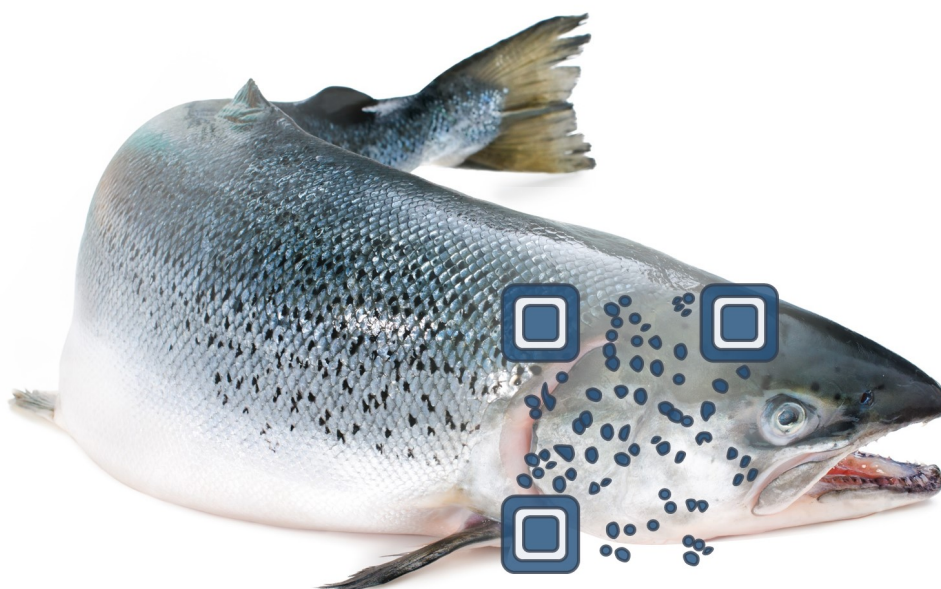
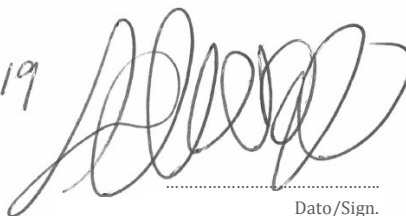


Verifikasjon av algoritmer for biometrisk identifikasjon av lakseindivider – SalmID

(Fase 2)



«Ett unikt mønster for hvert individ»

25/4-19 

Dato/Sign.
Prosjektleder: Aleksander Lillienkiold

29/4-19 

Dato/Sign.
Kvalitetssikrer: Harry Westavik

29/4-19 

Dato/Sign.
Prosjekteier: Hanne Digre

FHF901474 / SITNEF302003585
Verifikasjon av algoritmer for biometrisk
identifikasjon av lakseindivider – SalmID (Fase 2)

Rapportnummer: 2019:00375 v.1
Dato: 2019.04.02

Forfattet av:

Aleksander Lillienkiold
Jonatan Dyrstad
Morten Bondø

ISBN: 978-82-14-06835-1

Prosjektstøtte gjennom: Fiskeri- og havbruksnæringens forskingsfond



Nøkkelord:

Biometri

Prikkeidentifikasjon

Unik identifisering

Algoritmeutvikling

Deep-Learning

Laks

Oppdrett

Slakteri

Prikkemønster

Mønstergjenkjenning

Deteksjon av unike markører

Markører for identifikasjon

Produktkontroll

Produktoversikt

Logistikk

Produksjonskontroll

Individuelle tilpasning

FORORD

I lakseindustrien er fleksibilitet svært viktig for å kunne levere et mangfold av produkter til konsumentene, på en effektiv og lønnsom måte. På samme tid ønskes det å utnytte råvaren (fisken) på optimal måte for størst mulig andel høykvalitetsprodukter ut til forbrukeren.

For å gjennomføre en fleksibel produksjon hvor hvert enkelt individ skal få en tilpasset behandling er det viktig å ha god sporing og identifikasjon på individene. Gjennom forløperen til dette prosjektet jobbet SINTEF med å etablere en metode for bruk av prikkemønster som et biologisk fingeravtrykk for individet. Resultatene var gode, og viste til at dette kunne være mulig.

Forprosjektet viset til at det datagrunnlaget som var samlet på laboratoriet og fisken som ble benyttet ikke var presentert på en representative måte for industrialisering. Det ble også satt spørsmålsteget ved antall fisk som ble brukt i gjenkjennings algoritmen som oppnådde 99.8% nøyaktighet. Flere og større forsøk ble foreslått.

Dette prosjektet var basert på verifikasjon av algoritmen utviklet under forprosjektet. Målet var å gjennomføre forsøk på anlegg, og etablere hvorvidt algoritmen var stabil og om dette kunne fungere som identifikasjon av unike mønster på laks—et mønster basert på fiskens «fregner».

Metoden og algoritmen ble testet ut på anlegg, og resultatene viser at dette er mulig. Det vises også til en 100% gjenkjenning på lakseindividene avbildet under forsøket. Men det er her viktig å legge merke til at dette er gjort på et utvalg bestående av omtrent 9.500 fisk, og det er ikke enda bekreftet at alle fisk har unike mønster om man ser variasjonen gjennom mange dager med full produksjon.

PREFACE

For the salmon processing industry flexibility is extremely important for preserve the ability to produce a large variety of products for the consumer in an efficient and profitable manor. While variety and flexibility in volume and type is key – it is still waiting and needed to utilize and optimize the potential of the raw materials in such a way that the consumer can receive the best products possible.

To enable such flexible productions it is important that each individual fish get a unique identifier making it possible for specific adaptive handling for optimized production and product care. In multiple project at SINTEF there has been done work in the area of quality analyses on both inner and outer visual ques.

Addressing the flexibility needed within the industry SINTEF has for many years seen tagging of individual fish as a key challenge in the operations. RFID tags and other physical methods of tagging has been looked at. However, this has been something the industry have been reluctant to utilize due to the impact on the fish and cost. This type of tagging done not only need physical contact with the fish, but it also leaves a mark, devaluing the product.

In this project we utilize an established algorithm, for a prior project, to verify and improve on the algorithm. Showing that the pattern is recognizable not only in lab condition but in the field as well. The algorithm had earlier a success rate of 99.8%, something which needed verification. Our goal was to determine whether or not dot-pattern on the side of the fish was unique and can be used as a signature, like a finger print.

Method and algorithm was tested in field, and the results show that it is possible. We get a 100% recognition of salmon induvial during our trails. Of course more work is needed to make the algorithm invariant and robust, however on our selection of approximately 9.500 fish, we can say that is it possible to identify the individual.

