

---

# BACHELORARBEIT

---

Frau

**Jane Marie Wachter**

**Schwellwertanalyse bezüglich der Limitierung von MediaPipe Face Mesh anhand von informationsreduziertem Ausgangsmaterial**

Mittweida, 2022

# **BACHELORARBEIT**

---

## **Schwellwertanalyse bezüglich der Limitierung von MediaPipe Face Mesh anhand von informationsreduziertem Ausgangsmaterial**

Autorin:

**Frau**

**Jane Marie Wachter**

Studiengang:

**Allgemeine und Digitale Forensik**

Seminargruppe:

**FO19w2-B**

Erstprüfer:

**Prof. Dr. rer. nat. Dirk Labudde**

Zweitprüfer:

**M.Sc. Sven Becker**

Einreichung:

**Mittweida, 13.09.2022**

Verteidigung/Bewertung:

**Mittweida, 2022**

## **BACHELOR THESIS**

---

# **Threshold analysis regarding the limitation of MediaPipe Face Mesh based on information-reduced data**

author:

**Ms.**

**Jane Marie Wachter**

course of studies:

**General and Digital Forensic Science**

seminar group:

**FO19w2-B**

first examiner:

**Prof. Dr. rer. nat. Dirk Labudde**

second examiner:

**M.Sc. Sven Becker**

submission:

**Mittweida, 13.09.2022**

defence/evaluation:

**Mittweida, 2022**

## **Bibliografische Beschreibung:**

Wachter, Jane Marie:

Schwellwertanalyse bezüglich der Limitierung von MediaPipe Face Mesh anhand von informationsreduziertem Ausgangsmaterial. - 2022. - IX, 83, A-21 S.

Mittweida, Hochschule Mittweida, Fakultät Angewandte Computer- und Biowissenschaften, Bachelorarbeit, 2022

## **Referat:**

Im Rahmen dieser Bachelorarbeit sollen die Schwellwerte der Limitierung von MediaPipe Face Mesh in Bezug auf informationsreduziertes Ausgangsmaterial analysiert werden.

# Inhalt

<b>Inhalt</b>	<b>I</b>
<b>Abbildungsverzeichnis</b>	<b>III</b>
<b>Tabellenverzeichnis</b>	<b>V</b>
<b>Abkürzungsverzeichnis</b>	<b>IX</b>
<b>1 Übersicht</b>	<b>1</b>
1.1 <i>Motivation</i>	1
1.2 <i>Zielsetzung</i>	1
1.3 <i>Kapitelübersicht</i>	2
<b>2 Grundlagen und Stand der Technik</b>	<b>3</b>
2.1 <i>Begriffsdefinition Maschinelles Lernen</i>	3
2.2 <i>Begriffsdefinition Face Detection</i>	3
2.3 <i>MediaPipe Face Mesh</i>	4
2.4 <i>Verwandte Literatur</i>	5
<b>3 Auswahl des Ausgangsmaterials</b>	<b>7</b>
3.1 <i>Kategorie Frontale Bilder</i>	7
3.2 <i>Kategorie Bilder mit Brille</i>	7
3.3 <i>Kategorie Bilder mit Sonnenbrille</i>	8
3.4 <i>Kategorie Laterale Bilder</i>	8
3.5 <i>Kategorie Bilder mit teilweiser Verdeckung</i>	8
<b>4 Methoden</b>	<b>9</b>
4.1 <i>Schema des Skripts zu MediaPipe Face Mesh</i>	10
4.2 <i>Methode zur Verpixelung des Ausgangsmaterials</i>	11
4.3 <i>Methode zur Veränderung der Helligkeit des Ausgangsmaterials</i>	13
4.4 <i>Methode zur Anwendung Gauß'scher Unschärfe auf das Ausgangsmaterial</i>	15
4.5 <i>Methode zur Anwendung von Interpolation auf das Ausgangsmaterial</i>	17
<b>5 Ergebnisse</b>	<b>19</b>
5.1 <i>Ergebnisse der Gesichtserkennung bei Verpixelung des Ausgangsmaterials in Bezug auf die verschiedenen Bildkategorien</i>	19

5.2	<i>Ergebnisse der Gesichtserkennung bei veränderter Helligkeit des Ausgangsmaterials in Bezug auf die verschiedenen Bildkategorien .....</i>	29
5.3	<i>Ergebnisse der Gesichtserkennung bei Anwendung Gauß'scher Unschärfe auf das Ausgangsmaterial in Bezug auf die verschiedenen Bildkategorien .....</i>	41
5.4	<i>Ergebnisse der Gesichtserkennung bei Anwendung von Interpolation auf das Ausgangsmaterial in Bezug auf die verschiedenen Bildkategorien.....</i>	54
<b>6</b>	<b>Diskussion der erzielten Ergebnisse .....</b>	<b>61</b>
6.1	<i>Auswertung der Ergebnisse bei Verpixelung.....</i>	61
6.2	<i>Auswertung der Ergebnisse bei veränderter Helligkeit .....</i>	66
6.3	<i>Auswertung der Ergebnisse bei Anwendung Gauß'scher Unschärfe.....</i>	71
6.4	<i>Auswertung der Ergebnisse bei Anwendung von Interpolation.....</i>	77
<b>7</b>	<b>Ergebnisse und Ausblick .....</b>	<b>80</b>
7.1	<i>Ergebnisse und Problemstellungen.....</i>	80
7.2	<i>Bewertung der Arbeit .....</i>	81
7.3	<i>Ausblick .....</i>	82
	<b>Quellenverzeichnis .....</b>	<b>84</b>
	<b>Anlagen .....</b>	<b>89</b>
	<b>Anlagen, Teil 1.....</b>	<b>A-1</b>
	<b>Anlagen, Teil 2.....</b>	<b>A-2</b>
	<b>Anlagen, Teil 3.....</b>	<b>A-3</b>
	<b>Anlagen, Teil 4.....</b>	<b>A-4</b>
	<b>Anlagen, Teil 5.....</b>	<b>A-5</b>
	<b>Anlagen, Teil 6.....</b>	<b>A-6</b>
	<b>Anlagen, Teil 7.....</b>	<b>A-7</b>
	<b>Anlagen, Teil 8.....</b>	<b>A-10</b>
	<b>Anlagen, Teil 9.....</b>	<b>A-16</b>
	<b>Anlagen, Teil 10.....</b>	<b>A-19</b>
	<b>Selbstständigkeitserklärung</b>	

# Abbildungsverzeichnis

Abbildung 1: Graphische Darstellung eines eindimensionalen Gauß'schen Kerns mit Gewichtung eines Pixels in Relation zur Entfernung zur Mitte des Gauß'schen Kerns [Do2022d] .....	16
Abbildung 2: Dargestelltes Gesicht auf dem in 64x64 Pixeln kein Gesichtsnetz erkannt werden konnte, entsprechend Bild 19_links_pixel_64_64.png.....	26
Abbildung 3: Darstellung der Ergebnisbilder mit bei (a) 512x512 Pixeln und (b) 80x80 Pixeln erkannten Gesichtsnetzen der Kategorie „links lateral“. Bild (a) entspricht dem Eingabebild 19_links.png, Bild (b) dem Eingabebild 19_links_pixel_80_80.png.....	26
Abbildung 4: Darstellung der Ergebnisbilder mit bei einer Pixelanzahl von (a) 25x25 (b) 30x30 und (c) 50x50 erkannten Gesichtsnetzen der Kategorie „frontal“. Bild (a) entspricht dem Eingabebild 35_frontal_pixel_25_25.png, Bild (b) dem Eingabebild 35_frontal_pixel_30_30.png und Bild (c) dem Eingabebild 35_frontal_pixel_50_50.png..	62
Abbildung 5: Darstellung der Ergebnisbilder mit erkannten Gesichtsnetzen der Kategorie „mit Brille“. Bild (a) entspricht dem Eingabebild 8_brille.png, Bild (b) dem Eingabebild 26_brille.png.....	63
Abbildung 6: Darstellung der in der Pixelanzahl veränderten Ergebnisbilder bei 16x16 Pixeln mit erkannten Gesichtsnetzen der Kategorie „mit Brille“. Bild (a) entspricht dem Eingabebild 8_brille_pixel_16_16.png, Bild (b) dem Eingabebild 26_brille_pixel_16_16.png .....	64
Abbildung 7: Darstellung der Ergebnisbilder mit bei (a) 1- und bei (b) 1,96-facher Helligkeit erkannten Gesichtsnetzen der Kategorie „links lateral“. Bild (a) entspricht dem Eingabebild 3_links.png, Bild (b) dem Eingabebild 3_links_heller_500.png .....	69
Abbildung 8: Darstellung der Ergebnisbilder mit bei jeweils 1,96-facher Helligkeit erkannten Gesichtsnetzen der Kategorie „teilweise verdeckt“. Bild (a) entspricht dem Eingabebild 10_verdeckt_heller_500.png, Bild (b) dem Eingabebild 4_verdeckt_heller_500.png und Bild (c) dem Eingabebild 43_verdeckt_heller_500.png.....	70
Abbildung 9: Darstellung der Ergebnisbilder mit bei jeweils 1-facher Helligkeit erkannten Gesichtsnetzen der Kategorie „teilweise verdeckt“. Bild (a) entspricht dem Eingabebild	

10\_verdeckt.png, Bild (b) dem Eingabebild 4\_verdeckt.png und Bild (c) dem Eingabebild 43\_verdeckt.png ..... 71

