

20. januar 2017

Undersøgelse omkring brug af billedanalyse til inspektion af sygdomme ved indvendig kødkontrol på kyllingeslagteri

Udarbejdet af
Jens Fagertun, Ph.d.
Anders Jørgensen, M.Sc.

ClassifEYE VetInspector



INTRODUKTION	3
BILLEDOPTAGELSEN	3
DATASÆT.....	6
MANUEL KLASSIFIKATION	6
EFTERBEHANDLING AF DATASÆT FOR LEVER	8
EFTERBEHANDLING AF DATASÆT FOR HJERTE.....	9
ILLUSTRATION AF VARIATIONEN I LEVER KATEGORIER	10
ILLUSTRATION AF VARIATIONEN I HJERTE KATEGORIER.....	12
METODE	13
VISION PIPELINE	13
SEGMENTERING AF ORGANER.....	13
PIXEL KLASSIFICERING AF ORGANER	13
KLASSIFICERING AF ORGANER PER BILLEDE	13
VALIDERING	15
RESULTATER	15
LEVER.....	15
HJERTE	17
SAMMENLIGNINGSGRUNDLAG TIL FORANALYSERAPPORTEN	18
CASE OMKRING FORVENTET FLOW I PRODUKTION	19
VURDERING AF Udstyr	19
DISKUSSION	20
USIKKERHEDER I DEN MANUELLE KLASSIFIKATION.....	20
USIKKERHEDER I SEGMENTERINGEN AF ORGANER.....	20
RESULTATER OG TEST AF SYSTEMET	21
ANDRE ÅBNE SPØRGSMÅL.....	21
KONKLUSION	22
REFERENCER	22
APPENDIKS A – ILLUSTRATIONER AF KATEGORIER IKKE BEHANDLET I RAPPORTEN.....	23
APPENDIKS B – ILLUSTRATIONER AF KATEGORIER BEHANDLET I RAPPORTEN	26
APPENDIKS C – BESKRIVELSE AF ALGORITMEN FOR SEGMENTERING AF LEVEREN	32
APPENDIKS D – BESKRIVELSE AF VISION PIPELINEN BEHANDLING AF LEVEREN.....	34
APPENDIKS E – UDREGNINGER TIL CASE OMKRING FORVENTET FLOW I PRODUKTION	36
APPENDIKS F – COLLAGER OVER LEVERSTASE OG NEKROTISERENDE LEVERBETÆNDELSE	37

Introduktion

I dag udføres den lovpligtige kødkontrol af indmadssæt på kyllingeslagterier af opkvalificerede slagterimedarbejdere under supervision af uddannet veterinærpersonale. Slagterimedarbejderne står ved slagtelinjen og tager de indmadssæt af, der enten er åbenlyst sygt eller hvor de er i tvivl, hvorefter en dyrlæge vil inspicere de aftagne sæt. Både indmadssæt og fuglens krop aftages, men kroppen kan hænges på linjen igen hvis veterinæren vurderer denne rask.

For at sikre en høj koncentration hos slagterimedarbejderne, så de ikke står for lang tid ved slagtelinjen roterer de position. Med stigende produktionshastigheder er der et øget behov for udvikling af udstyr til brug for automatisk visuel inspektion til at assistere eller aflaste den manuelle kontrol.

Denne rapport er en del af GUDP projektet ”Computer vision baseret kontrolstation til aut. veterinærkontrol af slagtekyllinger på slagteri”. Fokus for denne rapport er at belyse muligheder og problemstillinger ved at bruge computer vision ved den indvendige veterinærkontrol. I rapporten arbejdes med hele vision processen ved den indvendige kontrol - lige fra at optage målinger (billeder) og segmentering af vigtige organer til detektion/klassificering af sygdomme.

Rapporten bygger videre på den viden der er sammenfattet i ”Forundersøgelse for inspektion af indmadssæt vha. computer vision” [1]. I denne rapport vil der i resultatafsnittet blive draget paralleller til foranalyserapporten [2], der beskriver mulighederne ved computer vision i den udvendige veterinærkontrol.

Grundlaget for forundersøgelsen er en billedoptagelse udført hos Danpo A/S i Aars i perioden oktober-november 2016.

Billedoptagelsen

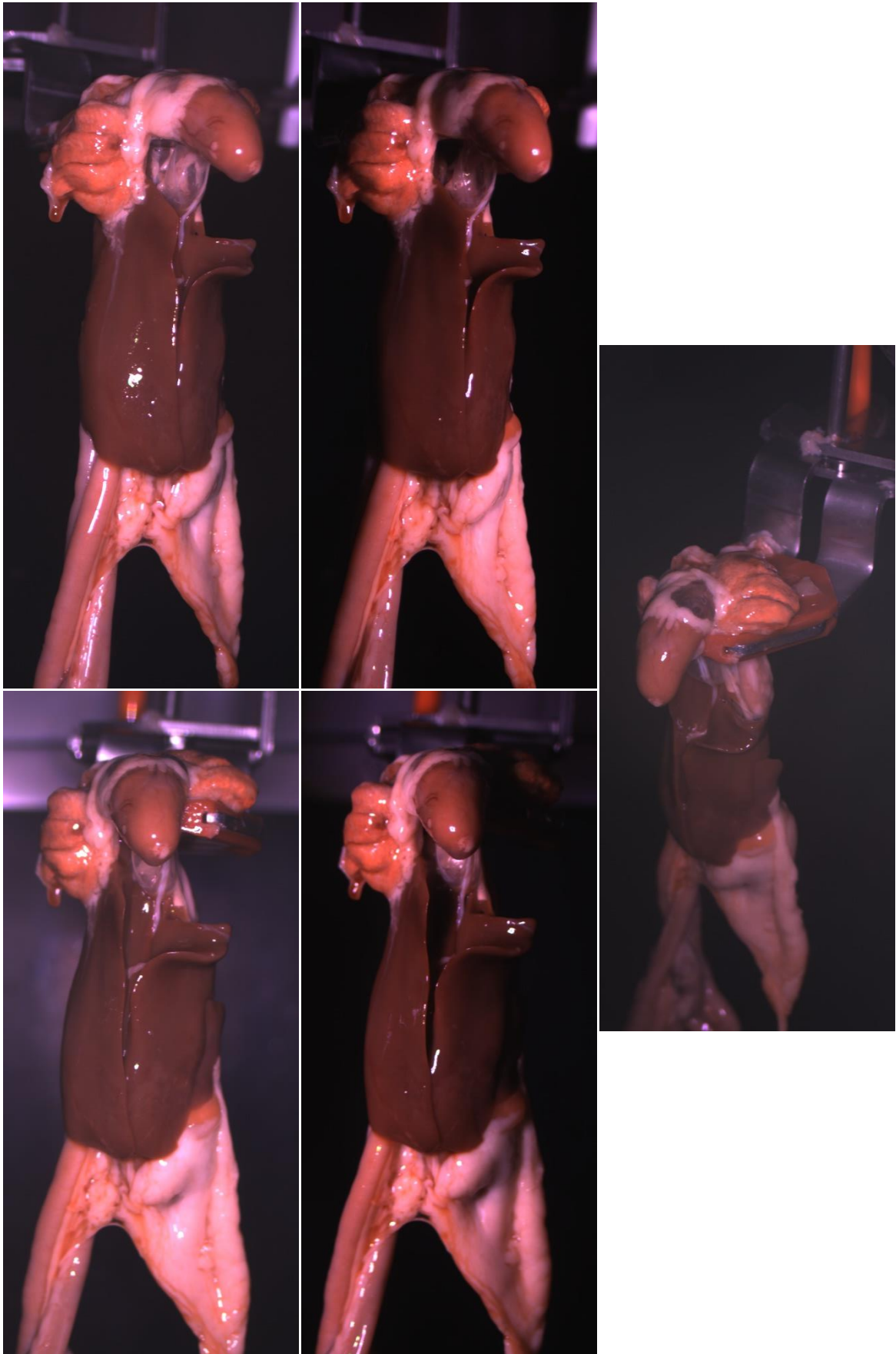
Til billedoptagelsen blev der anvendt et ophæng af tre Vision tuber med kamera og lys fra IHFood. Vision tuben indeholder kamera og led lys specielt designet ud fra forundersøgelsen [1]. De tre tuber sidder på samme side af linjen, hvor to sidder i niveau med indmadssættet for god eksponering af leveren og den sidste sidder oppe og kigger ned på krogen for eksponering af hjertet i de tilfælde hvor dette ikke kan ses godt fra de andre tuber. Tuber og billeder er navngivet henholdsvis Left, Right og Top. Se Figur 1 for illustration af forsøgsopstillingen. Ved et endeligt system vil der være tuber på begge sider af linjen. Forsøgsopstillingen er efterfølgende blevet afskærmet for at mindske lysgener til personale i området.

Da indmadssættene er meget reflekterende er det svært at undgå refleksioner i de endelige billeder. For at omgå dette problem belyses indmadssættet fra to vinkler for tuberne der kigger på leveren. Dette fjerner ikke refleksionerne, men flytter dem så man stadig kan se den fulde overflade af indmadssættet som er synlig fra tuben.

Billedoptagelsen foregik før den manuelle kontrol og en optisk sensor placeret efter den manuelle kontrol detekterede om indmadssættet var blevet taget af linjen. Det var ikke muligt at få diagnosen for de nedtagne indmadssæt, da dette ville forstyrre produktionen. Alt billedmateriale er derfor efterfølgende blev diagnosticeret af en dyrlæge. Se Figur 2 for eksempelbilleder af et indmadssæt.



Figur 1: Opstilling anvendt til billedoptagelse bestående af IHFood Vision tuber med farvekamera og lys monteret på den ene side af linjen. En sort baggrund er placeret på den anden side. Efterfølgende er der også blevet opsat afskærmning bag tuberne for at skærme for lysener.



Figur 2: Eksempel på billeder fra Left (øverst til venstre), Right (nederst til venstre) og Top tuberne (højre) af et rask indmadssæt.

Datasæt

Da forsøgsopstillingen sidder på linjen, er der adgang til en stor mængde billeder. Det detekteres automatisk om et indmadssæt er blevet taget ned af den manuelle inspektion. Så data er inddelt i "godkendt" og "til manuel inspektion af dyrlægen". For at få et godt grundlag at evaluere metoderne på er der derfor behov for at få billederne vurderet og klassificeret af nogle dyrlæger. Ønsket var at have tre dyrlæger til at gennemgå datasættet for at der kunne skabes en konsensus om hvordan de enkelte indmadssæt skulle klassificeres. Dette har kødkontrollen midlertidig ikke haft bemanning til, så billederne er blevet klassificeret af embedsdyrlæge ved kødkontrollen, Ole G. Nielsen. Det har den konsekvens at Ole alene har skulle beslutte tvivlstilfælde om hvordan sært tolkelige indmadssæt skulle klassificeres. Dette har bevirket at der ikke er et mål for usikkerhederne i datasættet, hvilket vil blive uddybet i afsnittet *Usikkerheder i den manuelle klassifikation*. Ole har selv givet udtryk for at han gerne ville gennemgå de svære eksempler igen, men det har der ikke været tid til. Det er vigtigt at understrege at Ole har ydet en kæmpe indsats under suboptimale forhold, både tidsmæssigt og i mangel af sparringspartnere.

Manuel klassifikation

Ole har inddelt datasættet i følgende sygdomstilstande, foruden rask og knyttet følgende kommentarer dertil:

- Leverstase
 - Som følge af bugvattersot
- Forandringer på leveroverfladen
 - De fleste havde blot varierende grader af fibrinbelægninger (ses ofte i forbindelse med bugvattersot). Kun enkelte havde egentlige betændelsesforandringer
- Nekrotiserende leverbetændelse
- Hjertesækbetændelse
 - Ofte kun fibrinbelægninger på overfladen og lidt mere væske i hjertesækken, ses i sammenhæng med bugvattersot. Når der er tale om tydelige belægninger er der oftest tale om kronisk hjertesækbetændelse
- Fedtlever
 - Fedtinfiltration i leveren ses i varierende grader som normalt fund. Jeg har sat enkelte på når der var tale om udprægede forandringer
- Andre forandringer
 - Ofte var der dog formodentlig tale om "*cholangiohepatitis*", men det kan være lidt vanskeligt at afgøre på billederne
- Gødningsforurening

Foruden sygdomstilstande kunne Ole også inddele i:

- Ukendt
 - Er ikke sikker på diagnose, men der er et eller andet galt
- Nonsens
 - Dette indmadssæt giver ingen mening, eks. leveren mangler
- Manuel inspektion nødvendig
 - Manuel inspektion nødvendig for korrekt diagnose

Det kan pointeres at disse kategorier er udarbejdet som følge af det Ole har set i datasættet. Så listen er måske ikke fuldstændig, samt at den ved fremtidig arbejde muligvis vil blive underinddelt eller "om inddelt" i nye kategorier.

For at blive fortrolig med data (billederne) gennemgik Ole først 155 eksemplarer som ikke er med i de efterfølgende undersøgelser, da Ole skulle se en del eksemplarer før han blev sikker på hvordan han skulle inddele datasættet. Ole har i alt klassificeret 3500 eksemplarer (hvor 155 ikke er brugt yderligere), med en fordeling som er opsummeret i Tabel 1.

	Kategorier	Indmadssæt
Nedtagne indmadssæt		2855
Rask		607
Syg		1266
Leverstase	687	
Forandringer på leveroverfladen	350	
Nekrotiserende leverbetændelse	248	
Hjertesækbetændelse	103	
Fedtlever	6	
Andre leverforandringer	2	
Gødningsforurening	9	
Sum	1405	
Ikke behandlet		982
Ukendt	177	
Nonsens	304	
Manual inspektion nødvendig	506	
Sum	987	
Ikke nedtagne indmadssæt		500
Rask		323
Syg		23
Leverstase	16	
Forandringer på leveroverfladen	5	
Nekrotiserende leverbetændelse	1	
Hjertesækbetændelse	1	
Sum	23	
Ikke behandlet		154
Nonsens	67	
Manual inspektion nødvendig	88	
Sum	155	

Tabel 1: Oversigt over embedsdyrlæge Ole G Nielsens klassifikationer af indmadssæt ud fra billedmateriale. Nedtagne indmadssæt er de sæt der er blevet nedtaget fra linjen af slagterimedarbejderne for inspektion af dyrlægerne. NB! Et indmadssæt kan have flere kategorier.

Af Oles klassifikation ses følgende interessante pointer:

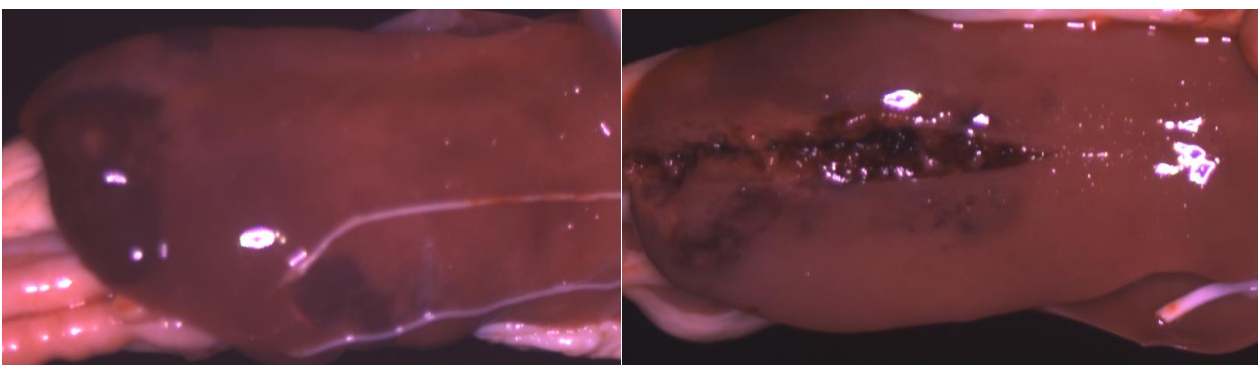
- Af de nedtagne indmadssæt er 21,3% raske (Der ses kun på fronten af indmadssættet og derfor kan dette tal godt være mindre i virkeligheden).
- Af de nedtagne indmadssæt føler Ole at han i 17,7% ikke kan klassificere eksemplaret uden at stå med det i hånden.
- Af de ikke nedtagne indmadssæt er 4,6% syge. (Der ses kun på fronten af indmadssættet og derfor kan dette tal godt være større i virkeligheden. Bemærk at tallet er beregnet ud fra en sample på 500 eksemplarer taget i et kort tidsrum samme dag, så der behæfter sig en vis usikkerhed på dette tal).

