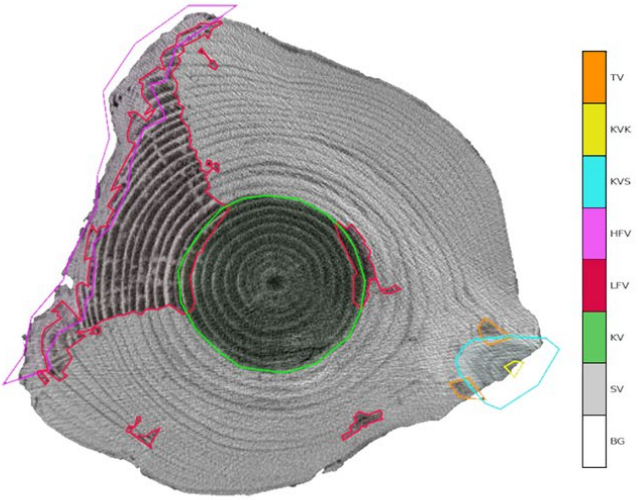
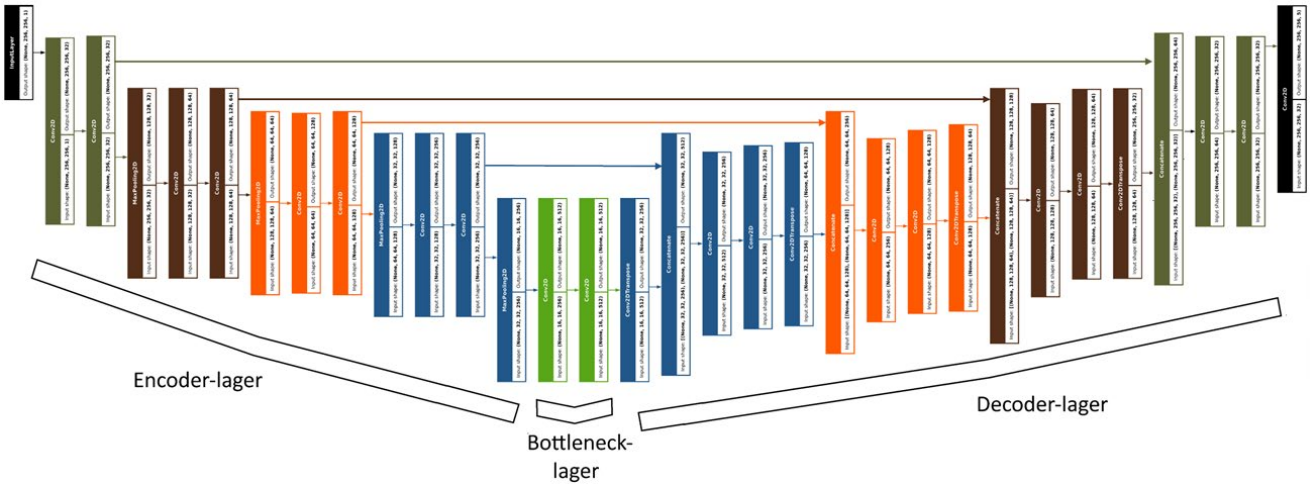
ARBETSRAPPORT 1248–2025

Neuralt nätverk för kvantifiering av törskaterelaterad fetved i CT-bilder av tall

Kari Hyll, Siri Westerblom, Mats Richardson, Sheng Joevenller, Johan J Möller





Mätning av törskateskadad tall, annotering av ett tvärsnitt av tallen i olika klasser, samt uppbyggnaden av det neurala nätverket.

skogforsk

LULEÅ TEKNISKA UNIVERSITET

Innehåll

skogforsk

Uppsala Science Park, 751 83 Uppsala
skogforsk@skogforsk.se
skogforsk.se

Kvalitetsgranskning (Intern peer review) har genomförts maj 2025 av Morgan Rossander, forskare. Därefter har Magnus Thor, Forskningschef, granskat och godkänt publikationen för publicering 23 september 2025.

Redaktör: Charlotte Hessulf, charlotte.hessulf@skogforsk.se
 ISSN 1404-305X

Förord

Många kloka sinnen har bidragit till det här arbetet. Bland kollegorna på Skogforsk tackar vi särskilt Liviu Ene för hans metodologiska rådgivning och Henrik Svennerstam som agerat bollplank vid analyserna. På Luleå Tekniska Universitets avdelning för Träteknik tackar vi Johannes Hübner, Micael Öhman och Olof Karlsson för deras insatser. Vi tackar Dick Sandberg för hans engagemang i projektet. Slutligen ger vi ett stort tack till Kempestiftelserna för deras finansiella bidrag.

Summary

A neural network for quantifying *Cronartium pini*-related resinwood in CT images of pine was developed. Previous studies have shown that CT scanning (three-dimensional X-ray scanning) can be used to detect *Cronartium*-related resinwood in pine. The detection is based on the CT scan's ability to distinguish between the higher density of healthy sapwood and the lower density of wood where resin has accumulated as a defense response to *Cronartium* infection. The goal of this work was to develop a conservative model that would rather underestimate than overestimate the presence of resinwood.

A total of 437 cross-sectional images from 36 logs were semi-manually annotated into eight classes: background (air), sapwood, heartwood, low-density resinwood, high-density resinwood, knot sapwood, knot heartwood, knot transition wood, and bark. The annotated images were used to train a neural network of the UNET architecture. The effect of including more than the first four classes was investigated, as well as variations in the loss function type and weighting, and data augmentation. A total of 14 model variants were evaluated.

The model variant where low-density resinwood and high-density resinwood were treated as a combined class performed worse than when the classes were treated separately. Separating out bark as its own class (from sapwood) also worsened the resinwood prediction. Data augmentation to increase the amount of input data to the model did not improve the outcome. Model weighting improved the results for high-density resinwood but worsened them for low-density resinwood, especially when the model included many classes.

The best model variant that included both types of resinwood correctly classified 79 percent of the low-density resinwood pixels and 41 percent of the high-density resinwood pixels. Notably, the model performed well on logs unaffected by *Cronartium pini*. In logs affected by *Cronartium pini* the model underestimated the amount of low-density resinwood by 14 percent and the amount of high-density resinwood by 32 percent. However, moderate- to high presence of resinwood was still detected.

All model variants predicted some amount of low-density resinwood even in logs assessed to be unaffected by *Cronartium pini;* on average 0.03 percent and at most 0.2 percent, as a share of the total log volume. The coordinates (position) of the resinwood perimeter differed quite a bit between the annotation and the prediction, according to the evaluation metric Dice. This is likely explained by an underestimation of the resinwood area sizes. However, the models did match the resinwood area centroids on average within ±2 pixels of the annotated value.

Overall, the developed model is useful for classification at the log level and meets the criterion of being conservative, thereby preserving timber value. To further increase classification accuracy and usability for sawmill applications, the model should be trained on more data — for example, through annotation and training in 3D.

A sub-study was conducted examining the correlation between *Cronartium* impact (stem cankers or internal resinwood) and shape parameters such as circularity, eccentricity, and off-center pith. As part of this, a separate neural network was trained to locate the pith in CT images, with a mean error of ±2 pixels. The study of the relationship between *Cronartium* infection and shape parameters showed that the three strongest statistical associations were between:

- Eccentricity and stem cankers (the more eccentric the cross-section, the more likely there are stem cankers),
- Pith displacement and stem cankers (greater pith displacement correlates with a higher likelihood of cankers),
- Circularity and stem cankers (the more oval the cross-section, the more likely there are cankers).

There was also a relatively strong relationship between eccentricity and the occurrence of centrally located low-density resinwood inside the log. However, the relationships between shape parameters and internal low-density resinwood were generally weaker than those between shape parameters and stem cankers.

Sammanfattning

Ett neuralt nätverk för kvantifiering av törskaterelaterad fetved i CT-bilder av tall har tagits fram. Tidigare studier har visat att CT-skanning (tredimensionell röntgenskanning) kan användas för att detektera törskaterelaterad fetved i tall. Detektionen grundar sig i att röntgenskanningen skiljer mellan den högre densiteten hos frisk splintved och den lägre densiteten hos ved som kådanrikats som försvar mot törskateangrepp. Målsättningen med arbetet var att ta fram en konservativ modell som hellre underskattade förekomsten av fetved än överskattade den.

437 tvärsnittsbilder från 36 stockar annoterades semimanuellt i 8 klasser: bakgrund (luft), splintved, kärnved, lågdensitetsfetved, högdensitetsfetved, kvistsplintved, kvistkärnved, kviststransitionsved, och bark. De annoterade bilderna användes för att träna ett neuralt nätverk av typen UNET. Påverkan av antalet klasser utöver de fyra första undersöktes, samt variationer i förlustfunktion (typ och viktning) och augmentering. Sammanlagt utvärderades 14 modellvarianter.

Den modellvariant där lågdensitetsfetved och högdensitetsfetved behandlades som en sammanslagen klass gav sämre resultat än när klasserna behandlades var för sig. Att separera ut bark som en egen klass (från splintved) försämrade också fetvedsprediktionen. Augmentering för att öka mängden indata till modellen förbättrade inte utfallet. Viktning av modellerna förbättrade resultatet för högdensitetsfetved men försämrade det för lågdensitetsfetved, framför allt när modellen innehöll många klasser.

Den bästa modellvarianten som inkluderade båda typerna av fetved rättklassade 79 procent av lågdensitetsfetvedens pixlar och 41 procent av högdensitetsfetvedens pixlar. Inte minst var modellen bra på klassning av stockar opåverkade av törskate. Hos stockar påverkade av törskate underskattade modellen mängden lågdensitetsfetved med 14 procent och mängden högdensitetsfetved med 32 procent. Allvarlig förekomst av fetved upptäcktes dock fortfarande.

Alla modellvarianter predikterade viss mängd lågdensitetsfetved även hos stockar som bedömts vara opåverkade av törskate; i medeltal 0,03 procent och som mest 0,2 procent, som andel av hela stockvolymen. Koordinaterna (positionen) hos fetvedens perimeter skiljde sig relativt mycket mellan annoteringen och prediktionen, enligt utvärderingsmåttet Dice. Detta förklaras troligen av att fetvedsområdenas storlek underskattas. Däremot prickade modellerna fetvedsområdenas masscentra i snitt inom ±2 pixlar från de annoterade.

Sammantaget är den framtagna modellen användbar vid klassning på stocknivå och uppfyller kriteriet att vara konservativ, vilket bevarar virkesvärde. För att ytterligare öka rättklassningen samt användbarheten för postning i sågverk bör modellen tränas på mer data, vilket exempelvis kan uppnås genom annotering och träning i 3D.

En delstudie genomfördes där korrelationen mellan förekomst av törskatepåverkan (yttre sår eller inre fetved) och formparametrar som cirkularitet, excentricitet och ocentrerad märg undersöktes. Som en del av detta tränades ett separat neuralt nätverk för att positionera märg i CT-bilder, med ett medelfel på ±2 pixlar. Studien av relationen mellan törskatepåverkan och formparametrar visade att de tre starkaste statistiska sambanden fanns mellan excentricitet och yttre stamsår (ju mer excentrisk tvärsnittsform desto troligare med stamsår) märgförskjutning och yttre stamsår (ju större märgförskjutning desto troligare med stamsår) och cirkularitet och yttre stamsår (ju mer oval tvärsnittsform desto troligare med stamsår). Ett relativt starkt samband fanns mellan

excentricitet och förekomst av lågdensitetsfetved av den centrala typen inuti stocken. Sambanden mellan formparametrar och förekomst av lågdensitetsfetved inuti stocken var generellt sett svagare än sambanden mellan formparametrar och yttre stamsår.

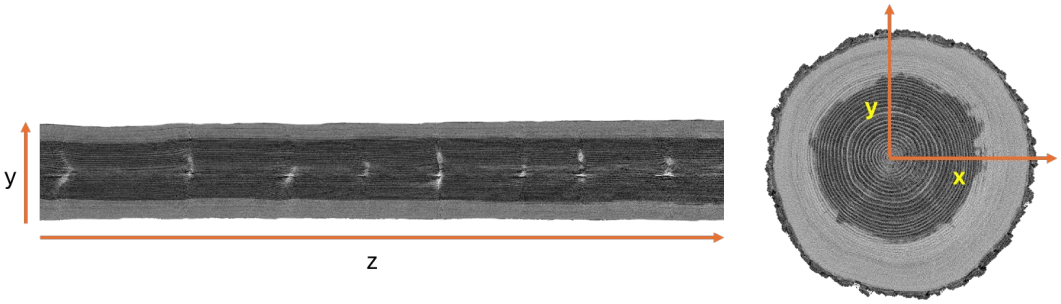
Bakgrund

CT-skanning av törskateinfekterade stockar

Studier har visat att CT-skanning (tredimensionell röntgenskanning) kan användas för att detektera törskaterelaterad fetved i tall (Hyll m.fl. 2022a, Hyll m.fl. 2022b, Hyll m.fl. 2024).

Törskate (*Cronartium pini* spp.) är en rostsvamp som angriper tall. Den tränger sig in från mantelytan genom kvisthål eller årsbarr. Trädet försvarar sig genom att anrika kåda och bilda fetved. Detta täpper till vätskekanalerna vilket torkar ut veden. Förlusten av vatten minskar vedens densitet, vilket kådanrikningen till viss del kompenserar för genom att öka densiteten. Fetved är oönskat hos sågtimmer eftersom de tilltäppta porerna inte tar åt sig målarfärg och att uttorkade områden på sågade trävaror kan leda till ojämn torkning.

För att upptäcka defekter och sortera stockar använder många sågverk sig av mätramar med diskret röntgen (en- eller tvådimensionell röntgen) eller CT-skanning (tredimensionell röntgen). De tre dimensionerna i CT-skanning visas i Figur 1.



Figur 1. De tre dimensionerna i CT-skanning av rundved: x, y och z (längdriktning). xy-riktningen kallas ofta för tvärsnitt.

Gemensamt för alla röntgenmätramar är att de främst detekterar densiteten hos provet. Den kådanrikade, uttorkade fetveden skiljer ut sig genom en lägre radiologisk densitet (mörka färger i bilderna) jämfört med den vattenrika, friska splintveden med högre densitet (ljusa färger i bilderna). Detta gäller om kådanrikningen är måttlig, vilket i denna rapport kallas lågdensitetsfetved (LFV). Lågdensitetsfetved har snarlik densitet som kärnved. Är kådanrikningen hög får fetveden en densitet som påminner om frisk splintved, även om texturen hos respektive vedtyp skiljer sig åt. Denna typ av fetved kallas i denna rapport för högdensitetsfetved (HFV).

I en tidigare studie togs en metod fram för att skilja mellan fetved och andra vedtyper i CT-bilder, baserad på klassisk bildanalys och maskininlärning (Hyll m.fl. 2024). Denna metod överskattade förekomsten av fetved. Detta innebär att friska stockar riskerar att klassas som törskatepåverkade, vilket inte är önskvärt om metoden ska användas vid sortering i sågverk.

En slutsats från det arbetet var att en industriellt användbar metod hellre bör vara konservativ, det vill säga underskatta förekomsten av fetved. Vidare bör metoden dels vara beräkningseffektiv (kunna prediktera på kort tid), dels fungera på bilder med låg spatial upplösning. En lämplig utgångspunkt för att uppnå detta är att träna en modell baserad på neurala nätverk.

Klassificering av bilder

Klassificeringsproblem brukar delas upp i binär klassificering och multiklass-klassificering. Vid binär klassificering finns endast två alternativ, till exempel ”fetved” eller ”ej fetved”. Den statistiska utvärderingen av binär klassificering är relativt enkel. Vid multiklass-klassificering finns tre eller fler klasser. Klassificeringsmetoder och metodparametrar måste ofta anpassas beroende på om uppgiften är binär eller multiklass.

Klassificering av bilder kan ske med olika detaljeringsgrad. Vid *bildklassificering* tilldelas hela bilden en klass (exempelvis ”tallobjekt" eller ”granobjekt”). Det extraheras ingen information om var i bilden olika motiv finns. Vid *objektdetektering* klassas och positioneras flera hela objekt i bilden med hjälp av begränsningsramar (bounding boxes). För ett tallobjekt skulle resultatet ge koordinaterna för tallarnas begränsningsram, men inte vilka pixlar inuti dem som är trädet och vilka som är mark eller luft. Vid *semantisk segmentering* klassificeras varje pixel i bilden, vilket skulle ge koordinaterna för alla tallpixlar i bilden. Däremot skiljer inte klassningen mellan olika instanser (förekomster) av samma klass (exempelvis skulle alla träd i klassen tall få samma etikett). Att klassificera olika vedtyper på pixelnivå i en CT-bild är en semantisk segmentering. *Instanssegmentering* är en utveckling av semantisk segmentering, men skiljer även på olika instanser av samma objekt. Varje träd i klassen tall skulle få en egen etikett, vilket gör att antalet tallar kan räknas.

Det kan noteras att det är mycket mer arbetskrävande att annotera (manuellt klassa) träningsdata till semantisk segmentering än till bildklassificering.

Neurala nätverk för klassificering av bilder

Virke delar många likheter med mänsklig vävnad i sin egenskap av biomassa och modeller framtagna för medicinskt bruk är goda startpunkter för modeller som ska tillämpas på virke. Neurala nätverk, särskilt Convolutional Neural Networks (CNN), har blivit vanliga inom medicinsk bildanalys. CNN är inspirerade av den mänskliga hjärnans synbark och består av flera lager som extraherar och analyserar mönster i bilder. Dessa lager inkluderar konvolutionella lager, som applicerar filter för att upptäcka mönster som kanter och texturer, och pooling-lager, som minskar bildens dimensioner för att effektivisera beräkningarna och hjälper modellen att generalisera och hitta övergripande mönster. Från storleken på indata och de konvolutionella lagren kan måttet träningsbara parametrar beräknas, som visar på modellens storlek och komplexitet.

Inför träning av en modell brukar datasetet splittras i träningsdata, valideringsdata och testdata, där träningsdata utgör majoriteten och validerings- och testdata brukar vara ungefär lika stor mängd. Träningsdata används för att lära det neurala nätverket att känna igen mönster genom att justera dess vikter. Valideringsdata används under träningen för att kontrollera hur väl modellen generaliserar till nya data och för att justera hyperparametrar utan att påverka träningen direkt. Testdata används efter träningen för att utvärdera modellens prediktionsförmåga på tidigare osedda data.

Förlust (loss) är ett mått på skillnaden mellan nätverkets prediktioner och de faktiska värdena. Målet för träningen av ett neuralt nätverk är att finna de vikter som minimerar förlusten. Övriga viktiga parametrar i denna process är aktiveringsfunktionen, optimeringsfunktionen, antalet epoker och batch-storleken.

Om modellen inte är förtränad skapas vikterna med en initialiseringsfunktion (ofta *Uniform Glorot*). När data matas genom nätverket multipliceras varje inputvärde med

tillhörande vikt, och summan av dessa produkter skickas sedan vidare genom en aktiveringsfunktion (exempelvis *ReLU* eller *sigmoid*) för att introducera icke-linjäritet. Utan icke-linjäritet skulle nätverket bara kunna representera linjära samband, vilket skulle begränsa den kraftigt när det gäller problem som bildigenkänning. Under träning beräknas förlusten som skillnaden mellan modellens prediktion och träningsdatas facit. Beräkningen sker med en förlustfunktion, till exempel *Categorical Crossentropy, Dice Similarity Coefficient* eller *Tversky*. Gradienten av förlustfunktionen beräknas, så kallad *backpropagation*, vilket säger åt vilket håll förlusten behöver justeras. En optimeringsalgoritm (exempelvis G*radient Descent* eller *ADAM*) beräknar hur mycket varje vikt bidrog till den totala förlusten och justerar vikterna med målet att minska förlusten. Den bestämmer också hur snabbt vikterna ska uppdateras (learning rate).

En epok är en fullständig genomgång av hela träningsdatasetet genom det neurala nätverket. Valideringsdata används för att utvärdera epokerna. Vissa parametrar ändras under epokkörningarna och ordningen på träningsdata i varje epok är slumpad. En batch är en mindre delmängd av träningsdata som används för att uppdatera vikterna. Att använda batcher gör att modellen kan tränas på stora dataset utan att kräva stora mängder minne. Små variationer i batcherna fungerar som en form av "brus", som hjälper modellen att undvika att fastna i lokala minima och förbättra generaliseringen till nya data. Genom att använda flera epoker lär sig modellen mönster genom upprepade viktuppdateringar, och konvergerar mot låg förlust.

Neurala nätverk som tillämpas på klassificering tjänar oftast på att tränas på så många annoterade bilder som möjligt. Ofta rekommenderas minst 1000 bilder per klass om en modell skall tränas från grunden (Google 2025). Beroende på tillämpning kan det dock räcka med ett tiotal (Schouten m.fl. 2021) alternativt 150-500 bilder (Shahinfar m.fl. 2020).

Det effektiva antalet bilder som modellen lär sig av kan ökas med augmentering, det vill säga att skapa modifierade kopior av de ursprungliga bilderna genom att exempelvis rotera, zooma eller spegelvända dem (t.ex. Shorten & Khoshgoftaar 2019).

