

# Machine Learning pipeline voor Facial Image Analysis in camera-gebaseerde gezondheidsmetingen: Schatten van leeftijd en geslacht voor het beoordelen van geestelijke gezondheid.

---

**Alexandra Stalmans.**

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

**Promotor:** Mevr. C. De Leenheer

**Co-promotor:** Dhr. T. Sanglet

**Academiejaar:** 2023–2024

**Eerste examenperiode**

**Departement IT en Digitale Innovatie .**

**HO  
GENT**



# Woord vooraf

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

# Samenvatting

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

# Inhoudsopgave

Lijst van figuren		vii
<b>1</b>	<b>Inleiding</b>	<b>1</b>
1.1	Probleemstelling	1
1.2	Onderzoeksvraag	2
1.3	Onderzoeksdoelstelling	2
1.4	Opzet van deze bachelorproef	3
<b>2</b>	<b>Stand van zaken</b>	<b>4</b>
2.1	Gezichtsdetectie	4
2.2	Feature extractie	5
2.2.1	Normalisatie	5
2.2.2	AAM	6
2.2.3	LBP	6
2.3	Feature dimensionaliteitsreductie	8
2.4	Bestaande uitdagingen	8
2.5	Bestaande Machine-Learning technieken	9
2.5.1	Extreme Gradient Boost	9
2.5.2	Adaptive Boosting	9
2.5.3	Random Forest Classifier	10
2.5.4	Support Vector Machines	11
2.5.5	Ensemble	11
2.5.6	Multi attribution model	12
2.6	Dataset	12
2.7	Voorspellen van geslacht en leeftijd	13
2.8	Performantiescores	13
<b>3</b>	<b>Methodologie</b>	<b>15</b>
<b>4</b>	<b>Conclusie</b>	<b>17</b>
<b>A</b>	<b>Onderzoeksvoorstel</b>	<b>19</b>
A.1	Introductie	20
A.2	Stand van zaken	20
A.2.1	Gezichtsdetectie	21
A.2.2	Bestaande Machine-Learning technieken	22
A.2.3	Bestaande uitdagingen	25

A.3	Methodologie . . . . .	25
A.3.1	Requirements . . . . .	25
A.3.2	Literatuurstudie . . . . .	25
A.3.3	Proof-of-concept . . . . .	26
A.3.4	Conclusie . . . . .	26
A.3.5	Afwerken scriptie . . . . .	26
A.4	Verwacht resultaat, conclusie . . . . .	26

# Lijst van figuren

2.1	Belangrijkste anthropometrische landmarks . . . . .	5
2.2	Gezichtsafbeeldingen voor normalisatie . . . . .	6
2.3	Gezichtsafbeeldingen na normalisatie . . . . .	7
2.4	AAM-features . . . . .	7
2.5	AAM-features . . . . .	8
2.6	Random forest voor houding van het hoofd . . . . .	10
2.7	SVM . . . . .	11
2.8	Robuust gezichtsherkenning systeem . . . . .	12
2.9	Confusion matrix . . . . .	14
A.1	Belangrijkste anthropometrische landmarks . . . . .	21
A.2	Gezichtsafbeeldingen voor normalisatie . . . . .	22
A.3	Gezichtsafbeeldingen na normalisatie . . . . .	23
A.4	Random forest voor houding van het hoofd . . . . .	24
A.5	Robuust gezichtsherkenning systeem . . . . .	24
A.6	Methodologie . . . . .	25
A.7	ROC curve . . . . .	28

# 1

## Inleiding

De onderzoeksvraag werd aangeboden door het bedrijf IntelliProve. IntelliProve biedt online gezondheidsoplossingen, een software die in staat is om binnen enkele seconden nauwkeurig gezondheidsparameters te bepalen, gebaseerd op een optische meting van het gezicht. Het doel van de bachelorproef is het ontwikkelen en implementeren van een robuust systeem voor het schatten van de leeftijd en het geslacht van personen op basis van gezichtsfoto's, met behulp van machine learning-technieken. Dit project is van bijzonder belang voor het verbeteren van de beoordeling van de geestelijke gezondheidszorg door middel van camera-gebaseerde gezondheidsmetingen. Het onderzoek beoogt bij te dragen aan de vooruitgang op dit gebied door gebruik te maken van geavanceerde algoritmen om leeftijd en geslacht nauwkeurig te voorspellen aan de hand van gezichtsbeelden. De literatuurstudie biedt een inzicht in facial analysis, de bestaande machine learning modellen en hun functionaliteiten. De proof-of-concept zal bestaan uit het ontwikkelen van een machine learning pipeline die in staat is om leeftijd en geslacht te voorspellen op basis van bestaande datasets. De pipeline omvat verschillende image preprocessing technieken om de dataset voor te bereiden op de modeltraining. Om betrouwbaarheid en accuracy te garanderen, worden de modellen verfijnd en geoptimaliseerd om de hoogst mogelijke nauwkeurigheid te bereiken bij het schatten van leeftijd en geslacht.

### 1.1. Probleemstelling

Het onderzoek wordt uitgevoerd voor het bedrijf IntelliProve. IntelliProve biedt online gezondheidsoplossingen, een software die in staat is om binnen enkele seconden nauwkeurig gezondheidsparameters te bepalen, gebaseerd op een optische meting van het gezicht. De bachelorproef vormt een opstap naar een applicatie die in de toekomst uitgewerkt zal worden.



## 1.2. Onderzoeksvraag

Het onderzoek beschrijft de ontwikkeling van een machine learning pipeline voor het schatten van leeftijd en geslacht. Specifiek voor het analyseren van gezichtsafbeeldingen in camera-gebaseerde gezondheidsmetingen. Voor het onderzoek werd volgende onderzoeksvraag opgesteld:

- Hoe kan een efficiënte machine learning-pipeline worden ontwikkeld en geoptimaliseerd voor het analyseren van gezichtsafbeeldingen in camera-gebaseerde gezondheidsmetingen, met als specifieke doelen het schatten van leeftijd en geslacht, toegepast om geestelijke gezondheid te beoordelen?

De onderzoeksvraag kan opgedeeld worden in enkele deelvragen:

1. Wat is de geschikte dataset voor het opgegeven probleem om zo veel mogelijk bias te vermijden?
2. Welke uitdagingen bestaan er al uit voorgaand onderzoek en moet rekening mee gehouden worden in het onderzoek?
3. Uit welke stappen bestaat de machine learning pipeline?
  - (a) Welke feature extractie en/of feature dimensionaliteitsreductie toepassingen zijn het meest geschikt voor het voorspellen van leeftijd en/of geslacht?
4. Welke machine learning modellen zijn er mogelijk voor het voorspellen van leeftijd en geslacht?
  - (a) Wordt er 1 model gemaakt om leeftijd en geslacht te voorspellen of worden er 2 modellen gebruikt die zich elk richten tot een specifieke taak?
  - (b) Welk model, uit een vergelijkende studie, geeft de beste resultaten?
5. Hoe kunnen we de performantie van een model meten?

## 1.3. Onderzoeksdoelstelling

Het resultaat van de bachelorproef is een proof-of-concept die zal bestaan uit het ontwikkelen van een machine learning pipeline die in staat is om leeftijd en geslacht te voorspellen op basis van bestaande datasets. De pipeline omvat verschillende image preprocessing technieken om de dataset voor te bereiden op de modeltraining. Om betrouwbaarheid en accuracy te garanderen, worden de modellen verfijnd en geoptimaliseerd om de hoogst mogelijke nauwkeurigheid te bereiken bij het schatten van leeftijd en geslacht. Er wordt naar een zo hoog mogelijk nauwkeurigheid gestreefd, waarbij de conclusies uit dit onderzoek het belangrijkste zijn. Er kan aangegeven worden waarom bepaalde modellen goed werken of juist niet en of er in de toekomst nog verder onderzoek vereist is.

## **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

Gezichtsanalyse bestaat uit het definiëren van menselijke gezichten in real time, met behulp van computeralgoritmen en machine learning technieken. Voor mensen en computersystemen bevat een gezichtsbeeld details zoals leeftijd, geslacht, stemming, afkomst, et cetera.