I det behandlede data er der ikke nok eksemplarer af andre leverforandringer, fedtlever og gødningsforurening til yderligere analyser og disse vil ikke blive behandlet mere i denne rapport. For "ukendt", "nonsens" og "manuel inspektion nødvendig" kategorierne vises eksempel billeder i Appendiks A, men behandles heller ikke yderligere i denne rapport.

Efterbehandling af datasæt for lever

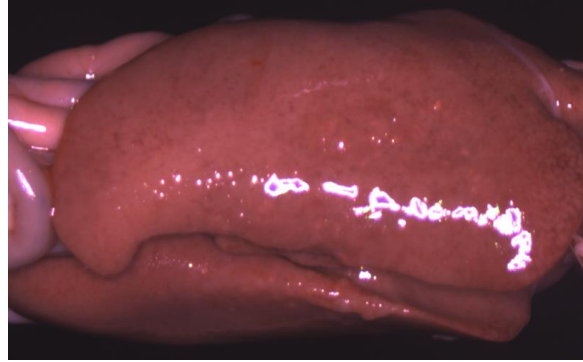
I gennem analyse af datasættet blev der fundet nogle levere der visuelt skilte sig ud. Efter samtale med Ole G. Nielsen blev det bestemt at disse sæt kunne placeres i to nye kategorier, nemlig blødninger/skader (se eksempler i Figur 3) og leverfibrose (se eksempel i Figur 4). Blødninger kan opstå i forbindelse med slagteprocessen, men kan også optræde i forbindelse med patologiske forandringer, det vil kræve yderligere vurdering af selve organet/slagtekroppen hvorvidt forandringen har nogen betydning. Leverfibrose¹ kan optræde i forbindelse med andre patologiske forandringer eksempelvis cholangiohepatitis/nekrotiserende tarmbetændelse med flere. Fra datasættet er der i alt udtrukket 115 billeder med blødninger/skader og 72 billeder med leverfibrose. Blødninger og skader ser også meget forskellige ud og bør på sigt deles i yderligere to kategorier. Det er valgt ikke at behandle disse kategorier i denne rapport.

Leversegmenteringen, beskrevet i afsnittet om *Segmentering af organer*, er trænet på et lille datasæt. Den trænedede model er brugt til at segmentere de klassificerede billeder. For at analyseresultaterne ikke kommer til at afhænge for meget af en dårlig segmentering, er billederne efterfølgende kigget igennem for at fjerne de sæt der er meget dårligt segmenteret. Der er fjernet 267 billeder fra kategorien *rask*, 93 fra *Leverstase*, 40 fra *Forandringer på leveroverfladen* og 22 fra *Nekrotiserende leverbetændelse*.



Figur 3: Eksempel på blødning og maskinskade.

¹ Hvoraf mange igen vil være såkaldt "cholangiohepatitis" (også kaldet stenlever).



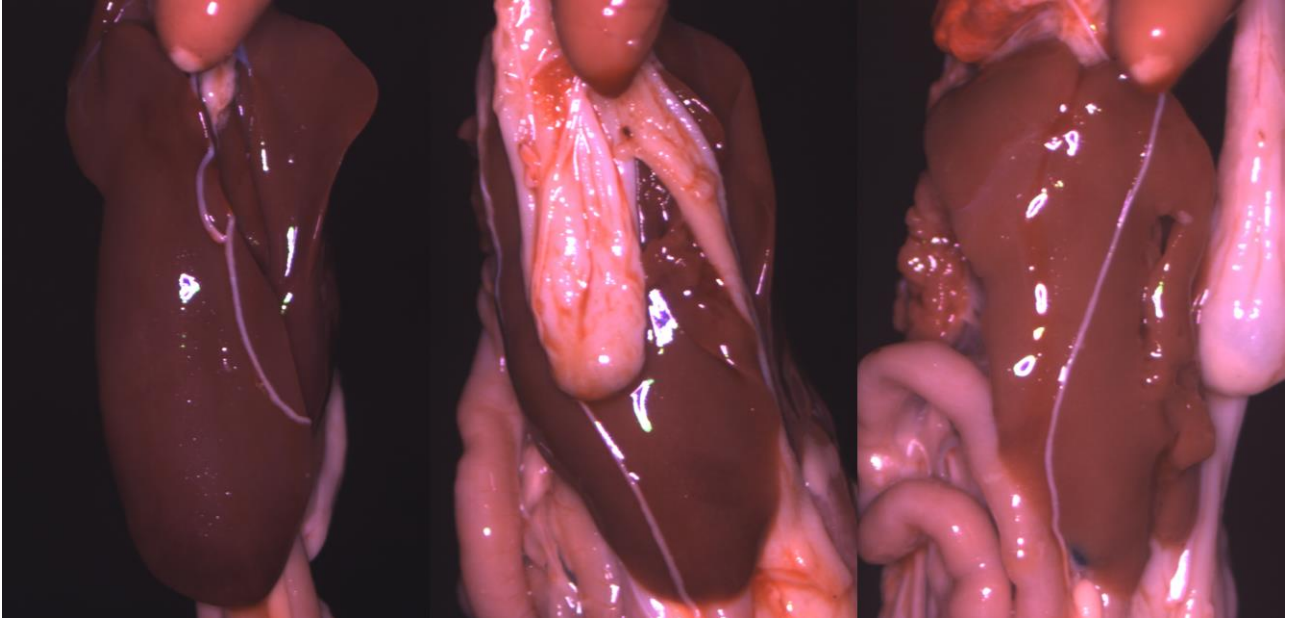
Figur 4: Eksempel på leverfibrose.

Efterbehandling af datasæt for hjerte

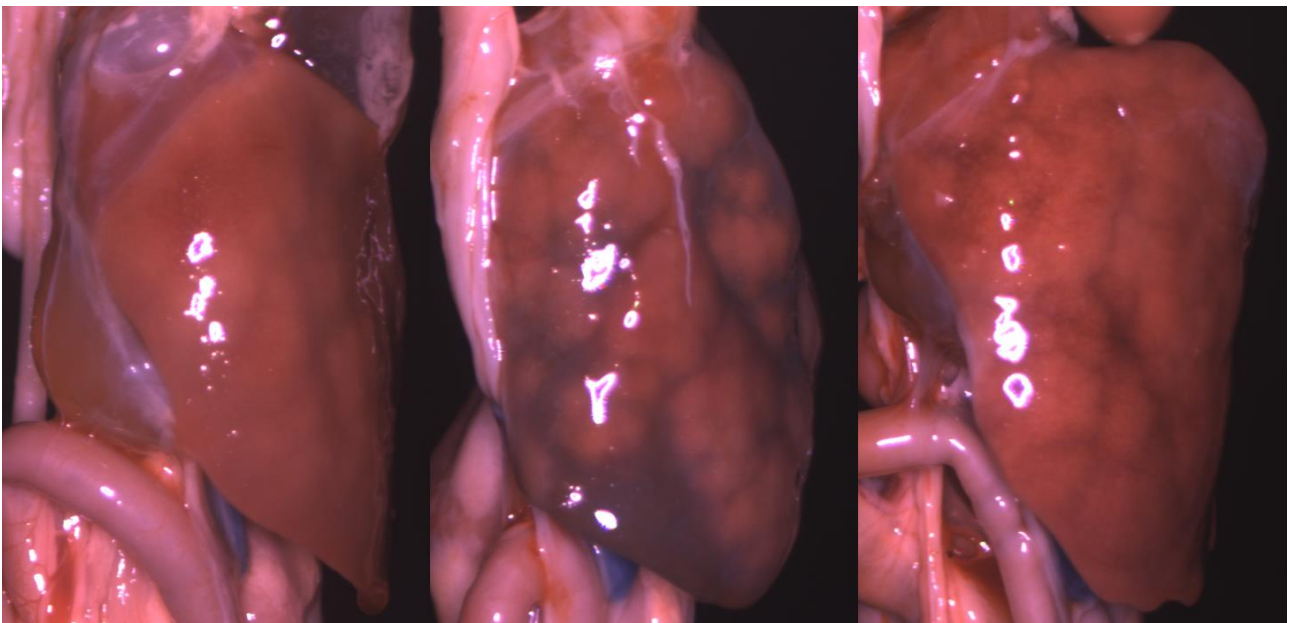
Den eneste efterbehandling af datasættet for hjertet er at de indmadssæt uden hjerte er blevet fjernet. Som følge af dette er der fjernet 8 eksemplarer.

Illustration af variationen i leverkategorier

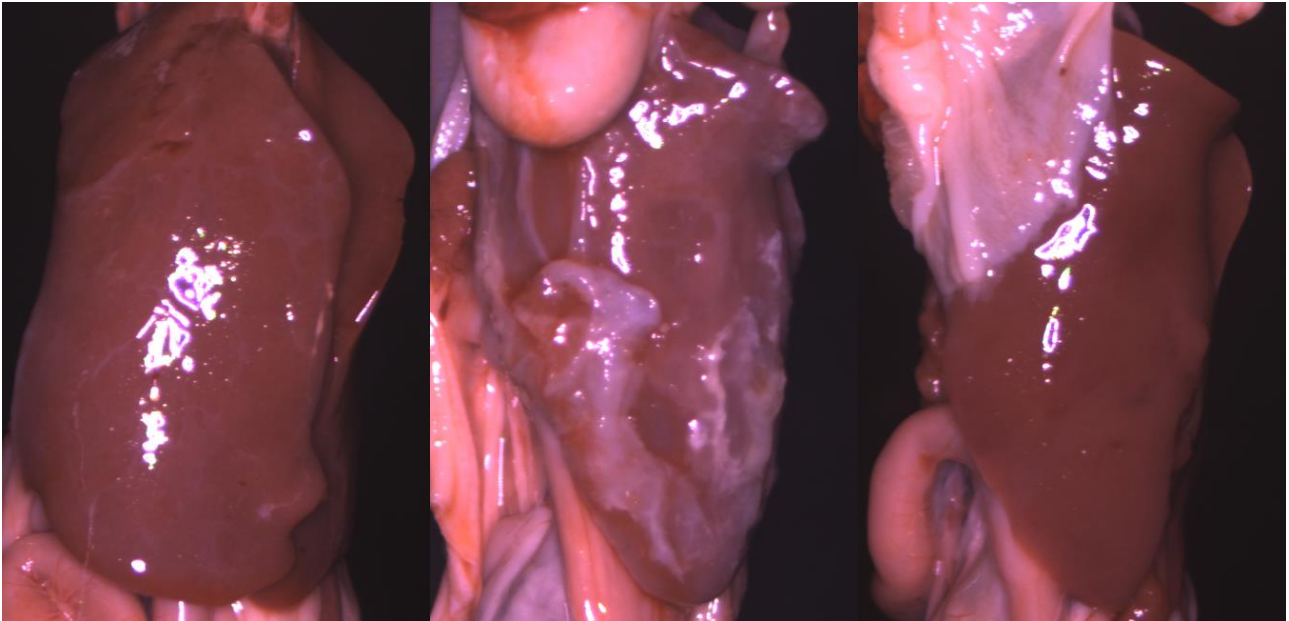
Variationen i leverkategorierne *Rask*, *Leverstase*, *Forandringer på leveroverfladen* og *Nekrotiserende leverbetændelse* ses i henholdsvis Figur 5, Figur 6, Figur 7 og Figur 8. For yderlig eksempler se Appendiks B.



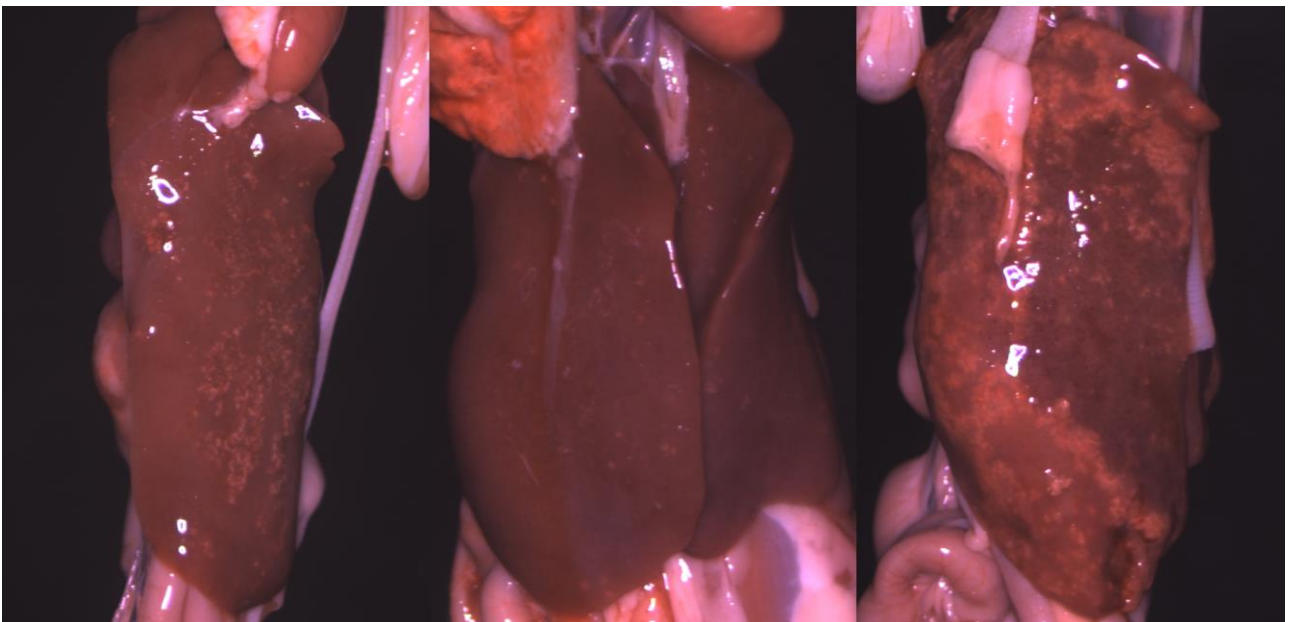
Figur 5: Eksempler fra leverkategorien: Rask.