Abbildung 10: Darstellung der Ergebnisbilder mit erkannten Gesichtsnetzen der Kategorie „frontal“. Bild (a) entspricht dem Eingabebild 22\_frontal.png, Bild (b) dem Eingabebild 49\_frontal.png..... 73

Abbildung 11: Darstellung der Ergebnisbilder mit bei Anwendung Gauß'scher Unschärfe mit Radius 30 erkannten Gesichtsnetzen der Kategorie „frontal“. Bild (a) entspricht dem Eingabebild 22\_frontal\_gaussian\_30.png, Bild (b) dem Eingabebild 49\_frontal\_gaussian\_30.png..... 73



# Tabellenverzeichnis

Tabelle 1: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	20
Tabelle 2: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	21
Tabelle 3: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	23
Tabelle 4: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	25
Tabelle 5: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	27
Tabelle 6: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	28
Tabelle 7: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	30
Tabelle 8: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	31
Tabelle 9: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	32

Tabelle 10: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	33
Tabelle 11: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	34
Tabelle 12: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	35
Tabelle 13: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	36
Tabelle 14: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	37
Tabelle 15: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	38
Tabelle 16: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	39
Tabelle 17: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	40
Tabelle 18: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	41
Tabelle 19: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	42

Tabelle 20: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	44
Tabelle 21: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	46
Tabelle 22: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	48
Tabelle 23: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	50
Tabelle 24: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	52
Tabelle 25: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf frontalen Bildern angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	54
Tabelle 26: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf Bildern mit Brille angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter.....	55
Tabelle 27: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf Bildern mit Sonnenbrille angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	56
Tabelle 28: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf links lateralen Bildern angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	57
Tabelle 29: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf rechts lateralen Bildern angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter .....	58

Tabelle 30: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf Bildern mit teilweiser Verdeckung angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter..... 59

# Abkürzungsverzeichnis

<b>FDDB</b>	Face Detection Data set and Benchmark
<b>FFHQ</b>	Flickr-Faces-HQ
<b>HoG SVM</b>	Histogram of oriented gradients Support Vector Machine
<b>OpenCV</b>	Open Source Computer Vision Library
<b>PIL</b>	Python Imaging Library
<b>RCNN</b>	Regional Convolutional Neural Network
<b>S<sup>3</sup>FD</b>	Single Shot Scale-invariant Face Detector

# 1 Übersicht

Im folgenden Kapitel werden die Motivation und die Aufgabenstellung der vorliegenden Bachelorarbeit dargelegt. Zugleich wird anhand eines kurzen Überblicks der Aufbau der Arbeit aufgezeigt.

## 1.1 Motivation

Heutzutage werden einige ermittlungsrelevante Orte in Deutschland videoüberwacht, so ist zum Beispiel die Anzahl an Videoüberwachungskameras pro 1000 Einwohner in Berlin von 2016 bis 2021 von 4,9 [DaKI2020] auf 6,3 [St2022] gestiegen. Dadurch kann Bild- und Videomaterial sichergestellt werden. Damit solches Material für die Ermittlungsarbeit verwendet werden kann, muss es eine ausreichende Qualität aufweisen. Dies ist allerdings häufig nicht der Fall.

Zur Identifizierung tatverdächtiger Person sind dabei Bilder, auf denen Gesichter abgebildet sind, besonders relevant. Hierfür muss zunächst geprüft werden, ob ein Gesicht auch auf einem Bild mit schlechter Qualität mittels geeigneter Algorithmen detektiert werden kann. Unter Zunahme von speziellen Algorithmen, die in der Lage sind, die erkannten Gesichter miteinander bzw. z. B. mit einer Datenbank abzugleichen, besteht anschließend die Möglichkeit die abgebildete Person zu identifizieren.

## 1.2 Zielsetzung

Im ersten Schritt der vorliegenden Arbeit soll getestet werden ab wann MediaPipe Face Mesh (Face Mesh) kein Gesicht an entsprechender Stelle detektieren kann. Dafür wird das Ausgangsmaterial einer Informationsreduktion durch Veränderung jeweils unterschiedlicher Faktoren unterzogen.

Das Ausgangsmaterial ist dafür in verschiedene Bildkategorien aufgeteilt, beispielsweise nach Position des Gesichts der abgebildeten Person oder ob z. B. eine Brille getragen wird. Zweck dieser Auswahl ist es zu überprüfen, ob die ausgewählten Faktoren der jeweiligen Bildkategorien die Ergebnisse der Gesichtserkennung beeinflussen.

Die Informationsreduktion erfolgt jeweils in mehreren Abstufungen, um so genau wie möglich die Limitierungen des ausgewählten Gesichtserkennungs-Algorithmus aufzuzeigen.

Ziel des genannten Vorgangs ist es eine Grundlage für weitere Schritte zu schaffen. Als ein solcher Schritt gilt z. B. die Personenidentifikation anhand von auf Bildern mit geringer Auflösung oder anderen informationsreduzierenden Bedingungen abgebildeten Gesichtern. Der Fokus der Untersuchung liegt darauf, festzustellen, wie zuverlässig die Ergebnisse der zugrundeliegenden Gesichtserkennung unter den entsprechenden Bedingungen sind, um beispielsweise in der forensischen Fallarbeit eine sichere Gesichtsdetektion und ggf. im weiteren Verlauf eine Personenidentifikation zu gewährleisten.

## 1.3 Kapitelübersicht

Die vorliegende Bachelorarbeit ist in sieben Kapitel unterteilt.

Nach der allgemeinen Einleitung wird in **Kapitel 2** ein Überblick über Begriffsdefinitionen, allgemeine verwendete Methoden und auch über den derzeitigen Stand der Technik gegeben.

Diesem Überblick folgt in **Kapitel 3** die Darlegung und Erläuterung der Auswahl des Ausgangsmaterials.

Anschließend werden in **Kapitel 4** die zur Informationsreduktion verwendeten Methoden dargestellt. Dabei werden sowohl die Auswahl der Kategorien als auch die Implementierung in Python erläutert.

Weiter werden in **Kapitel 5** die erreichten Ergebnisse der Gesichtserkennung auf Grundlage des informationsreduzierten Materials dargelegt. Festgestellt werden soll dabei, wie weit das Ausgangsmaterial verändert werden kann, sodass noch eine Detektion des Gesichts möglich ist.

Im Anschluss daran werden die dargestellten Ergebnisse in **Kapitel 6** diskutiert.

Zum Schluss werden in **Kapitel 7** die Resultate der vorliegenden Bachelorarbeit zusammengefasst und diskutiert. Zusätzlich erfolgt ein Ausblick auf mögliche Verbesserungen und Weiterentwicklungen.

## 2 Grundlagen und Stand der Technik

Für die Anwendung und anschließende Analyse eines Algorithmus zur Gesichtserkennung ist es ebenso essenziell die wichtigsten Begriffsdefinitionen zu kennen wie ein Grundverständnis für darin verwendete Methoden zu haben. Diese werden im folgenden Kapitel dargestellt und erklärt. Außerdem werden die Ergebnisse thematisch ähnlicher Arbeiten dargelegt, um später einen Vergleich ziehen zu können.

### 2.1 Begriffsdefinition Maschinelles Lernen

Der Ausdruck „maschinelles Lernen“ wird oftmals synonym zum Begriff „künstliche Intelligenz“ verwendet. Grundsätzlich lässt sich maschinelles Lernen als das Verwenden von Computeralgorithmen, um „autonom aus Daten und Informationen zu lernen“ [KIMu2020], definieren. Bei diesem Prozess können die Algorithmen selbstständig lernen und sich somit an die Aufgabenstellung anpassen. [KIMu2020]

Mit Hilfe der im Folgenden verwendeten Algorithmen wurden bereits im Vorfeld Modelle trainiert, sodass auf Basis eines solchen Modells auch mit noch unbekanntem Material Ergebnisse erzielt werden können.

### 2.2 Begriffsdefinition Face Detection

Das Ziel der Face Detection, zu Deutsch „Gesichtserkennung“, ist es durch eine automatische Suche zu bestimmen, ob ein Gesicht in einem Bild oder Video dargestellt ist oder nicht [ShMo2014].

Allgemein kann die Gesichtserkennung als ein Gebiet der Objekterkennung beschrieben werden, für welches bereits viele Methoden existieren. Bei diesen Methoden werden verschiedene Modelle des maschinellen Lernens verwendet. Die Anwendungsgebiete von Face Detection sind vielfältig, da die Gesichtserkennung für viele Verfahren den ersten und damit grundlegenden Schritt darstellt. So wird sie beispielsweise für kommerzielle Zwecke beim automatischen Fokus von Kameras sowie bei Fußgänger- und Fahrermüdigkeitserkennung in Autos und Zugriffs- oder Zutrittskontrollen verwendet. [ShMo2014] Des Weiteren wird sie bei automatischen Systemen wie Face Recognition (Gesichtswiedererkennung oder -identifikation), darunter auch zur Identifizierung von Personen in kriminellen Zusammenhängen [ShMo2014], Face Tracking und Erkennung von Gesichtsausdrücken verwendet [BeKe2006].



Bei Face Detection können primär zwei Arten von Fehlern auftreten, welche noch immer eine Herausforderung darstellen: die falsch positive und falsch negative Erkennung [AmSa2022]. Falsch positiv bedeutet im konkreten Zusammenhang mit Face Detection, dass auf einem Bild ein Gesicht erkannt wird, obwohl in der Realität keines abgebildet ist. Eine falsch negative Erkennung tritt auf, wenn kein Gesicht auf einem Bild erkannt wird, obwohl eines darauf abgebildet ist. [De2021] Eine mögliche Ursache dafür kann beispielsweise eine sehr niedrige Auflösung des entsprechenden Bildes sein [AmSa2022].

