be3d225297 🎽 🕨					
LIPLOAD IMAGE/S)						
OPEOND IMAGE(3)						
Jobs						
Indexing						
	.tif (83d2b00b.496b.410c-9882-42061c35e301)					
Labels						
New label	+					
0						
Settings will be used for next is	0					
	N					
<pre>"general": { "tile size px":</pre>						
						4

Figur 3.3. En grøn udfyldt cirkel under "Jobs" og "Indexing" viser, at billedet er indekseret. Ved klik på kort-symbolet, to pladser til højre for den grønne cirkel, tilgås det indekserede kort.

Definér mærkater

Man kan nu definere sine mærkater, altså betegnelser for objekter, som man ønsker at finde. Man skal som minimum angive én kategori af objekter som man ønsker at finde (fx fugle), samt én baggrundskategori. Det er målet, at baggrundskategorien indeholder så stor variation af baggrunde som muligt, men at de(t) objekt(er) man ønsker at finde, ikke også er indeholdt i baggrundskategorien. I eksemplet herunder er valgt mærkaterne "Troldand" og "Baggrund".

Angiv eventuelt særlige indstillinger

Config-feltet er en avanceret indstilling, som ofte kan ignoreres. Feltet kan ses nederst på Figur 3.2.

Under "Config" er der mulighed for at indstille størrelsen på de felter, der bliver analyseret, ved at indstille "tile_size_px" og "train_size_px" (som bør være ens). Standardværdien på 56 pixels er ofte et godt udgangspunkt, men afhængig af opløsning på billedet og objekternes størrelse, kan dette i nogle tilfælde være for meget eller for lidt. Ventetiden vil blive længere hvis man vælger en mindre værdi, og kortere, hvis man vælger en større værdi, da denne felt-størrelse bestemmer, hvor mange felter billedet bliver delt op i.

Teksten "tile_overlap" bestemmer hvor meget overlap der er mellem felterne, og standardværdien på 0.5 betyder, at der er 50 % overlap imellem felterne. Beregningstiden kan reduceres ved at vælge en mindre værdi, men medfører risiko for at der ikke vil være ét felt, som fuldt ud indeholder et ønsket objekt. En værdi tættere på 1.0 betyder markant flere felter der skal beregnes, hvilket forøger beregningstiden men giver mere præcis lokalisering af objekterne. I praksis er standardværdien på 0.5 et godt valg i næsten alle tilfælde.

Opbyg datasæt via visuel søgning

Tryk på kort-symbolet til højre for den grønne cirkel for at tilgå kortet (Figur 3.3). Hermed præsenteres man for det uploadede billede (Figur 3.4):



Figur 3.4. Det indekserede billede vises efter tryk på kort-symbolet.

Man kan zoome ind/ud på billedet ved at scrolle, og man kan panorere rundt ved at klikke og trække med musen. Klik på "i"-symbolet øverst til højre for at få vist hjælp (Figur 3.5):



Figur 3.5. Ved klik på "i"-symbolet kan man få en hjælpetekst frem.



Hold alt-tasten nede på Windows (option-tasten på Mac), og venstreklik på et objekt for at foretage en søgning efter visuelt lignende motiver (Figur 3.6).

Figur 3.6. Forud for dette trin er der klikket på et objekt (her en troldand), og udsnittet her viser (med de tre kvadrater) nogle af de visuelt lignende motiver, som Visual Search har fundet. I højre kolonne vises alle de 50 valgte visuelt lignende motiver

Brugeren præsenteres nu for de 50 bedste visuelle matches. Antallet kan justeres ved at vælge et andet tal for "result limit" øverst. Præsentationen foregår dels ved, at der tegnes en blå kasse omkring "hits" på kortet til venstre, og dels ved, at de relevante udklip vises i listen i højre side. Ved klik på pilen til venstre for hvert udklip, centreres kortet til det valgte objekt.

Målet er nu, at kigge listen igennem for at verificere, at alle hits er "korrekte", altså forestiller det objekt, der søges efter. Klik på det grønne flueben ud for den relevante mærkat (i dette eksempel "Troldand"). Derved bliver der automatisk sat flueben ved alle de fundne objekter i listen. Man kan herefter fjerne fluebenet ved de objekter, som eventuelt er "forkerte", altså ikke udgør et af de objekter som skal findes og tælles. Når listen er gennemgået og fluebenene er sat rigtigt, vælges "Add to training data" nederst, og et nyt objekt kan vælges ved at alt-klikke på kortet.