UNET

UNET är en typ av CNN som är framtagen för medicinsk bildklassificering (Ronneberger m.fl. 2015). Det har en symmetrisk encoder-decoder-struktur där encodern extraherar huvuddrag (feature-vektorer) från bilden och decodern rekonstruerar bilden från dessa huvuddrag. Dessa två strukturer möts i mitten i *bottleneck-lagret*; det djupaste lagret i nätverket. En viktig komponent i UNET är *skip connections* som kopplar samman motsvarande lager i encodern och decodern, vilket hjälper till att bevara detaljer och förbättra segmenteringsnoggrannheten. UNET har visat sig vara effektivt för segmentering av medicinska bilder, inklusive CT-bilder, där det kan klassificera olika anatomiska strukturer och patologier i flera klasser. Genom att träna UNET på stora mängder märkta CT-bilder kan nätverket lära sig att identifiera och segmentera olika organ och vävnader med hög precision.

UNET:s nätverksstruktur anses vara relativt bra på att prediktera från mindre dataset (Meehan 2018). UNET presterar dock mindre bra när det gäller att prediktera fina eller små strukturer eftersom dess lagerstruktur reducerar den spatiala informationen. Som ett alternativ till standard-UNET utvecklades därför varianten Attention UNET (Oktay m.fl. 2018). Här ersätts *skip connections* av *attention gates*. Detta hjälper modellen att vikta ner irrelevanta områden medan betydelsefulla områden viktas upp. Andra utvecklingar av

UNET inkluderar Res-UNET (Diakogiannis m.fl. 2019) och 3D-UNET, som tar 3D-data som input (t.ex. Çiçek m.fl. 2016, Zunair m.fl. 2020).

Utvärderingsmått

Ett flertal utvärderingsmått kan användas för att utvärdera resultatet av klassificering. Utvärderingsmått kvantifierar en eller flera aspekter av träffsäkerhet:

1. Hur pass korrekt är det predikterade klassvärdet?
    a. Hur vanligt är det att modellen felaktigt predikterar pixlar i en viss klass trots att de egentligen har andra klassvärden (False Positives)?
    b. Hur vanligt är det att modellen felaktigt predikterar pixlar i andra klasser än den rätta (False Negatives)?
2. Hur pass korrekt är positioneringen av de instanser som har klassvärdet?
    a. Hur väl överlappar instansens predikterade perimeter med den annoterade perimetern?
    b. Hur pass nära ligger instansens predikterade masscentra det annoterade masscentrat?
3. Hur pass korrekt är antalet instanser som har klassvärdet?

I denna studie är klassificeringsproblemet av typen semantisk segmentering, vilket innebär att samtliga utvärderingsmått är relevanta. Eftersom klassificeringsproblemet också är av typen multiklass behöver utvärderingsmåtten beräknas var klass för sig.

Klassificeringsmatrisen (confusion matrix) är användbar för att utvärdera aspekterna #1a och #1b. Måttet *Dice Similarity Coefficient* (Dice) och *Jacquards Index* (även kallat IoU eller F1-värde) kan användas för att utvärdera aspekt #2, det vill säga en sammanvägning av hur bra överlappet mellan den predikterade och den annoterade instansen är. Måtten *Average Symmetric Surface Distance* (ASSD) och *Hausdorff-avstånd* kan användas för att utvärdera aspekt #2a, medan masscentrumavståndet (Center of Mass Distance) kan användas för att utvärdera aspekt #2b. För att utvärdera noggrannheten i antalet instanser (aspekt #3) kan *Connected Component*-analys användas.

Rapportens upplägg

Huvudtexten av rapporten beskriver framtagandet av en UNET-modell för klassificering av törskaterelaterad fetved och andra vedtyper i CT-skannade tvärsnitt av tall. Med utgångspunkt i tidigare insamlat material har annoteringarna utökats och en grundlig analys av annoteringens betydelse för resultatet har genomförts. Vidare har stort vikt lagts vid olika utvärderingsmått som belyser resultaten ur olika perspektiv. Resultatsektionen visar olika utvärderingsmått för tvärsnitt (där de annoterade bilderna utgör referensdata) respektive hela stockar (här kallat produktionsdata). I diskussionen tas olika felkällor samt industriell tillämpbarhet och vidare arbete upp. I Appendix återfinns delstudien kring relationen mellan stamform och förekomst av törskaterelaterade yttre stamsår eller inre fetved.

Material och metoder

Grundannotering

Materialet beskrivs närmare i Skogforsk Arbetsrapport 1189-2024 (Hyll m.fl. 2024). I korthet bestod materialet av CT-bilder från 36 tallstockar (*Pinus sylvestris)* med en upplösning 0,5 × 0,5 × 0,5 mm^3. I samband med skanningen kalibrerades den radiologiska densiteten (bildvärdena) mot vatten och luft för att konvertera dem till materialdensitet (kg/m^3).

Totalt bestod datamängden av ca 600 000 st. tvärsnittsbilder, varav drygt 140 000 st. bedömdes ha central eller perifer fetved. I den refererade studien annoterades 436 tvärsnittsbilder med det internetbaserade verktyget CVAT (CVAT.ai Corporation). Två projektmedlemmar turades om med att annotera respektive kvalitativt granska varandras annoteringar. Annoteringarna exporterades och efterbehandlades till 8-bitars gråskalebilder, innehållandes fyra klasser:

1. Bakgrund
2. Splintved
3. Kärnved
4. Lågdensitetsfetved (LFV)

Denna kombination utgjorde basen (fortsättningsvis kallad grundannotering) även i denna studie.

Annoteringsexperiment

Annoteringen spelar en avgörande roll för resultatet på grund av de olika vedstrukturernas likheter i bildintensitet (densitet) och spatial fördelning. Exempelvis har HFV liknande densitet som frisk splintved, kvistkärnved har liknande densitet som kärnved, och bark har liknande densitet som LFV.

Ett annoteringsexperiment gjordes därför där olika klasser adderades till grundannoteringen, se Tabell 1. Kvistkärnved är kvistens kärnved, kvistsplintved är kvistveden som omger kärnan, och transitionsved är en övergångszon mellan kärnved och splintved eller kärnved och kvist.

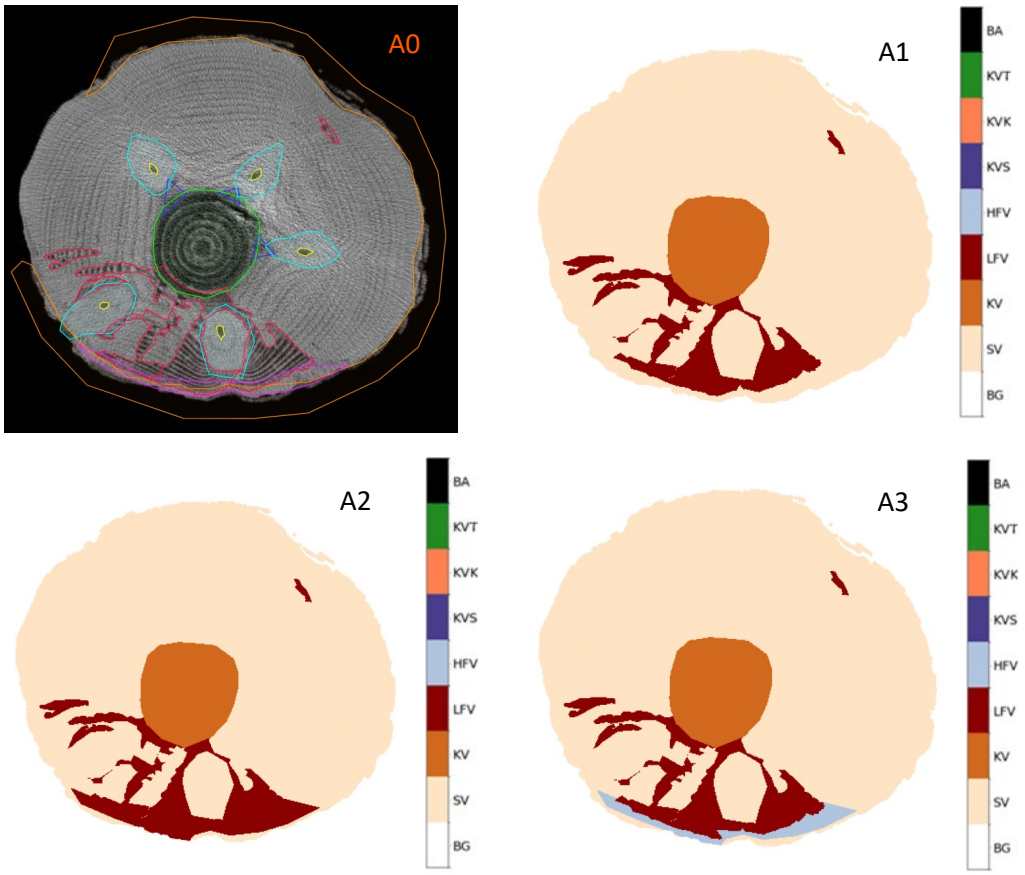
Tabell 1. Klasser och klassvärden i de annoterade bilderna, samt vilken grundklass som de nya klasserna tidigare inkluderades i.

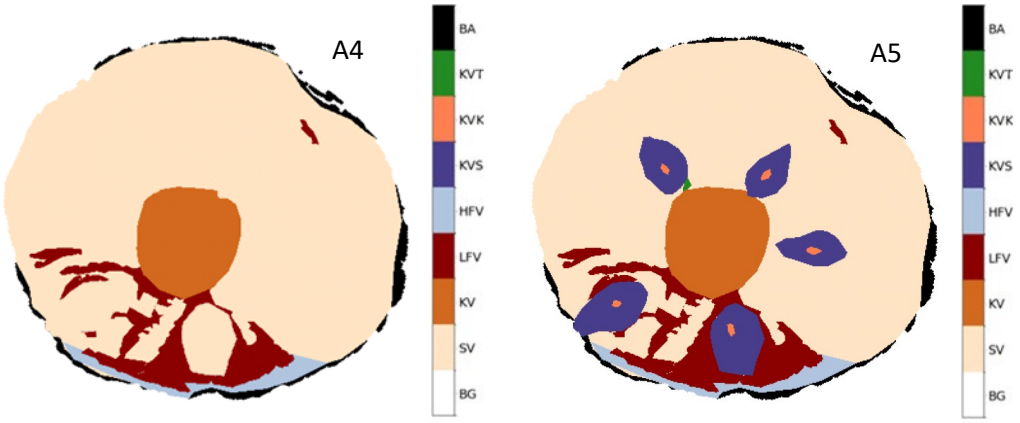
Klass	Förkortning	Klassvärde	Grundannotering
Bakgrund	BG	0	Egen klass
Splintved	SV	1	Egen klass
Kärnved	KV	2	Egen klass
Lågdensitetsfetved	LFV	3	Egen klass
Högdensitetsfetved	HFV	4	Splintved
Kvistsplintved	KVS	5	Kärnved om innanför kärnvedsradien, annars splintved
Kvistkärnved	KVK	6	
Transitionsved	TV	7	
Bark	BA	8	Splintved

För fetvedsklasserna gjordes flera variationer:

- HFV inkluderades i splintved (grundannotering)
- HFV bakades ihop med LFV till en gemensam klass
- Varsin klass för LFV och HFV

Exempel på olika annoteringsvarianter visas i Figur 2.



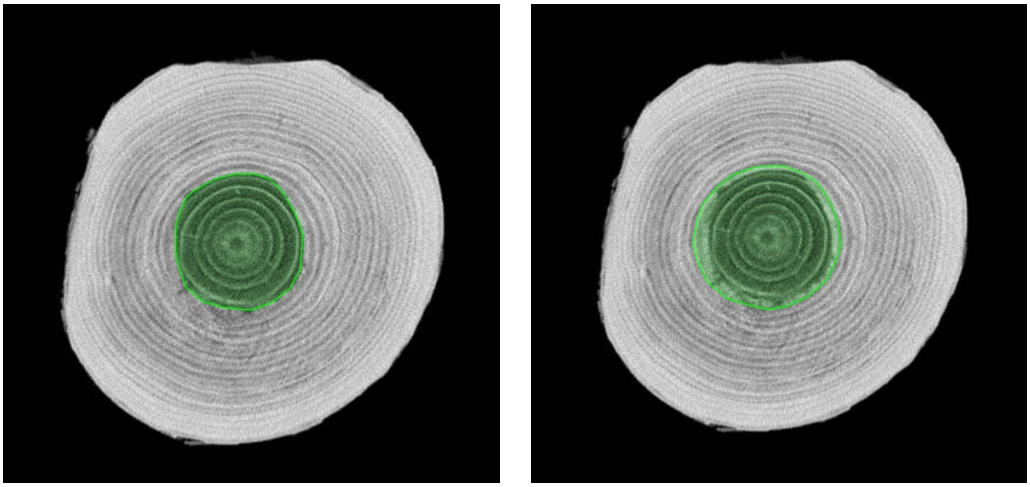


Figur 2. A0: originalbild med annoteringar innan export och bearbetning. A1: efterbehandling enligt grundannoteringen. A2: efterbehandling där HFV bakats ihop med LFV. A3: efterbehandling där HFV har adderats till grundannoteringen. A4: efterbehandling där HFV och BA har adderats till grundannoteringen. A5: efterbehandling där samtliga klasser adderats till grundannoteringen.

En ytterligare annoteringsvariation testades, i form av hanteringen av övergångszonen mellan kärnveden och splintveden:

A. Densitetsbaserad annotering, där polygonen följer densitetsgradienten (grundannotering)
B. Årsringsbaserad annotering, där kärnveden antas följa samma årsring i samtliga radiella riktningar

Exempel på de två annoteringsvarianterna visas i Figur 3.



Figur 3. Exempel på annotering (grön polygon) med densitetsbaserad övergångszon mellan kärnved och splintved (vänster) respektive årsringsbaserad övergångszon (höger).

Träningsexperiment

Baserat på grundannoteringen gjordes också fem modellvarianter där träningsrelaterade parametrar ändrades. I modell A1b ersattes förlustfunktionen Dice med förlustfunktionen Tversky. I modell A1c augmenterades bilderna med slumpmässig spegling. För modell A1d och övriga modeller med prefixet *d* användes en viktad förlustfunktion, med målet att få modellen att fokusera på klasser med få pixlar. Vikterna w beräknades klassvis enligt:

$$w_{cls} = \frac{1}{N_{cls}}$$

där N_{cls} var antalet pixlar i den aktuella klassen *cls*. De använda vikterna visas i Tabell 2. För övriga modeller sattes vikternas värden till ett för alla klasser.

Tabell 2. Värden på de vikter som användes i förlustfunktionen. Förkortningarna av klassnamnen förklaras i Tabell 1.

Modell	BG	SV	KV	LFV	HFV	KVS	KVK	TV	BA
A1d, A2d	0,02	0,22	0,90	2,87	-	-	-	-	-
A3d	0,007	0,061	0,270	0,833	3,830	-	-	-	-
A4d	0,007	0,062	0,238	0,831	3,724	-	-	-	1,138
A5d	0,001	0,008	0,031	0,101	0,433	0,325	4,754	3,208	0,140
Övriga modeller	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Sammanfattning av modellerna

Annoterings- och träningsexperimenten användes för att träna olika modellvarianter, se Tabell 3.

Tabell 3. Sammanfattning av modellvarianterna. Förkortningarna av klassnamnen förklaras i Tabell 1. Suffixet x kan stå för någon av modellvarianterna a, b, c eller d.

Modell	Annotering av övergångszon	Antal klasser	Förklaring klasser
A1x	Densitetsbaserad	4	Grund
A2x	Densitetsbaserad	4	Grund + (HFV == LFV)
A3x	Densitetsbaserad	5	Grund + HFV
A4x	Densitetsbaserad	6	Grund + HFV + bark
A5x	Densitetsbaserad	9	Alla klasser
B1a	Årsringsbaserad	4	Grund

Klassbalans

Fördelningen av antalet klasser både på tvärsnittsnivå (bildnivå) och pixelnivå är av stor betydelse för träningen av modellen, eftersom klasserna med flest pixlar tenderar att ges störst betydelse när modellen uppdaterar sina vikter. Fördelningen av klasser på tvärsnittsnivå visas i Tabell 4.

Tabell 4. Andel annoterade bilder som har minst en pixel av en given klass, per modell. Notera att andelen är 1,00 (100 %) för klasserna bakgrund, splintved och kärnved. Suffixet x kan stå för någon av modellvarianterna a, b, c eller d.

Klass/modell	A1x	A2x	A3x	A4x	A5x	B1a
Lågdensitetsfetved	0,52	0,52	0,52	0,54	0,52	0,52
Högdensitetsfetved	0	0	0,21	0,21	0,21	0
Kvistsplintved	0	0	0	0	0,59	0
Kvistkärnved	0	0	0	0	0,23	0
Transitionsved	0	0	0	0	0,31	0
Bark	0	0	0	0,73	0,71	0

Tabellen visar att datasetet är balanserat med avseende på förekomst av LFV, på tvärsnittsnivå. Det berättar dock inget om hur mycket fetved som finns i varje tvärsnitt med minst en pixel LFV. Detta visas i stället i Tabell 5.

Tabell 5. Procentandel annoterade pixlar, per klass och modell. Andelen är beräknad i relation till icke-bakgrundspixlar. Suffixet x kan stå för någon av modellvarianterna a, b, c, d eller e.

Klass/modell	A1x	A2x	A3x	A4x	A5x	B1a
Splintved	0,76	0,75	0,75	0,71	0,69	0,74
Kärnved	0,19	0,19	0,19	0,19	0,18	0,20
Lågdensitetsfetved	0,05	0,06	0,05	0,05	0,05	0,05
Högdensitetsfetved	0	0	0,01	0,01	0,01	0
Kvistsplintved	0	0	0	0	0,02	0
Kvistkärnved	0	0	0	0	0,001	0
Transitionsved	0	0	0	0	0,002	0
Bark	0	0	0	0,04	0,04	0

På pixelnivå är datasetet inte balanserat. Utöver bakgrunden (som utgör 88 procent av samtliga pixlar) är splintved är den dominerande klassen, följt av kärnved. Fem procent av pixlarna utgörs av LFV och en procent av HFV. Det är dock inte realistiskt att balansera datasetet på pixelnivå, utan målsättningen får bli att data är representativt för tallstockar med eller utan törskatepåverkan.

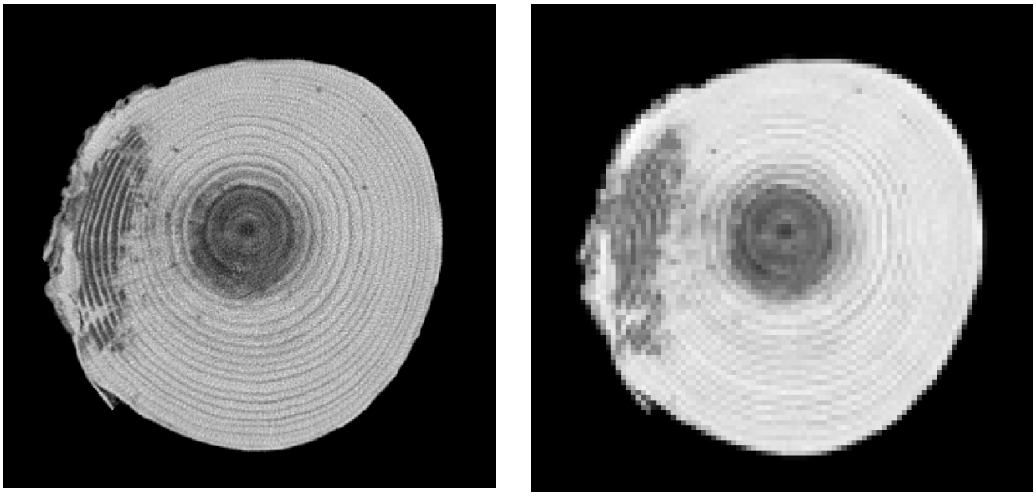
Förbehandling av originalbilder

Tröskling

Vissa av CT-bilderna hade artefakter med mycket höga intensitetsvärden, upp mot 7000 kg/m^3, att jämföra med tallens normalvärden på 800-1000 kg/m^3 (Träguiden 2017). Bilderna trösklades därför med medianen av träningsbildernas maxvärde, ca 2500 kg/m^3. Alla pixlar med värde över detta sattes till tröskelvärdet.

Förminskning

Neurala nätverk av den typ som användes i projektet kräver bilder av storleken $2^b \times 2^b$ pixlar, där b är 6, 8, 10, 12, och så vidare. Vanliga storlekar är 128 × 128, 256 × 256, eller 512 × 512 pixlar. I detta projekt användes storleken 256 × 256 pixlar, som en kompromiss mellan beräkningskrav och prestanda. Originalbilderna (900 × 900 pixlar) förminskades med Python-funktionen *cv2.resize*, där interpoleringsmetoden *cv2.INTER_AREA* användes för CT-bilderna och *cv2.INTER_NEAREST* användes för annoteringarna. Interpoleringsmetoden har stor betydelse då den ändrar både pixelvärdena och bildens textur. Träningsdata och produktionsdata (hela stockar) behöver förbehandlas på samma sätt så att förändringen av bilderna blir lika.



