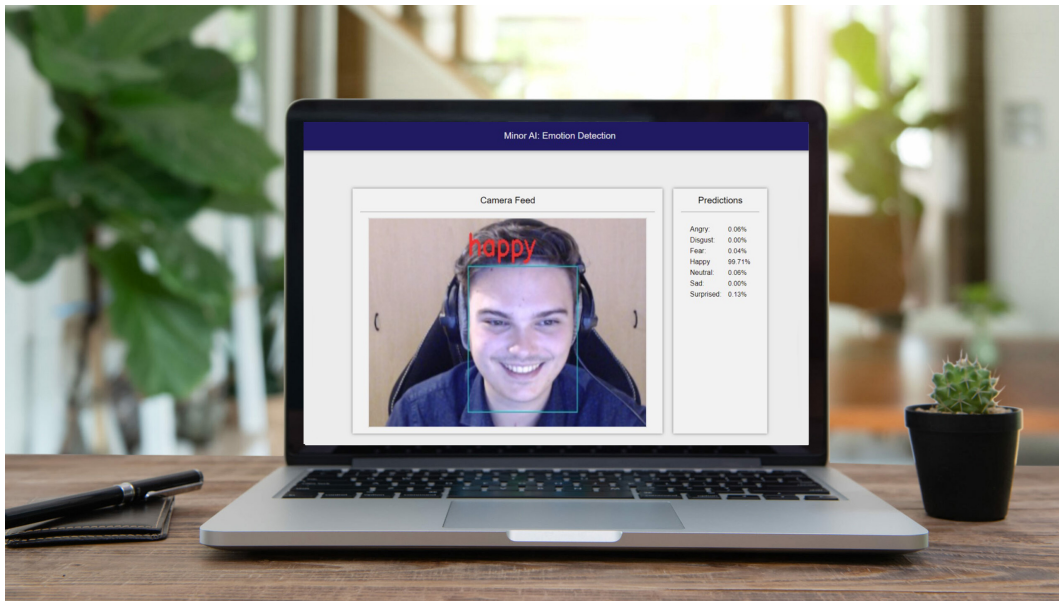


# Emotion Recognition

## Extended Abstract



BD02 Artificial Intelligence, HBO-ICT

by

Dwayne Debets (1329081), Roald Janssen (1519166), Bryan Kroon (1761978), Gavin Melis (1852787),

04-11-2021

Lectoraat Data Intelligence Zuyd Hogeschool

supervisor(s)

Koen Steeghs, Procesbegeleider  
Jim Bemelen, Opdrachtgever

# Abstract

Artificial Intelligence bestaat uit allerlei geavanceerde technologieën waardoor machines als het ware kunnen denken, leren en taken uit kunnen voeren zoals mensen. Er zijn talloze redenen waarom je AI voor MKB zou kunnen inzetten. Zo kan AI voorspellen, prospects opsporen, zorgen voor je klanten en inzicht geven in kansen. AI maakt MKB efficiënter en waardevoller dan ooit tevoren. Om MKB'ers helpen te laten realiseren wat de mogelijkheden van AI zijn heeft het lectoraat Data Intelligence van Zuyd Hogeschool de opdracht gegeven om een demonstrator te ontwikkelen die de werking van een Convolutional Neural Network in kaart brengt doormiddel van emotieherkenning. Om het project in goede banen te leiden werd er gebruik gemaakt van de Design Science Research methode van Hevner. Met het neergezette emotieherkenningsdemonstrator zijn verschillende mogelijkheden geëxploreerd met betrekking tot inzet binnen het bedrijfsleven. Met de resultaten van het project is het mogelijk gemaakt om het Artificial Intelligence algoritme Convolutional Neural Networks te demystificeren voor het MKB. Zo schijnen de projectresultaten niet alleen meer licht op Artificial Intelligence voor het MKB, maar kan ook het MKB aangespoord worden deze technologieën actief en succesvol toe te passen binnen het bedrijfsleven.