Efterhånden opbygges datasættet således. Det afhænger af datasættet, hvor mange træningseksempler, der er nødvendige for at opnå en tilstrækkelig præcision. Hvis man ønsker at lokalisere nogle visuelt meget tydelige/unikke objekter (fx fugle hhv. hjorte) på en ensartet baggrund, vil der skulle færre eksempler til, end hvis man ønsker at skelne visuelt meget lignende objekter (fx rådyr hhv. krondyr) på meget varierende baggrunde. En god tommelfingerregel kan dog være, at 1000 eksempler i hver kategori kan være et godt mål at sigte efter, hvis det er muligt. Med en "result limit" på 100 kræver dette dog også kun 10 iterationer for hver kategori, hvilket kan gøres på 5-10 minutter pr. kategori, hvis der ikke skal flyttes mange flueben.

Bemærk i øvrigt, at det er muligt at angive flere mærkater til det samme udklip, hvilket muliggør flere samtidige analyser, fx arts-bestemmelse samtidig med undersøgelse af underlag (fx om fuglen er på rede eller ej).

Træning af neuralt netværk

Når datasættet således er udformet, klikkes på "train model and run" øverst. Det anbefales at lade denne beregning køre natten over, da den er ganske tidskrævende.

Udlæsning af resultater

Efter endt træning, vises en grøn cirkel under Training. Klik på downloadsymbolet for at downloade resultaterne (Figur 3.7).

Den downloadede zip-fil kan nu udpakkes, og indeholder to filer: <classification.gpkg> er en vektor-fil der indeholder information om de felter der er fundet objekter i (deres mærkat og en score), samt <activation_map.tif>, som angiver hvilke pixels objekterne vurderes at være fundet i. Disse indlæses i det følgende i QGIS 3.24.1 (QGIS 2022).

Activities 🕴 Firefo	x Web Browser 🔻	May 4 08:12	🗢 🕸 🐨
🔡 Visual Search	× +		
$\leftarrow \ \rightarrow \ C$	O A https://naturfo	ndenvisualsearchdev.azurewebsites.net/projects/fdfc8e64-7989-44ff-9132-973cccc26eed	\$ ⊚ 🗶 😐 ≣
📊 Quantum 🗳 Doodl	e 👎 Create 📓 Visual Search		
LOG OUT PROJECTS			
Projects			
Select project			
Eskildsoe 👻			
Create new project			
Project name	CREATE		
Eskildsor	2		
ld: fdfc8e64-7989-44f	- 1.9132-973cccc26eed		
Images			
Name Uptoaded	Size id		
rai.dr 5/3/2022 3:0	9.05 PM 457052207 88C3801C-718	86-41C1-3003-840630553731	
UPLOAD IMAGE(S)			
Jobs			
Indexing			
• 🖬 🕅	fd1.tif (83d2b00b-496b-410c-988 Started: 5/3/2022 3:09:16 PM Ri	2-42061c35e301) un time: 12 min	
Training			
• i ±	fd1.tif (573ec50c-54d2-46 Started: 5/3/2022 4:26:06	63-9000-cf8adb65dc79) PM Run time: 75 min	
Imago			
fd1.tif 👻			
Job Select training ich	-		
Solot training job	-		

Figur 3.7. Når træningen er færdig, vises en grøn udfyldt cirkel under "Training". Resultaterne downloades ved klik på downloadsymbolet to pladser til højre for cirklen.

4 Analyse af resultater

Følgende afsnit forudsætter et basalt kendskab til QGIS, men er forsøgt tydeliggjort så meget som muligt. Afsnittet indeholder en del trin, som kan virke overvældende ved første gennemlæsning, men når man først er bekendt med processen, kan nedenstående udføres på få minutter.

I QGIS vælges Layer, Add Layer, Add Raster Layer, og stien til det oprindelige billede vælges som file source (Figur 4.1).



Figur 4.1. Resultaterne er uploadet til QGIS



Indlæs dernæst på samme vis activation_map.tif (Figur 4.2).

Figur 4.2. Sccreenshot af QGIS hvor filen activation_map.tif er uploadet.

Eventuel justering for inspektion

Dette afsnit kan evt. springes over, hvis man ikke ønsker at inspicere data og vælge en tærskelværdi, men for nye brugere anbefales dette trin, for at bibringe en bedre forståelse af dataene.

QGIS' standardvisning af dette lag er ikke nyttig, men zoomet ind kan laget forventes at se ud som i Figur 4.3, inden man foretager de nødvendige justeringer.



Figur 4.3. Standardvisning af filen activation_map.tif i QGIS, når der er zoomet ind og før de nødvendige justeringer er foretage.

Højreklik på activation_map nederst til venstre og vælg "Properties..." Der findes adskillige renderingsmetoder, der ville kunne bruges til dette formål, men følgende bør være et godt udgangspunkt:

Skift "Render type" til "Singleband pseudocolor".

Under "Band" vælges den relevante kategori (i dette eksempel Troldand).

Under Min / Max, Interpolation vælg Discrete.

