

4-11-2021

BD02 – Artificial Intelligence

## BD02 Emotion Recognition Portfolio

---

*Auteur(s):* Dwayne Debets, Roald Janssen, Bryan Kroon, Gavin Melis  
*Studentnummer(s):* 1329081, 1519166, 1761978, 1852787

*Module:* BD02 – Artificial Intelligence  
*Projectgroep:* 7

*Studie:* HBO-ICT  
*Document:* BD02 Emotion Recognition Portfolio  
*Versie:* 1.0  
*Datum:* 4-11-2021

---

## Inhoud

1.	Inleiding.....	2
2.	Methode .....	3
3.	Onderzoeksopzet .....	5
3.1.	Context .....	5
3.2.	Probleemomschrijving.....	5
3.3.	Doel.....	5
3.4.	Onderzoeksmethode .....	6
3.5.	Doelgroep .....	6
4.	Programma van Eisen .....	7
4.1.	Prioritering van de eisen.....	8
5.	Convolutional Neural Network .....	9
6.	Ontwerp .....	12
6.1.	Frameworks .....	12
6.2.	Component diagram .....	13
7.	Realisatie CNN.....	14
7.1.	Ontwikkel Proces .....	15
7.2.	Product .....	15
8.	Resultaten .....	19
9.	MVP-inzetbaarheden .....	23
10.	Testrapport .....	25
10.1.	Inleiding .....	25
10.2.	Scope .....	25
10.3.	Integration testing .....	26
10.4.	Acceptatie Test .....	27
10.5.	Testen van verschillende scenario's .....	28
11.	Conclusie .....	29
12.	Bibliografie .....	30

## 1. Inleiding

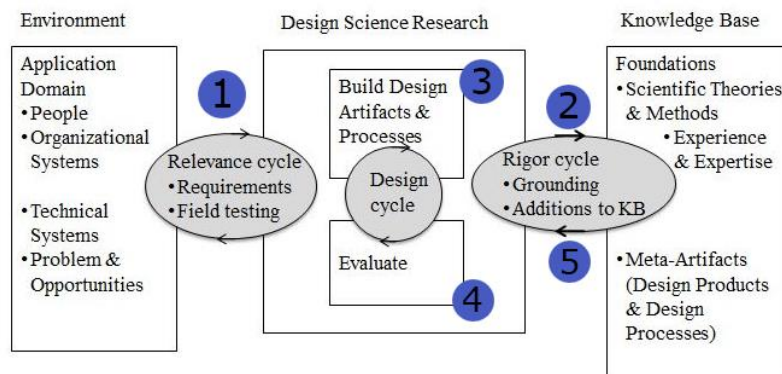
Artificial Intelligence (AI), oftewel kunstmatige intelligentie, bestaat uit allerlei geavanceerde technologieën waardoor machines als het ware kunnen denken, leren en taken uit kunnen voeren zoals mensen. AI wordt sterk bepalend voor onze toekomstige welvaart en welzijn. Door het generieke karakter 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 (Amirkhan & Tempelman).

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 (Boreau, 2020). Om MKB'ers helpen te laten realiseren wat de mogelijkheden van AI zijn werken de studenten aan een prototype, op een onderzoek gerichte manier. Dit prototype is in essentie een demonstrator die in staat is om aan de hand van een live beeld emoties te herkennen.

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 (IBM Cloud Education, 2020).

## 2. Methode

Om het project in goede banen te leiden wordt er gebruik gemaakt van de Design Science Research methode van Hevner (Hevner, 2007). 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.



*Figuur 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:



*Figuur 2 - Aangepaste (op maat gemaakte) Hevner methode*

### 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.

### 3. Onderzoeksopzet

#### 3.1. Context

Het lectoraat Data Intelligence van Zuyd Hogeschool heeft de opdracht gegeven aan de projectgroep om een demonstrator te ontwikkelen die de werking van een Convolutional Neural Network in kaart brengt.

Het doel van het huidige project is het ontwikkelen van een demonstrator die emoties kan herkennen op basis van live feed camera beelden. Deze tool wordt ingezet om de werking van een CNN te demonstreren.

#### 3.2. Probleemomschrijving

In de afgelopen jaren is het concept van Artificial Intelligence steeds meer populair en uitgebreid geworden. Door de groei worden ook mensen zonder kennis steeds meer bij het vakgebied betrokken.

Mensen zonder kennis van het vakgebied hebben vaak een ander idee van wat Artificial Intelligence is, hoe het werkt en hoe het kan worden ingezet. Dit kan leiden tot een andere visie van de resultaten en mogelijkheden van projecten.

Om dit te voorkomen wordt dit project uitgevoerd om informatie te verspreiden over de werking van Artificial Intelligence met als hoofddoel een Convolutional Neural Network. De doelgroep van dit project zijn studenten van Zuyd hogeschool en het MKB.

#### 3.3. Doel

Het doel van het onderzoek is kennis te vergaren en te delen over hoe een CNN kan toegepast worden om emoties te herkennen op basis van camera beelden. Met behulp van het onderzoeksrapport kunnen stakeholders en geïnteresseerde de argumentatie en redenering achter gemaakte keuzes vinden op basis van bevindingen.

Op basis hierop is de hoofdvraag van dit onderzoek “Hoe kunnen we met behulp van Python een CNN-algoritme opstellen die emoties kan herkennen aan de hand van een live webcam feed?”

De resultaten en bevindingen van het onderzoek worden gebruikt voor het ontwikkelen en ontwerpen van onze tool.

### 3.4. Onderzoeksmethode

In het onderzoek wordt er gebruik gemaakt van één onderzoeksmethode. Deze methode wordt uitgevoerd om antwoord te geven op de verschillende deelvragen. Bij de methode wordt een conclusie gegeven. Tot slot zal het onderzoek een algemene conclusie bevatten met daarin antwoord op de hoofdvraag. De conclusie wordt gebaseerd op de bevindingen van het onderzoek en de antwoorden op de deelvragen.

De onderzoeksmethode die wordt toegepast is een literatuuronderzoek. Bij deze methode worden gepubliceerde wetenschappelijke artikelen en onderzoeken bestudeerd. Uit deze artikelen en onderzoeken wordt informatie gehaald die van belang is voor het huidige project. Het doel van deze methode is het vergaren van informatie over Convolutional Neural Networks en hoe deze ingezet kunnen worden voor emotie herkenning.

Bij deze methode wordt eerst een zoekplan opgesteld. Met behulp van dit zoekplan worden de wetenschappelijke bronnen gezocht, geverifieerd of deze voldoen aan de eisen en bestudeerd. Op basis van bevindingen/resultaten wordt er antwoord gegeven op de onderstaande deelvragen:

- Wat is een CNN en hoe moet deze ingezet worden?
- Welke frameworks zijn er voor het gebruik van een CNN en hoe sluiten deze aan op emotieherkenning?

### 3.5. Doelgroep

De doelgroep van het onderzoek zijn een aantal stakeholders van het bijhorende project. Niet alle stakeholders hebben belang bij het onderzoek en worden hier niet benoemd. Ook is het onderzoek van belang bij geïnteresseerde personen in het vakgebied van Artificial Intelligence.

De eerste belanghebbende voor het onderzoek is het lectoraat Data Science van Zuyd Hogeschool. Het lectoraat Data Science is de opdrachtgever van het project. Ze hebben belang bij het onderzoek, omdat ze de redenen en argumentatie achter de gemaakte keuzes willen weten. Ook moeten ze inzicht hebben over de gemaakte keuzes aangezien ze de mogelijkheid willen hebben om verder te ontwikkelen.

De tweede belanghebbende zijn studenten van Zuyd hogeschool. Studenten van Zuyd hogeschool kunnen door middel van het onderzoek meer diepgang krijgen over het vakgebied van Artificial Intelligence. Hierdoor kunnen de studenten zich beter oriënteren en exploreren in verschillende vakgebieden. Er is vooral sprake van studenten die geïnteresseerd zijn of al bezig zijn met de opleiding ICT, dit sluit echter andere opleidingen niet uit.

De derde belanghebbende is het MKB. Het MKB is een organisatie die kleine tot middel bedrijven van 2 tot 50 mensen ondersteund in hun ontwikkeling en groei. Ze functioneren ook als netwerk waarin informatie gedeeld kan worden die van toepassing kunnen zijn voor deze bedrijven. De bedrijven die met Artificial Intelligence werken of in vakgebieden waarbij Artificial Intelligence kan worden toegepast. Hierdoor kan het onderzoek voor hun een introductie zijn in de werking van bepaalde aspecten van Artificial Intelligence.

## 4. Programma van Eisen

In de onderstaande tabel worden de betekenissen van bepaalde termen weergegeven om eventuele conflicten/miscommunicatie te voorkomen. Deze betekenissen worden in dit document gebruikt en hier wordt ook aan gehouden.

