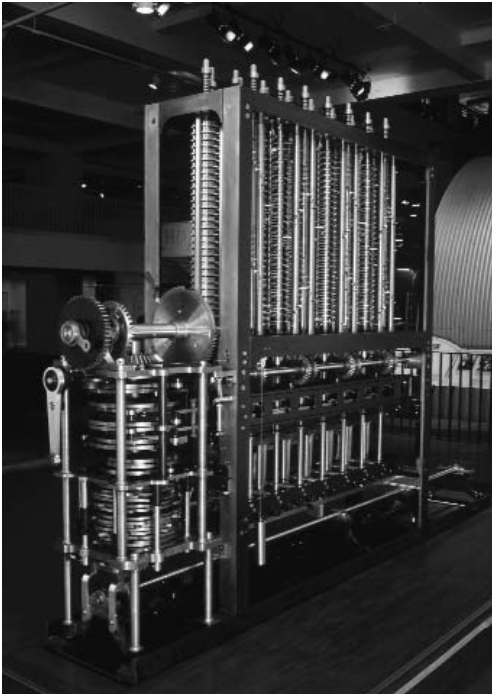


1

Tandwiel, stroom of lichtdeeltje.
Onder de motorkap van je computer

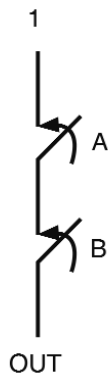


Babbage' *Difference Engine No. 2* in de collectie van
het London Science Museum.

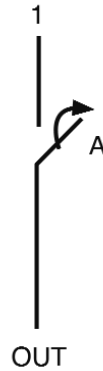
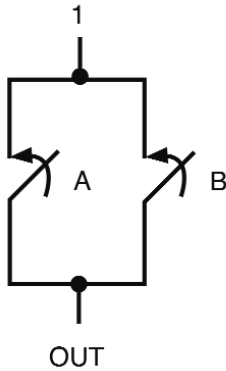


Twee tandwielen uit de *Difference Engine No. 2*.

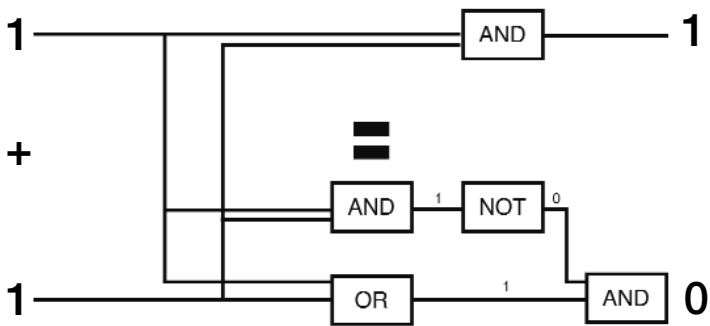
Zonder handen! Stroompjes in huidige computers



De bedrading voor een AND-poort: de stroom gaat alleen van bovenaf door naar de uitgang beneden als beide poortjes (A en B) dicht zijn, dus onder stroom staan.

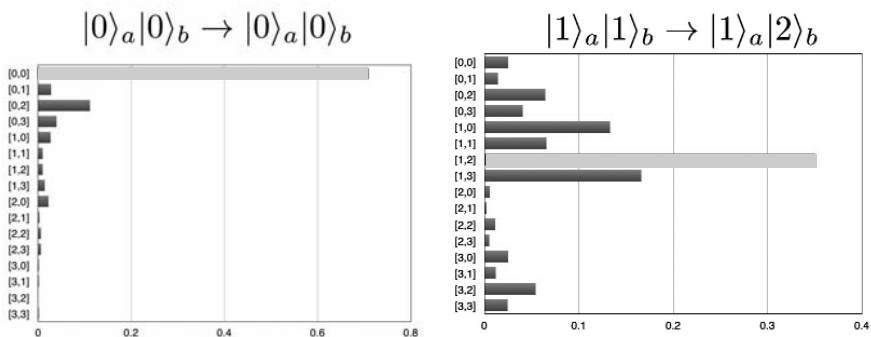


Links een OR-poort, met twee losse paden van invoer tot uitvoer.
 Rechts een NOT-poort, waarbij de A-schakelaar juist dichtgaat als er géén stroom door draadje A loopt.



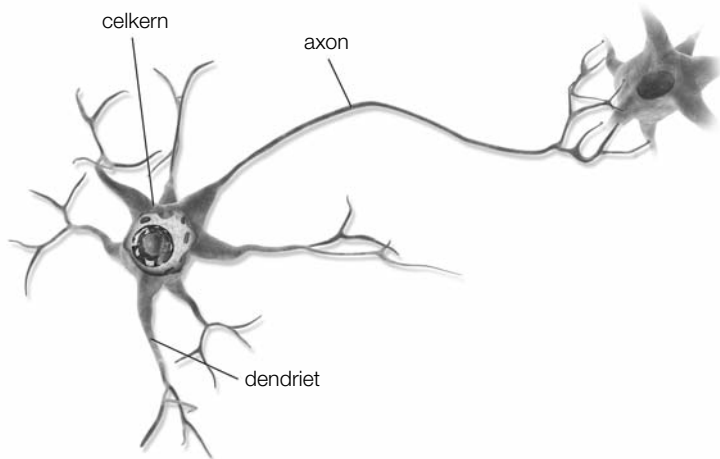
$1 + 1 = 10$ in een processor. Links komt de stroom binnen (beide wel, want twee enen) en via de poortjes komt rechts de juiste uitkomst: boven wel stroom (1) en onder niet (0). Na iedere poort staat of er wel of geen stroom in de draad te vinden is met een 1 of 0.

De stroom voorbij: kort over kwantumcomputers



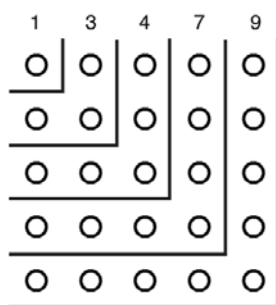
De uitkomsten voor $0 + 0 = 0$ (links) en $1 + 1 = 2$ (rechts) van een werkende kwantumcomputer.