Figur 6: Eksempler fra leverkategorien: Leverstase



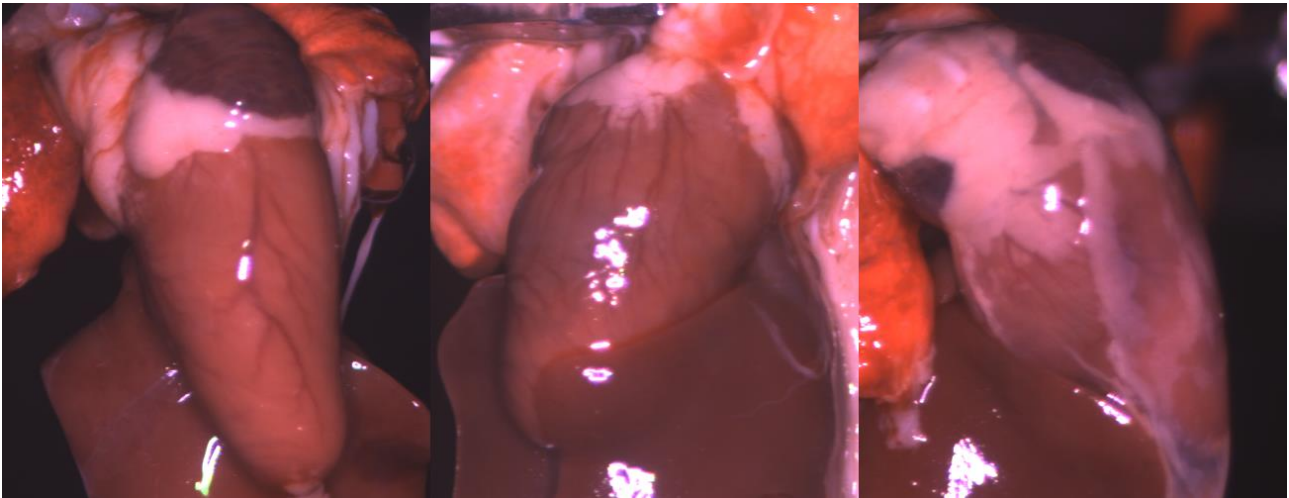
Figur 7: Eksempler fra leverkategorien: Forandringer på leveroverfladen.



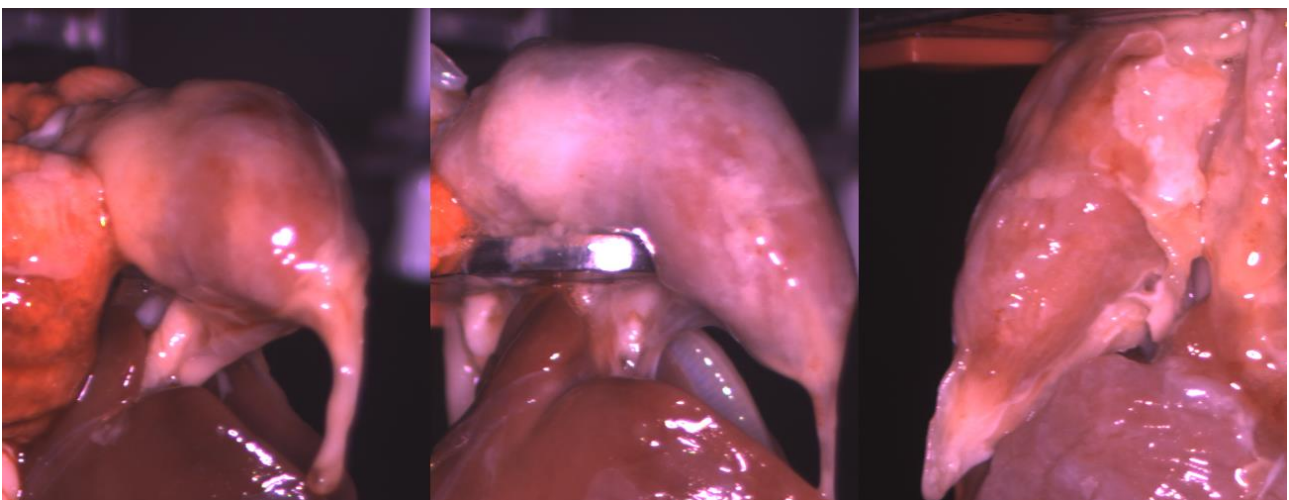
Figur 8: Eksempler fra leverkategorien: Nekrotiserende leverbetændelse.

Illustration af variationen i hjertekategorier

Variationen i hjertekategorierne *Rask* og *Hjertesækbetændelse* ses i henholdsvis Figur 9 og Figur 10. For yderligere eksempler se Appendiks B.



Figur 9: Eksempler fra hjertekategorien: Rask.



Figur 10: Eksempler fra hjertekategorien: Hjertesækbetændelse.

Metode

Vision pipeline

For systemet vil vision pipelinen ser således ud:



Denne rapport vil dog ikke gå yderligere ind i farvekalibrering og klassificering af indmadssæt ud fra flere billeder. Dette beror på den begrænsede tid der har været til rådighed for at behandle data.

Segmentering af organer

Da alle registrerede sygdomme i denne undersøgelse findes på lever og hjerte, er det disse der ønskes segmenteret. Her bør knyttes en kommentar fra embedsdyrlæge ved kødkontrollen Ole G. Nielsen "I Danmark kigger vi meget på lever og hjertes, men i princippet burde et computerbaseret system nok indrettes efter at man kigger på hele tarmsættet. F.eks. kan salmonellainfektioner medføre forandringer på milten og selvfølgelig også tarmen, men vi ser ikke salmonella hos kyllinger i DK".

Leveren segmenteres via maskinlæringsalgoritmer bl.a. "Simple Linear Iterative Clustering" og "Conditional Random Field". Dette resulterer i en maske der markerer hvor leveren befinder sig. Som det ses i Figur 11 lokaliseres leveren af algoritmen, men der kan stadig tages for meget eller for lidt med. For en overordnet beskrivelse af algoritmen se Appendix C.

Hjertet segmenteres via en modificeret "naive bayes" pixel klassifikationsalgoritme. Dette resulterer ligesom i leverens tilfælde i en maske der markerer hjertet. Se i Figur 11 for et eksempel. For en beskrivelse af algoritmen se [3].

Pixel klassificering af organer

På leveren klassificeres pixels via et specialtrænet neuralt netværk, der for den ovenstående segmentering kommer med en sandsynlighed for om denne pixel tilhører en af disse klasser: Rask, Leverstase, Forandringer på leveroverfladen, Nekrotiserende leverbetændelse. Dette samles til fire sandsynlighedsbilleder - en for hver klasse. Se eksempler i Figur 12.

Hjertesæksbetændelse kan se ud på mange måder og derfor er der på nuværende tidspunkt fulgt en strategi om kun at detektere gode hjertes. Hvis der derfor ikke findes et godt hjerte er det enten fordi hjertet ikke er der, eller det er sygt. Der pixelklassificeres derfor ikke yderligere på hjertet og de allerede fundne segmenteringer benyttes direkte.

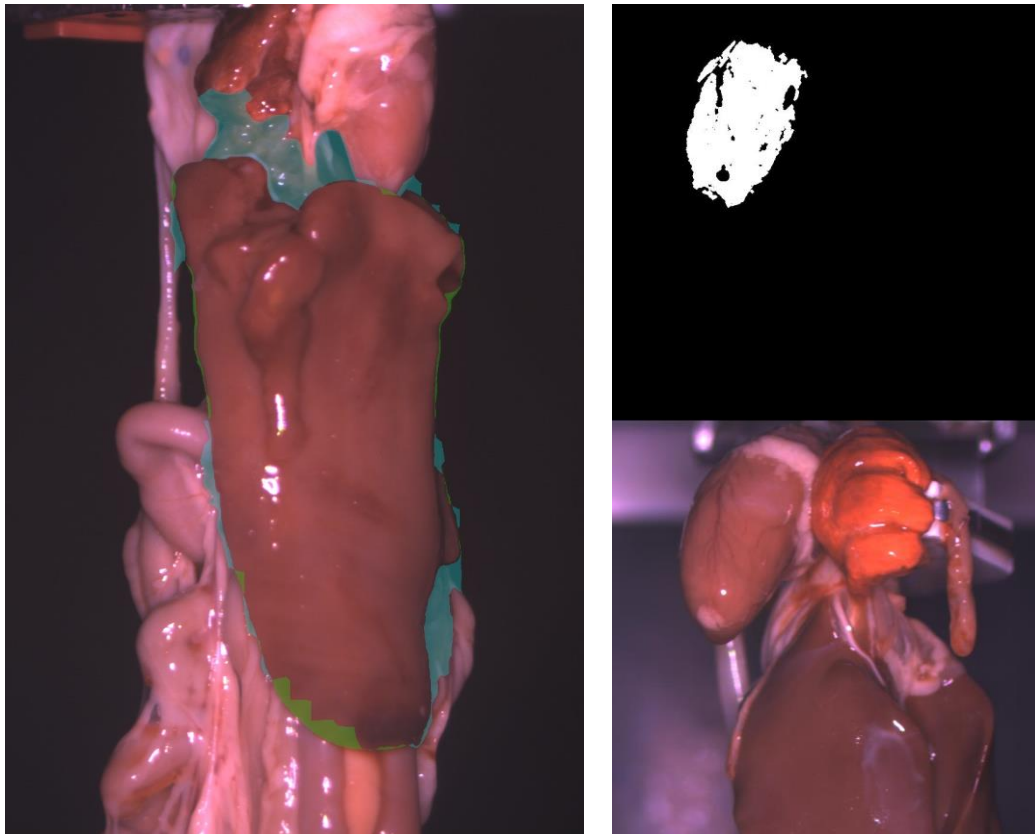
Klassificering af organer per billede

For leveren udtrækkes features fra de fire sandsynlighedsbilleder. Det er tal som hvor mange pixels de indeholder, hvor høje pixelværdierne er osv. Disse features puttes ind i en "Binomial Gradient Boost Classifier" der klassificerer hvert billedsæt og altså kommer med den samlede klassifikation.

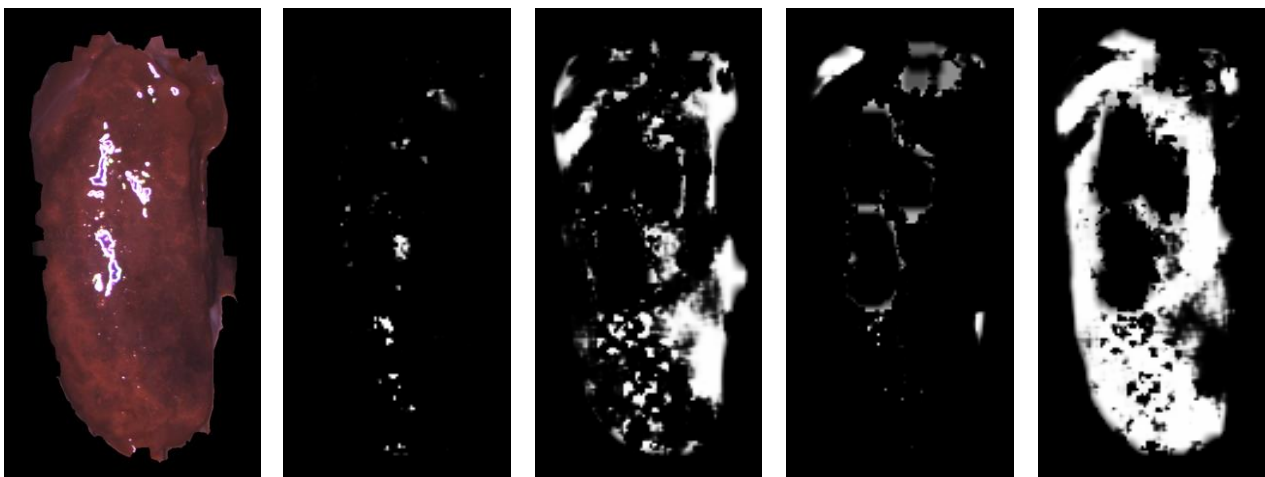
Hjertet behandles på en lignende måde hvor features udstrækkes fra segmenteringen. Til klassificering bruges en "Linear Discriminant Analysis Classifier".

En beskrivelse af hvordan en lever behandles i vision pipelinen er beskrevet i Appendix D. De endelige klassificeringerne er beskrevet i afsnittet

Resultater.



Figur 11: Eksempel på segmentering af leveren (venstre) og hjertet (højre). På leveren er også fejlsegmenteringerne markeret med grøn for lever der ikke er segmenteret og cyan for oversegmenteringer.



Figur 12: Eksempel de fire sandsynlighedsbilleder for en lever med leverbetændelse. Billederne er fra venstre: Segmenteret lever, Risk, Leverstase, Forandringer på leveroverfladen, Nekrotiserende leverbetændelse.

Validering

For at sikre at performance resultaterne bliver rapporteret for det generelle tilfælde, er det vigtigt at der opdeles i trænings- og testsæt. Et trænings sæt bruges som grundlag til at beregne de statistiske modeller. Efterfølgende anvendes testsættet så til at beregne, hvor korrekt modellerne kan skelne syg fra rask på ny data.

For den sidste del af analyserne i *"Klassificering af organer per billede"* er der yderligere benyttet *"stratified Five-fold Cross-validation"*. Dette sikrer yderligere at der ikke ved et tilfælde er valgt *"nemme"* eksempler i testsættet da alle eksempler skiftevis er i testsættet. Det foregår efter følgende procedure:

1. Datasættet opdeles i fem grupper, hvor der tilstræbes at fordelingen af sygdomme i hver gruppe minder om fordelingen i det samlede datasæt
2. Første gruppe tages ud som testsæt til evaluering, og modellen trænes på de resterende fire grupper
3. Trin 2 gentages for alle fem grupper, og giver derved en samlet evaluering for hele sættet uden at det enkelte indmadssæt er brugt til træning og test samtidig

Resultater

Ved brug af de beskrevne billedanalysemetoder er der foretaget en klassificering for hver sygdom. For at have et sammenligningsgrundlag til foranalyserapporten om den udvendige kontrol [2] laves de samme opdelinger i de tre udsorteringer; godkendt, manuel kontrol og kassation sidst i afsnittet.

