



Design optimization of hydraulic energy storage and conversion system for wave energy converters

Wang, Dong; Lu, Kaiyuan

Published in:
Protection and Control of Modern Power Systems

DOI (link to publication from Publisher):
[10.1186/s41601-018-0080-6](https://doi.org/10.1186/s41601-018-0080-6)

Creative Commons License
CC BY 4.0

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Wang, D., & Lu, K. (2018). Design optimization of hydraulic energy storage and conversion system for wave energy converters. *Protection and Control of Modern Power Systems*, 3(1), Article 7.
<https://doi.org/10.1186/s41601-018-0080-6>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

DEEP LEARNING

– et gennembrud inden for kunstig intelligens

Forfatterne

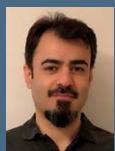


Morten B. Jensen er ph.d.-studerende
mboj@create.aau.dk



Chris H. Bahnsen er ph.d.-studerende
cb@create.aau.dk

Morten og Chris forsker i anvendelsen af deep learning til automatisk trafikanalyse.



Kamal Nasrollahi er lektor og forsker i computer vision og deep learning-metoder til automatisk analyse af ansigter.
kn@create.aau.dk



Thomas B. Moeslund er professor og leder af forskningsgruppen.
tbm@create.aau.dk

Alle arbejder i Visual Analysis of People Laboratoriet ved det tekniske fakultet for IT og Design på Aalborg Universitet. Forskningen fokuserer på automatisk analyse af billeder og video ved brug af computer vision og machine learning-metoder.



I løbet af de seneste 10 år er kunstige, neurale netværk gået fra at være en støvet, udstødt teknologi til at spille en hovedrolle i udviklingen af kunstig intelligens. Dette fænomen kaldes deep learning og er inspireret af hjernens opbygning.

Hvordan kan en computer vinde over verdensmesteren i GO, hvor der er flere mulige kombinationer på spillepladen end atomer i universet?

Hvordan kan en bil forstå, at der er en fodgænger foran den og selv bremse?

Svaret på denne type spørgsmål er *intelligente computersystemer*, der lærer ved at analysere data – rigtig meget data. Den nyeste metode indenfor dette forskningsområde kaldes *deep learning*. Metoden har på få år revolutioneret store dele af forskningsverdenen og er nu på vej ud i alle grene af samfundet, hvor den forventes at få afgørende betydning.

Et gammelt ordsprog siger, at viden er magt! Måske er data et af de vigtigste elementer i dannelsen af viden, men hvordan man styrer og udnytter data, er endnu vigtigere.

Derfor har forskere altid forsøgt at udvikle avancerede måder til indsamling af data for derefter at udnytte det bedst muligt. For at finde inspiration til at udvikle bedre databehandlingsteknikker har forskere kigget på hjernens opbygning og opførsel i håb om at kunne opnå en forståelse, der efterfølgende kan implementeres i computere. Dette forskningsområde kendes også som kunstig intelligens. På grund af hjernens komplekse struktur har det altid været meget udfordrende at forstå hjernens grundlæggende funktionalitet for derefter at opbygge et hjerne-lignende system. På trods af, at ingeniører længe har formået at konstruere systemer, der kan efterligne hjernen ved simple opgaver, så har forskere stødt hovedet mod muren, når det kom til at konstruere systemer, der er i stand til at løse mere udfordrende opgaver, for eksempel genkendelse af objekter.

Imidlertid har nylige fremskridt inden for dataindsamling og rå processe-

ringskraft gjort det muligt at bygge systemer baseret på kunstig intelligens, der kan løse komplekse problemer som objekt-detektion, -genkendelse og tracking. Systemerne er nu så gode, at de i nogle tilfælde klarer sig bedre end menneskelige eksperter.

Disse systemer bliver trænet ved hjælp af massive datamængder gennem matematiske algoritmer, der er bedre kendt under paraplybetegnelsen *deep learning*. Før vi kommer nærmere ind på det, må vi en tur omkring hjernen for at få en grundlæggende forståelse af disse systemer.