Keywords:

Identification
Unique identifiers
Algorithm development
Deep-Learning

Salmon
Fish farming
Fish slater

Dot patterns
Pattern recognition
Detection of unique markers
Markers for identification

Production control
Production logistics
Logistics
Individual adaptation

INNHOILDSFORTEGNELSE

Forord.....	3
Preface	4
Unik identifikasjon av lakseindivider	6
Mål og formål	6
Oppdatering av konseptet	6
Algoritmen i et nøtteskall	9
i. Deteksjon av prikker.....	9
ii. Fra prikker til konstellasjoner	10
iii. Unik beskrivelse av individer	11
Gjenkjenning av tidligere sett/identifisert fisk	12
Distorsjon	13
Forsøksoppsett, bilde innsamling.....	14
Validering av algoritmen.....	16
Algoritme-forsøk.....	17
Algoritmen har flere justerbare parametere	17
De evaluerte målekriteriene benyttet var	17
Konstellasjonstreffrate/gjenkjenning	18
Resultater.....	18
Forsøk 2).....	19
Forsøk 3).....	21
Diskusjon & Konklusjon	22
Veien videre	23



Unik identifikasjon av lakseindivider

Dette prosjektet tar for seg verifikasjon av algoritmer utviklet i *Identifikasjon av lakseindivider: Biometri fase 1 (FHF901263)*. Utfordringen som var fremsatt i det forprosjektet var^[1]: Er det mulig å unik identifikasjon lakseindivider uten fysisk inngripen/kontakt med fisken?

Dersom dette er mulig ville en laks kunne følges gjennom hele prosessen. Det kan gi en bedre forståelse av hvor lenge et enkelt individ har vært til utblødning, hvor lang transporttid før den går til filetering, sikkerhet i at kvalitetsdata følger denne enkelte laksen og at individet som passer til filet går dit, mens dem som egner seg som hel rund går til det. Altså et helautomatisk system som identifiserer laksen—beholder unik informasjon om produksjon, kvalitet, vekt, lengre etc. uten å gjøre en fysisk kontakt.

1. Det anbefales å les gjennom tidligere rapport av Eilertsen et al. (2017)—*Identifikasjon av lakseindivider: Biometri fase 1*

Mål og formål

Gjennom dette prosjektet har målet vært å vise til hvorvidt algoritmer fra *Identifikasjon av lakseindivider: Biometri fase 1 (FHF901263)* [Eilertsen et al. – 2017] ville fungerer på større datasett samlet ute på anlegg i virkelig omgivelser.

Ved å vise til industrinært billedtakningsoppsett vil det være med på å avduke mulige utfordringer som må tas hensyn til ved et eventuell neste steg.

Oppdatering av konseptet

For å kunne gjøre en identifikasjon eller verifikasjon basert på biometri, kreves det å finne fysiologiske karaktertrekk. For den norske laksen har vi etablert at melaninprikkene på siden av skinnen, på hodet og kroppen vil være et slikt fysiologisk trekk.



Enhver laks har et prikkemønster, håpet er at dette mønsteret er tilstrekkelig unikt til å kunne brukes som et type digitalt «fingeravtrykk». Forsøk har vært gjennomført som indikere at dette er mulig. Denne rapporten utvidere testdatasettet og søker et svart på hvorvidt dette bør gå over til industrialisering.

En algoritme for å differensiere basert på prikkemønster er utviklet ved SINTEF gjennom tidligere prosjekt.^[1]

Algoritmen som ble etablert baser seg på deep-learning metoder for prikke deteksjon, og en metode som laget konstellasjoner av fire-og-fire prikker. Metoden som ble bruket var originalt tiltenkte identifikasjon av stjernekonstellasjoner på natte himmelen.

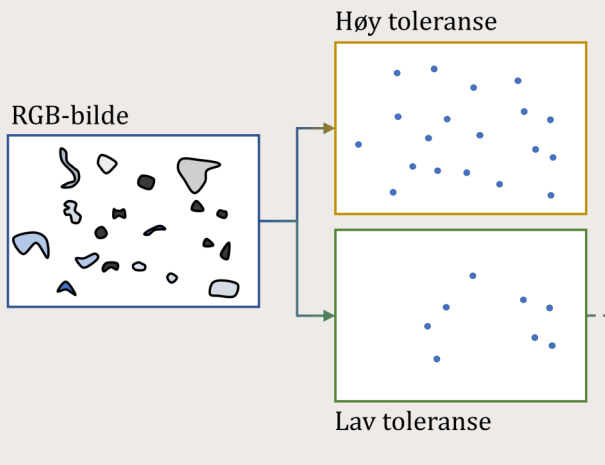
Basert på det tidligere prosjektet ble det besluttet at det ikke var godt nok grunnlag for å gå videre med en full industrialisering. Det ble etablert at det fremdeles var usikret hvorvidt prikkemønsteret er unikt nok. I dette prosjektet er det jobbet med å samle inn

en større database for å verifisere koden som er utviklet, samt vurdere unikheten i prikke mønstrene. Prosjektet har utvidet fra rundt 600 fisk til nesten 10.000 fisk.

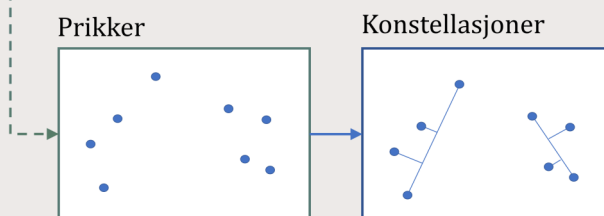
Konstellasjonene som skaptet i tidligere prosjekt lette etter ulikhet blant alle prikker på hver enkelt fisk, men det var tydelig at en del konstellasjoner var lett å finne igjen i mange fisk. Dette har det nå blitt jobbet med å håndtere bedre. Videre er det jobbet med maskinlært prikkeidentifikasjon. Teknikken er blitt bearbeidet for å kunne justere på sensitiviteten til prikkedeteksjonen, da det viste seg fra tidligere prosjektet at ikke alle prikker var så lett å definere klart som prikk eller støy.

Algoritmen i et nøtteskall:

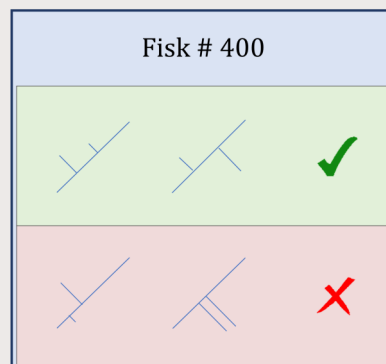
i. Deteksjon av laksens prikker i et fargebilde. Hva som skal telles som prikker kan justeres med en toleranse-parameter.



ii. Sammensetting av prikker til beskrivende mønstre/konstellasjoner. Et sett med regler brukes for å gjøre om prikkene til konstellasjoner av 4 og 4 prikker.



iii. Konstruksjon av en unik beskrivelse av individet ved bruk av konstellasjonene fra forrige hovedtrinn. Lakseindividet blir beskrevet med en liste som forteller hvilke konstellasjoner som er tilstede på laksen og hvilke som ikke er det.



