

Prosjektrapport i DAT255 – Deep Learning Engineering

---

# Klassifisering av menn og kvinner med dyp læringsteknikker

---

Dato: 24/04/2025

Christian Vidal (kand.nr. 235)

Jonas Eide Skaslien (kand.nr. 411)

Jeg bekrefter at arbeidet er selvstendig utarbeidet, og at referanser/kildehenvisninger til alle kilder som er brukt i arbeidet er oppgitt, jf. Forskrift om studium og eksamen ved Høgskulen på Vestlandet, § 10.

# Problembeskrivelse

Motivasjonen og målet bak prosjektet er å forbedre et tidligere arbeid. For forrige semester, lagde gruppen en maskinlæringsmodell som skulle predikere på bilder av en enkelt person, der modellen skal bestemme om personen i bildet er en mann eller dame. Så for dette prosjektet vil gruppen ta dette arbeidet, og forbedre modellen med bruk av dyplæringsteknikker. Målet er å lage en bedre og mer robust modell, med høyere nøyaktighet. Den gamle modellen hadde en nøyaktighet på rundt 70%. Vi vil forbedre dette, med bruk av dyplæringsteknikker. Her har vi tenkt å bruke CNN for å lage flere modeller i Keras.

For løsningen, skal det deployes til en nettside med Gradio hvor en bruker kan laste opp et bilde, som blir gitt til modellen, og det vil bli returnert et svar om personen på bildet er en mann eller dame.

Tidligere har gruppen benyttet tradisjonelle metoder for kjønnsdeteksjon basert på ansiktstrekk, men disse tilnærmingene har flere svakheter. Et gjennomgående problem var at trekk som rynker ofte ble tolket som maskuline, noe som førte til at eldre personer (60+) nesten konsekvent ble klassifisert som menn – uavhengig av faktisk kjønn. I tillegg hadde modellen store utfordringer med å skille mellom kjønn hos barn, hvor ansiktstrekk ofte er mindre distinkte. Videre tok modellen ikke høyde for variasjoner i bildene, som rotasjon, speiling eller endringer i lysstyrke, noe som reduserte robustheten og generaliseringsevnen betydelig.

## Data

Dataene som blir brukt i dette prosjektet er UTKface-dataset [1], human images dataset - men and women [2], Men/Women Classification [3] og Humanface\_of\_various\_age\_groups [6].

UTKFace-datasettet er en omfattende samling av over 20 000 bilder av menneskeansikter, varierende i alder, kjønn, etnisitet og ansiktsposisjon. Bildene er tatt i ulike vinkler og lysforhold, noe som gir datasettet stor variasjon og realisme. Dette gjør det godt egnet for å gjenkjenne kjønn da datasettet kan eliminere problemer som at modellen bare klarer å predikere kjønn for en etnisitet eller aldersgruppe. Den største ulempen med UTKFace-datasettet er at majoriteten av bildene er tatt i svært nær avstand til ansiktet. Dette innebærer at bakgrunn og øvrige kroppsdeler i liten grad er representert i bildene. Som en konsekvens kan modeller trent på dette datasettet ha utfordringer med å generalisere til situasjoner der hele kroppen, eller større deler av den, er synlig. Dette gjør datasettet mindre egnet for oppgaver som krever kontekst utover ansiktet.

Human images-datasettet er et mindre datasett på rundt 1600 bilder. Bildene i datasettet består hovedsakelig av voksne mennesker i ulike alder, etnisitet og med varierende ansiktsposisjoner. En klar fordel med datasettet er at bildene er tatt under forskjellige lysforhold og vinkler, noe som gir økt variasjon og realisme. I tillegg er det et visst mangfold i etnisk bakgrunn, som kan bidra til mer rettferdige og robuste modeller. Likevel har datasettet noen betydelige begrensninger. For det første er det relativt lite, kun omtrent en tjuendedel av størrelsen til UTKFace-datasettet, noe som kan påvirke modellens evne til å generalisere. For det andre består nesten alle bildene av personer i alderen 20 til 70 år. Dette gjør at datasettet i liten grad representerer barn og eldre, og er dermed mindre egnet for oppgaver som krever pålitelig kjønnsestimering på tvers av hele aldersspekteret.