## 2.3 MediaPipe Face Mesh

Wie zuvor erwähnt sollen die Grenzen der Gesichtserkennung mittels MediaPipe Face Mesh in Bezug auf unterschiedliche Faktoren der Informationsreduktion von Bildern getestet werden. Face Mesh wurde zum einen deswegen ausgewählt, weil es anders als die üblichen Algorithmen nicht nur 68 Landmarken [AmSa2022] sondern 468 3D-Landmarken schätzt, um die Oberflächengeometrie des menschlichen Gesichts abzuleiten und damit ein genaueres Ergebnis erzielt werden kann. Zum anderen wurde Face Mesh ausgewählt, weil lightweighted Modellarchitekturen zusammen mit einer GPU-Beschleunigung verwendet werden, sodass es theoretisch auch in Echtzeit beispielsweise auf Videomaterial eines Überwachungssystems angewendet werden kann. [Go2020] „Lightweighted“, zu Deutsch „leicht“ oder „leichtgewichtig“, bedeutet in diesem Zusammenhang, dass weniger Speicher benötigt wird, da keine detaillierten Informationen oder Metadaten enthalten sind [Pt2022].

MediaPipe Face Mesh verwendet maschinelles Lernen, um die 3D-Gesichtsoberfläche anhand der geschätzten 468 3D-Landmarken abzuleiten. Für diesen Prozess ist eine einzige Kameraeingabe ausreichend. [Go2020]

Zudem ist Face Mesh mit dem Face-Transform-Modul verknüpft, welches die Lücke zwischen der Erkennung von Gesichtsmarkmalen und nützlichen AR-Anwendungen (Augmented Reality) in Echtzeit schließt. Dafür erstellt es einen metrischen 3D-Raum und nutzt die Bildschirmpositionen der Gesichtslanmarken, um eine Gesichtstransformation innerhalb dieses Raumes abzuschätzen. [Go2020]

Außerdem wird eine ML Pipeline verwendet, welche auf zwei Neuronalen Netzwerken, die miteinander arbeiten, basiert: ein Detector, der auf dem vollständigen Bild arbeitet und Gesichtspitionen berechnet sowie ein 3D-Landmarken-Modell, das auf diesen Positionen arbeitet und die ungefähre 3D-Oberfläche mittels Regression vorhersagt. [Go2020]

In Face Mesh werden drei verschiedene Modelle genutzt. Bei dem Modell für die Detektion eines Gesichts handelt es sich um dasselbe „BlazeFace“ Modell, welches in MediaPipe Face Detection verwendet wird. Neben diesem wird ein Modell benötigt, um die Gesichtslanmarken vorauszusagen. Für diese 3D-Gesichtslanmarken wurde Transfer-Learning eingesetzt und ein Netzwerk mit mehreren Zielen trainiert. Zu diesen Zielen gehören sowohl

die Voraussage der 3D-Orientierungskordinaten auf synthetisch gerenderten Daten als auch von semantischen 2D-Konturen auf kommentierten realen Daten. [Go2020]

Das 3D-Landmarken-Netzwerk erhält als Eingabe einen zugeschnittenen Video-Frame bzw. ein statisches Bild ohne weitere Tiefeneingabe. Das Modell ist sowohl in der Lage die Positionen der 3D Punkte als auch die Wahrscheinlichkeit der Präsenz eines angemessen ausgerichteten Gesichts auszugeben. [Go2020]

Als Ergänzung zum Face Landmark Modell kann ein weiteres Modell verwendet werden, welches den Fokus auf semantisch bedeutsame Gesichtsregionen legt. Damit können Landmarken um die Lippen, Augen und Iris genauer vorhergesagt werden, allerdings auf Kosten von mehr Rechenleistung. [Go2020]

## 2.4 Verwandte Literatur

Besonders in den letzten Jahren wurde vermehrt im Bereich der Face Detection und Face Recognition geforscht. Dabei wurden unterschiedliche Verfahren und Algorithmen verwendet und entwickelt. Zheng et al. [ZhRa2010] stellen in ihrer Arbeit die Ergebnisse in Bezug auf Face Detection anhand einer Variation des Viola-Jones Detectors unter Zunahme von Color Census Transform und 9-bit Modified Census Transform vor. Bei einer Anzahl von beispielsweise 8x8 Pixeln (kleinste Bildelemente), welche das Gesicht auf dem Bild repräsentieren, kann bei ersterem eine Erkennungsrate von 95 % und unter Zunahme von 9-bit Modified Census Transform eine Erkennungsrate von 93,5 % erreicht werden. Insgesamt konnte eine maximale Erkennungsrate von 99,5 % erzielt werden. Diese Ergebnisse beziehen sich auf Bilder aus dem Georgia-tech-face-Datensatz [GTFD2022]. Grundsätzlich konnten Zheng et al. [ZhRa2010] feststellen, dass die Ergebnisse unter Zunahme der vorgestellten Color Census Transform allgemein besser als die unter Zunahme der Modified Census Transform waren, besonders dann, wenn nicht ausreichend Informationen zur Struktur vorliegen, um eine zuverlässige Erkennung durchzuführen. [ZhRa2010]

Zhou et al. [ZhLi2018] überprüfen in ihrer Arbeit das Robustheitslevel von vier Face Detection Algorithmen in Abhängigkeit verschiedener Faktoren. Bei den vier untersuchten Algorithmen handelt es sich um den Haar-like Adaboost cascade, auch als Viola-Jones Detector bezeichnet, HoG-SVM, Faster R-CNN und S<sup>3</sup>FD. Verwendet wurde der benchmark FDDB-Datensatz [JaLe2010]. Bei durch Gaussian Blur (Gauß'sche Unschärfe) veränderten Bildern konnte festgestellt werden, dass keiner der Algorithmen robust genug ist. Dabei konnte für S<sup>3</sup>FD beispielsweise bei höchster Veränderung eine True Positive Rate (entspricht der Erkennungsquote) von etwa 80 % und für HoG-SVM von nur ca. 35 % erreicht werden. Bei künstlich generiertem Rauschen der Bilder fiel auf, dass alle Modelle sehr stark von diesem Faktor beeinflusst werden. Besonders sticht hervor, dass bei einer Varianz von eins alle Modelle kein Gesicht mehr detektieren können. Zhou et al. [ZhLi2018] beschreiben dabei, dass ein menschlicher Betrachter höchstwahrscheinlich noch immer ein Gesicht von dem Hintergrund unterscheiden könnte. Weiterhin geben sie an, dass bei geringem Kontrast und

reduzierter Helligkeit die verschiedenen Modelle eine höhere Robustheit aufweisen. Dies begründen sie mit der Normalisierung während des Testens. Mit diesen Einstellungen kann S<sup>3</sup>FD bei größter Informationsreduktion noch immer eine Erkennungsquote von etwa 95 % erreichen. Mit Faster R-CNN kann an dieser Stelle ein Wert von etwas unter 90 % erzielt werden, mit HoG SVM etwas unter 70 %. Dagegen wird bei dem Haar AdaBoost-Algorithmus nur eine Erkennungsquote von ca. 28 % erreicht. [ZhLi2018]

## 3 Auswahl des Ausgangsmaterials

In diesem Kapitel wird die Auswahl für die in Kapitel 4 zur Informationsreduktion verwendeten Ausgangsbilder erläutert. Die Bilder wurden vor ihrer Bearbeitung in sechs verschiedene Kategorien aufgeteilt: frontal, mit Brille, mit Sonnenbrille, links lateral, rechts lateral und teilweise verdeckt. Dies dient dazu zu überprüfen, ob beispielsweise eine Sonnenbrille Einfluss auf die Gesichtserkennung bei Informationsreduktion des Originalbildes hat.

Alle Bilder des Ausgangsmaterials sind dem Flickr-Faces-HQ Dataset (FFHQ) [Ro2020] entnommen. Dieser Datensatz wurde unter anderem ausgewählt, da auf vielen Bildern Accessoires, wie z. B. Brillen oder Sonnenbrillen, von der abgebildeten Person getragen werden. [Ro2020]

### 3.1 Kategorie Frontale Bilder

Frontal abgebildete Gesichter auf Bildern gelten zunächst als „optimale“ Voraussetzung für eine Gesichtserkennung. Dies liegt u. a. daran, dass einige Gesichtserkennungsalgorithmen auf eben solchen Bildern äußerst effektiv arbeiten können, so z. B. auch der Viola-Jones Detector (Haar AdaBoost) oder HoG SVM. [ZhLi2018]

Beispiele des Ausgangsmaterials entsprechend dieser Bildkategorie können Anlagen, Teil 1 entnommen werden.

### 3.2 Kategorie Bilder mit Brille

Viele Personen tragen in ihrem Alltag heutzutage eine Brille. In Deutschland sind es beispielsweise 41,1 Millionen der Personen ab 16 Jahren [ZVA2022]. Aus diesem Grund wurden zusätzlich zu frontal ausgerichteten Gesichtern jene Bilder untersucht, auf denen die abgebildete Person eine Brille trägt. Anhand dessen soll geprüft werden, ob und in welchem Ausmaß eine Brille bei Informationsreduktion Einfluss auf die Gesichtserkennung nimmt. Diese Bilder wurden so ausgewählt, dass die Gesichter an sich so frontal wie möglich zur Kamera positioniert sind, um nach Möglichkeit den Unterschied der genannten Kategorien, abgesehen von der Brille, gering zu halten.