# 1

## Introductie

### Aanleiding

Artificial Intelligence (AI), oftewel kunstmatige intelligentie, bestaat uit allerlei geavanceerde technologieën waardoor machines als het ware kunnen denken, leren en taken uit kunnen voeren zoals mensen. AI wordt sterk bepalend voor onze toekomstige welvaart en welzijn. Door het generieke karakter zal AI impact hebben op alle bedrijfssectoren, ons privéleven en de maatschappij. Alleen door een proactieve inzet op en van AI kan Nederland mede bepalen hoe AI wordt toegepast voor sociale en economische uitdagingen en daar voordeel uit halen. Dit alles vereist een collaboratieve ketenaanpak, waarbij overheid, onderwijs- en onderzoeksinstellingen, bedrijfsleven en burgers samenwerken [2]. Er zijn talloze redenen waarom je AI voor MKB zou kunnen inzetten. Zo kan AI voorspellen, prospects opsporen, zorgen voor je klanten en inzicht geven in kansen. AI maakt MKB efficiënter en waardevoller dan ooit tevoren [3]

Een van de toepassingen van Artificial Intelligence binnen het MKB is beeldherkenning. Mensen kunnen intuïtief emoties van andere mensen herkennen. Binnen Artificial Intelligence is het mogelijk om machines de wereld te kunnen laten zien zoals mensen dat kunnen. Door een algoritme genaamd Convolutional Neural Networks (CNN) is het mogelijk om doormiddel van beelden te herkennen welke emotie een persoon uit. De architectuur van het CNN is gebaseerd op het menselijk brein, met name de visuele cortex. Individuele neuronen reageren op stimulansen die uiteindelijk leiden tot een trigger in het brein waarbij een beeld herkend kan worden.

Om MKB'ers helpen te laten realiseren wat de mogelijkheden van AI zijn heeft het lectoraat Data Intelligence van Zuyd Hogeschool de opdracht gegeven om een demonstrator te ontwikkelen die de werking van een Convolutional Neural Network in kaart brengt. Deze demonstrator geeft het lectoraat de mogelijkheid om het MKB bewust te stellen van de mogelijkheden die Artificial Intelligence beschikbaar stelt. Om dit prototype te kunnen bouwen wordt de volgende onderzoeksvraag gehanteerd: "Hoe kunnen we met behulp van Python een CNN-algoritme opstellen die emoties kan herkennen aan de hand van een live webcam feed?"

### Doelstelling

Het doel van het lectoraat Data Intelligence is om het MKB aan te kunnen sporen vaker en verstandiger om te gaan met AI-technieken, zodat niet alleen meer licht wordt geschinen op Artificial Intelligence, maar ook om deze actief en succesvol toe te passen binnen het bedrijfsleven. Om het doel van het lectoraat te behalen wordt gewerkt aan een demonstrator die in staat is om aan de hand van een live beeld de emotie kan herkennen. Hierbij dient het Convolutional Neural Network, de achterliggende algoritme die dit mogelijk maakt, zelf worden opgebouwd en getraind. Om meer kennis te vergaren met betrekking tot het neurale netwerk wordt een onderzoek opgesteld die zoveel mogelijk benodigde informatie kan geven. Aan de hand van de verzamelde informatie tijdens het onderzoek wordt een ontwerp voor de uiteindelijke oplossing opgesteld. Met de resultaten uit het onderzoek en het ontwerp wordt de oplossing gerealiseerd. Deze oplossing moet met Python gerealiseerd worden en losstaand kunnen functioneren als een demonstrator. Ook zal een overdrachtsportfolio moeten worden geleverd die de werking van de oplossing beschrijft.

# 2

## Theoretisch kader

De unieke opdracht heeft als overkoepelend doel om het MKB bewust te maken van de mogelijkheden die Artificial Intelligence brengt samen met het aansporen om deze technologie te gebruiken. Zo zal binnen dit project niet alleen de nadruk liggen op het bouwen van een demonstrator, maar wordt ook verder onderzocht hoe het lectoraat Data Intelligence van Zuyd Hogeschool geholpen kan worden om het MKB zich bewust(er) te stellen van Artificial Intelligence.