Det siste datasettet er Men/Women Classification. Dette datasettet består av ca. 3300 bilder av menn og kvinner i forskjellige posisjoner og vinkler. I motsetning til de andre datasettene inneholder dette datasettet flere bilder hvor både ansikt og kropp er synlige. Dette gir modellen mulighet til å lære seg sammenhenger mellom kroppslige trekk og kjønn, noe som potensielt kan forbedre prediksjonene i situasjoner der ansiktet alene ikke gir nok informasjon. Ved å inkludere hele eller deler av kroppen kan modellen oppnå en mer helhetlig forståelse av visuelle kjønnsindikatorer. En stor ulempe med dette datasettet er at det inneholder gruppebilder, noe som gjør det mer utfordrende å bruke i tradisjonelle klassifiseringsmodeller som forventer ett ansikt per bilde. Gruppebilder krever at man først identifiserer og isolerer hvert individ i bildet før man kan gjennomføre kjønnsestimering.

En siste stor fordel med alle datasettene er at de allerede er markert som menn og kvinner. Dette sparer mye tid da gruppen slipper å manuelt gå over alle bildene og markere dem manuelt.

Det finnes mange alternative datasett, ett av dem er CelebA. CelebA består av over 200 000 bilder av kjente personer markert som menn og kvinner. En stor fordel med CelebA er den betydelige størrelsen. Nevrale nettverk presterer ofte bedre når de trenes på store mengder data, da dette bidrar til bedre generalisering og reduserer risikoen for overtilpasning. Dersom CelebA hadde blitt benyttet, kunne den store datamengden potensielt ha forbedret modellens nøyaktighet og robusthet. Men den store mengden data er også en av grunnene til at gruppen valgte å ikke benytte CelebA-datasettet. På grunn av størrelsen krever datasettet betydelig lagringsplass, noe som kan by på praktiske utfordringer, spesielt ved lokal lagring eller begrensede ressurser. I tillegg vil trening av en modell på et så omfattende datasett ta vesentlig lengre tid, noe som ikke er passende med prosjektets tidsrammer. Dermed ble det vurdert at kostnadene ved å bruke CelebA, i form av tid og lagring, ikke veide opp for fordelene i denne sammenhengen.

Gruppen ønsker å sammenligne modellen med en tidligere modell laget av gruppen. Denne modellen var bare trent på Human images-datasettet, som igjen skal brukes i denne modellen.

Dataene i de tre datasettene som brukes er strukturert på ulike måter når det gjelder kjønnsmerking. I det ene datasettet angis kjønn gjennom filnavnet, der tallet 0 representerer menn og 1 representerer kvinner. I de to andre datasettene er bildene sortert i separate mapper, en for menn og en for kvinner. For å kunne bruke dataene effektivt i en og samme modell, må de være merket på en samme måte. Gruppen valgte derfor å standardisere datastrukturen ved å bruke mappestruktur som felles løsning. Denne strukturen gjør det enklere å laste inn og behandle dataene i treningsfasen, og er godt støttet av vanlige biblioteker som Keras. Som automatisk kan dele dataene inn i klasser basert på filstrukturen.

Siden en stor andel av bildene i datasettet viser frontale ansikter, kan modellen ha utfordringer med å håndtere bilder som er tatt i skrå eller uvanlige vinkler. For å øke modellens robusthet og evne til å generalisere, har gruppen valgt å inkludere rotasjon for datagenerering. Ved å rotere bildene i ulike grader under treningsprosessen skaper det variasjon i bildene, noe som kan bidra til at modellen lærer å kjenne igjen ansikter fra flere perspektiver.