Für diese Bildkategorie kann beispielhaftes Ausgangsmaterial in Anlagen, Teil 2 eingesehen werden.

### **3.3 Kategorie Bilder mit Sonnenbrille**

Die Auswahl der Kategorie „mit Sonnenbrille“ begründet sich ähnlich wie die der Kategorie „mit Brille“. Sonnenbrillen verdecken allerdings in vielen Fällen die Augen der Person vollständig. Da die Augenpartie eine wichtige Rolle in der Gesichtserkennung spielt, wurde überprüft, wie sehr eine Sonnenbrille Einfluss auf diese bei Informationsreduktion nimmt.

Ausgangsmaterial ist beispielhaft in Anlagen, Teil 3 dargestellt.

### **3.4 Kategorie Laterale Bilder**

Sowohl links als auch rechts laterale Bilder sind der übergeordneten Kategorie „lateral“ zuzuordnen. Grundsätzlich wurden die Bilder in dieser Kategorie ausgewählt, um mögliche Unterschiede in der Gesichtserkennung aufgrund der Positionierung des Gesichts aufzudecken. Oftmals sind beispielsweise Überwachungskameras so positioniert, dass frontale Aufnahmen nicht oder nicht immer möglich sind. Von welcher Seite eine Person allerdings zu sehen ist, kann durchaus variieren. Aus diesem Grund wurde die Kategorie „lateral“ in links und rechts aufgespalten. Dabei wird ein möglicher Unterschied zwischen diesen beiden lateralen Kategorien überprüft. Bei der Auswahl der Bilder wurde darauf geachtet, dass die abgebildeten Personen beispielsweise keine Brille oder Sonnenbrille tragen.

Beispielbilder können für links laterale Bilder bzw. Gesichter Anlagen, Teil 4 und für rechts laterale Anlagen, Teil 5 entnommen werden.

### **3.5 Kategorie Bilder mit teilweiser Verdeckung**

Als letzte Bildkategorie wurden Bilder ausgewählt, auf denen das dargestellte Gesicht teilweise verdeckt ist. Insbesondere bei Überwachungsaufnahmen, die in der forensischen Fallarbeit ausgewertet werden, ist nicht nur die Qualität des Video- bzw. Bildmaterials von Bedeutung, sondern auch die Tatsache, ob ein Gesicht vollständig abgebildet wird oder ein Teil davon verdeckt ist.

Als teilweise verdeckt gilt dabei beispielsweise, wenn ein Gegenstand oder eine Hand vor einer Gesichtspartie befindet oder dass das Gesicht stark geschminkt ist, wie es u. a. bei Karnevalsschminke der Fall ist. Auch hier soll der Einfluss auf die Ergebnisse der Gesichtserkennung bei Anwendung von Informationsreduktion getestet werden. Beispielhaftes Ausgangsmaterial ist in Anlagen, Teil 6 einzusehen.

## 4 Methoden

Im Folgenden sollen die verwendeten Methoden zur Informationsreduktion der Bilder und deren Implementierung erläutert werden. Davor werden diese Aspekte ebenfalls für die Gesichtserkennung Face Mesh von MediaPipe erläutert.

Für jedes Python-Skript müssen die benötigten Bibliotheken und Module importiert und gegebenenfalls vorher installiert werden. Letzteres kann z. B. mit „pip install <name>“ in der Kommandozeile erfolgen. Anschließend wird der Pfad der einzulesenden Dateien angegeben, woraufhin die dort gespeicherten Dateien mit ihrem vollständigen Pfad in einer Liste gespeichert werden [Se2022]. Ähnliches gilt für die Festlegung der neuen Dateinamen der eingelesenen Dateien. Hier wird die neue Bezeichnung als Zeichenkette der Kombination des angegebenen Speicherorts, der Stelle in der Dateinamenliste und die zusätzliche genaue Bezeichnung des Bildes festgelegt. Dieser letzte Schritt variiert im Skript zur Gesichtserkennung mit Face Mesh leicht, da dort nur ein einziger Pfad zur Speicherung der Bilder angegeben wird.

Außerdem ist zu erwähnen, dass das gesamte Ausgangsmaterial darauf überprüft wurde, ob grundsätzlich mit Hilfe von Face Mesh jeweils ein Gesicht darauf detektiert werden kann. Diese Überprüfung fiel positiv aus, was bedeutet, dass eine Nicht-Erkennung auf die entsprechende Informationsreduktion zurückzuführen ist.

## 4.1 Schema des Skripts zu MediaPipe Face Mesh

Die folgende Darstellung zeigt das Schema des Python-Skripts zu MediaPipe Face Mesh.

```
Import
|- cv2
|- mediapipe
|- os
|- Path (from pathlib)
|
Angabe der zu verarbeitenden Dateien
|- Pfad angeben
|- alle Dateien aus Pfad in Liste speichern
|
Verarbeitung der Bilder/Erstellung der Gesichtsnetze
|- mit MediaPipe Face Mesh-Funktion:
|-- jedes Element (/Bild) aus Liste einlesen
|-- Bild in RGB-Format überführen und anschließend
|   Gesichtsnetzfunktion anwenden
|-- wenn kein Ergebnis, dann:
|--- Dateiname des Bildes in Liste für nicht erkannt speichern
|--- mit nächstem Bild fortfahren
|-- sonst:
|--- Dateiname des Bildes in Liste für erkannt speichern
|- erkanntes Gesichtsnetz auf Bild zeichnen
|- Bild mit erkanntem Gesichtsnetz in angegebenen Pfad mit
|   Bezeichnung speichern
```

Die Grundfunktionen dieses Skriptes sind der offiziellen Internetdokumentation entnommen [Go2020]. Es wurden lediglich Aspekte, wie das Speichern der Dateinamen und auch der Ergebnisbilder, eingefügt bzw. angepasst. Weitere Informationen zu Bestandteilen von MediaPipe Face Mesh sind Kapitel 2.3 zu entnehmen.

## 4.2 Methode zur Verpixelung des Ausgangsmaterials

In der forensischen Fallarbeit liegt als Ausgangssituation oftmals Bild- und Videomaterial von Überwachungssystemen vor. Aufgrund der Tatsache, dass das Bildmaterial der meisten Überwachungssysteme eine geringe Auflösung aufweist, ist der zu analysierende Bereich häufig verpixelnt. Aus diesem Grund wurde „Verpixelung“ als eine Kategorie zur Informationsreduktion ausgewählt.

In diesem Schritt wird das Ausgangsmaterial bestehend aus 50 Bildern pro Kategorie mit Hilfe der Python-Bibliothek OpenCV (Open Source Computer Vision Library) in mehreren Abstufungen verpixelnt [StOv2022]. OpenCV beinhaltet eine Vielzahl an Computer Vision-Algorithmen. Besonders relevant für die Verpixelung ist das Modul zur Bildverarbeitung (Image Processing). Mit diesem werden die Bilder auf die entsprechende Pixelanzahl gebracht. [Do2022a] Nach und nach wird die Pixelanzahl immer weiter verringert, sodass die Qualität der Bilder zunehmend verschlechtert wird.

Grundsätzlich wurde sich für die in diesem Kapitel vorgenommenen Abstufungen an den gängigen Abstufungen wie 64x64, 32x32 und 16x16 Pixeln orientiert. Besonders zwischen den Stufen 16x16 und 32x32 Pixeln wurden jedoch noch einmal mehr Abstufungen mit jeweils geringem Abstand vorgenommen, da bei diesen Stufen, ausgehend von vorausgegangenen Versuchsreihen, die meisten Unterschiede erwartet werden. Die Versuche haben gezeigt, dass zwischen 32x32 und 64x64 Pixeln solche feinen Abstufungen nicht notwendig sind, da die Erkennungsquote in den meisten Fällen zwischen diesen noch nicht signifikant sinkt.

Bezüglich der folgenden Abstufungen ist zu erwähnen, dass als untere Grenze der Pixelanzahl der Wert festgelegt wurde, bei dem auf jedem Bild kein Gesicht mehr erkannt werden konnte. Zumeist war dies bei 16x16 Pixeln der Fall. Des Weiteren wurden weitere Abstufungen vor 64x64 Pixeln vorgenommen, sofern an dieser Stelle noch keine Erkennungsquote in Bezug auf die relative Anzahl erkannter Gesichter von 100 % erreicht werden konnte.

Die folgende Darstellung zeigt das verwendete Python-Skript in seinen wesentlichen Strukturen.



```

Import
|- cv2
|- os
|- Path (from pathlib)
|
Festlegen der einzulesenden Dateien
|- Pfad angeben
|- alle Dateien aus Pfad speichern
|
Festlegen der Bezeichnung der veränderten Bilder und damit auch
Speicherort
|- so lange i <= Anzahl der einzulesenden Bilder
|- neuer Dateiname + Speicherort
|
Verarbeitung der Bilder
|- Bild jeweils einlesen
|- Höhe und Breite mit shape() detektieren und speichern
|- gewünschte Pixelanzahl definieren
|- Bild jeweils auf definierte Pixelanzahl skalieren
|- temporäres Bild wieder auf ursprüngliche Größe skalieren
|- erfolgreich verpixeltes Bild mit festgelegter Bezeichnung
|   speichern
|

```

Damit die Bilder mittels der Python-Bibliothek OpenCV ver- bzw. bearbeitet werden können, müssen diese zuerst entsprechend mit der Methode `imread()` eingelesen werden. Daraufhin kann mit der Methode `shape()` die aktuelle Höhe und Breite in Form eines Tupels der Anzahl an Reihen und Spalten gespeichert werden [Do2022b]. Mit der definierten Pixelanzahl kann dann jedes Bild mit der Methode `resize()` auf die entsprechende Größe bzw. Anzahl an Pixeln gebracht werden. Dabei wird lineare Interpolation verwendet, welche standardmäßig bei dieser Methode ausgewählt ist. Dieses bearbeitete Bild wird jeweils zwischengespeichert. Im nächsten Schritt wird dieses temporäre Bild ebenfalls mittels Interpolation, allerdings mit der Methode „nearest“, zurück auf die ursprüngliche Größe gebracht. Zuletzt wird das erfolgreich verpixelte Bild mit entsprechender Bezeichnung gespeichert.