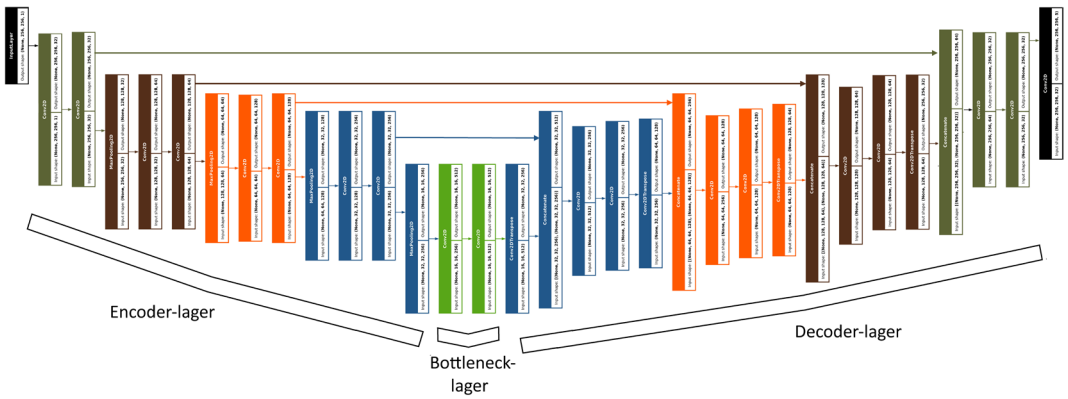
Figur 4. Vänster: tvärsnitt i originalstorleken 900 x 900 pixlar. Höger: tvärsnitt förminskat till 256 x 256 pixlar. Notera att årsringarnas utseende påverkas av interpoleringen.

Neuralt nätverk

UNET

En UNET-modell byggdes från grunden i Pythons ramverk Tensorflow-Keras. Modellen hade ett input-lager, fyra encoder-lager, ett bottleneck-lager, fyra decoder-lager och ett output-lager, se Figur 5. Antalet epoker var 50, batch-storleken 20, initialiseringsfunktionen *HeUniform* och optimeraren *ADAM*. För alla modeller utom A1b och A1e användes en förlustfunktion som var ett medelvärde av Dice Similarity Coefficient och Categorical Crossentropy. För modellerna A1b och A1e testades en förlustfunktion som var medelvärdet av Dice Similarity Coefficient och Tversky. Output-lagret hade aktiveringsfunktionen *softmax* medan övriga hade aktiveringsfunktionen *relu*. Totalt hade modellen 7 763 076 träningsbara parametrar.

Datasetet splittrades i träningsbilder (80 %), valideringsbilder (10 %) och testningsbilder (10 %). Träning och prediktion utfördes på en Azure Virtual Machine med Linux (ubuntu 22.04) Standard NV18ads A10 v5 (18 vcpus, 220 GiB minne).

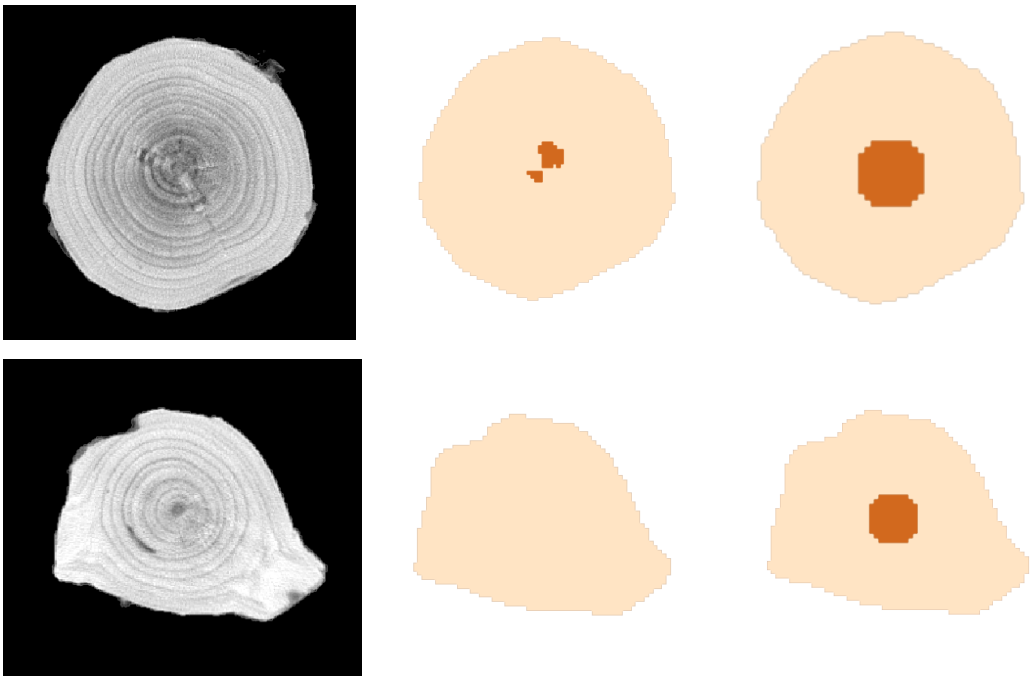


Figur 5. Lagerstrukturen hos modellen, vars nedåtgående och uppåtstigande struktur bildar den för UNET karaktäristiska ”U-formen”.

Efterbehandling av prediktioner

Expandering av kärnved hos klena stockar

Det förekom att modellen felklassade tvärsnitt från klena stockar (toppstockar från gallring) som enbart bestående av splintved. Detta berodde vanligen på en svag kärnvedsbildning i tvärsnittet. Därför tillämpades ett efterbehandlingssteg på de tvärsnitt där mindre än 5 procent av arean var kärnved. Tvärsnittets genomsnittliga radie $\bar{R}$ beräknades och en kärnvedscirkel med radien $R_{KV} = 0{,}25 \cdot \bar{R}$ placerades i tvärsnittets geometriska mittpunkt, se Figur 6. Ett typiskt värde på R_{KV} för klena stockar var elva pixlar.



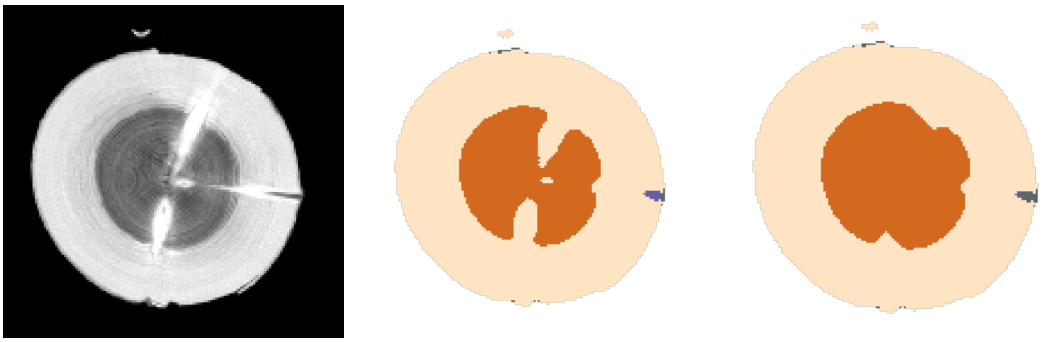
Figur 6. Exempel på korrigering av kärnved hos klena stockar. Vänster kolumn: CT-bilder. Mittenkolumn: prediktion utan efterbehandling. Höger kolumn: resultat av efterbehandling.

Fyllning av hål i kärnved

Modellen lämnade i flera fall hål i kärnveden, som i stället klassades som splintved (Figur 7). Detta beror vanligen på skanningsartefakter eller kvistar. Felklassningen åtgärdades genom att morfologisk stängning (*dask_image.ndmorph.binary_closing*) tillämpades på alla pixlar som tillhörde kärnveden, se Figur 7. Antalet iterationer N för stängningen bestämdes som halva den ekvivalenta radie R_{eq} som kärnveden skulle haft om den varit en rak cylinder:

$$N = \frac{R_{eq}}{2} = \frac{V_{KV,vox}}{2 \cdot \pi \cdot L_{vox}}$$

där $V_{KV,vox}$ är kärnvedens volym i voxlar och L_{vox} är stockens längd i voxlar. Ett typiskt värde på N för avverkningsmogna stockar var nio pixlar.



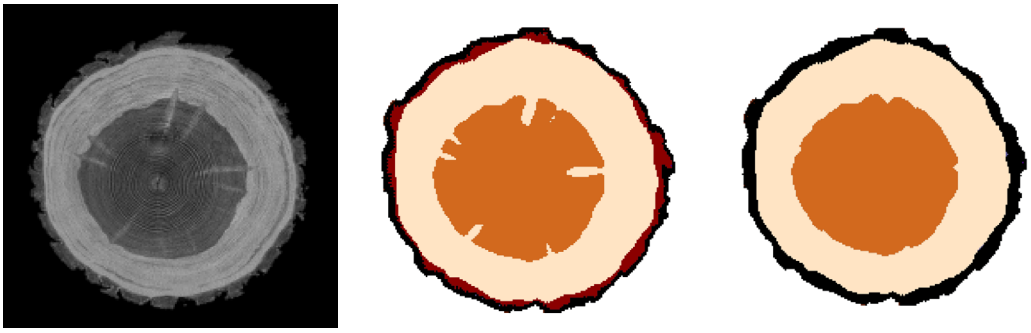
Figur 7. Exempel på fyllning av kärnvedshål. Vänster: originalbild. Mitten: modellens prediktion. Höger: prediktionen efter fyllning av kärnvedshål.

Omklassning av predikterad kärnved i barken

Modellerna hade en tendens att blanda in fetvedspixlar i barken. En morfologisk operation gjordes därför som ersatte predikterad fetved med bark, om fetvedspixlarna uppfyllde följande två villkor:

1. Var grannar med barkpixlar
2. Inom ett avstånd $s = 0{,}25 \cdot \bar{D}$ från mantelytan, där $\bar{D}$ var medelvärdet av avståndet mellan kärnveden och mantelytan

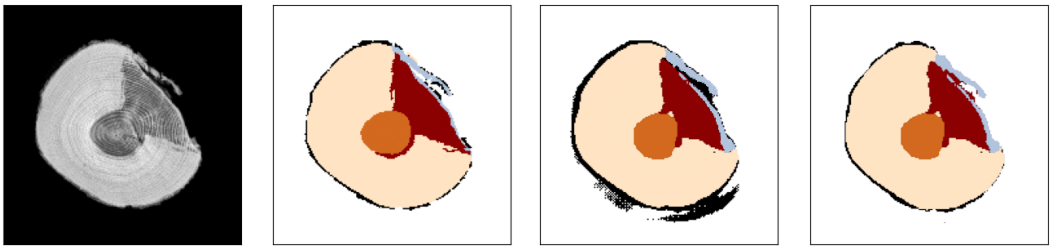
Ett typiskt värde på s för avverkningsmogna stockar var tio pixlar.



Figur 8. Exempel på omklassning av felklassad fetved vid barken. Vänster: originalbild. Mitten: modellens prediktion. Höger: prediktionen efter omklassning.

Borttagning av predikterade bakgrundspixlar

Som ett sista efterbehandlingssteg gavs alla pixlar som överlappade med originalbildens bakgrund värdet noll (0), se Figur 9. Denna felklassning utgjordes vanligen av bark som predikterats av de viktade modellerna (d-modellerna).



Figur 9. Yttre vänster: originalbild. Mitten vänster: annotering. Mitten höger: modellens prediktion. Yttre höger: prediktionen efter borttagning av pixlar i bakgrunden.

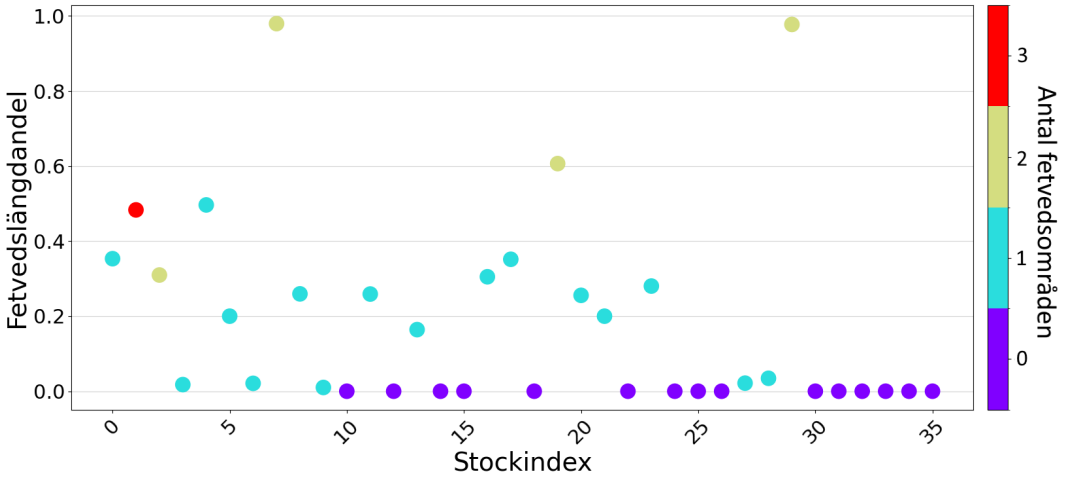
Utvärderingsmått

Den avsedda tillämpningen för projektets resultat är sortering och postning i sågverk. Då är det relevant att utvärdera både exaktheten i klassbestämningen och den spatiala noggrannheten, alltså hur väl dess koordinater har predikterats. Gemensamt avgör dessa hur väl arean/volymen av fetveden och andra vedtyper har predikterats. Därför valdes utvärderingsmåtten Dice och masscentrumsavstånd (COM), som visualiserades som låddiagram (*seaborn.catplot.box*). Utöver dessa beräknades sammanblandningsmatriser (pycm.ConfusionMatrix, Haghighi m.fl. 2018) på tvärsnittsnivå samt fetvedsandelen (LFV, HFV, LFV+HFV, i relation till icke-bakgrundspixlar) på tvärsnitts- och stocknivå. Utvärderingsmåtten beräknades på tvärsnitt ur testdatamängden, det vill säga tvärsnitt som varken tränat eller validerat modellen.

I detta projekt gick det inte att få ett mått på den sanna andelen fetved på stocknivå, eftersom endast ett fåtal tvärsnitt annoterades. Som jämförelsemått på stocknivå användes därför *fetvedslängdandelen* (FLA), definierad som den totala längden på stockens LFV som andel av stocklängden (L_{stock}), viktad för att lägga högre vikt vid central fetved jämfört med perifer fetved:

$$FLA = \frac{\sum_{n=0}^{N} 0{,}8 \cdot L_{LFV,cen} + 0{,}2 \cdot L_{LFV,peri}}{L_{stock}}$$

där N är totala antalet fetvedsområden i stocken, $L_{LFV,cen}$ är längden hos de centrala fetvedsområdena och $L_{LFV,peri}$ är längden hos de perifera fetvedsområdena. Längden och antalet på fetvedsområdena bedömdes via visuell inspektion av CT-bilderna, se Hyll m.fl. (2022a). Värdena på fetvedslängdandelen visas i Figur 10.



Figur 10. Fetvedslängdandel (FLA) för varje stock, färgkodad med antalet visuellt bedömda fetvedsområden i stocken.

Den totala predikterade fetvedsandelen på stocknivå visualiserades som ett stripdiagram (*seaborn.stripplot*), där dataspunkterna grupperades i påverkade och opåverkade stockar.

Simulering av industriella mätramar

Tidigare studier har simulerat hur resultat framtagna med CT-skanner skulle bli för diskreta röntgenmätramar (Skog & Oja 2007). Simuleringsprogrammet var dock framtaget för tidigare modeller av CT-skannrar och var inte tillämpbart i detta projekt på grund av föråldrade dataformat.

För att efterlikna de industriella mätramarnas upplösning (Tabell 6) tillämpades ett lokalt medelvärdesfilter (*scipy.nimage.uniform_filter)*, vilket anses vara en god approximation av diskreta röntgenmätramar och en acceptabel approximation av industriell CT-skanner (Johan Oja, LTU/Norra, pers. komm. 2025-03-24). För att simulera en större utsuddning av detaljer testades även ett lokalt gaussfilter (*scipy.ndimage.gaussian_filter*). Filterstorleken för medelvärdesfiltret sattes till (21, 3, 3), medan sigma-värdet för det gaussfiltet sattes till (12, 2, 2), baserat på att sigma bör vara ca 0,6 gånger den önskade nedsamplingen. För att simulera diskret röntgenmätram skapades digitala tvärsnitt i det geometriska planet x-z, alternativt både x-z och y-z.

Tabell 6. Publicerad spatial upplösning hos olika typer av röntgenmätramar för rundvirke. Notera att värdena för 1D/2D-mätramar endast gäller vissa modeller.

	CT-skanner i laboratorium	**Industriell CT-skanner**	**2D diskret röntgen**	**1D diskret röntgen**
Upplösning z-led	0,5 mm	10 mm	13 mm	13 mm
Upplösning xy-led	0,5 mm	1,0 mm	0,8 mm	0,4 mm
Källa	Ursella (2021)	Ursella (2021)	Skog (2013)	Skog (2013)

Resultat

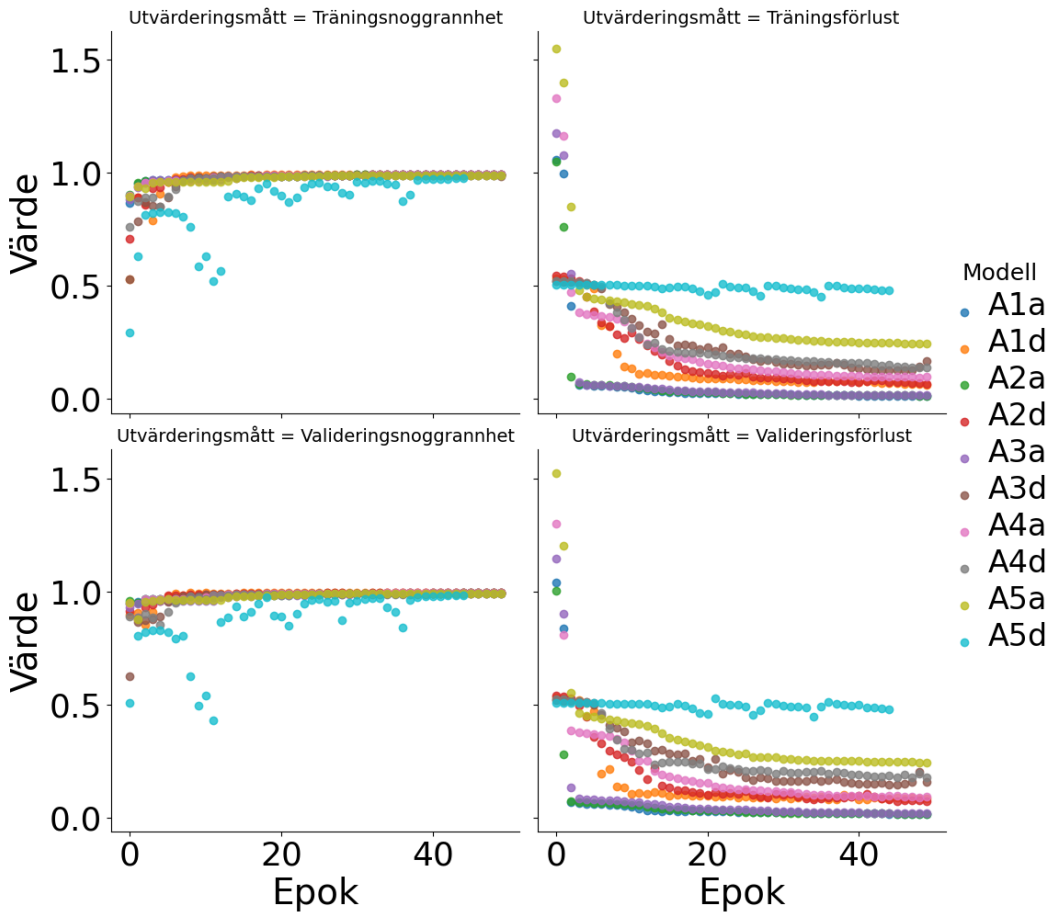
Träning, validering och testning av modellerna

Utvärderingsmått som genererades i samband med träning, testning och validering av modellerna visas i Tabell 7.

Tabell 7. Modellutvärderingsmått vid sista epoken bestående av noggrannhet och förlust vid träning, validering och testning av de olika modellerna. Suffix: b) Tversky-förlustfunktion, c) augmentering, d) viktad förlustfunktion. e) både Tversky-förlustfunktion och viktning.

Modell	Version	Tränings-noggrannhet	Tränings-förlust	Validerings-noggrannhet	Validerings-förlust	Test-noggrannhet
A1	a	0,995	0,045	0,995	0,051	0,995
	b	0,995	0,039	0,995	0,038	0,995
	c	0,994	0,048	0,995	0,053	0,994
	d	0,993	0,094	0,994	0,075	0,993
	e	0,993	0,076	0,993	0,073	0,993
A2	a	0,994	0,044	0,993	0,053	0,993
	d	0,993	0,079	0,992	0,079	0,990
A3	a	0,993	0,085	0,992	0,092	0,993
	d	0,991	0,125	0,988	0,242	0,990
A4	a	0,992	0,100	0,993	0,092	0,993
	d	0,991	0,138	0,990	0,178	0,991
A5	a	0,989	0,027	0,991	0,023	0,989
	d	0,970	0,362	0,975	0,408	0,971
B1	a	0,992	0,018	0,994	0,014	0,993

Vid en sammanvägning av modellutvärderingsmåtten presterar modell A1a (grundannotering) bäst, följt av modell A1b (grund med förlustfunktionen Dice ersatt av Tversky) och A1c (grund + augmentering). Generellt försämras värdet på utvärderingsmåtten ju fler klasser som inkluderas i modellen. Särskilt modell A5d sticker ut, med signifikant lägre utvärderingsmått. A5d hade dessutom svårigheter att stabilisera sig under träningen, se Figur 11.