ALGORITMEN I ET NØTTESKALL

Algoritmen er bygget på nytt i dette prosjektet, basert på tidligere arbeid, er det nå utviklet en trestegs-løsning. Algoritmen er delt opp i forskjellige oppgaver for å kunne lettere gjenbruke metoder og videre utvikle/bygge på koden ved behov.

Oppdelingen i av oppgaven er som beskrevet til, og illustrasjonen under viser hvordan hver av disse oppgaven (simplifisert) opererer og levere som utgangssignal.

Algoritmen består av følgende tre hoved steg:

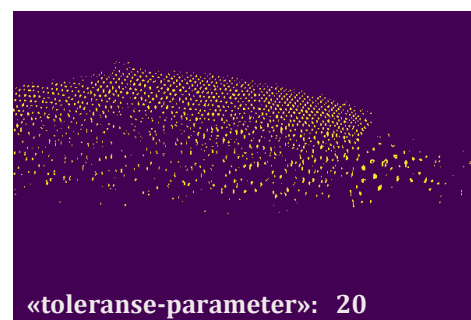
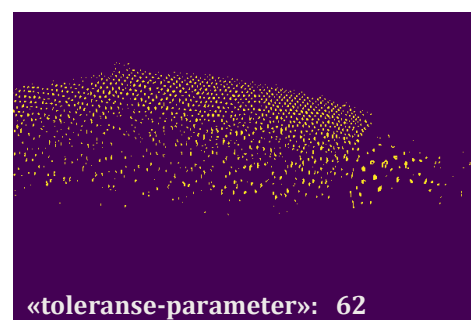
- i. **Deteksjon av laksens prikker i et fargebilde.**
- ii. **Sammensetting av prikker til beskrivende mønstre/konstellasjoner.**
- iii. **Konstruksjon av en unik beskrivelse av individet ved bruk av konstellasjonene.**

Mer utdypende om hvert steg kommer i på følgende kapitler.

i. Deteksjon av prikker

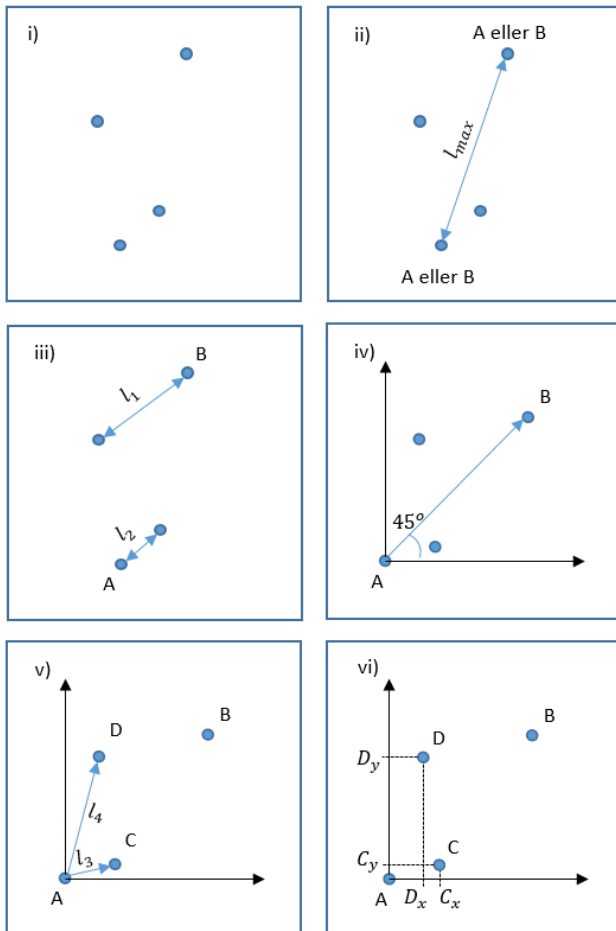
Prikkedeteksjonen gjøres ved bruk av et nevralt nettverk. Det nevrale nettverket er trent til å produsere et binærbilde hvor alt som ser ut som prikker settes lik 1 og alt som ikke ser ut som prikker settes lik 0. Det nevrale nettverket kan justeres slik at det kun beholder de tydeligste prikkene, eller slik at det beholder alt som ser «prikkeaktig» ut ved å justere «toleranse-parameteren». **Illustrasjon til høyre viser** resultat ved to (2) forskjellige «toleranse-parametere» og samme inngangsbilde.

Deretter gjøres binærbildet om til en liste med xy-koordinater basert på massesentrene til de detekterte prikkene.



ii. Fra prikker til konstellasjoner

En konstellasjon består av fire prikker og prikker settes sammen til konstellasjoner ved bruk av et sett med regler. Disse reglene gir hver prikk et navn, enten A, B, C eller D. Deretter brukes prikk A og B til å definere et koordinatsystem og prikk C og D beskrives i dette koordinatsystemet. For en gitt prikk genereres en konstellasjon med en 6-steps prosess:



- i) Av alle laksens prikker, lokaliseres de tre som ligger nærmest den valgte prikken og sammen skal disse fire prikkene danne en konstellasjon.
- ii) Avstanden mellom alle prikkene måles og de to prikkene som er lengst fra hverandre skal defineres som enten A eller B.
- iii) Avstanden fra de to A-/B-kandidatene til de to andre prikkene måles og den kandidaten som ligger nærmest en annen prikk blir definert som A, den andre som B.
- iv) Prikkene skaleres og roteres slik at avstanden mellom A og B er lik 1 og linjen mellom A og B er 45° på x-aksen.
- v) De to resterende punktene skal nå defineres som enten C eller D. Hvilken prikk som blir hva avgjøres ved å måle avstanden fra C-/D-kandidatene til A, og den av kandidatene som ligger nærmest A defineres som C, den andre D.
- vi) Til slutt defineres et koordinatsystem vha. A og B og prikkene C og D beskrives i dette koordinatsystemet med en x- og y-komponent.

Disse reglene er laget slik at et gitt sett med prikker alltid produserer den samme konstellasjonen uavhengig av fiskens avstand til og rotasjon i forhold til kameraet. Dette er viktig fordi fiskens posisjon og orientering i forhold til kameraet ikke er kjent og er også vanskelig å finne nøyaktig for ikke-rigide objekter slik som fisk, særlig hvis fisken også skal prosesseres og potensielt endre fasong og fortsatt være gjenkjennbar i etterkant (eks. hodekapping og fjerning av innmat).