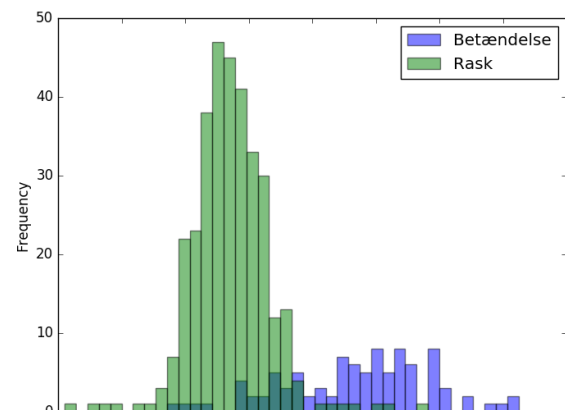
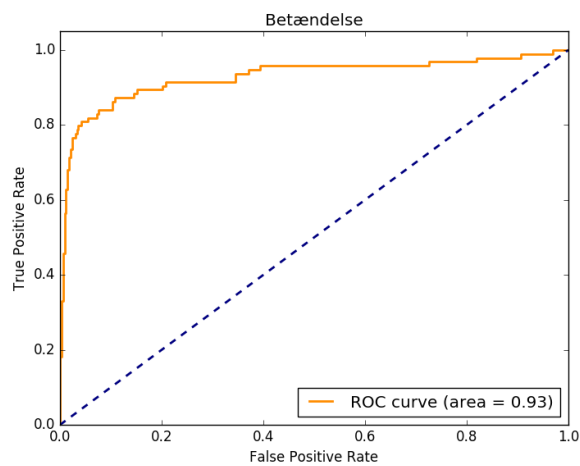
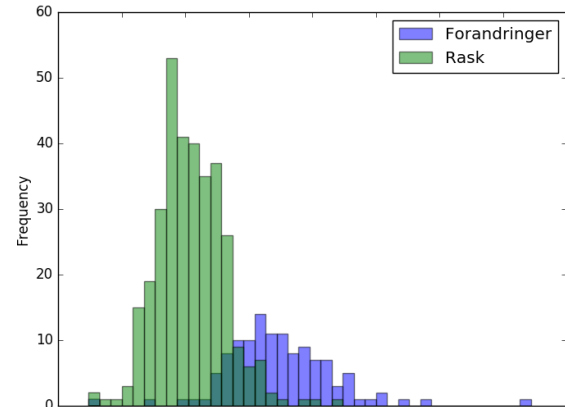
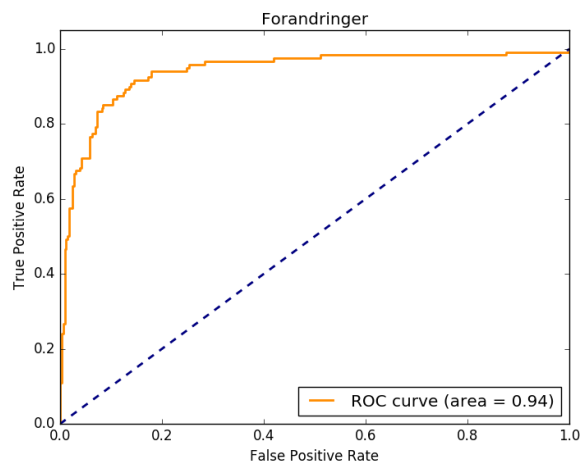
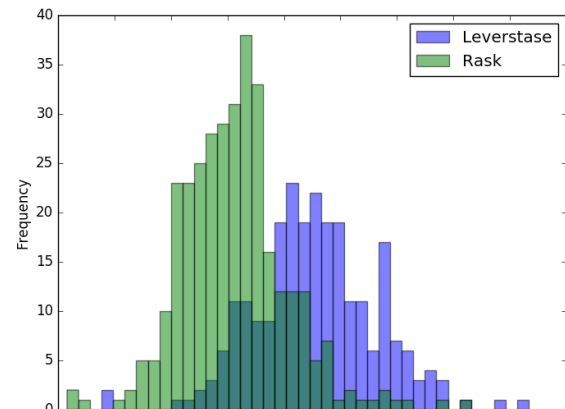
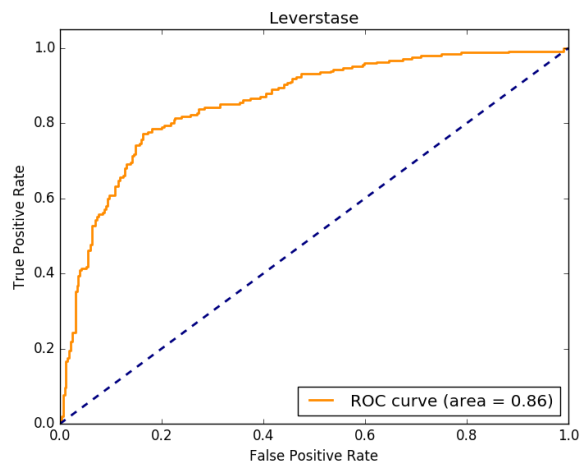
Lever

De samlede resultater for den bedste opdeling i leverklasserne er samlet i en *"Confusion matrix"*, se Tabel 2, der beskriver den samlede performance for hele vision systemet mht. leversygdomme.

		Predictions					True Positive
		Rask	Leverstase	Forandringer	Betændelse		
Targets	Rask	256	55	18	2	331	77,3%
	Leverstase	63	176	4	4	247	71,3%
	Forandringer	34	8	77	1	120	64,2%
	Betændelse	17	12	4	61	94	64,9%
		370	251	103	68	570	
						Total	72,0%

Tabel 2: Confusion matrix for klassifikation af leversygdomme. Tal i diagonalen (mørkegrønne felter) er korrekt klassificeret, tal i de lysegrønne felter er syge klassificeret til en anden sygdom og tal i de gule felter er forkert klassificeret.

For mere præcist at beskrive det trade-off der er mellem sande og falske positive i et klassificeringsproblem kan man benytte *"Receiver operating characteristic"* (ROC) kurver. Disse beskriver hvor mange procent raske skal udsorteres for at få en vis mængde syge udsorteret. Ligeledes kan sværhedsgraden af klassifikationsproblemet vises med klasse histogrammer der ved større og større overlap bliver et sværere og sværere problem. ROC kurver og tilhørende klassehistogrammer vises i Figur 13 for hver af de tre leversygdomme.



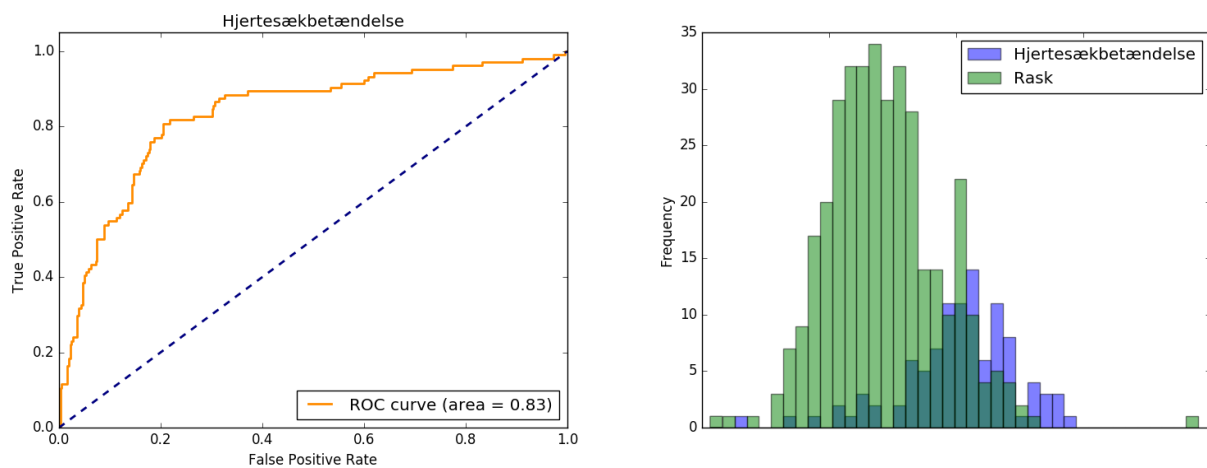
Figur 13: Receiver operating characteristic kurver og klassehistogrammer for hver af de tre leversygdomme: Leverstase, Forandringer på leveroverfladen og Nekrotiserende leverbetændelse. ROC kurverne beskriver den trade-off der er mellem sande og falske positive. Klassehistogrammer viser overlappet mellem rask og den aktuelle leversygdom.

Hjerte

Ligesom for lever gives resultaterne for hjerte i en "confusion matrix" i Tabel 3 og en ROC kurve med klassehistogram i Figur 14. I modsætning til lever laves der ingen manuelle rettelser til hjertesegmenteringerne, hvilket giver mere støj og også kan ses i sammenligninger af klasse histogrammerne.

		Predictions			
		Rask	Betændelse		
Targets	Rask	314	48	362	True Positive 86,7%
	Betændelse	44	60	104	57,7%
		358	108	374	
		Total			80,3%

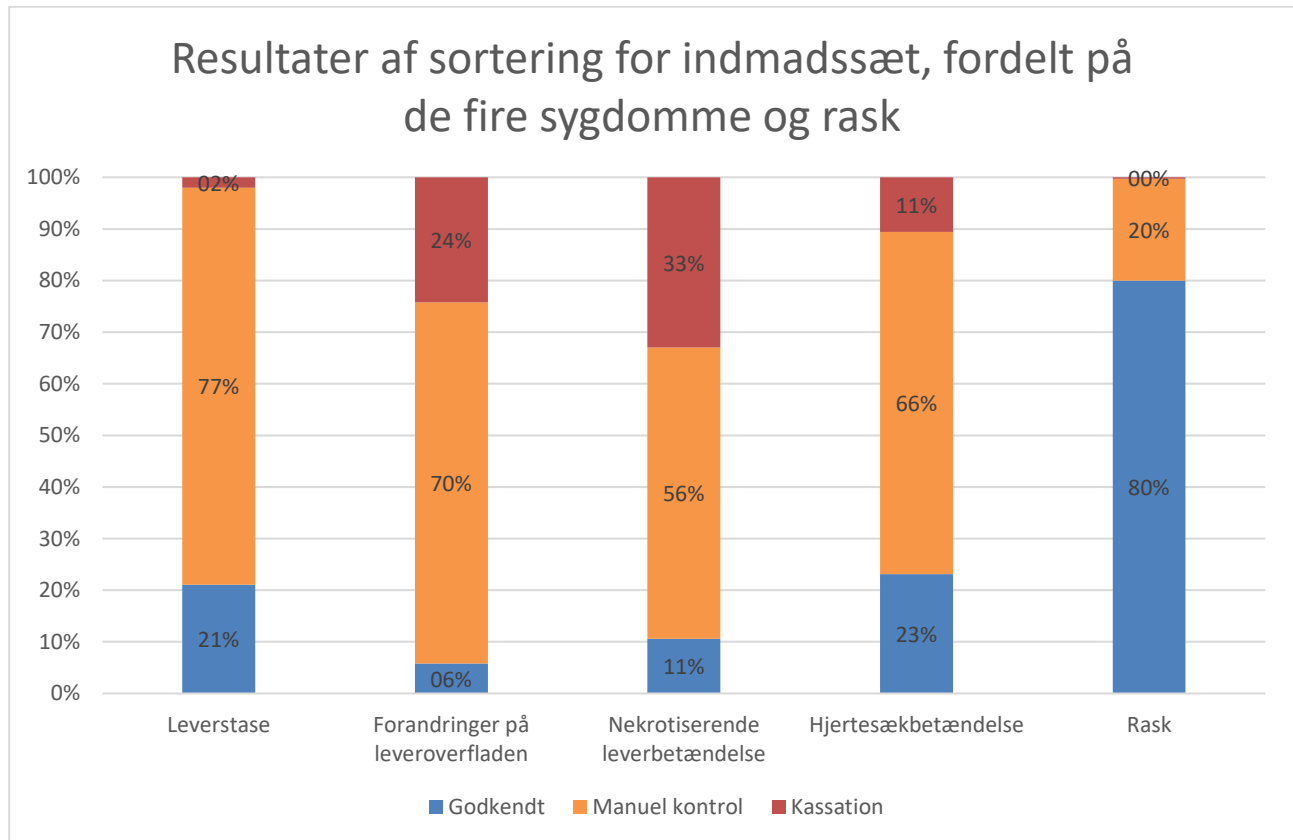
Tabel 3: Confusion matrix for klassifikation af hjertesækbetændelse. Tal i diagonalen (mørkegrønne felter) er korrekt klassificeret og tal i de gule felter er forkert klassificeret.



Figur 14: Receiver operating characteristic kurve og klassehistogram for hjertesækbetændelse. ROC kurven beskriver den trade-off der er mellem sande og falske positive. Klassehistogrammer viser overlappet mellem rask og hjertesækbetændelse.

Sammenligningsgrundlag til foranalyserapporten

For at have et sammenligningsgrundlag til foranalyserapporten om den udvendige kontrol [2], laves de samme opdelinger i de tre udsorteringer; godkendt, manuel kontrol og kassation. For at dette nemmest kan sammenlignes, gennemtvinges den samme fordeling af raske indmadssæt som raske fugle i [2] hvor der udtages ~20% af raske fugle til manuel kontrol og kun kasseres nogle få promiller. Resultater for denne sortering af indmadssæt ses i Figur 15.

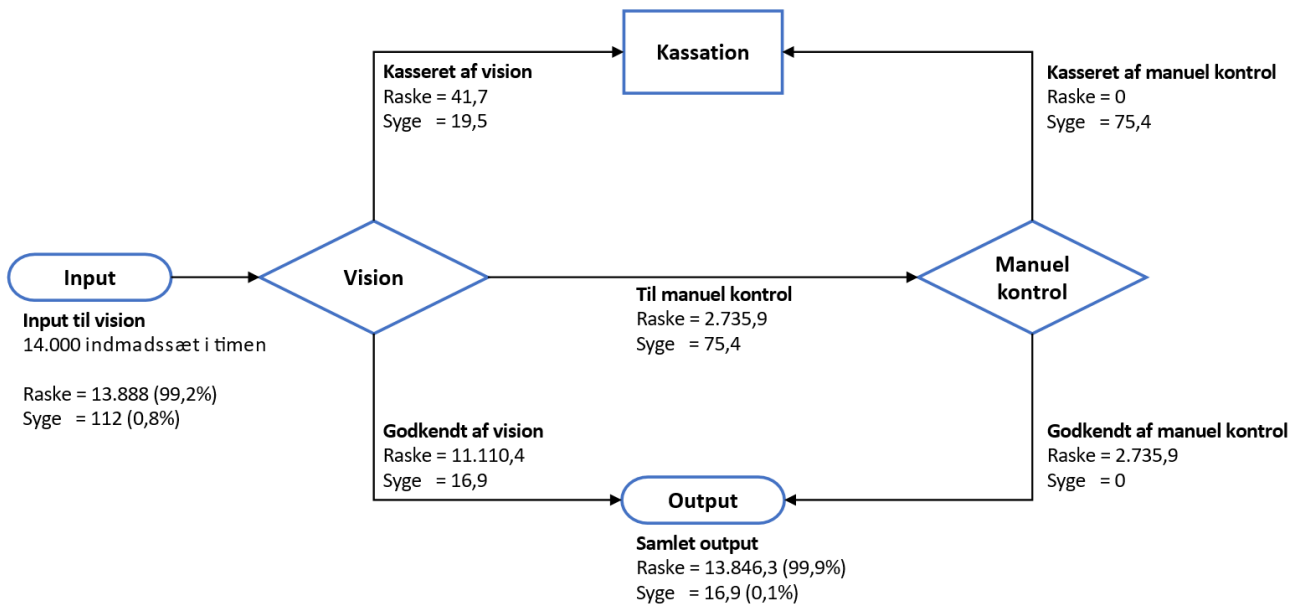


Figur 15: Diagram over sorteringen med samme forudsætninger for raske indmadssæt som i foranalyserapporten om den udvendige kontrol [2].

Case omkring forventet flow i produktion

Med udgangspunkt i en case, der minder om en normal produktion på slagteri, kan der beregnes et forventet flow. Produktionshastigheden sættes til 14.000 indmadssæt i timen, og der antages at være 0,8% syge indmadssæt ligeligt fordelt mellem de tre leversygdomme og den ene hjertesygdom.

Ud fra sorteringsresultaterne beskrevet i forrige afsnit er der lavet udregninger for estimeret flow som vist i Figur 16. Detaljer omkring udregningerne kan ses i Appendiks E.



Figur 16: Estimeret flow i produktionen baseret på de i nuværende analyser.

Vurdering af udstyr

Billederne er i passende opløsning til at kunne se det nødvendige og der er muligheder for at gå ned i opløsning. Farvebilleder er et klart krav for at kunne se de forskellige indikatorer der bruges til at stille diagnoserne, dette er også allerede behandlet i [1]. Den anvendte lukketid på 0.75 ms er passende hurtigt til at give skarpe billeder og undgå mærkbar "motion blur" – også i produktion med 14.000 indmadssæt i timen.

De anvendte kameravinkler kan muligvis optimeres yderligere, men nuværende opsætning er et godt startpunkt. Det er ikke blevet undersøgt hvor mange kameraer der er behov for på den anden side af linjen (alle nuværende kameraer hænger på samme side og ser derved lidt over halvdelen af indmadssættet). Det forventes ikke at der skal bruges mere end to kamera på den anden side af linjen.

