



Faculteit Bedrijf en Organisatie

Is het mogelijk om onherkenbare handgeschreven geschriften door middel van een convolutional neural network te herkennen, het correcte schriftsysteem vast te stellen en zo het proces te versnellen voor letterkundigen

Vercleyen Vital

Scriptie voorgedragen tot het bekomen van de graad van
professionele bachelor in de toegepaste informatica

Promotor:
Steven Van Impe
Co-promotor:
Glenn Vanlooveren

Instelling: —

Academiejaar: 2018-2019

Tweede examenperiode

Faculteit Bedrijf en Organisatie

Is het mogelijk om onherkenbare handgeschreven geschriften door middel van een convolutional neural network te herkennen, het correcte schriftsysteem vast te stellen en zo het proces te versnellen voor letterkundigen

Vercleyen Vital

Scriptie voorgedragen tot het bekomen van de graad van
professionele bachelor in de toegepaste informatica

Promotor:
Steven Van Impe
Co-promotor:
Glenn Vanlooveren

Instelling: —

Academiejaar: 2018-2019

Tweede examenperiode

Woord vooraf

Voor u ligt een onderzoek naar het herkennen van schriftsystemen door middel van deep learning. Dit is geschreven in het kader van mijn afstudeerrichting Toegepaste Informatica aan de Hoge School Gent. Van februari 2019 Tot en met mei 2019 ben ik bezig geweest met het onderzoek en het schrijven van mijn bachelorproef.

Samen met mijn promotor, Steven Van Impe, heb ik de onderzoeksvraag voor deze bachelorproef bedacht. Ik had veel baat aan de technische hulp die ik kreeg van Glenn Van Looveren, hij beantwoorde steeds mijn vragen waardoor ik steeds verder kon met mijn onderzoek.

Bij dezen wil ik graag mijn begeleiders bedanken voor de fijne begeleiding en hun ondersteuning tijdens dit traject. Ook wil ik alle respondenten bedanken die mee hebben gewerkt aan dit onderzoek. Zonder hun medewerking had ik dit onderzoek nooit kunnen voltooien.

Tevens wil ik ook mijn stageCollega Joris Opsommer bedanken die mij vaak heeft bijgestaan bij dit onderzoek. Bovendien ben ik ook dankbaar voor de wijze raad van mijn vrienden en familie.

Tot slot wil ik zeker mijn ouders bedanken die ervoor zorgden dat ik aan mijn studie kon beginnen en zal afmaken.

Ik wens u veel leesplezier toe.

Vercleyen Vital

Samenvatting

Inhoudsopgave

1	Inleiding	13
1.1	Probleemstelling	14
1.2	Onderzoeksvraag	14
1.3	Onderzoeksdoelstelling	14
1.4	Opzet van deze bachelorproef	14
2	Stand van zaken	17
2.1	Schriftsystemen en hun groepen	17
2.2	Artificiële intelligentie	19
2.3	Machine Learning en Deep Learning	20
2.4	Convolutional Neural Network	23
2.4.1	convolutional laag	23
2.4.2	pooling laag	25

2.5	Karakterherkenning	26
2.6	Schriftherkenning	27
3	Methodologie	29
4	Conclusie	31
	Bibliografie	33

Lijst van figuren

2.1	Lineair onderscheidbaar	22
2.2	Architectuur van een Perceptron	22
2.3	Het convolutional proces	24
2.4	Het max-pooling proces	25

Lijst van tabellen

1. Inleiding

Naast het gesproken woord en de verschillen in talen hierin zijn er ook verschillende soorten schriftsystemen die gebruikt worden in deze talen. Het onderscheiden van deze systemen is niet altijd vanzelfsprekend, vele van deze zijn vaak zeer verschillend maar een aantal zijn ook zeer gelijkaardig. Het herkennen hiervan is altijd al een onderwerp geweest in deep learning, echter is er niet veel onderzoek uitgevoerd naar het onderscheiden van specifieke karakters door middel van een convolutional neural network. Bovendien is het belangrijk om te weten welk soort schrift er wordt gebruikt bij een meertalige omgeving die ook meerdere schriften bevat vooraleer een karakter herkenner kan gebruikt worden.

Het automatisch herkennen van schriftsystemen is zeer gewild bij letterkundigen aangezien er bij de forensische taalkunde waarbij er gevraagd wordt om een bepaald document geschreven in een onbekend schrift te lokaliseren. Heel vaak gaat het daarbij vaak om heel onbekende of geheime schriftsystemen, of gewoon gecodeerde taal.

Ook is het classificeren van karakters uit verschillende schriftsystemen ook toepasselijk bij automatische transcriptie of het vinden van documenten op het internet die geschreven zijn in een specifiek schriftsysteem.

In dit onderzoek wordt er nagegaan of het mogelijk is om door middel van een convolutional neural network verschillende schriftsystemen te classificeren en zodanig het proces te versnellen voor letterkundigen die trachten een bepaald schriftsysteem te identificeren.

1.1 Probleemstelling

Het probleem dat wordt aangepakt en verder ook onderzocht wordt is vooral het helpen van deskundigen in de forensische letterkunde, deze krijgen af en toe een taak opgelegd om een bepaald schriftsysteem te herkennen en zodanig te lokaliseren. Als het mogelijk is om met behulp van een convolutional neural network bepaalde software te ontwikkelen kunnen deze taalkundigen sneller hun taak volbrengen en dan ook de gekoppelde zaak versnellen. Meer specifiek gaat het hier over Chris Bulcaen, de curriculum manager bij Taal Letterkunde aan de Ugent. Hij concludeerde dat dit inderdaad een meerwaarde zou geven in de forensische letterkunde en dat hierdoor het werk zeker versneld zou worden.