Hjernen

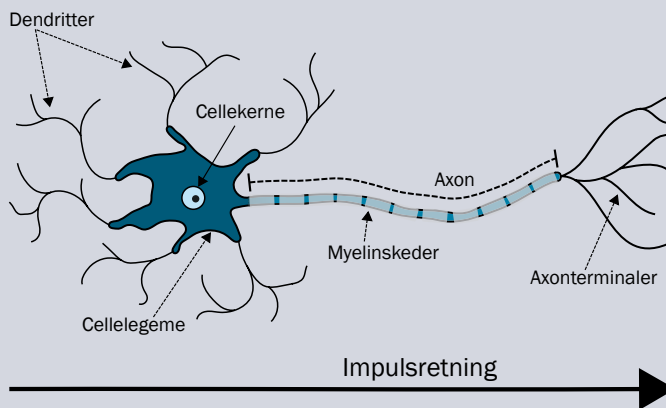
Hjernen er en af de mest komplekse strukturer, vi kender. Den er opbygget af 100 milliarder celler kaldet neuroner, og der er cirka samme antal neuroner i hjernen, som der er stjerner i Mælkevejen.

En gruppe af disse små hjerneneuroner, der er internt forbundet med

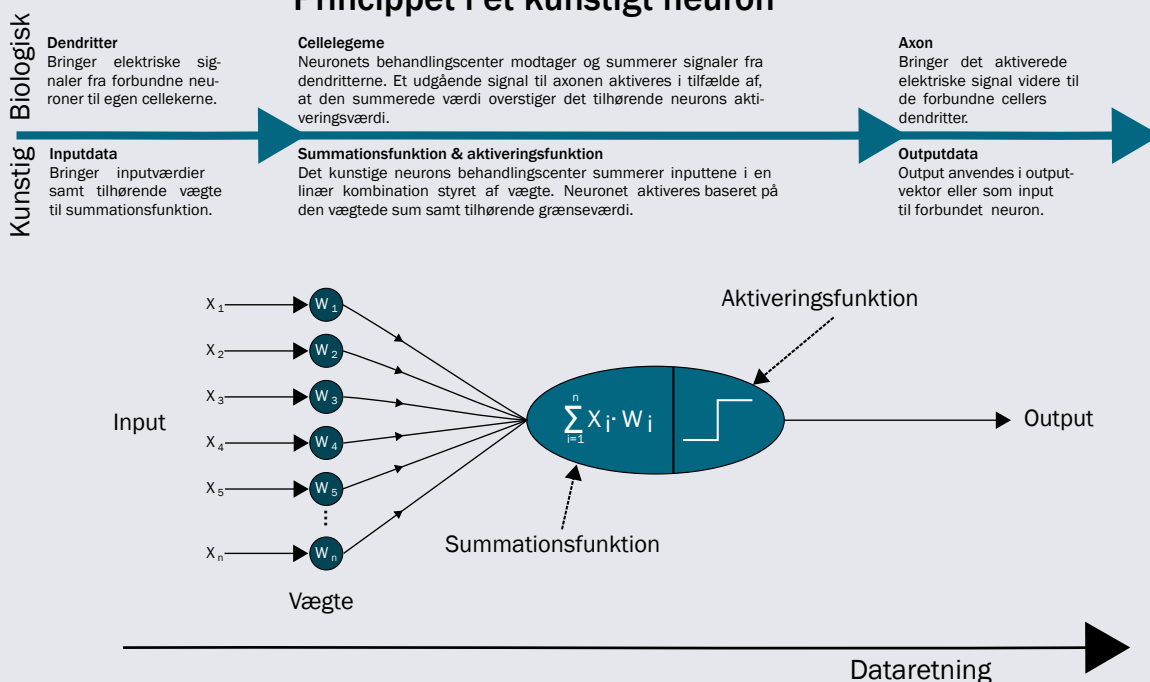
Neuroner er hjernens byggesten

Hvert neuron har:

- et cellelegeme, indeholdende kernen, som er neuronets behandlingscenter.
- et sæt indgangsforbindelser, dendritter, som bringer signaler fra de andre neuroner til kernen i det nuværende neuron.
- en axon, som overfører resultaterne af behandlingen af indgangssignalerne i kernen til de neuroner, der er forbundet til det aktuelle neuron via sine udgangsforbindelser (axonterminaler).