Et annet potensielt problem med datasettet er at de fleste bildene viser ansiktene i full bredde og på relativt lik avstand. Dette kan gjøre modellen mindre robust i situasjoner der ansiktet er delvis synlig eller tatt fra svært nært hold. For å gjøre modellen mer tilpasningsdyktig til slike variasjoner, har gruppen valgt å implementere både zoom- og beskæring for datagenerering. Ved å zoome inn og beskære forskjellige deler av ansiktet under treningen, blir modellen eksponert for et bredere spekter av ansikt og avstander. Dette kan forbedre modellens evne til å håndtere nærbilder, samt bilder med delvis synlige eller beskærede ansikter.

## Implementering av modell

For dette problemet er convolutional-neural-network (CNN) brukt som hovedfokus. CNN-er laget for bildegjenkjenning, noe som er det gruppen ønsker å gjennomføre. De ulike arkitekturene gruppen har valgt å se på er: VGG-16, Resnet og en arkitektur gruppen har tilpasset selv.

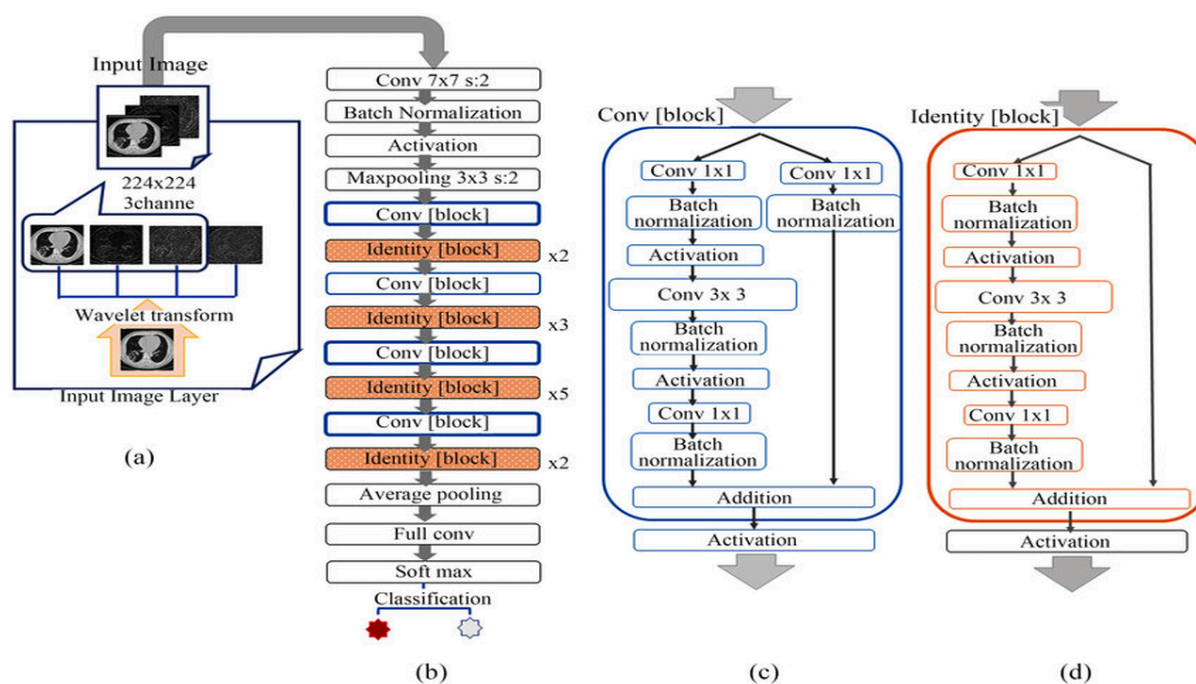
VGG16 ble tidlig vurdert som en potensiell arkitektur for prosjektet. Denne modellen er forhåndstrent og består av 16 lag, hvorav 13 er konvolusjonslag og 3 er fullt tilkoblede lag. En av fordelene med VGG16 er fleksibiliteten til å trene enten alle lagene eller kun de siste lagene, noe som gir muligheten til å tilpasse modellen til spesifikke oppgaver. Til tross for

denne fleksibiliteten, viste det seg under testing at VGG16 ikke leverte tilfredsstillende resultater for vårt spesifikke problem. Modellen hadde en tendens til å underprestere, og treningstiden var betydelig lengre sammenlignet med andre arkitekturer.

