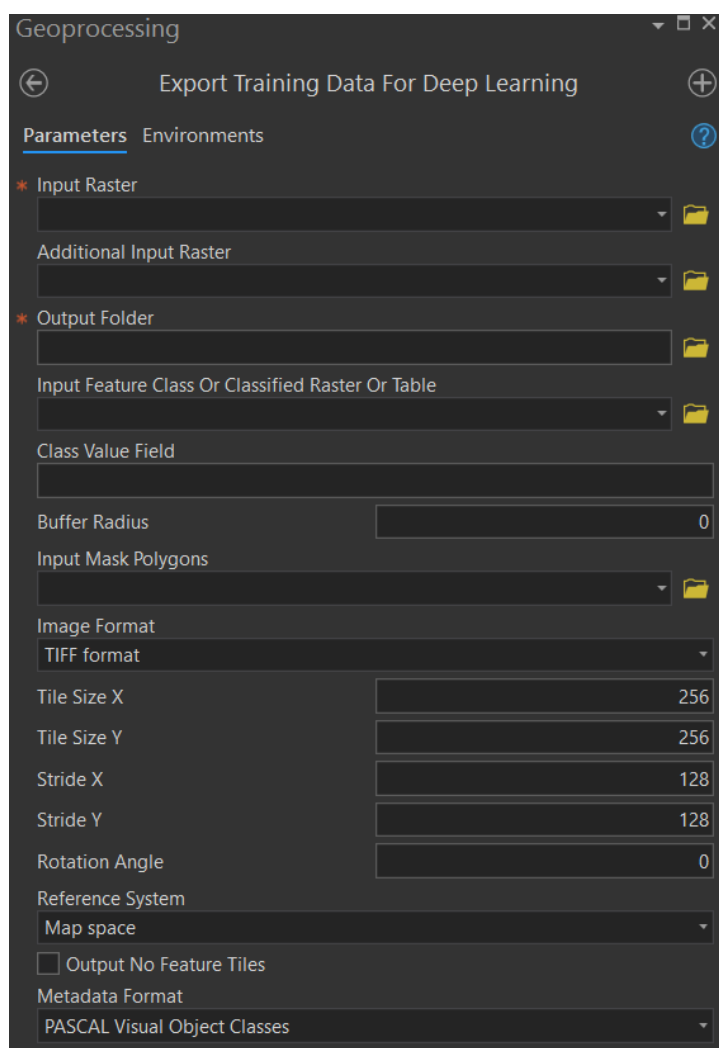


Vedlegg A – Dyplæringsverktøy i ArcGIS

Formålet med dette vedlegget er å redegjøre for de tekniske verktøyene som anvendes i oppgaven. Vedlegget er ment som en forklaring dersom leseren ønsker å etterse eller teste fremgangsmåten som er benyttet. Det er tre dyplæringsverktøy som beskrives, samt et verktøy som måler nøyaktigheten til deteksjonene.

For å lage treningsdata anvendes verktøyet *Export Training Data For Deep Learning*. Dette er det første dyplæringsverktøyet som benyttes på veien mot objektdeteksjon. Verktøyet lager såkalte *chips* som er mindre TIFF-bilder av rasteret med tilhørende informasjon. Det er disse chipsene som brukes videre i prosessen (ESRI, no date c). *Figur 1* viser hvilke parametere verktøyet inneholder.

