

UiO • Institutt for informatikk

Det matematisk-naturvitenskapelige fakultet

INF5860 - Maskinlæring for bildeanalyse

Introduksjon



Overordnede temaer

Forelesninger

Maskinlæring

Dyp læring og bildeanalyse

Dyp læring i praksis

Anvendelser

Gruppetimer

Numpy og bildeanalyse

Klassifikasjonsalgoritmer

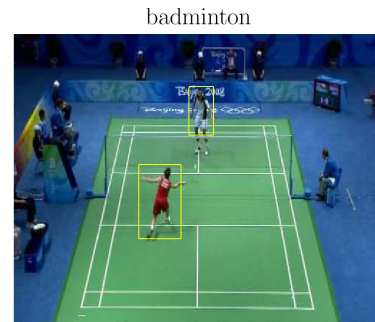
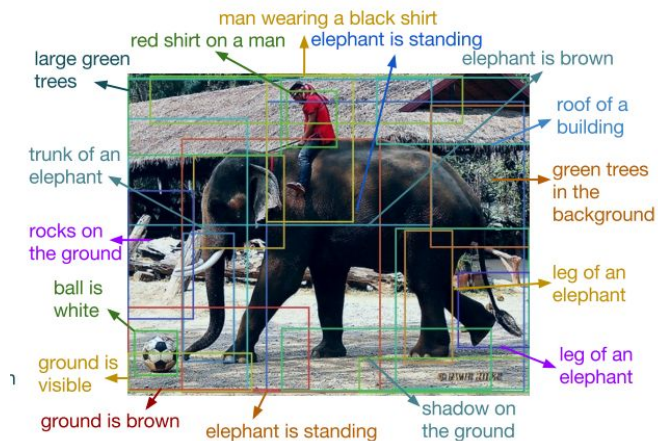
Tensorflow

Visualisering

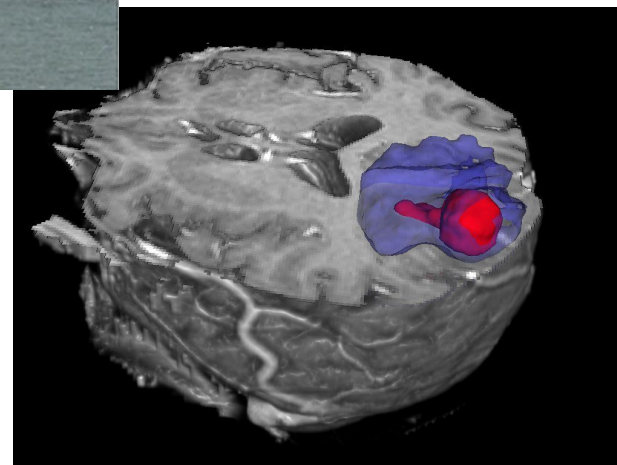
Hvorfor skal du lære dette?

Hvorfor bildeanalyse?

“Mange oppgaver går ut på å hente informasjon fra bilder.”



badminton



3 bildeanalyse konkurranser akkurat nå



Data Science Bowl 2017

Can you improve lung cancer detection?

Featured · 3 months to go · 192 kernels

\$1,000,000

691 teams



The Nature Conservancy Fisheries Monitoring

Can you detect and classify species of fish?

Featured · 3 months to go · Entered · 214 kernels

\$150,000

1,203 teams



Dstl Satellite Imagery Feature Detection

Can you train an eye in the sky?

Featured · 2 months to go · 88 kernels

\$100,000

136 teams



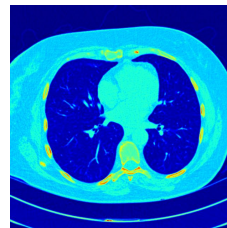
Two Sigma Financial Modeling Challenge

Can you uncover predictive value in an uncertain world?