Term	Betekenis
<b>Functionele eis</b>	Een functionele eis is een eis/requirement dat het gewenste gedrag van het systeem weergeeft. Alles wat nodig is om het systeem correct te laten werken behoort tot functionele eisen.
<b>Niet-functionele eis</b>	Een niet-functionele eis is een eis/requirement dat een kwaliteitseis is waaraan het systeem moet voldoen. Niet-functionele eisen zijn <u>niet</u> noodzakelijk om het systeem te laten werken.
<b>MoSCoW-analyse</b>	<p>De MoSCoW-methode is een wijze van prioriteiten stellen van eisen. De eisen aan het resultaat van een project worden ermee ingedeeld.</p> <p>De letters M, S, C, W staan voor:</p> <ul style="list-style-type: none"><li>- M = Must have</li><li>- S = Should have</li><li>- C = Could have</li><li>- W = Won't have</li></ul> <p>De kleine letters 'o' in de afkorting hebben geen betekenis, maar maken de afkorting makkelijker te onthouden.</p>
<b>Must (MoSCoW)</b>	Als een eis deze prioriteit krijgt is het noodzakelijk dat deze eis voltooid wordt, zonder deze eis is het product niet bruikbaar. Ook vallen deze eisen onder de acceptatiecriteria voor het Minimal Viable Product (MVP).
<b>Should (MoSCoW)</b>	Als een eis deze prioriteit krijgt is het gewenst dat deze eis voltooid wordt, maar zonder is het product wel bruikbaar.
<b>Could (MoSCoW)</b>	Als een eis deze prioriteit krijgt is het ook gewenst dat deze eis voltooid wordt, maar worden alleen uitgevoerd als er tijd en budget over is.
<b>Won't (MoSCoW)</b>	Als een eis deze prioriteit krijgt is het niet aangeraden om deze eis op dit moment te voltooien, maar kan eventueel in de toekomst uitgevoerd worden als de situatie ervoor geschikt is.



De eisen zijn gesteld uit gesprekken met experts/procesbegeleiders en de analyse die uitgevoerd is door de projectgroep.

#### 4.1. Prioritering van de eisen

De eisen zijn geprioriteerd door middel van de MoSCoW methode. Deze onderstaande tabellen staan op volgorde van hoe belangrijk de eisen zijn.

##### Must

F/NF	Eis
<b>F1</b>	Het systeem moet losstaand functioneren als een demonstrator.
<b>NF1</b>	De demonstrator moet met Python gebouwd worden. (3.9)
<b>F2</b>	De demonstrator moet als webapplicatie aan te bieden zijn of moet te integreren zijn met een webapplicatie.
<b>NF2</b>	De applicatie moet te draaien zijn als een of meer Docker containers.
<b>F3</b>	De demonstrator moet in staat zijn om aan de hand van een live beeld minimaal twee verschillende emoties te herkennen.
<b>F4</b>	De Convolutional Neural Network dient zelf opgebouwd en getraind te worden zonder het gebruik van een dienst.

##### Should

F/NF	Eis
<b>F5</b>	De demonstrator moet de percentages van waarschijnlijkheid op emoties kunnen weergeven.
<b>F6</b>	De demonstrator geeft de mogelijkheid om meerdere gedetecteerde emoties (per gezicht) weer te geven.
<b>F7</b>	De demonstrator kan emoties van meer dan één gezicht binnen de live-feed detecteren.
<b>F8</b>	Het Convolutional Neural Network moet tenminste zeven verschillende emoties kunnen detecteren.

##### Could

F/NF	Eis

##### Won't

F/NF	Eis
<b>F9</b>	De demonstrator geeft de mogelijkheid om weer te geven waarop het Convolutional Neural Network is getraind.

## 5. Convolutional Neural Network

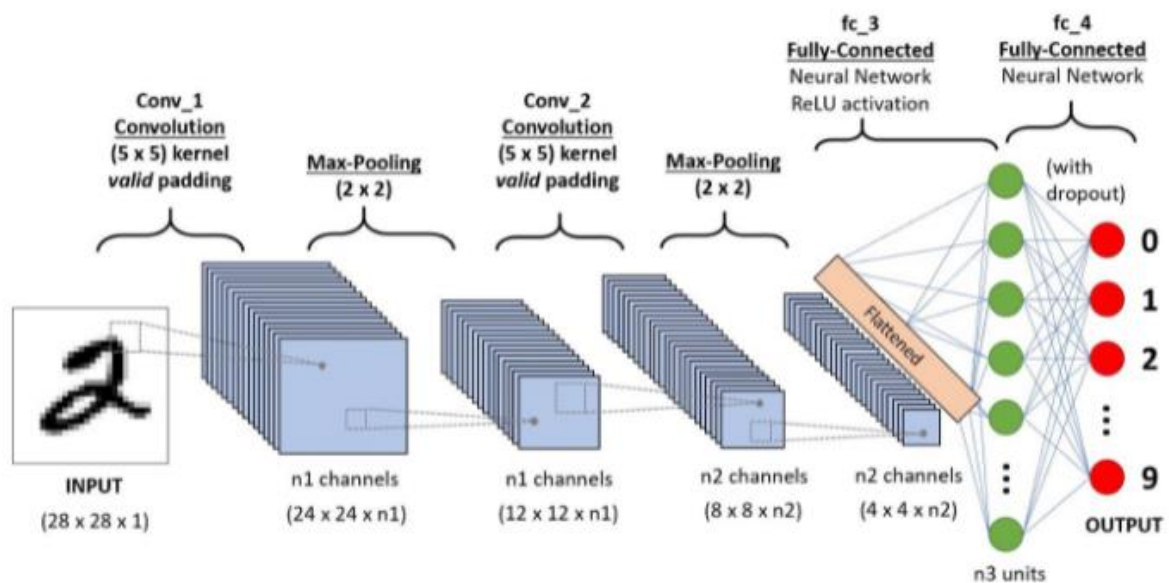
Mensen kunnen intuïtief emoties van andere mensen herkennen. Binnen Artificial Intelligence is het mogelijk om machines de wereld te kunnen laten zien zoals mensen dat kunnen. Door een specifiek algoritme genaamd Convolutional Neural Networks (CNN) is het mogelijk om doormiddel van beelden te herkennen welke emotie een persoon uit. Binnen dit hoofdstuk wordt dieper ingegaan op de werking van een CNN.

Een CNN is een Deep Learning algoritme 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 (Saha, 2018).

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.

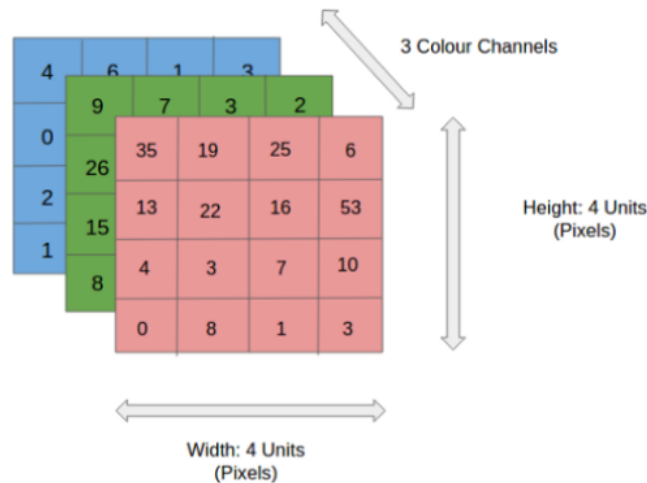
Het grootste verschil tussen een Convolutional Neural Network en een standaard feed-forward Artificial Neural Network (ANN) is dat een CNN het mogelijk stelt om succesvol relevante filters te kunnen genereren voor het classificeren van beeld. De CNN-architectuur geeft een betere fitting voor datasets 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.



Figuur 3 - Een CNN sequentie om handgeschreven nummers te herkennen

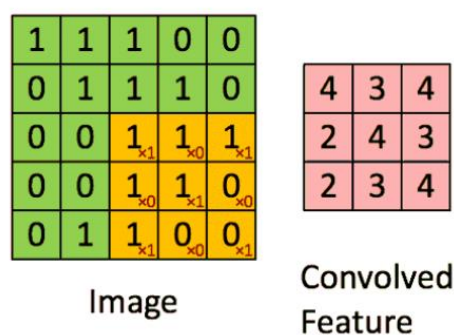
De eerste laag van een CNN is de Input Image. De onderstaande afbeelding laat zien hoe een RGB-beeld wordt onderverdeeld door zijn drie kleuren: rood, groen en blauw.



*Figuur 4 - 4x4x3 RGB-afbeelding*

Het is in te denken hoe een machine veel moeite kan hebben met het berekenen van beelden met een hoge resolutie. Daarbij beschrijft de opdracht ook dat aan de hand van een live-beeld emoties herkend moeten kunnen worden. Met een hoge resolutie zou dit voor een machine lang kunnen duren, waardoor het “live-effect” wellicht impact kan ervaren. Het CNN is bedoeld om deze grote beelden en de grootte van de calculaties die hierbij komen kijken in te korten, naar een vorm die makkelijker is om te processen. Op deze manier verliezen we geen features die anders cruciaal zouden zijn om een goede voorspelling te kunnen maken. Dit ontzettend belangrijk op het moment dat een architectuur ontworpen wordt die niet alleen verschillende features kan leren, maar ook schaalbaar is voor grotere datasets. Een vorm van deze grotere datasets is bijvoorbeeld meerdere emoties die het CNN moet kunnen herkennen.