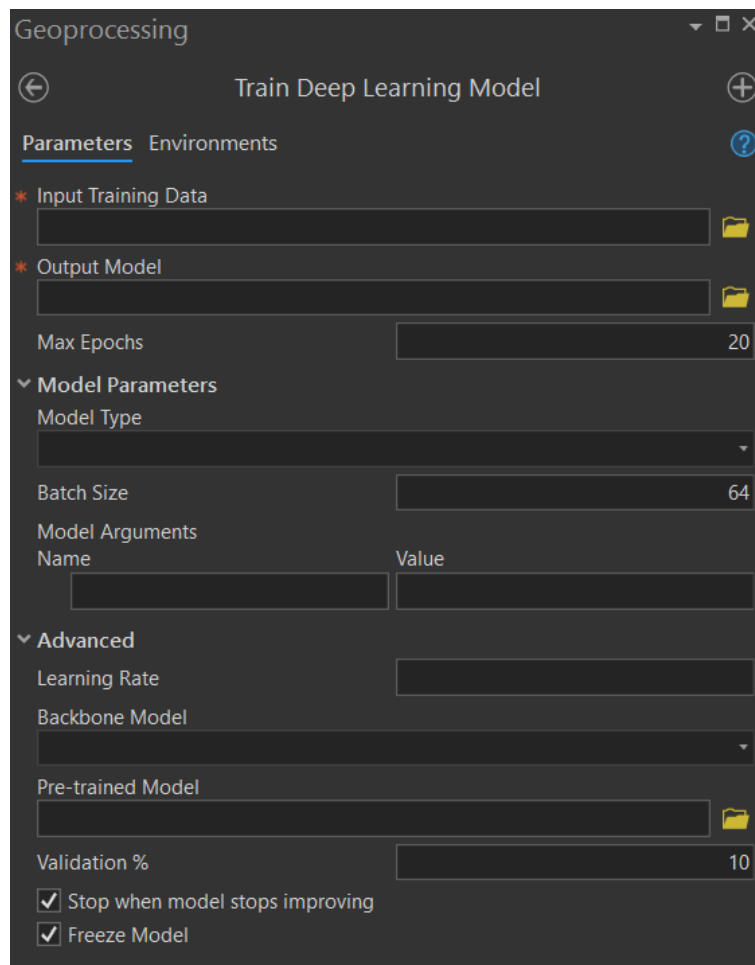


Figur 1: *Export Training Data For Deep Learning*.

- *Input Raster*: Her legges rasteret som skal benyttes for å lage treningsdata inn.

- *Additional Input Raster (Valgfri)*: Tilleggsraster som kan legges inn. Benyttes dersom treningsdataen skal inneholde ulike raster over et område.
- *Output Folder*: Filplasseringen til produktet.
- *Input Feature Class Or Classified Raster Or Table*: Data som skal benyttes til å trene på, eksempelvis FKB-data eller N50-data.
- *Class Value Field (valgfri)*: Attributfelt som inneholder verdier for de ulike klassene i datasettet.
- *Buffer Radius*: Her kan det legges til en buffer på dataen som skal benyttes til å trene på, brukes for eksempel hvis det benyttes vektordata i form av punkter.
- *Input Mask Polygons (valgfri)*: En polygon som bestemmer hvilket område treningsdataen lages på. Dersom dette ikke defineres lages den treningsdata på hele rasteret.
- *Image Format*: Bildeformatet treningsdataen lages i.
- *Tile Size X/ Tile Size Y (Valgfri)*: Angir hvilken størrelse hver chip skal være.
- *Stride X/ Stride Y (Valgfri)*: Angir hvor langt chipsen skal bevege seg før neste chip lages. Hvis denne er halvparten av Tile Size så vil det være 50% overlapp.
- *Rotation Angle*: Her kan en velge rotasjonsvinkel hvis det er ønskelig at det skal være flere bilder over samme treningsdata, med forskjellige vinkler. Dette gir flere bilder av samme objekt, men samtidig mer datamengde som må prosesseres.
- *Reference System*: Referansesystem til chipsene som lages.
- *Output No Feature Tiles*: Dersom treningsdataen skal inneholde bilder som ikke inneholder relevant vektordata må denne hukes av.
- *Metadata Format*: Formatet til treningsdataen.

Etter å ha laget treningsdata anvendes denne dataen til å lage en ny modell og trene denne, eventuelt retrene en eksisterende modell. For å gjøre dette benyttes verktøyet *Train Deep Learning Model* (ESRI, no date h), som er vist i *Figur 2*.