Resnet 50 i likhet med VGG16 er en forhåndstrent arkitektur som er trent på ImageNet-datasettet. Men i motsetning til VGG16, som består av en tradisjonell sekvens av konvolusjons og fullt tilkoblede lag, benytter Resnet 50 en mer avansert struktur med residual connections [5]. Dette gjør at Resnet i motsetning til VGG16 håndterer “vanishing gradient” problemet bedre og det gjør at selv om Resnet 50 har 50 lag så kan den fortsatt trene raskt.

En siste potensiell arkitektur er en hjemmelaget arkitektur laget av gruppen. Ved å bruke en egen arkitektur kan gruppen selv velge dybden og størrelsen på hvert lag. Med å lage en mindre arkitektur kan modellen trenes mye raskere enn VGG16 og Resnet. Dette lar gruppen teste flere hyperparametere i en rimelig tid. Ulempen med en mindre arkitektur er at den sannsynligvis ikke vil prestere like godt som de større arkitekturene.

Modellarkitekturen som ble valgt for prosjektet, var Resnet 50. Denne modellen ble valgt på grunn av dens raskere treningshastighet sammenlignet med VGG16, samt dens overlegne ytelse. Selv om ResNet50 ikke trente raskere enn gruppens egenutviklede modell, viste den betydelig bedre resultater på testsettet som gruppen hadde laget.



Figur 1 RESNET arkitektur

Det er flere hyperparametere som potensielt kan gjøre modellen bedre. En av dem er antall epoker modellen trenes med. Å velge riktig antall epoker kan være utfordrende da for mange kan føre til overfitting, mens for få kan gjøre at modellen ikke klarer å lære mønstrene i dataen. Vanligvis så ville antall epoker vært satt høyt med en stopp funksjon, som stopper treningen når tapsfunksjonen begynner å øke i valideringsettet. Et høyt antall epoker har vist seg å være tidkrevende. Derfor er valget av antall epoker basert på hvor lang tid det tar å trene modellen. Dette førte til at modellen ble dårligere enn den kunne potensielt ha vært, men treningen tok mindre tid.

En annen hyperparameter som kan påvirke modellen er batch size. Batch size bestemmer hvor mange bilder modellen skal se på før den oppdaterer vektene sine. Ved å ha en større batch kan en spare tiden det tar å trene en modell, da vektene ikke oppdateres like ofte. Men for stor batch kan føre til dårligere generalisering da modellen ser for mye på gjennomsnittet av bildene i stedet for individuelle bilder. Derfor har gruppen valgt å bruke en batch size på 32 som er en vanlig størrelse. Dette burde la modellen lære å generalisere uten at treningen tar for lang tid.

En siste hyperparameter som er viktig er learning rate eller læringsrate. Læringsrate sier noe om hvor mye vektene skal oppdateres. Dette vil si at hvis du har en læringsrate på 0.1 og gradienten av loss funksjonen er 10, så vil vekten oppdateres med  $0.1 * 10 = 1$ . Dette betyr at en høy læringsrate fører til raskere oppdateringer, noe som kan gjøre treningen raskere. Samtidig øker risikoen for at modellen overser optimale løsninger eller blir ustabil. En for lav læringsrate kan derimot føre til at modellen lærer veldig sakte, eller stopper opp før den når et godt resultat. Siden gruppen har problemer med lang treningstid og ikke bruker mange epoker så vil gruppen bruke en litt høyere treningsrate for å komme raskere til et godt svar.

Optimaliseringsalgoritmen som benyttes i treningsprosessen er Adam. Valget av Adam ble gjort fordi algoritmen automatisk justerer læringsraten for hver vekt og bias basert på tidligere gradienter. Dette gjør at modellen kan oppnå rask og stabil konvergens, selv ved begrenset treningstid. En annen sentral grunn til å bruke Adam er at den vanligvis krever mindre hyperparameter-tuning enn andre algoritmer, som for eksempel SGD. Adam er mindre sensitiv for valg av læringsrate og andre innstillinger, noe som reduserer behovet for omfattende eksperimentering og gjør det enklere å oppnå gode resultater uten omfattende justeringer.