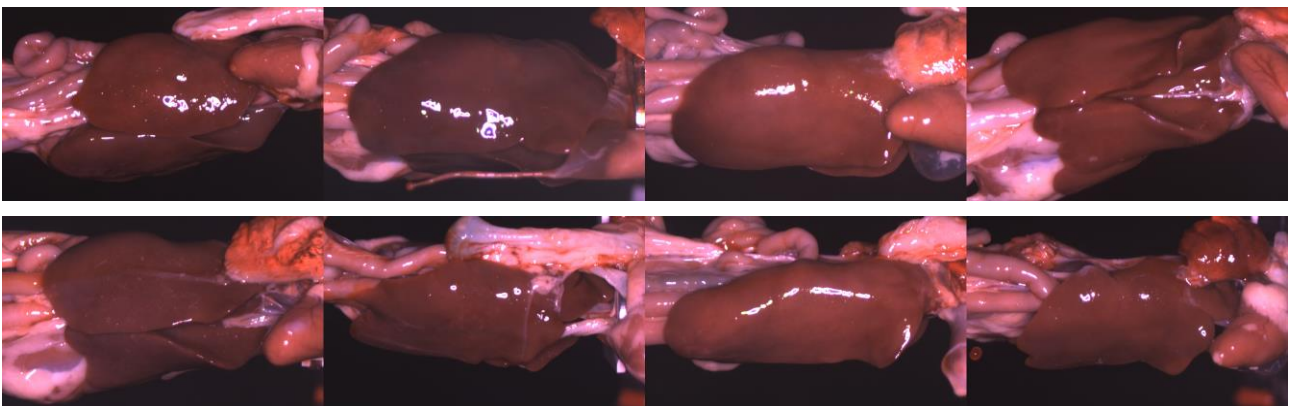
Fordelen ved de anvendte vision tubes er den kompakte form, der nemt tillader at de placeres i forskellige vinkler, samt at der kan placeres flere kameraer ved samme inspektionspunkt på kæden. Et forbedringspunkt med den nuværende opsætning er belysningen. Da indmadssættene er meget reflekserive er det måske ikke optimalt med lys indbygget i tuben og alternativer bør undersøges nærmere.

Diskussion

Usikkerheder i den manuelle klassifikation

AI analyse i denne rapport er blevet fortaget ud fra den klassifikation fortaget af embedsdyrlæge ved kødkontrollen, Ole G. Nielsen. Ole er bl.a. involveret i kødkontrollens arbejde med at uddanne slagterimedarbejdere og kvalitetssupervisere fødevarerstyrelsens personale (dyrlæger og teknikere) til at inspicere kyllingekroppe og indmadssæt konsistent, således at der mellem de enkelte medarbejdere på et slagteri og mellem forskellige slagterier inspiceres på samme måde og der kasseres så ens som muligt. Der er dog stadig en variation i hvornår den enkelte medarbejder kasserer en krop eller et indmadssæt. Med andre ord er der en del af indmadssættene hvor der formodeligt ikke er konsensus om hvordan de skal klassificeres. Omfanget af dette er ukendt, men bør estimeres ved at sammenholde flere dyrlægers vurderinger, således at det kan fastslås hvordan et endeligt system bør performe.

En kravspecifikation til systemet bør indeholde overvejelser om hvad der skal gøres med en evt. del af populationen hvor der ikke er menneskelig konsensus. For at illustrere at menneskelig konsensus kan være svær at opnå se Figur 17, hvor det illustreres at det kan være svært at bestemme hvor skillelinjen mellem syg og rask går. Yderligere, i Appendiks F vises to collager over leverstase og nekrotiserende leverbetændelse fra klart sygt til rask.



Figur 17: Eksempler på leverer hvor Ole G. Nielsen har vurderet de tre til venstre som syge og den til højre som rask, for hver række. Øverste række er af sygdommen Leverstase, og nederste række er af sygdommen nekrotiserende leverbetændelse.

Usikkerheder i segmenteringen af organer

I modsætning til den udvendige kontrol er segmenteringen på den indvendige kontrol markant sværere. Den biologiske variabilitet i formen på organerne er stor og sammenholdt med at placeringen af organerne i indmadssættet også har en stor variabilitet gøres problemet endnu sværere. De to primære organer der er vigtige at segmentere er lever og hjerte. Ved de undersøgelser der er fortaget virker det som om leveren er nemmere at segmentere end hjertet.

En bedre segmentering er vigtig for at forbedre systemets overordnede performance. I dette arbejde er de værste fejlsegmenteringer for leveren fjernet. Disse er ikke fjernet for hjertet, hvilket resulterer i en dårligere performance og understreger vigtigheden at der arbejdes på at forbedre segmenteringen af organerne.

Resultater og test af systemet

Resultaterne viser potentiale for muligheden at benytte vision som et værktøj i kødkontrollen. Dog er resultaterne ikke gode nok til hvad der er ønskeligt af et endeligt system.

På den tekniske side skal alle billederne af det enkelte indmadssæt indgå i analyserne, hvilket af tidsmæssige årsager har det ikke været muligt i denne rapport. Segmenteringen skal som allerede beskrevet forbedres. Ydermere kan meget formodentligt forbedres i udtagelsen af billedfeatures der benyttes af mønstergenkendelsesalgoritmerne. De billeder der ikke er taget i betragtning i forbindelse med analyserne skal håndteres og inkluderes i beregningerne. Her tænkes på blødninger/skader og leverfibrose på leveren.

I forbindelse med test af systemet bør det afklares hvor tungt grænsetilfælde bør veje. Er det lige slemt ikke at sortere meget syge som mindre syge indmadssæt fra?

Andre åbne spørgsmål

I forbindelse med arbejdet i denne rapport er der formuleret disse åbne spørgsmål, som der fremadrettet også skal arbejdes med:

- Hvad skal der ske hvis, der intet hjerte er i indmadssættet? (Det sidder i kroppen eller det er faldet af eller det sidder skjult i indmadssættet)
- Hvad skal der ske hvis, der ingen lever er i indmadssættet?
- Hvad skal der ske hvis, noget af leveren er revet af? (Sygdom kunne jo befinde sig der)
- Er det eneste kriterie, til om man skal kassere et indmadssæt at det sygt? Kan der være eksempler hvor man syntes at indmadssættet er så forskelligt fra "normalen" at man ikke ønsker at processere det selvom det ikke udgør en sundhedsrisiko?
- Der er ikke kigget på realtidsaspektet i denne rapport (Altså kan alle beregninger nås på ca. 250 ms i snit?)
- Ved leverbetændelse kan et meget lille område af leveren være betændt. Dette åbner for støj hvis denne lille del misses - hvad vil konsekvensen af dette være?
- Er der variation der ikke er set endnu? Har foder for eksempel betydning for udseendet og farve af indmadssættet?
- Hvordan skal sygdomme med meget lav frekvens håndteres? F.eks. sås der kun to eksempler på leverforandringer i de 3500 eksempler der blev behandlet
- Det er uafklaret hvad der "gemmer" sig i de eksemplarer der er i kategorien "manuel inspektion nødvendig". Her skal der først indsamles de korrekte diagnoser direkte fra dyrlægerne ved linjen, så der fremadrettet kan arbejdes videre med disse eksemplarer

Konklusion

Denne undersøgelse har beskæftiget sig med tre leversygdomme og en hjertesygdom. Disse sygdomme er valgt ud fra at vi minimum har 100 eksempler af hver i datasættet. Det har til en vis grænse vist sig muligt at kunne detektere de ønskede sygdomme ved brug af statistisk billedanalyse. Dog er der et stykke vej før analyserne er gode nok til at kunne bruges i et system i produktion. I en case omkring forventet flow i produktion, estimeres det at der er 0,1 % der bør blive sorteret fra som ikke gør det. Samtidigt sendes 20% af raske indmadssæt til manuel kontrol, hvilke må betragtes som for mange.

Alle analyser beror på billeder fra en enkelt vinkel, hvilket vil sige der kun bruges en 1/3 del af den data der er til rådighed på nuværende tidspunkt. Det har ikke været muligt tidsmæssigt at tage mere med i denne rapport, men det er et af de første steder at starte for at forbedre systemet. Bland andet kan nævnes at hjertet nogen gange ses bedst forfra og nogen gange bedst oppe fra fordi hjertet "ligger" på toppen af kroge. Nå vi som i denne rapport kun vælger én betragtningsvinkel vil dette gå ud over performance. Alt i alt bør de rapporterede performance målinger ses som et startpunkt - der kan blive markant forbedret.

Udover de tekniske udfordringer er der også blevet afdækket en del forståelsesmæssige udfordringer, som skal behandles yderligere for at kravspecifikationen kan gøres færdig. Her tænkes specielt på de usikkerheder der er i den manuelle klassifikation. Indtil nu har det kun været muligt at få klassificeret det brugte datasæt af én dyrlæge. Det virker på forfatterne som om at der er en del af populationen hvor der ikke er en konsensus om hvordan de skal klassificeres. På nuværende tidspunkt er omfanget af denne del af populationen ukendt, men bør estimeres ved at sammenholde flere dyrlægers vurderinger.

Lidt forenklet kan man sige at kravspecifikationen skal tage stilling til hvordan systemet skal håndtere at eksperterne (dyrlægerne) kan være uenige eller rettere ikke kan klassificere på præcis den samme måde. Ydermere hvis kravspecifikationen lægger op til meget lave fejlratere kræver det at systemet testes på eksempler som dyrlægerne kan klassificere meget konsekvent.

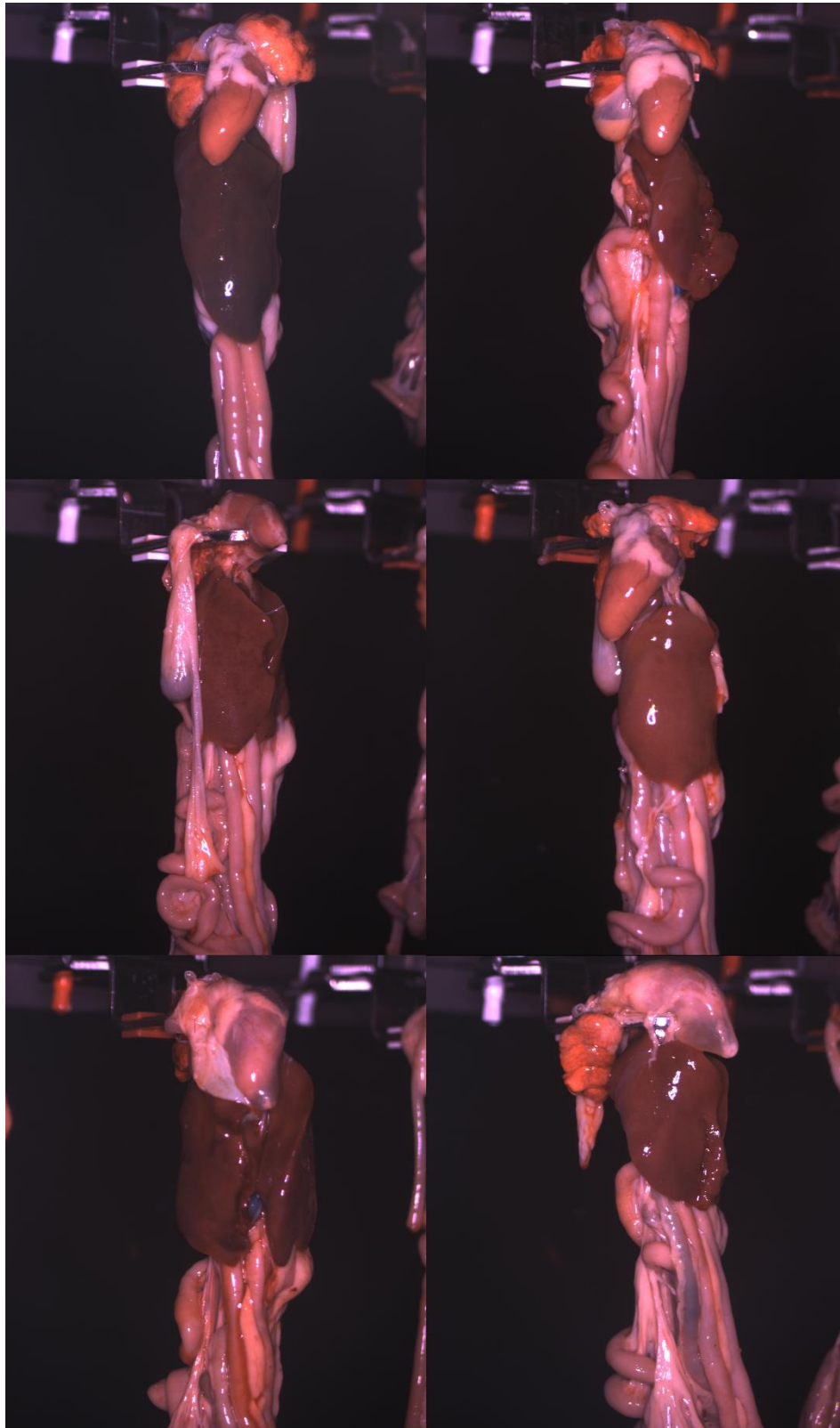
Samlet set er forfatterne optimistiske for at man med mere arbejde kan opnå en performance der er tilfredsstillende. Her er det vigtigt at understrege at for at det skal være muligt, er der både tekniske udfordringer der skal løses, samt det skal afdækkes hvordan performance skal forstås (og måles) før det giver mening som et værktøj i veterinærkontrollen på slagterierne.