Klik på farvegradienten ud for Color ramp.

Det anbefales at lave en farveskala, som går fra helt-gennemsigtig til en farve med halv gennemsigtighed. Klik på de små pile under farvegradienten og vælg Delete Stop, så der til sidst kun er én pil længst til venstre, med Opacity 0%, og én pil længst til højre, med Opacity 50% og en farve, i dette eksempel mørkerød. Color ramp bør nu ligne den der er vist i Figur 4.4.



Figur 4.4. Det er vigtigt at lave en farveskala, som går fra helt-gennemsigtig til en farve med halv gennemsigtighed.

Det anbefales at man gemmer denne farvegradient ved at klikke på pil-ned til højre for farvegradienten, og vælge Save Color Ramp. Fremover vil den gemte gradient kunne findes i samme menu under All Color Ramps (Figur 4.5).



Figur 4.5. Den gemte Color Ramp kan fremover findes i menuen.

Klik Classify nederst til venstre, for automatisk at sætte en tærskelværdi (i dette eksempel endte værdien på 97) - klik OK.

De pixels i activation_map, som har en værdi på over tærskelværdien (i eksemplet pt. 97), vises nu i gennemsigtig rød farve (Figur 4.6).



Figur 4.6. Pixels med en værdi på over tærskelværdien vises i gennemsigtig rød farve, der dog i første omgang ikke dækker meget af de fundne troldænder.

Valget af den ideelle tærskelværdi for hvad der betragtes som et "hit" (altså et fundet objekt) kan dog desværre variere fra datasæt til datasæt, og kræver en smule eksperimentering. I eksemplet på Figur 4.6 lader det til, at de røde områder er lige små nok, og faktisk ikke dækker ret meget af de fundne Troldænder. Dobbeltklik igen på activation_map nederst til venstre for at indstille tærskelværdien, dvs. Vælg en anden værdi under "Value" - i eksemplet her vælges nu værdien 50 (Figur 4.7).



Figur 4.7. Tærskelværdien er i dette eksempel sat til 50 og de røde områder dækker nu fuglene på billedet.

Hvis de røde pixels hermed kom til at fylde for meget, ville man skulle vælge en højere værdi imellem de to afprøvede yderpunkter (97 og 50, så næste bud kunne være 75), men eksemplet på Figur 4.7 ser faktisk rimeligt ud. Med den valgte tærskelværdi på 50, kan vi nu polygonisere rasteren.

Polygonisering af raster

Under Raster vælges Raster Calculator (Figur 4.8). Vælg et filnavn til den resulterende tif-fil. I feltet "Raster Calculator Expression" indsættes "activation_map@1" > 50 (hvor 50 er den valgte tærskelværdi i dette eksempel) - klik OK.



Figur 4.8. Raster Calculator i QGIS.

QGIS' standard-visning ser måske igen lidt forvirrende ud pga. hullerne som vist nedenfor, men det skal ignoreres (Figur 4.9).



Figur 4.9. Standard-visning efter valg af Raster Calculator.



Hullerne udfyldes nu via Raster -> Analysis -> Fill nodata, klik Run. Herefter ser det ud som i Figur 4.10.

Figur 4.10. Skærmbillede af projektet efter valg af "Fill nodata".



Med det nye lag ved navn Filled valgt, vælg Raster -> Conversion -> Polygonize (Raster to Vector), klik Run (Figur 4.11).

Figur 4.11. Skærmbillede af projektet efter valg af "Polygonize (Raster to Vector)".

Den store baggrundspolygon kan fjernes vha. Layer -> Filter, under "Provider Specific Filter Expression" indsæt "DN" == 1

Resultatet viser nu en polygon for hvert objekt der er fundet (Figur 4.12):



Figur 4.12. Det store baggrundspolyygon er fjernet og der vises nu et polygon for hvert objekt der er fundet.

Det kan være nødvendigt at bortfiltrere støj i form af meget små polygoner. Vælg Vector -> Geometry Tools -> Add geometry attributes. Klik run. Vælg derefter Layer -> Filter, og indtast fx "area" > 0.05 (Figur 4.13).



Figur 4.13. Query Builder benyttes til at bortfiltrere støj i billedet.

Der kan uddrages summariske statistikker vha. Vector -> Analysis Tools -> Basic Statistics for Fields. Vælg area. Den resulterende HTML-side opsummerer diverse statistikker (Figur 4.14).



Figur 4.14. Med "Analysis Tools" kan man se hvor mange troldænder systemet har fundet, i dette eksempel 2.776.

For denne anvendelse er det umiddelbart mest interessante felt Count, hvor der i dette eksempel er fundet 2.776 troldænder, men bl.a. gennemsnit og standardafvigelse på arealer kan i andre tilfælde være lige så relevante.