Ook zouden meerdere bedrijven die bezig zijn met karakterherkenning en vertaling hier baan aan hebben. Deze zijn meer bezig met het herkennen van karakters in één schriftsysteem. Wanneer er een systeem zou bestaan dat eerst en vooral het schriftsysteem zou herkennen kan dit voordelig zijn wanneer er achteraf dan de correcte karakterherkenning toepast voor het juiste schriftsysteem.

1.2 Onderzoeksvraag

Is het mogelijk om onherkenbare handgeschreven geschriften door middel van een convolutional neural network te herkennen, het correcte schriftsysteem vast te stellen en zo het proces te versnellen voor letterkundigen.

1.3 Onderzoeksdoelstelling

Het beoogde resultaat van deze bachelorproef is het ontwikkelen van een convolutional neural network waarbij er een aantal verschillende n schriftsystemen zo accuraat mogelijk van elkaar kunnen worden onderscheiden. Een model waarbij alle soorten schriftsystemen herkend kunnen worden is niet het beoogde resultaat aangezien dit veel meer tijd zou kosten bij het verzamelen van data, dit wordt duidelijk in het onderzoek.

1.4 Opzet van deze bachelorproef

De rest van deze bachelorproef is als volgt opgebouwd:

In Hoofdstuk 2 wordt een overzicht gegeven van de stand van zaken binnen het onderzoeksdomein, op basis van een literatuurstudie.

In Hoofdstuk 3 wordt de methodologie toegelicht en worden de gebruikte onderzoekstechnieken besproken om een antwoord te kunnen formuleren op de onderzoeksvragen.

In Hoofdstuk 4, tenslotte, wordt de conclusie gegeven en een antwoord geformuleerd op

de onderzoeksvragen. Daarbij wordt ook een aanzet gegeven voor toekomstig onderzoek binnen dit domein.

2. Stand van zaken

In dit hoofdstuk wordt de stand van zaken besproken. Dit bevat wat er al onderzocht is rondom het gene wat er in deze bachelorproef besproken wordt. De stand van zaken zorgt er voor dat een lezer die niet op de hoogte is van het besproken onderwerp kan volgen met na het lezen hiervan.

2.1 Schriftsystemen en hun groepen

Onderzoek naar verschillende soorten schriftsystemen is makkelijk terug te vinden aangezien het beschouwd wordt als een echte wetenschap sinds de 20e eeuw. Alhoewel het onderzoek vooral aanwezig was in de westerse wereld was er een soort van discriminatie aanwezig in dit vak.

Een van de eerste boeken over de verschillen in schriftsystemen was het boek genaamd 'A Study of Writing', geschreven door I. J. Gelb in 1952. (Gelb, 1952) De schrijver trachtte een basis neer te leggen voor de ontwikkeling van een nieuwe wetenschap dat de naam 'grammatologie' zou gekregen hebben. Deze nieuwe wetenschap probeerde standaard principes vast te leggen over het gebruik en de evolutie van het schrift op een vergelijkende wijze. Dit was anders dan de voorgaande studies in het vak van schriftsystemen aangezien het een aantal systemen vergeleek met andere en niet enkel keek naar één systeem zijn geschiedenis en evolutie. Dit boek bracht duidelijkheid in andere schriftsystemen dan dat de westerse bewoner gewoon was. In dit boek legde de schrijver uit hoe taalkundigen de gesproken taal zagen als iets dat meer biologisch is ontwikkeld en dit al doorheen de hele evolutie van de mens. Daarentegen vermeldde hij dat schriftsystemen zijn ontwikkeld als een technologie en dit enkel in de voorbije duizenden jaren en duidelijk meer een culturele

dan een biologische schenking is.

In de jaren 60 waren er nog grote misopvattingen aanwezig bij de taalkundigen. Bijvoorbeeld beschouwde Jack Goody, een toenmalig taalkundige, het Chinese schrift als een gelimiteerd systeem aangezien hij vond dat het incapabel was om vele ideeën uit te drukken en hinderde de adoptie van westerse talen in deze cultuur.

Gelukkig is in de laatste jaren alles veranderd, de studie naar schriftsystemen is een gerespecteerde wetenschap in de taalkunde en is daarbij ook een zeer beoefende wetenschap. Met dank aan de globalisatie is er een groter begrip gevormd voor andere talen dan die in de westerse wereld. Aangezien het duidelijk is dat schriftsystemen ontwikkeld zijn uit de cultuur en niet uit de biologische geschiedenis is het interessanter om dit te bestuderen aangezien het nieuwer en flexibeler is.

Een schriftsysteem, technisch beschreven als een schrift of orthografie kan geclassificeerd worden onder verschillende groepen. Er zijn een aantal soorten groepen en allen zijn makkelijk van elkaar te onderscheiden maar in dit onderzoek wordt er gefocust op de 3 meest voorkomende.