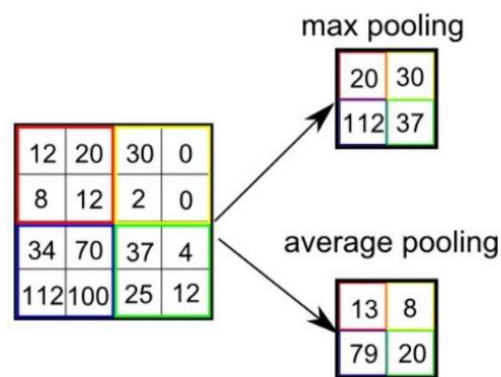
Binnen de Convolutional layer worden high-level features gehaald die essentieel zijn om beelden te kunnen classificeren.



*Figuur 5 – Convolution*

Convolutional Neural Networks hoeven niet beperkt te zijn om maar één Convolutional Layer te bevatten. De eerste Convolutional Layer is verantwoordelijk om low-level features te extraheren zoals randen, kleuren etc. Met meerdere Convolutional Layers is het mogelijk om ook high-level features te kunnen extraheren, wat de mogelijkheid biedt om beelden beter te kunnen begrijpen.

De Pooling layer is vergelijkbaar met de Convolutional layer, in de zin dat deze ook verantwoordelijk is om de spatial size te verminderen van de Convolved Feature. Het nut hiervan is om de rekenkracht van de computer te verminderen dat benodigd is om data te processen. Daarbij is deze laag ook uiterst relevant voor het extracten van dominante features die rotatie- en positioneel invariant zijn. Zo helpt deze laag ook bij het effectief trainen van het CNN-model. Er bestaan twee typen van Pooling: Max Pooling en Average Pooling. Max Pooling geeft de maximale waarde van een portie van het beeld wat bedekt is door de kernel. Average Pooling geeft de gemiddelde waarde van die portie.



*Figuur 6 - De twee typen van Pooling*

De Convolutional Layer en de Pooling layer vormen samen de i-th laag van een Convolutional Neural Network. Afhankelijk van de complexiteiten van de afbeeldingen zal wellicht de aantal lagen die de i-th laag vormen moeten worden uitgebreid. Zo is het mogelijk om low-level details te kunnen differentiëren van elkaar, maar dit kost natuurlijk meer rekenkracht voor de machine. Op het moment dat de i-th laag succesvol de verschillende features kan herkennen gaan we verder in het CNN-architectuur naar de Fully Connected Layer. Zo wordt de output van de i-th laag gebruikt als de input voor een normaal Artificial Neural Network voor de classificatiedoeleinden (Saha, 2018).

## 6. Ontwerp

### 6.1. Frameworks

#### TensorFlow

TensorFlow is een gratis open source framework voor het ontwikkelen van machine learning software. TensorFlow is een library die gebruikt maakt van dataflow en differentiële programmering om de mogelijkheid aan te bieden van het ontwikkelen van machine learning, met als focus het trainen en inferentie van deep neural networks. Het is een van de meest populaire en meest gebruikte frameworks voor machine learning op het moment, zeker met betrekking tot de python programmeer taal. Het wordt onder andere gebruikt door google voor onderzoek en productie.

#### PyTorch

Pytorch is ook een gratis open source framework voor het ontwikkelen van machine learning software. Het is alleen inzetbaar op python en c++. De focus van Pytorch ligt op het training en maken van Artificial Intelligence door middel van het gebruiken van de GPU, en deep neural network development. Het wordt met name ingezet voor het optimaliseren van AI-software die gebruikt maakt van CPU en GPU's.

#### Keras

Keras is een gratis open source machine learning framework die gebruikt maakt van tensorflow. De focus van Keras ligt op het makkelijk toepasbaar te maken van neurale netwerken voor beginners. Ze geven ook de mogelijkheid om Convolutional Neural Networks simpel en snel te ontwikkelen.

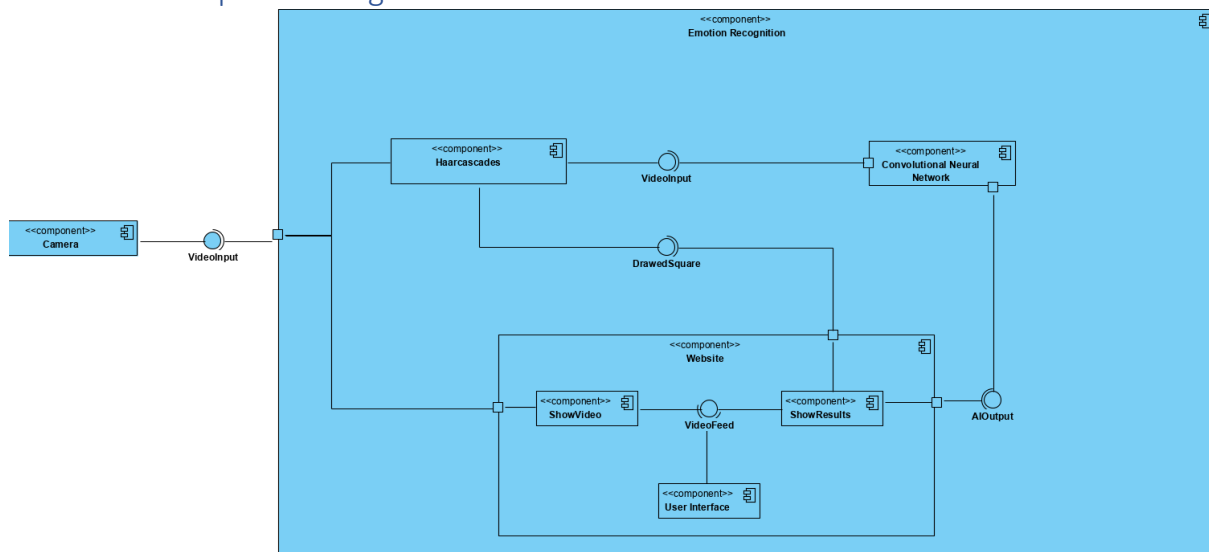
#### Scikit-learn

Scikit-learn is ook een gratis opensource machine learning framework. Het biedt machine learning functionaliteiten aan die met name nadruk heeft op dataverwerking zoals classificatie, regressie en clustering. Het is alleen bruikbaar in python.

#### Keuze framework

Aangezien de projectgroep een Convolutional Neural Network zal ontwikkelen en weinig ervaring heeft het met het ontwikkelen van machine learning software, zal er gebruik gemaakt worden van een combinatie van Keras en TensorFlow. 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.

## 6.2. Component diagram



*Figuur 7 - Component Diagram*

In het bovenstaande diagram is het component diagram te zien van de tool die ontwikkeld wordt. De tool heeft meerdere kleine componenten die afhankelijk zijn van elkaar.

De camera levert de video input die de tool nodig heeft om te weergeven en te analyseren. De component levert het videobeeld aan de webpagina en het haarcascades component.

De haarcascade component zorgt ervoor dat de gezichten in de video input worden gefilterd om te gebruiken voor het Convolutional Neural Network. De haarcascades zet het uitgefilterde gezicht om naar een gestandaard formaat van 48x48. Ook levert de haarcascades component een getekend vierkant om het gezicht in de video weergave, dit levert de component aan de webpagina, specifiek het ShowResults component.

De component genaamd Convolutional Neural Network bevat de AI die wordt ingezet voor het analyseren van de input. De component heeft een foto van een gezicht nodig om te functioneren. Deze krijgt de component van het haarcascades component. Het CNN analyseert het ontvangen gezicht en berekent hieruit een kans uit van welke emotie wordt getoond. Het percentage per emotie wordt gestuurd naar de webpagina component, specifiek de ShowResults component.

De websitecomponent bevat meerdere kleinere componenten. Het doel van de websitecomponent is de omvang van de hele website. De website verzocht meerdere functionaliteiten, vandaar wordt het opgesplitst in kleinere componenten. De Showvideo component verzorgt de video weergave die de camera voorziet. Deze wordt verbonden aan het user interface component om het te weergeven aan de gebruiker door middel van de interface. De ShowResults component verzorgt het weergeven van het getekende vierkant om het desbetreffende gezicht en de output die de AI heeft berekend voor het gezicht. Deze worden per gezicht berekend en gecreëerd. De component voorziet het User Interface component van deze resultaten.

De User Interface component verzorgt wat de gebruiker te zien krijgt. Deze aspecten bevatten, de AI-uitkomst, het getekende vierkant om de gezichten, het videobeeld en de interface waarmee de user interacteert.

## 7. Realisatie CNN

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

Om het CNN te kunnen trainen zal gebruik gemaakt moeten worden van trainingsdata. Om de load van het CNN te verminderen zijn deze afbeeldingen verwerkt naar een 48x48 pixels met een zwart-wit beeld. Deze tactiek wordt binnen het kader van het project gebruikt om (sneller) een CNN op te kunnen stellen zodanig dat deze acceptabel is voor het Minimal Viable Product (MVP). Wel kan in de toekomst geëxploreerd worden of een grotere resolutie en/of kleur succesvol ingezet kunnen worden binnen het CNN-model.