Sammenligning med ground truth

I dette eksempel er der manuelt udførte ground truth annoteringer, hvor der er angivet 2.292 troldænder (Figur 4.15).



Figur 4.15. Blå polygoner viser hvor Visual Search har lokaliseret troldænder, og de røde prikker angiver de manuelle annotationer.

På Figur 4.15 angiver de blå polygoner hvor Visual Search har lokaliseret troldænder, og de røde prikker angiver de manuelle annotationer. Kvalitativt

er der god overensstemmelse, om end der i dette udsnit er fundet flere troldænder, end systemet har lokaliseret.

Ved inspektion af billedet langs bredden opdages et antal grene i vandet, som i solen ser meget sort-hvide ud, og visuelt kan minde om troldænderne. Specielt for små stykker træ kan man godt se ligheden med troldænder. På Figur 4.16 er systemets (her fejlagtige) annotationer angivet i rød.



Figur 4.16. Systemet har her fejlagtigt forvekslet små stykker træ med troldænder.

Der er også lokaliseret et antal (fejlagtige) objekter på landjorden, særligt hvor der ligger sten, som igen for et utrænet øje godt kunne minde om troldænder (Figur 4.17).



Figur 4.17. Systemet har her fejlagtigt forvekslet sten med troldænder.

Der er flere måder at opnå en bedre præcision. En løsning er altid at skære sin tif-fil til, så den kun viser det område der skal optælles. I dette tilfælde er det kun troldænder på vandfladen, der er relevante og derfor minimeres fejlene ved at fjerne landjorden fra billedet, hvor troldænderne ikke er. Man skal desuden være omhyggelig med at fortælle systemet, hvad der ikke er troldænder. I dette eksempel gælder det især objekter i billedet der kan minde om troldænder, men ikke er det. Hvis man oplever, at der er fundet fejlagtige objekter, kan man alternativt løbe polygonerne igennem og fjerne de forkerte, hvilket kan være hurtigere end at skulle markere alle objekter fra bunden.

5 Efterprocessering

Efter de tidligere beskrevne trin er gennemført, haves et vektorlag, hvor der ideelt set er en polygon pr. faktisk objekt, og meget få falsk-negativer (oversete objekter), og meget få falsk-positiver (polygoner hvor der faktisk ikke var et objekt). I visse tilfælde, hvor træningen af det neurale netværk ikke har resulteret i en tilstrækkeligt høj grad af genkendelse, kan det dog være meget vanskeligt at vælge en tærskelværdi, som hverken over- eller underestimerer det fundne antal objekter. I en sådan situation vil der fx blive fundet for mange objekter (mange falsk-positiver), men når man justerer tærskelværdien så dette antal reduceres, begynder der også at mangle objekter, som ellers var blevet fundet (der opstår for mange falsk-negativer). Her kan følgende procedure være til hjælp:

Udgangspunktet er det vektorlag, der viser en polygon for hvert objekt, der er fundet (Figur 5.1).



Figur 5.1. Vektorlag, der viser en polygon for hvert objekt, der er fundet.

Når activation_map værdierne ligger tæt på den valgte tærskel-værdi, kan man komme ud for at et fundet objekt (her en troldand) er "fundet" som adskillige små pixel-grupper, fremfor ét sammenhængende område. Hvis man filtrerer små pixel-grupper væk, undgår man mange falsk-positiver (baggrundsstøj, der fejlagtigt vurderes som de eftersøgte objekter), men så forsvinder de objekter også, som ligger tæt på tærskelværdien. Man kan derfor udføre dette trin i et forsøg på at få det bedste af begge verdener:

Brug "Measure line"-værktøjet til at få en fornemmelse af den typiske mindste-afstand mellem objekterne (Figur 5.2).



Figur 5.2. "Measure line"-værktøjet i QGIS.

Vælg Vector -> Geoprocessing tools -> Buffer. Indtast en konservativ værdi som distance, som svarer til knap den halve målte afstand – hvis man fx er rimelig sikker på at objekterne oftest er 0,5 m fra hinanden, kan man fx skrive 0,2 m (Figur 5.3).



Figur 5.3. De fundne områder fra Figur 30 er her udvidet med 0,2 m.

De fundne områder er nu udvidet med den angivne værdi (fx 0,2 m) (Figur 31). Overlappende områder sammensmeltes nu vha. Vector -> Geoprocessing Tools -> Dissolve (Figur 5.4).



Figur 5.4. De overlappende områder er her smeltet sammen.

Efter en tilsvarende buffer med den samme værdi, men med negativt fortegn (i dette eksempel indtastes -0,2 m som buffer distance), ender man tilbage på noget der minder om det oprindelige, men hvor pixelgrupper nu er samlet (Figur 5.5).