Om hieraan te voldoen wordt gewerkt aan een type beeldherkenning dat onder Artificial Intelligence valt. Onder de motorkap wordt beeldherkenning aangedreven door een Convolutional Neural Network (CNN). Het CNN is in essentie een variant van Neurale Netwerken. Deze netwerken reflecteren het gedrag van het menselijk brein, wat ook de inspiratie is van deze technologie, om zodanig patronen te kunnen herkennen waardoor vraagstukken binnen AI beantwoordt kunnen worden. Neurale Netwerken bevatten, net zoals het menselijk brein, neuronen die corresponderen met andere neuronen waar uiteindelijk een soort uitkomst resulteert [5].

Het MKB kan deze beeldherkenning in de vorm van emotieherkenning inzetten voor verschillende mogelijke oplossingen. In het vakgebied van Psychologie kan de tool worden ingezet om te kijken welke emoties worden getoond van patiënten. Er kan een camera worden geïnstalleerd bij een kliniek van een psycholoog die dan de data per sessie naar de psycholoog stuurt. In de entertainmentindustrie kan de tool worden ingezet door te kijken hoe mensen reageren op bepaalde aspecten van media stukken. Zo kan er gekeken of mensen reageren op scenes zoals verwacht. Denk aan angst bij horrorfilms, verdrietig bij verdrietige films/scenes, vreugde bij comedy films. Dit is ook toepasselijk bij videogames. Ook kan het worden gebruikt om te kijken hoe kleine kinderen reageren bij het spelen met bepaalde stukken speelgoed. In de horeca kan de tool ingezet worden om te herkennen hoe klanten hun tijd ervaren in de zaak. Hier kan gekeken worden naar welke emoties de klanten hebben op het eten en de service. In clubs en bars kan de tool worden ingezet om de sfeer te controleren in de club of bar. Er ligt dan vooral een focus op negatieve emoties zoals boosheid en angst. Door deze emoties op te sporen kunnen bewakers en uitsmijters conflicten sneller voorkomen.

# 3

## Methode

### 3.1. Hevner based Design Science Research Process

Om het project in goede banen te leiden werd er gebruik gemaakt van de Design Science Research methode van Hevner [6]. Het doel van deze methode is om zowel de omgeving (Environment, via de Relevance Cycle) en de kennisbasis (Knowledge Base, via de Rigor Cycle) te betrekken bij de ontwikkeling van een product of dienst (Design Cycle) waarin onzekerheden weggenomen dienen te worden.

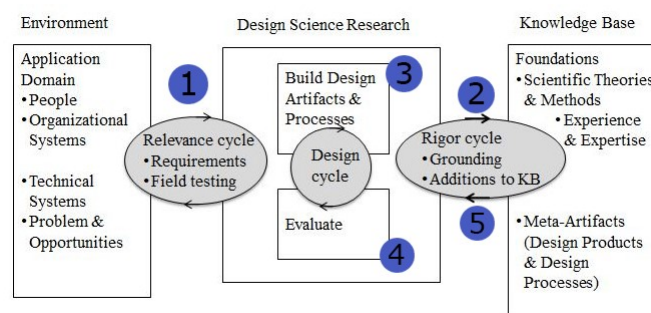


Figure 3.1: DSR: Hevner

Om deze methode te gebruiken zal het project moeten voldoen aan twee requirements:

- Aanwezigheid van een onderzoek waardige onzekerheid voor aanvang van het project.
- Het project heeft een doelstelling in oog, waarbij een artefact (product of dienst) gecreëerd dient te worden.