Featured · A month to go · 152 kernels

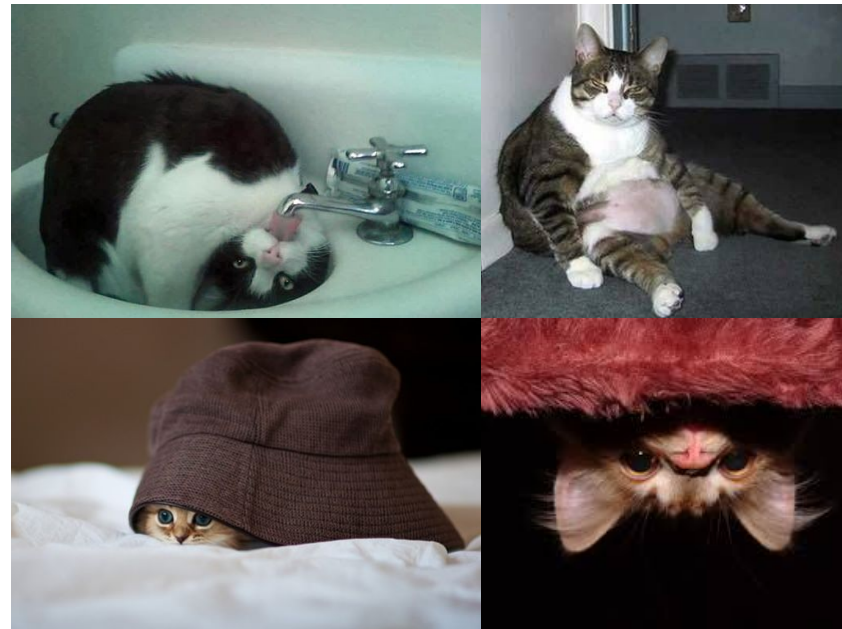
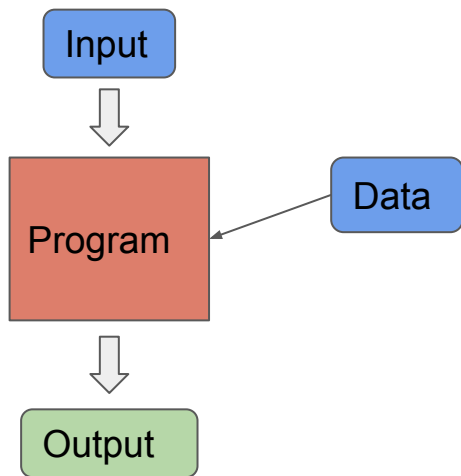
\$100,000

1,404 teams



Hvorfor maskinlæring?

“Tilpasser resultatet av programmet etter data, i stedet for å eksplisitt programmere alt.”



Hvorfor maskinlæring?

Bildeanalyse er vanskelig!

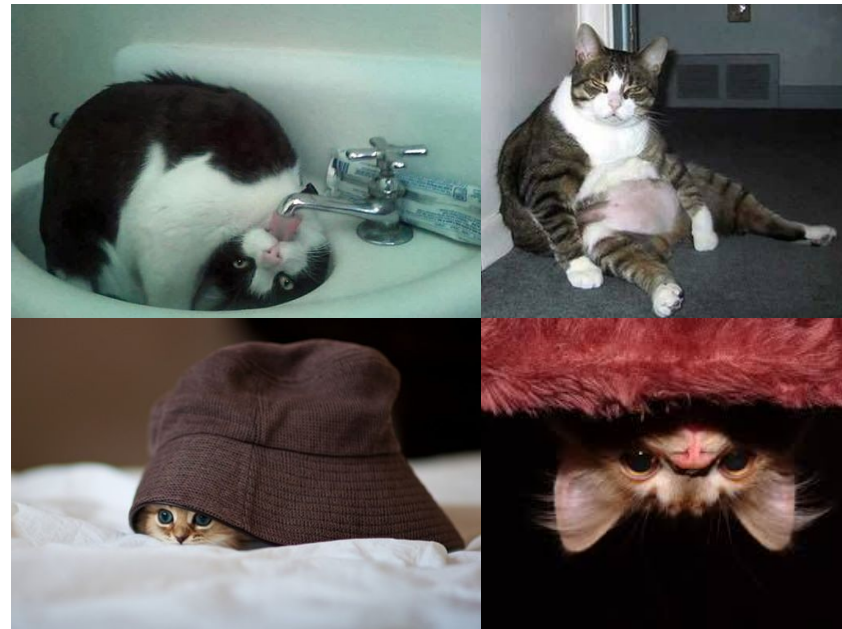


IN CS, IT CAN BE HARD TO EXPLAIN
THE DIFFERENCE BETWEEN THE EASY
AND THE VIRTUALLY IMPOSSIBLE.