Om de afbeeldingen met emoties, die zullen dienen voor de trainingsdata, op te kunnen halen is gebruik gemaakt van codering. Deze codering haalt in essentie de camerabeelden op vanuit de webcam. Het is essentieel om de belangrijkste features van die camerabeelden te converteren naar een 48x48 resolutieframe voor de training. Deze belangrijkste features zijn in het geval van het MVP puur het gezicht die een bepaalde emotie laat zien. Zo is het niet relevant om de gehele camerabeelden mee te nemen naar de trainingsdata voor het CNN-model, waarbij het CNN eventueel “verkeerde” filters zal toepassen die niet succesvol een emotie kunnen herkennen. Het probleem luidt dus hoe we puur van een webcam feed de belangrijkste aspecten kunnen halen die essentieel zijn voor het trainen van de CNN. Idealiter zou de webcam alleen het gezicht van mensen moeten kunnen weergeven. Om dit probleem 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 (Alalek, 2020).

Doormiddel van deze haarcascades is het mogelijk om het gezicht te isoleren. Daarbij wordt ook door codering het geïsoleerde gezicht van de webcam feed geconverteerd naar een 48x48 resolutie met zwart-witte kleuren. In conclusie is zodanig gebruik gemaakt van een (externe) dienst, maar het gebruik van deze dienst valt onder het isoleren van gezichten uit de webcam feed. De dienst wordt niet gebruikt om het CNN op te bouwen of te trainen, maar enkel om de trainings/testdata te kunnen verzamelen.

Het is relevant om te vermelden dat deze haar cascades niet voor commerciële doeleinden gebruikt mogen worden. In het kader van het project is dit geen probleem, aangezien het doel is om een demonstrator te maken om aan MKB'ers te kunnen laten zien wat de mogelijkheden van Artificial Intelligence zijn. Ook is overlegd met de opdrachtgever of het gebruik van deze haar cascades acceptabel is binnen de context van het project. Daarbij werd geantwoord dat het gebruik binnen deze context prima is. Wel zal met dit feit rekening gehouden moeten worden indien stakeholders besluiten om de demonstrator juist wel voor commerciële doeleinden te gebruiken.

### 7.1. Ontwikkel Proces

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 vormt 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.

### 7.2. Product

Een convolutional neural network is opgebouwd uit twee delen, de convolutie laag en de volledig verbonden laag (Mishra, 2020). In de convolutie laag wordt de input, in het geval van emotieherkenning een foto van een gezicht, verwerkt en gecomprimeerd om zoveel mogelijk van het plaatje in een zo klein mogelijk formaat te stoppen. Bij deze convolutie is het van belang dat zo min mogelijk data verloren gaat en de belangrijkste kenmerken van de foto naar voren komen. Een typisch kenmerk is dat zo'n convolutie meestal in serie wordt uitgevoerd, hoewel het theoretisch mogelijk is om deze parallel uit te voeren. De voornaamste reden voor het bestaan van de convolutie laag is echter het reduceren van de complexiteit van de foto, zodat deze sneller en effectiever verwerkt kan worden. Het vaststellen van een emotie is de taak van de volledig verbonden laag. In de volledig verbonden laag is de structuur van een standaard neurale netwerk te vinden. Hierin wordt als eerst de output van de convolutie laag getransformeerd naar een 1-dimensionale reeks aan getallen. Deze worden verwerkt volgens de werking van een algemeen artificial neural network, waaruit de uiteindelijk output (de emotie van een persoon) wordt uitgelezen.

De eerste fase hield zich bezig met de opzet van het convolutional neural network. De opbouw van het netwerk valt in twee onderdelen te verdelen. De training en de topologie van het netwerk.

De topologie van het netwerk valt onder te verdelen in deze van de convolutie laag en die van de volledig verbonden laag.

De volledig verbonden laag, of wel het standaard neurale netwerk, heeft een input laag, een verwerking (verborgen) laag en een output laag. Dit netwerk wordt gedefinieerd met de grote (het aantal knooppunten in het netwerk), diepte (het aantal lagen in het netwerk), breedte (het aantal knooppunten in een specifieke laag), en architectuur (de samenstelling aan knooppunten in het netwerk) van het netwerk. De samenstelling van knooppunten in een netwerk is direct verantwoordelijk voor de effectiviteit van het netwerk (Brownlee, 2019). Een knooppunt (met uitzondering van een input knooppunt) krijgt zijn waarde door de verbonden knooppunten van een eerdere laag op te tellen, ieder vermenigvuldigd met een gewicht, en daaraan toegevoegd een bias (toegevoegde waarde). Deze waarde wordt vervolgens eventueel door een transformatie of activatie functie heen gehaald, waarmee de output van het knooppunt tussen een bepaald bereik kan worden gehouden, en eventueel de lineaire aard van de originele verbinding kan worden doorbroken. Dit is dan ook de kracht van een neurale netwerk.

In de convolutie laag wordt multi-dimensionale informatie samengevat. Dit wordt gedaan middels een combinatie aan drie mechanismen: Het kernel, pooling en een activatiefunctie. Het kernel is een soort filter of lens die de lokale patronen van de foto samenvat (convolutie). De grote van zo'n kernel kan variëren. De belangrijkste aanrader voor het gebruik van kernels is symmetrie, en een zo klein mogelijke kernel gebruiken. Door de grootte klein te houden wordt de hoeveelheid nodige parameters niet te groot om te berekenen, en om hiervoor te compenseren kan een zo klein mogelijke kernel worden gebruikt om dezelfde



lokale patronen te vinden. De symmetrie van een kernel zorgt dat de grote van de output van de kernel gelijk kan blijven aan de input. Dit kan door de grote van de kernel gelijk te stellen aan een vlakte van een oneven aantal pixel voor in hoogte en breedte voor de input foto. Zo kan voor iedere pixel de omliggende pixels bij worden gevoegd (niet bestaande buurpixels zijn gelijkend aan een nul-waarde). Bij een even aantal wordt de input grote vervormd naar een andere grote in de output (Pandey, 2020). Voor de kernel zijn alternatieve manieren om de nevenpixels te selecteren voor een kernel dan directe burens van een pixel. Dit kan middels verwijding, door ruimte tussen burens te laten om grotere lokale patronen samen te vatten zonder hiervoor performance in te leveren. Echter leidt dit tot verlies van informatie, wat het doel van de convolutie tegen gaat: Zoveel mogelijk data van de input samenvatten in een zo klein mogelijke output om verwerking te bevorderen. Een andere techniek zou zijn door vervormbare convolutie, hoewel deze meer complex is een de afstand voor iedere pixel in de kernel toevoegt als extra parameter. Dit zorgt voor beter extractie van features, maar voegt ook een extra complexiteitslaag toe aan de kernel (Ganesh, 2019). De kernel kan de grote van de output veranderen ten opzichte van de input, hoewel dit niet hoeft, hiervoor kan pooling worden gebruikt om de uitkomst van een kernel te downsamplen. Zo wordt het aan parameters voor de convolutie verminderd en wordt tevens overfitting van de trainingsdata op het model tegengegaan. De meeste gebruikte en eenvoudigste pooling methoden zijn average pooling, max-pooling en mixed pooling. De pooling methoden hebben alle drie net als een kernel een specifieke grote, en hebben een stapgrote die kan afwijken van de grote van de pooling groep. Average pooling neemt een gemiddelde van een groep pixels, terwijl max-pooling alleen de grootste waarde selecteert uit deze groep. Mixed-pooling is een pseudo-random methode waarbij voor iedere groep apart wordt gekozen of deze gebruik maakt van average- of max-pooling, dit wordt meegenomen in het trainingsproces (Gholamalinezhad & Khosravi, 2020).

Het probleem met de combinatie van de kernel en pooling mechanisme is dat deze mogelijk lineair zijn zoals ook het probleem is bij een klassiek neurale netwerk. Om dit tegen te gaan is ook bij een Convolutional Neural Network een activatie functie belangrijk. De meeste gebruikte activatiefuncties zijn de sigmoid, tanh, Relu, LeakyRelu en softmax (Amor, 2020). De sigmoid en tanh worden gebruikt om de output te schalen tussen 0 en 1, en -1 en 1 respectievelijk. De sigmoid functie is het meest bruikbaar voor een probabilistische output, terwijl de tanh functie de data centreert en dus better werkt in de tussenlagen. Beide hebben echter het probleem dat bij hoge of lage waarde de leercurve te traag is. Relu stelt alle negatieve getallen gelijk aan 0 en laat zo alleen positieve waarde invloed hebben op het eindresultaat. Dit voorkomt de problemen met de leercurve binnen sigmoid en tanh, maar werkt niet als het netwerk overweldigend veel negatieve getallen gebruikt, waarbij negatieve waarde de leercurve niet beïnvloeden. Dit staat bekend als het strevende Relu probleem. Dit wordt opgelost met LeakyRelu met een vaste leercurve voor negatieve getallen in plaats van 0. Softmax is een speciale activatie functie die de totale som van de outputs naar een laag gelijkstelt aan 1, en daarmee iedere output net als een sigmoid functie schaalt naar een waarde tussen de 0 en 1. Dit maakt de softmax de beste oplossing voor een output voor meer dan 2 classificaties. Zo kunnen bijvoorbeeld meerdere emoties binnen een netwerk worden gedetecteerd, met de veronderstelling dat deze onderling afhankelijk zijn.