Princippet i et kunstigt neuron



hinanden, er ansvarlige for at udføre en specifik opgave. For eksempel udføres matematiske operationer i en bestemt del af hjernen, mens følelser opfattes af en anden gruppe neuroner. Ved løsning af specifikke opgaver viser de ansvarlige grupper af neuroner mere elektrisk aktivitet end resten af hjernen. Disse elektriske aktiviteter skyldes frigivelse af kemiske stoffer mellem neuronerne, der er internt forbundet med hinanden. Hvis summen af kemiske stoffer ved neuronet er større end et bestemt niveau, bliver neuronet aktiveret. I modsat fald forbliver det passivt.

Når vi som menneske prøver at lære en bestemt opgave, for eksempel

når en baby lærer at gå, gennemføres denne læring gennem adskillige forsøg. Under disse forsøg lærer hjernen, eller rettere: En specifik gruppe neuroner lærer, hvordan de skal aktiveres for at udføre den specifikke opgave. Mængden af de kemiske stoffer, der frigives mellem neuronerne, definerer graden af forbindelse, også kaldet vægtningen, mellem de tilsluttede neuroner.

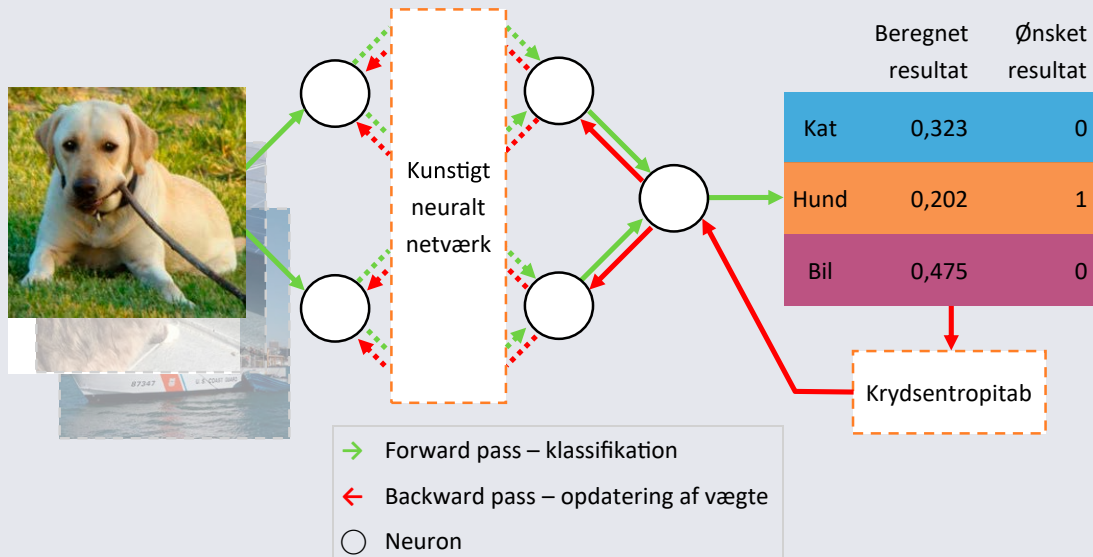
Man kan simulere det biologiske neuron med en matematisk funktion, der består af en lineær kombination af alle inputs til neuronet. Den lineære kombination styres af vægtene af de enkelte inputs. Denne sum af vægtede input svarer

til mængden af de kemiske stoffer, der kommer til et neuron. Derefter bestemmer en såkaldt aktiveringsfunktion, om neuronet skal aktiveres eller forblive passivt. Hvis den vægtede sum er større end en given grænseværdi, aktiveres neuronet.

Et kunstigt neuralt netværk består af en kombination af disse kunstige neuroner i forskellige lag, der er internt forbundet med hinanden gennem vægtede forbindelser. Antallet af lag beskriver *dybden* af netværket. Man betegner et neuralt netværk som dybt, hvis det indeholder tre eller flere lag.