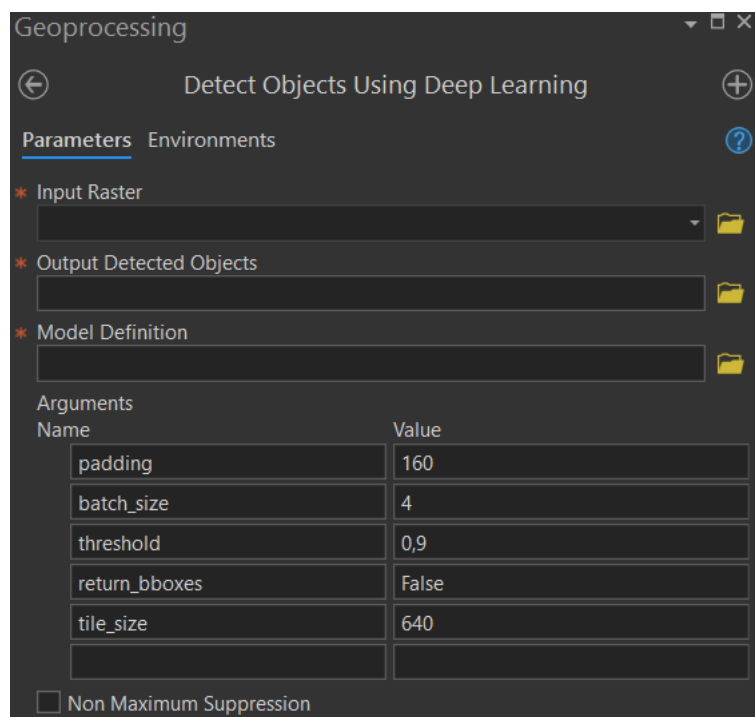
Figur 2: Train Deep Learning Model.

- *Input Training Data:* Resultatet fra *Export Training Data For Deep Learning*. Det vil si chipsene som beskrevet over.
- *Output Model:* Lokasjonen på produktet i filstrukturen.
- *Max Epochs:* Angir hvor mange ganger treningsdatasettet gjennomgås, altså antall treningsepoker.
- *Model Type:* Hva slags modelltype som skal brukes til å trene dyplæringsmodellen. Det finnes en rekke ulike modeller, med ulike egenskaper, og som dermed egner seg til ulike formål.
- *Batch Size:* Hvor mange prosesser maskinen skal kjøre samtidig. Avhenger av hva slags datamaskin som anvendes og størrelsen på treningsdataen.
- *Model Arguments:* Tilleggsparametere blir tilgjengelig under denne parameteren. Disse bestemmer hvordan modellen skal trenes.
- *Learning Rate (valgfri):* Definerer hvor mer vekt en nyere treningsepoke skal ha over en tidligere. Altså ved høyere Learning Rate vil treningsepoke 20 legges mer vekt på enn treningsepoke 1.

- *Backbone Model*: Det forhåndsconfigurerte kunstige nevralt nettverket som benyttes som arkitektur til å trene ny modell.
- *Pre-Trained Model (valgfri)*: En ferdigtrent modell som kan anvendes til å trene den nye modellen.
- *Validation*: Andelen treningsdata som blir brukt til å validere resterende treningsdata. Denne andelen blir kun brukt til validering.
- *Stop when model stops improving*: Dette er en parameter som stopper prosessen dersom modellen ikke forbedres. Hvis modellen blir ferdigtrent før den har nådd det maksimale antallet treningsepoker vil prosessen stoppe dersom denne er huket av.
- *Freeze Model*: Her velges det om *Backbone Model* skal få stå slik den er eller om det skal tillates at den tilpasses modellen.

En kan benytte treningsdata til å trene en ferdigtrent modell ytterligere. Dersom dette gjøres så fjernes enkelte parametere da dette legger en del begrensinger. Parameterne *Model Type*, *Arguments* og *Backbone Model* blir utilgjengelige fra verktøyet.

Detect Objects Using Deep Learning er et verktøy som bruker kunstige nevralt nettverk til å detektere objekter i et bilde. I *Figur 3* vises de ulike parameterne i verktøyet (ESRI, no date b):



Figur 3: Detect Objects Using Deep Learning.