Tapsfunksjonen som ble brukt i dette prosjektet er binary cross entropy. Binary cross entropy fungerer med å finne hvor langt unna en predikert verdi er fra den riktige klassen. For eksempel hvis en mann representeres som 0 og modellen predikerte et bilde av en mann

som 0.1 vil binary cross entropy være lav. Men hvis modellen predikerte 1.0 som er langt unna 0, så vil binary cross entropy være høy. I tilfeller der datasettet er ubalansert, for eksempel at det er veldig mange bilder av menn, men få bilder av damer. Så ville ikke binary cross entropy fungert like godt. Men siden gruppen har et balansert datasett så har gruppen valgt å bruke binary cross entropy.

For å forhindre overtilpasning ønsket gruppen å benytte early stopping som regulariseringsteknikk. Early stopping fungerer ved å stoppe treningen av modellen når valideringstapet begynner å øke, mens treningstapet fortsetter å synke. Dette indikerer at modellen har begynt å overtilpasse treningsdataene, og at den ikke lenger generaliserer godt til nye data. Etter testing fant imidlertid gruppen at modellen ikke nådde et punkt der valideringstapet begynte å øke, noe som gjorde det unødvendig.

I dette prosjektet har vi valgt å ikke benytte kryssvalidering som evalueringsmetode. I stedet har vi opprettet et eget testsett bestående av data som modellen aldri har hatt tilgang til under treningen, og som er helt adskilt fra treningsdataene. Testsettet består av bilder fra "Humanface\_of\_various\_age\_groups" [6], som inkluderer informasjon om alder, slik at vi kan analysere hvordan modellen presterer på ulike aldersgrupper. Dette testsettet ble brukt til å evaluere modellens ytelse etter at treningen var fullført. Denne tilnærmingen gir oss et mål på modellens evne til å generalisere til nye, usette data.

Dataene ble lastet inn ved hjelp av Keras-funksjonen `flow_from_directory()`, som automatisk kan organisere og tilordne klasser basert på mappestrukturen. For dette prosjektet ble det opprettet separate mapper for menn og kvinner. Deretter ble bildene plassert i de respektive mappene basert på enten filnavn eller mappestruktur, avhengig av hvordan dataene var organisert. Denne metoden gjør det enkelt å administrere store datasett og sikre at bildene automatisk blir tildelt riktig klasse under innlastingen.

## Evaluering

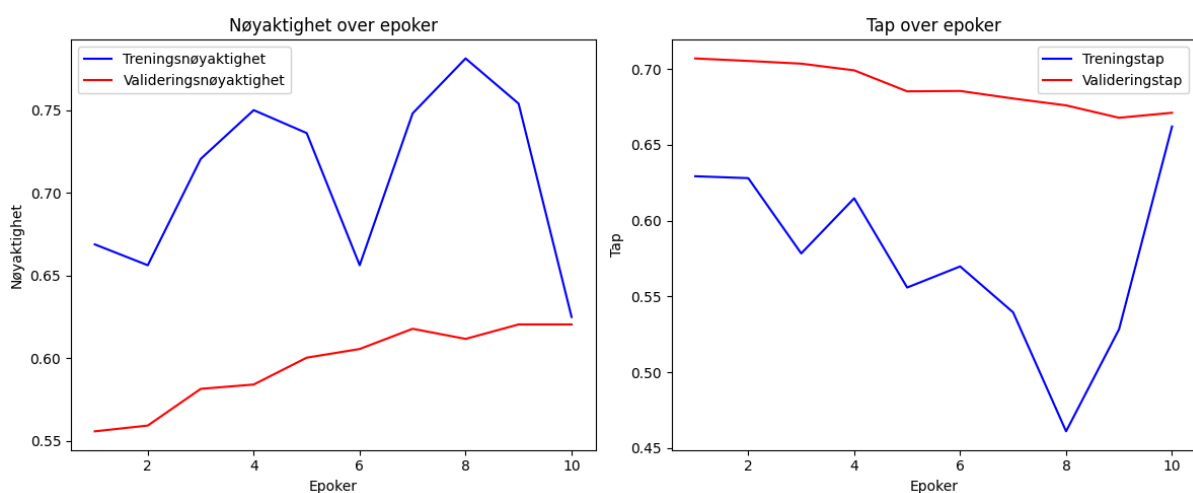
For å evaluere ytelsen til klassifiseringsmodellen har gruppen benyttet nøyaktighet som hovedmetrikk. Nøyaktighet måler andelen korrekte prediksjoner av totalt antall prediksjoner, og gir dermed en intuitiv indikasjon på hvor godt modellen klarer å klassifisere bildene i riktig kategori (mann eller kvinne). Evalueringen er gjennomført ved hjelp av et separat valideringssett, som ikke inngår i treningsdataene, for å sikre en objektiv vurdering av modellens generaliseringsevne.