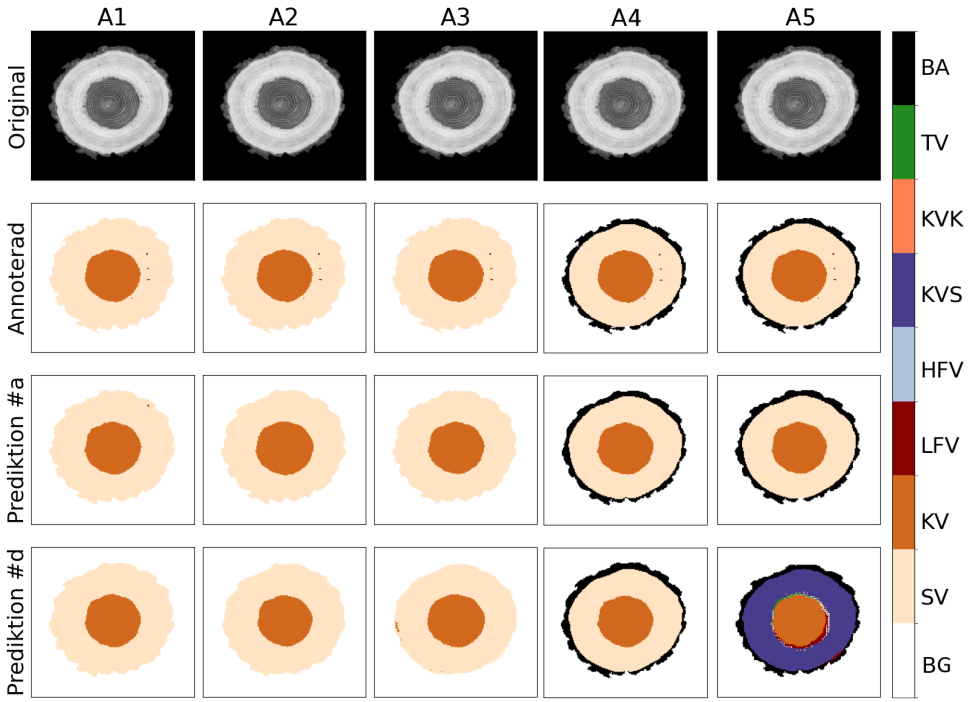
Figur 11. Modellutvärderingsmått per epok för de modeller som har a- och d-varianter. Övre vänster: träningsnoggrannhet. Övre höger: utvärderingsnoggrannhet. Nedre vänster: valideringsnoggrannhet. Nedre höger: valideringsförlust.

Modellutvärderingsmåtten domineras av de klasser som antingen har flest pixlar eller har viktats högst. Detta innebär att de inte ger så mycket information om lågdensitetsfetved, som är en mellanstor klass både i pixelantal och i vikt. Ytterligare mått behövs därför för att undersöka hur väl fetveden klassas, exempelvis sammanblandningsmatriser.

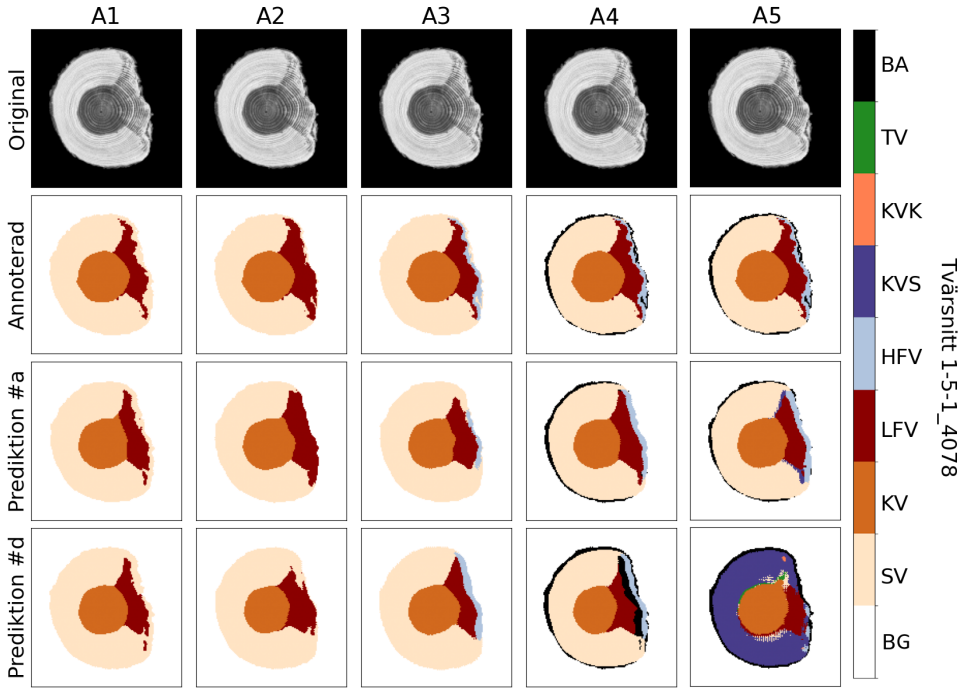
Utvärdering på tvärsnittsnivå

Exempelbilder

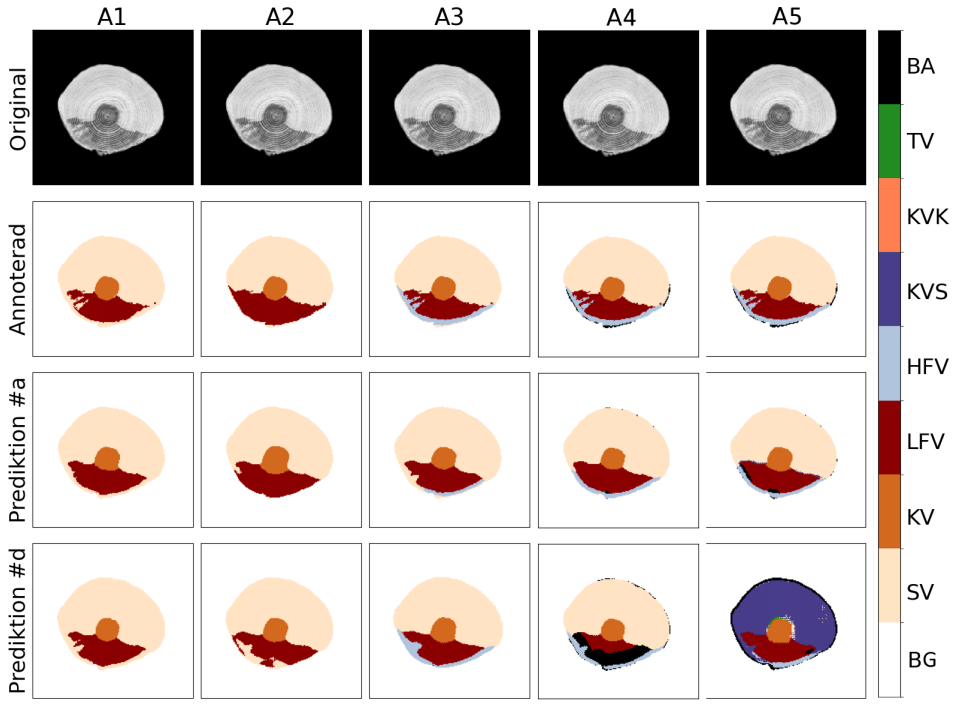
Ett urval av prediktioner visas i Figur 12 till Figur 17. Exemplen visar att modell A5d felaktigt predikterar splintved (SV) som kvistsplintved (KSV). Modell A4d predikterar i stället för mycket bark (BA). Det är dock svårt att avgöra vilken modell som är bäst genom att titta på bilderna. För det behövs sammanblandningsmatriser och utvärderingsmått.



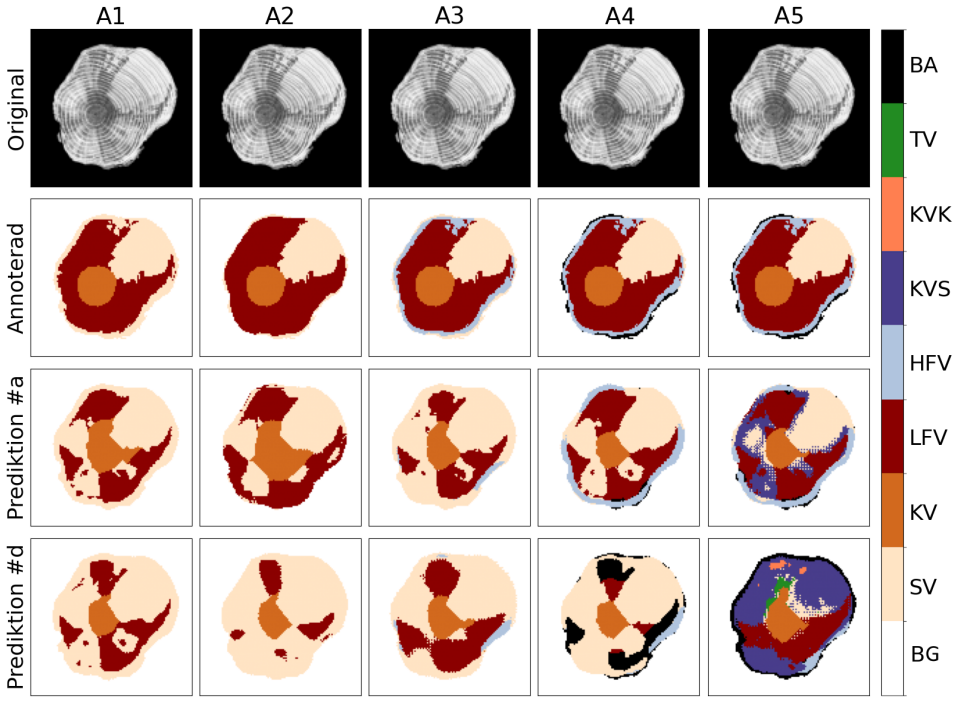
Figur 12. Exempel på indata och utdata för modellen för tvärsnitt från en förstastock opåverkad av törskate (stock 1-3-1, tvärsnitt 6548). Kolumnerna visar de olika modellerna, medan raderna visar originalbilden (överst), de annoterade bilderna, prediktionerna för a-varianten av modellen och prediktionerna för d-varianten i modellen (nederst). Förkortningarna förklaras i Tabell 1.



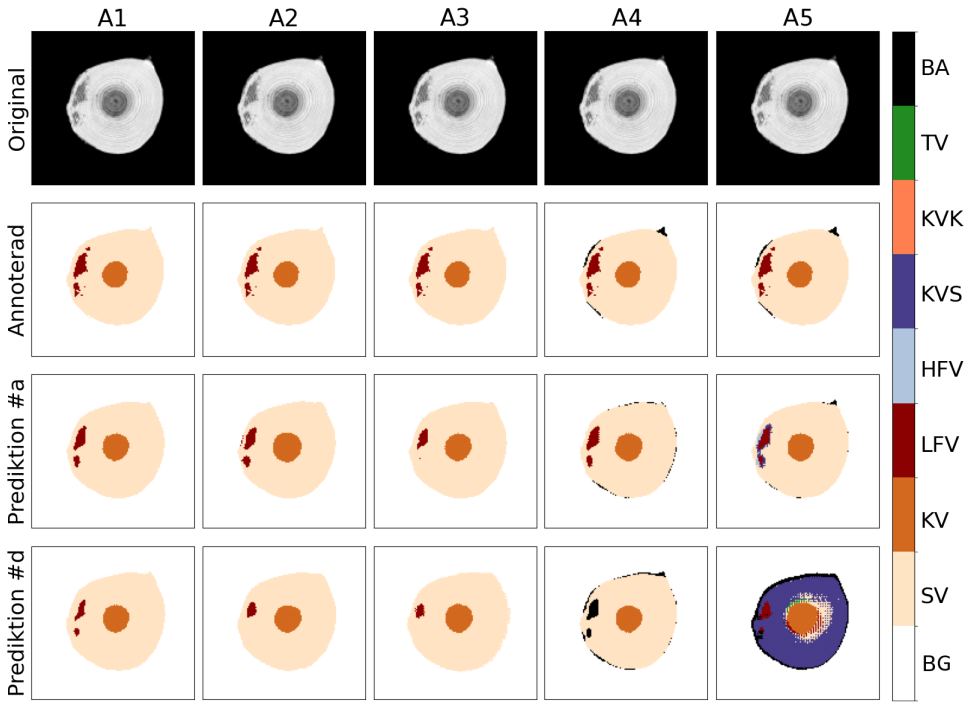
Figur 13. Exempel på indata och utdata för modellen för tvärsnitt från en förstastock påverkad av törskate (stock 1-5-1, tvärsnitt4078). Kolumnerna visar de olika modellerna, medan raderna visar originalbilden (överst), de annoterade bilderna, prediktionerna för a-varianten av modellen och prediktionerna för d-varianten i modellen (nederst). Förkortningarna förklaras i Tabell 1.



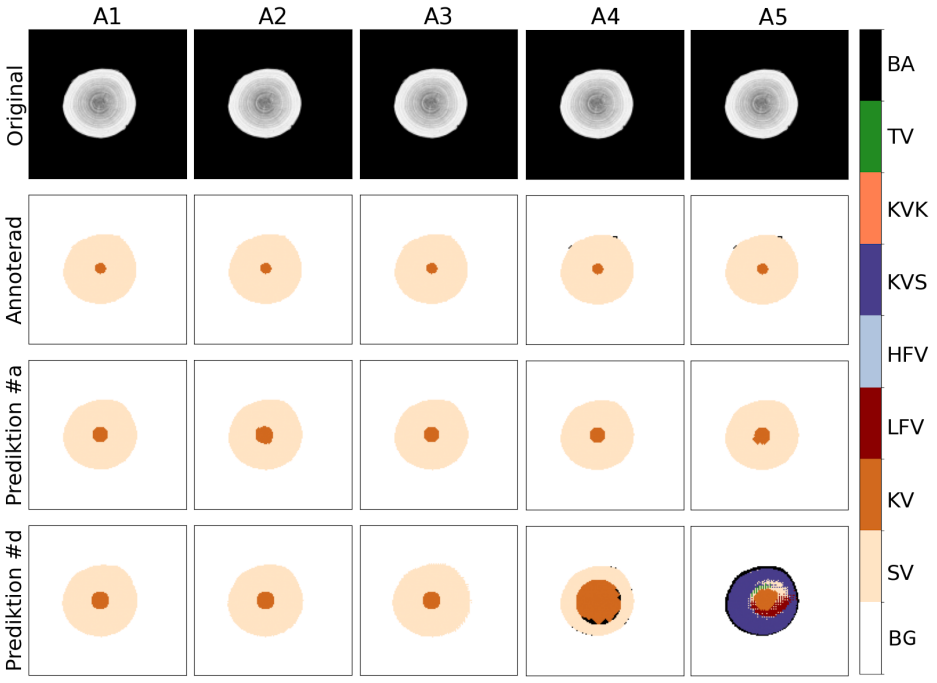
Figur 14. Exempel på indata och utdata för modellen för tvärsnitt från en andrastock påverkad av törskate (stock 1-1-2, tvärsnitt 6374). Kolumnerna visar de olika modellerna, medan raderna visar originalbilden (överst), de annoterade bilderna, prediktionerna för a-varianten av modellen och prediktionerna för d-varianten i modellen (nederst). Förkortningarna förklaras i Tabell 1.



Figur 15. Exempel på indata och utdata för modellen för tvärsnitt från en andrastock påverkad av törskate (stock 1-5-2, tvärsnitt 4589). Kolumnerna visar de olika modellerna, medan raderna visar originalbilden (överst), de annoterade bilderna, prediktionerna för a-varianten av modellen och prediktionerna för d-varianten i modellen (nederst). Förkortningarna förklaras i Tabell 1.



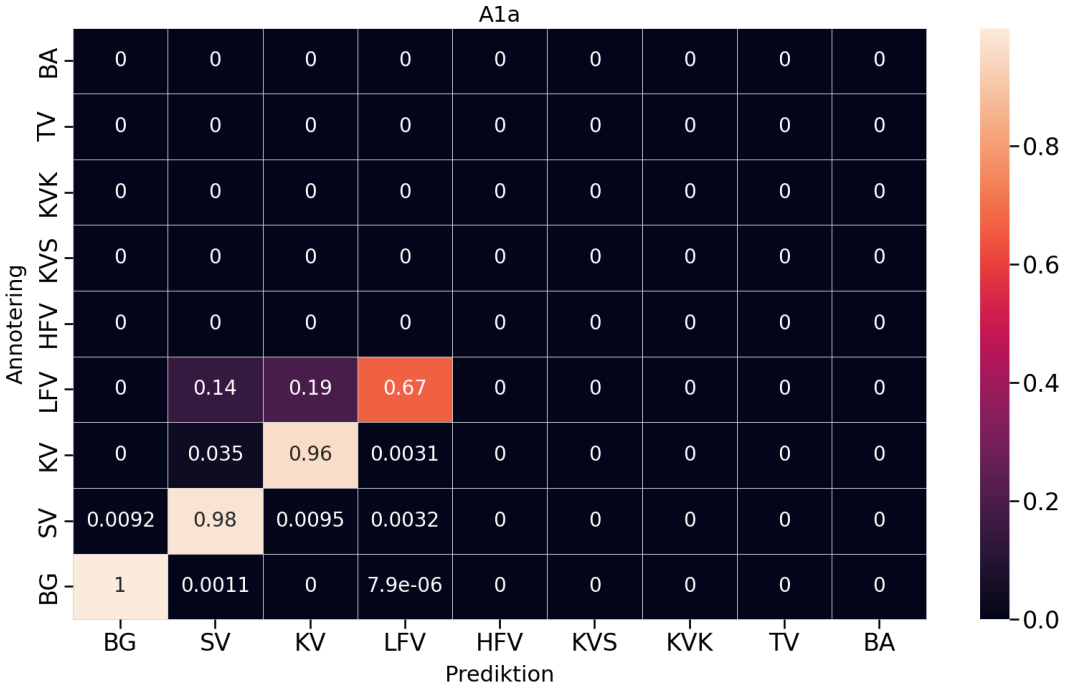
Figur 16. Exempel på indata och utdata för modellen för tvärsnitt från en tredjestock påverkad av törskate (stock 1-6-3, tvärsnitt 5708). Kolumnerna visar de olika modellerna, medan raderna visar originalbilden (överst), de annoterade bilderna, prediktionerna för a-varianten av modellen och prediktionerna för d-varianten i modellen (nederst). Förkortningarna förklaras i Tabell 1.



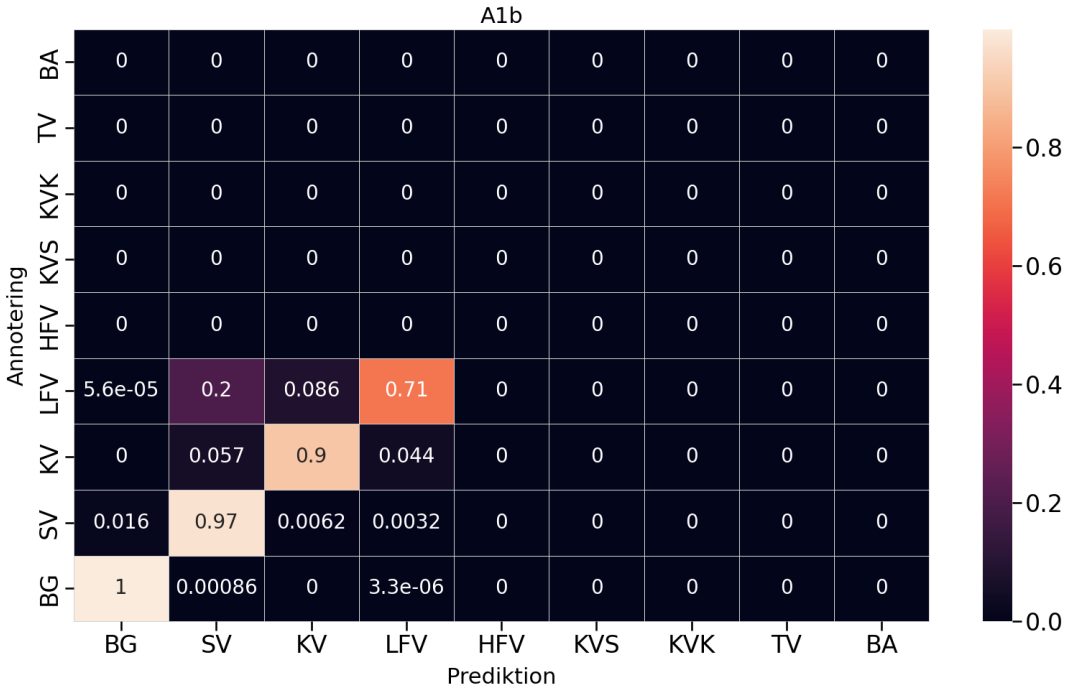
Figur 17. Exempel på indata och utdata för modellen för tvärsnitt från en tredjestock opåverkad av törskate (stock 1-3-3, tvärsnitt 4431). Kolumnerna visar de olika modellerna, medan raderna visar originalbilden (överst), de annoterade bilderna, prediktionerna för a-varianten av modellen och prediktionerna för d-varianten i modellen (nederst). Förkortningarna förklaras i Tabell 1.