Eén daarvan is het logografische schrift of anders verwoord, het beeldschrift. Bij deze groep zijn volledige woorden uitgeschreven als één volledig teken dat compleet op zijn eigen staat en geen hulp nodig heeft van andere tekens om het te kunnen lezen. Zoals hierboven reeds vermeld werd dit vroeger als een gelimiteerd schrift beschouwd aangezien het niet mogelijk is om een andere talen uit te drukken in dit schrift. Dit is niet correct aangezien alle logografische schriften nooit puur logografisch zijn. Bij deze schriften is er naast woord-gebaseerde tekens ook klank-gebaseerde tekens voor het representeren van woorden die niet eigen zijn aan de taal of het schrift. Het logografisch schrift wordt bekeken als een van de oudste groepen, de talen die onder deze groepen vallen zijn vaak ook zeer oude schriften waardoor ze tijd hadden om zich volledig te ontwikkelen en zijn daardoor zeer interessant in het heden.

Een andere groep is het syllabisch of lettergrepenschrift. Bij deze schriften stellen de symbolen een klinker of een combinatie van medeklinkers en klinkers voor, simpeler gezegd stelt elk symbool simpelweg een lettergreep voor. Lettergrepenschriften zijn vaak schriften die vroeger vooral werden gebruikt, in deze tijd komen deze schriften zelden voor. een lettergrepenschrift komt deze tijd meer voor bij een logografisch schrift, dit zijn de klank-gebaseerde tekens zoals hierboven vermeld.

Onder deze groep valt er een andere groep die nog vaak voorkomt en niet mag weggelaten worden. Deze groep is gekend als een consonantenschrift, deze schriften bestaan enkel uit medeklinkers waarbij de klinkers compleet worden genegeerd, hierdoor komt er bij het uitspreken van deze schriften wat gokken bij te pas en natuurlijk ervaring om de juiste klank te vinden.

De laatste en grootste is een groep dat vooral voorkomt in de westerse wereld en momenteel ook de meest gebruikte, het alfabetisch schrift. De reden waarom dit de meest voorkomende groep is is omdat het een van de simpelste groepen is in vergelijking met de andere. Het alfabetisch schrift representeert de fonologische structuur van de gesproken taal die dit

schrift gebruikt, de geschreven taal representeert de uitgesproken klank.

Hier zijn er natuurlijk verschillen in de klank tussen de talen die dit schrift gebruiken. De Engelse uitspraak voor het woord 'hand' is bijvoorbeeld zeer verschillend tegenover de Nederlandse uitspraak van dit zelfde woord. Deze verschillen komen voort uit het uitspreken van dit schrift, vaak zijn ze te vinden in hoe de lippen, tong, gehemelte en keel wordt gebruikt. Bij alle uitgesproken talen worden deze verschillend gebruikt en vormen daardoor steeds een andere soort klank.

Het alfabetisch schrift is er één die de andere schriften overtreft aangezien het fonologisch is, hiermee kunnen een groot deel van de reeds bestaande klanken worden uitgedrukt. Een voorbeeld hiervan is het universele fonologische schrift dat bestaat uit een aantal symbolen die weergeven hoe een bepaald woord moet worden uitgesproken. De klanken worden dan letterlijk uitgeschreven op papier en dit enkel op basis van gehoor.

(Rickandie2016) Het schrift is in de eerste vorm uitgevonden door Semitische volkeren met als bekendste de Feniciërs. Het alfabetisch schrift werd door veelvuldige contacten met omringende volken overgenomen door de Grieken, Hebreeërs en Arabieren. Via de Grieken namen ook de Romeinen het over. Gaandeweg hebben er echter veel aanpassingen plaatsgevonden tot het uiteindelijke eenvoudige schrift. De naam van het schrift is afgeleid van het Griekse alpha en bèta dat de eerste twee symbolen zijn van het Griekse alfabet.

2.2 Artificiële intelligentie

Aangezien het programma dat zal worden geschreven bepaalde beslissing moet maken en kunnen nadenken over deze beslissing wijst dit automatisch op een technologische term, namelijk artificiële intelligentie. Het mogelijk maken voor programma's om te denken, doen en leren als een mens. Of een meer genuanceerde definitie.

Artificiële of kunstmatige intelligentie is een interdisciplinair concept dat de mogelijkheid bestudeert voor het ontwikkelen van machines capabel om interactie aan te gaan met hun omgeving en in te werken op de ontvangen data op een manier die als intelligent beschouwd kan worden.

De term artificiële intelligentie werd als eerst gebruikt door John McCarthy in 1956, een Stanford onderzoeker. Hij bedacht de term en legde AI vast als een branche van de computerwetenschap.

Artificiële intelligentie is in onze tijd al zeer groot en blijft maar groeien, het is tegenwoordig overal te vinden in het leven van de doodgewone mens. Een eenvoudig voorbeeld hiervan is Siri of Alexa, een persoonlijke assistent. Deze en anderen zijn al in staat om onze stem te herkennen, de toegankelijke informatie te analyseren en een zo compleet mogelijk antwoord terug te geven. Deze assistenten leren voortdurend meer over hun gebruiker waardoor ze beter en gemakkelijker kunnen voldoen aan de persoonlijke eisen van de gebruiker.

(Katja, John, Allan, BaoBao & Owain, 2018) Experts voorspellen dat artificiële intelligentie binnen het volgende decennia de mens zal overtreffen in het vervullen van simpele opdrachten zoals het vertalen van buitenlandse talen, het schrijven van verslagen of het besturen van een wagen. Maar het kunnen schrijven van een hooggewaardeerd boek of de taken overnemen van een chirurg zou nog iets wat langer kunnen duren. Voor deze twee is er verwacht dat deze vaardigheden mogelijk zullen zijn in respectievelijk 2049 en 2053.

