



Notat

Detektion og håndtering af bylder

Jon Raunkjær Søndergaard

7. januar 2025

Proj.nr. 2011293

Version 1

Init. JRSO/MT/KRGR

Formål

Projektet har haft til formål at udvikle en metode til automatisk detektion af overskårne bylder efter tredelingen af halve slagtekroppe samt alarmere relevante medarbejdere. Det kan reducere driftsforstyrrelser og kontaminering af kød og udstyr. Projektet har, ved at kombinere dataindsamling, billedbehandling og computervision, søgt at forbedre hygiejnen og effektiviteten i slagteriprocesser.

Baggrund

Forekomst af skjulte bylder i slagtesvin kan give udfordringer på slagterierne i forhold til hygiejne og driftsforstyrrelser, især i de automatiske processer, da der her kan gå længere tid, før overskårne bylder opdages, hvilket medfører kontaminering af mere udstyr og kød. I dette projekt er der arbejdet med at finde overskårne bylder i forbindelse med den automatiske tredeling af halve slagtekroppe. Ved brug af billeder taget af skærefladerne på forender og skinker, umiddelbart efter tredelingen, kan modeller (neurale netværk) trænes til at genkende overskårne bylder. Når en byld detekteres, skal dette kommunikeres til relevante medarbejdere ved hjælp af en alarm, og der skal evt. foretages et stop på slagtelinjen. Dette vil medføre hurtigere reaktion på overskårne bylder og dermed mindre kontaminering af udstyr og kød.

Dataindsamling

De halve slagtekroppe bliver delt i tre stykker, forende, midterstykke og skinke, i den automatiske tredeler. Herefter kører de tre stykker videre ad hver sit transportbånd. Her blev der opsat to kameraer, som tog billeder af skærefladerne på forenden og skinken. Dette skete ved hjælp af et specialudviklet trigger-program, som tog billeder, når kødstykkerne var midt i billedet. På dage med optimal kørsel gav det mellem ti- og fjortentusinde billeder. Skitse af kameraopsætning ses i figur 1.



Figur 1. Kameraposition og synsvinkel er indikeret på billederne med grønt for overvågning af skinker (venstre) og med gult for overvågning af forender (højre).

Efter dataindsamlingen blev alle billederne scoret af en funktion, som ledte efter bylder og gav en score baseret på, hvor mange pixels af et billede som var del af en byld. En score på nul betyder derfor, at funktionen vurderede, at der ikke var en overskåret byld på billedet. Det startede med en funktion, som ud fra et filter på

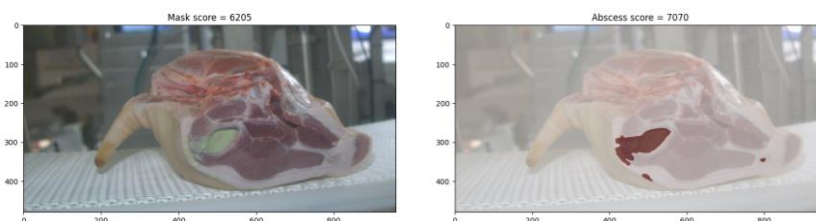
hvert billedes tre farvekanaler talte antallet af grønne pixels, da mange bylder er grønne. Det var ved hjælp af denne funktion, at de første bylder blev detekteret. Billederne med de højeste scorer blev derefter udtaget til annotering. Her blev billederne individuelt vurderet, og byldeerne markeret i et annoteringsprogram. Herefter var billederne klar til at indgå i modeludviklingen. Senere blev modellen, der ledte efter grønne pixels, skiftet ud med segmenteringsmodeller udviklet i modeludviklingsfasen.

Modeludvikling

Til selve detektionen af bylder blev forskellige typer af modeller undersøgt og sammenlignet. På baggrund af denne analyse blev en model valgt, som anvender et neuralt netværk baseret på U-net-arkitekturen med en ResNet18 backbone. U-net er særligt velegnet til billedsegmenteringsopgaver, hvilket gør det ideelt til præcist at identificere og segmentere bylder på billeder af skinker og forender. Modellen segmenterer billederne ved at identificere pixels, der tilhører bylder. En pixel klassificeres som en "byld-pixel", hvis U-net-modellen giver den en score på mere end 0,5 i et interval, der går fra 0 til 1.