Hvorfor maskinlæring?

Bildeanalyse er vanskelig!

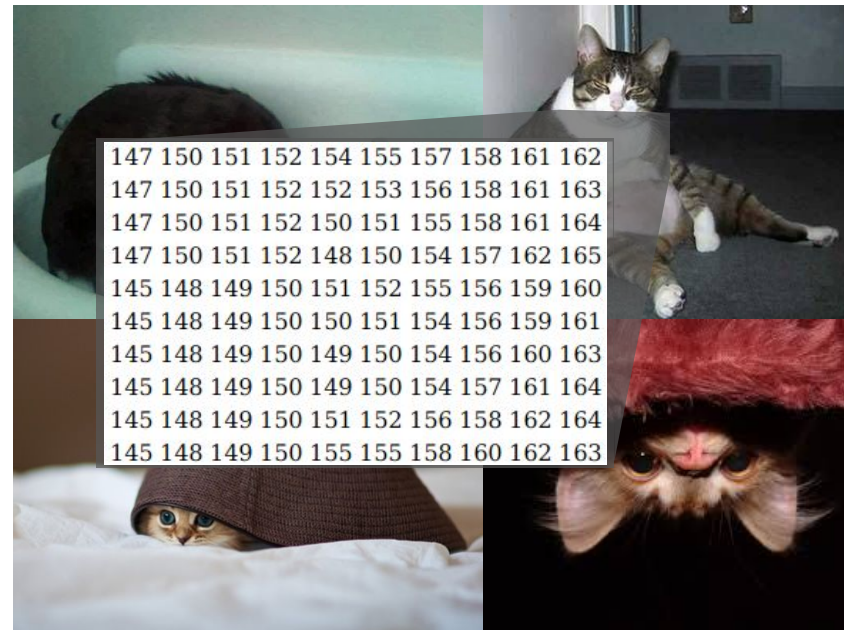
- Nesten umulig å spesifisere alle variasjoner av et objekt
- Å skrive ny kode for hver ting man skal finne er mye jobb



Hvorfor maskinlæring?

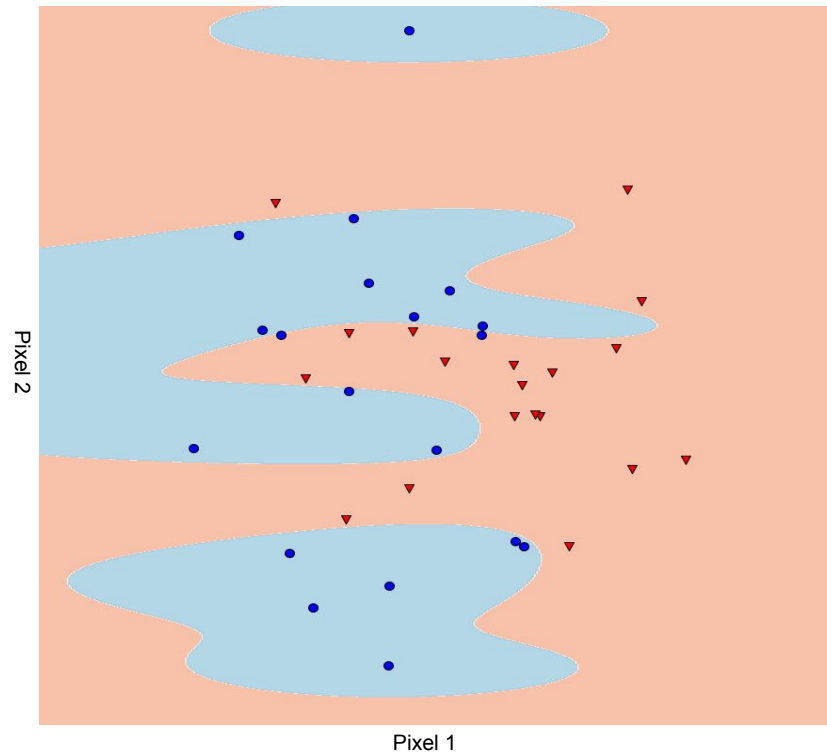
Bildeanalyse er vanskelig!