Informatie in de hersenen



De opbouw van een neuron.

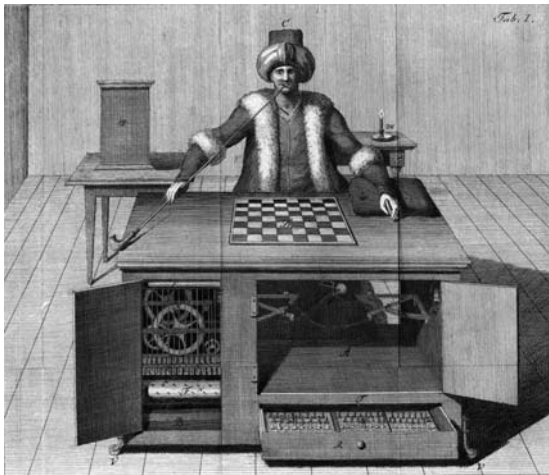
Twee soorten wiskunde



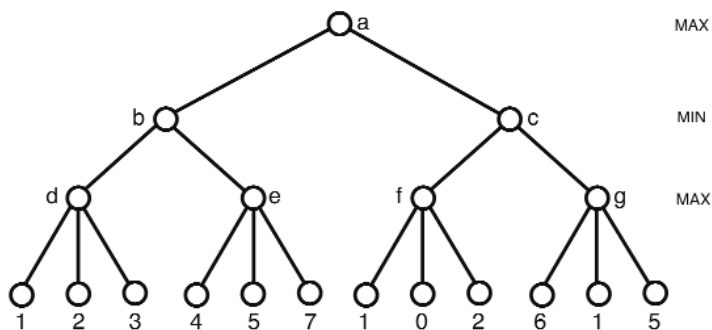
De som van de oneven getallen.

2

Zoek je een ongeluk. Het verleden van kunstmatige intelligentie



De mechanische Turk.

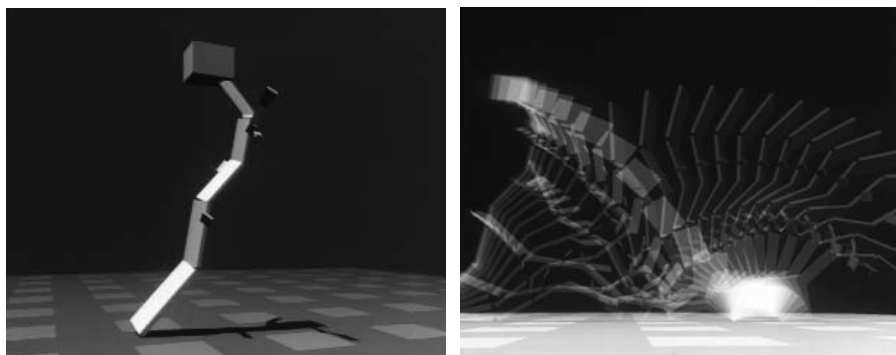


De verschillende uitkomsten waar je vanuit *a* op uit kunt komen.
 Minimax zoekt de beste zet die je bovenaan kunt maken.

Regel-matige therapie

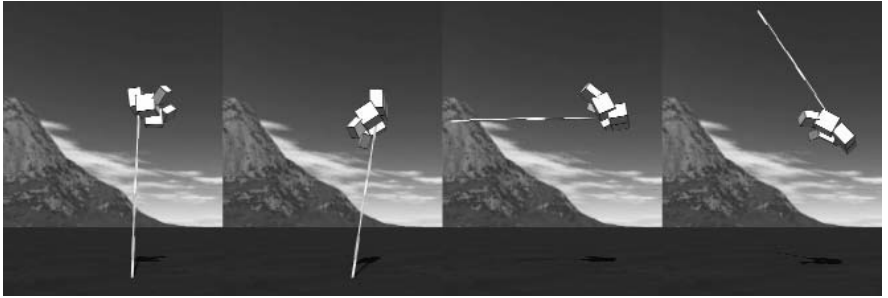
De hoogtepunten en ondergang van
 de expliciete regels

Met vallen en opstaan: evolutionaire algoritmes



De beste manier om je voort te bewegen volgens een
 evolutionair algoritme.

Overall competitie



Polsstokhoogspringen, maar dan anders.

3

Neurale netwerken in beeld



school bus 1.0 garbage truck 0.99 punching bag 1.0 snowplow 0.92



motor scooter 0.99 parachute 1.0 bobsled 1.0 parachute 0.54



fire truck 0.99 school bus 0.98 fireboat 0.98 bobsled 0.79

Computers kunnen simpele voorwerpen nauwelijks herkennen als je ze een beetje draait.

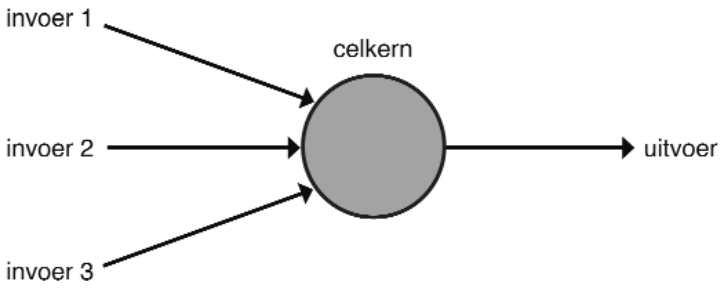
Met kunstmatige neuronen handschrift ontcijferen