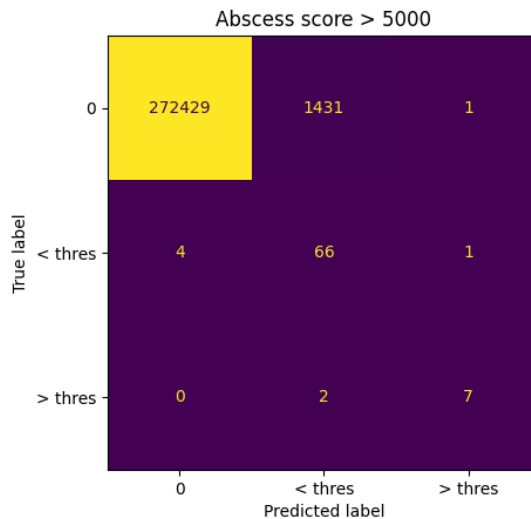
Modellerne blev trænet på de indhentede data og gentrænet i takt med, at datamængden voksede. Modellerne blev dermed bedre og bedre til at detektere bylder, da der løbende kom flere og flere forskellige bylder i datasættet. Hver gang modellen blev forbedret, blev alle billeder scoret igen, for at se om der var bylder, som den foregående model ikke havde detekteret. Den nye, optimerede model blev herefter anvendt til at score alle nye billeder, indtil en ny og endnu bedre model blev udviklet.

For at afgøre om et billede indeholder en byld, anvendes en tærskelværdi for antallet af byld-pixels: Hvis antallet af segmenterede byld-pixels overstiger denne tærskelværdi, klassificeres billedet som "byld", ellers som "ingen byld". Flere forskellige tærskelværdier blev testet: 100, 200, 500, 1000, 2000 og 5000 pixels. Figur 2 viser billeder af en skinke med en overskåren byld og de detekterede pixels tilhørende bylden. Modellerne blev udviklet, trænet og testet på metodeniveau for at sikre, at den nødvendige detektionsnøjagtighed kunne opnås.



Figur 2. Billede af en skinke med en overskåren byld (venstre) og detekterede pixels, markeret med rød, tilhørende bylden (højre).

Modellerne blev optimeret mod at have så få falske positive som muligt, da det er omkostningsfyldt at stoppe produktionen, og man derfor ikke vil have for mange falske alarmer. Som det ses i figur 3, resulterede modeludviklingen i en model, som på omkring 274.000 billeder kun havde én falsk positiv.



Figur 3. Confusion-matrix af detekterede bylder på over 5000 pixels.

Prototypetest

I prototypefasen i 2024 blev algoritmerne implementeret til online-kørsel, og der blev fastlagt en hardwarearkitektur, der lever op til de operationelle krav. Løsningen blev implementeret online på værtsslagteriet, og muligheder for dataintegration med relevante slagterisystemer blev undersøgt for at sikre, at alarmer vil kunne håndteres effektivt, fx ved at stoppe produktionen og informere de relevante medarbejdere. Der blev også gennemført en afsluttende prototypetest af systemet.

Prototypetesten er evalueret ud fra følgende metrikker: sande positive (TP), falske positive (FP), sande negative (TN) og falske negative (FN). *Recall* måler andelen af faktiske positive tilfælde, der korrekt identificeres, hvilket er afgørende for at vurdere, hvor effektive modellerne er til at opdage bylder. *Precision* indikerer andelen af korrekte positive klassifikationer ud af alle positive klassifikationer, mens *False Discovery Rate* (FDR) angiver andelen af falske positive ud af alle positive klassifikationer.