Sammanblandningsmatriser

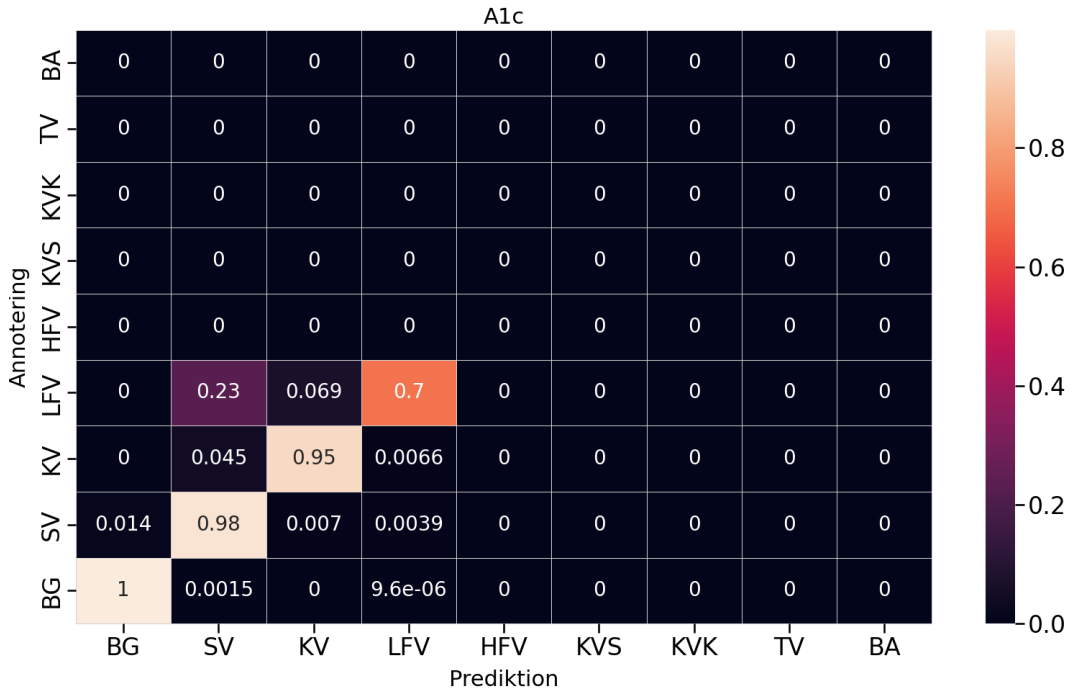
Sammanblandningsmatriser som visar träffsäkerheten för modellerna visas i Figur 18 till Figur 31.



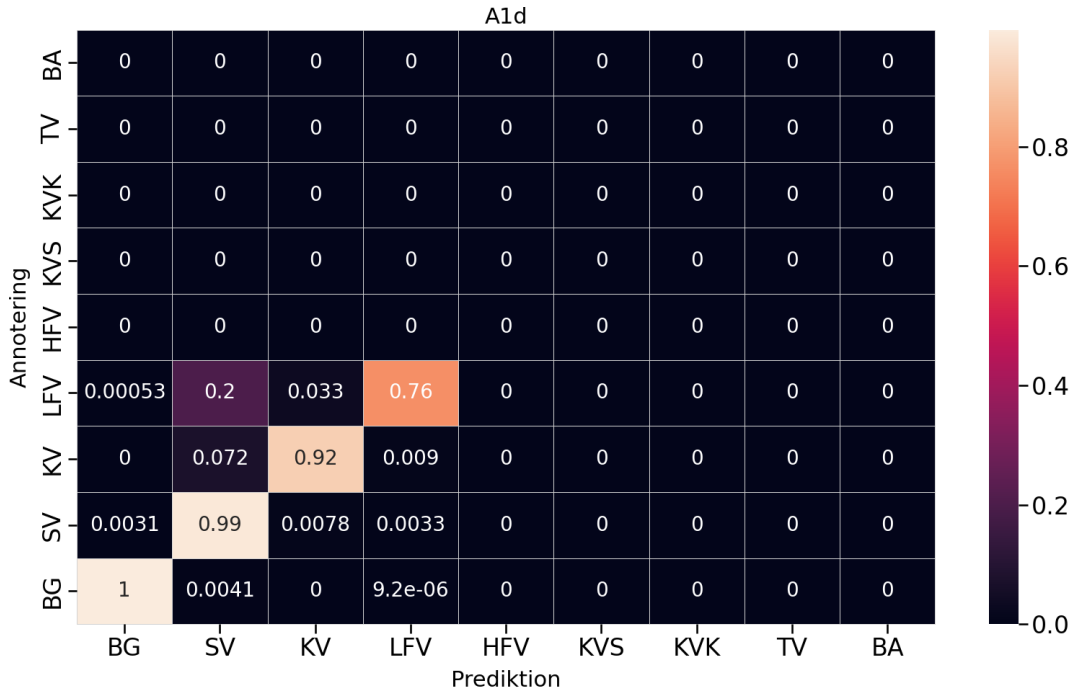
Figur 18. Normaliserad sammanblandningsmatris för modell A1a (fyra klasser, HFV inkluderad i splintvedsklassen), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



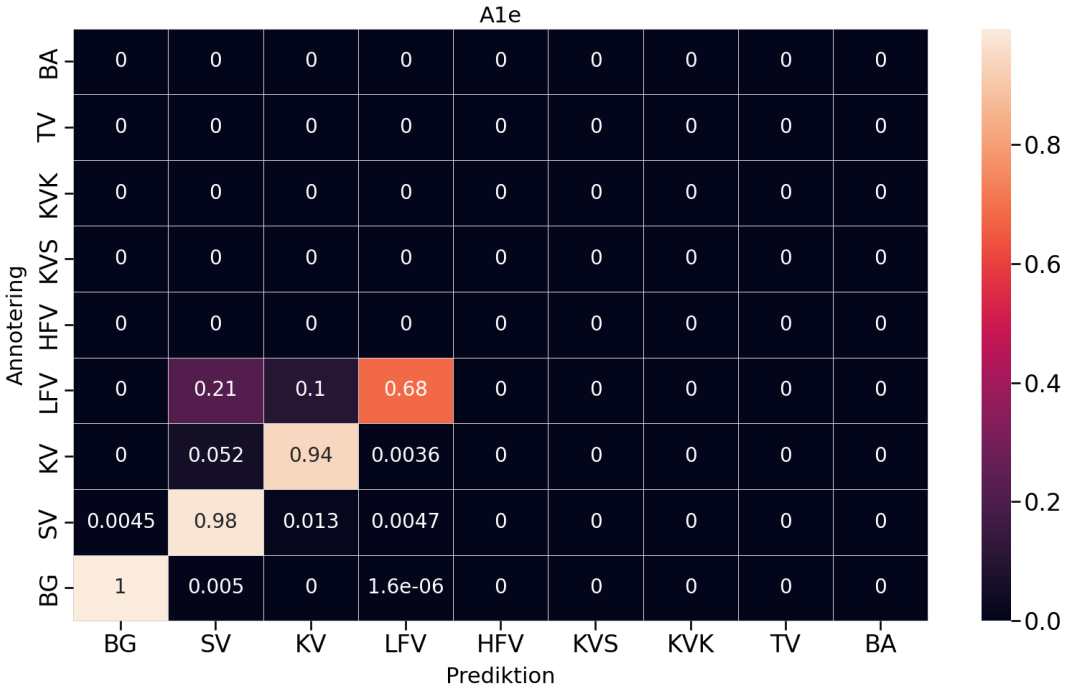
Figur 19. Normaliserad sammanblandningsmatris för modell A1b (fyra klasser, men med förlustfunktionen Dice utbytt mot Tversky), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



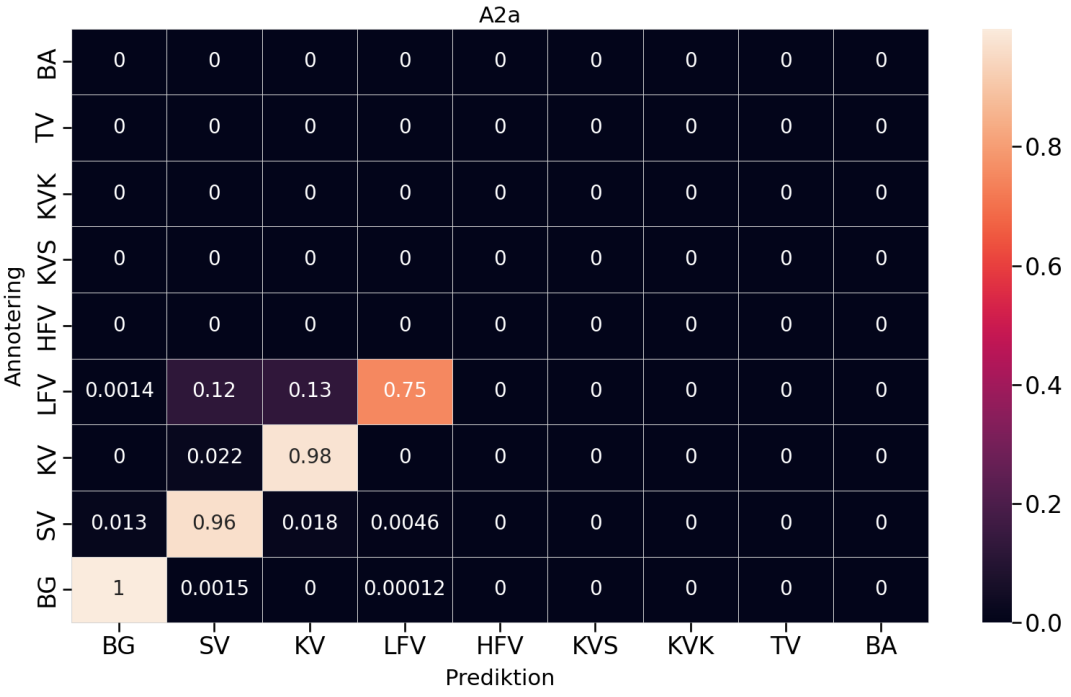
Figur 20. Normaliserad sammanblandningsmatris för modell A1c (A1a samt augmentering av originalbilderna), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



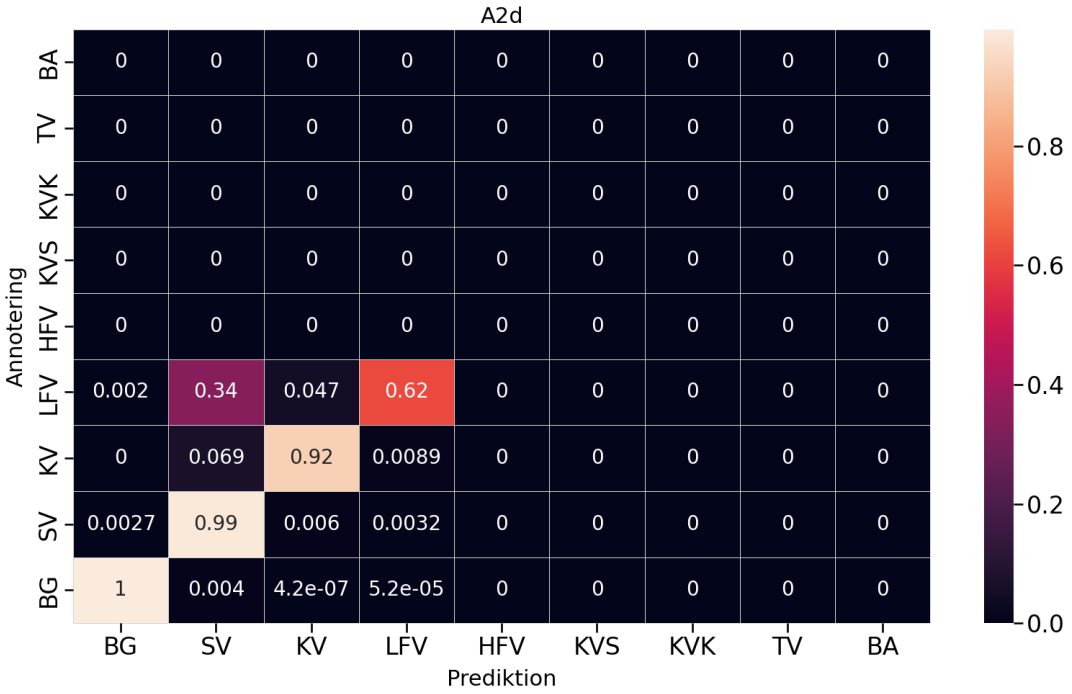
Figur 21. Normaliserad sammanblandningsmatris för modell A1d (A1a samt viktad förlustfunktion), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



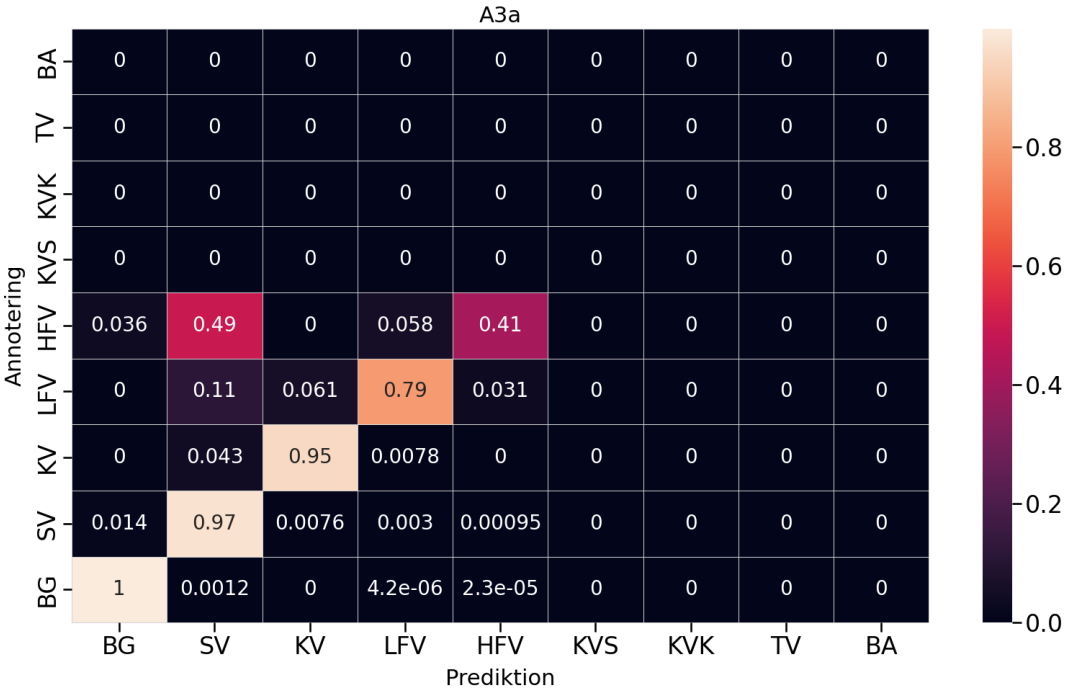
Figur 22. Normaliserad sammanblandningsmatris för modell A1e (A1a samt viktad Tversky-förlustfunktion), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



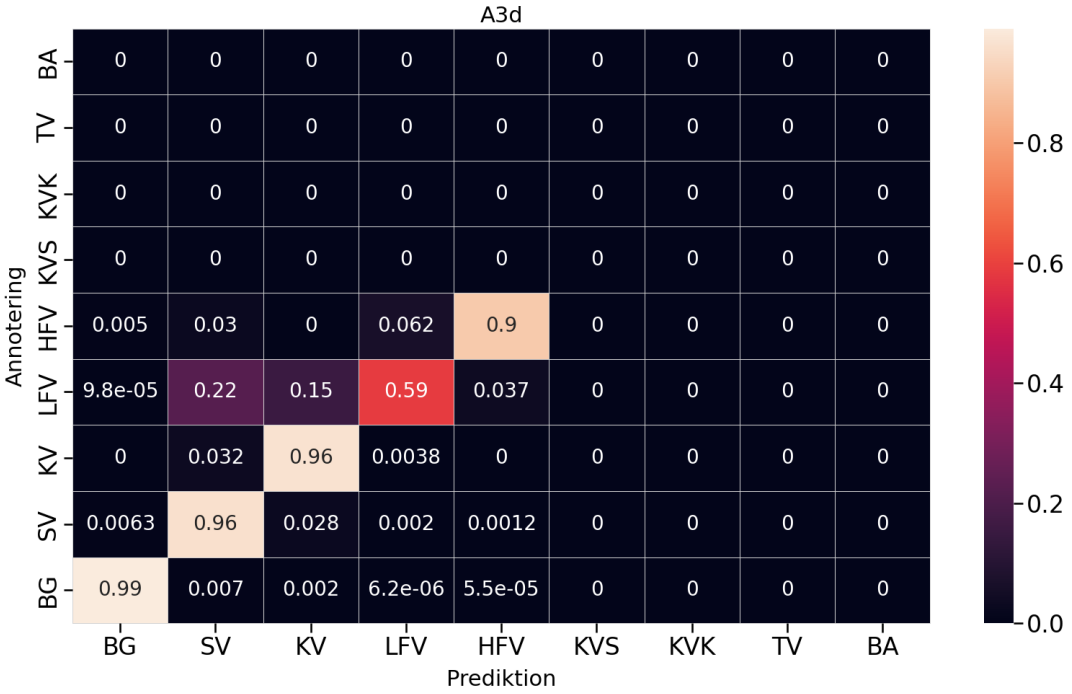
Figur 23. Normaliserad sammanblandningsmatris för modell A2a (fyra klasser; HFV hopbakad med LFV), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



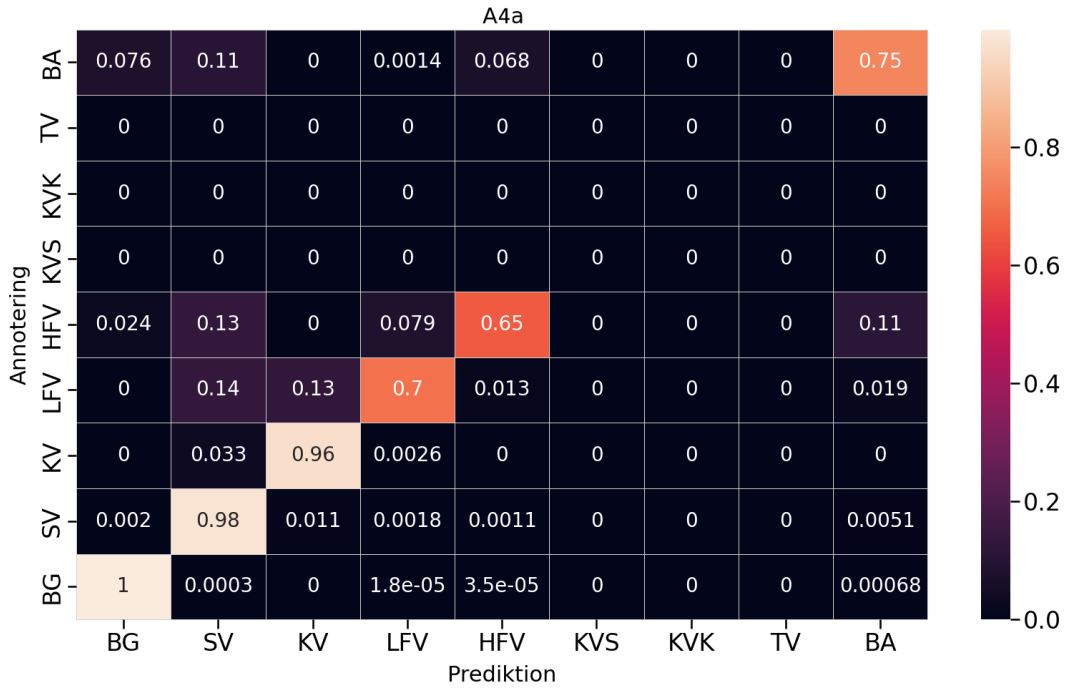
Figur 24. Normaliserad sammanblandningsmatris för modell A2d (fyra klasser; HFV hopbakad med LFV, samt viktad förlustfunktion), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



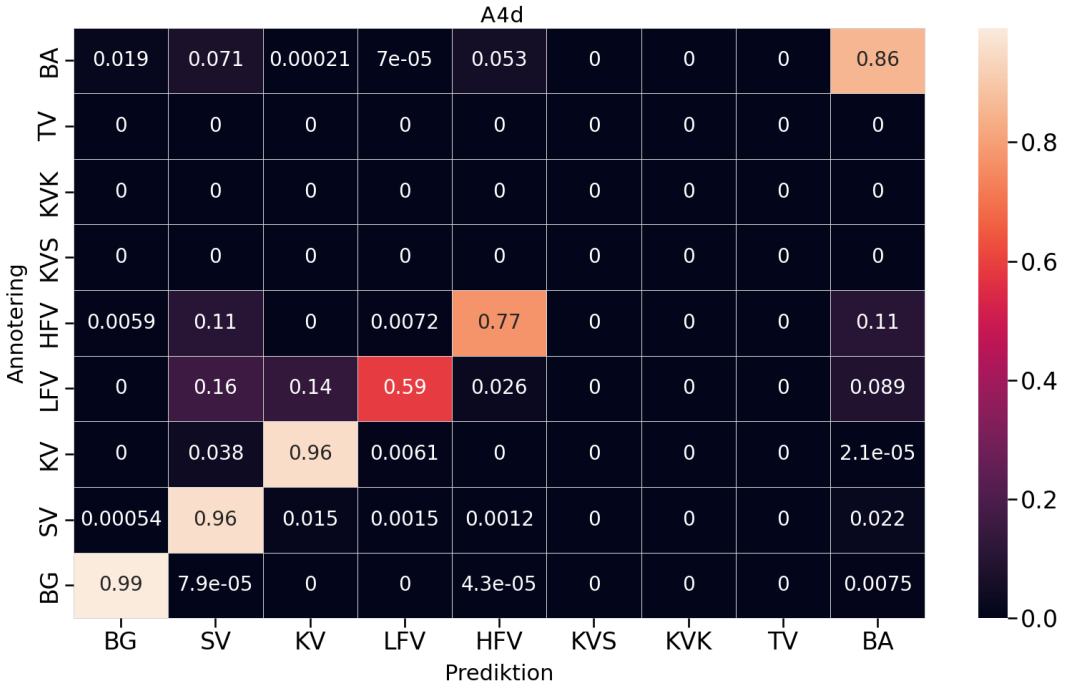
Figur 25. Normaliserad sammanblandningsmatris för modell A3a (fem klasser; grundannotering + HFV), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