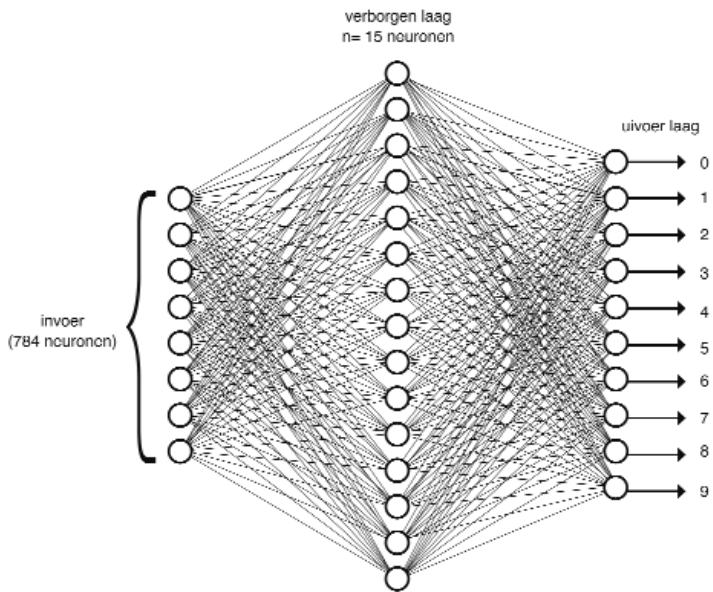
Voorbeelden waarmee het neurale netwerk leert om geschreven getallen te herkennen.



Vier verschillende patronen waarvoor een neurale netwerk (tegelijktijd) gevoelig kan zijn.

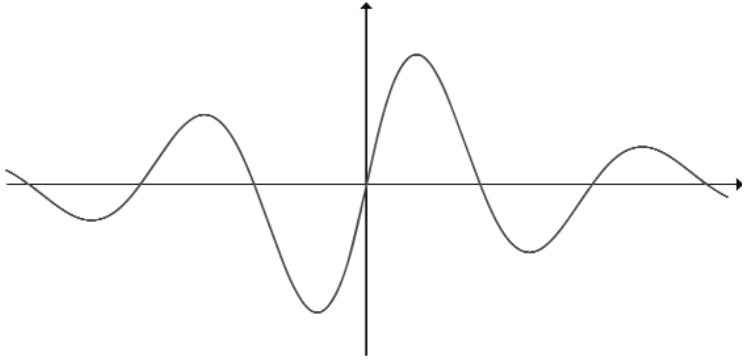


Een kunstmatig neuron met invoer (linkerpijlen), rekengedeelte (cirkel) en uitvoer (rechterpijl).



Een neuraal netwerk voor het herkennen van geschreven getallen.

Kleine stapjes naar beneden:
hoe een neurale netwerk leert



De leermethode achter neurale netwerken: kleine stapjes naar beneden zetten totdat het netwerk in een dal uitkomt (met de hoogte gelijk het aantal fouten van het netwerk).

Beperkte generalisaties in neurale netwerken

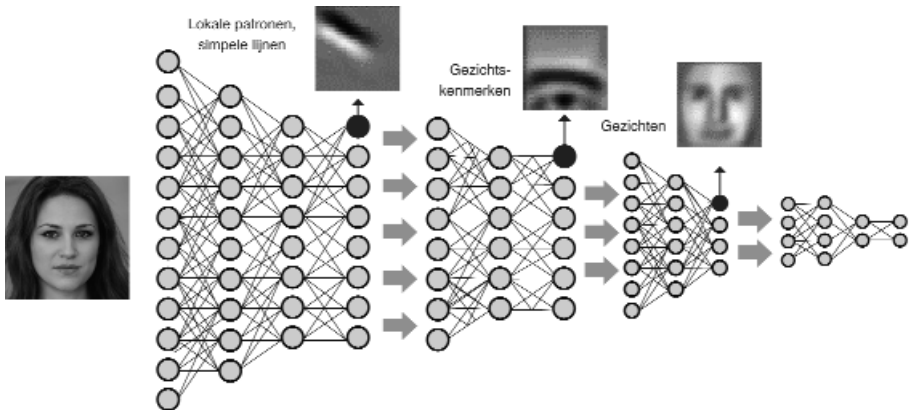


Een stopbord met vier stickers erop. Het gevolg: geen zelfrijdende auto die nog snapt dat je moet stoppen.



Gezichten gezien door apen (links) en de reconstructie op basis van de stroompjes in hun hersenen (rechts). Door 205 neuronen in de gaten te houden was de reconstructie vrijwel perfect.

Hersenen imiteren op de computer



Gezichtsherkenning met een convolutional neuraal netwerk. Van heel kleine, lokale patronen links gaat het algoritme stapsgewijs (door herhaalde patroonherkenning) naar meer complexe, globale patronen.

Een CNN, gebruikt voor van alles



Welk recept zou een neuraal netwerk hiervoor opstellen?

4

Voorspelbare gesprekken. Neurale netwerken en taal

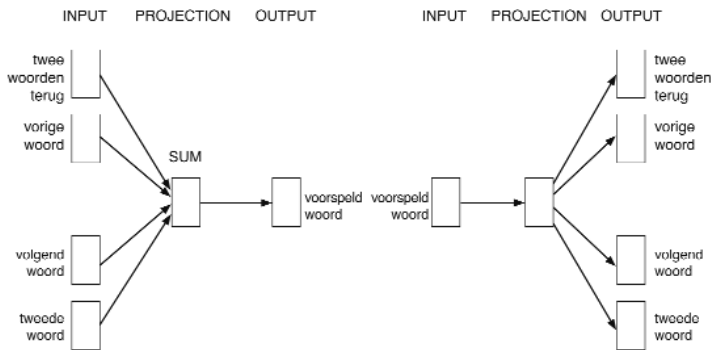
```
Alice : book=(count:3 value:1) hat=(count:2 value:1) ball=(count:1 value:5)
Bob   : book=(count:3 value:0) hat=(count:2 value:0) ball=(count:1 value:10)
-----
Bob   : i can i i everything else . . . . .
Alice : balls have zero to me to me to me to me to me to me to me to me to
Bob   : you i everything else . . . . .
Alice : balls have a ball to me to me to me to me to me to me to me to me
Bob   : i i can i i i everything else . . . . .
Alice : balls have a ball to me to me to me to me to me to me to me to me
Bob   : i . . . . .
Alice : balls have zero to me to me to me to me to me to me to me to me to
Bob   : you i i i i everything else . . . . .
Alice : balls have 0 to me to me to me to me to me to me to me to me to
Bob   : you i i i everything else . . . . .
Alice : balls have zero to me to me to me to me to me to me to me to me to
```

De ‘eigen taal’ ontwikkeld door het neurale netwerk van Facebook: een gevolg van een op hol geslagen trainingssessie met zichzelf.