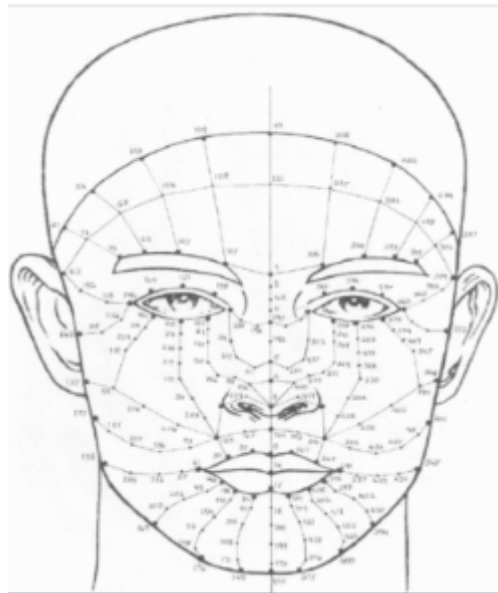
Gezichtsanalyse omvat dan ook het lokaliseren en meten van gezichtskenmerken in een afbeelding. Vervolgens wordt het gezichtsbeeld geanalyseerd voor het extraheren van kenmerken (**Sanil2023**). Gezichtsanalyse speelt dan ook een grote rol in real world applicaties, zoals animatie, identiteitsverificatie, medische diagnose, et cetera. Ondanks het bestaande onderzoekswerk over dit onderwerp, is gezichtsanalyse nog steeds een lastige taak vanwege verschillende factoren zoals veranderingen in hoek, gezichtsuitdrukkingen en achtergrond (**Siddiqi2022**).

Volgens onderzoek van **Basystiuk2023** **<empty citation>** is beeldherkenning een simpele procedure die uit slechts 3 stappen bestaat.

1. Preprocessing: We voegen filters toe aan de afbeelding om het geschikter te maken voor herkenning.
2. Feature Extractie: We identificeren belangrijke data en behouden deze om mee verder te werken. Dit wordt verder besproken in [2.1](#).
3. Classificatie: het analyseren en identificeren van de data na de feature extractie.

### 2.1. Gezichtsdetectie

Gezichtsdetectie is de eerste stap in het bepalen van de gezichtsfeatures. Deze features zijn de interessante delen van het gezicht , bijvoorbeeld ogen, neus en mond

**Figuur (2.1)**

De belangrijkste anthropometrische landmarks(Sanil2023).

en worden ook wel eens landmarks genoemd (Coppens2018). In dit onderzoek raadt de auteur 3 gezichtsdetectieservices aan: Microsoft Face API, Face++ en Kairos. Deze zijn alle 3 ook geschikt om geslacht en leeftijd te voorspellen. Er wordt met 3 services tegelijk gewerkt om elkaars beperkingen op te vangen. In het onderzoek van Sanil2023<empty citation> berekenen ze de landmarks op 2 verschillende methoden: Euclidische afstand en Geodetische afstand.

Sanil2023<empty citation> stelt 468 landmarks visueel voor, in plaats van 68, om de accuracy te verbeteren. Dit door gebruik te maken van de Mediapipe library (Zubair2021<empty citation>). Op figuur 2.1 staan de belangrijkste antropometrische landmarks aangeduid. Het is noodzakelijk om de belangrijkste features te identificeren die kunnen leiden tot het correct voorspellen van de leeftijd of het geslacht. Om de leeftijd te voorspellen kunnen we bijvoorbeeld de rimpels op de gezichtsfoto analyseren (Kwon1994).

## 2.2. Feature extractie

### 2.2.1. Normalisatie

In onderzoek van (Chen2011) voor het voorspellen van leeftijd op basis van afbeeldingen werden de gezichtsafbeeldingen genormaliseerd voor de feature extraction plaatsvond. De geometrie van de afbeeldingen kan worden genormaliseerd op basis van de gedetecteerde features, zoals oog of mond coördinatie. De auteur maakt gebruik van een mask om de pixels die zich niet in de ovaal van de typische gezichtsregio bevinden te verwijderen. Dit gaat dan bijvoorbeeld over haar en hemdskragen. Zo blijven enkel de belangrijkste features van het gezicht over.



**Figuur (2.2)**

Gezichtsafbeeldingen voor normalisatie(Chen2011).

Figuur 2.2 geeft een voorbeeld van een dataset met gezichtsfoto's. In figuur 2.3 worden deze afbeeldingen genormaliseerd, waardoor er minder variaties overblijven in de afbeeldingen en we bijvoorbeeld al geen achtergrond overhouden. De gezichtsafbeeldingen kunnen ook genormaliseerd worden op basis van de coördinaten van de ogen, zodat het centrum van beide ogen van de afbeeldingen op een vaste positie liggen (Chen2011).

### 2.2.2. AAM

In het onderzoek van Lakshmiprabha2016 worden de Active Appearance Model (AAM) features gebruikt in de feature extractie stap. Deze features zijn geschikt om het effect van veroudering weer te geven voor het voorspellen van geslacht. De features die we van AAM krijgen hebben zowel vorm- als textuurinformatie die geschikter is voor veroudering. Figuur 2.4 geeft een representatie van de FG-NET database na het toepassen van de AAM-features. Het resultaat is vergelijkbaar met de normalisatie van Chen2011 op figuur 2.3.

### 2.2.3. LBP

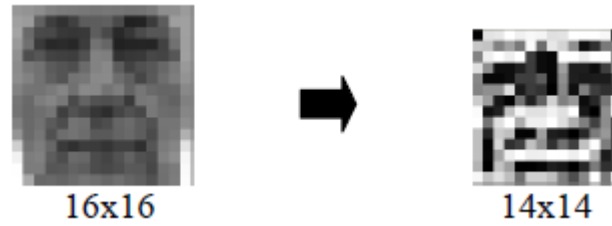
Local Binary Patterns (LBP) is een textuurdescriptor die gebruikt kan worden om gezichten te representeren, aangezien een gezichtsbeelding kan worden beschouwd als een samenstelling van microtextuurpatronen (SanchezLopez2010). De LBP operator kent een label toe aan elke pixel van een afbeelding door een threshold te nemen van de 3x3 burens met de centrale pixelwaarde en het resultaat te beschouwen als een binair getal. Het resultaat na de toepassing van LBP in het onderzoek

**Figuur (2.3)**

Gezichtsafbeeldingen na het toepassen van normalisatie(**Chen2011**).

**Figuur (2.4)**

Cohn Kanada database na toepassen van AAM (**Lakshmi Prabha2016**).



**Figuur (2.5)**

Toepassen van LBP op een 16x16 gezichtsafbeelding resulteert in 14x14 label afbeelding (**SanchezLopez2010**).

van **SanchezLopez2010** is weergegeven op figuur 2.5.

In het onderzoek van **Chen2011** wordt LBP toegepast om de gezichtsfeatures te extraheren uit elk blok die de leeftijdsinformatie bevat. Het resultaat hiervan is een feature dimensie vector van 59. Deze features worden dan samengevoegd in een holistische vector.

### 2.3. Feature dimensionaliteitsreductie

De herkenning van gezicht focust zich voornamelijk op het detecteren van individuele features in het gezicht, zoals ogen, hoofdometrek, mond of het definiëren van het model van gezicht door positie, grootte of relatie tussen de features (**Lin2006**). Het extraheren van de features speelt een grote rol in de preprocessing fase.

De Principal Component Analyse (PCA) wordt veelal gebruikt voor gezichtsherkenning. PCA is een unsupervised, statistische techniek die gebruikt wordt voor dimensionaliteitsreductie in machine learning (**SalihHasan2021**). Het reduceert het hoge aantal dimensies in een grote dataset naar een lager aantal dimensies om de opslag en het verwerkingsproces te versnellen. Hierdoor is de data makkelijker te interpreteren en sneller te verwerken. De techniek behoudt het grootste aantal informatie en verwijdert redundante ruis en data. Het grote nadeel van PCA is dat de lineaire transformatie die wordt uitgevoerd om de grootste varianties te behouden, niet altijd goed geschikt is voor het op te lossen probleem, namelijk de classificatie van geslacht (**Wang2010**).