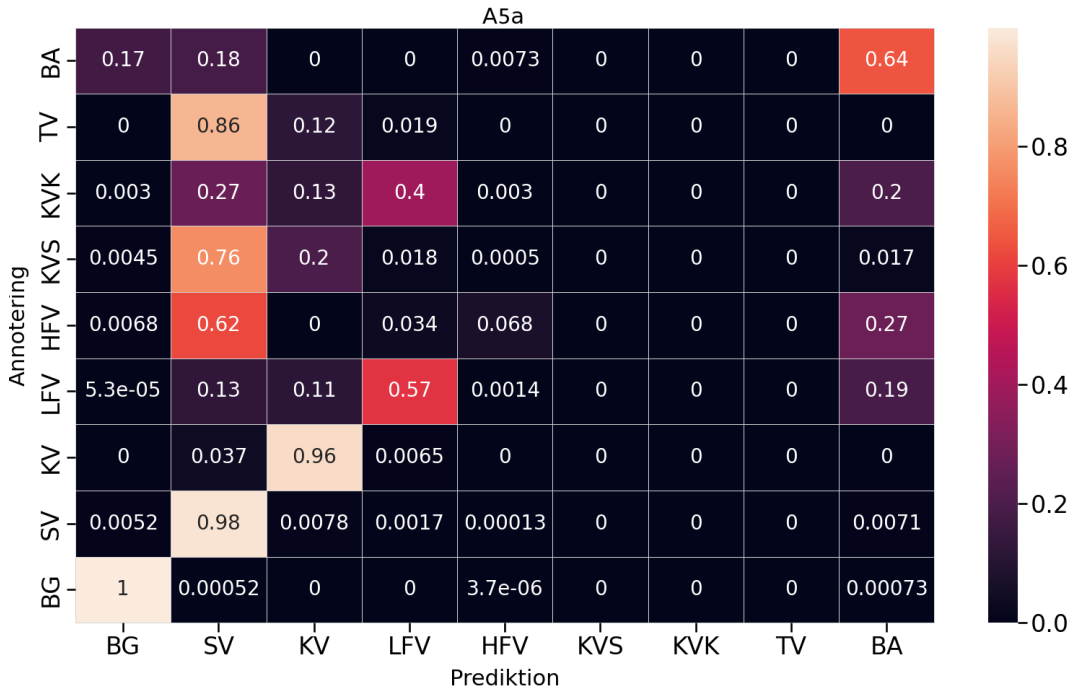
Figur 26. Normaliserad sammanblandningsmatris för modell A3d (fem klasser; grundannotering + HFV, viktad förlustfunktion), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



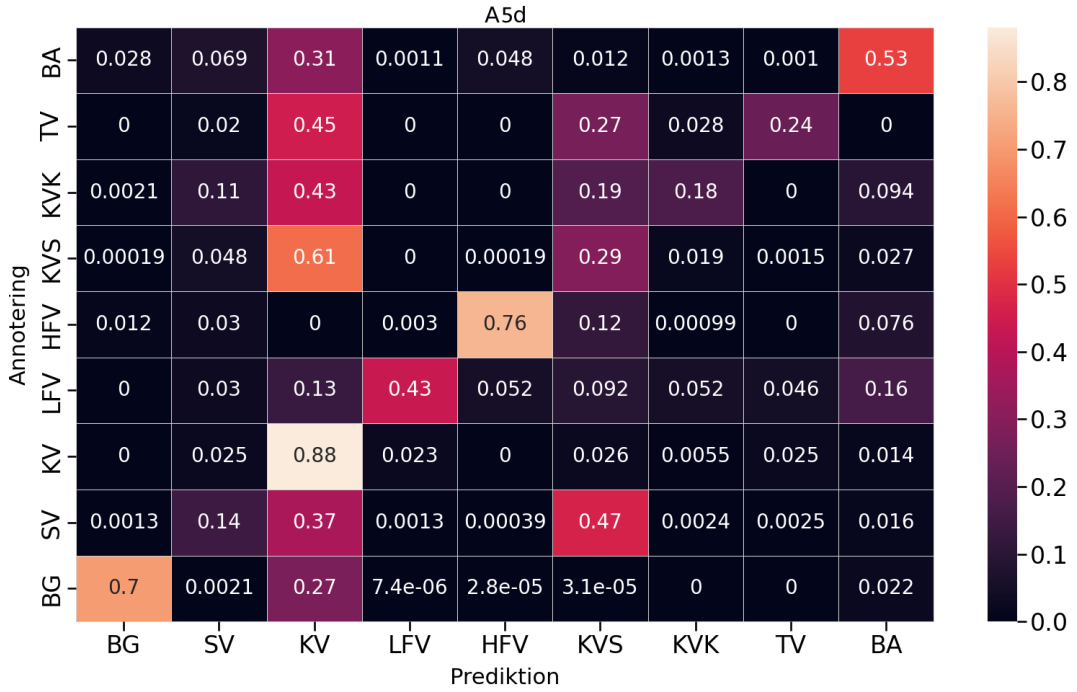
Figur 27. Normaliserad sammanblandningsmatris för modell A4a (sex klasser; grundannotering + HFV + BA), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



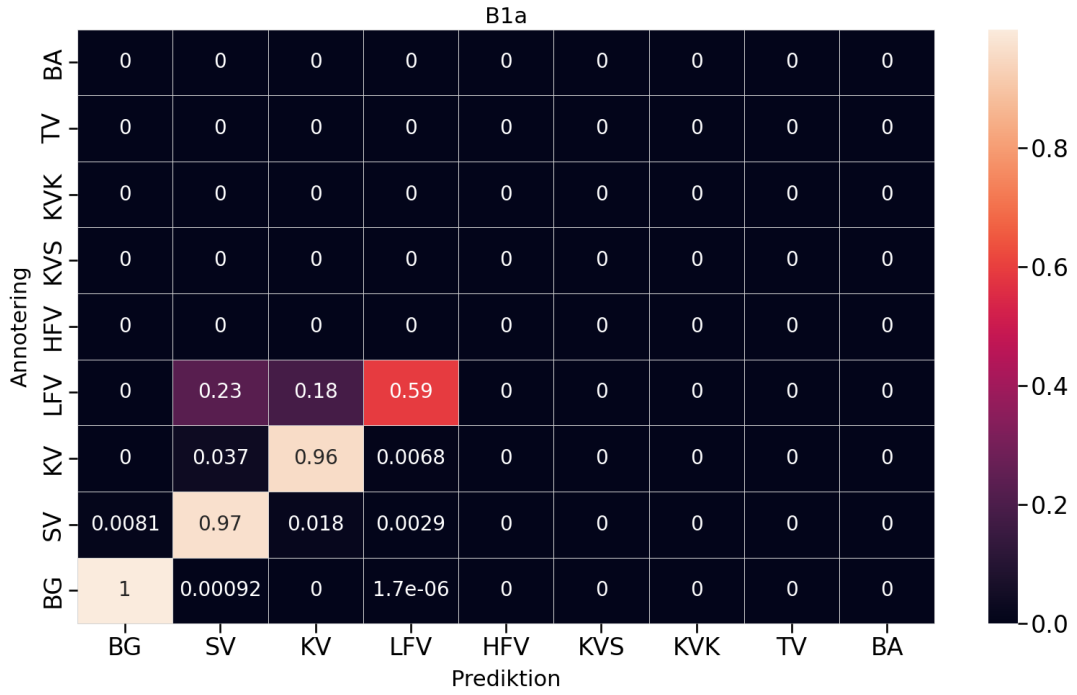
Figur 28. Normaliserad sammanblandningsmatris för modell A4d (sex klasser; grundannotering + HFV + BA, viktad förlustfunktion), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



Figur 29. Normaliserad sammanblandningsmatris för modell A5a (alla nio klasser), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



Figur 30. Normaliserad sammanblandningsmatris för modell A5d (alla nio klasser, viktad förlustfunktion), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.



Figur 31. Normaliserad sammanblandningsmatris för modell B1a (grundannotering men med kärnveden annoterad enligt årsringsprincipen), medelvärdesbildad över samtliga annoterade tvärsnitt. Förkortningarna av klassnamnen förklaras i Tabell 1.

Modell A1a är bäst på att prediktera LFV, med en rättklassningsgrad på 81 procent. Bland de modeller som inkluderar HFV är modell A3d bäst, med en rättklassningsgrad om 59 procent för LFV och 90 procent för HFV (Tabell 8, kolumnen näst längst till höger). Om större vikt läggs vid rättklassningen av LFV, som är betydligt vanligare klass än HFV, blir den bästa modellen i stället A4a (Tabell 8, kolumnen längst till höger).

Tabell 8. Andel korrekta klassträffar för fetved för modellvarianter som har både LFV och HFV, samt två kombinerade värden.

Modell	Variant	LFV	HFV	Medelvärde (LFV + HFV)	Viktat medelvärde (0,7 × LFV + 0,3 × HFV)
A3	a	0,79	0,41	0,60	0,67
	d	0,59	0,90	0,75	0,67
A4	a	0,70	0,65	0,68	0,69
	d	0,59	0,77	0,68	0,64
A5	a	0,57	0,07	0,29	0,42
	d	0,43	0,76	0,60	0,54

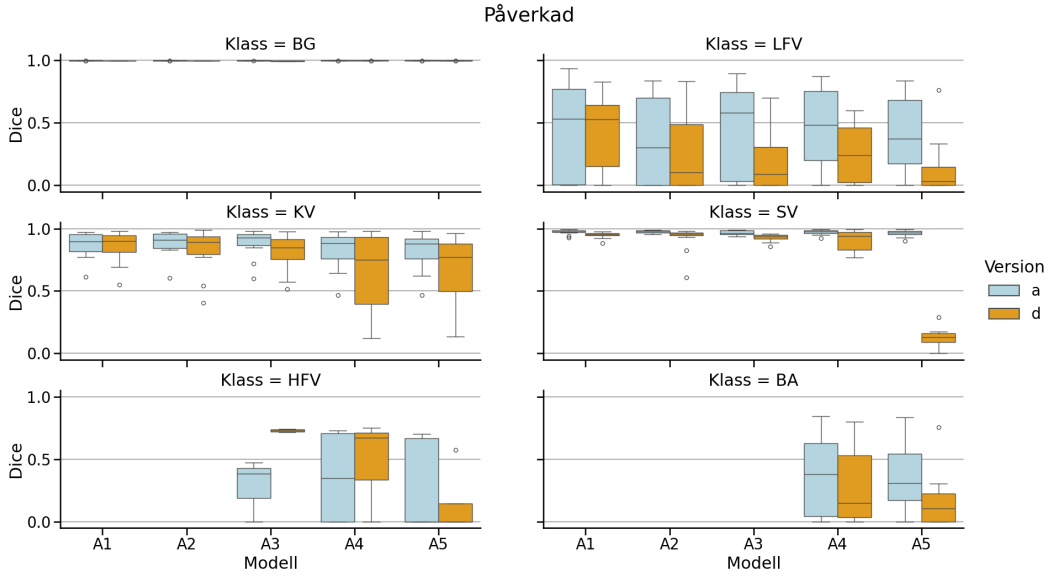
För de oviktade modellerna blir prediktionen av HFV sämre när ytterligare klasser inkluderas, jämför modell A5a med A4a. Detta kan delvis motverkas genom att vikta förlustfunktionen (d-modellerna). Viktning ger dock bäst resultat för om klasserna inte är för små, jämför modell A3b och A5d. När alltför små klasser viktas upp blir prediktionen av de större klasserna i stället sämre.

Bland modellerna med grundannotering är det noterbart att augmentering av indata (A1c) inte förbättrar prediktionen. En sammanhållen klass för LFV och HFV (A2a) förbättrar däremot prediktionen av LFV jämfört med grundannoteringen. Likaså blir resultatet bättre om förlustfunktionen Dice byts ut mot Tversky, och om förlustfunktionen viktas. Däremot blir modellen inte bättre om både Tversky och viktad förlustfunktion används (A1e). Den alternativa annoteringen av kärnved (B1a) försämrar klassningen av kärnved jämfört med grundannoteringen (A1a).

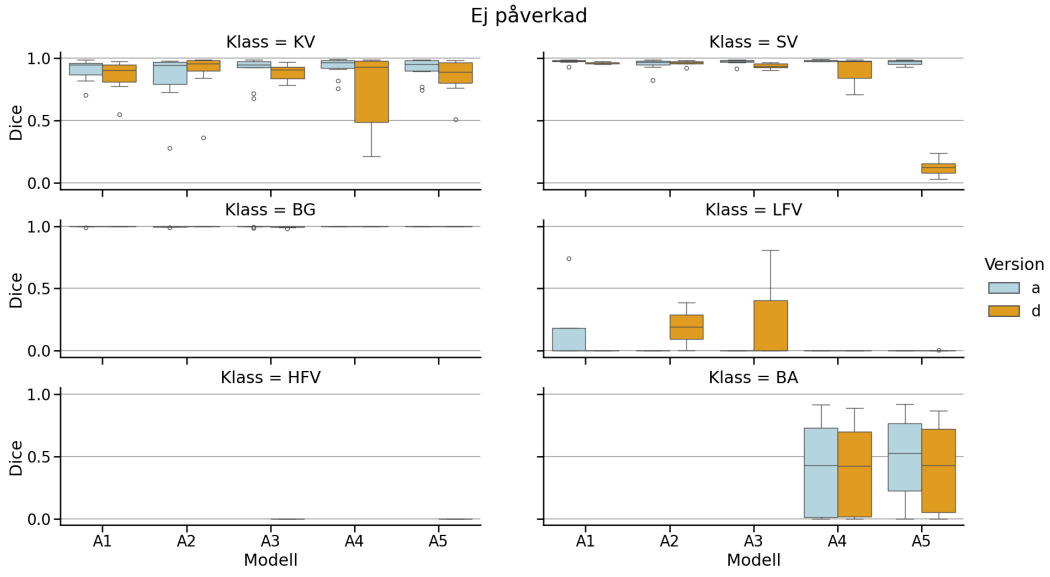
Ingen av de oviktade modellerna predikterade KVS eller KVK, sannolikt på grund av att pixelmängden var för liten.

Utvärderingsmåtten Dice och masscentrumsavstånd

Utvärderingsmåttet Dice visas för stockar påverkade av törskate (Figur 32) och för stockar opåverkade av törskate (Figur 33). Ju närmare värdet ett (1) som Dice ligger desto bättre överensstämmer annoteringen och prediktionen för den givna klassen. Notera att vissa värden ligger så nära noll eller ett att de knappt syns i figurerna.



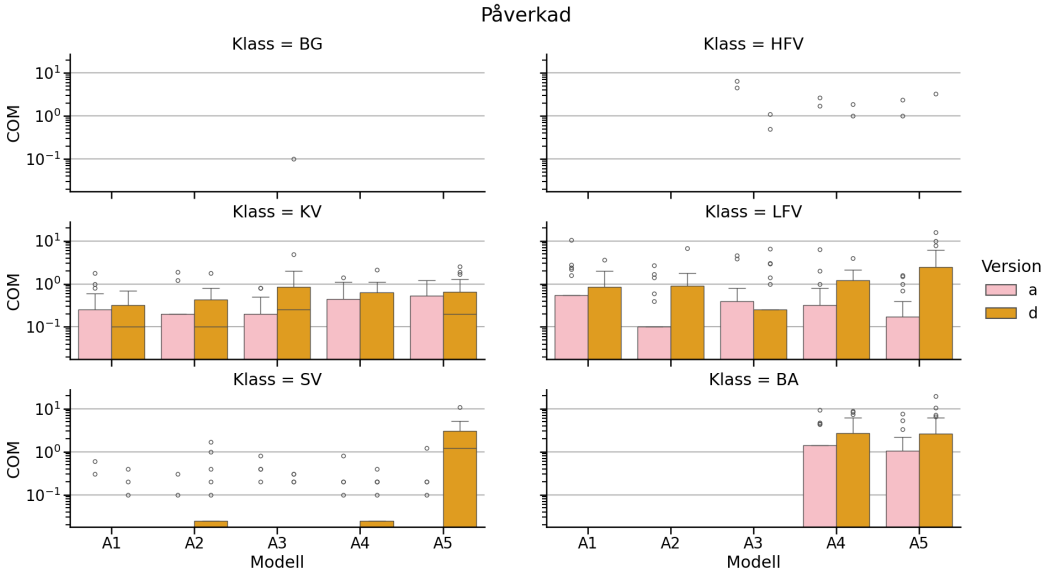
Figur 32. Värde på utvärderingsmåttet Dice för tvärsnitt från stockar påverkade av törskate, uppdelat på klass, modell och modellvariant (a eller d). Dice-värdet för varje stock är medelvärdet av de annoterade tvärsnitten. Felstaplarna är konfidensintervaller (95 %), medan cirklarna är outliers (värden bortom konfidensintervallet).



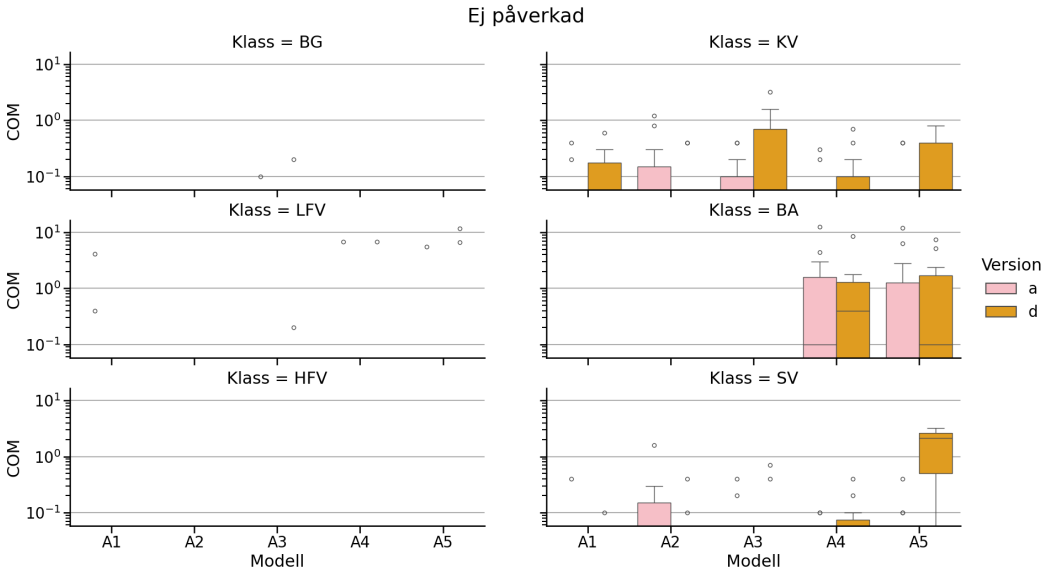
Figur 33. Värde på utvärderingsmåttet Dice för tvärsnitt från stockar opåverkade av törskate, uppdelat på klass, modell och modellvariant (a eller d). Dice-värdet för varje stock är medelvärdet av de annoterade tvärsnitten. Felstaplarna är konfidensintervaller (95 %), medan cirklarna är outliers (värden bortom konfidensintervallet). Notera att dessa stockar inte bör innehålla LFV eller HFV såvida inte dessa är felaktigt predikterade av modellen.

Modell A3a ger den bästa Dice-koefficienten för LFV, tätt följd av A4d. Modell A3d är bäst för HFV hos påverkade stockar, men presterar i gengäld dåligt för LFV på opåverkade stockar. Att Dice-koefficienterna för fetved är måttliga (kring 0,5) kan bero på att storleken på fetvedsområdena underskattas.

Masscentrumsavstånd (COM) visas i Figur 34 (påverkade stockar) respektive Figur 35 (opåverkade stockar). Ju lägre värde desto närmare ligger annoteringens och prediktionens masscentra varandra. Notera att y-axeln är logaritmerad och att vissa värden ligger så nära noll att de knappt syns i figurerna.



Figur 34. Värde på utvärderingsmåttet masscentrumsavstånd (pixlar) för stockar påverkade av törskate, uppdelat på klass, modell och modellversion. Värdet för varje stock är det logaritmerade medelvärdet av de annoterade tvärsnitten. Felstaplarna är konfidensintervaller (95 %), medan cirklarna är outliers (värden bortom konfidensintervallet).



