# Data

In dit onderzoek wordt gebruikgemaakt van twee typen data: video’s en foto’s.

## Video’s

De input voor het CAKE-framework bestaat uit video’s. Holland Mechanics heeft voor dit onderzoek drie instructievideo’s gemaakt. Deze video’s behandelen de montage van de machinecover, de stabilizer pressure control en de uitloopgoot. De duur van de video’s bedraagt respectievelijk 27, 5 en 13 minuten.

In de video’s wordt stap voor stap uitgelegd hoe de verschillende onderdelen gemonteerd dienen te worden, volgens de richtlijnen van Holland Mechanics. In Figuur … en … zijn screenshots weergegeven uit de video’s over de stabilizer pressure control en de uitloopgoot, waarin specifieke montage-instructies worden toegelicht.

Afbeelding met persoon, overdekt, vinger, nagel

Door AI gegenereerde inhoud is mogelijk onjuist.

## Foto’s

Het YOLOv12-model moet in staat zijn om onderdelen en gereedschappen uit de video’s correct te detecteren. Om dit te bereiken, wordt er een YOLOv12 model getraind op foto’s van relevante onderdelen en gereedschappen die in de video’s voorkomen. Deze onderdelen en gereedschappen zijn geselecteerd door de monteur die betrokken was bij de opnames van de instructievideo’s.

In samenwerking met deze monteur zijn bij Holland Mechanics foto’s gemaakt van de geselecteerde objecten. Vervolgens zijn deze foto’s geannoteerd met behulp van de Roboflow-annotatiesoftware (BRON UIT PVA). Elk relevant object is voorzien van een bounding box en het bijbehorende label. De eerste annotaties per object zijn samen met de monteur uitgevoerd, waarna de overige annotaties zelfstandig zijn aangebracht op basis van zijn instructies.

Bij het annoteren is als richtlijn gehanteerd dat een object alleen werd gemarkeerd wanneer het op basis van de context van de foto herkenbaar was. Dit is met name van belang bij objecten die slechts gedeeltelijk zichtbaar zijn. Na het annoteren zijn alle foto’s gecontroleerd op fouten, er zijn geen fouten aangetroffen, al kan niet met zekerheid worden gesteld dat de dataset volledig foutloos is.

De uiteindelijke dataset bestaat uit 741 foto’s, met 37 klassen met in totaal 1787 annotaties. Vijf foto’s bevatten geen objecten (UITLEG WAAROM). De dataset is willekeurig gesplitst in drie sets: 60% voor training, 20% voor validatie en 20% voor testen. In Figuur … is de verdeling van de labels binnen de trainingset weergegeven. De klasse ‘bout’ komt het meest voor met in totaal 185 voorbeelden. Het ‘afdekplaatje voor het ventieleiland’ is het minst vertegenwoordigd met slechts 5 voorbeelden.

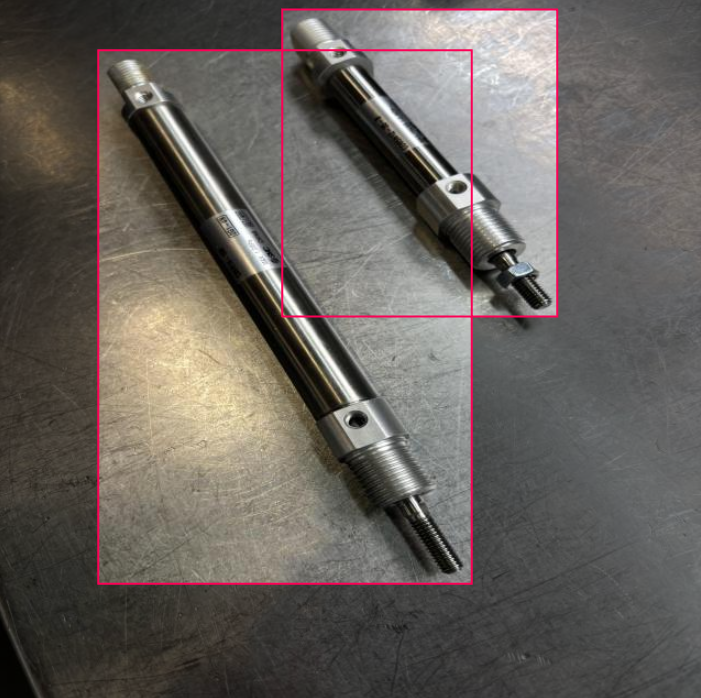
In Figuur … en … zijn foto’s uit de trainingsset te zien. In Figuur … zijn een ringsteeksleutel, schroevendraaier en momentsleutel zichtbaar. In Figuur … zijn twee pneumatische cilinders te zien.

Om de variatie in de trainingsdata te vergroten, is data augmentatie toegepast op de foto’s. Hiermee worden automatisch per epoch nieuwe versies van bestaande foto's gegenereerd, waarbij bijvoorbeeld de helderheid, kleurtoon, oriëntatie of positie van objecten licht gewijzigd is, zie paragraaf … voor de specifieke instellingen van deze hyperparameters. Deze verrijking van de dataset helpt het model om objecten te herkennen onder verschillende omstandigheden, zoals wisselende belichting of kijkhoeken. In figuur … zijn de foto’s uit de eerste batch van de eerste epoch te zien waarbij data augmentatie is toegepast.

Afbeelding met tekst, schermopname, Lettertype, diagram

Door AI gegenereerde inhoud is mogelijk onjuist.

Afbeelding met handgereedschap, Handgereedschap voor metaalbewerking, Huishoudelijke ijzerwaren, Snijgereedschap

Door AI gegenereerde inhoud is mogelijk onjuist.

Afbeelding met schermopname, collage, paars, Kleurrijkheid

Door AI gegenereerde inhoud is mogelijk onjuist.