De feature reductie vindt typisch plaats na de feature extractie fase en voor de eigenlijke classificatie (**Lakshmiprabha2016**).

### 2.4. Bestaande uitdagingen

Ondanks het uitgebreide onderzoekwerk, heersen er nog vele uitdagingen op het vlak van gezichtsanalyse. Deze uitdagingen zijn het gevolg van verschillende factoren zoals gezichtsuitdrukkingen, ruis, belichting, et cetera. Om de nauwkeurigheid van de gezichtsherkenning te verbeteren, is het belangrijk om taken van de

gezichtsanalyse met elkaar te correleren. Het is bijvoorbeeld zeer waarschijnlijk dat mannen een baard of snor kunnen hebben, maar vrouwen en kinderen niet (**Siddiqi2022**).

Ook is het detecteren van de locatie waar de gezichten zich bevinden op de afbeelding een uitdaging. Dit wordt vaak meegenomen in de preprocessing stap van het analysesysteem (**Jiang2008**).

De problemen situeren zich niet enkel op het vlak van de afbeelding, maar ook bij de machine learning modellen. Eén van de grootste problemen hierbij is overfitting. Dit gebeurt wanneer we het algoritme te veel trainen, waardoor het te hard lijkt op de trainingsdata. Hierdoor zal het systeem falen wanneer we nieuwe data willen classificeren (**Coppens2018**).

## 2.5. Bestaande Machine-Learning technieken

Supervised learning is de methode die input mapt naar een output, gebaseerd op een voorbeeld (**Rustam2018**). Supervised learning is gericht op het voorspellen van de waarde van variabelen of labels, op basis van de input features. Het wordt gewoonlijk gebruikt voor classificatie, benadering, modelleren of identificatie. Het groeperen van mannen of vrouwen op basis van gezichtsherkenning is een classificatieprobleem. Er zijn 2 opties: man of vrouw, ofwel 0 of 1. Dit maakt supervised learning de geschikte methode voor de voorspelling van geslacht. Bij supervised learning bestaat de trainingsdata uit een reeks trainingsvoorbeelden, waarbij elk voorbeeld bestaat uit een input en een verwachte output waarde (**Shah2012**). Het algoritme analyseert de trainingsdata en voorspelt dan de correcte output categorie voor de gegeven data input.

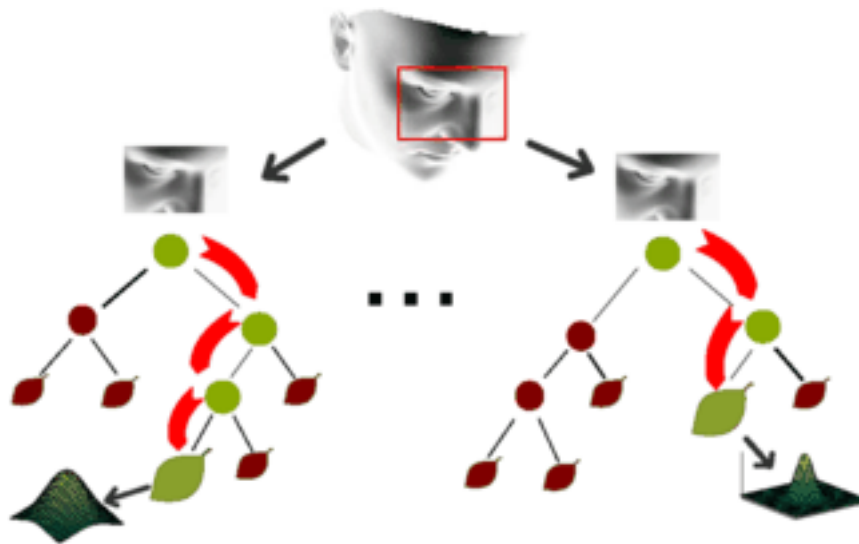
### 2.5.1. Extreme Gradient Boost

Extreme Gradient Boost, ofwel XGBoost, is een supervised learning model, waar trainingdata, met meerdere features,  $x_i$  wordt gebruikt om target waarde  $y_i$  te voorspellen (**XGBoost2023**). Uit onderzoek van (**Chen2023**), waarbij vermoeidheid werd voorspeld op basis van gezichtsafbeeldingen met XGBoost, wordt een boom aangemaakt op basis van de geleverde features. Dit model gaf een hoge accuracy en is minder gevoelig voor noise. Uit onderzoek naar gezichtsdetectie van **Sanil2023<empty citation>** bleek dat Extreme gradient boosting de beste resultaten gaf, met een accuracy van 78%, gevolgd door Adaptive Boosting (77%) en Random Forest (75%).

### 2.5.2. Adaptive Boosting

De Adaptive Boost, of AdaBoost, is typisch een classificatie tussen twee klassen (**Guo2001**). Deze ML-techniek kan dus toegepast worden op de voorspelling van man of vrouw (0 of 1). Om het herkenningsprobleem met meerdere klassen op te lossen, kan een majority voting (MV) strategie worden gebruikt om alle paarsgewijze classificatieresultaten te combineren. Hierbij kiezen we de meest voorko-





**Figuur (2.6)**

Voorbeeld van een random forest voor het schatten van de houding van het hoofd(**Fanelli2012**).

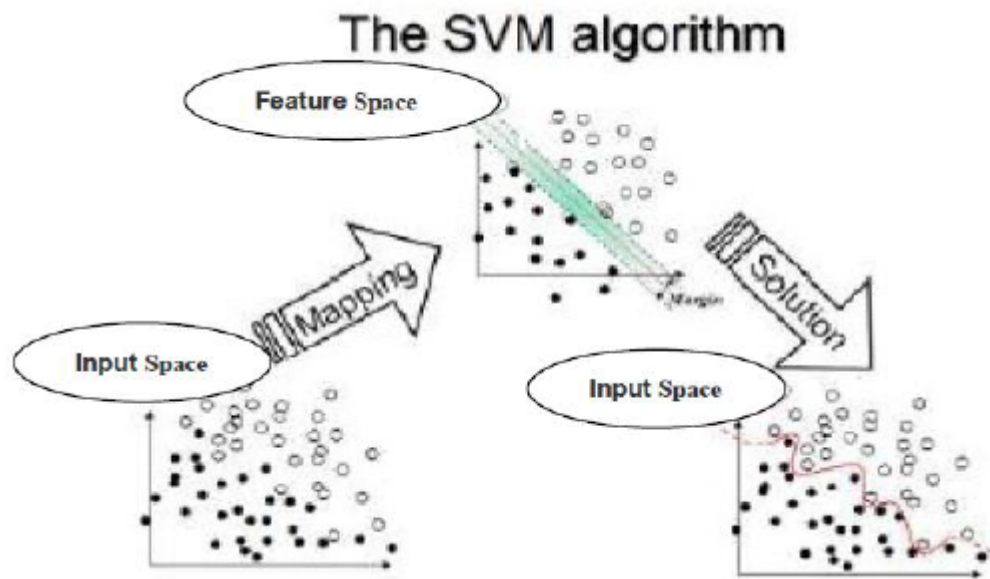
mende klasse als voorspelling. AdaBoost is een adaptief algoritme om een reeks classificeerders te boosten, in die zin dat de gewichten dynamisch worden bijgewerkt op basis van de fouten in eerdere leerresultaten.

### 2.5.3. Random Forest Classifier

Random forests zijn verzamelingen decision trees die elk getraind zijn op een willekeurig gekozen deelverzameling van de beschikbare data. Deze decision trees worden willekeurig door de trainingsvoorbeelden die aan elke boom worden gegeven, maar ook door een willekeurige subset van tests die beschikbaar zijn voor optimalisatie op elke node (**Fanelli2012**). Wanneer we een input vector aan het model geven, gaat het model bewegen door de gepaste boom (**Chen2011**). De uiteindelijke classificatie is gebaseerd op een majority voting over alle bomen. Uit het onderzoek van **Wang2010**<sup><empty citation></sup> blijkt dat Random Forest ook geschikt is voor feature selectie.