2.3 Machine Learning en Deep Learning

Terwijl artificiële intelligentie, Machine Learning en Deep Learning vaak door elkaar worden gebruikt zijn er een aantal verschillen dat deze drie termen van elkaar onderscheid. Een manier om de relatie tussen deze te visualiseren is door middel van concentrische cirkels. artificiële intelligentie is hierbij de overkoepelende cirkel en dan ook de buitenste laag die het gehele domein van de studie voorstelt, Machine Learning is dan een cirkel binnen in artificiële intelligentie. Als laatste heb je Deep Learning, de kleinste cirkel, dat een verfijning is van Machine Learning en representeert de meeste artificiële intelligentie applicaties die vandaag worden gebruikt.

Machine Learning is simpel uitgelegd het proces van het ontwikkelen van machines die in staat zijn datasets op te halen, algoritmes uit te voeren op deze data en vervolgens zichzelf te trainen en uiteindelijk waardevolle inzichten te ontwikkelen gebaseerd op deze datasets. Het grootste verschil tussen Machine Learning en artificiële intelligentie is dat Machine Learning niet expliciet afhangt van de geschreven code van zijn ontwikkelaar. Machine Learning gebruikt eerder geschreven code als een startpunt en halen daarna data en informatie op waarop ze kunnen studeren, veel gelijkend op hoe een student zou studeren voor een examen.

(Arthur Samuel; 1959) "The ability to learn without being explicitly programmed."

De term Machine Learning werd voor het eerst gebruikt in 195 door Arthur Samuel, was in de jaren 80 daarentegen kreeg het vak meer aandacht en werd er meer onderzoek in gestoken waardoor er een grote groei plaatsvond die nog altijd even sterk is in het heden. (Keith D., 2019) Momenteel wordt Machine Learning vooral gebruikt bij het herkennen van gezichten, stemcommando's en het vertalen van talen. Vooral terug te vinden bij persoonlijke assistenten zoals Siri of Alexa, zoals hierboven vermeld.

Een voorbeeld van Machine Learning is het voorspellen van de prijs van een wagen. De enigste taak dat Machine Learning moet volbrengen is het geven van een best passend antwoord maar met invloed van vele parameters, een model dat de prijs zou kunnen voorspellen van een bepaalde wagen gaat als volgt. Eerst en vooral kijkt het model naar een hoop data van andere wagens waar de prijs al van bekend is. In deze data komen vooral parameters terug die invloed hebben op de prijs van de auto zoals het merk, het jaar van productie, het aantal al afgelegde kilometers, de motor, de wielen, enzovoort. Uit deze parameters ontwerpt het model een voorspellingsfunctie, wanneer een prijs moet worden

voorspelt van een bepaalde wagen worden alle nodige parameters ingevoerd in deze functie en hieruit wordt een best passende prijs aan gegeven.

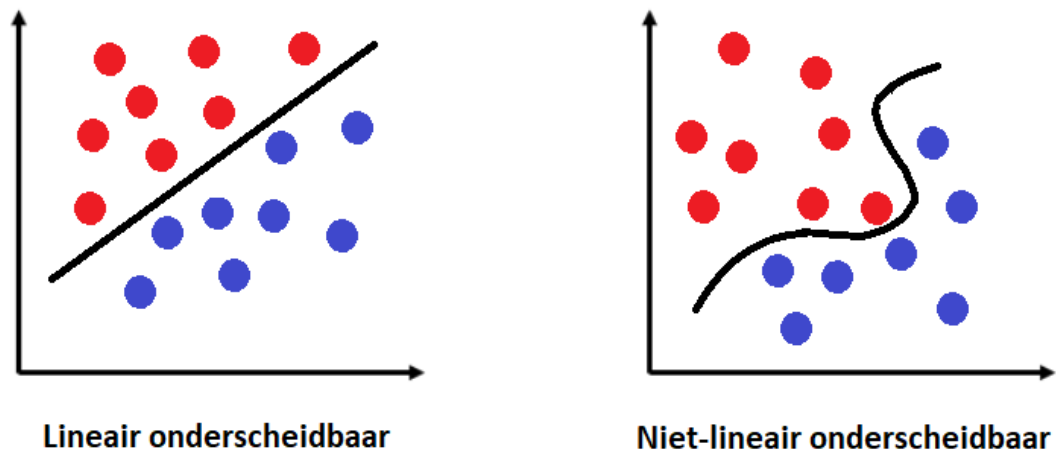
Dit soort van Machine Learning staat bekend onder gesuperviseerd leren waarbij er data is gegeven, je heb daarnaast ook niet-gesuperviseerd leren waarbij er geen data is gegeven en het model correlaties legt tussen data. Bij het voorbeeld van de wagens zou het dus kunnen vaststellen bij 3 verschillende wagens dat wagen 1 meer gelijkenissen heeft met wagen 2 dan met wagen 3. Als laatste onder Machine Learning heb je ook Reinforcement Learning waarbij een model beloningen krijgt voor bepaalde acties die het onderneemt. Bij een slechte actie krijgt het een negatieve beloning en bij een goede actie krijgt het een positieve beloning, hieruit kan het model leren wat het wel en niet zou moeten doen.