Figur 5.5. Der indsættes en buffer med den samme værdi men med negativt fortegn, hvilket samler pixelgrupperne.

Nu kan man bortfiltrere støj i form af meget små polygoner, uden at miste for mange objekter der lå omkring tærskelværdien. Med Identify Features-værktøjet valgt, kan man klikke på nogle typisk falsk-positiver, og også nogle små sandt-positiver, for at få en fornemmelse af hvor små arealer falsk-positiverne typisk har, fx 0,05. Vælg Vector -> Geometry tools -> Multipart to singleparts. Vælg Vector -> Geometry Tools -> Add geometry attributes. Klik run. Vælg derefter Layer -> Filter, og indtast fx "area" > 0.05 (Figur 5.6).



Figur 5.6. Query Builder med filteret "area" > 0.05.

6 Valg af tærskelværdier og filtrering af areal

Nærværende kapitel foreslår en metode, som bør gøre det lettere at vælge fornuftige tærskelværdier (se kapitel 4) og filtrere ikke-ønskede arealer fra.

Valg af tærskelværdier

For at udlede antal fugle fra det dybe neurale netværks output, udvælges først et område med et antal objekter som er overskueligt manuelt at optælle, vist i gul her (Figur 6.1.).



Figur 6.1. Et område på billedet er valgt til manuel optælling.

Da det bør være entydigt om et objekt er inden for eller udenfor dette referenceområde, bør man ikke blot sætte et rektangel, som ofte vil gennemskære objekter.

Der oprettes et nyt vektorlag med geometritypen Point, og der sættes manuelt et punkt pr. objekt, her hættemåger. I dette tilfælde blev der manuelt annoteret 162 fugle (Figur 6.2).



Figur 6.2. Fuglene i polygonet er her manuelt annoteret.

Man kan dække dronebilledet med et sort baggrundslag, fx vha. Vector -> Research Tools -> Extract Layer Extent, og vælge dronebilledet som input. Det resulterende lag ved navn Extent kan så sættes til sort farve (Figur 6.3).



Figur 6.3. Dronebilledet er dækket med et sort baggrundslag.

Ved visning af activation_map-laget, og justering af tærskelværdien for hvilke pixels der betragtes som "hits", kan man nu bedre vurdere, hvilken tærskelværdi der giver flest sand-positiver (overlap mellem de lyseblå "ground truth"-cirkler og de røde fundne områder).



Figur 6.4. Visning af activation_map-laget.

Efter gennemgang af konverteringsproceduren, for at omdanne raster til vektor, og efterprocessering med Buffer og Dissolve, som beskrevet tidligere, kan man med Vector -> Geometry Tools -> Centroids generere ét punkt pr. fundet objekt, som fx kan visualiseres som blå prikker med ground truth vist som røde cirkler (Figur 6.5).



Figur 6.5. Alle fundne objekter visualiseres som blå prikker.

Referenceområdet muliggør desuden automatisk optælling vha. Vector -> Geoprocessing Tools -> Clip (for at beskære de fundne objekter til referenceområdet). Optællingen fås vha. F6 (attribute table), idet man scroller til bunden for at se hvor mange objekter der er fundet. Det bemærkes dog, at falskpositiver og falsk-negativer i nogen grad udligner hinanden, og det samlede antal er derfor ikke nødvendigvis en ideel metrik (idet man fx kunne have 0 i afvigelse men et stort antal falsk-positiver, så længe der er lige så mange falsknegativer).

Filtrering af landareal

Særligt på vandet findes der mange objekter, hvor det er uklart om de bør betragtes som sandt- eller falsk-positiver, fx i udsnittet på Figur 6.6.



Figur 6.6. Flere objekter kan skyldes artefakter i overgangen mellem to sammensatte billeder.

Dette reflekterer nok til dels, at det også ved træningssættets tilblivelse kan være vanskeligt at beslutte, hvornår et udklip skal betragtes som "hættemåge" eller hvornår det er et artefakt fra billed-sammensyningen. Såfremt man udelukkende ønsker at finde fugle på land, simplificeres opgaven dog noget, og kan forventes at give bedre resultater, hvis træningssættet udelukkende indeholder fugle på land.

For at bortfiltrere fugle på vand, kan man tegne en grov kontur omkring øen (Figur 6.7).



Figur 6.7. En grov kontur er tegnet omkring øen, så vandarealet kan blive fjernet.

Igen bruger man Vector -> Geoprocessing Tools -> Clip til at fjerne objekter, som er udenfor det angivne polygon.