Figur 35. Värde på utvärderingsmåttet masscentrumsavstånd (pixlar) för stockar opåverkade av törskate, uppdelat på klass, modell och modellversion. Värdet för varje stock är det logaritmerade medelvärdet av de annoterade tvärsnitten. Felstaplarna är konfidensintervaller (95 %), medan cirklarna är outliers (värden bortom konfidensintervallet).

Modellerna ger låga masscentrumsavstånd för de flesta klasser. Bortsett från modellerna A4d och A5d är skillnaden i masscentrumsavstånd mellan annoteringen och prediktionen är mindre än 2 pixlar (1 mm). Det visar att modellerna positionerar de predikterade pixlarna väl.

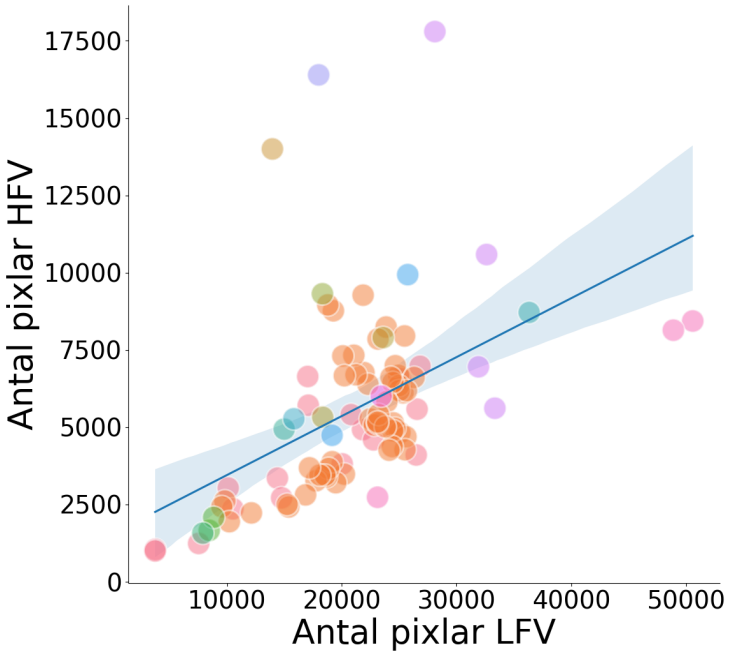
Missar modellerna fetved?

Hur mycket av fetveden detekterar modellerna? Tabell 9 nedan visar att modell A3a underskattar mängden LFV med 14 procent och mängden HFV med 32 procent, medan modell A4a underskattar mängden LFV med 19 procent och överskattar mängden HFV med 16 procent.

Tabell 9. Procentandel (%) felklassning av antalet pixlar baserat på de annoterade test-tvärsnitten, medelvärdesbildat över alla stockar och uttryckt per modell och klass. Ett positivt värde innebär att modellen underskattar pixelantalet, medan ett negativt värde innebär att modellen överskattar pixelantalet. Förklaring av förkortningarna finns i Tabell 1.

Modell	Version	BG	SV	KV	LFV	HFV	BA
A1	a	0,1	-1	-7	28		
	d	-0,1	-1	4	12		
A2	a	-0,1	1	-10	18		
	d	0,4	-7	4	32		
A3	a	-0,1	-1	-0,3	14	32	
	d	1	-5	-16	33	-47	
A4	a	-0,1	0,1	-2,4	18,6	-15,6	-11,4
	d	0,1	3,4	-24,7	62,4	10,34	-88,6
A5	a	-0,1	-2,4	-9,4	30,6	-7,4	6,2
	d	0,1	93,9	-105,3	12,5	32,1	-80,0

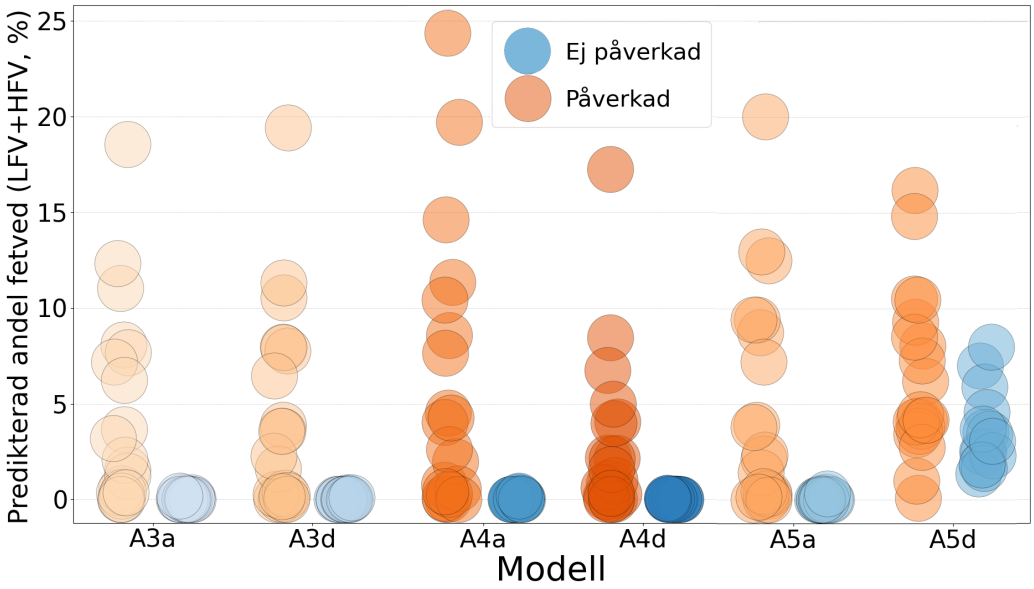
Som tidigare nämnt är en underklassning att föredra framför en överklassning (som kan leda till att opåverkade stockar vrakas). Mängden HFV är positivt korrelerad med mängden LFV, se Figur 36. HFV förekommer endast när det också finns mycket LFV. Att mängden HFV underskattas får därmed ingen betydelse för sortering. För postning av stockar kan det däremot få viss betydelse om HFV felaktigt klassas som splintved, och ett fetvedsområde därmed bevaras i sågningen i tron att det är normalved.



Figur 36. Antal pixlar som annoterats som LFV mot antalet pixlar som annoterats som HFV, från tvärsnitt som innehåller HFV. Den blå linjen visar en linjär anpassning, medan ljusblå zonen är ett konfidensintervall (95 %). Cirklarna är färgkodade efter vilken stock data kommer ifrån.

Utvärdering på stocknivå

En av målsättningarna är att modellerna ska prediktera en låg andel fetved i stockar opåverkade av fetved. Prediktionen av fetved på stocknivå för stockar med eller utan törskatepåverkan summeras i Figur 37. Här innebär ”påverkad stock” att stocken har minst ett tvärsnitt som annoterats med fetved, oavsett mängd fetved.



Figur 37. Stripdiagram över predikterad andel fetved per stock, fördelad på modell och uppdelad mellan med respektive utan törskatepåverkan (oavsett mängden fetved). Varje cirkel representerar en stock. Notera att värdet på andelen fetved ligger i cirklarnas mittpunkt. Färgernas mättnadsgrad har

endast syftet att särskilja modellerna från varandra. Parametern *dodge* sattes till True, vilket innebär att överlappande cirklar spreds ut genom en slumpmässig förskjutning i x-led.

Figur 37 visar att modellerna är olika känsliga i avseendet hur mycket fetved de predikterar, där modell A4a predikterar mest fetved och modell A4d minst. Alla modeller predikterade viss mängd fetved även hos stockar som bedömts vara opåverkade, men bortsett från modell A5a är den felpredikterade mängden fetved låg, se Tabell 10.

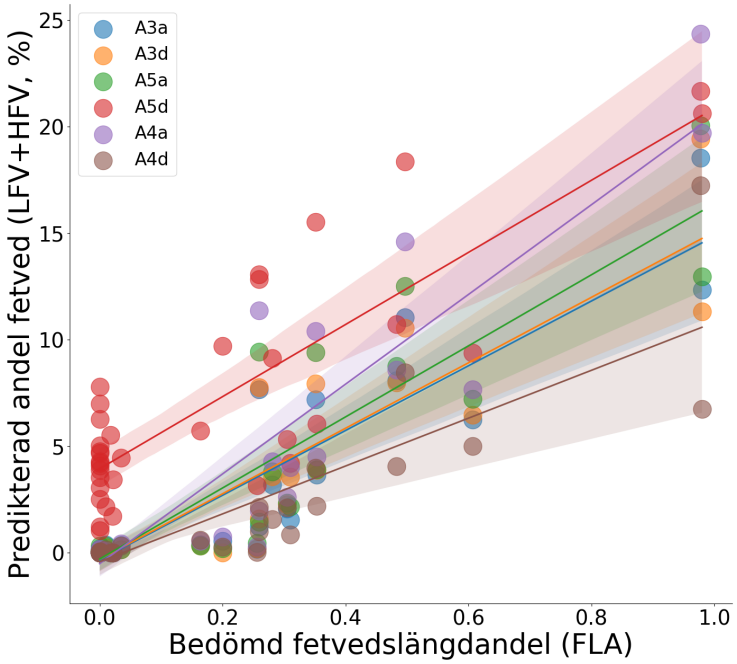
Tabell 10. Procentandel (%) predikterad fetved i hela stockar som bedömts vara opåverkade av törskate, fördelat på modell. Nyckeltalet beräknades baserat på antalet pixlar i klasserna LFV och HFV delat med totala antalet icke-bakgrundspixlar.

Modell	Version	Medelvärde	Minsta värde	Största värde
A1	a	0,07	0,01	0,42
	d	0,04	0	0,21
A2	a	0,10	0	0,83
	d	0,01	0	0,02
A3	a	0,03	0	0,20
	d	0,04	0	0,16
A4	a	0,03	0	0,18
	d	0,02	0	0.06
A5	a	0,05	0	0,34
	d	4,23	1,05	7,79

De flesta modellerna predikterar mindre än 0,1 procent fetved hos opåverkade stockar. Om algoritmen skulle användas vid automatisk kvalitetsbedömning på stocknivå kan detta därför vara en bra tröskel, där stockar med högre andel fetved sorteras ut och stockar med lägre andel accepteras som sågbara.

Jämförelse mellan prediktion och referensmått på stocknivå

På stocknivå är det svårare att jämföra det predikterade resultatet med ett referensmått. Därför togs referensmåttet fetvedslängdandel (FLA) fram, som summerar längden och allvarlighetsgraden hos stockarnas fetved. Den totala andelen predikterad fetved i en stock visas mot värdet på FLA i Figur 38.



Figur 38. Predikterad andel fetved (%) vs. fetvedslängdandel på stocknivå, färgkodat efter modell.

Den predikterade fetvedsandelen korrelerar med den visuellt bedömda fetvedslängdandelen, förutom för modell A5d. Det tyder på att den predikterade fetvedsandelen är användbar för att utvärdera hur allvarlig förekomsten av törskaterelaterad fetved är i stockar.

Slutsats

Sammantaget visar utvärderingen på tvärsnittsnivå att modellerna har olika styrkor och svagheter. Eftersom LFV har högre antal pixlar än HFV (se Tabell 4) är det viktigare att LFV predikteras väl än HFV. Det är även bättre att modellen underskattar mängden fetved än överskattar den. Slutsatsen blir därför att modell A3a är bäst lämpad för uppgiften i denna studie.

Diskussion

Helhetsintryck

Den bästa modellen, A3a, uppvisade viss underskattning av mängden LFV och betydande underskattning av mängden HFV. Inte minst perifer fetved (i stort sett alltid LFV) underskattades. Samtidigt överensstämde positionen (masscentrumavståndet) hos den fetved som faktiskt predikterade väl med den annoterade. Eftersom ett mål med modellen var att den skulle vara konservativ får denna målsättning sägas vara uppfylld.

Annotering

Annoteringen av LFV och HFV baserar sig skillnader i densitet som är visuellt urskiljbara i CT-bilderna, i många fall kombinerat med närhet till uppmätta yttre törskaterelaterade stamsår. Förekomsten av fetved säger inget om lokal förekomst av levande törskatesvamp i veden. Det går inte heller att utesluta att delar av den annoterade ”fetveden” består av låg- eller högdensitetsskador (reaktionsved) med annan orsak än törskate. Denna andel bedöms dock vara relativt liten.

Mängden träningsdata

Modellen som redovisas i denna rapport är annoterad och tränad på relativt få tvärsnitt, vilket sannolikt har stor inverkan på resultaten. Den modesta mängden träningsdata beror på den stora tidsåtgången för annotering. En möjlig väg framåt är att annotera bilderna i 3D, exempelvis med mjukvaran 3D Slicer, och anpassa modellen att kunna tränas på 3D-data. Annotering i 3D skulle också ge modellen mycket mer data utan den stora arbetsinsats som krävs vid 2D-annotering, vilket troligen skulle ge förbättrad prediktion. Det skulle även möjliggöra att utvärdera prediktionens spatiala noggrannhet på stocknivå, vilket hade varit önskvärt för att bedöma tillämpbarhet vid postning i sågverk.

Bildförminskning

Av resurstekniska skäl förminskades CT-bilderna från 900 × 900 pixlar till 256 × 256 pixlar inför träningen av modellen. Sannolikt hade resultaten förbättrats något om de i stället förminskats till 512 × 512 pixlar. Denna ändring kan enkelt göras om mer processorkraft finns att tillgå.

Tidseffektivisering

Den tränade modellen tog i snitt 10 minuter på sig att prediktera på varje tvärsnitt i en stock, se Tabell 11. Genom att endast använda var 10:e tvärsnitt (varannan centimeter) kortas tiden ner till 1,2 minuter utan nämnvärda konsekvenser för klassning på stocknivå. Detta beror på att fetveden förändras långsamt i stockens längdriktning. Genom att endast använda var 20:e tvärsnitt (varje centimeter) kortades tiden ner till 45 sekunder, och om prediktionen gjordes var 200:e tvärsnitt (varje decimeter, motsvarande diametermätningen i skördare) gick processeringen på 3 sekunder. Sannolikt kan prediktionstiden kortas ytterligare genom att använda en kraftfullare processor eller parallellprocessering. Filhanteringen och efterbehandlingsstegen är kodade i Python-ramverket *dask* som är utvecklat för parallellprocessering.

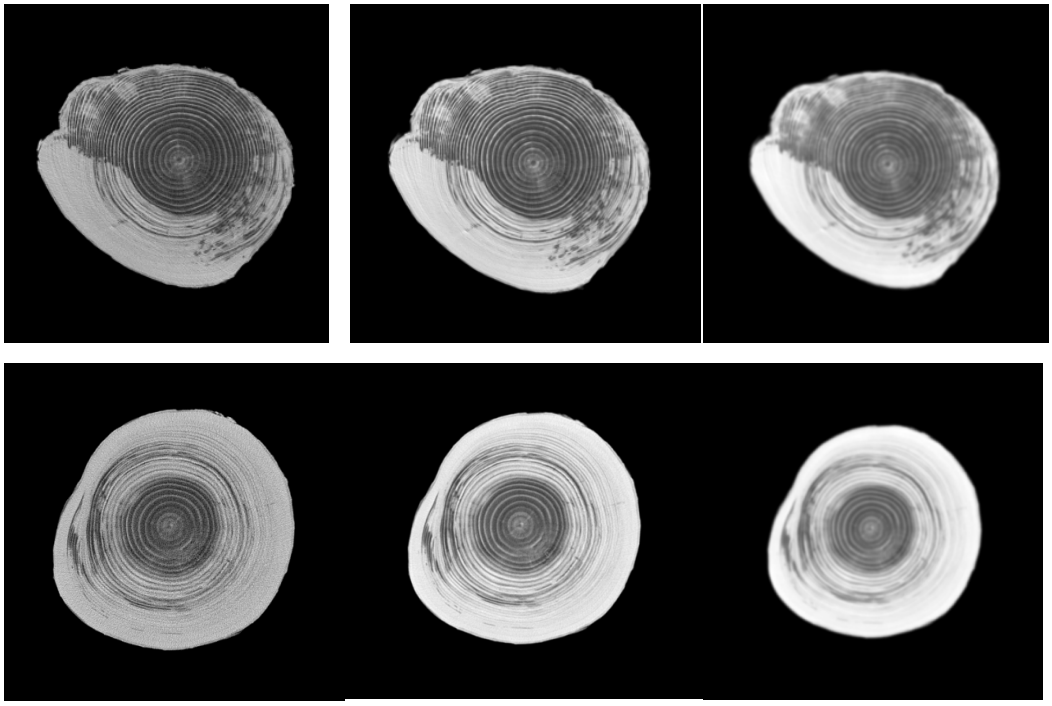
Tabell 11. Tidsanalys över olika steg i prediktionen och efterbehandlingen.

Operation	Genomsnittlig tid per operationssteg (s)
Ladda in stockdata i minnet (från zarr-fil)	13
Förminska stocken till 256x256 pixlar	8
Prediktera på varje tvärsnitt i stocken	510
Addera saknad kärnved i klena stockar	1
Fylla igen hål i den predikterade kärnveden	60
Ta bort bark i bakgrundspixlar	1
Spara resultatet (till zarr-fil)	14
Totalt	602 sekunder (10 minuter)

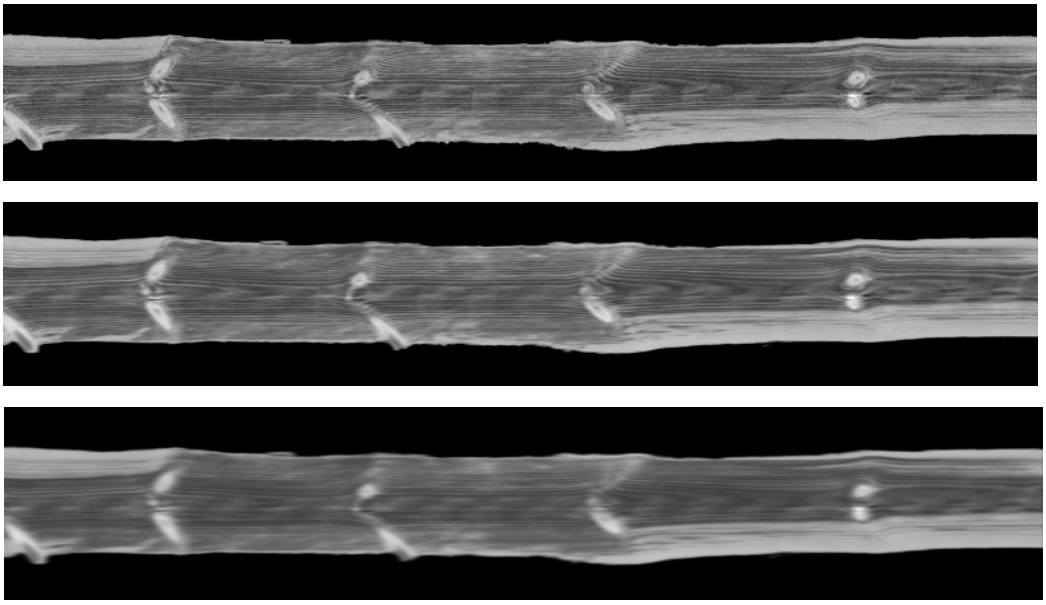
Tillämpbarhet på industriella mätramar

Industriella CT-skannrar har goda förutsättningar att detektera fetved till följd av törskateinfektion, då skillnaden i upplösning jämfört med laboratoriebaserad CT-skanner är liten.

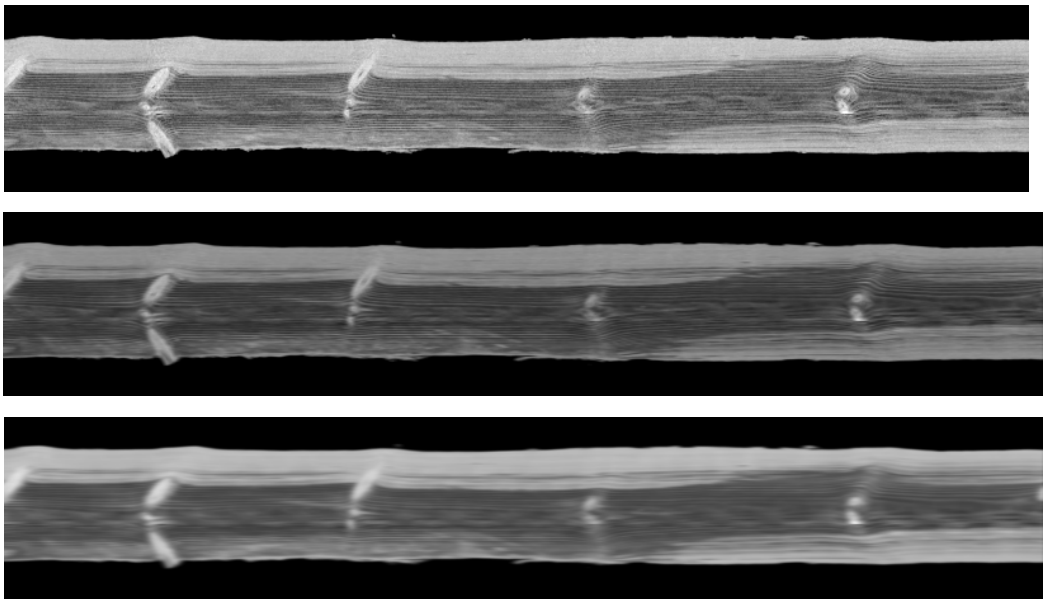
Diskreta röntgenmätramar kan troligen detektera LFV. Det verkar däremot svårt att skilja mellan LFV och HFV. Att detektera högdensitetsfetved är inte nödvändigt för sortering på stocknivå. Frågan som kvarstår är därför om detektion av HFV är nödvändigt för postning av stocken.