Een voorbeeld voor het gebruik van Random Forest in gezichtsanalyse is figuur 2.6. Hierbij wordt een boom getoond voor het schatten van de houding van het hoofd.

Het onderzoek van **Mady2018**<sup><empty citation></sup> gebruikt random forest classifier voor de herkenning en detectie van gezichten. De classifier wordt toegepast op het gezicht om dit te herkennen uit de databank. Hier wordt aangehaald wanneer we het aantal bomen in de random forest verhogen, zal ook de nauwkeurigheid verhogen, maar is het ook veel tijdrovender. De random forest is dus een trade-off tussen accuracy en tijd. Het toevoegen van Local Binary Pattern (LBP) en Histogram of Oriented Gradient(HOG) verhoogde de algemene accuracy van de classifier.

**Figuur (2.7)**

Support Vector Machine voor niet-lineair scheidbare data

#### 2.5.4. Support Vector Machines

Een Support Vector Machine (SVM) is een ML-techniek die gebruikt wordt voor patroonclassificatie en regressieanalyse (**Chen2011**). Het is gebaseerd op het zoeken naar een lineaire grens tussen twee klassen van patronen. Het SVM mechanisme gaat op zoek naar de optimale classifier die de data in 2 verschillende klassen opdeelt (**Rustam2018**). De optimale classifier wordt de hyperplane genoemd. Dit kunnen we vinden door de marge te maximaliseren en zal zorgen voor zo min mogelijk misclassificatie.

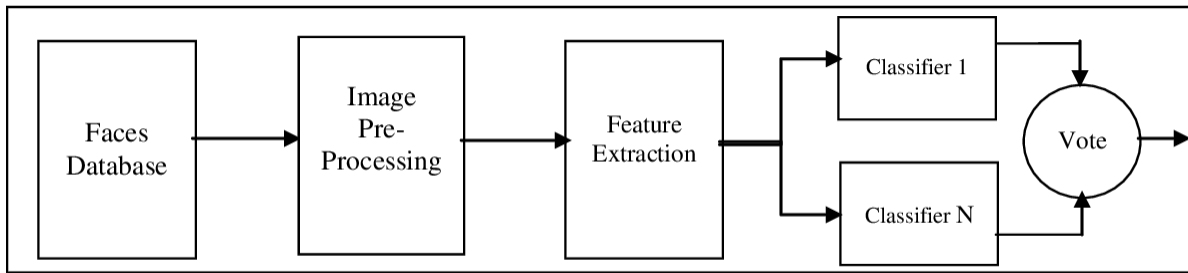
Voor niet-lineair scheidbare data mappen we alle punten naar een feature ruimte door gebruik te maken van een kernel (**Shah2012**). Na het splitsen, kunnen we de punten terug mappen naar de input ruimte met een kromme hyperplane, zoals op figuur 2.7.

Een SVM is zeer effectief wanneer een hoog aantal dimensionele ruimtes zijn. Ze zijn ook veelzijdig, aangezien we meerdere kernels kunnen gebruiken. De keuze van kernel is afhankelijk van de requirements van het model.

In het onderzoek van **Rustam2018** wordt SVM gebruikt voor de classificatie van leeftijdsgroepen en bereikt de maximale accuracy van 100 %.

#### 2.5.5. Ensemble

Een set van 30 Lineaire Discriminant Analyse (LDA) dienen als basis classifiers en dragen bij tot het ensemble in het model van **Khan2017**. Het ensemble gebruikt de gecombineerde classificatiecapaciteiten van de base classi-

**Figuur (2.8)**

Robuust Gezichtsherkenning systeem(Khan2017).

ficiers, zo verbeteren we de algehele prestaties. De volledige pipeline voor de gezichtsanalyse vinden we in figuur 2.8.

### 2.5.6. Multi attribution model

In het onderzoek van (Gupta2022) wordt een multi-attribution model uitgewerkt. Dit voorspelt leeftijd en geslacht door middel van slechts 1 model. Het multi-attribution model voorspelt dus beide: leeftijd en geslacht. Het model van (Guo2014), dat gebruik maakt van de Biologically Inspired Feature (BIF), gaf betere resultaten in vergelijking met individuele feature modellen. Een CNN is het meest gebruikte model voor multi-attribution, maar dit valt buiten de scope van deze bachelorproef.

## 2.6. Dataset

Uit onderzoek van Karkkainen2021<empty citation> bleek dat bestaande, publieke datasets sterk bevoordeeld zijn naar blanke mensen. Modellen die getraind worden op enkel 1 publieke dataset vertonen dan ook slechte en inconsistente resultaten. Kärkkäinen gebruikt in zijn onderzoek een zelf samengestelde dataset, met 108501 afbeeldingen die verdeeld zijn over elk ras. De resultaten van het onderzoek op de dataset vertonen een betere accuraatheid over verschillende rassen en leeftijdsgroepen.

Ook in het onderzoek van Buolamwini2018<empty citation> wordt een eigen dataset aangemaakt. Hier wordt niet enkel gekeken naar ras, maar bijvoorbeeld ook naar dikte en hoeveelheid haar. Het model moet zodoende dezelfde accuraatheid geven over verschillende leeftijdsgroepen, geslachten, rassen, et cetera. Uit het onderzoek bleek ook dat resultaten bij modellen die gebruik maken van Microsoft's Face Detect of Face++ een slechtere precisiescore vertonen, mogelijks door bias in trainingsdata.

## 2.7. Voorspellen van geslacht en leeftijd

Het eerste baanbrekende onderzoek op het vlak van classificatie op basis van gezichtsafbeeldingen dateert uit 1994. **Kwon1994<empty citation>** voorspelde slechts 3 leeftijdsgroepen: baby's, tieners en volwassenen. Het onderzoek focuste zich op het vinden van de primaire features in het gezicht, zoals ogen, mond en neus, en combineerde dit met secundaire features, zoals rimpels. Het onderzoek toonde hiermee aan dat de gezichtsfeatures een grote rol spelen in het classificeren van leeftijd.

Menselijke gezichten zijn onderhevig aan veroudering en groei (**Gupta2022**). Deze verandering in uiterlijk kan verschillen van persoon en is het gevolg van verschillende factoren, zoals gezondheid, levensstijl, ras, roken, et cetera.

Vrouwen en mannen verouderen anders, omdat ze een verschillend verouderingspatroon in het gezicht hebben. Dit is afkomstig van het gebruik van make-up, haarstijl, accessoires van vrouwen of snor en baard bij mannen. Vrouwelijke gezichten ogen dan ook vaak jonger dan mannelijke gezichten.

Het voorspellen van leeftijd kan ingedeeld worden in 2 categorieën: leeftijdsgroep classificatie, waarbij we de leeftijd opdelen in ranges (bijvoorbeeld 10-18 jaar) of als regressieprobleem, waarbij we een exact getal gaan voorspellen (**Gupta2022**). In de proof-of-concept van deze bachelorproef zal een leeftijdsgroep worden voorspeld.

## 2.8. Performantiescores

Er kunnen verschillende evaluaties uitgevoerd worden om de performantie van het model te testen (**Sanil2023**). Er kan hiervoor een confusion matrix worden opgesteld, zoals weergegeven in figuur 2.9. Deze bevat de True Positives (TP), True Negatives (TN), False Positives (FP) en False Negatives (FN), waarbij TP en TN de correcte voorspellingen weergeven en FP en FN de incorrecte voorspellingen. De confusion matrix werkt voor binaire voorspellingen, ideaal voor de voorspelling van geslacht. Het onderzoek van **Sanil2023<empty citation>** berekent op basis van de opgestelde matrix 5 scores.