- *Input Raster*: Rasteret over analyseområdet legges inn.

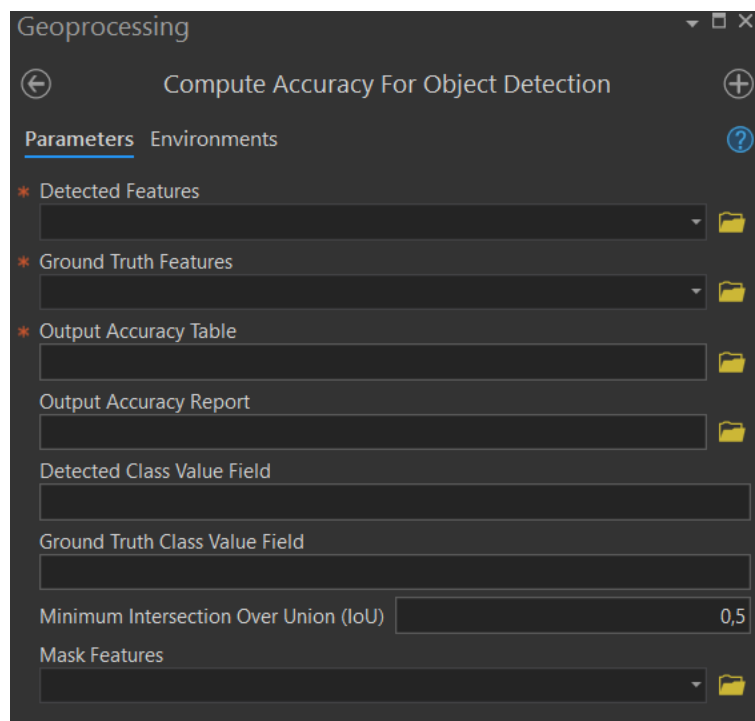
- *Output Detected Objects*: Ønsket filplassering til produktet.
- *Model Definition*: Modellen som benyttes til å detektere objekter, altså modellen som ble laget i forrige verktøy, eller en nedlastet ferdig modell.
- *Arguments*: Parametere som muliggjør endringer av deteksjonens gjennomføring.

Disse dukker opp etter at en modell legges inn:

- *Padding*: Antall piksler som skal ignoreres fra kanten i en chip.
- *Batch size*: Hvor mange prosesser som skal gå samtidig.
- *Threshold*: Konfidensnivå på resultatet. Default er 0,9 som tilsvarer 90%. Det vil si at kun deteksjoner som modellen kan bekrefte med 90% sikkerhet vil være med i produktet.
- *Return bbox*: Her bestemmes om det er ønskelig med en polygon som dekker objektet eller kun omrisset.
- *Tile Size*: Størrelsen på hver chip som skal detektere objekter.

For å evaluere deteksjonen benyttes verktøyet *Compute Accuracy For Object Detection*. I dette verktøyet sammenliknes deteksjonen med et datasett definert som «sannheten».

Produktet kommer ut som en PDF-fil med ulike verdier som beskriver deteksjonen. Verktøy vises i *Figur 4*:



Figur 4: Compute Accuracy For Object Detection.

- *Detected Features*: Datasettet som inneholder deteksjonen.
- *Ground Truth Features*: Datasettet som her representerer «sannheten».

- *Output Accuracy Table*: Filplasseringen av resultatene i form av en tabell.
- *Output Accuracy Report (valgfri)*: Filplassering på PDF-filen som inneholder rapporten.
- *Detected Class Value Field (valgfri)*: Hvilket attributfelt som skal evalueres i deteksjonsdatasettet.
- *Ground Truth Class Value Field (valgfri)*: Hvilket attributfelt som deteksjonsdatasettet skal evalueres mot.
- *Minimum Intersection Over Union (IoU)*: Her velges hvor mye overlapp det skal være mellom deteksjonen og «sannheten» for at verktøyet skal markere den som True Positive.
- *Mask Features (valgfri)*: Et avgrenset område som analysen skal gjøres på.