Binnen de onderzoeksopzet is de onderzoek waardige onzekerheid vastgesteld, met de vraag “Hoe kunnen we met behulp van Python een CNN-algoritme opstellen die emoties kan herkennen aan de hand van een live webcam feed?” Het artefact wat gemaakt dient te worden is een product in de vorm van een demonstrator die live emoties kan herkennen. Binnen de Knowledge contribution framework wordt dit project uitgevoerd binnen het kader ‘Exaptation’, wat inhoudt dat bekende oplossingen gebruikt kunnen worden om nieuwe problemen op te lossen. De volgende stappen komen aan bod bij de best-practise (DSR) die binnen het project worden gebruikt:



Figure 3.2: DSR: Hevner

**1. Requirements**

In deze stap wordt het doel en de scope van het product (artefact) in samenspraak met stakeholders gespecificeerd. Binnen deze stap wordt de minimale acceptatiecriteria geschetst van de opdrachtgever. Elicitatie technieken in de vorm van interviews zullen worden ingezet om achter de eisen en wensen te komen van de opdrachtgever.

**2. Grounding**

In deze stap wordt, doormiddel van onderzoek, de onzekerheden weggenomen die zijn vastgesteld in het theoretisch kader. Dit wordt gedaan aan de hand van deskresearch.

**3. Design Build & Evaluate**

Binnen deze stap wordt de demonstrator gebouwd. De focus ligt eerst op het bouwen van een Minimal Viable Product (MVP), waarbij het product voldoet aan de minimale acceptatiecriteria. Aan de hand van gesprekken met de opdrachtgever wordt het product verder bevorderd aan de hand van extra eisen en wensen in de vorm van Should Have's. Zo wordt het product ook verder geoptimaliseerd zodat het aan alle generieke eisen voldoet.

**4. Dissemination**

In deze laatste stap wordt nieuwe kennis aan de Knowledge Base toegevoegd. De opgedane resultaten tijdens de uitvoering van het project zal daarmee gedocumenteerd worden, zodat deze ook bruikbaar is voor andere onderzoekers en belanghebbende.

# 4

## Resultaten

### 4.1. Onderzoekresultaten

Uit het onderzoek is gebleken dat een Convolutional Neural Network (CNN) een Deep Learning algoritme is dat een input afbeelding kan pakken, belangrijke aspecten van die afbeelding extraheren om vervolgens verschillen te kunnen herkennen tussen verschillende afbeeldingen. Een CNN is een classificatie algoritme, waarbij er op basis van data verschillende typen klassen voorspeld kunnen worden. Het pre-processen van een CNN is lager ten opzichte van andere classificatie algoritmen, met genoeg training kan een CNN zelf filters of karakteristieken leren [7]. Het grootste verschil tussen een Convolutional Neural Network en een standaard feed-forward Artificial Neural Network (ANN) is dat een CNN het mogelijk stelt om succesvol relevante filters te kunnen genereren voor het classificeren van beeld. De CNN-architectuur geeft een betere fitting voor datasets met beeld, mede door een minder groot aantal parameters die erbij betrokken worden samen met de mogelijkheid voor hergebruik van gewichten. Kortgezegd kan een CNN beter de verschillende features herkennen van een bepaald beeld, waar puur naar de essentie wordt gekeken voor het classificeren van data. Het CNN-architectuur bestaat uit verschillende lagen: Input, Convolution layer, Pooling layer en de Fully connected layer. De onderstaande afbeelding geeft een voorbeeld van een high-level CNN-architectuur die handgeschreven nummers kan herkennen.