- Nesten umulig å spesifisere alle variasjoner av et objekt
- Å skrive ny kode for hver ting man skal finne er mye jobb



Maskinlæring - lage beslutningsgrenser

- Pikksele kombinasjoner gir forskjellige klasser
- Vanskelig å se hvordan dette er nyttig for bilder
- Høy dimensjonalt rom
- Bilder fra samme klasse er veldig forskjellige



Bilder er vanskelig input

- Alle piksler kan potensielt være kattepixels
- Veldig mange mulige bilder $256^{(M \cdot N)}$
- Veldig liten andel av disse bildene er katter
- Katter kan være langt unna hverandre i dette rommet

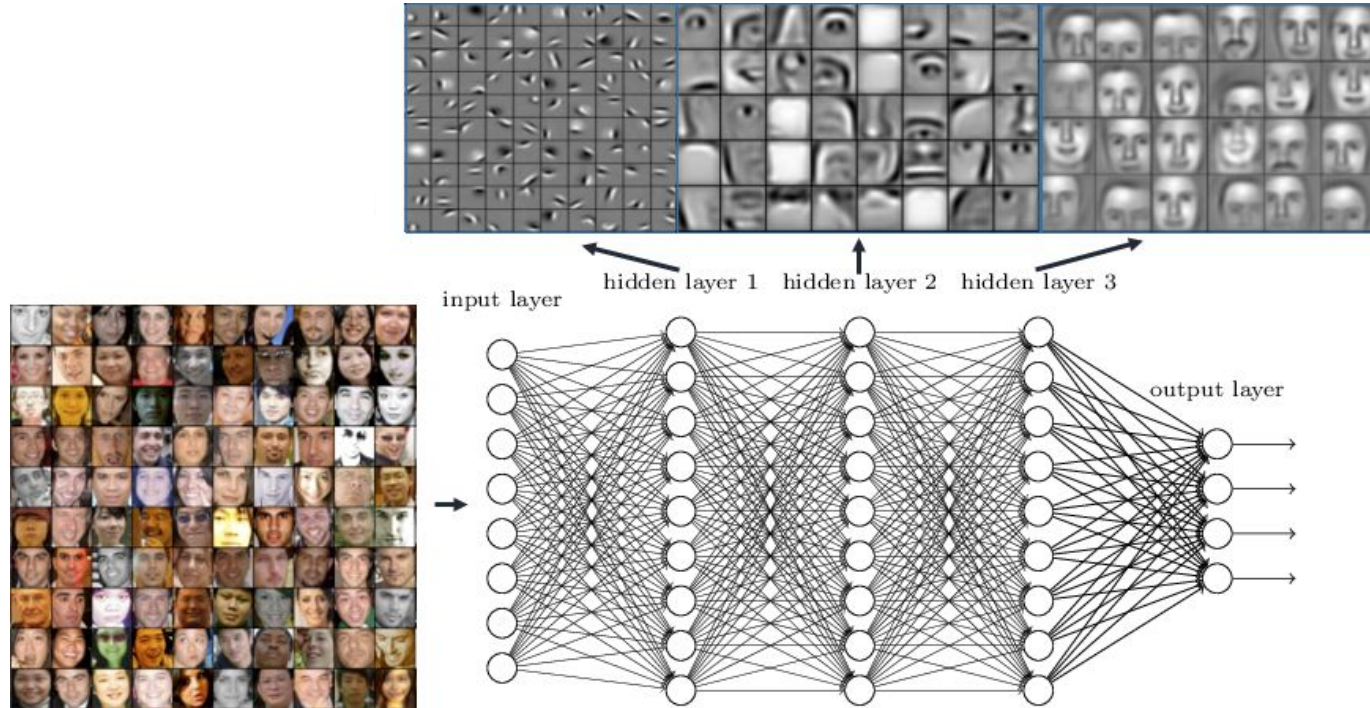


Hvorfor er dypt læring spesielt nyttig for bildeanalyse?