iii. Unik beskrivelse av individer

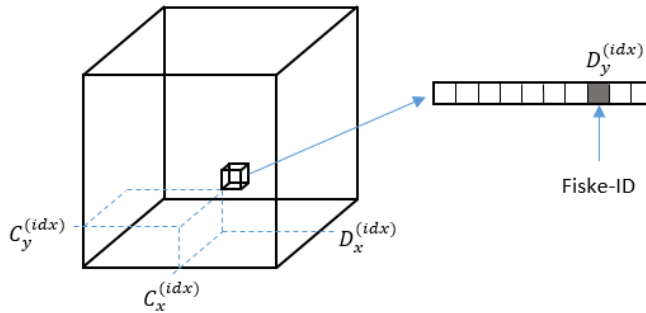
Alle fiskeindividerne lagres i en database som består av en 4-dimensjonal liste med dimensjoner $Ls \times Ls \times Ls \times Ls$. Når en ny fisk skal legges til i databasen benyttes konstellasjonene fra forrige steg. For hver av laksens prikker har det blitt generert en konstellasjon som kan beskrives på formen:

$$\mathcal{V}_{konst} = [C_x, C_y, D_x, D_y]$$

Konstellasjonsvektoren, $\mathcal{V}_{konst}$, gjøres om til en indekseringsvektor for 4D-listen ved å multiplisere den med listestørrelsen, Ls , og avrunde elementene til nærmeste heltall:

$$\mathcal{V}_{konst}^{(idx)} = \text{round}(\mathcal{V}_{konst} * Ls)$$

Det gjøres et oppslag i listen på lokasjonen gitt av indekseringsvektoren og fiske-ID-en legges til her. På denne måten blir fiske-ID-en lagt til på like mange lokasjoner i listen som det er prikker på fisken.



4D-indeksering av fiske ID, figuren illustrere hvordan fiske-IDer er lagret i en fire-dimensjonal matrise. Dette gir en form for oppslags verk, hvor identifiseringen av en ukjent fisk rett og slett hentes ut fra en plass i denne matrisen.

GJENKJENNING AV TIDLIGERE SETT/IDENTIFISERT FISK

Gitt at det er bygget en database med fisk kan vi forsøke å identifisere fisk på følgende måte:

1. ta bilde av fisken du ønsker å gjenkjenne og detekter prikker (som tidligere),
2. sett sammen prikkene til konstellasjoner (som tidligere),
3. slå opp i databasen på lokasjoner gitt av konstellasjonene, men: I stedet for å legge fiskens ID inn på gitte lokasjoner, hent ID-en til all fisk som allerede finnes på disse lokasjonene.

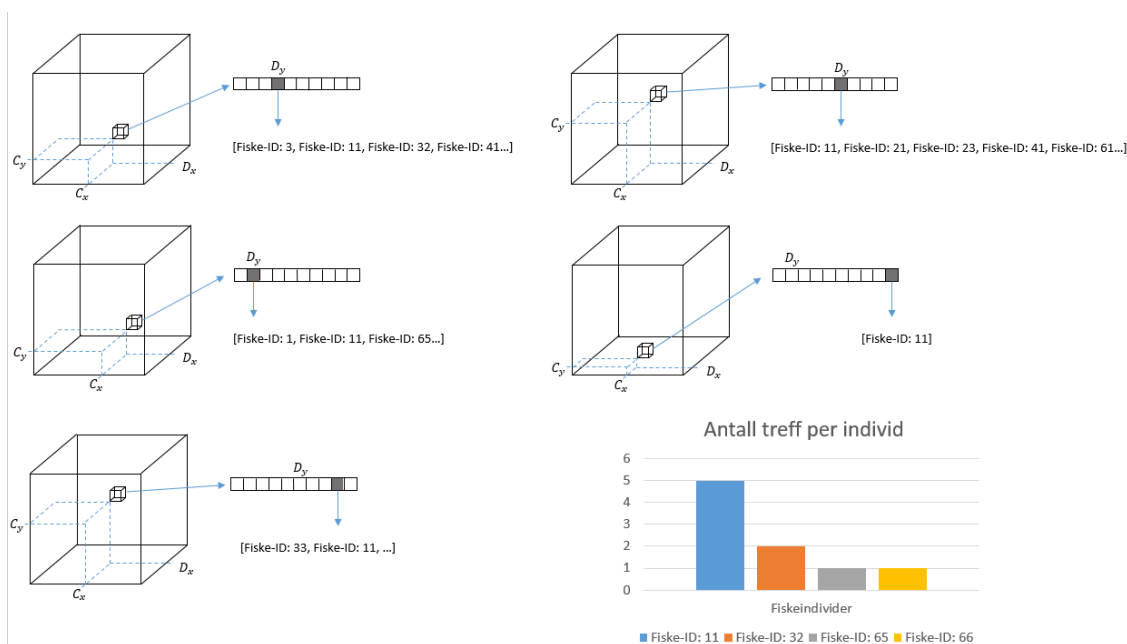
Resultatet av dette er et sett med fisk med forskjellige antall "treff". Noen ID-er vil kun bli funnet på et par lokasjoner, mens andre vil bli funnet på flere og forhåpentligvis vil én fisk skille seg ut som den fisken med klart flest treff.

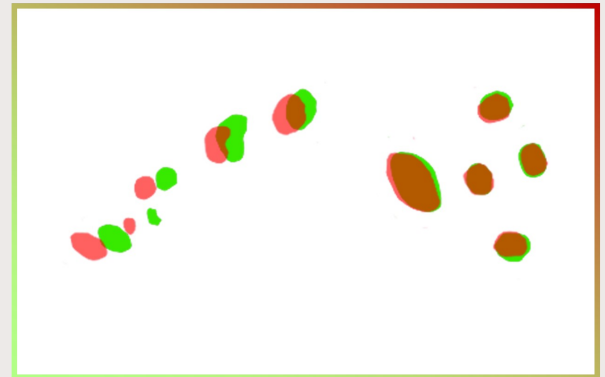
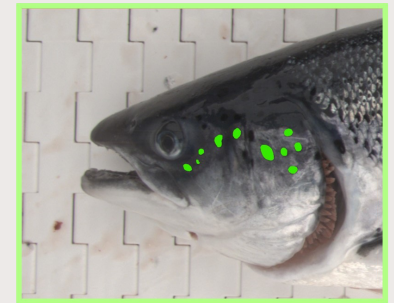
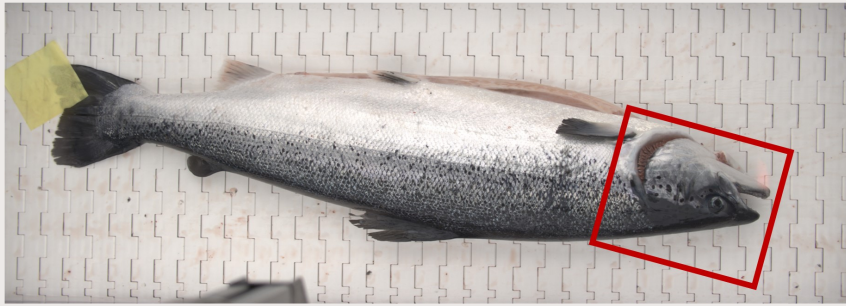
I tillegg til dette blir det gjort en sjekk basert på forskjellen i antall prikker på fiskene som sammenliknes. Hvis en fisk har veldig mange flere prikker enn

en annen, over en grense *cut-off* (i %), vil ikke disse fiskene telles som en match selv om de skulle ha mange treff. Begrensning på hvilke fisk som skal sammenliknes basert på antallet prikker på fisken.