De topologie gebruikt voor het convolutioneel neural netwerk bestaat uit de volgende setup:

De convolutie laag bestaat uit 4 series aan convoluties. Na iedere convolutie wordt de output genormaliseerd zodat de verdeling van de output niet wordt beïnvloed door te grote waarde verschillen. Daarna wordt dit resultaat door een Relu activatie gehaald om negatieve resultaten te negeren, en worden de meeste actieve output waarden opgehaald met max-pooling volgens een 2 bij 2 groepering. Om overfitting te voorkomen wordt een 25% kans toegevoegd dat een pool-waarde op 0 wordt gezet. Voor de convolutie worden meerdere filters gebruikt om het resultaat van een enkele convolutie te berekenen. Iedere filter is verantwoordelijk voor het zoeken naar een eigen set aan features. De eerste

convolutie combineert het resultaat van 64 filters met een kernel grote van 3 bij 3; de tweede convolutie combineert 128 filters met een kernel grote van 5 bij 5; de derde convolutie combineert 512 filters met een kernel grote van 3 bij 3; en de vierde convolutie combineert ook 512 filters met een kernel grote van 3 bij 3.

De volledig verbonden laag vormt het resultaat van de laatste convolutie naar een 1-dimensionale set aan data. Deze wordt vervolgens verbonden aan een reeks aan 2 verborgen lagen, waarna nog een output laag. De output van de verborgen lagen worden achteraf genormaliseerd en daarna door een Relu activatie gehaald om negatieve resultaten te negeren. Hierna wordt ook weer overfitting tegengegaan met een 25% kans om een waarde naar 0 te zetten zoals gedaan bij de convolutie. De eerste verborgen laag heeft 256 knopenpunten en de tweede verborgen laag heeft 512 knooppunten. De output wordt vervolgens bepaald met een output van 7 eindknopen, bepaald volgens de softmax activatie, waarmee de kans dat een emotie mogelijk is wordt gegeven. De totale kans van alle 7 output is daarmee 100%.

Het trainen van een neuraal netwerk wordt gedaan door deze eerst te initialiseren met willekeurige gewichten voor alle knopenpunten en kernels. Het netwerk wordt iteratief op een dataset getraind waarbij de volledige dataset meerdere keren wordt doorlopen. Iedere keer wordt het neurale netwerk een klein beetje bijgesteld om te zorgen de output van het neurale netwerk meer aansluit op de verwachte output van de dataset. Hoeveel keer de dataset wordt doorlopen kan op meerdere manieren worden bepaald: Dit kan vooraf worden vastgezet op een specifiek aantal, of het trainingsalgoritme kan tussendoor kijken of verder trainen van toegevoegde waarde is. Tijdens het trainen is sprake van een leercurve, hoe sterk deze leercurve is kan vooraf worden ingesteld. De leercurve bepaald hoe extreem de iteratieve aanpassing aan de gewichten in het netwerk zijn. Deze leercurve ook dynamisch kan worden ingesteld om te zorgen dat deze minder extreme aanpassing na mate het neurale netwerk beter wordt. Dit heeft als voordeel dat het neurale netwerk dichter een mogelijke optimale opstelling komt, hoewel het ook kan dat het netwerk vast komt te zitten een lokaal optimum waardoor het netwerk niet meer verbetert terwijl deze wel een mogelijke verbetering kan doormaken.

Voor de training is gebruik gemaakt van een leercurve van 0.0005 en een early-stop algoritme die de training stopt als het netwerk in de vorige 15 volledige iteraties niet is verbeterd met een accuraatheidsgraad van minstens 0.001. Bij de leercurve is gekeken wat voor een impact de dynamische aanpassing heeft op het trainen van een netwerk. Dit heeft geen verschil gemaakt in het generen van een optimaal netwerk, aangezien hier de gemiddelde aantal iteraties 42 bleef, zowel met, als zonder dynamische leercurve. Niet alleen bleef het aantal iteraties hetzelfde, maar ook bleef de accuraatheidsgraad hetzelfde. De training van het netwerk vindt plaats buiten de demonstrator. Bij het trainen is gebruik gemaakt van hardware acceleratie middels een grafische kaart, omdat het trainen zeer intensief is. Dit is niet mogelijk binnen een Docker container. Dit heeft op kleine datasets weinig effect, maar naar mate het aantal foto's (de dataset) toenam heeft de inzet van de grafische kaart de trainingstijd verminderd naar 1/4<sup>de</sup> van de originele trainingstijd.

Bij het trainen op meerdere emoties is het netwerk getraind om boze, walgende, angstige, blij, verdrietige, veraste en neutrale gezichten te herkennen. Het neurale netwerk de meeste problemen gehad met het herkennen van walging. Dit was voornamelijk wegens dat niet veel data beschikbaar was voor walging. Twee oplossingen zijn hierop losgelaten: Undersampelen en oversampelen. Deze twee oplossingen zorgen dat evenveel datapunten beschikbaar zijn voor alle emoties. Door te undersampelen werd het netwerk meer accuraat dan bij het oversampelen. Dit proces is drie keer doorlopen en bleef hetzelfde. Een probleem met het undersampelen is daarentegen wel dat ondanks dat deze gevalideerd werd beter te zijn, mist deze ook een gedeelte van de dataset. Dit leidt tot grote problemen wanneer een schaarste aan datapunten

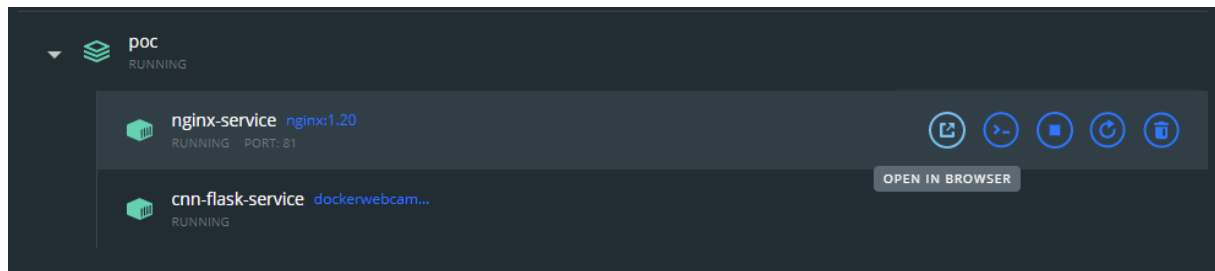
voordoet zoals bij de undersampling voor dit netwerk. Dit was ook terug te zien wanneer het netwerk in de praktijk werd gebruikt.

Voor de demonstrator moest een python webapplicatie worden gerealiseerd waarin het convolutional neural network bepaald aan de hand van een webcam feed welke emotie personen op de webcam feed hebben.

Voor het web-framework is Flask gebruikt, waarin met OpenCV een webcam feed is opgehaald. Met OpenCV zijn middels Haar-cascades gezichten gezocht op de feed. Per frame is iteratief ieder gezicht door het neurale netwerk beoordeeld, en de het beste antwoord is vervolgens getoond boven ieder bijhorend gezicht. Met javascript werd vervolgens in de webbrowser de webcam feed ingeladen met de gemaakte voorspelling.

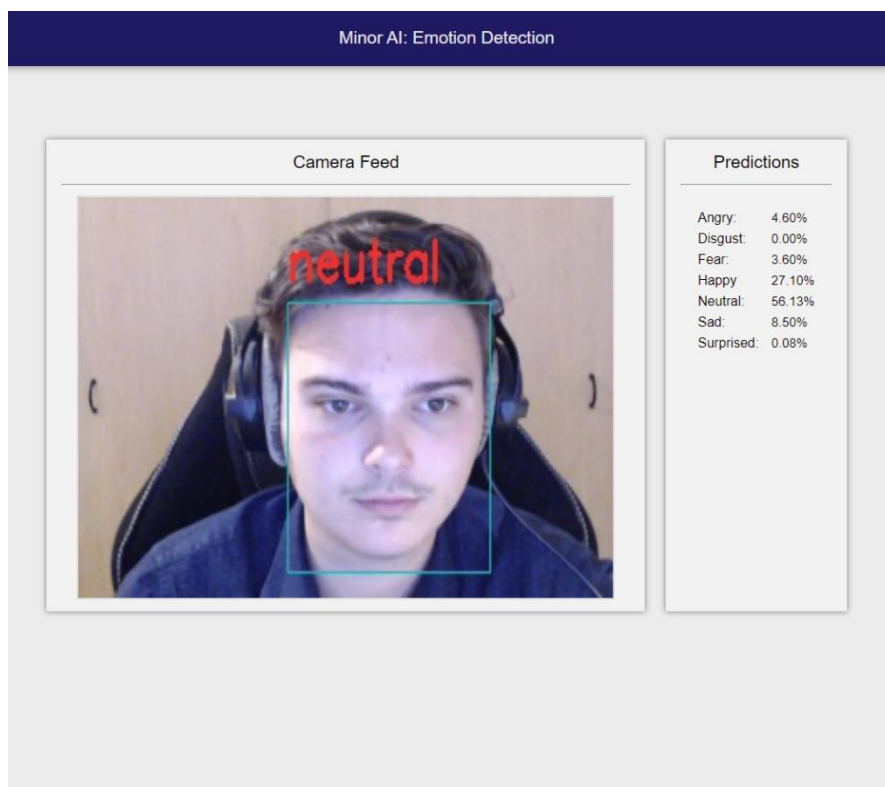
Een probleem met deze aanpak is dat deze applicatie niet functioneert in een Docker container. De webcam feed werd namelijk van de camera van de server genomen en niet van de webclient. Om dit werkend te krijgen is het Flask framework uitgebreid naar het Flask-Websocket-io framework. Met javascript een websocket geopend en via dit socket wordt een frame van de webcam gestuurd die de Flask server vervolgens verwerkt terugstuurt. Deze applicatie is in een afgeschermd Docker container gestopt die wordt gemanaged door een nginx container. Deze oplossing werkt over lange afstanden op het internet en kan vanaf iedere computer en moderne webbrowser worden gebruikt (hoewel sommige webbrowsers een ssl-certificaat vereisen om webcam toegang toe te staan). Een laatste optimalisatie van de webapplicatie was gemaakt in de resolutie die de webbrowser tekent voor de webcam feed. Een hoge resolutie zorgt dat de frames niet getekend werden. Dit is opgelost door de resolutie te beperken tot 320x180 pixels. Dit heeft als extra voordeel dat een enkele openstaande verbinding maar 1kb per seconde gebruikt, waardoor bandwidth gebruik van de server minimaal blijft.