Deep Learning is geïnspireerd door de structuur van het menselijke brein, het menselijk brein bestaat uit een complex netwerk van miljarden neuronen die met elkaar communiceren door middel van synapsen. Deze neuronen sturen voortdurend elektrische impulsen waardoor het nodige informatie kan doorsturen naar andere delen van het lichaam, zoals de spieren of het hart. Dit is gelijkaardig bij Deep Learning aangezien het gebruik maakt van artificiële neurale netwerken. In zo een netwerk is elk neuron capabel om een antwoord te geven op simpele ja/nee-vragen. Door vele neuronen in zo een dergelijk neurale netwerk te implementeren kan zo een netwerk deftige antwoorden terug geven zonder het aanpassen van de geschreven code.

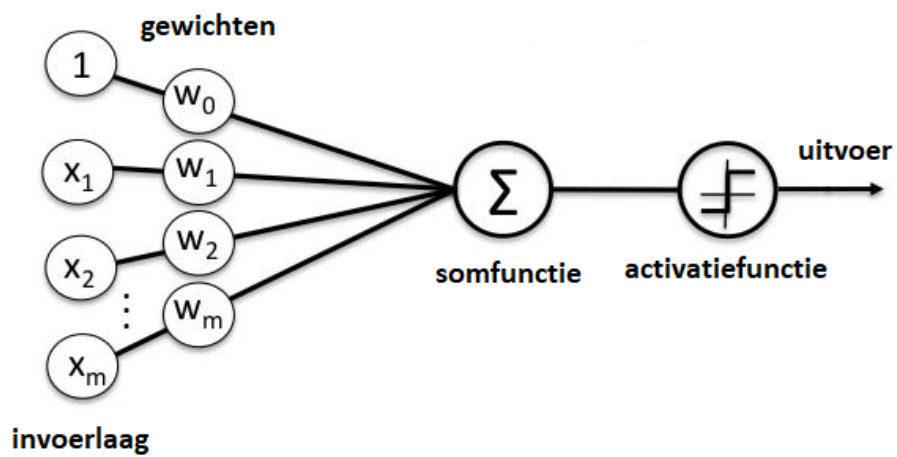
Het simpelste en oudste voorbeeld van een neurale netwerk is een perceptron, een netwerk dat lineair classificeert dat wordt gebruikt voor binaire voorspellingen. Dit betekent dat de data lineair onderscheidbaar moet zijn. (Figuur 2.1) Een perceptron, eigenlijk één neuron, bestaat uit één enkele laag, een invoer laag, bijkomende gewichten, een somfunctie en een activatiefunctie. De invoerlaag krijgt de invoer en vermenigvuldigt deze met de bijkomende gewichten, deze worden dan ingevoerd in de somfunctie die een uitvoer doorgeeft naar de activatiefunctie die vervolgens bepaald of de finale uitvoer 0 of 1 bevat. (Figuur 2.2) Trainen van een perceptron gebeurt door middel van het invoeren van veel data met de gekende output, naarmate wijzigt het perceptron de gewichten zodanig dat een zo best passende output kan worden gegenereerd.

Een meer geavanceerd neurale netwerk is een netwerk met meerdere lagen, waarbij elke laag een groot aantal aan neuronen bevat, de eerste laag is altijd de invoerlaag en de laatste laag is altijd de uitvoer laag, de gekende data wordt ingevoerd in de invoerlaag en vervolgens wordt er een resultaat berekend dat terug te vinden is in de uitvoerlaag. De daar in tussen liggende lagen worden 'hidden layers' genoemd, hierin krijgen de neuronen bepaalde invoer en produceren ze een gepaste uitvoer door middel van een activatie functie. Elk neuron in een bepaalde laag is vervolgens verbonden met elk neuron in de daaropvolgende laag waardoor de neuronen genoeg informatie krijgen om te beslissen welke output ze moeten doorgeven naar de volgende laag.

Neurale netwerken kunnen gebruikt worden bij fotoherkenning. Neem nu een simpel voorbeeld van een artificieel neurale netwerk dat moet beslissen of een foto een appel of een banaan bevat. Het netwerk heeft drie verschillende vragen.



Figuur 2.1: Lineair onderscheidbaar



Figuur 2.2: Architectuur van een Perceptron

- Is het voorwerp in de foto rond?
- Is het voorwerp in foto geel?
- Heeft het voorwerp in de foto een steel?

Bij een foto van een banaan zouden de neuronen antwoorden met respectievelijk neen, ja en neen. Voor een appel zou het antwoorden met respectievelijk ja, neen en ja. Met het gebruik van binaire getallen zou het netwerk aanleren dat de output voor een banaan 010 bevat en voor een appel 101. Wanneer dit voorbeeld wordt uitgeschreven over een groot aantal neuronen lagen is het mogelijk om veel complexere problemen aan te pakken.

Deep Learning is een redelijk nieuwe term aangezien het pas voor het eerst werd gebruikt rond het jaar 2000. Sinds is het een veelgebruikte term en is het een regelmatig besproken onderwerp binnen het vak van artificiële intelligentie. (Goff, 2018)

2.4 Convolutional Neural Network

Een Convolutional Neural Network, ook bekend onder CNN of ConvNet, is een neuraal netwerk dat zich specialiseert in het verwerken van data dat roostervormig is, zoals een afbeelding.