- Begynner enkelt
 - Det finnes veldig få katter blant alle mulig bilder
 - Kanter finnes det flere av
- Drar nytte av mange flere bilder enn bare katebilder
- Kaster bort uvesentlig informasjon lag for lag
- Skaper “invarians”



Et dypt nett bygger mer komplekse representasjoner



Eksempler på bruksområder

Klassifisering

- ImageNet
- Mange oppgaver kan formuleres som klassifiserings oppgaver
 - Sjekke om noe er tilstede
 - Telle noe
 - Ta beslutninger

Image classification

Easiest classes



Hardest classes



[Identity mappings in deep residual networks](#)

Segmentering

Knytte enkelt piksler til objekter

[MSCOCO](#) er standard benchmark

Kan brukes til:

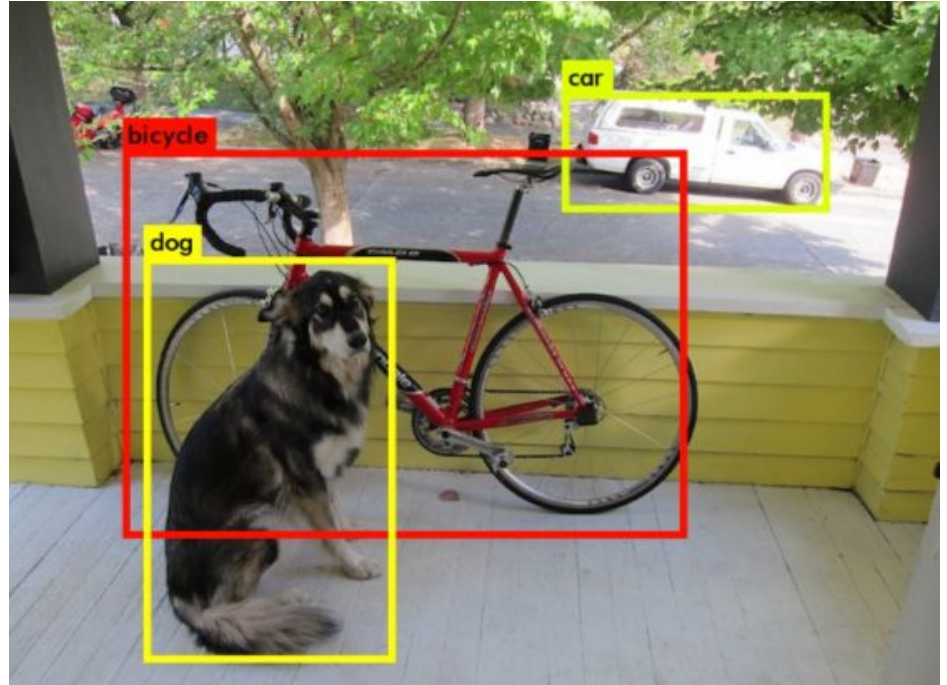
- Finne ut hvor man kan kjøre
- Måle kalorier i mat
- Måle hjertevolum
- osv.



Deteksjon

- Når man ikke har segmentering
- Brukt som en del av enkelt eksempel segmentering

[You only look once: Unified, real-time object detection](#)



Keypoint detection

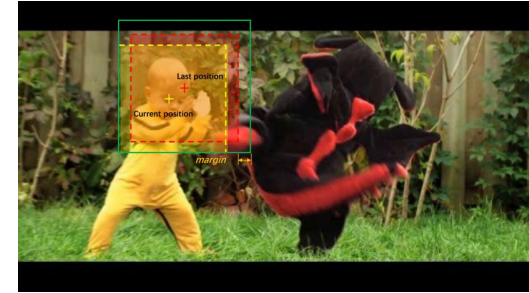
Finne igjen kjente punkter

[Towards Accurate Multi-person Pose Estimation in the Wild](#)



Tracking, re-identification and odometry

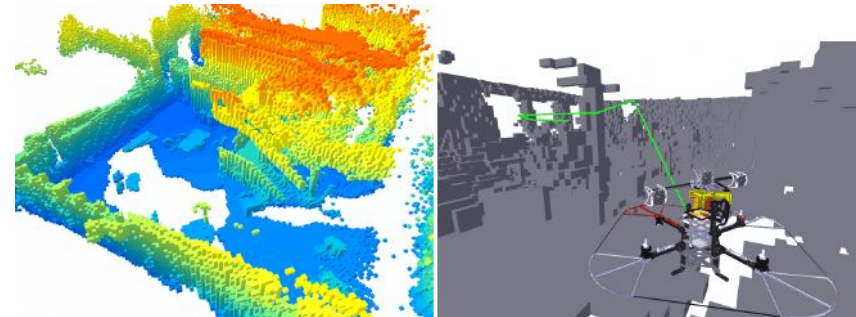
- Følge eller gjenkjenne “vilkårlige” objekter eller punkter
- Gjenkjenne mennesker
- Lage kart fra en video, ved matching av piksler