Aktiveringsfunktioner spiller en

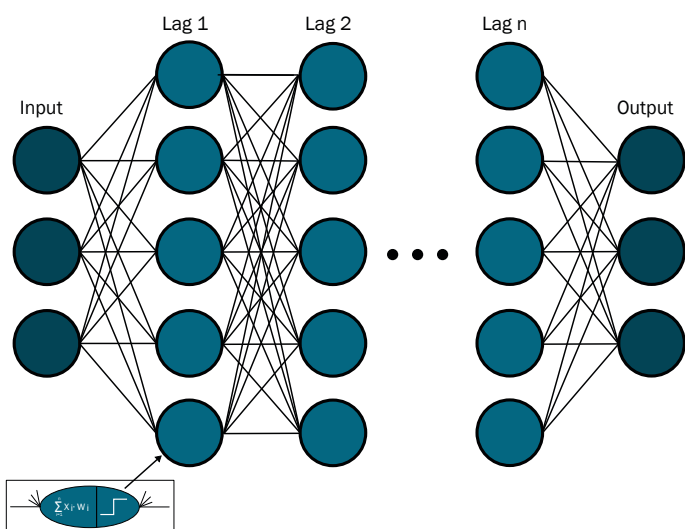
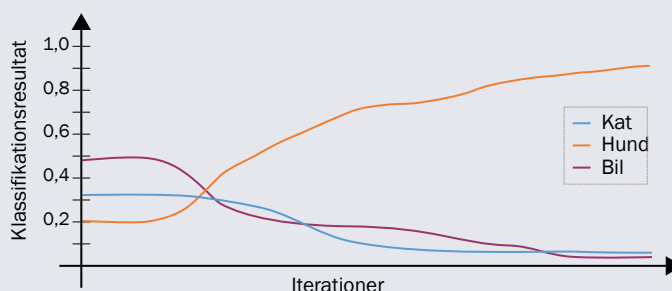
Træning af et kunstigt neuralt netværk



Netværket trænes ved at udregne dets respons for en række billeder (grønne pile), hvor vi på forhånd har defineret det ønskede resultat (annoteret data). Forskellen mellem det ønskede resultat og det beregnede resultat beregnes i det såkaldte krydsentropitab, som føres baglæns gennem netværket for at opdatere dets vægte (røde pile).

To af nøgleordene bag denne læringsalgoritme er differentiability og kædereglen for differentiering af sammensatte funktioner. Alle de neuroner, som et kunstigt neuralt netværk er sammensat af, er grundlæggende (stykvis) differentierbare funktioner. Det betyder, at vi kan flytte det samlede netværks opførsel ved, neuron for neuron, at finde gradienten for den partielt differentierede funktion, opdatere funktionens vægte på baggrund heraf, og føre gradienten videre til de neuroner, som funktionen er forbundet til. Denne proces gentages for hver iteration, indtil alle neuroner er opdateret.

Iterativ læring af et kunstigt neuralt netværk. I starten resulterer netværket ikke i andet end støj, men jo flere annoterede billeder, der køres igennem netværket, jo bedre bliver det til at klassificere billeder af hunde som "hund".



I et Multi-Layer Perceptron (MLP) netværk er alle neuroner forbundet til hinanden imellem lagene.



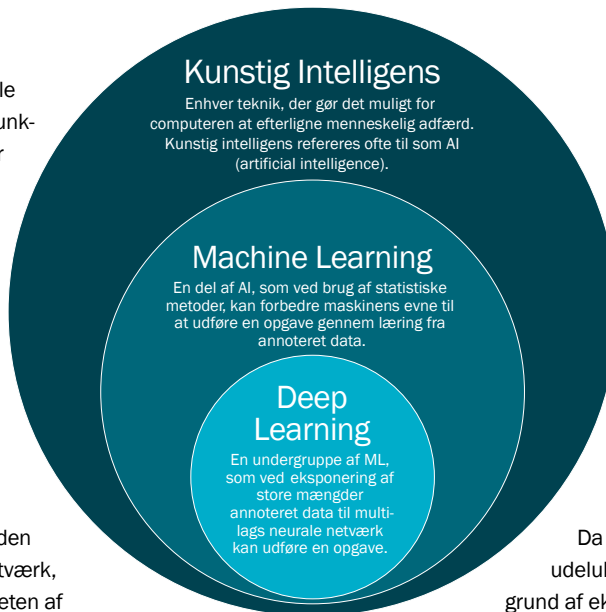
Overlejrerede annoteringer af objekter i trafikken, hvor hver farve indikerer en kategori. Denne type data er fx vigtig for træningen af selvkørende biler.

nøglerolle i kunstige neurale netværk. Hvis aktiveringsfunktionen udelukkende består af lineære funktioner, kan det kunstige neurale netværk udelukkende beskrive lineære fænomener, og dets samlede funktion kan grundlæggende beskrives af én stor matrix. Hvis aktiveringsfunktionen derimod ikke kan beskrives som en lineær kombination af dens input, er det kun dybden af det kunstige neurale netværk, der begrænser kompleksiteten af de funktioner, som netværket kan beskrive.