Bemærk at hvis man *ikke* ønsker at filtrere til dyr på land, anbefales det stadig at udføre en clipping af fundne objekter til det aktuelle billedområde, da systemet desværre pt. har en tendens til også at lokalisere objekter udenfor billeddækningen. En sådan clipping maske kan nemt genereres vha Raster Calculator, indtast fx "egaa_engsoe@1">0 som udtrykket (i eksemplet her er billedlagets navn egaa_engsoe).

Kvalitativt ser resultatet rimeligt ud, med langt de fleste hættemåger lokaliseret, om end stadig med et antal falsk-positiver (Figur 6.8).



Figur 6.8. Systemet har lokaliseret langt de fleste hættemåger. Dog er der et antal falsk-positive, men justering af tærskelværdien resulterede i 99 % præcision.

Med denne metode og de valgte tærskelværdier findes der sammenlagt 5.643 hættemåger på landjorden. For fuld præcision bør resultatet dog gennemgås særligt mhp. at fjerne falsk-positiver, men dette bør stadig være tidsbesparende ift. fuld manuel annotation. Ved manuel optælling var der fundet 5.015 individer, svarende til en overestimering på 12,5 % før manuel korrektion. Det bemærkes dog, at tærskelværdi og buffer-afstand ville kunne vælges, så man er mere tilbøjelig til at underestimere end overestimere, når dette er at foretrække. Et mere konservativt valg af en tærskelværdi på 40 fremfor 30 (samt 0.1 m buffering og areal-filtrering på 0.05 m2) resulterede fx i et resultat på 4.953 individer, svarende til 99 % præcision.

Når man opnår erfaring med systemet, bør det kunne give solide estimater med under 10 % afvigelse fra resultatet ved manuel optælling, som man i øvrigt vil kunne ramme nøjagtigt med manuel tilretning. En sådan tilretning koster naturligvis tid, men må forventes at tage markant kortere tid, end en fuld manuel annotering.

7 Optælling af hjorte på termografiske billeder

Visual Search er et alsidigt værktøj, som kan bruges til at finde visuelt lignende objekter, i princippet underordnet hvad de forestiller. Træningen af specialiserede neurale netværk for særlig god genkendelse af bestemte objekter, forudsætter dog et betydeligt antal træningseksempler. Det er umuligt på forhånd at give et præcist antal for hvor mange eksempler, der behøves, da dette bl.a. afhænger meget af variationen inden for de objekter som søges, og af hvor forskelligt det man søger, er fra alt det, man *ikke* søger. Erfaringen viser dog, at 1.000 eksempler er et godt udgangspunkt, som ofte giver gode resultatet. I nogle situationer kan man måske opnå samme kvalitet med væsentligt færre eksempler, og andre gange behøves flere.

For fuglekolonier, som ofte tæller mange tusind individer, er dette ikke nødvendigvis et problem, særligt idet træningssættet netop opbygges via visuel søgning, så man med forholdsvis få klik (ned til 3 klik hvis man er heldig) kan annotere et betydeligt antal eksempler (fx 100 stk.) samtidig. For populationer af hjortevildt er antallet dog markant lavere, hvilket udgør en udfordring for deep learning-baserede løsninger. Den oprindelige visuelle søgning, som genkender visuelt lignende objekter via et generelt neuralt netværk (som altså ikke er trænet til at genkende netop det man klikker på), bør være velegnet til at kunne lokalisere selv få individer på store billeder, men den efterfølgende træning, som har til formål et forbedre genkendelsen, og gøre det muligt at udlede hvilke pixels der indgår i et objekt, har behov for et stort antal træningseksempler.



Figur 7.1. Hjortevildt er her lokaliseret på et dronebillede fra Ovstrup hede. Pilene til venstre i højrekolonnen panorerer til udklippenes oprindelsespunkt.

Som Figur 7.1 viser, formår den visuelle søgning at lokalisere udklip med hjortevildt. Umiddelbart kan udklippenes standard-størrelse på 56 pixels se ud til at være lige store nok i dette tilfælde, hvor dyrene kan være ned til ca. 5 pixels i bredden, men omvendt er baggrunden også forholdsvis heterogen. Når man som menneske foretager den visuelle søgning for at opbygge et datasæt, og støder på et udklip man er i tvivl om, kan man ofte afgøre sagen ved at klikke på pilen (Figur 7.1) for at panorere til udklippets oprindelsespunkt, hvor man netop kan se den omkringliggende baggrundsintensitet og -mønster. Af denne grund kan det være en god idé her at anvende lidt "for store" felter, som fx 56 pixels, da de afslører mere af omgivelserne. Samtidig forøges beregningstiden ved mindre udklip, da billedet opdeles i flere udklip.

I Holm m.fl. (2020) er det vist, at dyrenes køn, art og alder kan estimeres ved bl.a. at se på dyrenes areal og omkreds. Disse parametre vil også kunne udtrækkes af den klassifikations-tif-fil, som resulterer fra træningen. Da der ikke er tale om ægte instance segmentation, vil det dog ikke være muligt at "skille" dyr ad, som rører hinanden, hvilket i visse tilfælde også vil være vanskeligt for et menneske at gøre.