1. De accuracy geeft weer hoe goed het model voorspellingen maakt. Dit wordt voorgesteld door het aantal correcte voorspellingen (**Narkhede2018**).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

2. De precision stelt de hoeveelheid correcte positieve voorspellingen voor.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

**Figuur (2.9)**

Confusion matrix(Narkhede2018).

3. De recall, ook wel sensitivity genoemd, stelt de kans op een positieve klasse voor.

$$\text{Recall} = \frac{TP}{TP+FN}$$

4. De specificity, stelt de kans op een negatieve klasse voor.

$$\text{Specificity} = \frac{TN}{TN+FP}$$

5. De F1-score is het gemiddelde van de precision en recall scores, waarbij 1.0 de maximale score is.

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

# 3

## Methodologie

Etiam pede massa, dapibus vitae, rhoncus in, placerat posuere, odio. Vestibulum luctus commodo lacus. Morbi lacus dui, tempor sed, euismod eget, condimentum at, tortor. Phasellus aliquet odio ac lacus tempor faucibus. Praesent sed sem. Praesent iaculis. Cras rhoncus tellus sed justo ullamcorper sagittis. Donec quis orci. Sed ut tortor quis tellus euismod tincidunt. Suspendisse congue nisl eu elit. Aliquam tortor diam, tempus id, tristique eget, sodales vel, nulla. Praesent tellus mi, condimentum sed, viverra at, consectetur quis, lectus. In auctor vehicula orci. Sed pede sapien, euismod in, suscipit in, pharetra placerat, metus. Vivamus commodo dui non odio. Donec et felis.

Etiam suscipit aliquam arcu. Aliquam sit amet est ac purus bibendum congue. Sed in eros. Morbi non orci. Pellentesque mattis lacinia elit. Fusce molestie velit in ligula. Nullam et orci vitae nibh vulputate auctor. Aliquam eget purus. Nulla auctor wisi sed ipsum. Morbi porttitor tellus ac enim. Fusce ornare. Proin ipsum enim, tincidunt in, ornare venenatis, molestie a, augue. Donec vel pede in lacus sagittis porta. Sed hendrerit ipsum quis nisl. Suspendisse quis massa ac nibh pretium cursus. Sed sodales. Nam eu neque quis pede dignissim ornare. Maecenas eu purus ac urna tincidunt congue.

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

Maecenas non massa. Vestibulum pharetra nulla at lorem. Duis quis quam id lacus dapibus interdum. Nulla lorem. Donec ut ante quis dolor bibendum condimentum. Etiam egestas tortor vitae lacus. Praesent cursus. Mauris bibendum pede at elit. Morbi et felis a lectus interdum facilisis. Sed suscipit gravida turpis. Nulla at

lectus. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Praesent nonummy luctus nibh. Proin turpis nunc, congue eu, egestas ut, fringilla at, tellus. In hac habitasse platea dictumst.

Vivamus eu tellus sed tellus consequat suscipit. Nam orci orci, malesuada id, gravida nec, ultricies vitae, erat. Donec risus turpis, luctus sit amet, interdum quis, porta sed, ipsum. Suspendisse condimentum, tortor at egestas posuere, neque metus tempor orci, et tincidunt urna nunc a purus. Sed facilisis blandit tellus. Nunc risus sem, suscipit nec, eleifend quis, cursus quis, libero. Curabitur et dolor. Sed vitae sem. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Maecenas ante. Duis ullamcorper enim. Donec tristique enim eu leo. Nullam molestie elit eu dolor. Nullam bibendum, turpis vitae tristique gravida, quam sapien tempor lectus, quis pretium tellus purus ac quam. Nulla facilisi.

# 4

## Conclusie

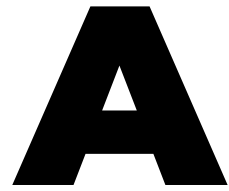
Curabitur nunc magna, posuere eget, venenatis eu, vehicula ac, velit. Aenean ornare, massa a accumsan pulvinar, quam lorem laoreet purus, eu sodales magna risus molestie lorem. Nunc erat velit, hendrerit quis, malesuada ut, aliquam vitae, wisi. Sed posuere. Suspendisse ipsum arcu, scelerisque nec, aliquam eu, molestie tincidunt, justo. Phasellus iaculis. Sed posuere lorem non ipsum. Pellentesque dapibus. Suspendisse quam libero, laoreet a, tincidunt eget, consequat at, est. Nullam ut lectus non enim consequat facilisis. Mauris leo. Quisque pede ligula, auctor vel, pellentesque vel, posuere id, turpis. Cras ipsum sem, cursus et, facilisis ut, tempus euismod, quam. Suspendisse tristique dolor eu orci. Mauris mattis. Aenean semper. Vivamus tortor magna, facilisis id, varius mattis, hendrerit in, justo. Integer purus.

Vivamus adipiscing. Curabitur imperdiet tempus turpis. Vivamus sapien dolor, congue venenatis, euismod eget, porta rhoncus, magna. Proin condimentum pretium enim. Fusce fringilla, libero et venenatis facilisis, eros enim cursus arcu, vitae facilisis odio augue vitae orci. Aliquam varius nibh ut odio. Sed condimentum condimentum nunc. Pellentesque eget massa. Pellentesque quis mauris. Donec ut ligula ac pede pulvinar lobortis. Pellentesque euismod. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent elit. Ut laoreet ornare est. Phasellus gravida vulputate nulla. Donec sit amet arcu ut sem tempor malesuada. Praesent hendrerit augue in urna. Proin enim ante, ornare vel, consequat ut, blandit in, justo. Donec felis elit, dignissim sed, sagittis ut, ullamcorper a, nulla. Aenean pharetra vulputate odio.

Quisque enim. Proin velit neque, tristique eu, eleifend eget, vestibulum nec, lacus. Vivamus odio. Duis odio urna, vehicula in, elementum aliquam, aliquet laoreet, tellus. Sed velit. Sed vel mi ac elit aliquet interdum. Etiam sapien neque, convallis et, aliquet vel, auctor non, arcu. Aliquam suscipit aliquam lectus. Proin tincidunt magna sed wisi. Integer blandit lacus ut lorem. Sed luctus justo sed enim.



Morbi malesuada hendrerit dui. Nunc mauris leo, dapibus sit amet, vestibulum et, commodo id, est. Pellentesque purus. Pellentesque tristique, nunc ac pulvinar adipiscing, justo eros consequat lectus, sit amet posuere lectus neque vel augue. Cras consectetur libero ac eros. Ut eget massa. Fusce sit amet enim eleifend sem dictum auctor. In eget risus luctus wisi convallis pulvinar. Vivamus sapien risus, tempor in, viverra in, aliquet pellentesque, eros. Aliquam euismod libero a sem. Nunc velit augue, scelerisque dignissim, lobortis et, aliquam in, risus. In eu eros. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Curabitur vulputate elit viverra augue. Mauris fringilla, tortor sit amet malesuada mollis, sapien mi dapibus odio, ac imperdiet ligula enim eget nisl. Quisque vitae pede a pede aliquet suscipit. Phasellus tellus pede, viverra vestibulum, gravida id, laoreet in, justo. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Integer commodo luctus lectus. Mauris justo. Duis varius eros. Sed quam. Cras lacus eros, rutrum eget, varius quis, convallis iaculis, velit. Mauris imperdiet, metus at tristique venenatis, purus neque pellentesque mauris, a ultrices elit lacus nec tortor. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent malesuada. Nam lacus lectus, auctor sit amet, malesuada vel, elementum eget, metus. Duis neque pede, facilisis eget, egestas elementum, nonummy id, neque.



## Onderzoeksvoorstel

Het onderwerp van deze bachelorproef is gebaseerd op een onderzoeksvoorstel dat vooraf werd beoordeeld door de promotor. Dat voorstel is opgenomen in deze bijlage.