Fordi algoritmen baserer seg på å telle antall treff og hver fisk har svært mange konstellasjoner, vil den kunne bli robust mot støy. Ingen del av fisken er vesentlig for å gjenkjenne individet og det kan derfor være mulig å kjenne igjen fisk, selv hvis noen av prikkene er dekket av skitt eller liknende. Det vil også kunne fungere selv om fisken er hel i et bilde og uten hode og hale i et annet.

En annen fordel med systemet er at gjenkjenningen av et individ er uavhengig av antallet fisk som individet skal sjekkes mot. Prosessen går kun ut på å slå opp i en tabell og dette gjør systemet skalerbart til sanntidsapplikasjoner i industrisetting.





Forvrengning

Geometrisk forvrengning oppstår når et objekt er avbildet fra forskjellige vinkler. Et prikkemønster ekstrahert fra et bilde tatt med halen nærmest kameraet vil ikke se likt ut som et prikkemønster ekstrahert fra et bilde tatt med hodet nærmest kameraet. Som vist i dette eksemplet finnes det ingen rotasjon og skalering av dette prikkemønsteret som kan få alle prikkene til å overlappe samtidig. Algoritmen er ikke designet for å takle slike forvrengninger, men et forsøk ble gjort for å teste hvor sårbart systemet er for denne typen perspektivendringer.

Perspektivforskjeller

Bildene over viser samme fisk, avbildet to ganger fra forskjellige vinkler. Det skaper perspektivforskjeller og forvrengning i prikkemønsteret. Det vil si, som nederste bildet viser, mønstrene er nå ikke identiske. Uten å vite mer om 3D formen vil det ikke være mulig å si at dette er samme mønster—da det vi ser kan være naturlig variasjon mellom to fisk. Men 3D-informasjon ville det potensielt vært mulig å fjerne forvrengningen som kommer fra flate (2D) bilder.

FORSØKSOPPSETT, BILDEINNSAMLING

Det ble lagd et kameraoppsett bestående av 3 stykk FLIR Blackfly S (BFS-U3-200S6C-C) med 20MP IMX183-sensor. Da kamera har rullende lukker (rolling shutter) måtte oppsettet lages med pulserende lys (strobe). For å synkronisere triggering av lys og kamera ble det benyttet en Gardasoft LFC420 kontroller.

Det ble benyttet 4 stykk hvite LED-paneler (4000K) på 120x60cm, som ble satt opp i en vifteform rundt fisken for å gi jevn belysning med minimalt med reflekser uten behov for å bruke polarisering av lyset.

Kameraene ble montert slik at de ga en liten forskjell i perspektiv. De ble koblet opp til en datamaskin via USB 3.0 og bildene ble lagret ved hjelp av et LabVIEW-program. Kameraoppsettet ble montert over et transportbånd hvor fisken transporteres fra graderingen til filetlinjen. Fisken ble ikke rørt under transport.

Det ble gjennomført en datainnsamling 19.-20 juni 2018 der det ble tatt 3 bilder av 3400 fisk. Etter å ha analysert disse bildene ble det gjort noen justeringer på oppsettet med finjustering av bildevinkler og belysning.

Deretter ble det gjennomført en ny datainnsamling 18.-20. september 2018 hvor det ble tatt 3 bilder av 9100 fisk. Denne gangen ble enkelt fisk vilkårlig plukket fra linjen og kjørt 3 gang gjennom forsøks-riggen. Totalt bl 90 fisk kjørt gjennom 3 ganger hver, slik at man totalt har 9 bilder for hver av disse fiskene.





Forsøks oppsett over samlebånd på anlegg

Det ble satt opp fire led-lys som en tunnel, med kamera tittende inn fra forskjellige vinkler fra toppen. Fisken passert uforstyrret gjennom belysningstunellen.

VALIDERING AV ALGORITMEN

Det ble utført tre ulike forsøk for å verifisere algoritmen og for å finne begrensningene ved systemet. I alle forsøkene ble det tatt to forskjellige bilder av et sett med fisk ved at fisken ble markert med en lapp og kjørt gjennom kamerariggen vist på bildet på forrige side.



Forsøk 1) Bilder tatt med fisk samme vei: Sjekker algoritmen med minst mulig distorsjon av prikkemønsteret som følge av perspektivendringer og liknende.



Forsøk 2) Bilder tatt med fisk samme vei, men med ett bilde rotert i software slik at et av bildene har rygg opp, og et har rygg ned: Sjekker om algoritmen er rotasjonsinvariant uten å tilføre distorsjon som følge av perspektiv-endringer. (Bildet under t.h. er bildet over t.h. rotert med 180 grader).



Forsøk 3) Bilder tatt med fisk liggende forskjellig vei på transportbåndet: Sjekker om algoritmen er perspektivinvariant. Fordi det benyttede kameraet står i vinkel i forhold til båndet fører dette til distorsjon av prikkemønsteret.

Algoritme-forsøk

For hvert av forsøkene beskrevet på forrige side ble 9162 ukjente fisk lagt til i databasen sammen med 158 kjente fisk. Forsøkene ble utført ved at et nytt bilde av de kjente 158 fiskene ble benyttet for å prøve å gjenkjenne den riktige fisken blant de totalt 9320 fiskene i databasen.

Et perfekt resultat på testen vil derfor være 158 treff av 158 mulige og sannsynligheten for å finne riktig fisk med ren gjetning er på 0,01%.

Algoritmen har flere justerbare parametere

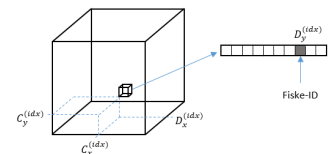
- ◆ **Listestørrelsen:** L_s – Dette beskriver oppløsningen på 4D-tabellen i databasen. (Beskrevet under iii. Unik beskrivelse av individer i Algoritmen i et nøtteskall.)
- ◆ **Toleranseparameteren:** tp – Dette forteller hva som skal telle som prikker i fargebildet. En lav tp betyr at det meste teller som prikker, mens en høy tp betyr at kun de tydeligste prikkene brukes. Det vil si en type filetering på «støy» i bildet.
- ◆ **Matchende prikkeantall:** $cut-off$ – Denne parameteren setter en terskel på hvor likt et antall prikker på to fisk må være for at de skal sammenliknes. Hvis $cut-off$ er lik 0.2 betyr det at kun fisk med $\pm 20\%$ forskjell i totalt prikkeantall sammenliknes.

For å validere algoritmen ble det kjørt et «grid search» over alle de justerbare parametrene, slik at et sub-sett av alle mulige kombinasjoner ble testet.