For at teste Visual Search på termografiske hjortebilleder, blev et træningssæt på ca. 250 dyr og 300 baggrunde etableret. Dette forarbejde tager ca. 15 minutter. En træning af datasættet blev herefter igangsat. Indekseringen af det store billede tog ca. 8 timer, mens træningen på dette begrænsede datasæt kun tog 26 minutter. Udklip fra resultaterne er vist i Figur 7.2 og 7.3.



Figur 7.2. Test af Visual Search på termografiske hjortebilleder.



Figur 7.3. Test af Visual Search på termografiske hjortebilleder.

Da der er langt imellem grupper af hjortevildt, som fordeler sig over et stort område, blev der i dette eksempel kun benyttet billeder, hvor der var dyr på. Pga. de enkelte billeders georeferering, blev den resulterende mosaik alligevel et meget stort billede, hvor over 99 % af pixels var "no data". Disse udelades dog fra søgningen, og den visuelle søgning fungerer fint også under disse omstændigheder.

Som billederne viser formår systemet i de fleste tilfælde at lokalisere hjortene med dette datasæt. Jo flere data systemet trænes med, jo mindre usikkerhed nås i forhold til det faktiske tal. Hvis man medtager alle billeder fra et datasæt, og ikke manuelt udvælger netop de billeder hvor der er dyr på som i eksemplet her, kunne der være langt imellem grupperinger af dyr, og ikke mindst her kunne det være hjælpsomt med en automatisk indikation af relevante områder til optælling, selv hvis den blev foretaget manuelt.

Et delformål med den termografiske monitorering af hjortevildt er at kunne udlede bl.a. areal og omkreds på hjortene, som kan benyttes til at estimere art, køn og alder på dyrene. Til dette formål viser disse resultater, at et træningssæt på ca. 250 positive eksempler er for lidt til at skabe tilstrækkelig robusthed til, at de præcise pixels lokaliseres. I en videre udvikling vil det være relevant at undersøge, hvor mange træningseksempler der skal til, for at en tilpas præcis kortlægning af de relevante pixels kan foretages.

8 Konklusion

Nærværende første version af Visual Search formår at lokalisere visuelt lignende objekter på store dronebilleder af forskellig karakter, fx RGB-billeder af fuglekolonier og termografiske billeder af hjortevildt. Den visuelle søgning kan benyttes til hurtigt at etablere et træningsdatasæt bestående af fx 1.000 billedeksempler, som ved et klik på én knap, bruges til at træne et specialiseret neuralt netværk, til netop at lokalisere de valgte objekter. I samme automatiske trin anvendes det resulterende netværk til at analysere det samlede billede, og tildeler en "score" til hver enkelt pixel i dronebilledet.

Dette output kan via standard-operationer i et geoinformationssystem (fx QGIS) efterbehandles for at konvertere til lokationer (og evt. arealer og omkredse) på de fundne dyr. Systemets endelige præcision afhænger delvist af menneskets erfaring ift. at vælge fornuftige tærskelværdier, men foreløbige undersøgelser indikerer, at man kan ramme indenfor 10 procents afvigelse fra det faktisk antal, inden der evt. udføres manuel korrektion. Med manuel korrektion kan fejlen komme ned på 0 %, men kræver naturligvis mere arbejdstid – dog forventeligt stadig væsentligt mindre, end hvis alle dyr manuelt skulle annoteres enkeltvis.

Systemet giver for nuværende muligheder for at genbruge trænede modeller på fremtidige billeder, så hvis man opnår en model der generaliserer godt, og som er særligt god til at genkende en bestemt art, kan man drage nytte af den igen og igen. Det vil desuden teknisk set være muligt i fremtiden at udbygge systemet med yderligere funktionalitet, der gør det muligt at pulje datasæt på tværs af projekter, hvilket kan forventes at forbedre både generalisering og præcision.

Systemet har ligeledes været testet på termografiske dronebilleder af hjorte. De foreløbige resultater viser, at visuel søgning kan bruges til at lokalisere hjortene, og at der kan opbygges et datasæt på samme måde som det har været gjort med fugle på RGB-billeder. Resultatet af træningen var et neuralt netværk, som formåede at lokalisere hjorte med rimelig præcision, men for arealog omkreds-baseret analyse kræves meget præcise segmenteringsmasker, som kræver betydeligt større mængder træningsdata. Dette kan være svært at opnå på et enkelt dronebillede, da dyrenes antal i sagens natur er lavere end fx ynglefuglenes, men det er fortsat vurderingen, at det med tilstrækkelige datamængder vil være teknologisk muligt at opnå.