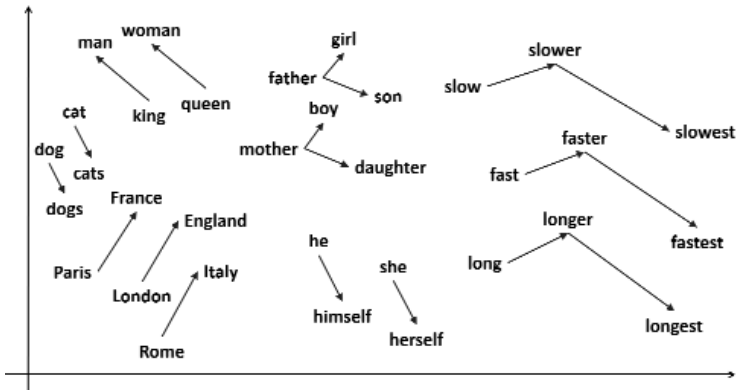
Woordenschat en swipen



Drie potentiële matches op Tinder, met rechtsonder hoe er geswipt is.

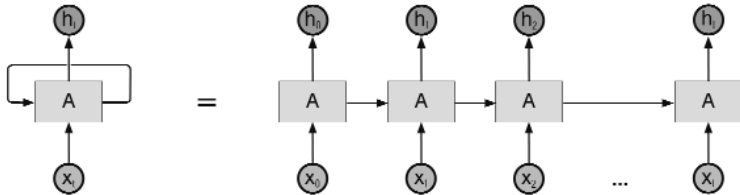


De twee versies van Word2Vec: voorspel het missende woord aan de hand van de rest van de zin (links) of de rest van de zin aan de hand van een woord (rechts).



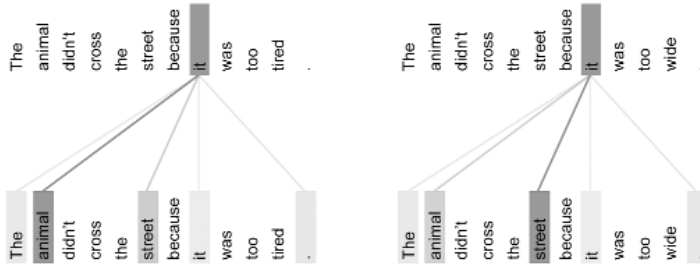
Onderlinge woordverhoudingen bij Word2Vec. De afstand tussen de getallen voor koning en man is hetzelfde als die tussen de getallen voor koningin en vrouw enzovoort.

Nepnieuws uit een chip



Een recurrent neurale netwerk. Links zie je de invoer die vanaf onder binnenkomt en verwerkt wordt door het netwerk om er bovenaan als uitvoer uit te komen. Tegelijkertijd wordt het resultaat teruggestuurd als extra invoer voor de volgende ronde. Rechts zie je hoe dat werkt van woord tot woord: de berekening A van woord x_0 telt mee voor de berekening bij x_1 enzovoorts.

Het tempo van moderne kunstmatige intelligentie: taalsprongen



Waar verwijst 'it' naar? In de linkerzin naar het dier, in de rechterzin naar de straat. Dit stukje van de Transformer pikt dat dankzij de extra context op. 'Animal' is links donkerder gekleurd en 'street' rechts; die woorden (getallenreeksen in het netwerk) wegen zwaarder mee.

De afhankelijkheid van data



Ground truth: Soap Nepal, 288 \$/month

Azure: food, cheese, bread, cake, sandwich

Clarifal: food, wood, cooking, delicious, healthy

Google: food, dish, cuisine, comfort food, spam

Amazon: food, confectionary, sweets, burger

Watson: food, food product, turmeric, seasoning

Tencent: food, dish, matter, fast food, nutriment



Ground truth: Soap UK, 1890 \$/month

Azure: toilet, design, art, sink

Clarifal: people, faucet, healthcare, lavatory, wash closet

Google: product, liquid, water, fluid, bathroom accessory

Amazon: sink, indoors, bottle, sink faucet

Watson: gas tank, storage tank, toiletry, dispenser, soap dispenser

Tencent: lotion, toiletry, soap dispenser, dispenser, after shave

Twee soorten zeep. Computers snappen alleen de meest gangbare vorm in rijke landen.

Wat je al niet kan met kansen!

Filosofisch intermezzo: woorden en de wereld

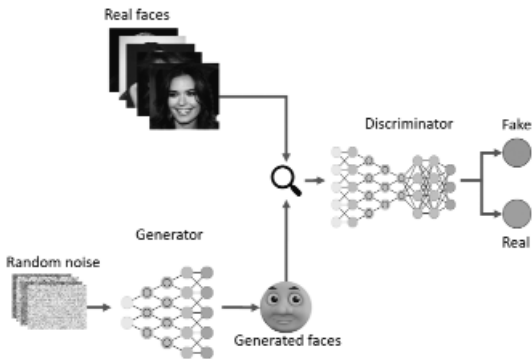
Kunnen we betekenis ooit inprogrammeren?

5 Creativiteit uit een chip



Een gat in de linkerfoto, opgevuld door verschillende neurale netwerken, uit 2009 (links) tot aan 2019 (uiterst rechts).