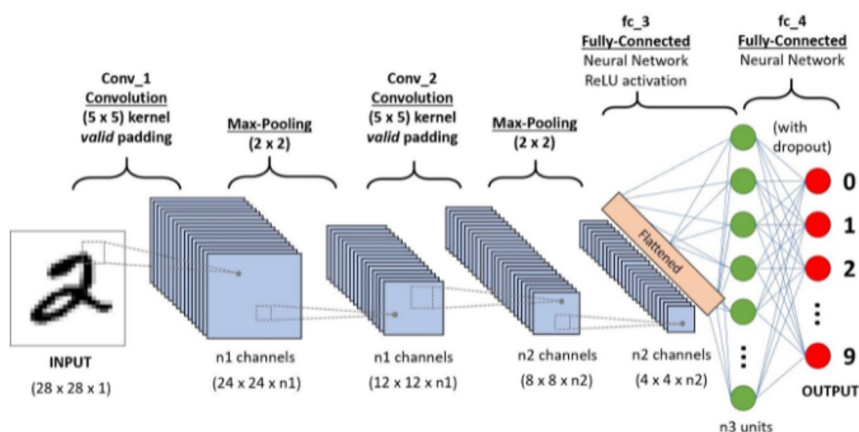


Figure 4.1: Een CNN sequentie om handgeschreven nummers te herkennen

Ook is gebleken uit het onderzoek dat het Convolutional Neural Network het beste gebouwd kon worden middels een combinatie van frameworks als Keras en Tensorflow. Dit aangezien de projectgroep, voor aanvang van het project, weinig ervaring heeft met het ontwikkelen van machine learning software. Keras is gebruiksvriendelijk en ondersteunt het ontwikkelen van CNN's voor beginners zeer goed. Aangezien Keras gebruikt maakt van TensorFlow wordt ook TensorFlow toegepast in de applicatie. PyTorch wordt niet

ingezet aangezien in de huidige ontwikkelingsfase optimalisatie niet van groot belang is, Dit kan wel een mogelijkheid zijn voor de toekomst. Ook Scikit-Learn zal niet gebruikt worden omdat de focus van dit framework ligt op dataverwerking en dat niet van toepassing is bij het huidige project.

## 4.2. Beschrijving Artefact

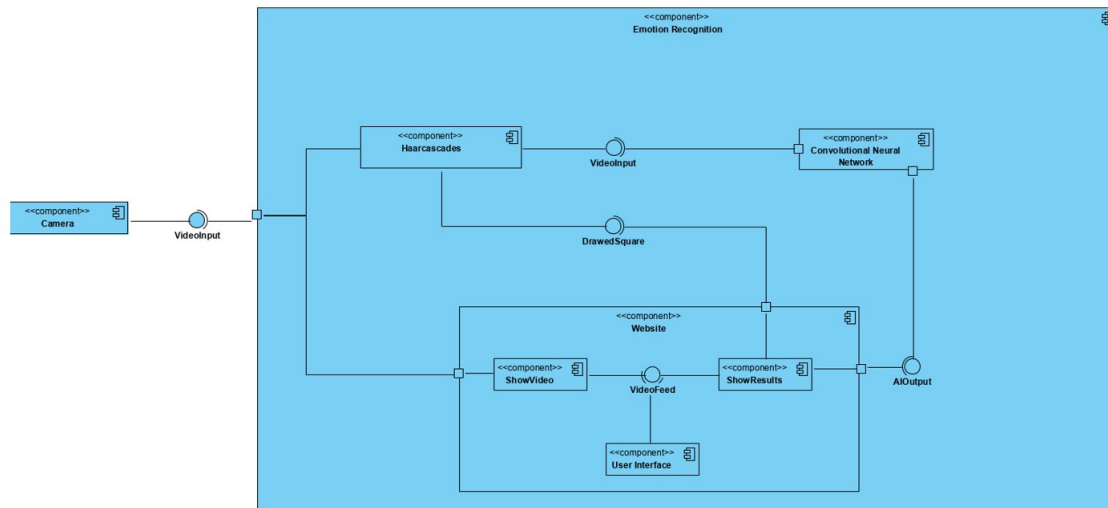


Figure 4.2: Component Diagram

Aan de hand van de onderzoeksresultaten is het bovenstaande component diagram gemaakt, wat diende als het ontwerp. De oplossing heeft meerdere kleine componenten die afhankelijk van elkaar zijn.