Læring

For at lære et kunstigt neuralt netværk at udføre en specifik opgave kræves en læringsalgoritme, hvis formål er at finde de rette vægte mellem netværkets neuroner. Vægtene læres gennem adskillige iterationer, hvor det neurale netværk præsenteres for store mængder træningsdata. Hver enkelt stykke data er annoteret, det vil sige, at det er parret med den ønskede respons fra det neurale netværk – for eksempel at et billede af en hund hører til kategorien "hund", hvis formålet med det neurale netværk er at genkende objekter i billeder. Når billedet er kørt igennem det neurale netværk, giver netværket dets bud på hvilken kategori, billedet tilhører. Herefter udregnes forskellen mellem det beregnede og det ønskede resultat, hvilket kaldes krydsentropitabet. Det beregnede krydsentropitab fødes herefter baglæns ind i det neurale netværk og opdaterer vægtene i retning af det ønskede resultat.

I starten resulterer det neurale netværk ikke i andet end støj. Men ganske langsomt, iteration for iteration, lærer netværket at tilpasse sig det pågældende træningsdata. Når det beregnede resultat konvergerer mod det ønskede resultat, er træningen afsluttet og netværket er nu specialiseret i at klassificere datasættet. Hvis datasættet indeholder tilstrækkeligt mange annoterede billeder



og er repræsentativt for de ønskede kategorier, for eksempel hunde og katte, har man nu en udmærket hunde- og kattedetektor.

Hvorfor først nu?

Kunstige neurale netværk går tilbage til 1940'erne. Disse netværk var imidlertid ikke særlig populære i samtiden på grund af den beregningsmæssige kompleksitet og manglende træningsdata.

Den beregningsmæssige kompleksitet skyldes, at et netværk, der kan løse praktiske problemer, indeholder mindst tre lag med mange neuroner i hvert lag, der er internt forbundet med hinanden. Denne type netværk er kendt som en Multi-Layer Perceptron (MLP). I en MLP er hvert neuron i et lag forbundet til alle neuroner i det næste lag af netværket. Dette fænomen, der er kendt som fuldt forbundne netværk, resulterer i store matricer, der beskriver vægtningen af neuronernes forbindelser. De store matricer fører igen til beregningsmæssigt krævende læringsalgoritmer, der i mange år var for store til, at computere kunne håndtere dem. Dette ændrede sig dog med introduktionen af såkaldte Graphics Processing Units (GPU) i 1990'erne, som tilbød hurtig og parallel databehandling. Brugen af GPU'er har gjort det muligt at implementere neurale netværk i praksis, hvorefter deres popularitet kun er steget. Faktisk er neurale netværk nu blandt de allerbedste værktøjer, der er i stand

Sammenhæng mellem Kunstig Intelligens, Machine Learning og Deep Learning.

til at løse meget komplicerede problemer som billedbaseret objektgenkendelse. Denne succes skyldes imidlertid ikke kun udviklingen af bedre GPU'er, men også tilgængeligheden af enorme mængder data.

Da kunstige neurale netværk udelukkende kan lære på baggrund af eksempler, er tilgængeligheden af eksempeldata kritisk. Jo mere data, jo bedre er læringsprocessen. Imidlertid var store databaser ikke så almindelige for blot 10 år siden. Men siden 2010 er enorme databaser gradvist blevet opbygget. Et eksempel er ImageNet, der består af cirka 14 millioner annoterede billeder, inddelt i mere end 20.000 forskellige kategorier.