Under treningsfasen er Binary Cross Entropy (BCE) benyttet som tapsfunksjon. BCE beregner forskjellen mellom modellens predikerte sannsynligheter og de faktiske klassene, og brukes

av modellen for å justere vektene sine og forbedre ytelsen over tid. Mens BCE brukes som et internt treningsverktøy, er det nøyaktigheten som benyttes som hovedindikator for hvor godt modellen presterer i praksis.

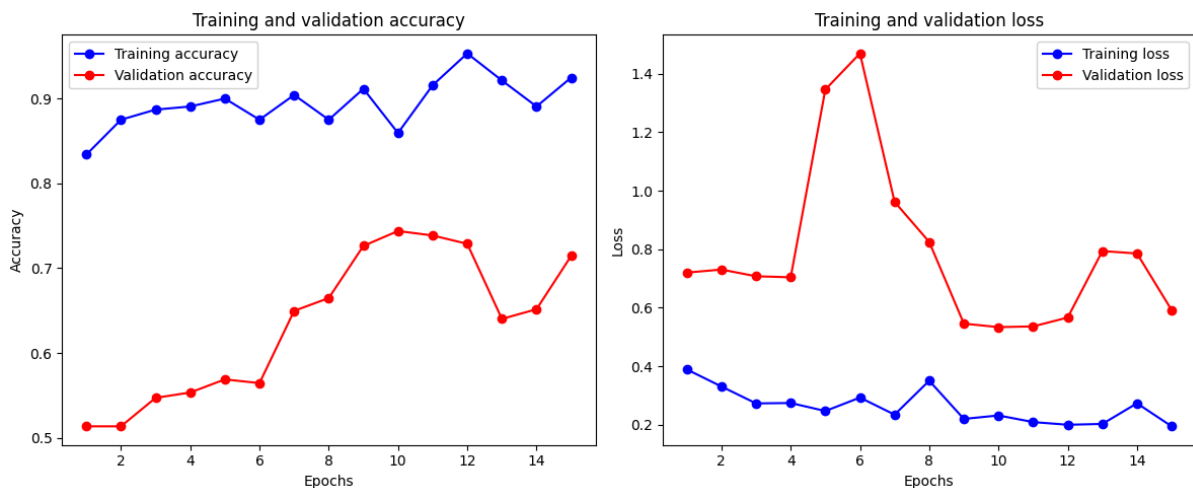
Som baseline for evaluering av den nye modellen benyttes tidligere implementasjoner fra tidligere arbeid, samt VGG-16-arkitekturen, som fungerte som utgangspunkt og prototype i prosjektets tidlige faser. Ved å sammenligne ytelsen til den nye modellen med disse baseline-modellene, kan vi vurdere om endringene og tilpasningene har ført til forbedret ytelse.

Sammenligningen baseres hovedsakelig på nøyaktighet og tapsverdi. En forbedring i disse målene vil indikere at den nye modellen generaliserer bedre og gir mer presise prediksjoner enn de tidligere versjonene. Dette gir et solid grunnlag for å vurdere hvorvidt den videreutviklede modellen representerer en faktisk forbedring over baseline-løsningene.