Doormiddel van de onderzoeksresultaten en het ontwerp kan de uiteindelijke oplossing in de vorm van de demonstrator gerealiseerd worden. Het ontwikkelen van de demonstrator is gedaan in een viertal fasen. De eerste stap was het ontwikkelen van een netwerk die een tweetal emoties kon herkennen. Dit vormde de basis van de demonstrator en geeft ruimte om te exploreren hoe best-practices voor het implementeren van een convolutional neural network een dergelijk netwerk beter maken ten opzichte van andere mogelijke oplossingen. De tweede stap was het trainen van meerdere emoties in het convolutional neural network. Dit moet de werking van het netwerk aantonen op een grotere schaal. De derde stap was het omzetten van de demonstrator in een Docker container. De vierde stap was het verbeteren van de visualisatie van het convolutional neural network [4].

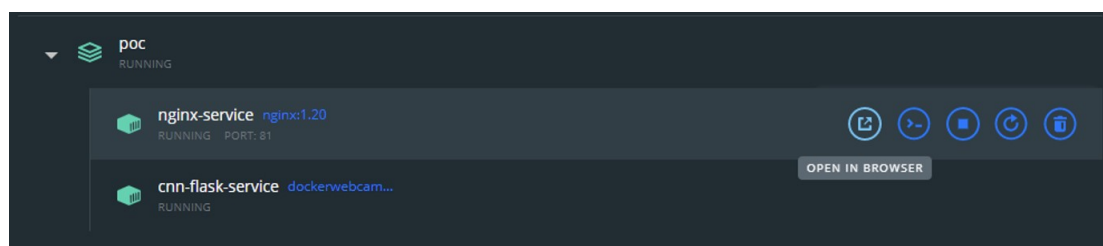


Figure 4.3: Docker weergave

In Figuur 4.3 is de Docker container te zien die ook aan het runnen is. Als de gebruiker op de knop drukt om in de browser te openen dan is het mogelijk om de demonstrator te gebruiken.