Een digitale afbeelding is een binaire representatie van visuele data. Het bevat een reeks van pixels geordend in een rooster waarbij de pixelwaarden aantonen hoe licht en welke kleur de pixel moet weergeven.

Een CNN bestaat, net zoals bij een normaal neuraal netwerk, uit een invoerlaag, een uitvoerlaag en een aantal tussenliggende lagen. Maar wat een CNN een CNN maakt zijn de tussenliggende lagen waarvan sommige convolutional of pooling zijn, de twee belangrijkste lagen in een CNN.

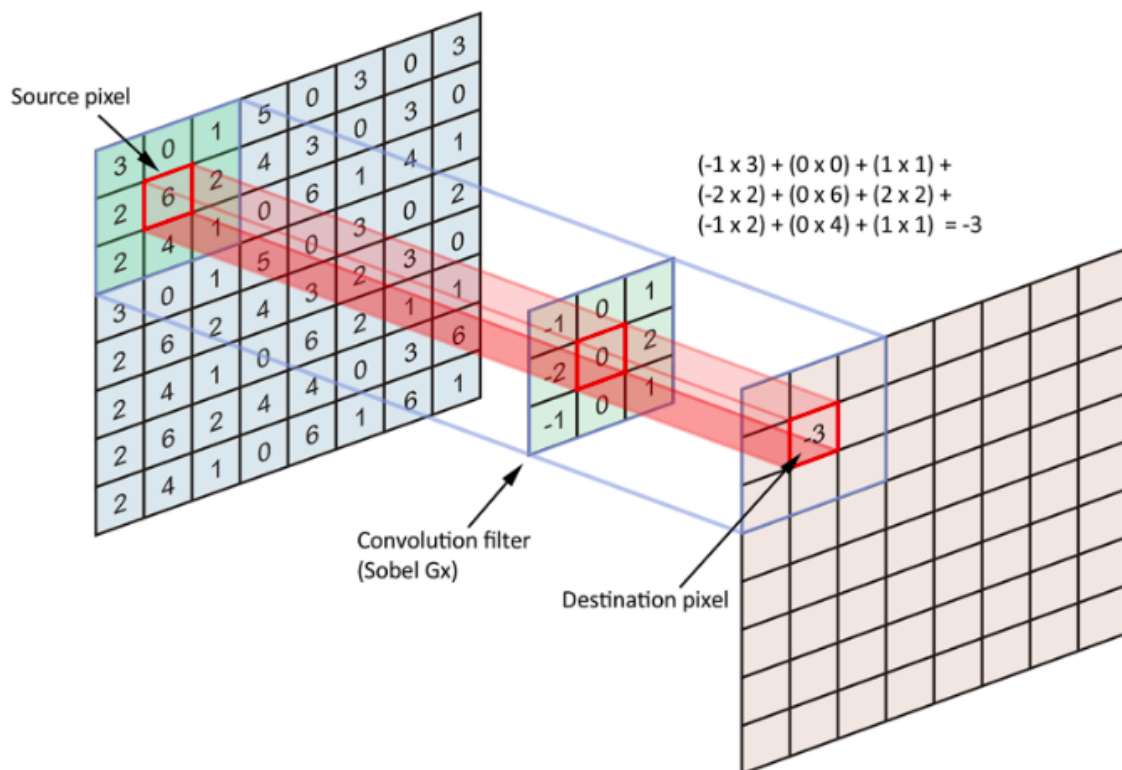
2.4.1 convolutional laag

Het verschil tussen een convolutional laag en een normale laag is dat convolutional lagen lokale patronen in kleine roosters van twee dimensies, details dus, aanleert en een normale laag globale patronen aanleert over het geheel van de afbeelding.

Het grootste doel van een convolutional laag is om randen, lijnen, vormen, enz. te herkennen. Vanaf dat het zo een karakteristiek op een specifieke plaats heeft aangeleerd kan het het later in elke andere plaats herkennen. Terwijl een normale laag de karakteristiek opnieuw zou moeten aanleren.

Wanneer een tweede convolutional laag volgt op een andere convolutional laag kan de tweede laag aangeleerde patronen gebruiken van de voorgaande laag. De tweede laag kan daardoor veel geavanceerde patronen aanleren, zoals een oog of een voetbal.

De convolutional lagen werken met 3 dimensionale roosters, twee assen hiervan tonen de

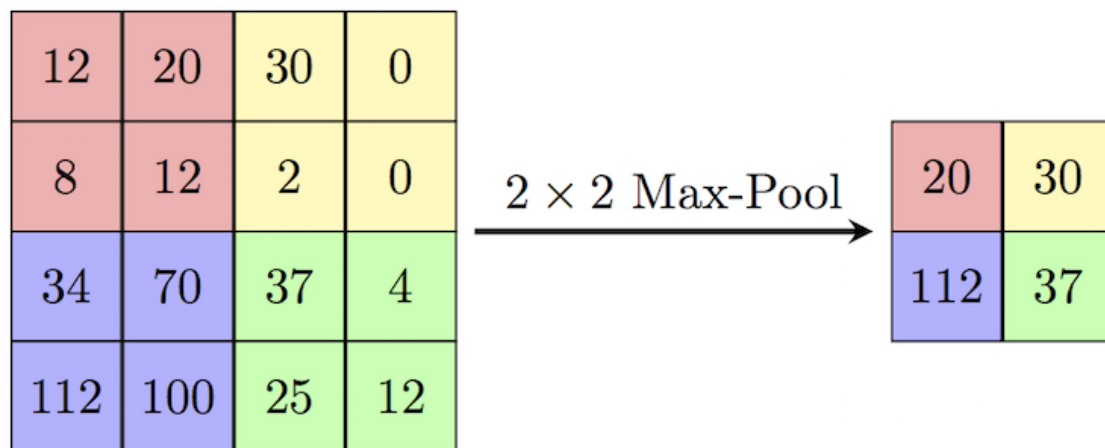


Figuur 2.3: Het convolutional proces

hoogte en breedte aan en de derde de diepte. Wanneer een afbeelding kleuren bevat zal deze diepte 3 bevatten, bij een zwart-wit foto zal deze 1 bevatten.