I prototypetesten var der i alt 82 skinker og 16 forender med bylder. Tabel 1 og 2 opsummerer modellernes præstation givet ved metrikkerne beskrevet ovenfor. Tabel 1 viser, at modellen for skinker opnåede høj sensitivitet ved lave tærskelværdier (recall = 0,97 for tærskel = 100 pixels), hvilket betyder, at den korrekt identificerede en stor andel af de faktiske bylder. Denne tærskel resulterede dog i mange falske positive, hvilket viste sig i en høj FDR (FDR = 0,99 for tærskel = 100 pixels). Når tærskelværdien øgedes, faldt antallet af falske positive (FDR = 0,14 for tærskel = 5000), men sensitiviteten reduceredes tilsvarende (recall = 0,14 for tærskel = 5000), hvilket indikerede en lavere evne til at opdage alle bylder. For forender viser tabel 2, at modellen ikke fandt nogen af de 16 bylder uanset tærskelværdien, dvs. for alle tærskelværdier: TP = 0, FN = 16, og som følge deraf er recall = 0. Der var samtidig en meget lille andel falske positive (for tærskel = 100: FP = 168 ud af TP+TN = 202219). Modellen for forender præsterede dermed væsentligt dårligere, end hvad der blev observeret under udviklingsforløbet, og baggrunden herfor er ikke umiddelbar klar. Det vil kræve en nærmere gennemgang af datamaterialet at finde årsagen, hvilket ikke var muligt at gennemføre så sent i projektet. Det bemærkes også, at der var et usædvanligt lavt antal bylder på forender under prototypetesten. En forudsætning

for at lave en grundig vurdering af modellens præstation i en fremtidig analyse er, at der indgår et større antal bylder i data.

Tabel 1. Algoritmens præstation for skinker ved forskellige pixel-tærskelværdier givet ved antallet af TN, TP, FN og FP samt ved metrikkerne accuracy, precision, recall og FDR.

Skinke							
Tærskel	TN	TP	FN	FP	precision	recall	FDR
100	255421	74	2	7143	0,010	0,974	0,990
200	258698	69	7	3866	0,018	0,908	0,982
500	261838	57	25	726	0,073	0,695	0,927
1000	262438	43	39	126	0,254	0,524	0,746
2000	262547	26	56	17	0,605	0,317	0,395
5000	262562	12	70	2	0,857	0,146	0,143

Tabel 2. Algoritmens præstation for forender ved forskellige pixel-tærskelværdier givet ved antallet af TN, TP, FN og FP samt ved metrikkerne accuracy, precision, recall og FDR.

Forender							
Tærskel	TN	TP	FN	FP	precision	recall	FDR
100	202051	0	16	168	0	0	1
200	202135	0	16	84	0	0	1
500	202212	0	16	7	0	0	1
1000	202217	0	16	2	0	0	1
2000	202219	0	16	0	N/A	0	N/A
5000	202219	0	16	0	N/A	0	N/A

Fremtidige perspektiver

Ud fra alle indsamlede data på værtsslagteriet blev cost-benefit-analysen udarbejdet igen. Den opdaterede analyse viste, at nedetiden i forbindelse med produktionsstop ved detektion af bylder på slagtelinjerne var lavere end forventet. I gennemsnit var nedetiden 2,9 min per stop, og den gennemsnitlige nedetid per uge var 6,9 min, svarende til ca. 6 timer per linje per år. Den økonomiske gevinst ved projektet er derfor betydeligt lavere end først antaget, og som følge heraf blev det besluttet at afslutte udviklingsarbejdet inden årets udgang og fokusere arbejdet på en grundig afrapportering, så projektet kan genoptages, hvis problemstillingen bliver relevant i fremtiden. Detektion af bylder ville muligvis også kunne kombineres med anden teknologi til overvågning af tredelingen.

Konklusion

Automatisk detektion af overskårne bylder viste sig muligt, og modellen opnåede et lavt antal falske positive på store mængder af data. Dette var dog på bekostning af sensitiviteten af modellen. Modellen er relativt lille i størrelse og kan derfor score alle kødstykker i realtid uden de store krav til GPU eller andre accelerators. Den opdaterede cost-benefit-analyse viste, at driftsforstyrrelserne forårsaget af bylder er begrænset i en sådan grad, at automatisk detektion af bylder ikke har mulighed for at begrænse nedetiden. Der vil derfor ikke være nogen økonomisk gevinst ved at finde bylderne hurtigere, end man gør i forvejen. Det forventes heller ikke, at den mindskede kontaminering vil kunne spare slagterierne for mere end omkostningerne ved et system til automatisk detektion af bylder.