Vervalser en agent: de werking van een GAN

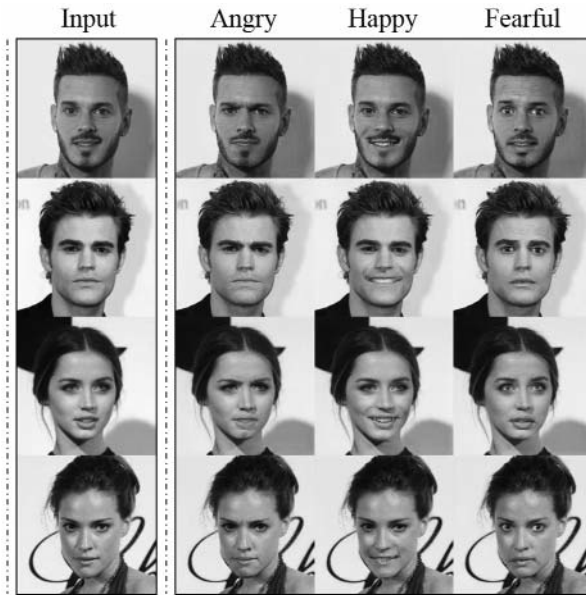


De interne werking van een generative adversarial network (GAN). Linksonder het deel ('generator') dat nieuwe plaatjes creëert. Die worden vergeleken met bestaande foto's door het rechneretwerk ('discriminator').

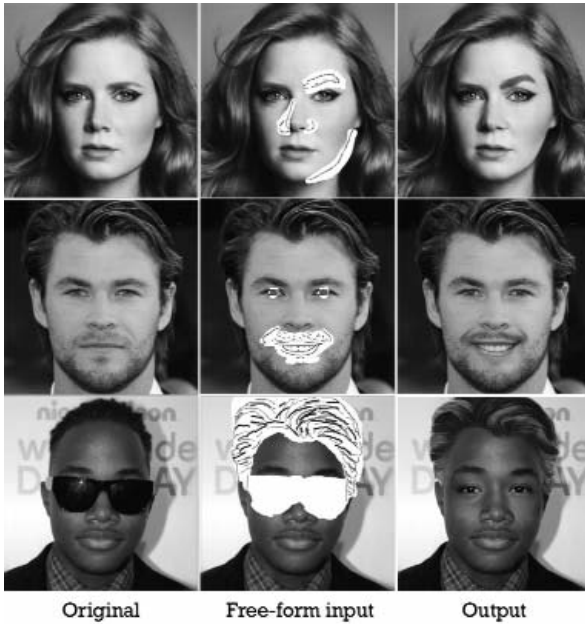


Drie foto's van katten, gegenereerd door hetzelfde programma dat de twee portretten aan het begin van dit hoofdstuk creëerde.

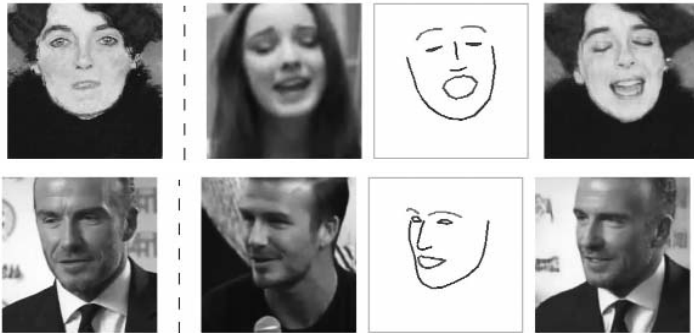
Gerichte manipulatie



Vier voorbeelden van de veranderingen in gezichtsuitdrukking die mogelijk zijn met StarGAN. De linkerkolom is de originele (echte) portretfoto, de portretten in de drie kolommen rechts zijn gemaakt door het netwerk.

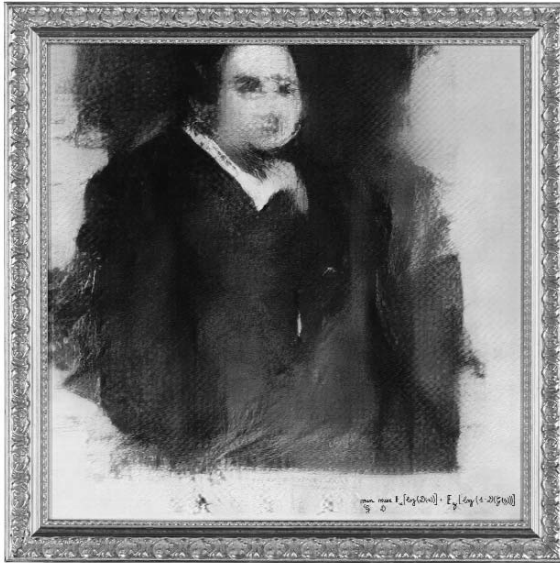


Originele foto's links, die op basis van de schetsen in het midden aangepast worden tot de neppe, door het GAN bedachte, resultaten rechts.



Op basis van één foto (links) kan een GAN, met behulp van filmpjes van mensen die erop lijken (midden), een clip maken van de persoon op de foto (helemaal rechts beelden uit de resulterende clip).

Computerkunst!?



Het portret van de fictieve Edmond Belamy, gesignd met de formule die zegt dat de vervalsers moeten zorgen dat de politieagent zo min mogelijk verschil ziet tussen echte kunst en de vervalsingen.

Gebrekkige transparantie

Creativiteit buiten de kunst: hoe nabij is de robot-apocalyps?



Een originele satellietfoto (links), een door de computer gegenereerde kaart (midden) en een reconstructie van de satellietfoto op basis van de middelste kaart (rechts).



Links een schematische kaart met op allerlei plekken kleine aanpassingen (in het midden beter te zien), die als extra geheugen gebruikt worden door het algoritme. Waardoor de linkerkaart gereconstrueerd wordt tot de rechterfoto.

6

Kunstmatige intelligentie in de (toekomstige) samenleving

De zwakke kanten van neurale netwerken

Datadrang



WHEN YOU TRAIN PREDICTIVE MODELS
ON INPUT FROM YOUR USERS, IT CAN
LEAK INFORMATION IN UNEXPECTED WAYS.