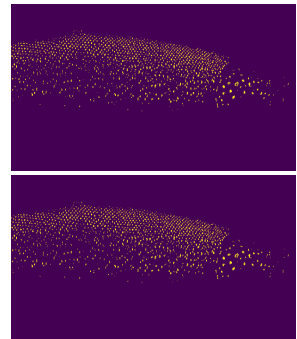
De evaluerte målekriteriene benyttet var

- ⇒ **Treffprosent:** Hvor ofte finner algoritmen den riktige fisken.
- ⇒ **Topp N treffprosent:** Hvor ofte finner algoritmen den riktige fisken med N forsøk.
- ⇒ **Konstellasjonstreffrate ifht. nærmeste fisk:** Hvis riktig fisk ble funnet, hvor nært var det at algoritmen bommet.

Listestørrelse



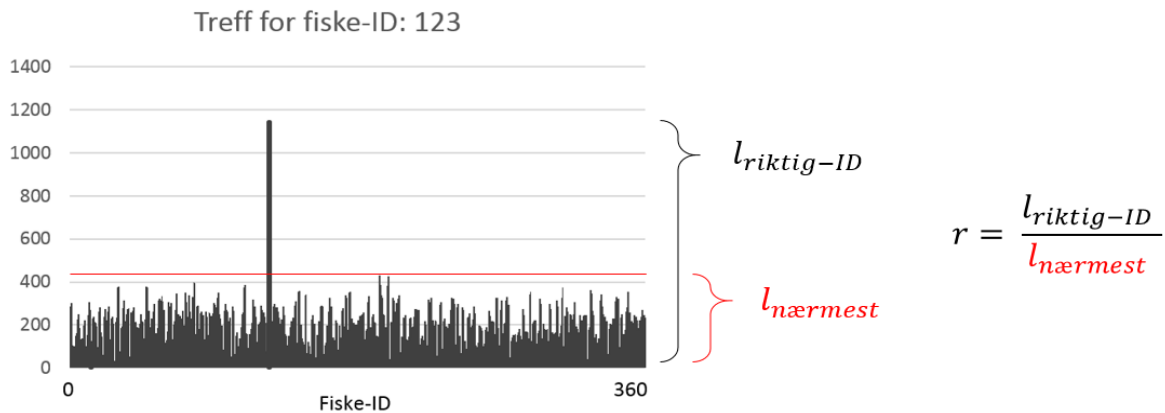
Toleranse



Konstellasjonstreffrate/gjenkjenning

Fisk gjenkjennes ved å sammenlikne konstellasjoner på fisken man ønsker å gjenkjenne og all fisken i databasen. Gitt at algoritmen fant riktig fisk kan man kvantifisere hvor nære det var at den bommet på gitt fisk ved å finne forholdet r mellom antall treff på den riktige fisken $l_{riktig-ID}$ og antallet treff på den fisken med nest flest treff $l_{nærmest}$.

Under testing ble konstellasjonstreffraten r funnet for alle fisk. I hver test ble den fisken med lavest r kalt r_{min} , den fisken med høyest r kalt r_{max} og gjennomsnitts- r ble i testen ble også funnet som $r_{gj.snitt}$.



Illustrasjon av sammenlikning av felles konstellasjoner.

Det viser sammenhengen mellom nærmeste like individ og individet som faktisk gjenkjennes. I denne illustrasjonen er det lett å se hvilket individ som er gjenkjent. Men slik er det ikke alltid, og det er derfor r er så viktig å følge med på.

RESULTATER

Det var gjennomført flere forsøk, disse er beskrevet under VALIDERING AV ALGORITMEN. Forsøkene ble delt opp, slik som beskrevet, for å vurdere robustheten av maskinsyn/billedtakningen. For alle testen ble det også gjennomført en tuning/justering av parametere for finne beste gjenkjenningen av enkelt individene. Her presenteres resultatene fra **Forsøk 2)** og **3)**. **Forsøk 1)** er kun en enklere versjon av Forsøk 2); og er valgt å ikke bli presentert her da Forsøk 2) (den mer utforende oppgavene) ga så gode resultater og Forsøk 1) ikke ville gitt noe mer til beskrivelsen av resultatene.

Forsøk 2)

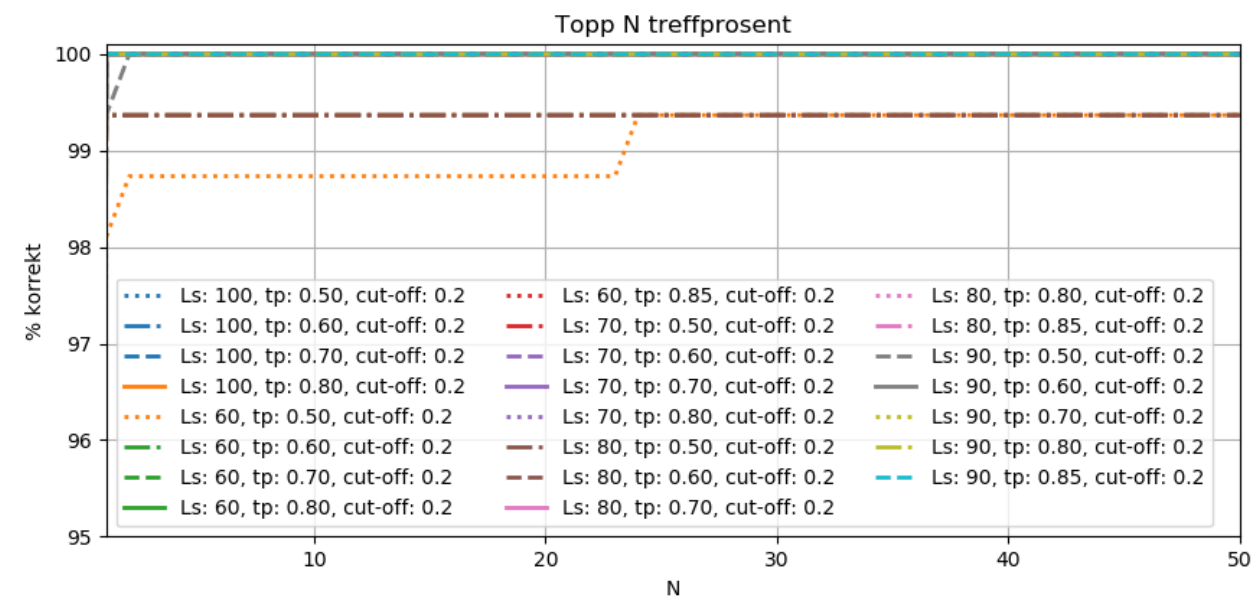
Som vist i tabellen gav alle testene svært gode resultater. Algoritmen identifiserte riktig fisk nesten uavhengig av listestørrelsen Ls og toleranseparameteren tp. Dårligst ut kom algoritmen med Ls=60 og tp=0.5, men selv da ble 155 fisk identifisert korrekt og 3 fisk feilidentifisert.

I grafen ser vi topp N treffprosent. Av dette kan vi lese at algoritmen med Ls=90 og tp=0.50 får 100% treff med to forsøk, mens Ls=80, tp=0.5, Ls=70, tp=0.5, og Ls=60, tp=0.5 får 99.37 % selv med 50 forsøk, som tilsvarer 1 av 158 feilidentifiserte fisk.