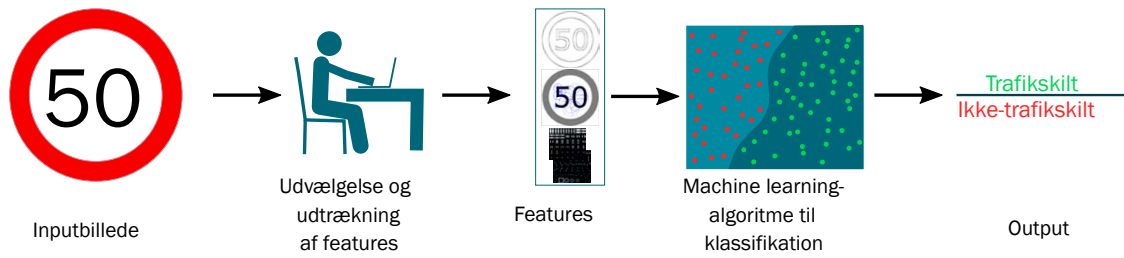
De fleste billeder i ImageNet indeholder kun én annotering, det vil sige at hele billedet tilhører én kategori. En anden og mere omfattende måde at annotere billederne på er at definere det specifikke område i billedet, der indeholder et givent objekt – for eksempel fodgængere, cyklister, biler og trafikskilte.

Fra machine learning til deep learning

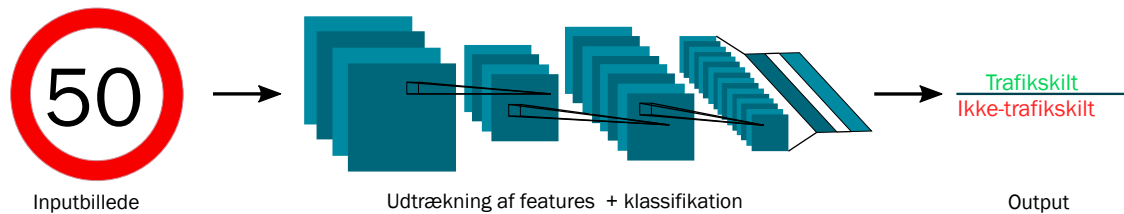
Traditionelle machine learning-teknikker er baseret på at udvikle og udvælge specifikke karaktertræk, også kaldet features, ved de objekter, man ønsker at finde og genkende. Det har betydet, at forskere tidligere har brugt tid på manuelt at definere features, som efter deres vurdering var unikke og gav en god repræsentation af de ønskede objekter.

Til detektion af trafikskilte vil man i traditionel machine learning udvælge features, der kan beskrive skiltets cirkulære form og dets karakteristiske røde kant. Herefter udvælger man manuelt en eller flere metoder, der kan konvertere de ønskede fea-

Traditionel machine learning



Deep learning



Sammenligning af workflowet for traditionel machine learning og deep learning. I modsætning til traditionel machine learning kræver deep learning ikke manuel udvælgelse af features.

tures til en matematisk repræsentation. Disse features bruges til at træne en machine learning-algoritme, som benytter de udregnede features til at skelne mellem trafikskilte og ikke-trafikskilte.

Til trods for, at det i sidste ende er computerens algoritmer, der udregner det endelige resultat, indebærer traditionel machine learning en del manuelt arbejde med at definere hvilke features, der er relevante. Deep learning har til forskel fra machine learning ingen behov for menneskelig indblanding i forbindelse med udvælgelse og udformning af features.

Det kræver dog stadig menneskelig indblanding, når deep learning-netværkene skal designes. Det gøres for eksempel ved at definere netværkets størrelse, det vil sige hvor mange lag, netværket skal bestå af.

Et lag består af en række funktioner. Den vigtigste funktion i moderne neurale netværk er en såkaldt convolution (på dansk en foldning), og derfor kaldes disse netværk også Convolutional Neural Networks (CNN'er). Convolution er en matematisk operation, der benytter sig af et filter. Filtrenes overordnede funktion er at trække features ud af inputbilledet, og et moderne neuralt

netværk indeholder rigtig mange filtre, der er grupperet i flere convolution-lag. Populære deep learning-netværk som AlexNet, VGG, GoogLeNet og Microsoft ResNet er alle CNN'er, og de indeholder henholdsvis 8, 19, 22 og 152 lag. Outputtet fra convolution kaldes et feature map, og det er udgangspunktet for et andet meget anvendt lag i neurale netværk kaldet pooling. Denne funktion reducerer størrelsen på de feature maps, der genereres fra convolution-laget, således at man opnår en mere kompakt repræsentation.