Allgemein ist das Konzept der Interpolation, bekannte Daten zu nutzen, um Werte für unbekannte Datenpunkte zu schätzen. Bildinterpolation arbeitet in zwei Richtungen und versucht dabei die beste Annäherung eines neuen Pixels in Bezug auf seine Farbe und Intensität auf Grundlage der Werte der umliegenden Pixel zu erreichen. [Ca2022]

Bei Interpolation kann grundsätzlich zwischen vier verschiedenen Methoden gewählt werden: `area`, `linear`, `cubic` oder `nearest`. Erstere verwendet die Bereichsbeziehung zwischen Pixeln für das Resampling, welches das Verfahren zum Umwandeln eines abgetasteten Bildes von einer beliebigen Koordinatenstruktur in eine andere Struktur bezeichnet [PaS2018]. Die zweite Methode wirkt sich auf die 4x4 benachbarten Pixel des Bildes aus. Dabei wird der Gewichtungsdurchschnitt der 16 Pixel genommen, um das neue interpolierte Pixel zu erzeugen. Bei der dritten Methode, der linearen Interpolation, handelt es sich um

ein ähnliches Vorgehen wie bei der bikubischen (cubic) Interpolation. Allerdings werden anstatt 4x4 lediglich 2x2 benachbarte Pixel verwendet, um den gewichteten Durchschnitt für das interpolierte Pixel zu berechnen. Nearest Interpolation verwendet das „Nächster-Nachbar-Prinzip“, das bedeutet, dass lediglich das nächste benachbarte Pixel für die Interpolation berücksichtigt wird. [TL2021]

### **4.3 Methode zur Veränderung der Helligkeit des Ausgangsmaterials**

Über eine geringe Auflösung hinaus weist das Material nicht selten eine zu hohe oder niedrige Helligkeit auf, beispielsweise bei Nachtaufnahmen, die ohne eine dafür geeignete Kamera aufgenommen wurden. Es gilt dementsprechend zu untersuchen inwieweit dies Auswirkungen auf die Gesichtserkennung mittels MediaPipe Face Mesh hat.

Nachdem die Bilder schrittweise durch Verpixeln verschlechtert wurden, wurde die Helligkeit der Ausgangsbilder zunächst immer weiter erhöht und anschließend verringert. Auch für dieses Szenario wurde zur Durchführung dieser Schritte Methoden aus der Python-Bibliothek OpenCV verwendet. Die wesentliche Funktion zur Veränderung der Helligkeit wurde einem bereits existierenden Skript [Ge2020] entnommen. Dieses ist in der Lage einen Schieberegler über einem bestimmten ausgewählten Bild anzuzeigen, mit dem individuell die Helligkeit des Bildes verändert werden kann [Ge2020]. Das Skript wurde jedoch so angepasst, dass es ausreicht den gewünschten Helligkeitswert vor Ausführung der Funktion „controller“ einzustellen bzw. anzugeben. Grundsätzlich wird der Helligkeitswert eines Bildes mit Hilfe der Funktion `addWeighted()` der OpenCV-Bibliothek angepasst. Erneut wird bei jedem Schritt überprüft, ob ein Gesicht in den verschiedenen Kategorien mit Hilfe von Face Mesh erkannt wird oder nicht. Auf der folgenden Seite ist das Python-Skript in seiner grundlegenden Struktur dargestellt.

```

Import
|- cv2
|- os
|- Path (from pathlib)
|
Festlegen der einzulesenden Dateien
|- Pfad angeben
|- alle Dateien aus Pfad speichern
|
Festlegen der Bezeichnung der veränderten Bilder und damit auch
Speicherort
|- so lange i <= Anzahl der einzulesenden Bilder
|- neuer Dateiname + Speicherort
|
Definieren der Funktion „Controller“
|- Helligkeitswert als Integer des Ergebnis der Formel defi-
| nieren
|- wenn neuer Helligkeitswert ungleich und größer als null,
| dann:
|-- Speichern des Helligkeitswerts in „shadow“
|-- max auf 255 festlegen
|- sonst:
|-- shadow auf 0 setzen
|-- max als 255 + Helligkeitswert festlegen
|- alpha als (max-shadow) / 255 definieren
|- shadow in gamma speichern
|- alpha und gamma auf Bild anwenden
|- wenn neuer Helligkeitswert gleich null, dann:
|-- Rückgabewert auf Eingabebild festlegen
|
Verarbeitung der Bilder
|- Bild jeweils einlesen
|- Kopie des eingelesenen Bildes erstellen, um Veränderung
| darauf anzuwenden
|- Funktion „controller“ auf kopiertem Bild mit angegebenem
| Helligkeitswert anwenden
|- erfolgreich in der Helligkeit verändertes Bild mit
| festgelegter Bezeichnung speichern
|

```

Der nachstehenden Formel zur Berechnung der neuen Helligkeit wird der jeweils angegebene Helligkeitswert übergeben. Diese Angabe ist demnach ausschlaggebend für die Gesamthelligkeit des Bildes, da sie anhand dieses Wertes anschließend verändert wird. Mit Hilfe dieser Formel wird der Helligkeitswert normalisiert.

$$brightness = int((brightness - 0) * (255 - (-255)) / (510 - 0) + (-255)) \text{ [Ge2020]}$$

Nach der Berechnung folgen weitere Wertedefinitionen, um schlussendlich Werte für alpha und gamma zu definieren. Diese werden wiederum mit Hilfe der Funktion `addWeighted()`

auf das Bild angewendet. Die Funktion `addWeighted()` berechnet die gewichtete Summe zweier Arrays [Do2022c].

Der Funktion `addWeighted()` werden fünf Parameter übergeben. Die zwei Input-Arrays, welche jeweils durch das eingelesene Bild dargestellt werden, einem alpha-, beta- und einem gamma-Wert. Der beta-Wert wird dabei allerdings auf 0 gesetzt, weswegen dieser außer Acht zu lassen ist. Er wird auf 0 gesetzt, da es sich bei beiden Arrays um denselben Input handelt. Der alpha-Wert gibt die Gewichtung des ersten Arrays bzw. dessen Elemente an. Mit der Angabe eines gamma-Wertes wird ein Skalar-Wert angegeben, welcher zu jeder Summe der Arrays hinzugerechnet wird. [Ge2020]

In der darstellenden Technik wird beispielsweise von einem Gamma-Wert gesprochen in dem Zusammenhang, dass die menschliche Helligkeitswahrnehmung als nicht linear gilt. Dies wiederum bedeutet, dass „die empfundene Helligkeit nicht linear mit der tatsächlichen Helligkeit ansteigt“ [IT2022]. Eine dementsprechende Anpassung kann man als Gamma-korrektur bezeichnen. Im Grunde genommen handelt es sich dabei um eine optische Anpassung für das menschliche Auge. [IT2022] Die Anwendung dessen begründet sich auf der Tatsache, dass das originale Skript [Ge2020] die Veranschaulichung des Ergebnisbildes vorsieht.

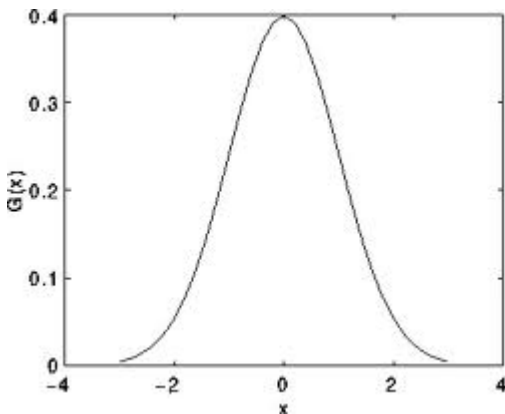
## 4.4 Methode zur Anwendung Gauß'scher Unschärfe auf das Ausgangsmaterial

Mit Hilfe der Gauß'schen Unschärfe kann unter anderem durch Bewegung einer Person erzeugte Unschärfe simuliert werden. Des Weiteren kann sie nicht fokussierte Kameraaufnahmen nachstellen. Aus diesen Gründen wurde die Robustheit der Gesichtserkennung auch bei dieser Kategorie der Informationsreduktion überprüft.

Um eine solche Unschärfe zu erhalten, wurden im Folgenden aus der Python Imaging Library (PIL) die Module `Image` und `ImageFilter` verwendet. Das `Image`-Modul bietet unter anderem eine Reihe an Werksfunktionen, beispielsweise Funktionen zum Laden von Bildern aus Dateien und zum Erstellen neuer Bilder [LuCl2022a]. Aus dem Modul `ImageFilter`, welches allgemein Definitionen für einen vordefinierten Satz von Filtern bietet, kann der Filter „GaussianBlur“ mit Hilfe der `Image.filter()`-Methode auf die gewünschten Bilder angewendet werden [LuCl2022b]. Der Parameter „radius“ beschreibt dabei den Unschärfe-Radius, welcher die Intensität der angewendeten Gauß'schen Unschärfe festlegt [Ge2019].