Fremtidige forbedringsmuligheder

Undervejs i udviklingsprojekter opstår der ofte flere nye idéer til forbedringer og udvidelser, end der kan realiseres indenfor en i forvejen fastlagt ramme. Nedenfor er angivet en liste over forbedringsforslag:

 Automatisering af QGIS-trin. Pt. foretages efterprocesseringen manuelt i QGIS, hvilket tager tid. Selvom der kræves menneskelig vurdering af visse parametre, skulle det være muligt at automatisere disse processer, i første omgang direkte i QGIS, men på sigt kunne det også gøres som en del af Visual Search cloud-applikationen.

- Genvejstaster under visuel søgning for hurtigt af sætte/flytte flueben ved mærkater. Dette vil kunne gøre det markant hurtigere at opbygge træningssæt.
- Udvidelse af zoom-niveauer. Pt. er det begrænset hvor langt man kan zoome ind og ud, og det kunne somme tider være rart med en større rækkevidde.
- Indbygget vejledning. En "hjælper" eller interaktiv guide der fortæller hvad man skal gøre, hvor mange eksempler man bør sigte efter, om dyrene skal være helt i billedet etc.
- Stitching direkte i systemet. Ville gøre det muligt at uploade georefererede enkeltbilleder, uden forudgående stitching.
- Understøttelse af ikke-georefererede billeder. I tilfælde, hvor information om areal og omkreds ikke er nødvendig, kunne billeder evt. håndteres uden georeferering.
- Understøttelse af enkeltbilleder, så der kan søges på tværs af billeder, om de så er sammensatte eller ej.
- Tværgående datasæt, så et datasæt kan samles fra flere forskellige projekter, og således udbygges, når der fx kommer et nyt dronebillede med en given art.
- Standardmodeller. Når der fx er trænet en god model til en bestemt art eller type billeder, kan den vælges som udgangspunkt, og bruges i indekseringen, hvilket gør den visuelle søgning stærkere. Gør det desuden muligt for en bruger slet ikke at opbygge et datasæt, hvis der skal tælles en velundersøgt art, fordi der måske ligger en standardmodel klar.

9 Referencer

Holm, T.E., Pedersen, C.L. & Jørgensen, H. 2020. Brug af drone med termisk kamera til overvågning af hjortevildt. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 30 s. - Teknisk rapport nr. 169 http://dce2.au.dk/pub/TR169.pdf

Holm, T.E. & Bregnballe, T. 2020. Kortlægning af ynglefugle i Filsø, Birkesø og ved Gyldensteen Strand i 2020 ved brug af drone. Notat fra DCE - Nationalt Center for Miljø og Energi; Nr. 2020 | 58, 25 s.

Holm, T.E. & Bregnballe, T. (red.) 2019. Overvågning af ynglefugle ved brug af droner. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 70 s. - Videnskabelig rapport nr. 311 http://dce2.au.dk/pub/SR311.pdf

Holm, T.E., Kanstrup, N., Riddervold, M., Jensen, L.Ø. & Bregnballe, T. 2019. Droner forstyrrer ikke ynglende fugle. Modelflyvenyt, Bind 43, Nr. 4, 08.2019, s. 4-7.

Holm, T.E., Kanstrup, N., Riddervold, M., Jensen, L.Ø. & Bregnballe, T. 2018. Brug af droner til overvågning af ynglende vandfugles reaktioner på menneskelig færdsel. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 32 s. - Teknisk rapport nr. 129 <u>http://dce2.au.dk/pub/TR129.pdf</u>

QGIS. 2022. QGIS Geographic Information System. QGIS Association. http://www.qgis.org

VISUAL SEARCH

Et system til optælling af fugle og pattedyr på dronebilleder

Aarhus Universitet har siden 2017 udviklet metoder til droneovervågning af ynglende vandfugle og hjortevildt. De nye metoder har vist sig at være tidsbesparende i felten, at give bedre data og forstyrre mindre end traditionelle optællingsmetoder. Den efterfølgende optælling fra dronebilleder har til gengæld vist sig at være tidskrævende, da dette gøres manuelt. For at afhjælpe dette, har vi i projektet "Visual Search" udviklet en prototype på et system, der ved hjælp af deep learning kan lokalisere og optælle dyr på billeder. Formålet har været at erstatte den nuværende manuelle og tidskrævende optælling på billeder og dermed nedsætte tidsforbruget med op til 90%. En sådan besparelse kan være med til at gøre droneovervågning billigere og dermed betydeligt mere udbredt end i dag. Denne tekniske rapport beskriver udviklingen og brugen af Visual Search-prototypen.

ISBN: 978-87-7156-690-1 ISSN: 2244-999X