		Listestørrelse (Ls=)				
		60	70	80	90	100
Toleranse parameter (tp =)	0.50	98.1 %	99.4 %		99.4 %	100.0 %
	0.60	99.4 %	100.0 %	100.0 %	100.0 %	100.0 %
	0.70	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %
	0.80	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %
	0.85	100.0 %		100.0 %	100.0 %	

Tabell for suksessfulle treff av individer for forsøk 2).

Denne tabellen lister suksessen for gjenkjenning i Forsøk 2). Dog det er mange metoder som gir 100% er det viktig å forstå hvor godt fiske individene var skilt fra hverandre. Se tabell under, jf. illustrasjonen under kapitelet for *Konstellasjons-treffrate*.

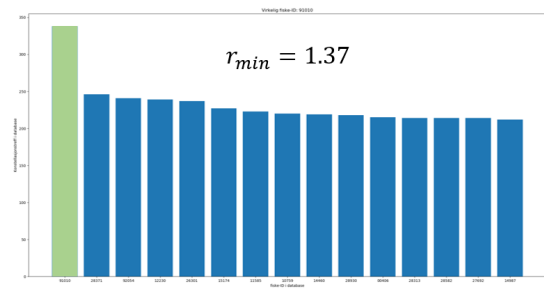
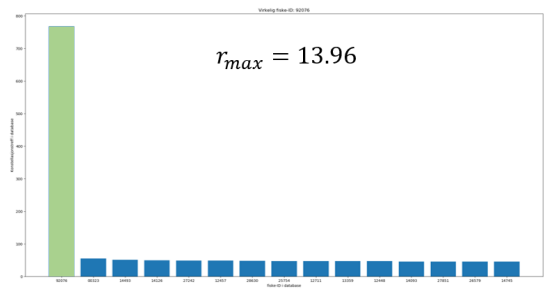


Treffprosent Forsøk 2)—Hvor mange forsøk må algoritmen ha for å finne rett individ.

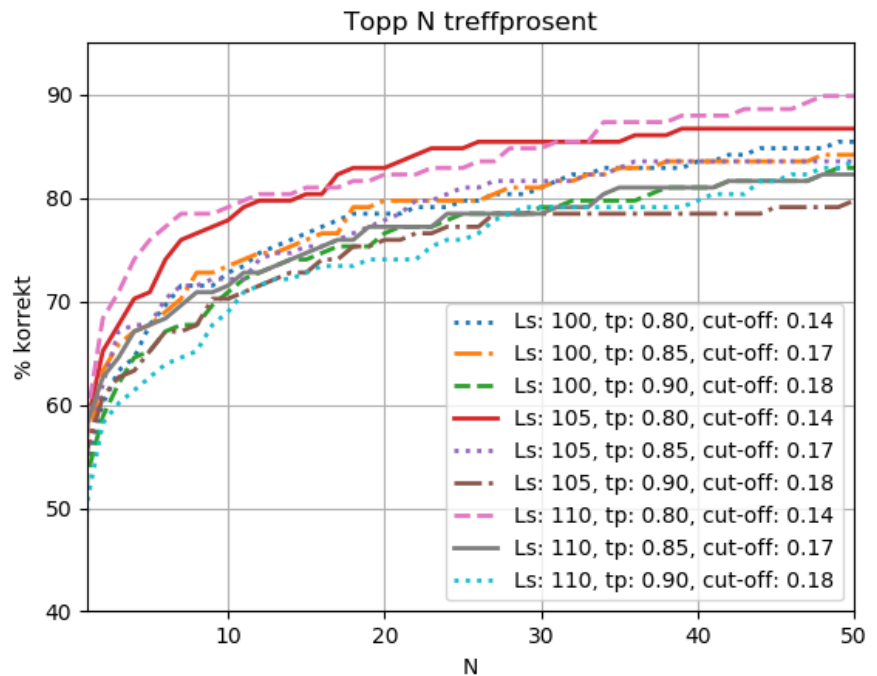
Selv om flere av testene gav resultater med 100% treffrate, er det forskjell på hvor godt fisken skilles fra hverandre med ulike parametere. Dette kvantifiseres med $r_{\text{n rmest}}$, $r_{\text{gj.snitt}}$ og r_{max} , som beskrevet p  forrige side. Av tabellen til h yre ser vi at testen som skilte den vanskeligste fisken best var med $L_s=100$, $tp=0.7$, og den testen som i snitt skilte fisk best var med $L_s=100$, $tp=0.8$.

Dette er enda bedre representert i plottet som viser r_{max} og r_{min} for testen med listest rrelse $L_s=100$, og toleranseparameter $tp=0.70$. Den gr nne stolpen indikerer antall treff p  korrekt fisk, mens stolpene til h yre viser de 14 fiskene i databasen med h yest antall konstellasjonstreff.

			Listest�rrelse (L_s =)					Toleranseparameter (tp =)		Tabell for treffrate/ unikhet av individer for Fors�k 2)
			60	70	80	90	100			
0.50	r_{min}		-	-	-	0.98	1.05			
	$r_{\text{gj.snitt}}$		1.67	2.05	2.49	2.96	3.41			
	r_{max}		3.00	4.20	5.52	6.59	7.52			
0.60	r_{min}		-	1.01	1.06	1.10	1.24			Denne tabellen kvantifiserer hvor forskjellig/like individene var i konstellasjonen deres. Jo s�rre tall bo mer distinkt var et prikkem�nster.
	$r_{\text{gj.snitt}}$		1.87	2.36	2.87	3.43	3.94			
	r_{max}		3.69	5.17	6.43	7.59	8.70			
0.70	r_{min}		1.02	1.11	1.29	1.33	1.37			
	$r_{\text{gj.snitt}}$		2.12	2.70	3.36	4.03	4.66			
	r_{max}		5.09	7.32	9.87	12.87	13.96			
0.80	r_{min}		1.11	1.12	1.18	1.11	1.23			
	$r_{\text{gj.snitt}}$		2.44	3.16	3.93	4.71	5.48			
	r_{max}		6.76	9.48	12.70	16.13	17.95			
0.85	r_{min}		1.07		1.04	1.09				
	$r_{\text{gj.snitt}}$		2.65		4.31	5.19				
	r_{max}		8.86		16.06	22.09				



Konstellasjonstreffrate— $L_s=100$, $tp=0.70$. Gr nne stolpen indikerer antall treff p  korrekt fisk, stolpene til h yre viser de 14 fiskene i databasen med h yest antall konstellasjonstreff.



Treffprosent Forsøk 3)—Hvor mange forsøk må algoritmen ha for å finne rett individ.

Forsøk 3)

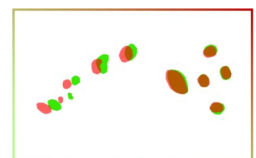
I tabellen til vises treffprosenten for algoritmen på datasettet med perspektivendringer. Algoritmen er ikke beregnet for å håndtere dette, men har allikevel over 50% treff på alle testene.