Figur 2 VGG16 treningsdata





Figur 3 RESNET treningsdata

Etter en grundig evaluering av ulike modeller konkluderte gruppen med at Resnet 50 ga en betydelig forbedring i både nøyaktighet og tap sammenlignet med de andre modellene. Resnet 50 ble derfor valgt som den modellen gruppen ønsket å fokusere på, og oppnådde en nøyaktighet på rundt 90% og et tap på rundt 0,2. Dette er et sterkt resultat, spesielt ettersom det overgår baseline-løsningene.

Resnet 50 har et stort potensial, spesielt med sin dype arkitektur og residual læring, som gjør at modellen kan lære komplekse funksjoner uten å lide av problemer som vanlige dype nettverk kan støte på, som vanishing gradients. Dette gir Resnet 50 en fordel i forhold til andre nettverksmodeller, spesielt når det gjelder å lære fra store datasett og oppnå god generalisering. Modellen er også kjent for å være svært effektiv i overføring av læring, noe som gjør den egnet for en rekke bildeklassifiseringsoppgaver, selv med begrensede treningsdata.

Imidlertid kan det påpekes at gruppen ikke fullt ut har utnyttet det potensialet Resnet 50 har å tilby. Den tidkrevende treningen av modellen førte til at gruppen ikke hadde nok tid til å utforske alle muligheter for finjustering og optimalisering. Videre er Resnet 50 trent på et bredt spekter av objekter, ikke spesifikt på menneskebilder, noe som gruppen kunne dra nytte av ved å legge til en kategori. Denne kategorien ville representere ikke-mennesker, noe som ville hjelpe modellen med å filtrere ut objekter og dyr som ikke er mennesker. Dersom mer tid hadde vært investert i treningen, og dersom modellen hadde fått tilpasset vektene sine ytterligere, ville det vært mulig å oppnå ytterligere forbedringer både i nøyaktighet og tap.

Utover klassiske evalueringsmål som nøyaktighet og tap, finnes det praktiske hensyn som også er viktige ved implementering og bruk av modellen. Dette inkluderer blant annet lagringsplass, kjøretid og ressursbruk under trening og drift.

Siden modellen er deployert via Gradio som en nettbasert applikasjon, er det viktig at modellen ikke er for stor eller krevende å kjøre, slik at responstiden for brukerne forblir akseptabel. Ressurshensyn som modellens størrelse og lastetid blir dermed relevante metrikker for den endelige brukeropplevelsen i web-applikasjonen.

## Deployment

Prosjektgruppen har valgt å benytte Gradio som plattform for deployering av maskinlæringsmodellen. Gradio er et brukervennlig rammeverk som gjør det mulig å lage en interaktiv webapplikasjon med et enkelt brukergrensesnitt, og dermed tilgjengeliggjøre modellen for testing og bruk uten behov for lokal installasjon. Modellen er deployert direkte til Gradio, og en offentlig lenke genereres automatisk. Denne kan benyttes av brukere som ønsker å teste modellen ved å laste opp et bilde. Modellen forsøker deretter å klassifisere bildet som enten mann eller kvinne, og presenterer resultatet direkte i brukergrensesnittet.

Gradio har vist seg å ha god funksjonalitet for vårt brukstilfelle. Det tilbyr en lavterskel løsning for deployering av maskinlæringsmodeller, samtidig som det gir et profesjonelt og responsivt grensesnitt. Rammeverket støtter også enkel integrasjon av brukerinteraksjon, noe som vil være nyttig for innsamling av tilbakemeldinger i en videreutviklingsfase.

For å sikre vedlikehold og forbedring av modellen over tid, planlegger vi å implementere en funksjon der brukeren kan gi tilbakemelding på om klassifiseringen var korrekt. I tilfelle modellen gjør en feilklassifisering, vil det være mulig å lagre en kopi av bildet sammen med den riktige etiketten. Denne prosessen vil gjøre det mulig å oppdatere og forbedre modellen ved å bruke de lagrede bildene i fremtidige treningsrunder. Denne tilbakemeldingssløyfen vil kunne bidra til en kontinuerlig forbedring av modellens nøyaktighet og generaliseringsevne.