Im Allgemeinen wird ein Gauß'scher Filter angewendet, indem jeder Punkt im Eingangsarray mit einem Gauß'schen Kernel „gefaltet“ und dann alle Punkte summiert werden, um das Ausgangsarray zu erzeugen. Wenn man beispielsweise ein Bild als eindimensional annimmt, wird das Pixel in der Mitte des Gauß'schen Kerns am höchsten gewichtet. Gleichzeitig werden die entsprechenden Nachbarpixel mit zunehmender Entfernung zum mittleren

Pixel immer weniger gewichtet. Diese Aspekte sind in dem Graphen in Abbildung 1 visuell dargestellt. [Do2022d]



**Abbildung 1: Graphische Darstellung eines eindimensionalen Gauß'schen Kerns mit Gewichtung eines Pixels in Relation zur Entfernung zur Mitte des Gauß'schen Kerns [Do2022d]**

Gauß'sche Weichzeichnung oder auch Unschärfe, wendet eine mathematische Funktion auf ein Bild an, um eine Unschärfe zu erzielen [Ad2022]. Ein Gauß'scher Weichzeichner ist ein „Tiefpassfilter, der ungleichmäßige Pixelwerte in einem Bild glättet, indem er extreme Ausreißer entfernt“ [Ad2022].

Aufgrund der Voreinstellung des adaptierten Python-Codes [Ge2019] wird mit einem Radius-Wert von fünf begonnen. Diese Radius-Einstellung gibt in etwa die Intensität der Gauß'schen Unschärfe an [Ge2019]. Um große Sprünge in der Erkennungsquote möglichst zu vermeiden, wurden die meisten Abstufungen in Einser-Schritten vorgenommen. Bei den letzten Abstufungen wurden allerdings ab einem Wert von 20 Fünferschritte ausgewählt. Dies begründet sich darauf, dass vorherige Beobachtungen zur Auswahl dieser Werte gezeigt haben, dass an diesen Stellen auch mit mehr Zwischenschritten eine teilweise größere Differenz bezogen auf die Erkennungsquote nicht vermieden werden kann und somit keinen Mehrwert bringen. Zudem fällt bei den meisten Bildkategorien auf, dass ab einer Einstellung von 30 die Erkennungsquote nur noch 0 % beträgt. Tritt dieser erwartete Fall ein, so werden höhere Werte nicht zusätzlich überprüft. Im Folgenden erfolgt die Darstellung der grundlegenden Struktur des entsprechenden Python-Skripts.

```

Import
|- cv2
|- numpy as np
|- os
|- Path (from pathlib)
|- Image, ImageFilter (from PIL)
|
Festlegen der einzulesenden Dateien
|- Pfad angeben
|- alle Dateien aus Pfad speichern
|
Festlegen der Bezeichnung der veränderten Bilder und damit auch
Speicherort
|- so lange i <= Anzahl der einzulesenden Bilder
|- neuer Dateiname + Speicherort
|
Verarbeitung der Bilder
|- Bild jeweils einlesen
|- Filter „GaussianBlur“ auf eingelesenes Bild anwenden
|- Bild mit angewendetem Filter als numpy-Array speichern
|- als Array gespeichertes Bild in RGB-Farben konvertieren
|- konvertiertes Bild mit festgelegter Bezeichnung speichern
|

```

Zunächst werden auch hier die zu bearbeitenden Bilder in einer Liste gespeichert. Zudem werden in einer weiteren Liste die späteren Dateinamen hinterlegt. Sobald dies geschehen ist, können die Bilder einzeln eingelesen werden. Auf jedes dieser Bilder wird dann der Filter „GaussianBlur“ aus dem ImageFilter-Modul der Python Imaging Library angewendet. Um die Bilder im Anschluss daran mittels der Funktion `imwrite()` aus OpenCV speichern zu können, muss das Bild jeweils in ein numpy-Array und auf Grund dessen in RGB-Farben konvertiert werden. Wenn dies durchgeführt wurde, kann das Bild wie zuvor gespeichert werden.

## 4.5 Methode zur Anwendung von Interpolation auf das Ausgangsmaterial

In der forensischen Fallarbeit wird mit Bildern in unterschiedlichsten Größen gearbeitet. Aufgrund der Tatsache, dass die Größe eines Bildes nicht immer ausreichend ist, um damit weiterzuarbeiten, kommt es vor, dass diese erhöht werden muss. Im einfachsten Fall würde eine solche Vergrößerung mittels Interpolation durchgeführt werden, wobei die Qualität des Bildes mitunter stark leidet. Aus diesem Grund wurde untersucht, ab welchem Vergrößerungsverhältnis mittels Interpolation der Bilder MediaPipe Face Mesh nicht mehr in der Lage ist Gesichter zuverlässig zu detektieren. Dafür wurden die Bilder zunächst mittels Interpolation verkleinert und anschließend auf ihre Originalgröße von 512x512 Pixeln gebracht.

Die Verkleinerung und Vergrößerung mittels Interpolation erfolgt hier unter Anwendung derselben Funktion (`resize()`) wie bei den Zwischenschritten des Verpixelns des Ausgangsmaterials. Die nachfolgende Darstellung zeigt das zugrundeliegende Schema des Python-Skriptes.

```
Import
|- cv2
|- os
|- Path (from pathlib)
|
Festlegen der einzulesenden Dateien
|- Pfad angeben
|- alle Dateien aus Pfad speichern
|
Festlegen der Bezeichnung der veränderten Bilder und damit auch
Speicherort
|- so lange i <= Anzahl der einzulesenden Bilder
|- neuer Dateiname + Speicherort
|
Verarbeitung der Bilder
|- Bild jeweils einlesen
|- Bild mit resize() neu skalieren
|- erfolgreich in der Größe verändertes Bild mit festgelegter
| Bezeichnung speichern
|
```

Demnach können die Bilder, nach Einlesen und Festlegen des Speicherorts der entsprechenden Dateien, mit der Funktion `resize()` aus OpenCV auf die gewünschte Größe skaliert und anschließend gespeichert werden. Dieser Vorgang muss pro Eingabebild insgesamt zweimal durchgeführt werden, im ersten Schritt zum Verkleinern und im zweitem zum Vergrößern des Bildes.

## 5 Ergebnisse

In diesem Kapitel werden jeweils die Ergebnisse der Gesichtserkennung in Bezug auf die verschiedenen Informationsreduktions-Kategorien dargelegt und erläutert. Dabei erfolgt eine Unterscheidung der verschiedenen Bildkategorien des Ausgangsmaterials.

Die Abbildungen in diesem Kapitel bzw. die darin veranschaulichten Bilder sind zum Zweck der besseren Lesbarkeit nicht in ihrer Originalgröße dargestellt. Stattdessen betragen sie etwa ein Drittel ihrer eigentlichen Größe.

### 5.1 Ergebnisse der Gesichtserkennung bei Verpixelung des Ausgangsmaterials in Bezug auf die verschiedenen Bildkategorien

In diesem Abschnitt werden die Ergebnisse der Gesichtserkennung bei Informationsreduktion durch Verpixeln der Bilder in den jeweiligen Bildkategorien diskutiert.

#### Frontale Bilder

In Tabelle 1 und in entsprechender Abbildung in Anlagen, Teil 7 werden die Ergebnisse der Versuche in Bezug auf das Verpixeln der frontalen Bilder gezeigt. Zu frontalen Bildern ist grundsätzlich zu sagen, dass diese bezüglich der Gesichtserkennung als „ideal“ angesehen werden. Dies begründet sich dadurch, dass bei frontaler Positionierung des Gesichts zur Kamera, ohne zusätzlich möglicherweise einschränkende Merkmale wie das Tragen einer Brille oder Sonnenbrille, alle Landmarken erkannt werden müssten.



**Tabelle 1: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil erkannter Gesichter
Pixelanzahl = 64x64	50	100 %
Pixelanzahl = 50x50	50	100 %
Pixelanzahl = 32x32	49	98 %
Pixelanzahl = 30x30	50	100 %
Pixelanzahl = 29x29	50	100 %
Pixelanzahl = 28x28	50	100 %
Pixelanzahl = 25x25	47	94 %
Pixelanzahl = 21x21	26	52 %
Pixelanzahl = 19x19	7	14 %
Pixelanzahl = 18x18	16	32 %
Pixelanzahl = 17x17	2	4 %
Pixelanzahl = 16x16	1	2 %
Pixelanzahl = 15x15	1	2 %
Pixelanzahl = 14x14	1	2 %
Pixelanzahl = 13x13	0	0 %
Durchschnitt	26,666 $\approx$ 26,7	53,333 $\approx$ 53,3 %

Die in Tabelle 1 dargestellten Ergebnisse zeigen, dass Face Mesh bei einer Pixelanzahl von 16x16 nur auf einem Bild ein Gesicht erkennt. Eine Erkennungsrate von 0 % wird erst bei einer Pixelanzahl von 13x13 erreicht. Auffällig ist, dass bei einer Pixelanzahl von 19x19 entgegen der erwarteten Tendenz weniger Gesichter erkannt werden als bei 18x18 Pixeln, weshalb die Differenz zwischen den Ergebniswerten bei 21x21 und 19x19 Pixeln relativ groß ist. Eine noch höhere Differenz wird z. B. zwischen 25x25 und 21x21 Pixeln ersichtlich. Weiterhin fällt auf, dass bis zu einer Pixelanzahl von 28x28 auf fast allen Bildern ein Gesicht erkannt wird. Jedoch gibt es bei 32x32 Pixeln einen minimalen Abfall von 2 %. Die beiden anderen genannten Auffälligkeiten werden bei den anderen Bildkategorien erneut aufgegriffen.