## 8. Resultaten



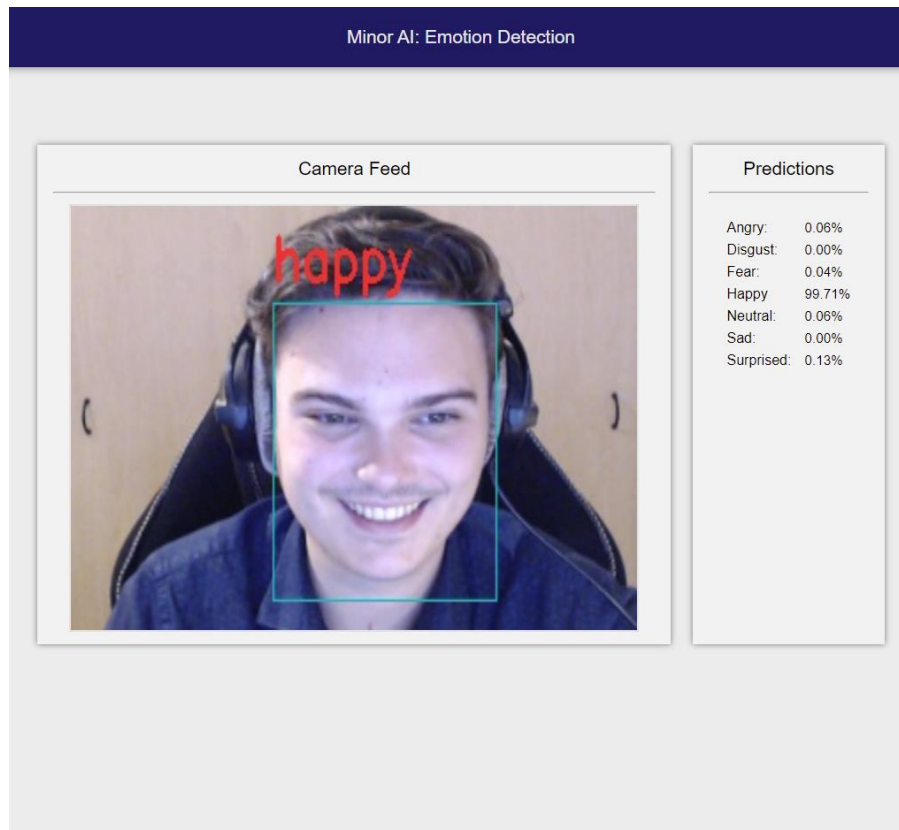
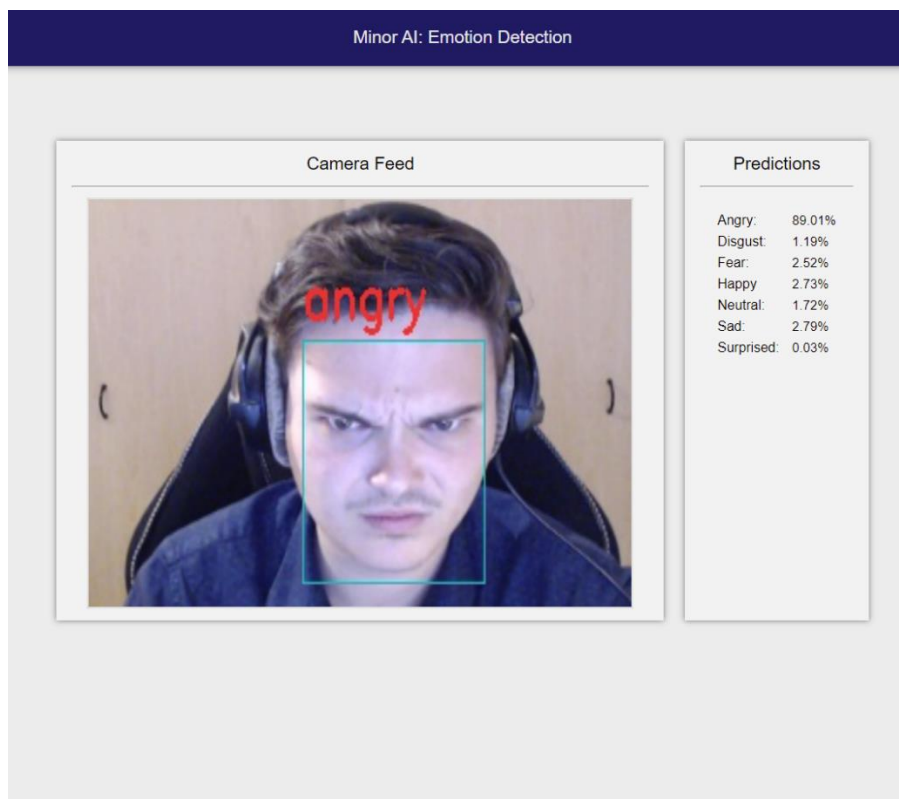
Figuur 8 – Docker

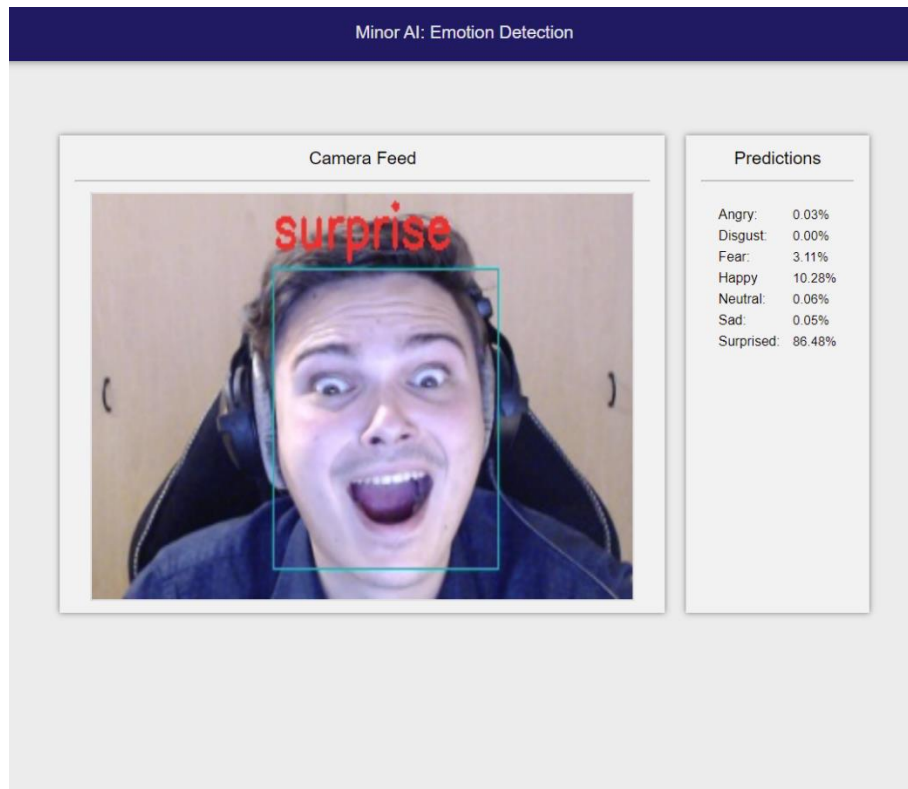
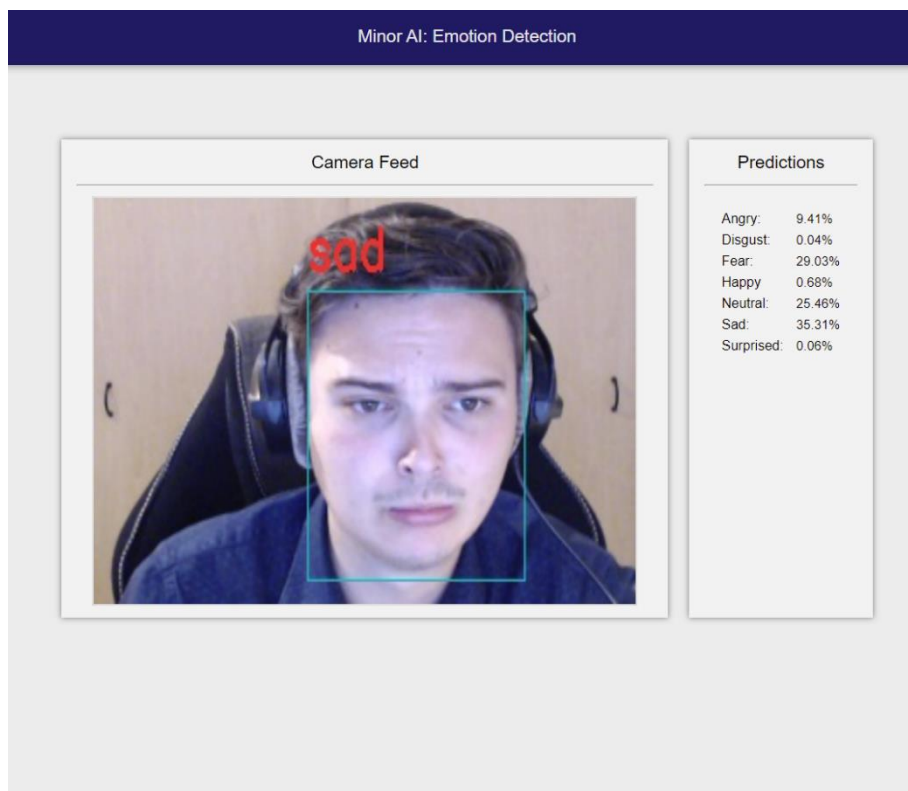
In de bovenstaande afbeelding is de Docker container te zien die ook aan het runnen is. Als de gebruiker op de knop drukt om in de browser te openen dan volgt het product.

