

Manetgjenkjenning

Av

Even Bakke, Bjørn Instefjord, Kamil Matyjaszczyk & Richard Persson

15.11.2024

Problemet

Prosjektet skal lage en enkel manetidentifikator for å raskt finne ut hvilken manet man har funnet, uten å måtte søke manuelt. Ytelsen måles med treffsikkerhet (accuracy score): 49,44%.

Scope

Målet med prosjektet er å utvikle en modell for å identifisere maneter som kan brukes av alle ved å laste opp et bilde av den aktuelle maneten. Dette vil være nyttig for både forskere, lærere, elever og allmennheten som ønsker å lære mer om maneter. Løsningen kan også videreutvikles for å bli tilgjengelig på ulike plattformer. I dag finnes det løsninger som Google Bilder og enkelte maskinlæringsprogrammer dedikert til å gjenkjenne dyrearter, men disse kan være begrenset når det gjelder spesifikke manetarter.

Modellen vår er en del av et større system som består av tre trinn: bearbeiding av data, modellens prediksjoner og brukergrensesnittet. Hvis vi gjør endringer i ett av trinnene, for eksempel i hvordan vi behandler data, kan det påvirke hvor godt modellen fungerer. Derfor må vi ofte justere de andre delene for at hele systemet skal fungere best mulig sammen.

Interessentene våre inkluderer allmennheten og folk som er på stranden og vil identifisere arter som maneter. Vi jobbet hele tiden med prosjektet til det var ferdig, uten å dele det inn i spesifikke faser eller bruke formelle arbeidsmodeller. Planleggingen skjedde i uke 1, der vi blant annet valgte databasen vi skulle bruke. Uke 2 ble brukt til andre prioriteringer, og uke 3 gikk til å trene og forbedre modellen. Vi trengte noen ressurser for å få alt til, som GPU, utviklere, og et datasett med bilder av maneter.

Metriker

For at prosjektet skal være en suksess, er minimumsytelsen definert som en treffsikkerhet på minst 80 %. Dette er en "business metric" som reflekterer hvor godt modellen klarer å gi korrekte prediksjoner og dermed hvor nyttig den vil være i praktisk bruk. Vi har ikke oppnådd dette, så modellen krever videreutvikling.

Vi bruker også andre maskinlæringsmetriker som kvadratisk avvik (mean squared error), latens og throughput for å evaluere systemets ytelse. MSE er på 3,3667 som tyder på at

modellen tar betydelig feil når den skal kategorisere mellom 6 ulike maneter. Gjennomsnittlig latens målt til ca 0,009 sekunder pr prediksjon sikrer at modellen gir raske svar, noe som er viktig for brukervennligheten. Throughput, med kapasitet på rundt 11257 bilder pr sek, viser hvor effektivt modellen kan skaleres. Disse metrikker vil gi oss et godt bilde av systemets totale ytelse og sammenheng med det definerte forretningsmålet.

Data

Vi bruker bilder av seks forskjellige manettyper: barrel, blue, compass, lions_mane, mauve_stinger og moon_jellyfish. Bildene konverteres til gråtoner og skaleres til 26x26 for standardisering. Dataene er hentet fra åpne kilder, og vi sikrer at labels er konsistente ved manuell verifisering.

Bildene renses og flates ut før bruk i modellen for å sikre at de er kompatible med våre trenings- og prediksjonsprosesser.

Modellering

Vi startet med å prøve CatBoost, men endte opp med å bruke RandomForest på grunn av bedre ytelse og kjappere trening. Det var også mye mindre «knot» ved å bruke RandomForest contra CatBoost. Økning i pixelstørrelsen virket ikke å ha noe påvirkning på modellens treffsikkerhet.

Deployment

Modellen settes i drift via en Gradio-webapp som gjør det mulig for brukere å laste opp bilder og få en umiddelbar prediksjon om hvilken type manet det er.

Refferanser

A. Tanwar. «Jellyfish Image Dataset (Version 16)» (2023) Nettside:
<https://www.kaggle.com/datasets/anshtanwar/jellyfish-types/data>