Automatische suggesties van Google kunnen (ook) gevoelige informatie onthouden.

Het menselijke gevaar van kunstmatige intelligentie



Portretfoto's van echte mensen (links) en de reconstructie van hun gezicht op basis van de stem, gemaakt door Speech2Face (rechts).

Computers in plaats van werknemers?

Tot slot

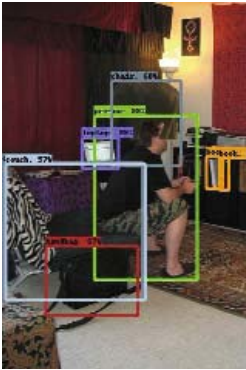
Fotokatern



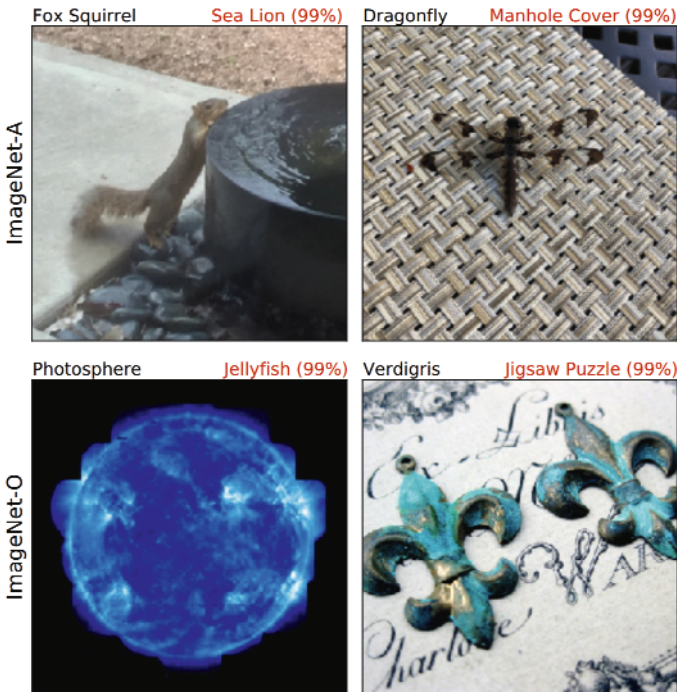
1. Reese Witherspoon (links) wordt normaal correct herkend door neurale netwerken. Voeg een speciaal gekleurde bril toe (midden) en de computer is ervan overtuigd dat het Russell Crowe (rechts) is.



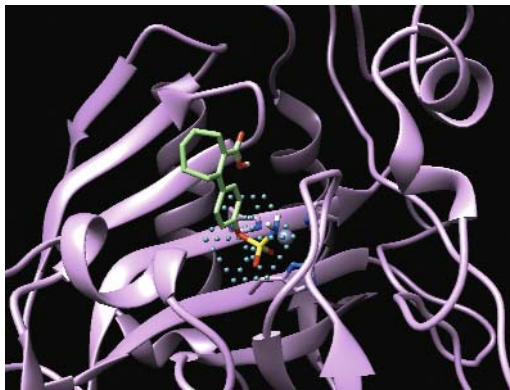
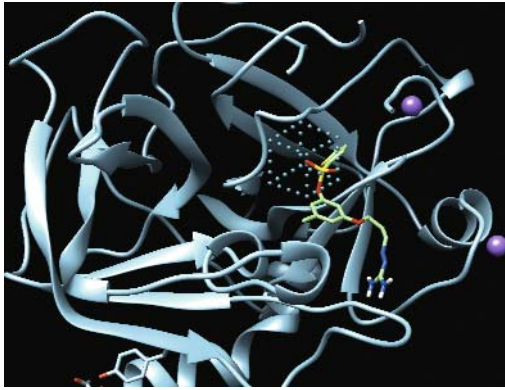
2, 3. Met een (niet zo stijlvolle) bril op ben je voor de computer een filmster én een supermodel, Milla Jovovich in dit geval.



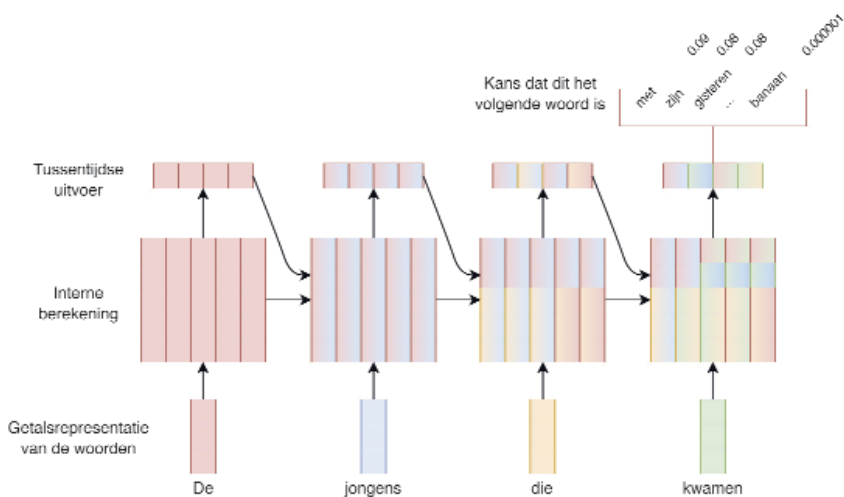
4, 5. Links een huiskamer met een olifant, die blijkbaar een stoel is (de computer herkent ook een aantal dingen, zoals een mok naast de bank, niet meer). Rechts een dubbele kat, met een voorpoot die op de een of andere manier op een hond lijkt.



6-9. Computers herkennen plaatjes niet door de vorm (linksboven en waarschijnlijk rechtsonder), en/of niet door de kleur (linksonder en rechtsboven).