Referencer

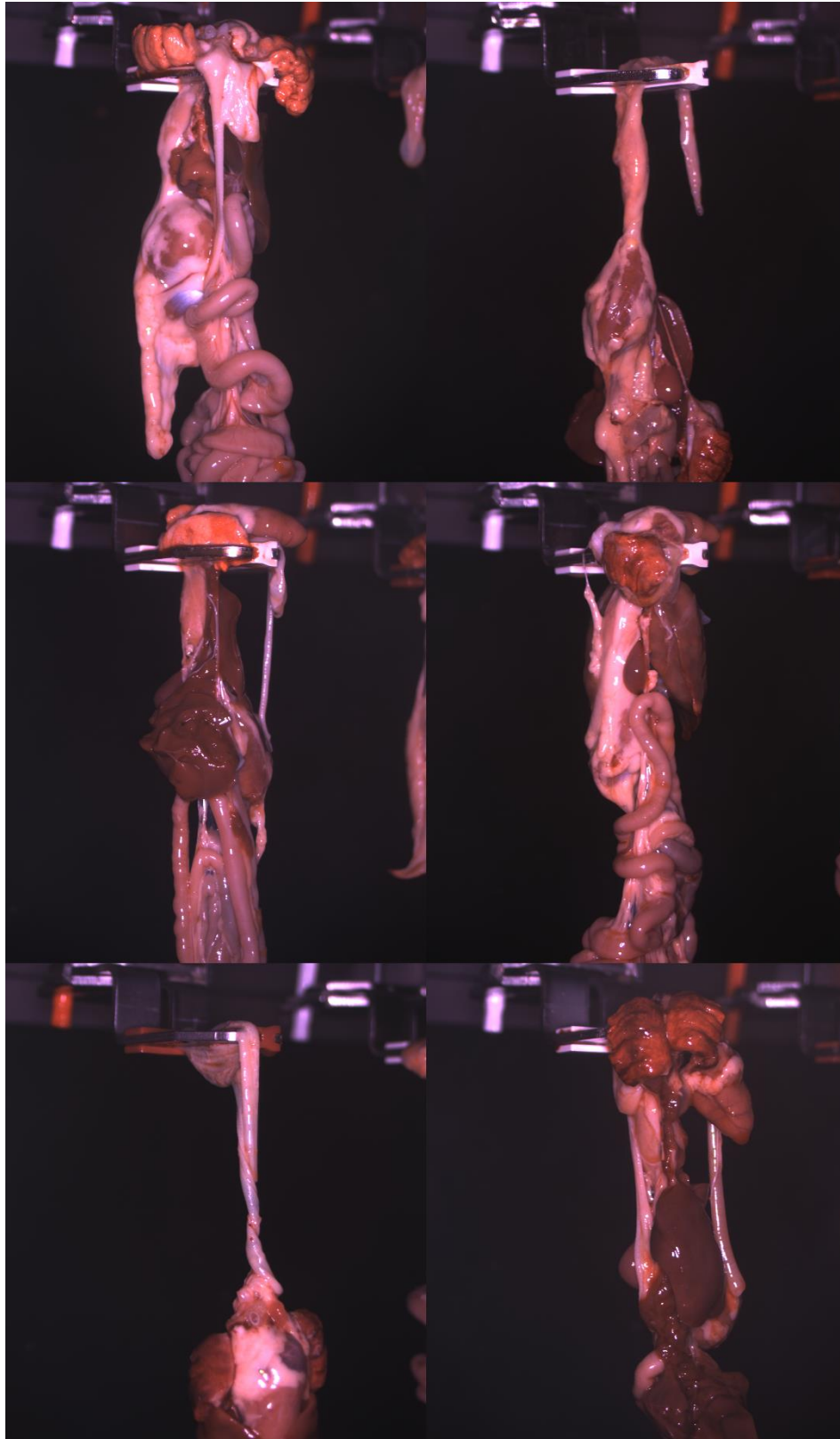
- [1] A. Jørgensen, »Forundersøgelse for inspektion af indmadssæt vha. computer vision,« 2016.
- [2] E. M. Jensen, "Forundersøgelse ved brug af billedanalyse til inspektion af sygdomme ved udvendig kødkontrol på kyllingslagteri," 2014.
- [3] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning, Springer, 2010.

Appendiks A – Illustrationer af kategorier ikke behandlet i rapporten

I det følgende vises illustrationer af kategorierne "ukendt", "nonsens" og "manuel inspektion nødvendig" som ikke bliver behandlet yderligere i rapporten, i henholdsvis Figur 18, Figur 19 og Figur 20.



Figur 18: Eksempler fra kategorien ukendt, som skal forstås at selvom dyrlægen har disse i hånden er en diagnose måske ikke mulig.



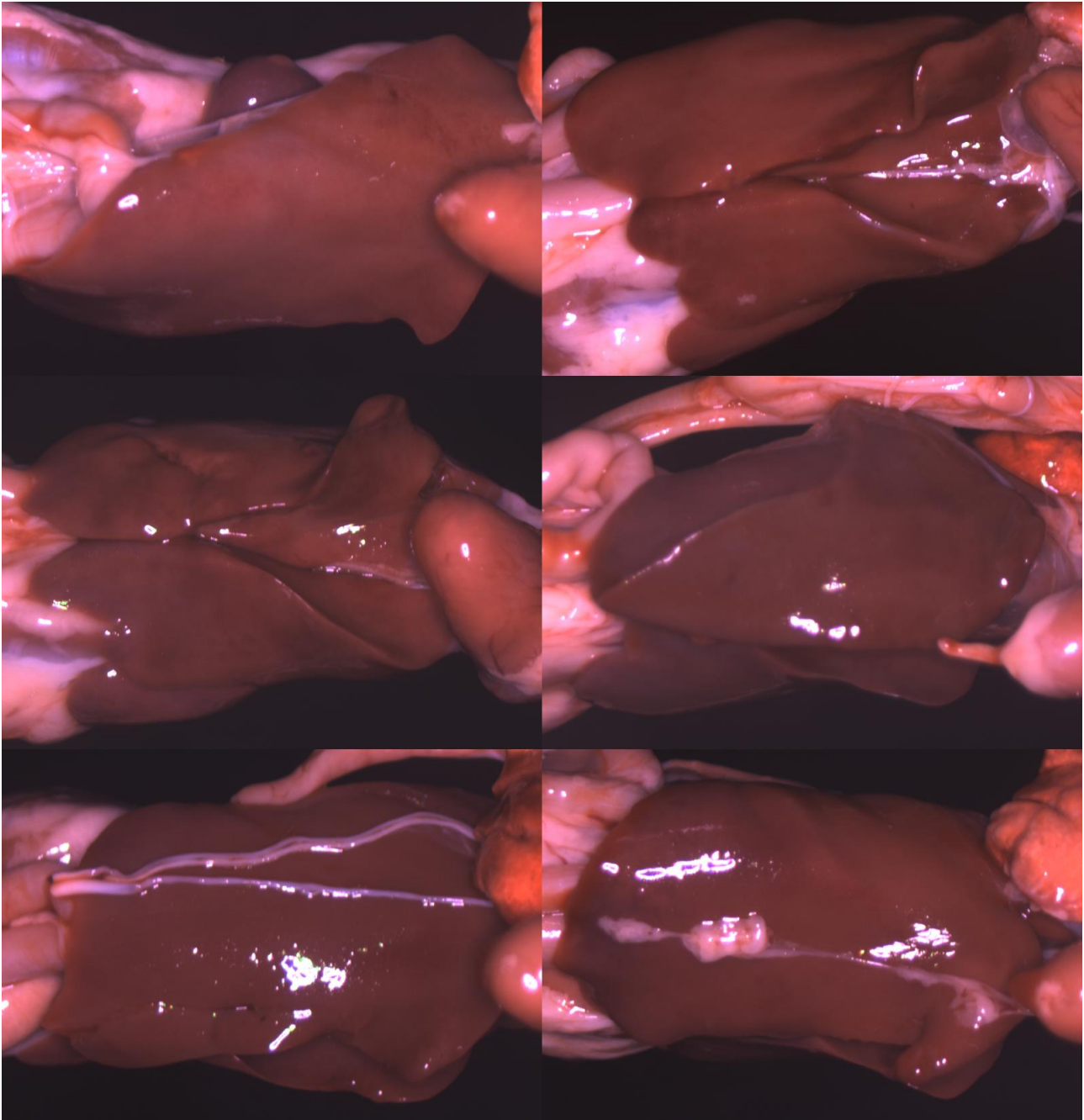
Figur 19: Eksempler fra kategorien nonsens, som skal forstås som eksemplarer hvor analyse ikke giver mening.



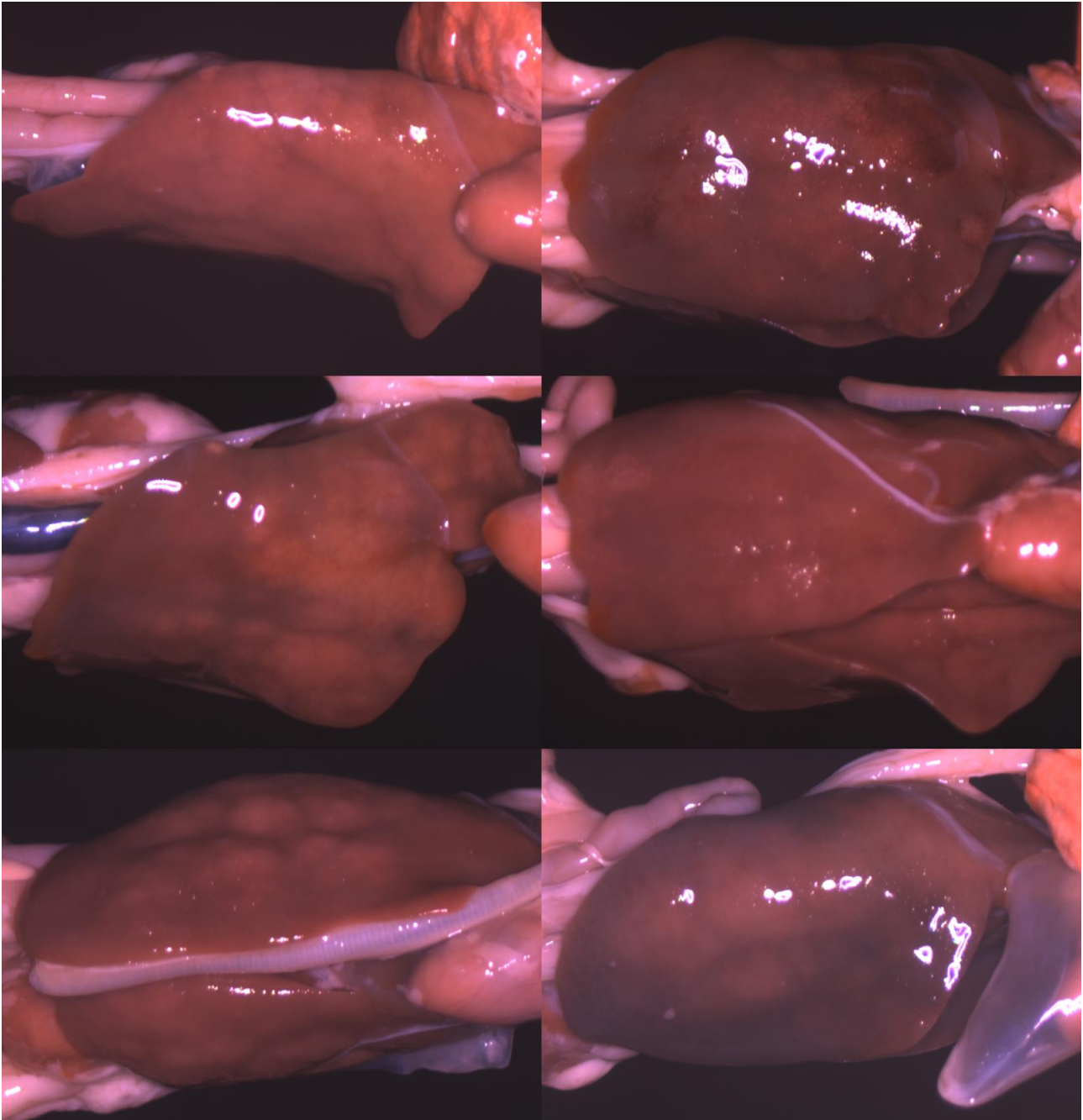
Figur 20: Eksempler fra kategorien manuel inspektion nødvendig, som skal forstås som dyrlægen ud fra billederne ikke føler en diagnose kan stilles.

Appendiks B – Illustrationer af kategorier behandlet i rapporten

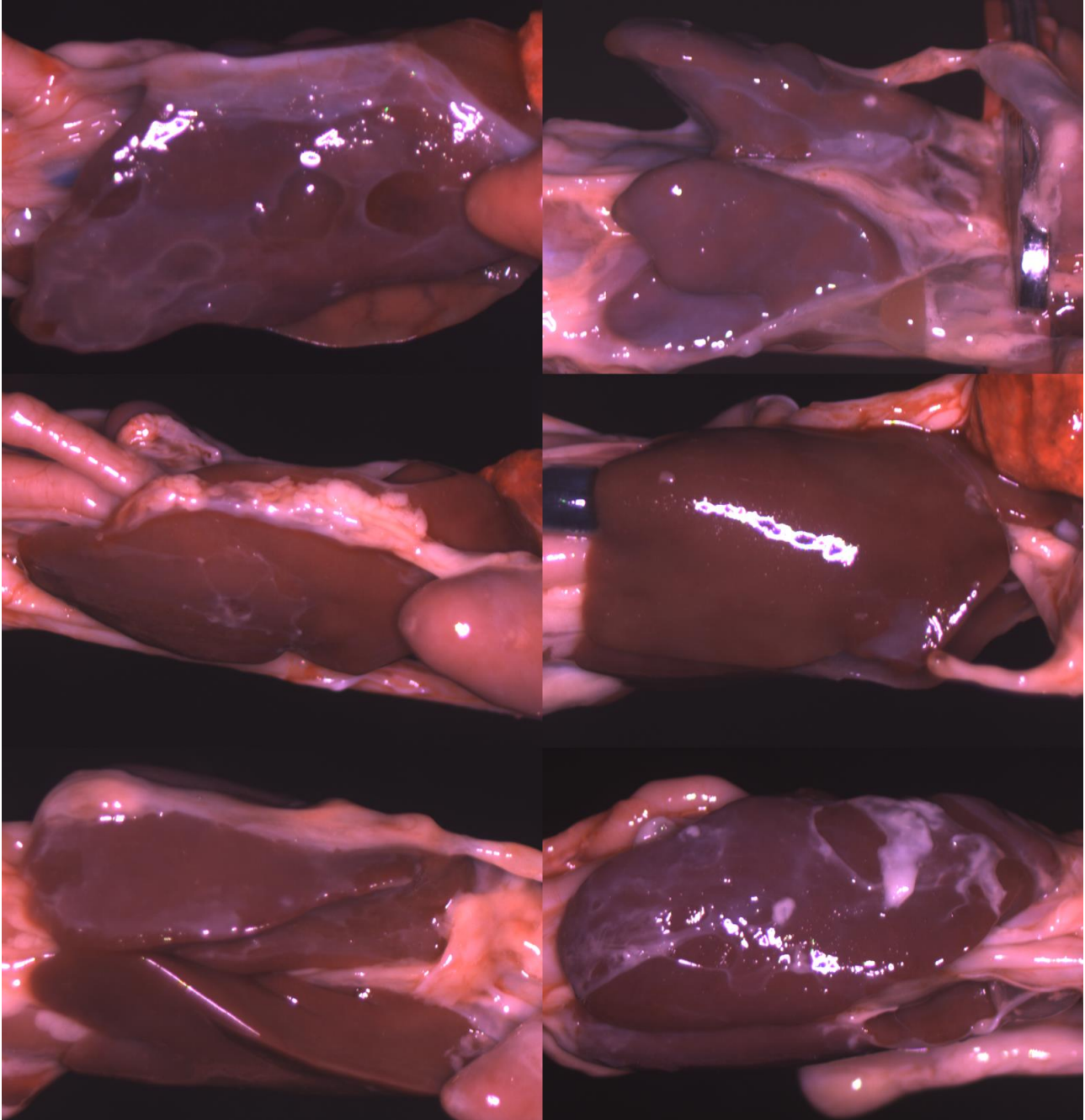
I det følgende vises illustrationer af kategorierne Rask lever, Leverstase, Forandringer på leveroverfladen, Nekrotiserende leverbetændelse, Rask hjerte og Hjertesækbetændelse, i henholdsvis Figur 21, Figur 22, Figur 23, Figur 24, Figur 25 og Figur 26.