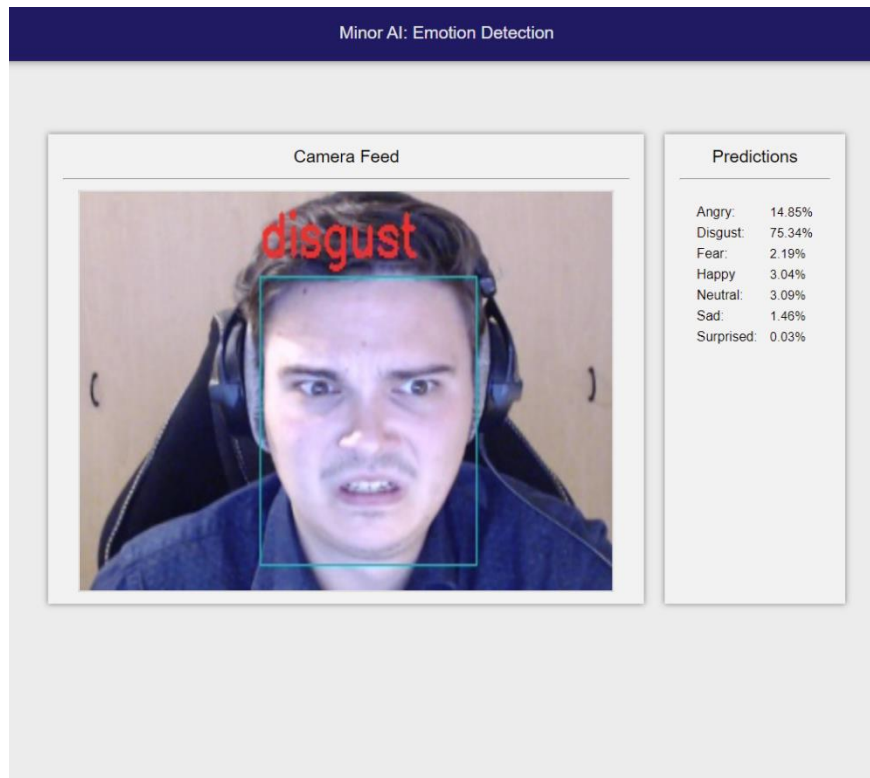


Figuur 9 - Neutrale emotie

De bovenstaande afbeelding geeft de webpagina weer van het product. 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. De volgende afbeeldingen geven allen de verschillende soorten emoties weer.

*Figuur 10 - Blijde emotie**Figuur 11 - Boze emotie*

*Figuur 12 - Verrassende emotie**Figuur 13 - Verdrietige emotie*



*Figuur 14 - Walgend emotie*

De huidige versie van de webbrowsers die gebruikt zijn, waar de demonstrator het doet, zijn:

- Chrome: Versie 95.0.4638.69 (Officiële build) (64-bits)
- Edge: Versie 95.0.1020.40 (Officiële build) (64-bits)
- Firefox: 94.0 (64-bit)



## 9. MVP-inzetbaarheden

### Psychologie

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.

Hiermee kan de psycholoog dieper ingaan op hoe een patiënt reageert en waar de problemen liggen. Het kan zijn dat de psycholoog iets over het hoofd ziet dat de tool wel zal herkennen. Dit zal het herkennen van problemen bij de patiënt versnellen en verbeteren.

Echter zal het in dit vakgebied minder van toepassing aangezien psychologen getraind zijn op het herkennen van emoties en reacties bij patiënten. Echter zal de tool wel nog als ondersteuning kunnen worden toegepast.

### Entertainmentindustrie (films/games/speelgoed/etc.)

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.

Hiermee kunnen ontwikkelaars van het stuk media snel zien waar de problemen in het stuk media liggen. Dit zal de scores van de stukken media verhogen, kost minder personeel en kost minder om te onderzoeken dan een uitgebreid onderzoek.

### Horeca

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. Ook kan er naar de data gekeken worden wat de correlatie is tussen de tevredenheid van klanten en de hoeveelheid fooi die ze geven.

Hiermee kunnen horecazaken sneller herkennen waar de klanten niet tevreden mee zijn zonder een uitgebreid onderzoek uit te laten voeren of extra personeel in te huren.

Echter heeft de tool camera feed nodig. Dit houdt in dat een camera moeten worden gericht worden op de klanten van de zaken. Dit kan klanten anders laten gedragen, sneller ontevreden maken of afstoten van de zaak.



## Clubs/bar

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.

## UI/X design

Bij het ontwerpen van UI van applicaties/software kan de tool worden ingezet om de te kijken hoe gebruikers reageren op bepaalde aspecten van de User Interface. Er kan gekeken worden hoe gebruikers reageren op bepaalde figuren, kleuren, plaatjes, etc. Op basis hiervan kan de website aangepast worden naar zijn doelstelling.

Hiermee kunnen UI/X ontwerpers hun fouten opsporen waarmee gebruikers en opdrachtgevers minder tevreden zouden zijn. Dit zou testers van de interfaces zeer goed ondersteunen in hun taak om de acceptatie van een UI's te verbeteren.

## 10. Testrapport

### 10.1. Inleiding

Er wordt een testplan opgesteld om de tool systematisch te testen. Het doel van het testen is het opsporen van fouten en problemen die in de tool zitten. Dit zal zorgen voor hogere kwaliteit van de tool. In het test rapport wordt beschreven welke test soorten er worden gebruikt, waarom deze test soorten en worden gebruikt.

### 10.2. Scope

Eerst wordt er gekeken naar de scope van het testrapport. Hierin wordt beschreven wat er wel en niet getest zal worden, en waarom.

Eerst wordt er beschreven wat wel getest zal worden daarna zal gekeken worden naar wat niet getest wordt.

Er zijn voor twee test methodes gekozen. De eerste test methode die wordt toegepast is Integration Testing. Deze test methode focust zich op het testen van de communicaties tussen verschillende modules. Het doel van de tool is het realiseren van een AI die emoties kan herkennen. Om dit modulair in te zetten bij andere applicaties, moet dus getest worden of de tool de correcte output weergeeft en doorgeeft aan andere onderdelen.

De tweede test methode die wordt toegepast is System testing. Het doel van deze test methode is het testen van een systeem in zijn geheel. Hierbij wordt getest of de tool in zijn geheel correct functioneert. De tool moet de correcte output leveren op de input die de camera levert.

De derde test methode die wordt toegepast is acceptatie test. Het doel van een acceptatie test is om te controleren of het ontwikkelde product acceptabel volgens de eisen die gesteld zijn door de opdrachtgever. Hierbij wordt er dus gekeken of alle must have requirements zijn voldaan en correct functioneren.

Ook wordt de tool/ai getest op verschillende scenario's. Mensen kunnen verschillende kenmerken hebben of de omgeving waarin de tool/ai wordt ingezet kan verschillen. Denk bijvoorbeeld aan lichtinval, huidskleur van de persoon, accessoires zoals brillen. Om te kijken dat de AI/tool niet selectief mensen eruit filtert, hier wordt dus op getest.

Er wordt niet getest op de werking van de tool met andere applicaties of software. Dit is van toepassing bij door ontwikkeling van de tool, echter is dit niet toepasbaar in de huidige ontwikkelingsfase van de tool.