10. Een aantal moleculen, waar de gestippelde kubus laat zien dat een patroonherkenningsstukje van het neurale netwerk consequent dezelfde moleculen opspoorde.



11. Het verwateren van het geheugen van een recurrent neuraal netwerk. Je start links en naarmate het meer woorden heeft gezien (verder naar rechts) is het steeds moeilijker om die originele rode/roze kleur te onthouden.



12, 13. Twee portretfoto's gecreëerd door een neurale netwerk.
Geen van beide personen bestaat.

A small bird with brown and white feathers, and a red head



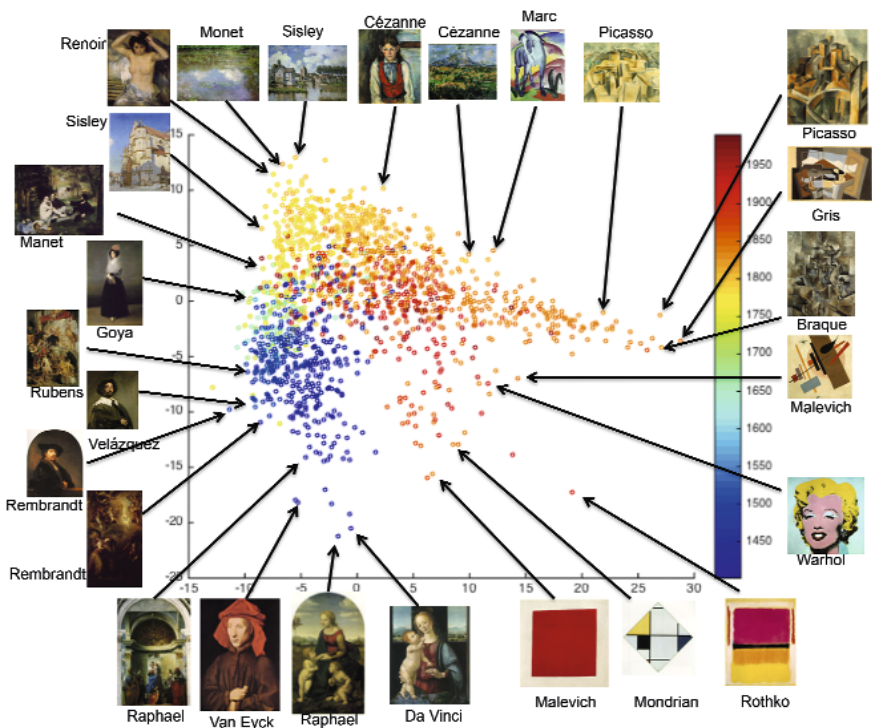
14. Een neurale netwerk kreeg de bovenstaande zin en puur op basis daarvan ontstonden deze foto's.



15. Een kopje espresso gemaakt van een hond, want waarom niet?



16. Schilderijen van CAN die de proefpersonen het meest waardeerden.



17. De verschillende kleuren vertegenwoordigen de kunststijlen. Chronologisch begin je linksonder, met donkerblauw. Vervolgens ontwikkelt de kunst zich met de klok mee, via groen, geel, oranje, rood, totdat ze onderaan uitkomt bij Mondriaan en Malevich.

Links

Hoofdstuk 1

- 1 De werking van Babbages Difference Engine: https://youtu.be/vdra5Ms__9s

Hoofdstuk 2

- 2 Praat met ELIZA: <https://www.masswerk.at/elizabot/>
- 3 De survey van Operation Match: <https://blogs.lib.unc.edu/hill/wp-content/uploads/2017/01/Operation-Match-Survey.pdf>
- 4 De evolutionaire pogingen tot 'lopen': goo.gl/4h7TG8
- 5 En een evolutionair algoritme dat 'springt': <https://youtu.be/N9DLEiakEs>

Hoofdstuk 3

- 6 Artikel over aanpassingen van één pixel: <https://arxiv.org/pdf/1710.08864.pdf>
- 7 Facebook maakt automatisch kaarten: <https://mapwith.ai/>

Hoofdstuk 4

- 8 Dagelijkse portie algoritmisch nepnieuws: <https://thisarticledoesnotexist.com/>
- 9 Schrijf je eigen teksten met een algoritme: <https://gpt2.apps.allenai.org/>

- 10 en: <https://talktotransformer.com>
- 11 Het verhaal van Ronald Giphart, samen met een algoritme geschreven: <https://asibot.nl/giphart>
- 12 Schotse spraakherkenning: <http://youtu.be/XQCHOKAq9xA>
- 13 Schrijf jij als een computer?: <http://gltr.io/dist/index.html>

Hoofdstuk 5

- 14 Neppe portretfoto's: <https://thispersondoesnotexist.com/>
- 15 Kattenfoto's: <https://thiscatdoesnotexist.com/>
- 16 en: <https://thesecatsdonotexist.com/>
- 17 Artikel met make-up-GAN: <https://bit.ly/2JMVJm7>
- 18 Top tien deepfakes: <https://www.youtube.com/watch?v=-QvIX3cY4lc>
- 19 Espresso-honden en andere exotische plaatjes zelf maken: <https://artbreeder.com/>
- 20 Verbeeld je huisdier als andere beesten: <http://nvidia-research-mingyuliu.com/ganimal>
- 21 Automatisch opsporen van deepfakes: <https://www.getsherlockai.com/>
- 22 De nieuwe 'Rembrandt' gemaakt door een computer: <https://www.nextrembrandt.com/>
- 23 componeer zelf muziek met een neurale netwerk: <https://openai.com/blog/musenet/>
- 24 Doe de Bach Challenge: <http://bachbot.com/>