Een convolutional laag is convolutional aangezien het een convolutional operatie uitvoert. Stel nu dat na een invoerlaag er een convolutional laag is geplaatst dan krijgt de convolutional laag de niet verwerkte invoer binnen. De laag maakt één of meerdere vensters aan, dat een filter wordt genoemd, met meegegeven hoogte en breedte, en per vak krijgt het een willekeurige waarde. Er kan veronderstelt worden dat deze filter verschuift over de gehele afbeelding. Voor elke positie dat de filter kan aannemen op de afbeelding is er een verbonden neuron. Wanneer het venster bepaalde waarden in de afbeelding opneemt verwerkt hij dit met de willekeurige waarden en de uitvoer wordt in een nieuw rooster geplaatst genaamd een 'feature map', de afbeelding wordt zodanig kleiner. (Figuur 2.3)

Dit proces wordt duidelijker met een voorbeeld. Gegeven is een afbeelding van grootte 32x32 (hoogte x breedte), De afbeelding wordt ingevoerd in de convolutional laag en deze maakt een filter aan met grootte 3x3. De aangemaakte filter krijgt zijn willekeurige waarden en begint links van boven bij de afbeelding, hier verwerkt de filter de waarden en plaatst ze in een nieuw rooster. Vervolgens schuift de filter met één stap naar rechts, wanneer deze filter niet meer naar rechts kan verschuiven, verschuift het met één stap naar beneden en begint opnieuw aan de linkerkant van de afbeelding. Wanneer de filter op alle unieke posities is geweest is het nieuwe rooster gevormd. Bij dit voorbeeld heeft het nieuwe rooster een grootte van 30x30 aangezien dit het aantal unieke posities is dat de filter kan aannemen. De convolutional laag gebruikt voor elk neuron dezelfde filter, één



Figuur 2.4: Het max-pooling proces

filter kan één enkele karakteristiek herkennen, daarom wordt er aangeraden om meerdere filters te gebruiken in één convolutional laag. Wanneer meerdere filters worden gebruikt kunnen er meerdere karakteristieken worden herkend en zullen er betere resultaten worden opgeleverd.

2.4.2 pooling laag

Naast de convolutional lagen zijn ook de pooling lagen belangrijk in het proces van een CNN. De pooling lagen worden meestal net na een convolutional laag geplaatst, zo een laag versimpelt de informatie die verkregen werd door de convolutional laag door een compactere versie te maken van deze informatie. Er zijn twee manieren waarop een pooling laag dit kan doen, max-pooling en average-pooling, de meeste gebruikte hiervan is de eerste.

De pooling laag maakt zoals bij een convolutional laag een filter van meegegeven grootte, vervolgens begint de filter in de hoek links van boven en neemt, bij max-pooling, de grootste waarde in de filter en plaatst deze vervolgens in een nieuw rooster. Wanneer dit is voltooid schuift de filter door naar een volgende positie maar overlapt niet met informatie waar het al is geweest, in tegenstelling tot de filter bij een convolutional laag. Bij average-pooling neemt de filter het gemiddelde van de waardes die te zien zijn in de filter. Stel nu dat de meegegeven hoogte en breedte 2 bevat, dan zal het verwerkte rooster half zo groot zijn als het originele rooster. (Figuur 2.4) Zoals hierboven vermeld heeft een convolutional laag vaak meerdere filters, wanneer we een pooling laag introduceren in het neurale netwerk zal het evenveel pooling filters bevatten als convolutional filters.

2.5 Karakterherkenning

Een onderwerp waarvoor Convolutional Neural Networks veelal voor gebruikt wordt is karakterherkenning of OCR (Optical Character Recognition), een technologie waarbij uit een afbeelding of tekstdocument alle tekens uit het bestand kunnen worden herkend en apart kan worden opgeslagen. Een voorbeeld hiervan is nummerplatherkenning bij flitspalen. waarbij een foto wordt getrokken, en vervolgens wordt opgeslagen onder de herkende nummerplaat. Karakterherkenning kan ook te pas komen bij het herkennen van handgeschreven karakters, een probleem dat hierbij voorkomt is dat geschriften vaak kunnen variëren in stijl maar dezelfde patronen komen vaak terug. Dit probleem kan dus opgelost worden met een Convolutional Neural Network aangezien dit deze vaak terugkomende patronen kan onthouden en gemakkelijk kan herkennen op andere plaatsen dan de originele gevonden plaats. Wanneer onderzocht wordt naar karakterherkenning met behulp van een CNN zullen de onderzoekers vaak trachten bepaalde karakters of woorden te herkennen uit één specifiek schriftsysteem. Even zoeken op het internet en vele artikels kunnen teruggevonden worden die dit specifiek onderzoek toepassen, zoals het herkennen van woorden in het Latijnse schrift (Aiquan, Gang, Po, Yanni & Xinting, 2012), het herkennen van Chinese karakters (Weixin, Lianwen, Zeeheng & Ziyong, g.d.), het herkennen van het Bengaals alfabet, een schriftsysteem uit Zuid-Azië (Mahbubar, M. A. H., Shahidul & Pintu, 2015), enzovoort. Deze onderzoek trachten verschillende karakters uit een gekozen schriftsysteem te herkennen en zodanig woorden te vormen en uiteindelijk deze te lezen. Dit proces is simpelweg het lezen van tekst door middel van een CNN, wat vooral gebruikt kan worden bij het lezen van onbekende schriftsystemen maar veronderstelt dat het gebruikte schriftsysteem al gekend is.