### Samenvatting

Gezichtsanalyse heeft de laatste jaren veel aandacht gekregen vanwege de brede toepassingen op verschillende gebieden, zoals gezondheidszorg, beveiliging en marketing. De focus van de bachelorproef ligt op het ontwikkelen en implementeren van een robuust systeem voor het schatten van leeftijd en geslacht op basis van gezichtsfoto's, met behulp van machine learning technieken. Dit onderzoek is van bijzonder belang voor het verbeteren van de beoordeling van geestelijke gezondheidszorg door camera-gebaseerde gezondheidsmetingen. Deze inspanningen dragen bij aan de doelstellingen van IntelliProve, een platform dat online gezondheidsoplossingen biedt. Er wordt onderzocht welke machine learning technieken de beste resultaten tonen en uit welke elementen de volledige pipeline zal bestaan. In de eerste fase van het onderzoek wordt een literatuurstudie uitgevoerd om uit bestaand onderzoek de gebruikte modellen te verkennen en hun sterktes, limitaties en performantie te achterhalen. In de proof-of-concept gebeurt eerst de pre-processing, feature extractie en gezichtsdetectie. Daarna worden de modellen uit de literatuurstudie getest op bestaande datasets en wordt een prestatie-evaluatie opgesteld op basis van metrieken zoals accuracy, precision, recall, F1-score en ROC curve. Het verwachte resultaat is een robuuste machine learning pipeline waarmee IntelliProve de leeftijd en het geslacht kan voorspellen op basis van een gezichtsafbeelding.

## A.1. Introductie

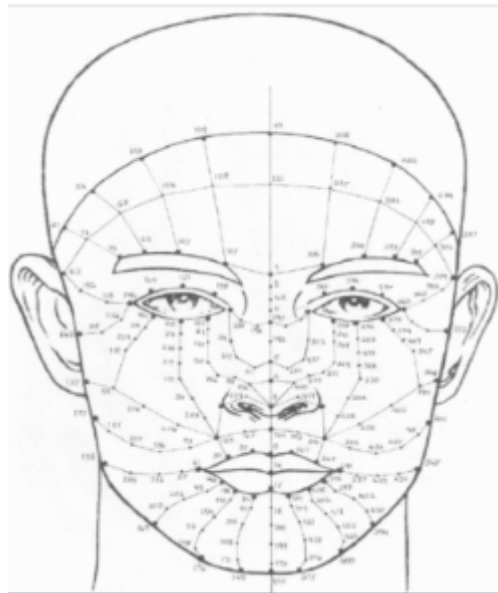
De onderzoeksvraag werd aangeboden door het bedrijf IntelliProve. IntelliProve biedt online gezondheidsoplossingen, een software die in staat is om binnen enkele seconden nauwkeurig gezondheidsparameters te bepalen, gebaseerd op een optische meting van het gezicht. Het doel van de bachelorproef is het ontwikkelen en implementeren van een robuust systeem voor het schatten van de leeftijd en het geslacht van personen op basis van gezichtsfoto's, met behulp van machine learning-technieken. Dit project is van bijzonder belang voor het verbeteren van de beoordeling van de geestelijke gezondheidszorg door middel van camera-gebaseerde gezondheidsmetingen. Het onderzoek beoogt bij te dragen aan de vooruitgang op dit gebied door gebruik te maken van geavanceerde algoritmen om leeftijd en geslacht nauwkeurig te voorspellen aan de hand van gezichtsbeelden. De literatuurstudie biedt een inzicht in facial analysis, de bestaande machine learning modellen en hun functionaliteiten. De proof-of-concept zal bestaan uit het ontwikkelen van een machine learning pipeline die in staat is om leeftijd en geslacht te voorspellen op basis van bestaande datasets. De pipeline omvat verschillende image preprocessing technieken om de dataset voor te bereiden op de modeltraining. Om betrouwbaarheid en accuracy te garanderen, worden de modellen verfijnd en geoptimaliseerd om de hoogst mogelijke nauwkeurigheid te bereiken bij het schatten van leeftijd en geslacht.

## A.2. Stand van zaken

Gezichtsanalyse bestaat uit het definiëren van menselijke gezichten in real time, met behulp van computeralgoritmen en machine learning technieken. Voor mensen en computersystemen bevat een gezichtsbeeld details zoals leeftijd, geslacht, stemming, afkomst, et cetera. Gezichtsanalyse omvat dan ook het lokaliseren en meten van gezichtskenmerken in een afbeelding. Vervolgens wordt het gezichtsbeeld geanalyseerd voor het extraheren van kenmerken (**Sanil2023**). Gezichtsanalyse speelt dan ook een grote rol in real world applicaties, zoals animatie, identiteitsverificatie, medische diagnose, et cetera. Ondanks het bestaande onderzoekswerk over dit onderwerp, is gezichtsanalyse nog steeds een lastige taak vanwege verschillende factoren zoals veranderingen in hoek, gezichtsuitdrukkingen en achtergrond (**Siddiqi2022**).

Volgens onderzoek van **Basystiuk2023<empty citation>** is beeldherkenning een simpele procedure die slechts uit 3 stappen bestaat.

1. Preprocessing: We voegen filters toe aan de afbeelding om het geschikter te maken voor herkenning.
2. Feature Extractie: We identificeren belangrijke data en behouden deze om mee verder te werken. Dit wordt verder besproken in [A.2.1](#)



**Figuur (A.1)**

De belangrijkste anthropometrische landmarks (Sanil2023).

3. Classificatie: het analyseren en identificeren van de data na de feature extractie.

### A.2.1. Gezichtsdetectie

Gezichtsdetectie is de eerste stap in het bepalen van de gezichtsfeatures. Deze features zijn de interessante delen van het gezicht, bijvoorbeeld ogen, neus en mond en worden ook wel eens landmarks genoemd (Coppens2018). In dit onderzoek raadt de auteur 3 gezichtsdetectieservices aan: Microsoft Face API, Face++ en Kairos. Deze zijn alle 3 ook geschikt om geslacht en leeftijd te voorspellen. Er wordt met 3 services tegelijk gewerkt om elkaars beperkingen op te vangen. In het onderzoek van Sanil2023 berekenen ze de landmarks op 2 verschillende methoden: Euclidische afstand en Geodetische afstand.

Sanil2023 stelt 468 landmarks visueel voor, in plaats van 68, om de accuracy te verbeteren. Dit door gebruik te maken van de Mediapipe library (Zubair2021). Op figuur A.1 staan de belangrijkste antropometrische landmarks aangeduid. Het is noodzakelijk om de belangrijkste features te identificeren die kunnen leiden tot het correct voorspellen van de leeftijd of het geslacht. Om de leeftijd te voorspellen kunnen we bijvoorbeeld de rimpels op de gezichtsfoto analyseren (Kwon1994).

### Normalisatie

In onderzoek van (Chen2011) voor het voorspellen van leeftijd op basis van afbeeldingen werden de gezichtsafbeeldingen genormaliseerd voor de feature extraction plaatsvond. De geometrie van de afbeeldingen kan worden genormaliseerd op basis van de gedetecteerde features, zoals oog of mond coördinatie. De auteur

**Figuur (A.2)**

Gezichtsafbeeldingen voor normalisatie(Chen2011).

maakt gebruik van een mask om de pixels die zich niet in de ovaal van de typische gezichtsregio bevinden te verwijderen. Dit gaat dan bijvoorbeeld over haar en hemdskragen. Zo blijven enkel de belangrijkste features van het gezicht over. Figuur A.2 geeft een voorbeeld van een dataset met gezichtsfoto's. In figuur A.3 worden deze afbeeldingen genormaliseerd, waardoor er minder variaties overblijven in de afbeeldingen en we bijvoorbeeld al geen achtergrond overhouden.

### A.2.2. Bestaande Machine-Learning technieken

Uit onderzoek naar gezichtsdetectie van Sanil2023<empty citation> bleek dat Extreme gradient boosting de beste resultaten gaf, met een accuracy van 78%, gevolgd door Adaptive Boosting (77%) en Random Forest (75%). Onderzoek van (Khan2017) gebruikt dan weer een Ensemble systeem.