Illustratieverantwoording

ILLUSTRATIES BINNENWERK

- H1 'Difference Engine No 2, designed by Charles Babbage, 1847-1849, Engine built by Science Museum.' door: Science Museum Group. (1992).
- H1 Twee tandwielen uit de *Difference Engine No. 2*, DPS Design
- H1 De bedrading voor een AND-poort, DPS Design
- H1 OR-poort en een NOT-poort, DPS Design
- H1 $1 + 1 = 10$ in een processor, DPS Design
- H1 'Experimental Fourier addition' uit: Devitt, S. (2016).
- H1 Blaus, B. Medical gallery of Blausen Medical 2014. 'Multipolar Neuron.' *WikiJournal of Medicine* 1 (2). (2014).
- H1 De som van de oneven getallen, DPS Design
- H2 Racknitz, J. 'The Mechanical Turk.' Public domain (1789).
- H2 De verschillende uitkomsten waar je vanuit a op uit kunt komen, DPS Design
- H2 'Exploiting potential energy to locomote.' uit: Lehman, J. et al. (2018).
- H2 'Exploiting potential energy to pole-vault.' uit: Lehman, J. et al. (2018).
- H3 'The Google Inception-v3 classifier.' uit: Alcorn, M. et al. (2019).
- H3 Patroonherkenning bij neurale netwerken, DPS Design

- H3 Vier verschillende patronen waarvoor een neuraal netwerk, DPS Design
- H3 Een kunstmatig neuron met invoer en uitvoer, DPS Design
- H3 Een neuraal netwerk voor het herkennen van geschreven getallen, DPS Design
- H3 De leermethode achter neurale netwerken, DPS Design
- H3 Een stopbord met vier stickers erop, DPS Design
- H3 De regio's van ons brein die visuele informatie verwerken, DPS Design
- H3 'Texforms [...] generated using a texture-synthesis model.' uit: Long, B. et al. (2018).
- H3 'Reconstruction of Facial Images Using Linear Regression.' uit: Chang, L. et al. (2017).
- H3 Gezichtsherkenning met een convolutional neuraal netwerk, DPS Design
- H3 Kazanjian, H., & Lucas, G. 'Darth Vader kills the Emperor.' *Star Wars: Return of the Jedi*. United States: Lucasfilm Ltd. (1983).
- H4 'The AI agents Alice and Bob developing a new English language.' Facebook Artificial Intelligence Research. (2017).
- H4 Karras et al. and Nvidia (2019) *This Person Does Not Exist*, <https://thispersondoesnotexist.com/image>. (2020).
- H4 De twee versies van Word2Vec, DPS Design
- H4 Zafrany, S. 'NLP with gensim (word2vec).' NLP With Gensim (Word2vec). *Samyazaf.com*. Beschikbaar op: <https://samyazaf.com/ML/nlp/nlp.html>. (2017).
- H4 Een recurrent neuraal netwerk, DPS Design
- H4 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł. and Polosukhin, I. 'The encoder self-attention distribution for the word "it" from the 5th to the 6th layer of a Transformer trained on English

to French translation (one of eight attention heads).
Transformer: A Novel Neural Network Architecture For
Language Understanding. *Google AI Blog*. Beschikbaar
op: <https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html>. (2017).

- H4 'Images of household items across the world, and classes recognized in these images by five publicly available image-recognition systems.' uit: Devries, T. et al. (2019).
- H5 'Qualitative comparisons with baselines on four datasets with different characteristics.' uit: Zeng, Y. et al. (2019).
- H5 Flarup, G. 'Generative Adversarial Network for image generation.' *Generative Adversarial Networks. Department of Physics and Astronomy*. (2019).
- H5 thiscatdoesnotexist.com. (2020).
- H5 'Multi-domain image-to-image translation results on the CelebA dataset via transferring knowledge learned from the RaFD dataset.' uit: Choi, Y. et al. (2018).
- H5 'Face image editing results by our system.' uit: Jo, Y. et al. (2019).
- H5 'The results of talking head image synthesis using face landmark tracks.' uit: Zakharov, E. et al. (2019).
- H5 'Portrait of Edmond Belamy'. Created by GAN. Christie's. (2018).
- H5 'Generation of a single target aerial photo x^* from two arbitrary maps y_0 . Note that (c) is amplified for visibility' uit: Chu, C. et al. (2017).
- H5 'Details in x are reconstructed in $G_F x$, despite not appearing in the intermediate map $f x$ ' uit: Chu, C. et al. (2017).
- H6 Monroe, R. 'Predictive Models.' *xkcd*. Beschikbaar op: <https://xkcd.com/2169/>. (2017).
- H6 'Qualitative results on the AVSpeech test set.' uit: Oh, T. et al. (2019).

FOTOKATERN

- 1 'An impersonation using frames.' uit: Sharif, M. et al. (2016).
- 2-3 'Examples of successful impersonation and dodging attacks' uit: Sharif, M. et al. (2016).
- 4 'Detecting an elephant in a room.' uit: Rosenfeld, A. et al. (2018).
- 5 'Effects of transplanting an object from an image into another location in the same image.' uit: Rosenfeld, A. et al. (2018).
- 6-9 'Natural adversarial examples from IMAGENET-A and IMAGENET-O' uit: Hendrycks, D. et al. (2019).
- 10 'Sulfonyl/sulfonamide detection with autonomously trained convolutional filters.' uit: Wallach, I. et al. (2015).
- 11 Het verwateren van het geheugen van een recurrent neuraal netwerk. DPS Design.
- 12-13 Karras et al. and Nvidia (2019) *This Person Does Not Exist*, www.thispersondoesnotexist.com/. (2020).
- 14 'Generated images on CUB compared with StackGAN.' uit: Zhang, Z. et al. (2018).
- 15 *Artbreeder*, www.artbreeder.com/. (2019).
- 16 'Example of images generated by CAN.' uit: Elgammal, A. et al. (2017).
- 17 'Modes of variations of the activation subspace showing smooth temporal transition and correlation with Wolf-flin's concepts.' uit: Elgammal, A. et al. (2017).

De uitgever heeft getracht alle rechthebbenden te achterhalen. Indien iemand meent als rechthebbende in aanmerking te komen, kan hij of zij zich tot de uitgever wenden.