Figur 39. Tvärsnitt (i xy-riktning, se Figur 1) från laboratoriebaserad CT-skanning i originalupplösning (vänster kolumn), medelvärdesfiltrerad (mittenkolumnen) och gaussfiltrerad (höger kolumn), för att efterlikna industriell CT-skanner.



Figur 40. Tvärsnitt (i xz-riktning, se Figur 1) från laboratoriebaserad CT-skanning i originalupplösning (överst), medelvärdesfiltrerad (mitten) och gaussfiltrerad (underst), för att efterlikna industriell CT-skanner.



Figur 41. Tvärsnitt (yz-riktning, se Figur 1) från laboratoriebaserad CT-skanning i originalupplösning (överst), medelvärdesfiltrerad (mitten) och gaussfiltrerad (underst), för att efterlikna industriell CT-skanner.

Simulering av industriella röntgenmätramar

Existerande mjukvaror för simulering av industriella röntgenmätramar baserat på laboratoriebaserad CT-data är utdaterade eller företagsinterna. Det finns ett stort utvecklingsbehov av ett simulationsprogram som kan användas för forskning.

Vidare utveckling

För att öka modellens generaliserbarhet vore det önskvärt att träna den med törskatepåverkade stockar från andra geografier, samt på fler opåverkade tallstockar, exempelvis från Luleå Tekniska Universitets stambank med CT-data. En ytterligare möjlighet är att ta fram olika modeller för olika diameterklasser eller stocknummer (förstastock, andrastock, etc.). Slutligen kan det vara intressant att ta fram en modell för stockar som är mindre färska och mer uttorkade än de som användes till denna studie.

Den spatiala sannolikhetsfördelningen för vissa av klasserna är känd på förhand. Till exempel så minskar sannolikheten för kärnved med avstånd från märgen ut mot mantelytan, medan sannolikheten för fetved minskar med avståndet från mantelytan in mot märgen. Sannolikheten för bark minskar skarpt med avståndet från mantelytan. Att kunna nyttja dessa kända sannolikheter skulle kunna förbättra modellen. En ytterligare utvecklingsmöjlighet är att utöka stocknivåklassificeraren med en modell som klassar stocken baserat på de olika modellernas utgångar och som i sig själv kan tränas. Den övergripande modellen kan potentiellt lära sig att olika delar av stocken är olika viktiga för att sortera stocken vidare på sågverket.

Tillgängliggörande av data

Dataset med annoterade stocktvärsnitt är ovanliga. En ambition är därför att göra de annoterade bilderna av såväl vedtyper som märgkoordinater (se Appendix) tillgängliga för andra forskare.

Referenser

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. & Ronneberger, O. 2016. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2016.

Diakogiannis, F. I., Waldner, F., Caccetta, P. & Wu, C. 2019. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. ArXiv, abs/1904.00592.

Google. 2025. Vertex AI - Prepare image training data for classification [Online].

Haghighi, S., Jasemi, M., Hessabi, S. & Zolanvari, A. 2018. PyCM: Multiclass confusion matrix library in Python. Journal of Open Source Software, 3**,** 729-.

Hyll, K., Joevenller, S., Svennerstam, H., Nordström, M., Broman, O., Oja, J. & Sandberg, D. 2022a. CT-skanning som verktyg för detektering av törskateangrepp på tall. Skogforsk Arbetsrapport 1126–2022. 1-37 s.

Hyll, K., Joevenller, S., Svennerstam, H., Nordström, M., Broman, O., Oja, J. & Sandberg, D. 2022b. X-ray computed tomography for the detection of damage in Scots pine trunks caused by blister-rust fungus Cronartium pini (Willd.). Wood Material Science & Engineering**,** 1-3.

Hyll, K., Joevenller, S., Svennerstam, H., Nysjö, F., Broman, O., Oja, J. & Sandberg, D. 2024. Bildanalysmetod för att i CT-bilder kvantifiera fetved och andra skador relaterade till törskateangrepp hos tall. Arbetsrapport 1189-2024, Skogforsk. 41 s.

Meehan, M. 2018. Semantic Segmentation of Small Data using Keras on an Azure Deep Learning Virtual Machine. Dev Blogs, [Online].

Öhman, M., Nyström, J. & Oja, J. 2005. Prediction of compression wood in sawn products of Picea abies by 3-D sequential outer shape medelling of the log. 5th workshop "Connection between forest resources and wood quality". Waiheke Island, Auckland, New Zealand.

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y. & Kainz, B. 2018. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F., eds. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 2015// 2015 Cham. Springer International Publishing, 234-241.

Schouten, J. P. E., Matek, C., Jacobs, L. F. P., Buck, M. C., Bošnački, D. & Marr, C. 2021. Tens of images can suffice to train neural networks for malignant leukocyte detection. Scientific Reports, 11**,** 7995.

Shahinfar, S., Meek, P. & Falzon, G. 2020. “How many images do I need?” Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. Ecological Informatics, 57**,** 101085.

Shorten, C. & Khoshgoftaar, T. M. 2019. A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6**,** 60.

Skog, J. 2013. Charaterization of sawlogs using industrial X-ray and 3D scanning. Doctoral thesis, Luleå Universitet, 1-162 s.

Skog, J. & Oja, J. 2007. Improved log sorting combining X-ray and 3D scanning - A preliminary study. In: Grzeskiewicz, M. (ed.) Quality Control for Wood and Wood Products. Warsaw, Poland: Faculty of Wood Technology, Warsaw University of Life Science.

Träguiden. 2017. Densitet - definitioner. Svenskt Trä. URL: https://www.traguiden.se/om-tra/materialet-tra/traets-egenskaper-och-kvalitet/densitet1/definitioner/ [Hämtad 2025-03-27].

Ursella, E. 2021. In-line industrial computed tomography applications and developments. PhD thesis, Ca'Foscari Venezia University, 101 s.

Zunair, H., Rahman, A., Mohammed, N. & Cohen, J. P. Uniformizing techniques to process CT scans with 3D CNNs for tuberculosis prediction. Predictive

Intelligence in Medicine: Third International Workshop, PRIME 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 8, 2020, Proceedings 3, 2020. Springer, 156-168.

Appendix - Relation mellan formparametrar och törskatepåverkan

Bakgrund

I en delstudie av detta projekt undersöktes relationen mellan inre och yttre formparametrar samt inre och yttre törskaterelaterad påverkan. Tidigare studier har funnit en korrelation mellan stockens yttre form och förekomst av törskatepåverkan i form av stamsår eller fetved (Öhman m.fl. 2005).

För att åstadkomma hanterbar beräkningskapacitet gjordes analysen på var 200:e tvärsnitt, vilket motsvarar ett värde per decimeter i stockens längdriktning. Detta är jämförbart med skördarens mätning av stockens längd och diameter, vilken också sker varje decimeter.

Metod

De undersökta parametrarna listas i Tabell 12. Törskatemåtten togs fram genom okulär inspektion av stockens utsida eller i CT-bilder där varje tvärsnitt klassades som opåverkad (värde noll) eller med förekomst av påverkan (värde ett).

Tabell 12. Undersökta törskatemått och formparametrar. Formparametrarna har kontinuerliga värden medan törskatemåtten har binära klasser (noll eller ett för opåverkad respektive påverkad).

Törskatemått	Formparametrar
Yttre stamsår	Cirkularitet
Inre perifer fetved	Excentricitet
Inre central fetved	Märgförskjutning

För att beräkna formparametrarna cirkularitet och excentricitet användes Python-funktionen *skimage.measure.regionprops_table* på tvärsnittets binära mask. Funktionen gav måtten storaxelns längd (S, enhet pixlar), lillaxelns längd (L, pixlar), omkrets (P, pixlar), area (A, pixlar²), excentricitet (E) och geometrisk centrumkoordinat (x_c, y_c). Cirkulariteten beräknades sedan som:

$$C = \frac{4\pi A}{P^2}$$

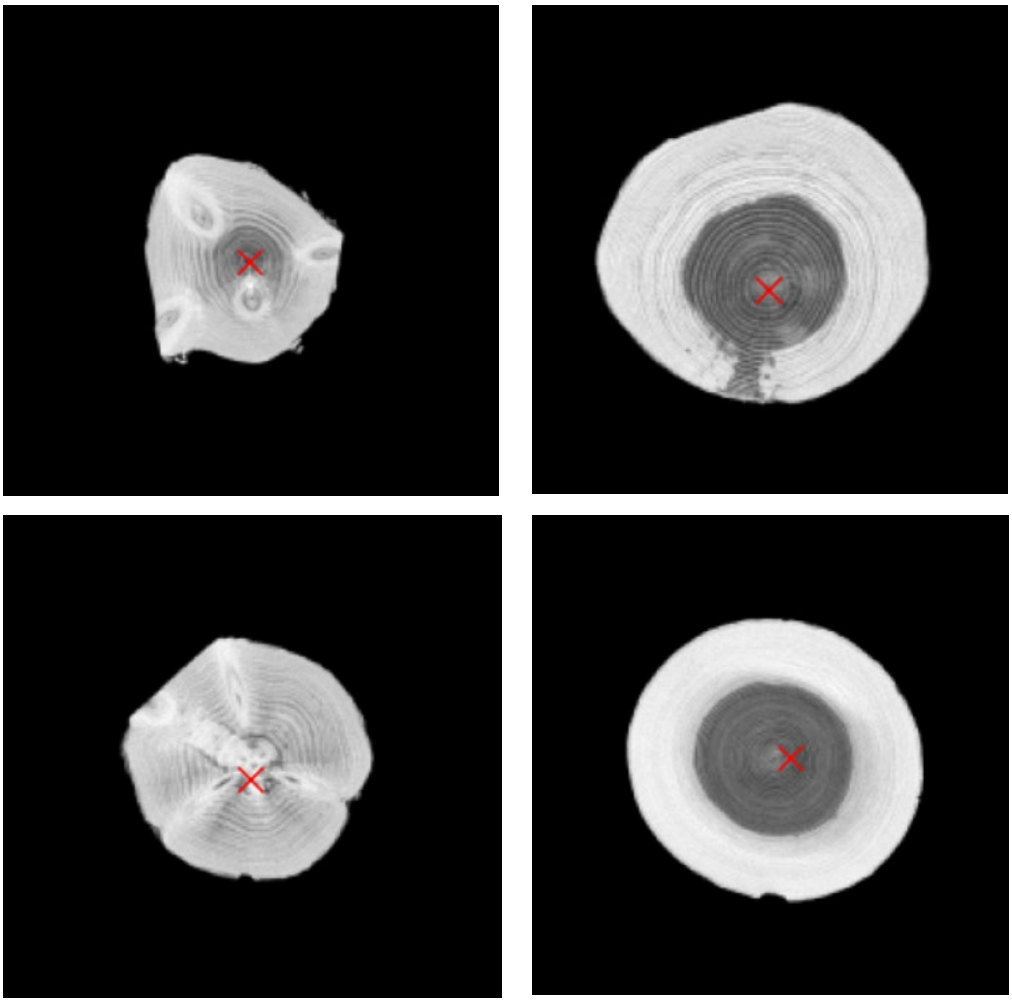
där A är arean och P är omkretsen. Värdet noll på cirkularitet innebär en helt cirkulär form och värdet ett en kvadratisk. Excentricitet definieras som:

$$E = \frac{\sqrt{S^2 - L^2}}{S}$$

Värdet noll på excentricitet innebär en cirkel och värdet ett innebär en linje.

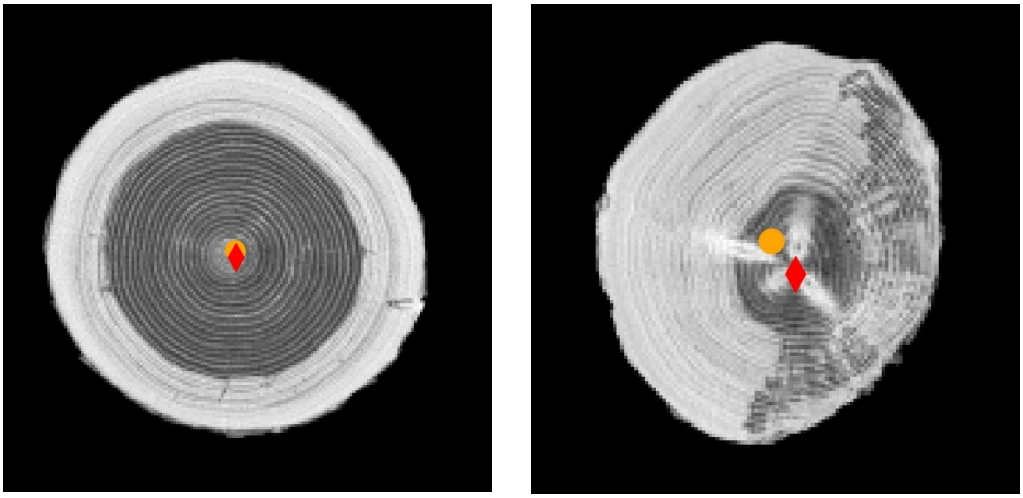
För att beräkna märgförskjutning relativt tvärsnittets geografiska mittpunkt behövdes en metod för att detektera och koordinatbestämma märgen. För detta tränades det neurala nätverket av typen *ResNet50* med 2257 st. annoterade tvärsnitt (60 st. från varje stock i datasetet). Annoteringen skedde med verktyget CVAT och det neurala nätverket

implementerades med Pytorch. Före träning beskars CT-bilderna från 900 × 900 pixlar till 756 × 756 pixlar. Därefter krymptes bilderna till 256 × 256 pixlar (*cv2.resize*, interpolering=*cv2.INTER_AREA*) för att matcha bildstorlekskravet hos ResNet50. 70 procent av bilderna användes för träning, 20 procent för validering och 10 procent för testning, och genererades grupperat på stocknivå så att tvärsnitt från samma stock hamnade i samma dataset (träning/validering/testning). Därefter augmenterades bilderna med slumpmässig rotation och spegelvändning. Medelkvadratfel (MSE) användes som förlustfunktion och modellen tränades under 1000 epoker. Förlusten uppgick till 0,60 vid träning, 0,91 vid validering och 0,62 vid testning. Omräknat till pixlar blir detta ett medelkvadratfel på 2,28 pixlar vid träning, 2,82 pixlar vid validering och 2,32 pixlar vid testning. Resultatet demonstreras i Figur 42.



Figur 42. Exempel på märgprediktioner (rött kryss) som har hög precision (övre raden), vilket är typfallet, samt märgprediktioner med lägre överensstämmelse (nedre raden).

För att få ett mått på ocentrerad märg beräknades det euklidiska avståndet mellan den av modellen predikterade märgkoordinaten och den tidigare beräknade geometriska centrumkoordinaten.



Figur 43.Exempel på masscentrumskoordinat (orange cirkel) och predikterad märgkoordinat (röd diamant).

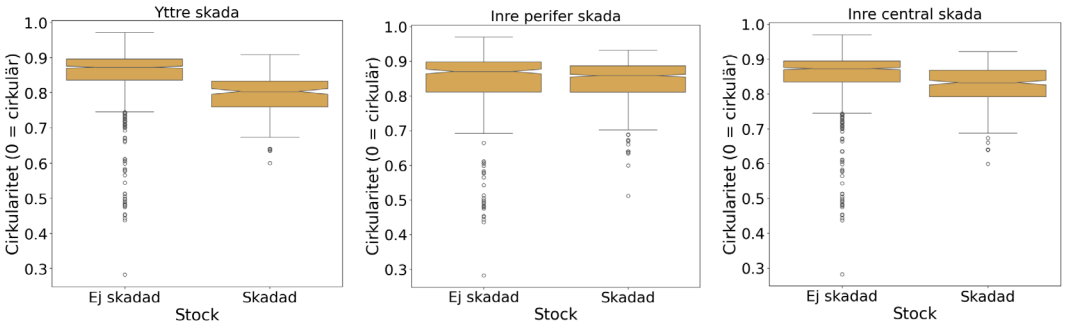
Värdena på de olika skademåtten standardiserades. För att utvärdera relationen mellan formparametrar och skademått anpassades en generell linjär modell (GLM) modell av typen Mixed Binomial Bayes (*statsmodels.BinomialBayesMixedGLM*) för varje formparameter, med skademåtten som fixa variabler. Vilken formparameter som hade den bästa förklaringsgraden utvärderades genom att jämföra absolutvärdet hos deras modellkoefficienter (Posterior Mean, μ_{FE}). Statistisk signifikans utvärderades genom konfidensintervaller och standardpoäng (z-score). 2σ-konfidensintervaller beräknades genom:

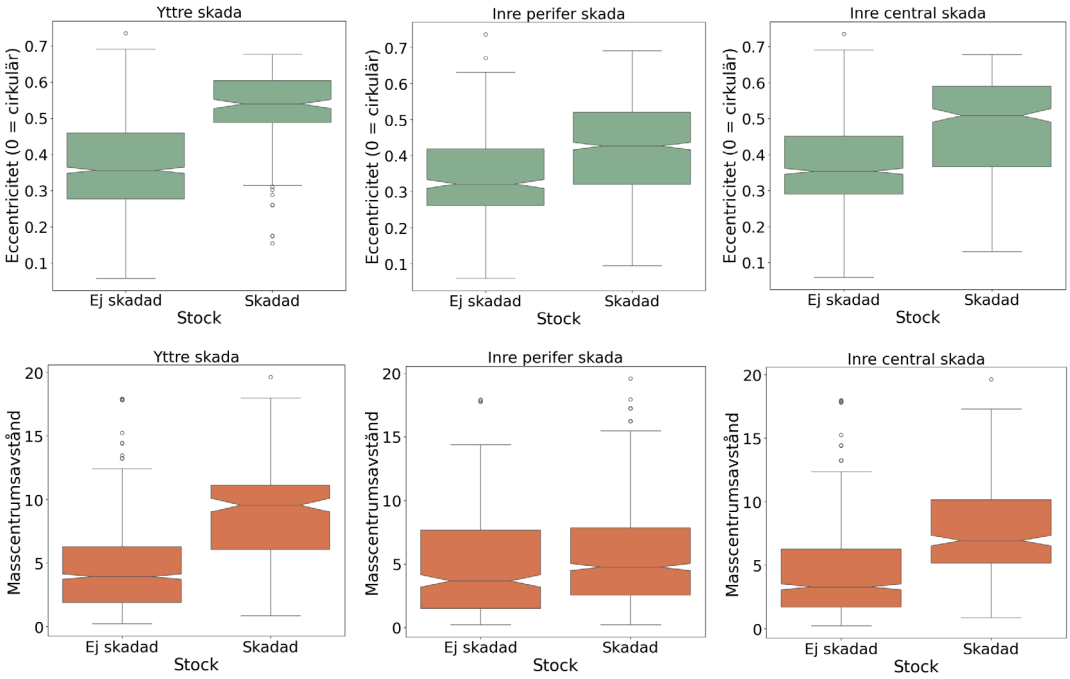
$$[\,\mu_{FE} - 1{,}96 \cdot \sigma_{FE}\,,\mu_{FE} + 1{,}96 \cdot \sigma_{FE}\,]$$

där μ_{FE}är medelvärdet av formparametrarnas modellkoefficienter och σ_{FE}är motsvarande standardavvikelse. Ett bra konfidensintervall bör inte inkludera värdet noll.

Resultat

Förhållandet mellan skademåtten och formparametrarna visualiseras i Figur 44.





Figur 44. Förekomst av påverkan för olika törskatemått (yttre, inre perifer, inre central) gentemot olika formparametrar (cirkularitet, excentricitet, märgförskjutning). Insnitten kring medianstrecket representerar konfidensintervall (95 %). Skillnaden mellan påverkad och opåverkad stock är signifikant om intervallen för insnitten i boxarna ej överlappar varandra.

Tabell 13. Värden på modellkoefficient, standardpoäng och p-värde för modellanpassningen mellan de olika formparametrarna och törskatemåtten.

Formparameter	**Skademått**	**Koefficient μ_{FE}**	**Konfidensintervall**
Cirkularitet	Yttre stamsår	-1,16	-1,36 : -0,95
	Inre perifer fetved	-0,09	-0,27 : 0,09
	Inre central fetved	-0,34	-0,51 : -0,16
Excentricitet	Yttre stamsår	1,43	1,24 : 1,61
	Inre perifer fetved	0,66	0,48 : 0,85
	Inre central fetved	1,09	0,93 : 1,24
Märgförskjutning	Yttre stamsår	1,17	1,01 : 1,35
	Inre perifer fetved	-0,48	-0,64 : -0,32
	Inre central fetved	0,70	0,53 : 0,86

Slutsats

De tre starkaste statistiska sambanden fanns mellan excentricitet och yttre stamsår (ju mer excentrisk tvärsnittsform desto troligare med stamsår), märgförskjutning och yttre stamsår (ju större märgförskjutning desto troligare med stamsår) och cirkularitet och yttre stamsår (ju mer oval tvärsnittsform desto troligare med stamsår). Sambanden mellan formparametrar och inre perifer fetved var svagare än sambanden mellan formparametrar och inre central fetved. Ett relativt starkt samband fanns mellan

excentricitet och inre central fetved. För relationen mellan cirkularitet och inre perifer fetved inkluderade konfidensintervallet värdet noll, vilket indikerar ett samband som ej är signifikant.