En af årsagerne til, at CNN fungerer så godt, er, at netværkene selv kan lære at sammensætte både simple og komplekse features i deres convolution-lag – uden at man som operatør specifikt beder dem om at gøre sådan.

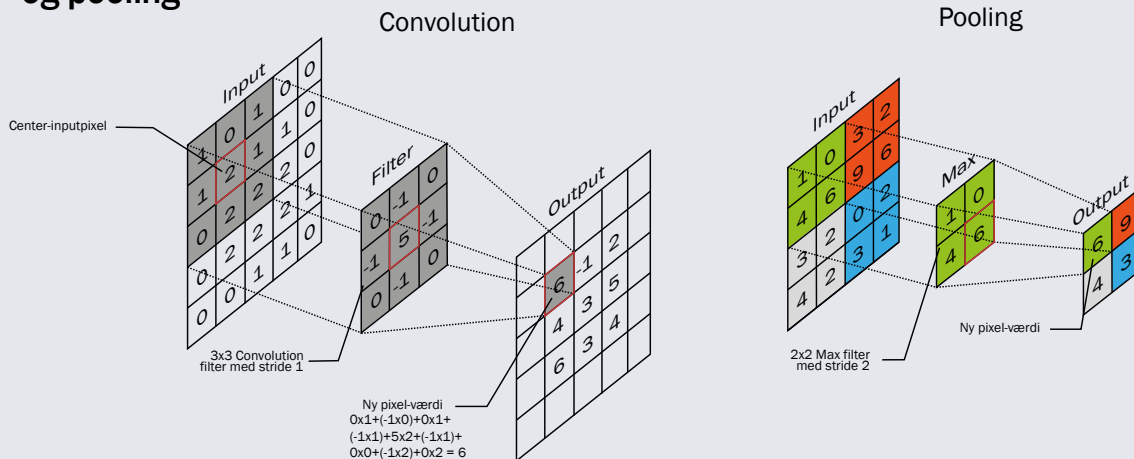
Tidkrævende træning

Hvert convolution-filter er specialiseret i at genkende et specifikt mønster, og der skal derfor en kombination af mange filtre til at genkende forskellige objekter. Det populære AlexNet-netværk, som startede deep learning-bølgen, består af fem convolution-lag. Det første convolution-lag indeholder 96 forskellige filtre, som hver har en størrelse på 11x11 pixels. Dette producerer 96 feature maps med dimensionerne

55x55 pixels, svarende til 290.400 neuroner. Hver enkelt af disse neuroner har $11 \times 11 \times 3 = 363$ vægte og 1 bias. Det første convolution-lag alene har dermed $290.400 \times 364 = 105.705.600$ forbindelser. Totalt set har de 5 convolution-lag i AlexNet over 600 millioner forbindelser med over 60 millioner parametre, der bliver justeret gennem de mange træningsiterationer med store mængder træningsdata. Det er de enormt mange forbindelser og parametre, der gør CNN'er i stand til at tilpasse sig mange forskelligartede opgaver, hvis bare de bliver trænet på tilstrækkelige store mængder af nøjagtigt og repræsentativt annoteret data.

Men denne træning er en tidskrævende opgave, der indebærer mange gentagne gennemkørsler af hele det annoterede data. På en nyere gamer-computer vil det derfor tage over en uge at træne AlexNet på ImageNet-databasens 14 millioner billeder. Det er i den forbindelse værd at nævne, at AlexNet er et ældre netværk med 8 lag, hvor Microsofts nyere ResNet indeholder 152 lag. Hvis man ønsker at træne dette netværk, må man gange træningstiden med 19. Store netværk, store datamængder og store krav til processeringskraften hænger derfor uløseligt sammen.