### 10.3. Integration testing

Integration testing focust zich op het testen van de communicaties tussen verschillende modules. Aangezien de ontwikkelde tool/AI in verschillende omgevingen geïmplementeerd kunnen worden, moet het mogelijk zijn dat verschillende componenten vervangen of verwijderd worden.

Aangezien het systeem zeer modulair moet zijn is dit een grote prioriteit om te testen. Dit zal de inzetbaarheid bij toekomstige applicatie verbeteren, wat voor de groei van het product een zeer grote impact heeft.

Het systeem wordt bottom-up getest. Dit houdt in dat de modules van low level naar high level werking wordt getest. Dit wordt gedaan om de losstaande componenten eerst te testen voordat ze samen worden gevoegd. Aangezien het CNN-component het belangrijkste component is voor de werking van de tool zal deze ook het meest in detail moeten worden getest. Als deze niet werkt zullen de boven liggende componenten ook niet werken.

De volgorde van testing is als volgende:

Component	Beredenering
1. CNN	Het CNN is het belangrijkste component voor de werking van de tool om aan de must have functionaliteiten te voldoen. Hiervoor is het belangrijk dat deze als eerst wordt getest zonder invloed van buitenstaande componenten.
2. Camera	De camera moet correct videobeeld leveren. Als er geen correct beeld wordt geleverd kunnen andere aspecten niet correct getest worden.
3. Haarcascades	Haarcascades zorgen voor input voor de CNN en de getekende vierkanten bij de gezichten.
4. Showvideo	Daarna wordt getest op het tonen van de video en resultaten op de website. Dit is het tonen van de werking aan de gebruiker. Deze twee componenten kunnen parallel getest worden.
5. Showresults	Daarna wordt getest op het tonen van de video en resultaten op de website. Dit is het tonen van de werking aan de gebruiker. Deze twee componenten kunnen parallel getest worden.
6. User Interface	Het User Interface component is de combinatie van het tonen van de video, resultaten en UI aan de gebruiker op de website. Dit is het minst belangrijk voor de werking van de tool en zal daarom als laatste getest worden.

### 10.4. Acceptatie Test

Bij een acceptatietest worden alle must have requirements geëvalueerd.

De requirements die getest worden zijn:

<b>F1</b>	Het systeem moet losstaand functioneren als een demonstrator.
<b>NF1</b>	De demonstrator moet met Python gebouwd worden. (3.9)
<b>F2</b>	De demonstrator moet als webapplicatie aan te bieden zijn of moet te integreren zijn met een webapplicatie.
<b>NF2</b>	De applicatie moet te draaien zijn als een of meer Docker containers.
<b>F3</b>	De demonstrator moet in staat zijn om aan de hand van een live beeld minimaal twee verschillende emoties te herkennen.
<b>F4</b>	De Convolutional Neural Network dient zelf opgebouwd en getraind te worden zonder het gebruik van een dienst.

De resultaten van de testen zijn:

Eis nummer	Voldaan?	Waarom/hoe?
F1	Ja	Het systeem is modulair ontwikkeld. Alle verschillende componenten kunnen vervangen worden of aangepast worden zonder dat een ander component daardoor wordt beïnvloed.
NF1	Ja	Alle software is ontwikkeld in Python.
F2	Ja	De tool is bereikbaar door een webpagina. Deze is ook ter beschikking gesteld met behulp van Docker
NF2	Ja	Docker is ingezet bij het implementeren van de tool.
F3	Ja	Het CNN wordt getraind om 7 verschillende emoties herkennen. Deze emoties zijn: <ul style="list-style-type: none"> <li>- Neutraal</li> <li>- Blij</li> <li>- Boos</li> <li>- Verrast</li> <li>- Verdrietig</li> <li>- Walgend</li> <li>- Beangstigend</li> </ul>
F4	Ja	Het CNN is zoals eerder beschreven in hoofdstuk 6 ontwikkeld door de projectgroep zonder gebruik te maken van externe diensten.

### 10.5. Testen van verschillende scenario's

Om ervoor te zorgen dat de tool kan werkt bij alle mensen wordt getest op verschillende scenario's die kunnen plaatsvinden. Om dit te testen worden groepen mensen van verschillende etniciteit geselecteerd en voeren ze dezelfde stappen uit. De groepen mensen zullen tien personen bevatten, waarvan vijf vrouwen en vijf mannen. Hierbij wordt ook gekeken naar mensen met brillen, petten, lichtinval en camera hoek.

Nadat de test is uitgevoerd wordt de data geanalyseerd om te kijken of de tool problemen heeft bij enkele van deze aspecten en kan de training van het Convolutional Neural Network hierop aangepast worden.

De stappen die elke groep zal verlopen is als volgt:

Stap nummer	Stap omschrijving	Verwachte resultaat	Daadwerkelijke resultaat
1.	De persoon toont een neutrale emotie	De tool herkent een neutrale emotie	
2.	De persoon toont een boze emotie	De tool herkent een boze emotie	
3.	De persoon toont een blijde emotie	De tool herkent een blijde emotie	
4.	De persoon toont een verdrietige emotie	De tool herkent een verdrietige emotie	
5.	De persoon toont een walgende emotie	De tool herkent een walgende emotie	
6.	De persoon toont een verassende emotie	De tool herkent een verassende emotie	
7.	De persoon toont een beangstigende emotie	De tool herkent een beangstigende emotie	

Deze stappen worden met de volgende scenario's meerdere keren uitgevoerd:

Scenario nummer	Scenario omschrijving
1.	Geen aanpassingen aan het scenario
2.	De persoon zal een bril dragen
3.	De persoon zal een pet dragen
4.	De camera zal van een in een andere invalshoek worden geplaatst voor een minder goed zicht op het gezicht.
5.	Het licht van de kamer waarin de test plaatsvindt wordt gedimd
6.	Er vindt een combinatie plaats van scenario 2 en 3.
7.	Er vindt een combinatie plaats van scenario 2 en 4.
8.	Er vindt een combinatie plaats van scenario 2 en 5.
9.	Er vindt een combinatie plaats van scenario 3 en 4.
10.	Er vindt een combinatie plaats van scenario 3 en 5.
11.	Er vindt een combinatie plaats van scenario 4 en 5.
12.	Er vindt een combinatie plaats van scenario 2, 3 en 4.
13.	Er vindt een combinatie plaats van scenario 2, 3 en 5.
14.	Er vindt een combinatie plaats van scenario 3, 4 en 5.
15.	Er vindt een combinatie plaats van alle scenario's samen.

## 11. Conclusie

De doelstelling van het project is om middels een Convolutional Neural Network, 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.

Verder zijn verschillende mogelijkheden geëxploreerd met betrekking tot inzet van het gebouwde emotieherkenningsdemonstrator. Deze zijn gehanteerd in de vorm van use-cases. Doormiddel van deze use-cases kan het MKB de relevantie voor hun (eventuele) gebieden evalueren voor eigen gebruik. 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.

## 12. Bibliografie

- Alalek. (2020, April 13). *Github Haarcascades*. Opgehaald van github.com: <https://github.com/opencv/opencv/tree/master/data/haarcascades>
- Amirkhan, F., & Tempelman, I. (sd). *Artificial Intelligence (AI)*. Opgehaald van mkb.nl: <https://www.mkb.nl/standpunten/artificial-intelligence-ai>
- Amor, E. (2020, Mei 29). *Understanding Non-Linear Activation Functions in Neural Networks*. Opgehaald van Medium: <https://medium.com/ml-cheat-sheet/understanding-non-linear-activation-functions-in-neural-networks-152f5e101eeb>
- Bureau, C. (2020, November 16). *AI voor het MKB - Bespaar tijd en geld*. Opgehaald van salesforce.com: <https://www.salesforce.com/nl/blog/2020/11/ai-voor-het-mkb.html>
- Brownlee, J. (2019, Augustus 6). *How to Configure the Number of Layers and Nodes in a Neural Network*. Opgehaald van Machine Learning Mastery: <https://machinelearningmastery.com/how-to-configure-the-number-of-layers-and-nodes-in-a-neural-network/>
- Ganesh, P. (2019, Oktober 18). *Types of Convolution Kernels : Simplified - Towards Data Science*. Opgehaald van Medium: <https://towardsdatascience.com/types-of-convolution-kernels-simplified-f040cb307c37>
- Gholamalizadeh, H., & Khosravi, H. (2020). Pooling Methods in Deep Neural Networks, a Review. *CoRR*, *abs/2009.07485*. Opgehaald van <https://arxiv.org/abs/2009.07485>
- Hervner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*(2). Opgehaald van <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1017&context=sjis>
- IBM Cloud Education. (2020, Augustus 17). *Neural Networks*. Opgehaald van ibm.com: <https://www.ibm.com/cloud/learn/neural-networks>
- Mishra, M. (2020, September 2). *Convolutional Neural Networks, Explained - Towards Data Science*. Opgehaald van Medium: <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>
- Pandey, S. (2020, Juli 1). *How to choose the size of the convolution filter or Kernel size for CNN?* Opgehaald van Medium: <https://medium.com/analytics-vidhya/how-to-choose-the-size-of-the-convolution-filter-or-kernel-size-for-cnn-86a55a1e2d15>
- Saha, S. (2018, December 15). *A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way*. Opgehaald van towardsdatascience.com: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>