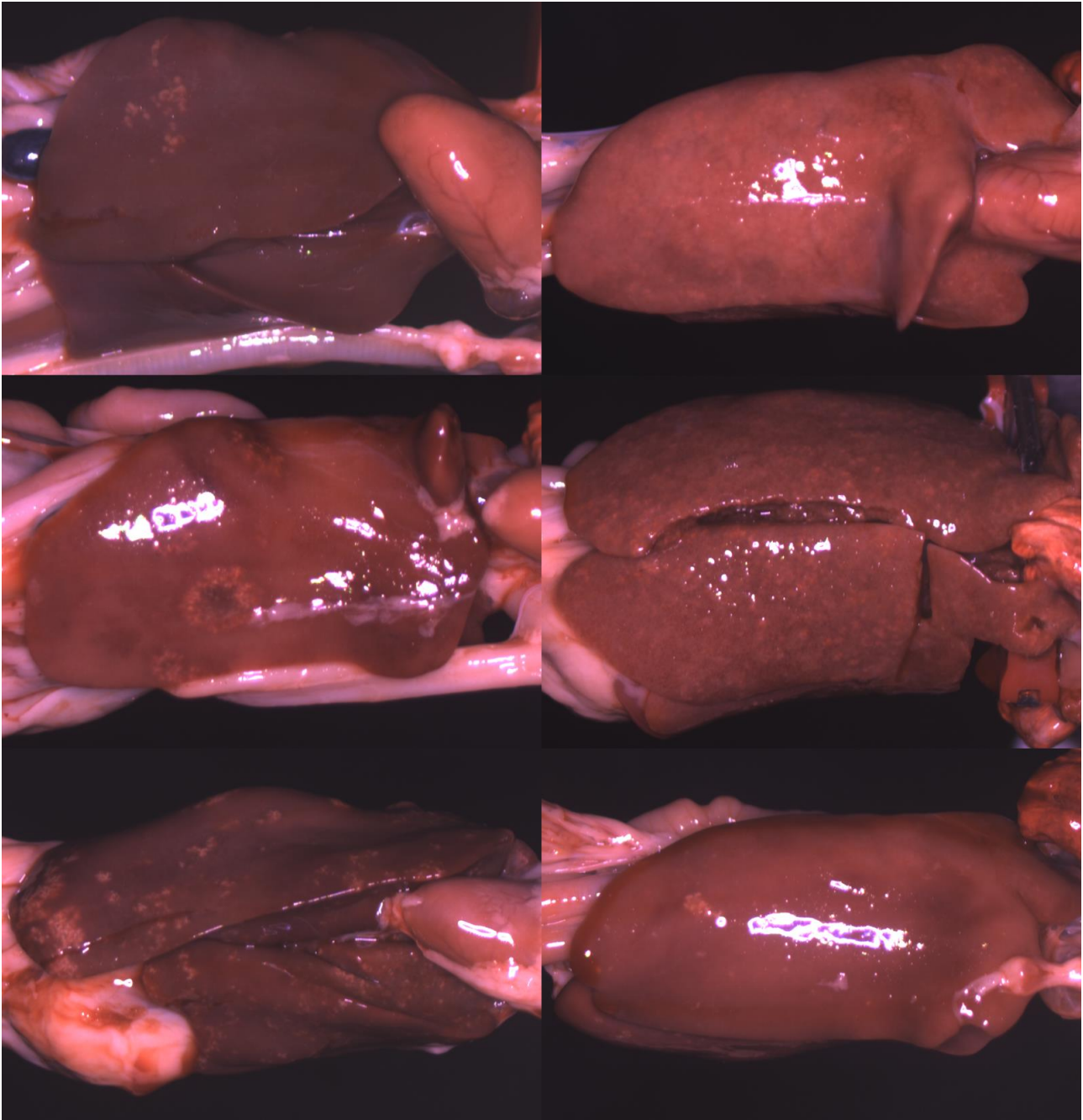
Figur 21: Eksempler på Rask lever.



Figur 22: Eksempler på Leverstase.



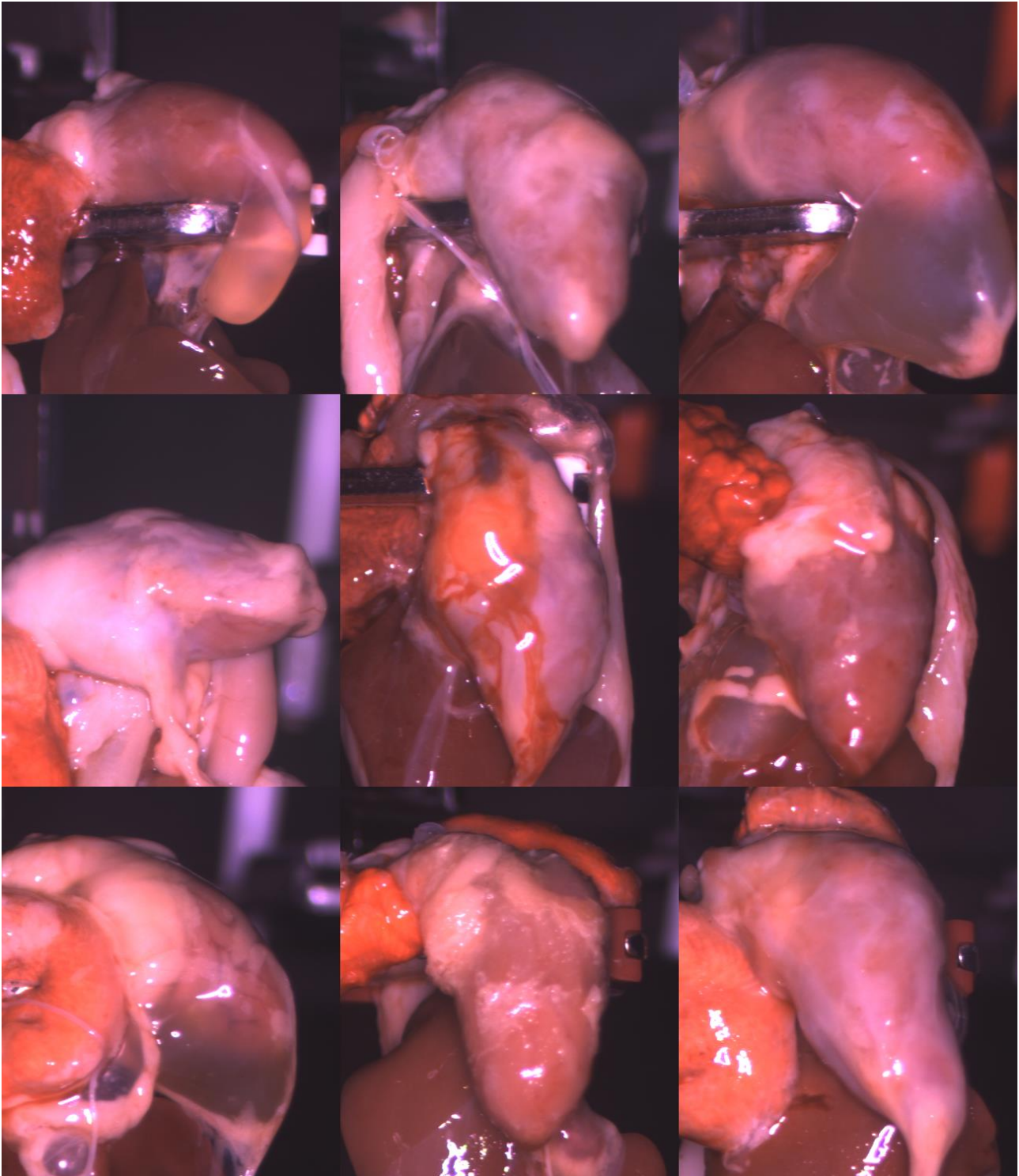
Figur 23: Eksempler på Forandringer på leveroverfladen.



Figur 24: Eksempler på Nekrotiserende leverbetændelse.



Figur 25: Eksempler på Rask hjerte.



Figur 26: Eksempler på Hjertesækbetændelse.

Appendiks C – Beskrivelse af algoritmen for segmentering af leveren

Af Jacob Velling Dueholm

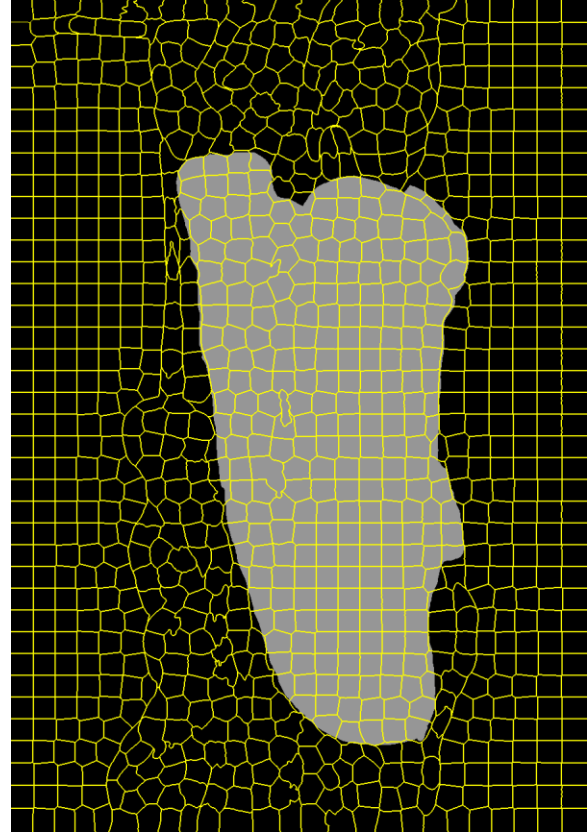
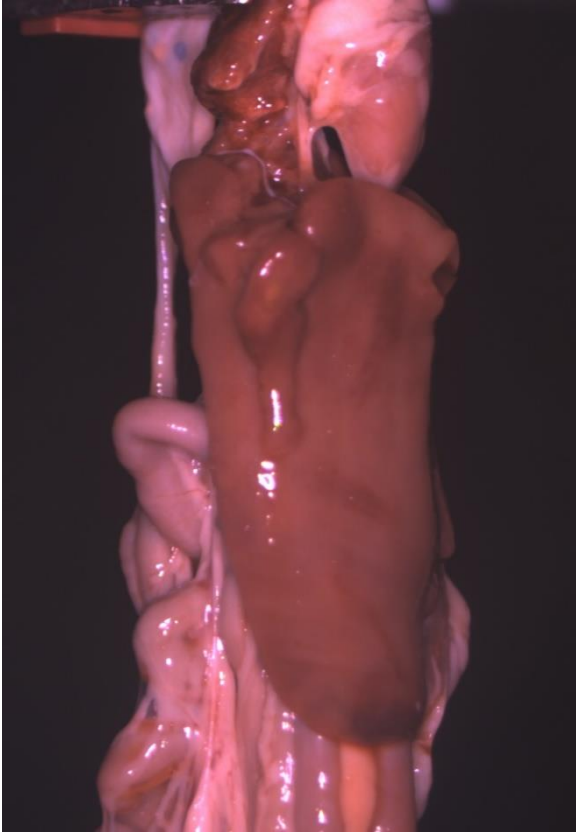
Denne metode går under maskinlærings kategorien *“supervised learning”*, hvor en model læres igennem eksempler. I denne applikation arbejdes der i første omgang med to klasser værende lever og alt det andet, hvor eksemplerne består af relativt store mængder pixel niveau annoterede billeder. Den følgende beskrevne algoritme er anvendt både under træning og til test af hidtil usete billeder.

Det første trin i algoritmen er en oversegmentering der grupperer ensartede pixels i såkaldte superpixels. Dette giver en hurtigere program eksekvering, men har også til fordel at visualisere grænser, som ses af figur Figur 27 overlagt den manuelle annotering. Her er anvendt *“Simple Linear Iterative Clustering”*, hvormed et kompromis mellem størrelsen og præcision af superpixels indgås, da mindre superpixels giver et mere præcist resultat på bekostning af længere beregningstider.

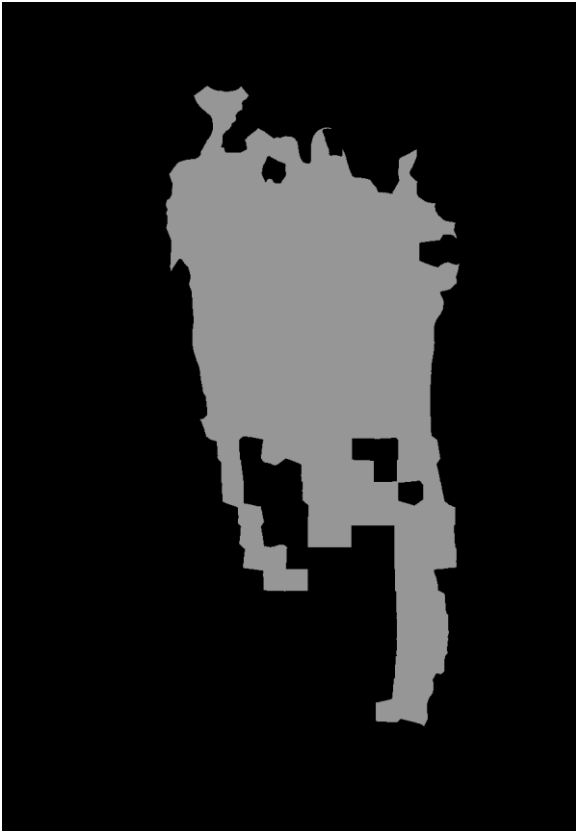
Det andet trin er udtrækning af karakteristiske features såsom farve og position i billedet. Dette gøres for hver enkel superpixel. De udvundne features anvendes af et klassificeringsmodul som har til opgave at afgøre om den enkelte superpixel tilhører lever eller baggrunds klassen. Siden der generelt er mindre lever end baggrund arbejdes der med ubalancerede klasser. Dette kompenseres for i en vægtning baseret på det annoterede trænings data. Til den egentlige klassificering anvendes en ensemble metode ved navn *“Random Forest”*, som består af en række mindre *“decision trees”*. Hvert træ bruger en forskellig kombination af features og grænseværdier i dets afgørelse mellem lever og ikke-lever. Afgørelserne for de mange træer summeres for det endelige resultat, værende en sandsynlighed for at tilhøre hver klasse. Den mest sandsynlige klasse kan bruges direkte til segmentering, som ses i Figur 28, hvorpå det også ses at masken er hullet.

Generelt kan det siges, at superpixels sandsynligvis tilhører samme klasse som de omkringliggende superpixels. Segmenteringen kan dermed forbedres yderligere ved at introducere kontekst i modellen, hvor der med kontekst menes information om de omkringliggende superpixels. Disse naboskaber tages med i betragtning ved brugen af *“Conditional Random Field”*, som dermed bruger både sandsynligheder fra *“Random Forest”*, samt forskellen i features af nabo-superpixels. En graf opbygges af sammenhænge mellem superpixels, hvor der for hver forbindelse udvindes nabo-features, værende forskelle i de oprindeligt fundne features. Kombinationen af sandsynligheder for den enkelte superpixel samt ensartetheden på tværs af grænser, gør det muligt at korrigere for fejl begået i forrige trin, og dermed udfylde huller i masken. Dermed opnås den endelige segmentering som ses i Figur 28.

Denne algoritme giver endvidere mulighed for udvidelser til flere klasser såsom hjerte, lunge m.m. Igen vil kontekst kunne bidrage med relativ position, som på videre sigt kan bruges til for eksempel at beskrive leverens placering i forhold til hjertet.



Figur 27: Billede af lever som skal segmenteres (Venstre). De tilhørende superpixels med den manuelle annotering af leveren overlagt (Højre).



Figur 28: Den mest sandsynlige lever superpixel ses til venstre, det endelige resultat når der også tages hensyn til naboskab vises til højre. Oversegmenteringer er markeret med cyan og undersegmenteringer er markeret med grøn.