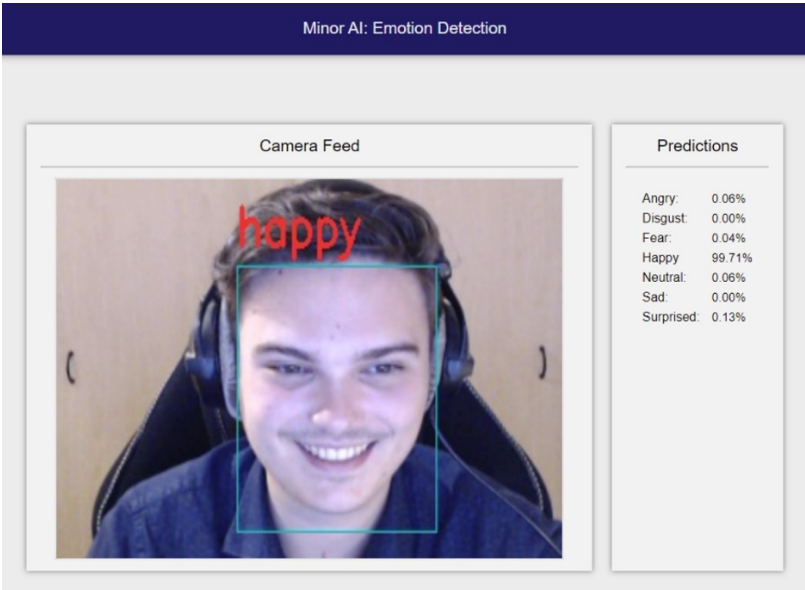


Figure 4.4: Demonstrator die een blijde emotie detecteert

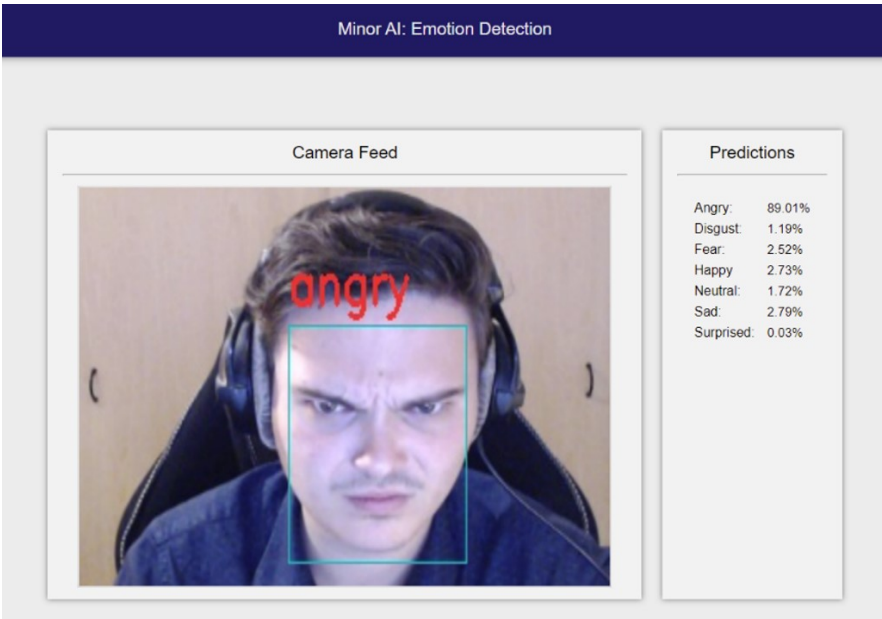


Figure 4.5: Demonstrator die een boze emotie weergave

Figuur 4.4 en Figuur 4.5 geven de webpagina weer van het product, wat dient als de demonstrator. Hierbij wordt de webcam feed van de gebruiker weergegeven samen met een vierkant over het gezicht en de gedetecteerde emotie. Voor duidelijkheid wordt alleen enkel de meest prominente emotie boven het vierkant weergegeven. Onder Predictions zijn alle zeven emoties te vinden, samen met het percentage van waarschijnlijkheid van desbetreffende emoties. Deze zeven emoties worden niet weergegeven op hoogste naar laagste probability, deze staan namelijk vast. Risico bestaat dat als dit wel op hoog naar laag weergegeven wordt, het voor de gebruiker moeilijker te lezen is.

# 5

## Discussie

Binnen het Software Requirements Specification is beschreven dat de Convolutional Neural Network (CNN) in staat moet zijn om aan de hand van een livebeeld minimaal twee verschillende emoties moet kunnen herkennen. Daarbij dient de CNN zelf opgebouwd en getraind te worden zonder het gebruik van een (externe) dienst[4].

Het is essentieel om de belangrijkste features van die camerabeelden te converteren naar een 48x48 resolutieframe voor de training, zodat de load van het CNN wordt verminderd. Deze belangrijkste features zijn in het geval van het MVP het gezicht die een bepaalde emotie laat zien. Zo is het niet relevant om de gehele camerabeelden mee te nemen als data voor het CNN-model. Het probleem is dan hoe we van een webcam feed de belangrijkste aspecten kunnen halen die essentieel zijn voor het trainen van de CNN.

Om dit op te lossen is gebruik gemaakt van een aantal haarcascade xml-bestanden die het mogelijk maken om het gezicht binnen camerabeelden te herkennen. Deze bestanden zijn afkomstig van GitHub onder OpenCV [1]. Doormiddel van deze haarcascades is het mogelijk om het gezicht van beeldmateriaal te isoleren. Daarbij wordt ook door codering het geïsoleerde gezicht van de webcam feed geconverteerd naar een 48x48 resolutie met zwart-witte kleuren.