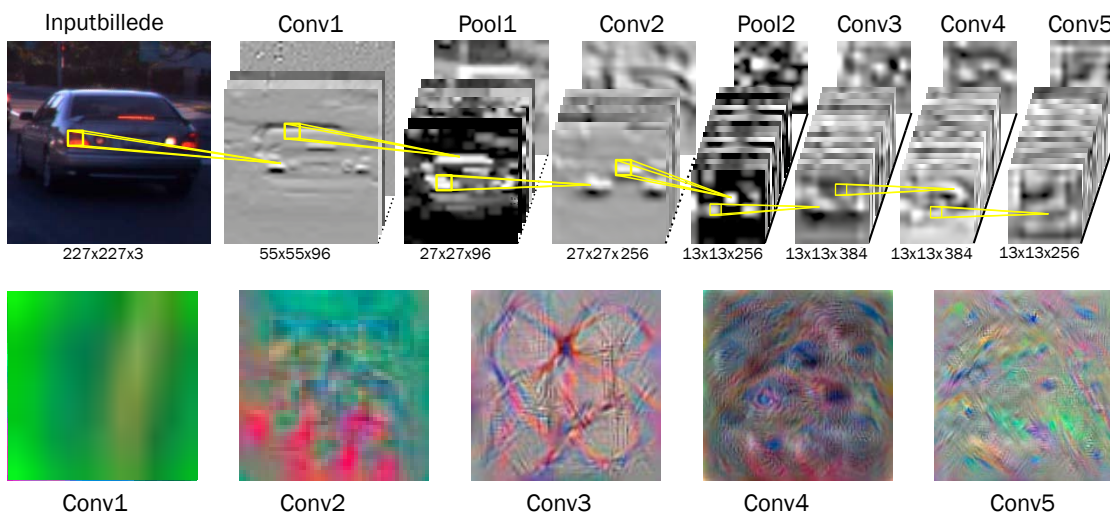
Convolution og pooling



De vigtigste funktioner i moderne neurale netværk er såkaldt convolution og pooling, som er illustreret i figuren. I eksemplet på convolution benytter vi et filter af størrelsen 3x3 pixels med udgangspunkt i den pixel der er markeret med rød ring. Convolution består i, at man anvender 3x3 filteret på den "røde" centerpixel samt dets nabopixels, illustreret med det grå område i inputmatricen. Den resulterende pixelværdi i outputmatricen opnås ved at gange filterets vægte på de tilhørende pladser i inputmatricen for derefter at summere resulta-

tet. Herefter rykker vi vores 3x3 filter én gang til højre og gentager udregningen.

I eksemplet på pooling bruges her et såkaldt max pooling-filter med en størrelse på 2x2 pixels og en stride på 2 pixels. En stride på 2 pixels betyder, at vi flytter filteret 2 pixels til højre efter hver operation. Maxfilteret undersøger alle værdierne i et 2x2 område, tager den højeste værdi heri og smider de øvrige værdier væk. Den højeste værdi udgør nu den nye pixelværdi i outputmatricen.



Eksempel på et AlexNet-inspireret neuralt netværk med 5 convolution-lag og 2 pooling-lag samt dertil hørende lærte features for hvert convolution-lag (øverst). Conv-lagene ændrer sig gradvis fra at være ret genkendelige omrids af bilen i conv1 til mere komplekse strukturer i conv5, som knap nok er genkendelige for det menneskelige øje.

Den nederste del af figuren visualiserer, hvad 1 tilfældigt udvalgt feature-map i hvert af de 5 convolution-lag reagerer på i billedet. Conv1-features repræsenterer kanter og farveforskelle i forskellige retninger, mens de senere convolution-lag repræsenterer komplekse, specialiserede mønstre.

Ikke begrænset af menneskelige sanser

Det ser i store træk ud til, at deep learning og kunstige neurale netværk virker som hjernen ved løsning af bestemte opgaver. Det betyder, at kunstig intelligens i princippet vil kunne klare det samme som et

menneske. Potentialet er dog endnu større for den kunstige intelligens, da dets input ikke er begrænset til de menneskelige sanser, men vil kunne opfatte verden gennem et utal af sensorer og have adgang til ufattelige mængder information. For at nå så langt kræves der dog bety-

delige fremskridt indenfor udvikling af læringsalgoritmer og håndtering af information. På vejen dertil vil deep learning ændre fremtiden i mange forskellige applikationer og sektorer, fra sikkerhed og finans til medicin og transport. Vi er glade for at være en del af dette spændende eventyr. ■