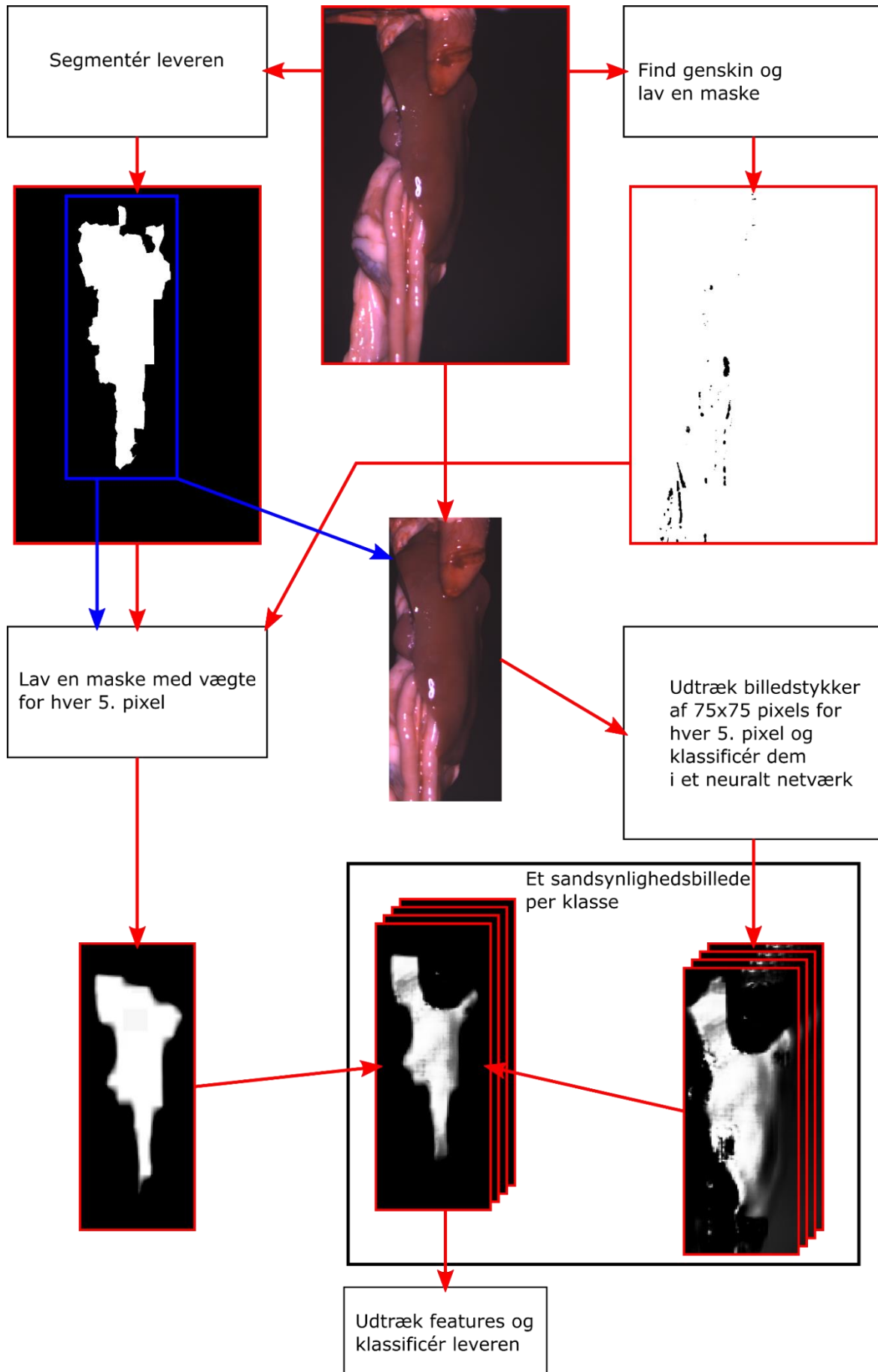
Appendiks D – Beskrivelse af vision pipelinen behandling af leveren

Lever segmenteringen foregår som beskrevet i Appendiks C. Hvorefter leverklassifikationen består overordnet af to trin. Først kigges der på regioner af leveren for at bedømme hvilken klasse denne region tilhører. Ud fra dette gives der et samlet bud på hvilken bedømmelse hele leveren skal have. Denne tilgang gør det nemmere at lede efter tekstur i regioner af varierende størrelse og form.

Inde i maven på metoden er der flere underordnede trin. Først skal pixels med genskin detekteres og sorteres fra så disse ikke forstyrrer klassifikationen. Resultatet af dette er en maske hvor pixels er 0 hvis den tilsvarende pixel i inputbilledet er genskin og 1 i alle andre pixels. Denne maske bruges sammen med masken fra segmenteringen til at lave en tredje maske der indeholder en vægt for hvor meget man stoler på resultatet af klassifikationen i denne region. Vægten er 1 hvis en region ikke indeholder nogle sorte pixels fra levermasken eller genskinsmasken og 0 hvis en region indeholder 10% eller flere sorte pixels. Det sidste er ensbetydende med at der slet ikke stoles på regionen og derfor kommer den ikke med i den samlede vurdering.

Bedømmelsen af regioner sker i et neuralt netværk. En region består af et område på 75 gange 75 pixels, hvilket bruges til at bestemme klassen for den midterste pixel i regionen. I det neurale netværk er det kun hver 5. pixel der bedømmes, hvilket betyder at resultatbilledet er nedskaleret med en faktor 5. Netværket giver som output et billede for hver klasse, der indeholder sandsynligheden for at den givne pixel tilhører netop denne klasse. Disse fire billeder ganges sammen med vægtmasken for at nedjustere de områder der ikke stoles på.

Der udtrækkes nu features fra de fire sandsynlighedsbilleder. Det er tal som hvor mange pixels de indeholder, hvor høje pixelværdierne er osv. Disse features puttes ind i en *"Binomial Gradient Boost Classifier"* der klassificerer hvert billedsæt og altså kommer med den samlede klassifikation på inputbilledet. Hele pipelinen med udgangspunkt i lever er skitseret i Figur 29.



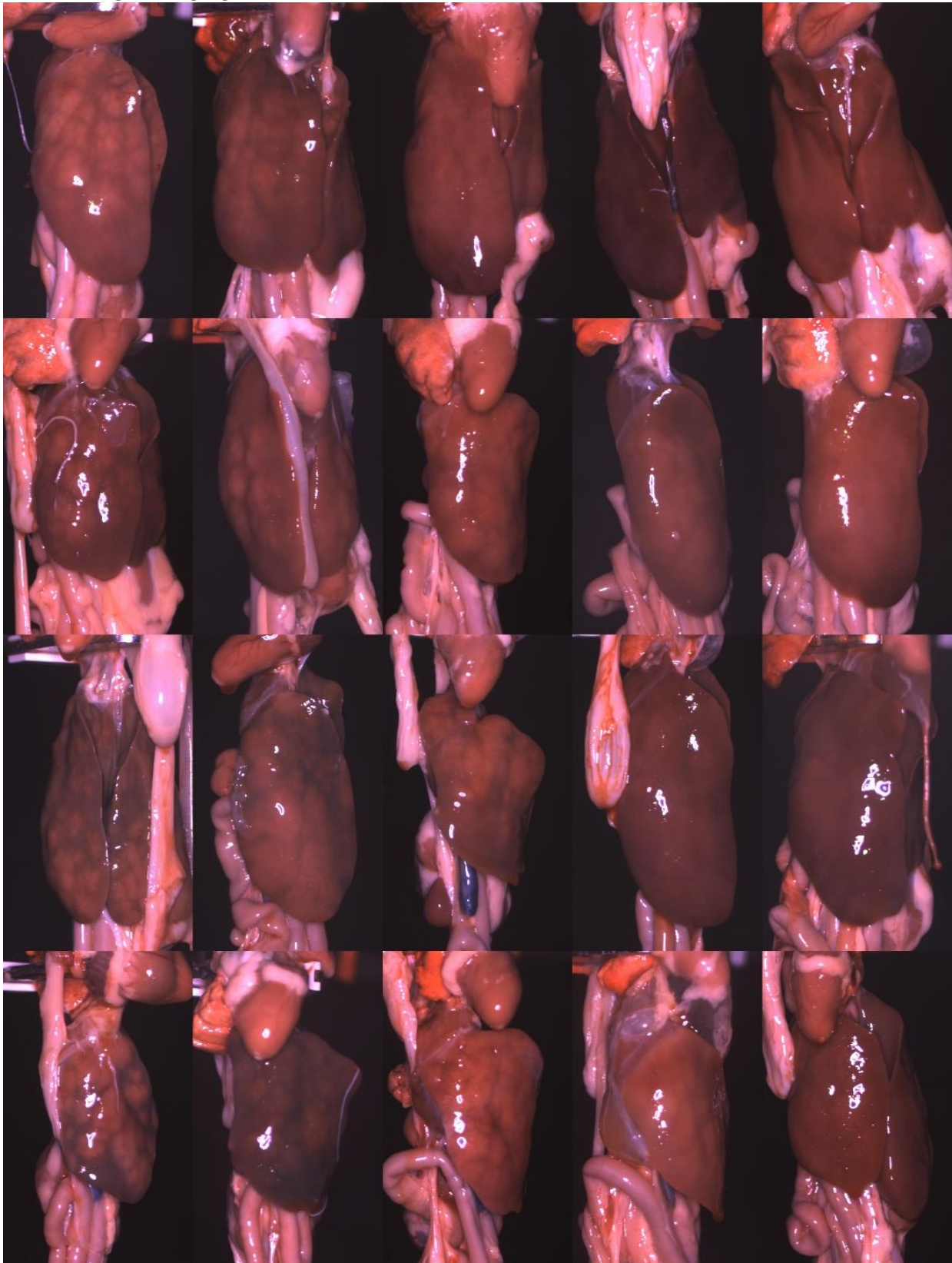
Figur 29: Oversigt over vision pipelinen for behandling af lever.

Appendiks E – Udregninger til case omkring forventet flow i produktion

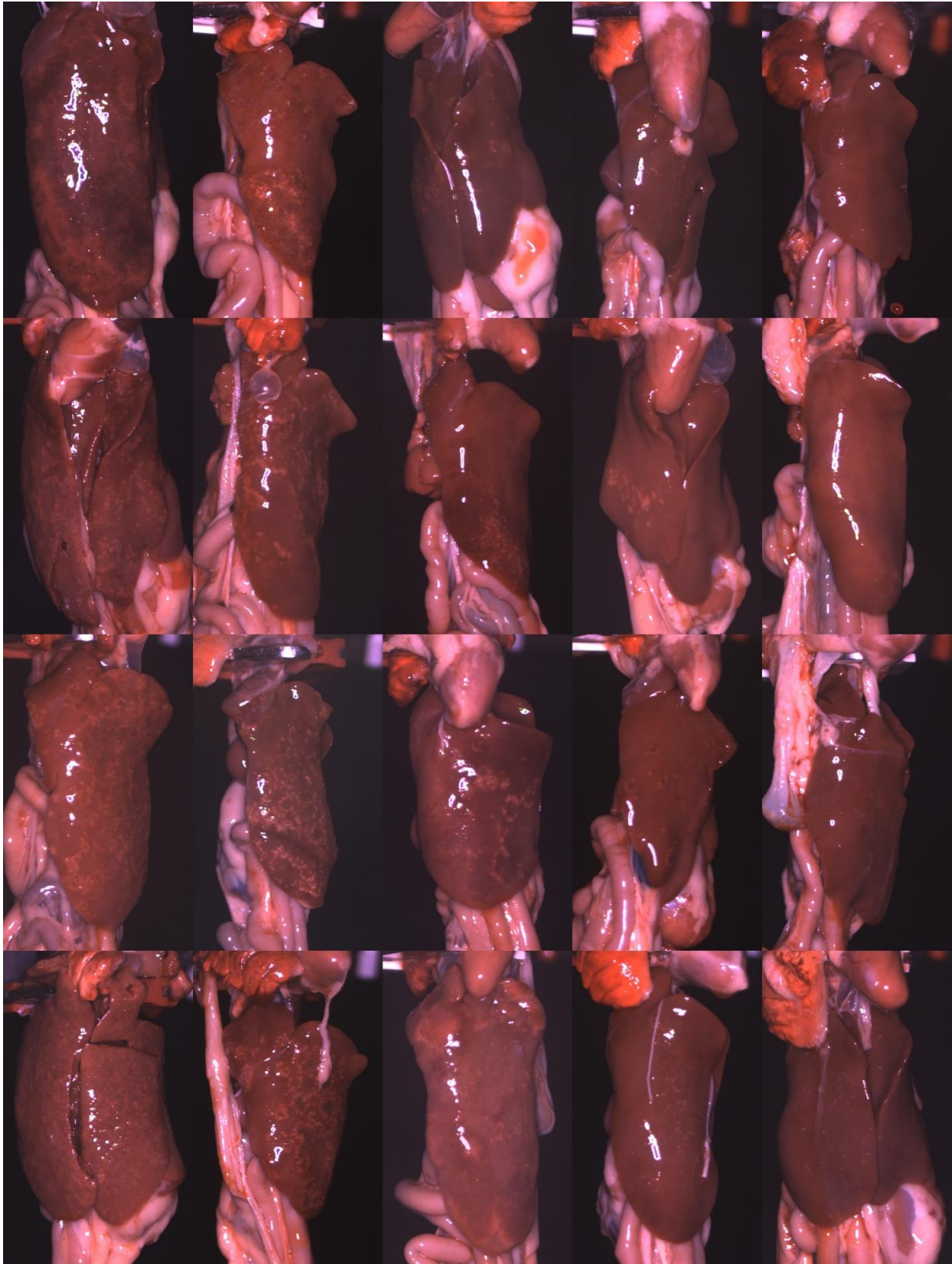
		<i>Udregning</i>	<i>Resultat</i> <i>(antal indmadssæt)</i>
Input			
<i>(14.000 indmadssæt)</i>	Raske (99,2%)	14.000 x 99,2%	13.888
	Syge (0,8%)	14.000 x 0,8%	112
Kassation			
	Raske	13.888 x 0,3%	41,7
	Leverstase	(112/4) x 2,0%	0,6
	Forandringer på leveroverfladen	(112/4) x 24,0%	6,7
	Nekrotiserende leverbetændelse	(112/4) x 33,0%	9,2
	Hjertesæksbetændelse	(112/4) x 10,6%	3,0
Manuel Kontrol			
	Raske	13.888 x 19,7%	2.735,9
	Leverstase	(112/4) x 77,0%	21,6
	Forandringer på leveroverfladen	(112/4) x 70,0%	19,6
	Nekrotiserende leverbetændelse	(112/4) x 56,0%	15,7
	Hjertesæksbetændelse	(112/4) x 67,3%	18,6
Godkendte			
	Raske	13.888 x 80%	11.110,4
	Leverstase	(112/4) x 21,0%	5,9
	Forandringer på leveroverfladen	(112/4) x 5,8%	1,6
	Nekrotiserende leverbetændelse	(112/4) x 10,6%	3,0
	Hjertesæksbetændelse	(112/4) x 23,1%	6,5
Output			
<i>(Godkendte + output fra manuel kontrol)</i>	Raske	11.110,4+2.735,9	13.846,3
	Syge	5,9+1,6+3,0+6,7	16,9
	Syge % af output	16,9 / (13.846,3+16,9)	0,1%

Appendiks F – Collager over leverstase og nekrotiserende leverbetændelse

I det følgende vises collager over leverstase og nekrotiserende leverbetændelse fra klart sygt til rask, i henholdsvis Figur 30 og Figur 31.



Figur 30: Collage over leverstase fra klart sygt (nederste vestre hjørne) til rask (øverste højre hjørne).



Figur 31: Collage over nekrotiserende leverbetændelse fra klart syg (nederste vestre hjørne) til rask (øverste højre hjørne).