Deze dienst is alleen gebruikt voor het isoleren van gezichten uit de webcam feed. De dienst wordt niet gebruikt om het CNN op te bouwen of te trainen, maar enkel ter ondersteuning van de input voor het CNN. De haarcascades mogen niet voor commerciële doeleinden gebruikt worden. In het kader van het project is dit geen probleem, aangezien dit streeft naar demonstratiedoeleinden. Dit is tevens overlegd met de opdrachtgever, waarbij werd benadrukt dat de demonstrator alleen als gratis educatieve oplossing wordt gebruikt.

Voor toekomstig gebruik is het belangrijk het detecteren van gezichten te vervangen met een eigen systeem als gewenst wordt om alsnog de demonstrator voor commerciële doeleinden te gebruiken. Dit kan door drie manieren:

- Een eigen set aan getrainde cascades bouwen.
- De dataset vervangen door een dataset van verschillende emoties verdeeld tussen verschillende aanzichten van het gezicht (bv. Met lichaam, alleen boven torso, zijaanzicht, etc.).
- Een tweede CNN die gezichten detecteert en daarbij de locatie van het gezicht bepaald (regressie) in plaats van classificatie zoals in dit EA besproken.

Dit laatste advies is af te raden wegens de complexiteit. Eigen cascades of een nieuwe dataset zou zowel kwalitatief als kwantitatief de beste resultaten moeten leveren.

# 6

## Conclusie

De doelstelling van het project is om middels Convolutional Neural Networks, met name emotieherkenning, het MKB bewust(er) te stellen van de mogelijkheden van Artificial Intelligence. De doelstelling nam twee probleemstukken mee:

- Hoe bouwen we een Convolutional Neural Network die verschillende emoties kan herkennen?
- Hoe kunnen we middels de emotie herkenning het MKB aansporen tot het gebruik van Artificial Intelligence?

Doormiddel van onderzoek werd antwoord gegeven op de hoofd- en deelvragen. Zo luidt het antwoord op de hoofdvraag “Hoe kunnen we met behulp van Python een Convolutional Neural Network algoritme opstellen die emoties kan herkennen aan de hand van een live webcam feed?” dat dit aan de hand gedaan kan worden van een combinatie van frameworks als Keras en Tensorflow.

Aan de hand van het opdrachtformulier, samen met gesprekken met de opdrachtgever, is een Software Requirement Specification opgesteld met minimale acceptatiecriteria waaraan het product moet voldoen voor desbetreffende acceptatie. Aan de hand van de resultaten uit het testrapport is gebleken dat het gebouwde demonstrator voldoet aan alle eisen. Zo is het gelukt om dit probleemstuk op te lossen.

Aan de hand van het opdrachtformulier, samen met gesprekken met de opdrachtgever, is een Software Requirement Specification opgesteld met minimale acceptatiecriteria waaraan het product moet voldoen voor desbetreffende acceptatie. Aan de hand van de resultaten uit het testrapport is gebleken dat het gebouwde demonstrator voldoet aan alle eisen. Zo is het gelukt om dit probleemstuk op te lossen.

# **Appendices**

# Bibliography

- [1] Alalek. Github haarcascades, April 2020. URL <https://github.com/opencv/opencv/tree/master/data/haarcascades>.
- [2] Feroz Amirkhan and invette Tempelman. Artificial intelligence (ai). URL <https://www.mkb.nl/standpunten/artificial-intelligence-ai>.
- [3] Camille Boreau. Ai voor het mkb - bespaar tijd en geld, November 2020. URL <https://www.salesforce.com/nl/blog/2020/11/ai-voor-het-mkb.html>.
- [4] Dwayne Debets, Roald Janssen, Bryan Kroon, and Gavin Melis. Bd02 emotion recognition portfolio, November 2021.
- [5] IBM Cloud Education. Neural networks, Augustus 2020. URL <https://www.ibm.com/cloud/learn/neural-networks>.
- [6] A.R. Hevner. A three cycle view of design science research., 2007. URL <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1017&context=sjis>.
- [7] Sumit Saha. A comprehensive guide to convolutional neural networks - the eli5 way, December 2018. URL <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.