I grafen vises topp N treffprosenten og den viser tydelig at perspektiv er et problem for algoritmen. Selv med 50 forsøk oppnår den beste testen kun 90% riktig identifikasjon.

		Listestørrelse (Ls=)		
		100	105	110
Toleranse parameter (tp =)	0.80	54.43 %	58.23 %	58.86 %
	0.85	55.70 %	55.06 %	58.23 %
	0.90	53.16 %	55.70 %	50.63 %

Tabell for suksessfulle treff av individer for forsøk 3).

Denne tabellen lister suksessen for gjenkjenning i Forsøk 3). Det vi ser er at algoritmen ikke håndterer store skift i prikkemønsteret. Dette er ikke overraskende. Jf. Distorsjons kapitlet.



Distorsjon

DISKUSJON & KONKLUSJON

Målet med dette prosjektet har vært å verifisere og finne begrensningene til algoritmen for gjenkjenning av fisk basert på prikkemønster som ble utviklet i Identifikasjon av lakseindivider: *Biometri fase 1* (Eilertsen et al. (2017)).

Testene som er gjort viser at algoritmen fungerer svært godt, selv i storskala industriforsøk hvor enkeltindivider kan differensieres fra tusenvis av andre fisk. Likevel er marginene i noen tilfeller små og algoritmen vil derfor trolig ikke være robust nok til å kunne håndtere mange titalls tusen fisk av gangen helt feilfritt. I tillegg er det tydelig at algoritmen er sårbar for perspektivendringer noe som betyr at det nåværende systemet egner seg best i kontrollerte miljøer slik som på samleband i prosesseringsanlegg hvor fisken ligger noenlunde flatt.

Det finnes flere åpenbare algoritme forbedringer som kan gjøres og som vil gi bedre resultater. Blant annet kan informasjon om prikkenes relative størrelse og orientering benyttes som ekstra dimensjoner i databasetabellen eller man kan benytte 3D-informasjon for å projisere prikkene i forskjellige plan. Hva som burde fokuseres på er avhengig av hva algoritmen skal brukes til.

Til bruk i kontrollerte miljøer, slik som i prosesseringsanlegg, kan det være naturlig å fokusere på å øke gjenkjenningsmarginene

(konstellasjonstrefferatene r) slik at hver fisk skiller seg mer fra alle andre, (f.eks. ved å bruke prikkenes relative størrelse og orientering).

Hvis algoritmen skal benyttes til å spore laks fra merdkanten (eller før) helt til middagstallerkenen vil det mest sannsynlig bli behov for å projisere prikkene annerledes med 3D-informasjon. Dette kan for eksempel gjøres ved bruk av stereo-kameraer på anlegg og ved å ta flere bilder av middagstallerkenen med en smarttelefon.

Forsøkene utført i dette prosjektet har vist at prikkemønsteret på laks kan benyttes til å gjenkjenne enkeltindivider. I de utførte forsøkene ble enkeltindivider gjenkjent i en database på 9320 fisk, men om dette skalerer slik at prikkemønsteret kan benyttes alene for å skille millioner av laks fra hverandre forblir foreløpig et åpent forskningsspørsmål.

En stor fordel med den utviklede algoritmen er at den ikke baserer seg på bruken av et globalt koordinatsystem på fisken. Den er også robust i form av at flere av prikkene kan mangle og fisken vil fortsatt kunne gjenkjennes. I praksis betyr dette at en fisk kan gjenkjennes for eksempel både før og etter hodekapping.

Det utviklede systemet har et enormt potensiale og kan muliggjøre full sporbarhet av laks gjennom hele verdikjeden. Man kan se for seg en applikasjon hvor sluttbrukeren kan ta et bilde av en laks i kjøledisken og få tilgang til all informasjon som er samlet opp om individet gjennom produksjonsprosessen.

PROSJEKTFASER

I søknaden til dette forprosjektet ble det skissert en mulig fremdriftsplan/steg for å komme til en mulig kommersialisering av ideen sammen med en skisse av fremtidens muligheter.

FØRSTE FASEN (DEN NÅVÆRENDE)

Testing av hypotese. Utstysrleverandører, sluttbruker og andre får tilgang til forskning og kunnskap vedrørende muligheten for bruk av biometrisk skanning av laks som hensiktsmessig teknologi for unik individuell identifikasjon.

ANDRE FASE

Utvikle prototype (opptil TRL-nivå 5). Utstysrleverandørene får kompetanseoverføring fra FoU-miljø og dermed tilgang til teknologi som kan gjøre biometrisk skanning tilgjengelig for norsk laksenæring.

TREDJE FASE

Industrialisering av teknologien. Biometrisk skanning integrert i et lakseslakteri gir bedre sporbarhet internt i lakseslakteriet med mulighet for bedre logistikk og mer fleksibel layout av produksjonslinjer og –utstyr.

FREMTIDEN

Det er håp om at biometrisk skanning av laks, og eventuelt annen fisk, kan revolusjonere sporbarheten i hele verdikjeden, på individnivå—helt fra hav til fat.

VEIEN VIDERE

Algoritmen som er utviklet og prikkemønsteret virker til å være egnet for sporing internt i en fabrikk. Det er vanskelig å garantere unikheter da det er snakk om biologisk materiale med en usikker variasjonsbredde. Forsøkene utført så langt viser til at dette kan og bør tas videre.

For anlegg og dag-til-dag produksjon ser vi det som gjennomførbart å bruke mønsteret som en markør for enkeltindivider. En markør som dette vil at kunne brukes ved vært dataregistreringspunkt, hvor ny kvalitativ informasjon om fisken er aggregert.

Data tilhørende enkeltindividene kan være med på å drive frem en mer tilpasset og optimal utnyttelse av dette fantastiske råstoffet.

Potensialet ved individuell gjenkjenning er enormt. Dette kan brukes til alt fra sporing gjennom tilvekst for bedre helse og kvalitet, for individinformasjon for forbruker, selger, tilbakemelding fra forbruker. Mulighetene gjennom livsløpet, dersom mønsteret er unik, et uant.

Vi mener derfor at det burde gjøres et større helhetlige studier på prikkemønsteridentifikasjonen for å kunne vurdere bruken av en slik ikke-invasiv gjenkjenning i flere ledd. Et slik system vil kunne revolusjonere flere ledd av industrien. Det vil være viktig at et slik studie blir stor nok og følger identifiserbare individer gjennom flere generasjoner/livsløp.

FHF901474 / SITNEF302003585
Verifikasjon av algoritmer for biometrisk
identifikasjon av lakseindivider – SalmID (Fase 2)

Rapportnummer: 2019:00375 v.1
Dato: 2019.04.02

Forfattet av:

Aleksander Lillienkiold
Jonatan Dyrstad
Morten Bondø

ISBN: 978-82-14-06835-1