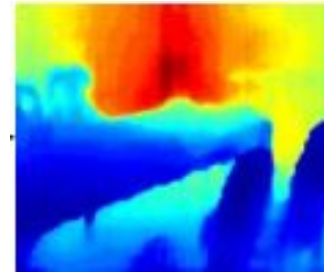
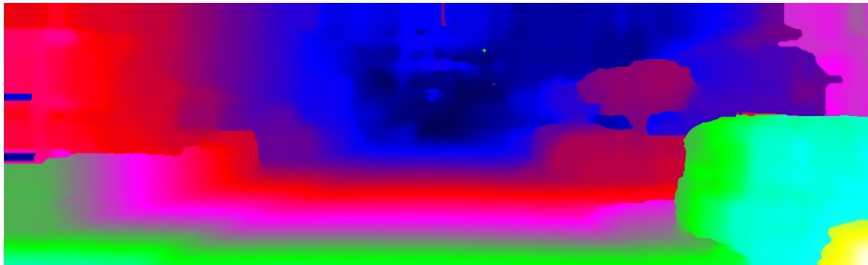
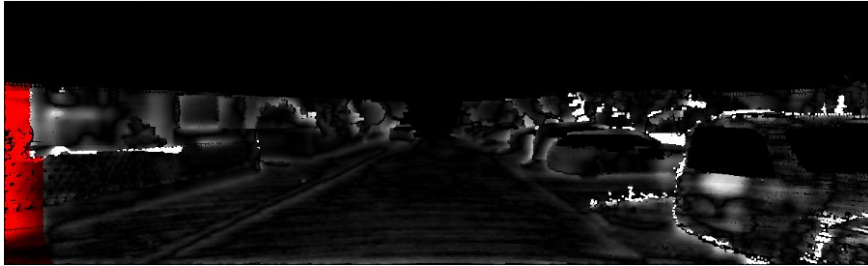
[Dual Deep Network for Visual Tracking](#)

[DeepVO: A Deep Learning approach for Monocular Visual Odometry](#)

[A Multi-task Deep Network for Person Re-identification](#)



Dybdeestimering fra ett eller to kameraer

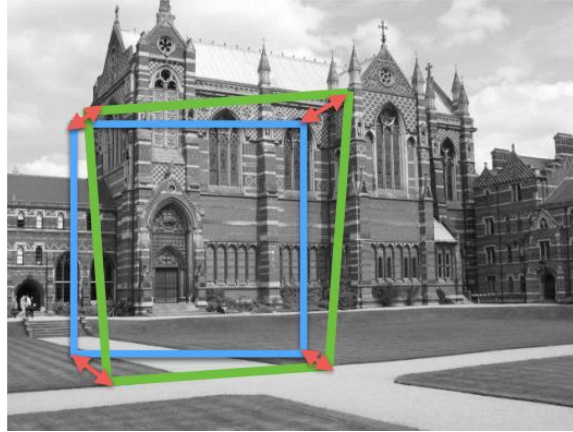


[Estimating Depth from Monocular Images as Classification Using Deep Fully Convolutional Residual Networks](#)

Posisjon/Homografi estimering

Direkte estimering av
projeksjons matrise

[Deep Image Homography Estimation](#)



Optisk flyt og prediksjon

Forutse neste bilde.

Enten for prediksjonen selv (vær etc.)

Eller for å forbedre optisk-flyt
estimering

[FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks](#)

[Deep multi-scale video prediction beyond mean square error](#)



Input frames



Adversarial result



Adversarial+GDL result

Beskrivelser

Mer avanserte arkitekturer

Image Captioning and Visual Question Answering Based on Attributes and External Knowledge



Attributes:

umbrella
beach
sunny
day
people
sand
laying
blue
green
mountain

Image Caption:

A group of people enjoying a sunny day at the beach with umbrellas in the sand.



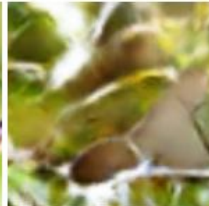





External Knowledge:

An umbrella is a canopy designed to protect against rain or sunlight. Larger umbrellas are often used as points of shade on a sunny beach. A beach is a landform along the coast of an ocean. It usually consists of loose particles, such as sand....