Videreutvikling av prosjektet kan gjøres på flere måter. En naturlig utvidelse vil være å trene modellen til å håndtere bilder med flere personer. Dette krever implementering av persondeteksjon og separasjon før individuell klassifisering. Videre kan brukeropplevelsen forbedres ved å legge til en ekstra klassifikasjonskategori for "ikke-mennesker". Ettersom brukere har mulighet til å laste opp vilkårlige bilder, bør systemet være i stand til å filtrere ut bilder som ikke inneholder mennesker, for eksempel bilder av dyr eller objekter. Dette vil kunne bidra til økt treffsikkerhet og en mer robust modell.

## Konklusjon

I dette prosjektet har vi hatt som mål å forbedre en eksisterende maskinlæringsmodell for kjønnsdeteksjon ved å implementere dyplæringsteknikker, spesielt ved bruk av Convolutional Neural Networks (CNN). Den opprinnelige modellen, som ble utviklet i forrige semester, hadde en nøyaktighet på rundt 70%. Vårt hovedmål var å forbedre denne ytelsen ved å lage en mer robust modell som kunne håndtere utfordringer som rynker, ansiktstrekk hos barn, og variasjoner i bildeforhold som lysstyrke, rotasjon og speiling.

Gjennom bruken av Resnet 50, en forhåndstrent dyp læringsmodell, og databehandlingsteknikker som bildeaugmentering, har vi klart å oppnå en modell som gir bedre resultater på testsettet og viser høyere generaliseringsevne. Med en nøyaktighet på 92% på treningssettet, 71% på valideringssettet og 83% på testsettet. Så var dette en god forbedring fra modellen som ikke brukte dyp læring. Modellen vi utviklet kunne skille mellom kjønn på et bredere spekter av bilder, og vi unngikk noen av de problemene som tidligere modeller hadde, som å feilklassifisere eldre personer eller barn.

Når vi ser på de valgte metrikkene, som nøyaktighet og tap, har de vært passende for å evaluere modellens ytelse. Metrikkene ga oss viktig innsikt i hvordan modellen presterer, både under trening og evaluering, og har hjulpet oss med å justere hyperparametere og gjøre forbedringer. Imidlertid kunne vi også ha benyttet flere metoder for evaluering, som presisjon og tilbakekalling, for å få et mer nyansert bilde av modellens styrker og svakheter.

Overordnet sett, har dyplæring vist seg å være en effektiv løsning for kjønnsdeteksjon på bilder. Den nye modellen har forbedret nøyaktigheten, og samtidig har vi bygget en løsning som er mer robust mot variasjoner i bildene. Den viktigste potensielle forbedringen for fremtiden vil være å trene modellen lenger med flere epoker.

## Kilder

<https://susanqq.github.io/UTKFace/> [1] hentet 17.04.2025

<https://www.kaggle.com/datasets/snmahsa/human-images-dataset-men-and-women/data> [2] hentet 17.04.2025

<https://www.kaggle.com/datasets/playlist/men-women-classification> [3] hentet 17.04.2025

<https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html> [4] hentet 17.04.2025

<https://www.baeldung.com/cs/transformer-networks-residual-connections> [5] hentet 24.04.2025

<https://www.kaggle.com/datasets/neeeeeear/humanface-of-various-age-groups> [6] hentet 24.04.2025

<https://github.com/672648/Dat255-Obliq> link til github

## Figur

Figur 1 : Eksempel på RESNET arkitektur

<https://www.thinkingstack.ai/blog/business-use-cases-11/a-comprehensive-guide-to-resnet-50-architecture-and-implementation-56#:~:text=ResNet50%20is%20a%20particular%20configuration%20of%20the%20ResNet,ResNet50%20denotes%20the%20plain%20network%20of%20the%20network>. hentet 24.04.2025

Figur 2 : VGG16 nøyaktighet og loss for hver epoke

Figur 3 : Resnet50 nøyaktighet og loss for hver epoke