#### Extreme Gradient Boost

Extreme Gradient Boost, ofwel XGBoost, is een supervised learning model, waar trainingdata, met meerdere features,  $x_i$  wordt gebruikt om target waarde  $y_i$  te voorspellen (XGBoost2023). Uit onderzoek van (Chen2023), waarbij vermoeidheid werd voorspeld op basis van gezichtsafbeeldingen met XGBoost, wordt een boom aangemaakt op basis van de geleverde features. Dit model gaf een hoge accuracy en is minder gevoelig voor noise.

#### Adaptive Boosting

De Adaptive Boost, of AdaBoost, is typisch een classificatie tussen twee klassen. Om het herkenningsprobleem met meerdere klassen op te lossen, kan een majo-

**Figuur (A.3)**

Gezichtsafbeeldingen na het toepassen van normalisatie(Chen2011).

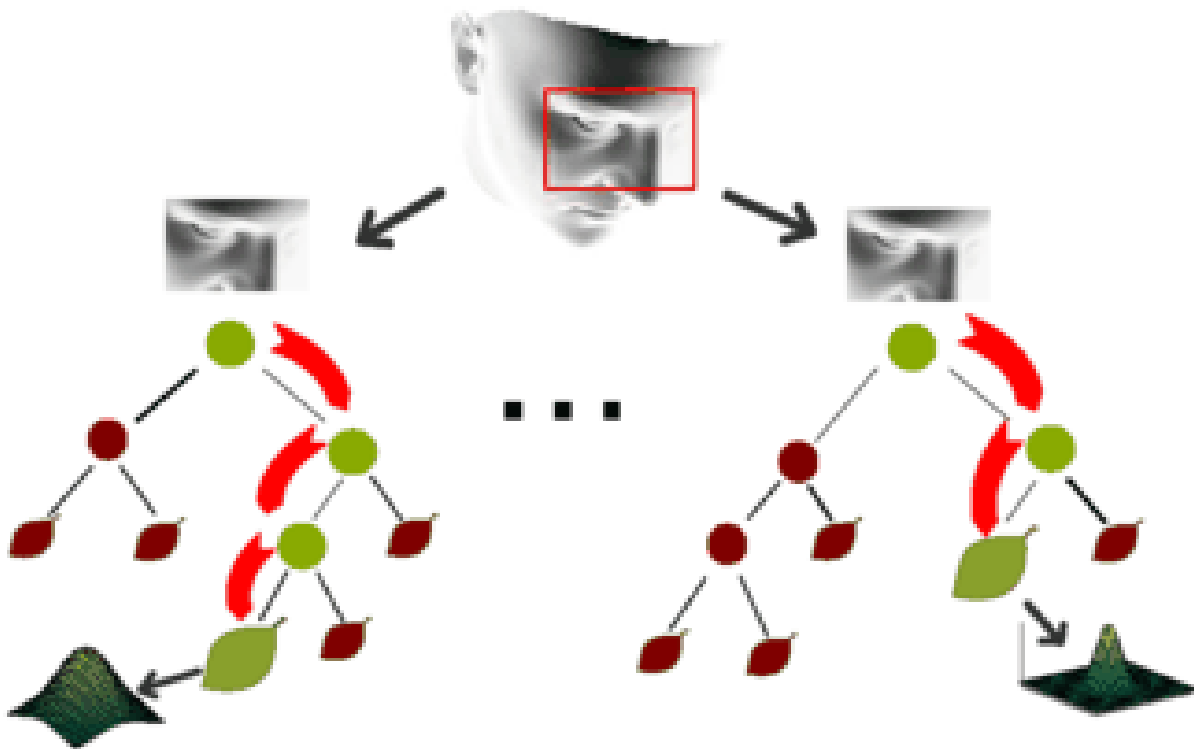
rity voting (MV) strategie worden gebruikt om alle paarsgewijze classificatieresultaten te combineren. Hierbij kiezen we de meest voorkomende klasse als voorspelling. AdaBoost is een adaptief algoritme om een reeks classificeerders te boosten, in die zin dat de gewichten dynamisch worden bijgewerkt op basis van de fouten in eerdere leerresultaten (Guo2001).

### Random Forest Classifier

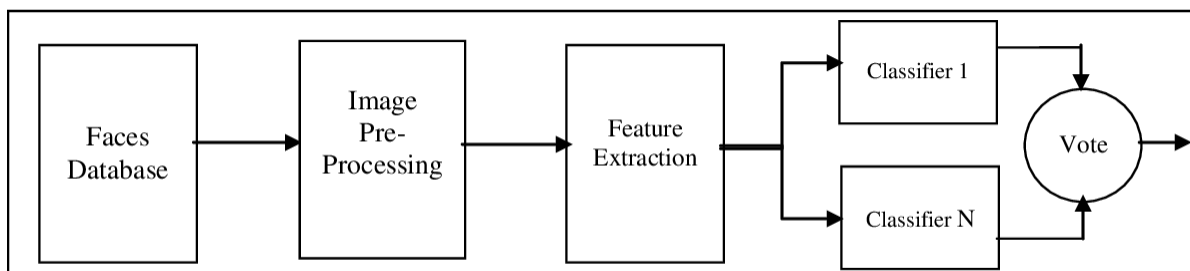
Random forests zijn verzamelingen decision trees die elk getraind zijn op een willekeurig gekozen deelverzameling van de beschikbare data. Deze decision trees worden willekeurig door de trainingsvoorbeelden die aan elke boom worden gegeven, maar ook door een willekeurige subset van tests die beschikbaar zijn voor optimalisatie op elke node (Fanelli2012). Een voorbeeld voor het gebruik van Random Forest in gezichtsanalyse is figuur A.4. Hierbij wordt een boom getoond voor het schatten van de houding van het hoofd.

### Ensemble

Een set van 30 Lineaire Discriminant Analyse (LDA) dienen als basis classifiers en dragen bij tot het ensemble in het model van Khan2017<empty citation>. Het ensemble gebruikt de gecombineerde classificatiecapaciteiten van de base classifiers, zo verbeteren we de algehele prestaties. De volledige pipeline voor de gezichtsanalyse vinden we in figuur A.5.

**Figuur (A.4)**

Voorbeeld van een random forest voor het schatten van de houding van het hoofd(**Fanelli2012**).

**Figuur (A.5)**

Robuust Gezichtsherkenning systeem(**Khan2017**).

**Figuur (A.6)**

Methodologie voor het onderzoek.

### A.2.3. Bestaande uitdagingen

Ondanks het onderzoekwerk, heersen er nog vele uitdagingen op het vlak van gezichtsanalyse. Deze uitdagingen zijn het gevolg van verschillende factoren zoals gezichtsuitdrukkingen, ruis, belichting, et cetera. Om de nauwkeurigheid van de gezichtsherkenning te verbeteren, is het belangrijk om taken van de gezichtsanalyse met elkaar te correleren. Het is bijvoorbeeld zeer waarschijnlijk dat mannen een baard of snor kunnen hebben, maar vrouwen en kinderen niet (**Siddiqi2022**). Ook is het detecteren van de locatie waar de gezichten zich bevinden op de afbeelding een uitdaging. Dit wordt vaak meegenomen in de preprocessing stap van het analysesysteem (**Jiang2008**). De problemen situeren zich niet enkel op het vlak van de afbeelding, maar ook bij de machine learning modellen. Eén van de grootste problemen hierbij is overfitting. Dit gebeurt wanneer we het algoritme te veel trainen, waardoor het te hard lijkt op de trainingdata. Hierdoor zal het systeem falen wanneer we nieuwe data willen classificeren (**Coppens2018**).

## A.3. Methodologie

Figuur A.6 geeft een visuele representatie van de opgestelde methodologie.

### A.3.1. Requirements

In de eerste week wordt nagevraagd aan belanghebbenden van IntelliProve aan welke criteria de modellen moeten voldoen. Alle data (gezichtsfoto's) wordt verzameld. Er wordt onder andere nagegaan over welke functionaliteiten de modellen moeten beschikken en wat de verwachte prestatievereisten zijn. Als resultaat verwerven we een lijst van alle functionele en niet-functionele requirements, geordend volgens belang.