Question Answering:

Q: Why do they have umbrellas? **A :** Shade.

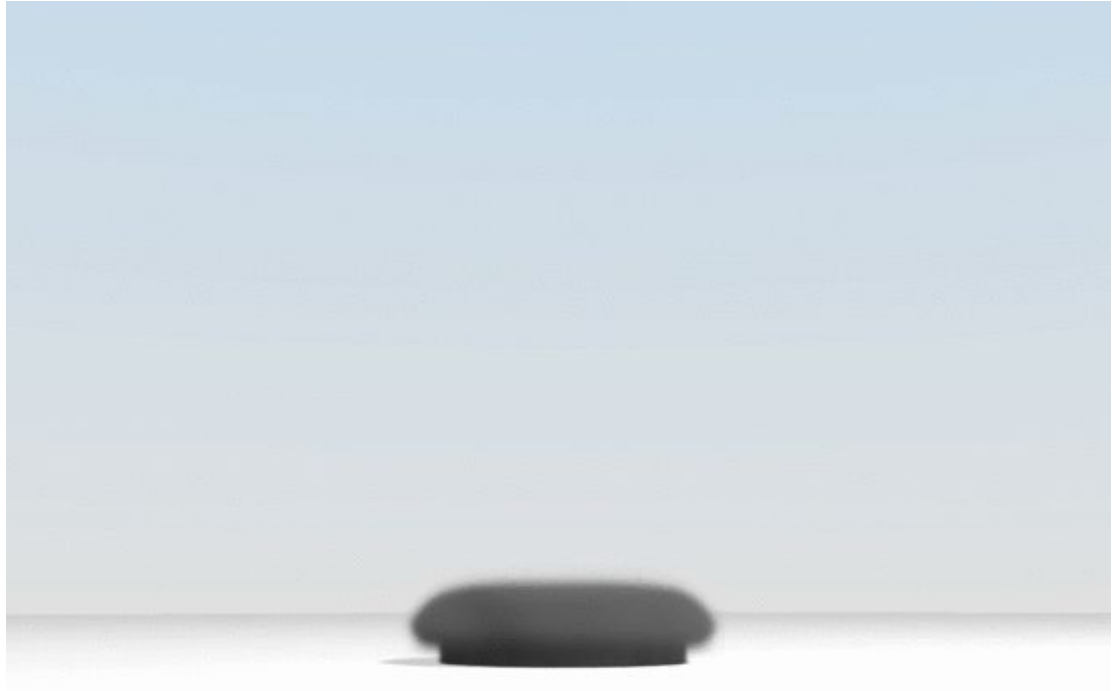
Bildgenerierung

Text description	This bird is blue with white and has a very short beak	This bird has wings that are brown and has a yellow belly	A white bird with a black crown and yellow beak	This bird is white, black, and brown in color, with a brown beak	The bird has small beak, with reddish brown crown and gray belly	This is a small, black bird with a white breast and white on the wingbars.	This bird is white black and yellow in color, with a short black beak
Stage-I images							
Stage-II images							