Die durchschnittliche Erkennungsquote beträgt etwa 53,3 %.

## Bilder mit Brille

Im nächsten Schritt wurden die Verpixelungsstufen auf Bilder angewendet, auf denen die abgebildete Person eine Brille trägt. Bei der Auswahl der Bilder wurde versucht möglichst frontal ausgerichtete Gesichter zu verwenden, damit der Vergleich zu den frontalen Bildern ohne Brille möglichst akkurat ist. Die hierbei erzielten Ergebnisse sind in Tabelle 2 und in entsprechender Abbildung in Anlagen, Teil 7 einzusehen.

**Tabelle 2: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil erkannter Gesichter
Pixelanzahl = 64x64	50	100 %
Pixelanzahl = 50x50	49	98 %
Pixelanzahl = 32x32	45	90 %
Pixelanzahl = 30x30	41	82 %
Pixelanzahl = 29x29	41	82 %
Pixelanzahl = 28x28	39	78 %
Pixelanzahl = 25x25	29	58 %
Pixelanzahl = 21x21	13	26 %
Pixelanzahl = 19x19	0	0 %
Pixelanzahl = 18x18	7	14 %
Pixelanzahl = 17x17	6	12 %
Pixelanzahl = 16x16	2	4 %
Pixelanzahl = 15x15	0	0 %
Durchschnitt	24,769 $\approx$ 24,8	49,538 $\approx$ 49,5 %

Die in Tabelle 2 dargestellten Ergebnisse zeigen, dass auch bei Bildern mit Brille bei einer Anzahl von 19x19 Pixeln im Vergleich zum folgenden Schritt weniger Gesichter erkannt werden. Hier kommt es zu dem Extremfall, dass überhaupt kein Gesicht erkannt wird. Ansonsten scheint die Erkennungsquote im Vergleich zu den frontalen Bildern ohne Brille langsamer zu sinken. Beispielsweise zeigt der Verlauf hier bei einer Pixelanzahl von 25x25 nur einen Anteil an erkannten Gesichtern von 58 %, wohingegen Tabelle 1 an dieser Stelle ein Wert von 94 % entnommen werden kann. Allgemein, mit Ausnahme des Wertes bei 19x19 Pixeln, ist die Tendenz sinkend oder gleichbleibend und weist keine weiteren

Auffälligkeiten auf. Beispielsweise kann bis zu einer Pixelanzahl von 32x32 stets eine Erkennungsquote von mindestens 90 % erreicht werden. Außerdem beträgt die Erkennungsquote, wie zunächst als Grundannahme vorausgesetzt, bei 64x64 Pixeln 100 %.

Aufgrund der Tatsache, dass bei einer Anzahl von 15x15 Pixeln bereits keine Gesichter mehr erkannt worden sind, wurden die Verpixelungsstufen von 14x14 und 13x13 Pixeln nicht zusätzlich überprüft. Diese weitere Abstufung war allerdings notwendig zu prüfen, da bei 16x16 Pixeln zwei Gesichter vermeintlich erkannt wurden.

Die durchschnittliche Erkennungsquote beträgt etwa 49,5 %.

## Bilder mit Sonnenbrille

Ähnlich zu den Bildern, auf denen die dargestellte Person eine Brille trägt, wurden auch solche untersucht, auf denen eine Sonnenbrille getragen wird. Es ist zu vermuten, dass leicht schlechtere Ergebnisse im Vergleich zu den Bildern mit Brillen erzielt werden, da die Sonnenbrille jeweils die Augen der dargestellten Person verbirgt. Tabelle 3 und entsprechende Abbildung in Anlagen, Teil 7 stellen die Ergebnisse dar.

**Tabelle 3: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil erkannter Gesichter
Pixelanzahl = 64x64	50	100 %
Pixelanzahl = 50x50	49	98 %
Pixelanzahl = 32x32	30	60 %
Pixelanzahl = 30x30	29	58 %
Pixelanzahl = 29x29	22	44 %
Pixelanzahl = 28x28	21	42 %
Pixelanzahl = 25x25	17	34 %
Pixelanzahl = 21x21	11	22 %
Pixelanzahl = 19x19	3	6 %
Pixelanzahl = 18x18	7	14 %
Pixelanzahl = 17x17	2	4 %
Pixelanzahl = 16x16	1	2 %
Pixelanzahl = 15x15	1	2 %
Pixelanzahl = 14x14	1	2 %
Pixelanzahl = 13x13	0	0 %
Durchschnitt	16,266 $\approx$ 16,3	32,533 $\approx$ 32,5 %

Tabelle 3 zeigt bei einer Pixelanzahl von 17x17 bis 13x13 eine langsam sinkende Erkennungsquote. Bei einer Pixelanzahl von 18x18 kann zuletzt eine Erkennungsquote von über 10 % erreicht werden. Darüber hinaus lässt sich sagen, dass die Erkennungsquote bei 19x19 Pixeln erneut sinkt und bei 18x18 Pixeln anschließend wieder steigt, allerdings

langsamer als beispielsweise bei den Eingabebildern mit Brillen. Im Allgemeinen werden die Ergebniswerte der Bilder mit Brille zumeist untertroffen. Beispielsweise kann bei 30x30 Pixeln mit 58 % eine Erkennungsquote von über 50 % erreicht werden. Mit jeweils einem Wert der Erkennungsquote von 98 % bzw. 100 % für die Pixelanzahlen 50x50 und 64x64 werden die gleichen Werte wie in Tabelle 2 erreicht.

Die durchschnittliche Erkennungsquote beträgt etwa 32,5 %.

## Links laterale Bilder

Um einen möglicherweise auftretenden Unterschied zwischen der Erkennung von links und rechts lateralen Gesichtern aufzudecken bzw. zu untersuchen, wurden lateral aufgenommene Bilder in „links“ und „rechts“ aufgeteilt. Damit ein möglichst akkurater Vergleich zu den anderen Kategorien gezogen werden kann, wurde darauf geachtet laterale Bilder zu verwenden, auf die kein anderes Kategoriemerkmal, wie z. B. das Tragen einer Brille, zutrifft. Die Ergebnisse der links lateral abgebildeten Gesichter sind in Tabelle 4 und in entsprechender Abbildung in Anlagen, Teil 7 dargestellt.

**Tabelle 4: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil erkannter Gesichter
Pixelanzahl = 80x80	50	100 %
Pixelanzahl = 70x70	49	98 %
Pixelanzahl = 64x64	49	98 %
Pixelanzahl = 50x50	48	96 %
Pixelanzahl = 32x32	42	84 %
Pixelanzahl = 30x30	42	84 %
Pixelanzahl = 29x29	42	84 %
Pixelanzahl = 28x28	39	78 %
Pixelanzahl = 25x25	30	60 %
Pixelanzahl = 21x21	15	30 %
Pixelanzahl = 19x19	1	2 %
Pixelanzahl = 18x18	2	4 %
Pixelanzahl = 17x17	2	4 %
Pixelanzahl = 16x16	0	0 %
Durchschnitt	29,357 $\approx$ 29,4	58,714 $\approx$ 58,7 %

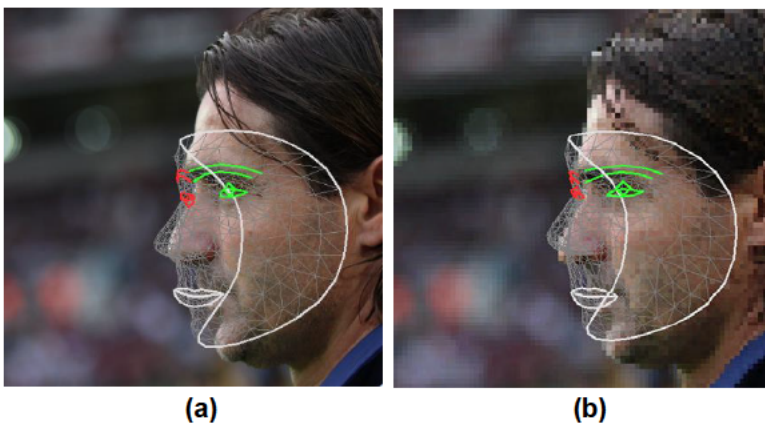
Bei den in Tabelle 4 dargestellten Ergebnissen ist festzustellen, dass erneut eine Steigung der Erkennungsquote zwischen 19x19 und 18x18 Pixeln vorliegt. Bei einer Pixelanzahl von 32x32, 30x30 und 29x29 scheint die Erkennungsquote zunächst zu stagnieren. Allerdings sinkt diese ab einer Anzahl von 28x28 Pixeln weiter. Im Vergleich zu den anderen

Ergebnissen kann in dieser Kategorie auch bei 64x64 Pixeln keine Erkennungsquote von 100 % erreicht werden. Abbildung 2 zeigt das Bild, auf dem bei einer Pixelanzahl von 64x64 kein Gesicht erkannt wird. Auf diesem ist die Person vollständig lateral zur Kamera positioniert. In Abbildung 3 ist zum einen dieses Bild bei einer Pixelanzahl von 80x80 sowie das entsprechend erkannte Gesichtsnetz, zum anderen das dazugehörige Referenzbild, ebenfalls mit dem darauf erkannten Gesichtsbild, dargestellt. Des Weiteren wurde getestet, ab wann auch auf diesem Bild ein Gesicht erkannt werden kann. Um eine unendliche Menge an Abstufungen zu vermeiden, wurden als dafür vorgesehene Pixelanzahlen zunächst der nächste Zehner und darauffolgende Zehner-Schritt(e) festgelegt. Dabei ist festzustellen, dass erst bei einer Pixelanzahl von 80x80 eine Erkennungsrate von 100 % erreicht wird. Eine weitere Beobachtung ist, dass es sich bei den nicht erkannten Gesichtern bei 64x64 und bei 70x70 jeweils um das gleiche Bild bzw. Gesicht handelt.