Deze veel voorkomende onderzoeken zijn uitermate zeer interessant aangezien het een beeld heeft van wat er mogelijk is en wat niet. De onderzoeken volgen meestal dezelfde werkwijze, eerst en vooral wordt er voldoende data verzameld van het specifieke schriftsysteem. Dit meestal uit al bestaande datasets. Wanneer men spreekt over datasets zijn er twee groepen te onderscheiden, online datasets en offline datasets. Bij een online handgeschreven karakter dataset zijn de opgeslagen karakters met wanneer de 'gebruikte pen' werd opgeheven en wanneer het werd neergezet, dit kan bijvoorbeeld komen door het gebruik van een smartphone waarbij de 'gebruikte pen' de vinger op het scherm is. Een voorbeeld hiervan is de CoMNIST dataset, een dataset bestaande uit het Latijnse schriftsysteem en het Cyrillische. Bij deze dataset kan iedereen participeren met de opslag van de handgeschreven karakters. De ontwerpers van de dataset maakten een website aan waarbij je een opgegeven letter krijgt, deze letter moet je vervolgens tekenen in een venster en deze wordt opgeslagen in de dataset. Een offline dataset daarentegen bestaat uit data die gelezen is uit documenten of afbeeldingen, deze zijn vaker niet handgeschreven maar niet altijd.

Om verder te gaan op de veelgebruikte werkwijze in de al bestaande onderzoeken is de volgende stap het verwerken van de datasets zodanig dat het gemakkelijk gebruikt kan worden voor de volgende stap. Hoe een dataset verwerkt wordt hangt af van welke technologieën er worden gebruikt en persoonlijke preferentie. In een bepaald onderzoek worden de afbeelding geroteerd, verschoven en uitgerekt om zodanig lokale vervorming te

creëren, hierdoor komt er meer variëteit in de data (Weixin e.a., g.d.). Wat vooral voorkomt is het veranderen van de grootte van de data specifiek voor de verwachte invoer van het gebruikte model (Aiquan e.a., 2012) (Mahbubar e.a., 2015). Aan het einde van de nodige aanpassingen van de data volgt er vaak een laatste stap, namelijk de normalisatie van de te gebruiken data. Normalisatie is het aanpassen van alle data zodanig dat de waarden van alle bijkomende variabelen op een zelfde schaal terecht komen.

De volgende stap is het aanmaken van het model, er wordt gezocht naar welke lagen de beste resultaten op zullen leveren en zorgen dat elke laag de correct invoer krijgt. Wanneer het model is aangemaakt wordt het getraind met de verwerkte datasets, wanneer dit is gebeurd wordt het model getest en de resultaten worden neergeschreven.

Bij elk onderzoek is er zoals verwacht een conclusie uitgeschreven, in de meeste conclusies wordt er uitgelegd dat het gelukt is om een model te schrijven voor hun karakterherkenning. Maar wat er in de conclusies vaak voorkomt is dat hun model nog kan worden verbeterd. Eén onderzoek blinkt hier wel in uit, dit onderzoek concludeert dat ze een zeer hoge accuraatheid hebben bereikt met hun model. Het verschil met de andere onderzoeken is dat zij niet enkel een CNN gebruiken maar andere technologieën gebruiken, dit zijn vooral technologieën die gebruikt worden voor hun datasets (Yann & Yoshua, g.d.).

2.6 Schriftherkenning

3. Methodologie

4. Conclusie

Bibliografie

- Aiquan, Y., Gang, B., Po, Y., Yanni, G. & Xinting, Z. (2012). Handwritten English Word Recognition based on Convolutional Neural Networks. *Nankai University*.
- Gelb, I. J. (1952). *A Study of Writing*.
- Goff, A. (2018). A History of Deep Learning.
- Katja, G., John, S., Allan, D., BaoBao, Z. & Owain, E. (2018). When will AI Exceed Human Performance? Evidence from AI Experts. *AI Impacts*.
- Keith D., F. (2019). A Brief History of Machine Learning. *Dataversity*.
- Mahbubar, R., M. A. H., A., Shahidul, I. & Pintu, C. (2015). Bangla Handwritten Character Recognition using Convolutional Neural Network. *MECS*.
- Weixin, Y., Lianwen, J., Zeeheng, X. & Ziyong, F. (g.d.). Improved Deep Convolutional Neural Network For Online Handwritten Chinese Character Recognition using Domain-Specific Knowledge. *South China University of Technology*.
- Yann, L. C. & Yoshua, B. (g.d.). Word-Level Training of Handwritten Word Recognizer Based on Convolutional Neural Networks.