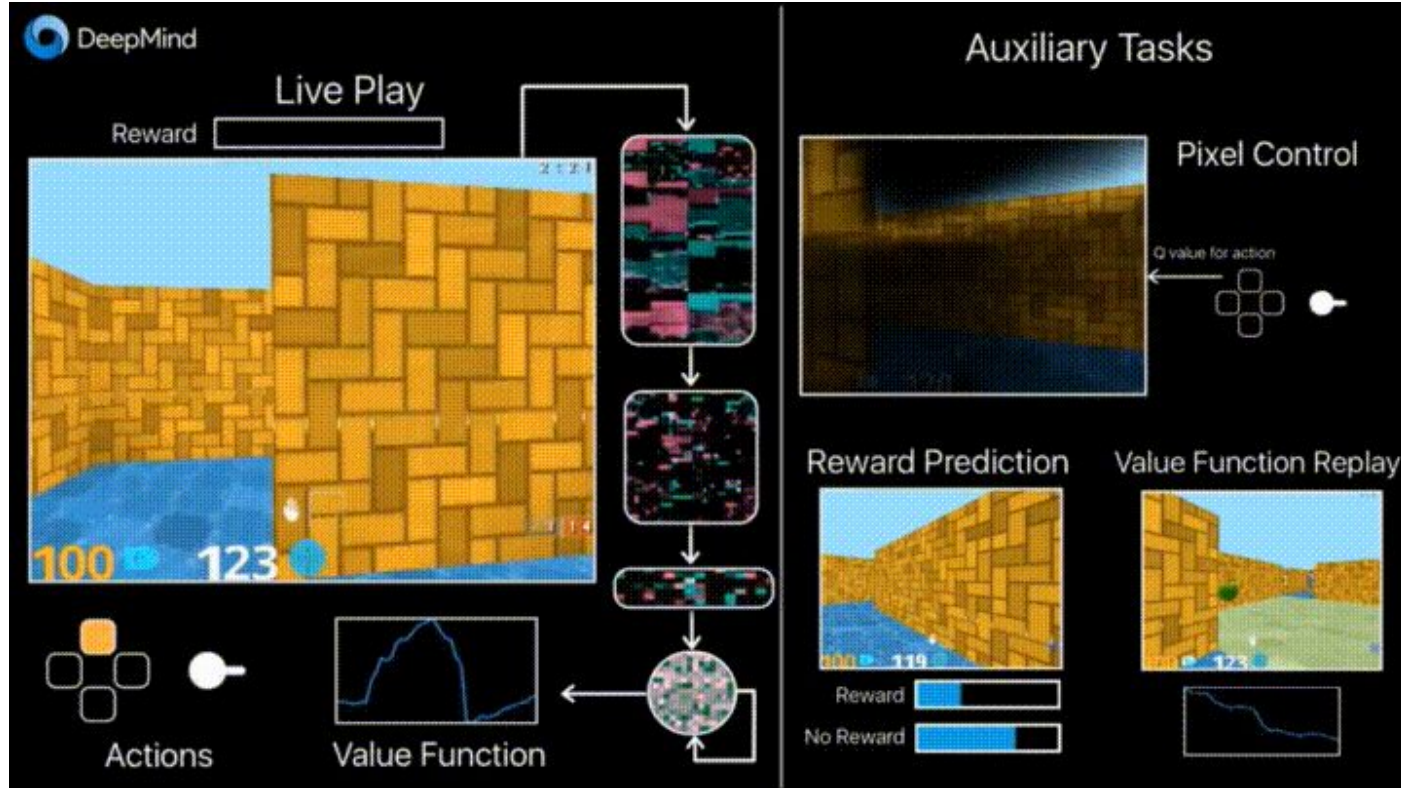
Simulering av veske og røyk

- Relativt nøyaktig
- Veldig raskt

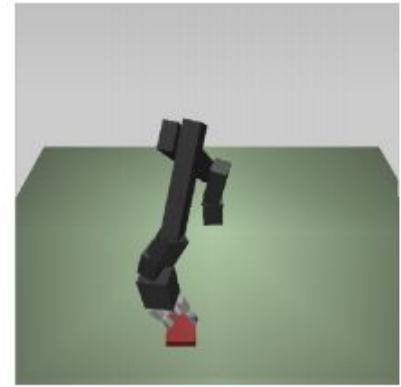
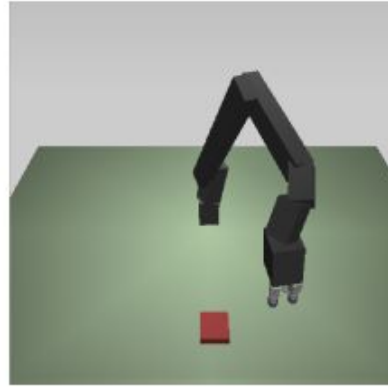
[Accelerating Eulerian Fluid Simulation With Convolutional Networks](#)



Dataspill med kun bilde som input



Direkte kontroll med bilde input



[Sim-to-Real Robot Learning from Pixels with Progressive Nets](#)

Maskinlæring kan ikke gjøre alt

- Tidkrevende å bevege seg ut i ukjent farvann
- Komplekse problemer kan være svært vanskelig å optimere, spesielt hvis de krever minne
- Ikke gode eksempler på å lære algoritmer
- Feil kan fort bygge seg større hvis man setter sammen mange nettverk
- Kan trenge domenekunnskap
 - Data kan endrer seg over tid
 - Korrelerte data
 - “Data augmentering”