### A.3.2. Literatuurstudie

De literatuurstudie omvat een diepgaande verkenning van facial analysis technieken en machine learning modellen. Deze fase biedt inzicht in de verschillende methoden voor het extraheren van gezichtskenmerken en image preprocessing technieken, specifiek met betrekking tot het schatten van leeftijd en geslacht. Het doel is om kennis uit bestaand onderzoek te vergaren om effectieve methodologieën te identificeren in de huidige benaderingen van facial analysis. Het eindresultaat van deze fase, die 3 weken duurt, is een samenvatting van de belangrijkste bevindingen



uit de literatuurstudie, die als basis zal dienen voor de proof-of-concept.

### **A.3.3. Proof-of-concept**

Deze fase start met het verzamelen en analyseren van de datasets. Technieken zoals normalisatie en scaling van de features, feature-extractie en data augmentation worden gebruikt om de dataset voor te bereiden op modeltraining. Vervolgens worden verschillende machine learning algoritmen geselecteerd op basis van de bevindingen uit de literatuurstudie. De modellen worden getraind en geëvalueerd op basis van de vooropgestelde requirements. De pipeline wordt beoordeeld op betrouwbaarheid en precisie op basis van de opgestelde requirements. Metrieken zoals accuracy, precision, recall, F1-score en ROC curves worden gebruikt om de modellen te evalueren. Er worden cross-validatie technieken gebruikt om de robuustheid van de modellen en de generalisatie van nieuwe data te testen. Deze fase vereist dan ook veel tijd en zal 6 weken duren. Het resultaat van deze fase is een proof-of-concept die bestaat uit een machine learning pipeline, die stappen voor het preprocessen van gegevens en getrainde modellen integreert voor het schatten van leeftijd en geslacht op basis van gezichtsfoto's.

### **A.3.4. Conclusie**

In de conclusiefase worden de resultaten van de evaluatie grondig geanalyseerd. De prestaties van de ontwikkelde modellen worden beoordeeld, waarbij hun sterke punten en beperkingen worden benadrukt. Het belang van een nauwkeurige schatting van leeftijd en geslacht bij de beoordeling van de geestelijke gezondheid en de implicaties voor de gezondheidszorg worden besproken. Er worden aanbevelingen gedaan voor mogelijke verbeteringen of toekomstige onderzoeksrichtingen op basis van de bevindingen en beperkingen van het project.

### **A.3.5. Afwerken scriptie**

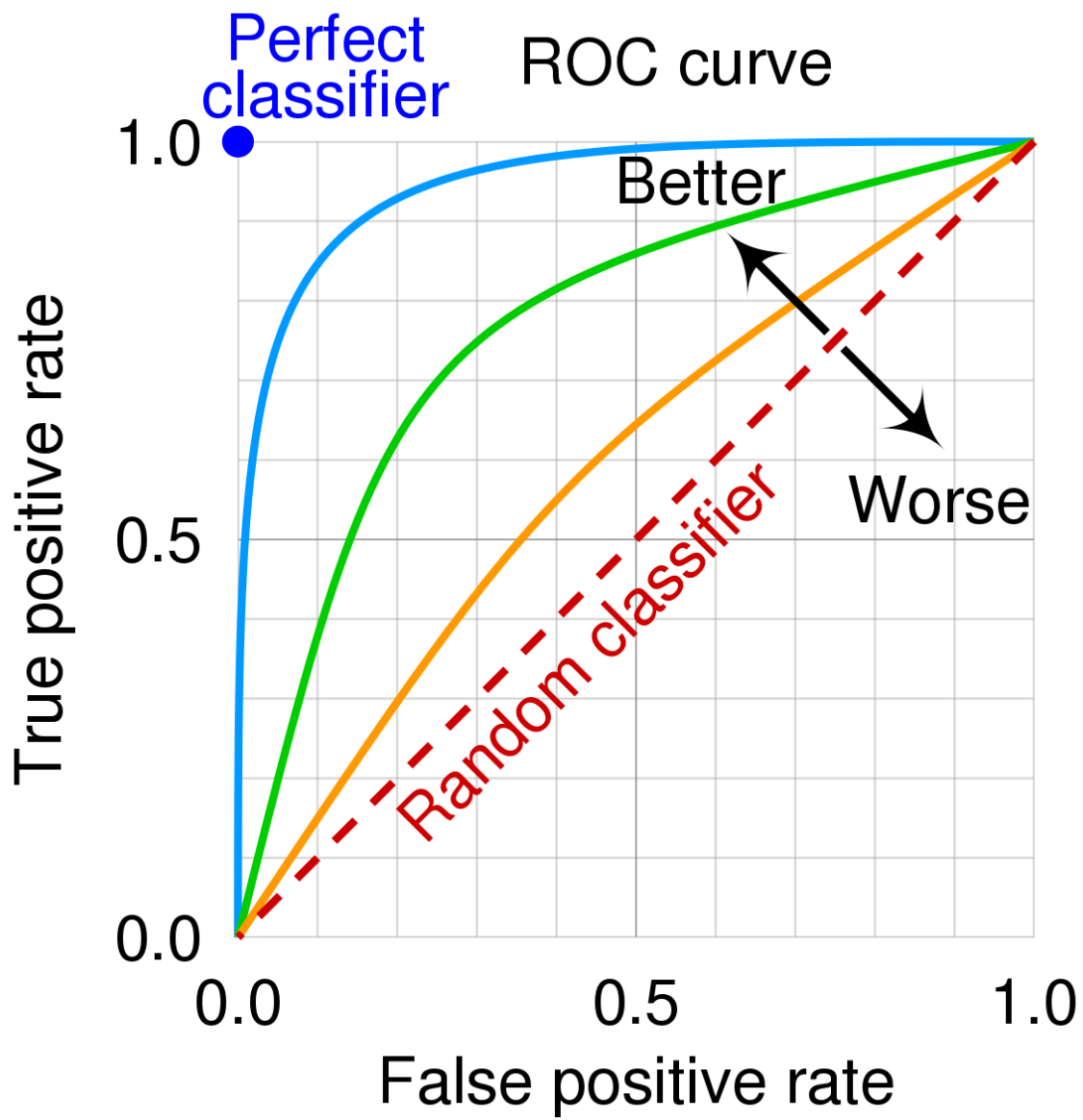
De laatste fase, die 2 weken duurt, omvat het afwerken van de bachelorproef. Dit is het eindresultaat van het geleverde onderzoek met een proof-of-concept die zal worden ingediend.

## **A.4. Verwacht resultaat, conclusie**

Als resultaat van het onderzoek wordt een proof-of-concept opgesteld die een robuuste machine learning pipeline implementeert. We streven naar een accuracy van 75%, wat al haalbaar bleek uit onderzoek van **Sanil2023<empty citation>**. Er wordt verwacht dat de modellen van Boosting, Random Forest en een ensemble de hoogste scores en accuracy opleveren. We geven naast de proof-of-concept ook een prestatie-evaluatie van de gebruikte modellen met uitgebreide analyse en vergelijking. De pipeline kan worden beoordeeld op betrouwbaarheid en precisie op

basis van de opgestelde requirements. Metrieken zoals accuracy, precision, recall, F1-score en ROC-curves worden gebruikt om de modellen te evalueren. Onderstaand een mock-up van de scores met op figuur A.7 een ROC-curve. Dit onderzoek is van bijzonder belang voor het verbeteren van de beoordeling van de geestelijke gezondheidszorg door middel van camera-gebaseerde gezondheidsmetingen en levert een bijdrage aan de vooruitgang op dit gebied door gebruik te maken van geavanceerde algoritmen om leeftijd en geslacht nauwkeurig te voorspellen aan de hand van gezichtsopnames. Het is een opstap naar verder onderzoek in gezichtsanalyse.

Model	Accuracy	Precision	Recall	F1-Score
Model A	0.72	0.50	0.80	0.82
Model B	0.52	0.25	0.90	0.39
Model C	0.61	0.70	0.40	0.51



**Figuur (A.7)**  
ROC curve(Gomede2023).