Die durchschnittliche Erkennungsquote beträgt etwa 58,7 %.



**Abbildung 2:** Dargestelltes Gesicht auf dem in 64x64 Pixeln kein Gesichtsnetz erkannt werden konnte, entsprechend Bild 19\_links\_pixel\_64\_64.png



**Abbildung 3:** Darstellung der Ergebnisbilder mit bei (a) 512x512 Pixeln und (b) 80x80 Pixeln erkannten Gesichtsnetzen der Kategorie „links lateral“. Bild (a) entspricht dem Eingabebild 19\_links.png, Bild (b) dem Eingabebild 19\_links\_pixel\_80\_80.png

## Rechts laterale Bilder

Tabelle 5 und entsprechende Abbildung in Anlagen, Teil 7 zeigen die Ergebnisse der rechts gerichteten Bilder bzw. Gesichter.

**Tabelle 5: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil erkannter Gesichter
Pixelanzahl = 64x64	50	100 %
Pixelanzahl = 50x50	49	98 %
Pixelanzahl = 32x32	41	82 %
Pixelanzahl = 30x30	42	84 %
Pixelanzahl = 29x29	38	76 %
Pixelanzahl = 28x28	34	68 %
Pixelanzahl = 25x25	29	58 %
Pixelanzahl = 21x21	8	16 %
Pixelanzahl = 19x19	6	12 %
Pixelanzahl = 18x18	2	4 %
Pixelanzahl = 17x17	1	2 %
Pixelanzahl = 16x16	0	0 %
Durchschnitt	25	50 %

Im Gegensatz zu den vier vorangegangenen Kategorien – frontal, mit Brille, mit Sonnenbrille und links lateral – ist bei diesen Ergebnissen keine Erhöhung bezüglich der Anzahl erkannter Gesichter zwischen einer Pixelanzahl von 19x19 und 18x18 zu beobachten. Allerdings fällt auf, dass in den ersten Stufen der Verpixelung weniger Gesichter erkannt werden als bei den vorangegangenen Durchführungen. Dennoch kann bei einer Pixelanzahl von 64x64 eine Erkennungsquote von 100 % erreicht werden. Es zeigt sich grundsätzlich eine sinkende Tendenz, mit Ausnahme einer Erhöhung der Erkennungsquote von 2 % bei einer Pixelanzahl von 30x30 im Vergleich zu 32x32 Pixeln.

Die durchschnittliche Erkennungsquote beträgt 50 %.



## Bilder mit teilweiser Verdeckung

Im Folgenden wird überprüft, ob eine teilweise Verdeckung des Gesichts die Gesichtserkennung bei einer Informationsreduktion durch Verpixeln der Bilder beeinflusst. Die Ergebnisse können Tabelle 6 und entsprechender Abbildung in Anlagen, Teil 7 entnommen werden.

**Tabelle 6: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verpixelten Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil erkannter Gesichter
Pixelanzahl = 70x70	50	100 %
Pixelanzahl = 64x64	49	98 %
Pixelanzahl = 50x50	46	92 %
Pixelanzahl = 32x32	33	66 %
Pixelanzahl = 30x30	36	72 %
Pixelanzahl = 29x29	29	58 %
Pixelanzahl = 28x28	25	50 %
Pixelanzahl = 25x25	23	46 %
Pixelanzahl = 21x21	9	18 %
Pixelanzahl = 19x19	0	0 %
Pixelanzahl = 18x18	1	2 %
Pixelanzahl = 17x17	1	2 %
Pixelanzahl = 16x16	0	0 %
Durchschnitt	23,231 $\approx$ 23,2	46,462 $\approx$ 46,5 %

Die Erhöhung der Erkennungsquote von 19x19 auf 18x18 ist auch hier zu erkennen, allerdings wurde in den nachfolgenden Verpixelungsstufen maximal ein Gesicht erkannt. Weiter ist zu beobachten, dass bei maximal 64x64 Pixeln eine Höchsterkennungsquote von 98 % und nicht von 100 % erreicht werden kann. Aus diesem Grund wurde, wie bei den links lateralen Bildern, eine weitere Abstufung mit einer Pixelanzahl von 70x70 überprüft. An dieser Stelle konnte dabei eine Erkennungsquote von 100 % erreicht werden. Deshalb wurde im Gegensatz zu den links lateralen Bildern kein Test der Pixelanzahl 80x80 vorgenommen. Die durchschnittliche Erkennungsquote beträgt 46,5 %.

## **5.2 Ergebnisse der Gesichtserkennung bei veränderter Helligkeit des Ausgangsmaterials in Bezug auf die verschiedenen Bildkategorien**

Zur Ermittlung der Werte wurde eine manuelle Abschätzung durch Beobachtungen vorgenommen, welche Werte sich als untere und obere Grenze eignen können. Bei Erhöhung der Helligkeit wurde beispielsweise ersichtlich, dass größere Schritte als etwa eine Erhöhung um 0,2 nicht sinnvoll für alle Abstufungen wären. Dies basiert wiederum auf der Beobachtung, dass grundsätzlich bei einigen aufeinanderfolgenden Abstufungen große Veränderungen der Ergebnisse zu beobachten sind.

In Folge der Beobachtungen zur Verminderung der Helligkeit fiel unmittelbar auf, dass die Helligkeit drastisch gesenkt werden muss, damit ein Abfall der Erkennungsquote ausgelöst wird. Aus diesem Grund wurde die obere Grenze bereits auf das etwa 0,1-fache des Ausgangswerts festgelegt. Andernfalls wären zahlreiche Abstufungen notwendig, um aussagekräftige Ergebnisse erzielen zu können.

## Frontale Bilder

Die Reihenfolge der Bildkategorien wurde bei veränderten Helligkeitswerten weiter beibehalten. Jedoch wurden die Ergebnisse der erhöhten bzw. verminderten Helligkeit jeweils in einer alleinstehenden Tabelle zusammengefasst. Die Ergebnisse der frontalen Bilder sind für erhöhte Helligkeit Tabelle 7 und entsprechender Abbildung in Anlagen, Teil 8 zu entnehmen. Tabelle 8 und entsprechende Abbildung in Anlagen, Teil 8 stellen die Ergebnisse für diese Bilder mit verminderter Helligkeit dar.

**Tabelle 7: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Erhöhung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 1,76	50	100 %
Helligkeit = 1,86	50	100 %
Helligkeit = 1,87	50	100 %
Helligkeit = 1,88	50	100 %
Helligkeit = 1,89	49	98 %
Helligkeit = 1,90	48	96 %
Helligkeit = 1,91	48	96 %
Helligkeit = 1,92	47	94 %
Helligkeit = 1,93	45	90 %
Helligkeit = 1,94	35	70 %
Helligkeit = 1,95	15	30 %
Helligkeit = 1,96	1	2 %
Helligkeit = 1,98	0	0 %
Durchschnitt	37,538 $\approx$ 37,5	75,077 $\approx$ 75,1 %

Bei Erhöhung der Helligkeit wird von einem bis zu einem 1,86-fachen Wert eine Erkennungsquote von 100 % erreicht und erst ab einem 1,94-fachen Wert sinkt die Erkennungsquote signifikant. Demnach sinkt die Erkennungsquote bei 1,76 bis 1,93-facher gar nicht bzw. gering. Da bei 1,96-facher Helligkeit noch immer ein Gesicht erkannt worden ist, wurde eine Überprüfung bei 1,98-facher Helligkeit vorgenommen. Für diesen Fall wird die untere Grenze der Erkennungsquote von 0 % schließlich erreicht. Die durchschnittliche Erkennungsquote beträgt 75,1 %.

**Tabelle 8: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verminderung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 0,1	50	100 %
Helligkeit = 0,06	49	98 %
Helligkeit = 0,05	49	98 %
Helligkeit = 0,04	48	96 %
Helligkeit = 0,03	43	86 %
Helligkeit = 0,02	12	24 %
Helligkeit = 0,008	0	0 %
Durchschnitt	35,857 $\approx$ 35,9	71,714 $\approx$ 71,7 %

Bei Verringerung der Helligkeit können selbst bei einem auf das 0,06-fach kleineren Helligkeitswert fast alle Gesichter erkannt werden. Auffallend ist dabei, dass bei 0,05- bis 0,03-facher Helligkeit die Erkennungsquote stets bei über 85 % bleibt, bei 0,05 und 0,04-facher sogar über 95 %. Die Differenz in Bezug auf die Erkennungsquote ist zwischen der 0,03- und 0,02-fachen Helligkeit mit 62 % beachtlich hoch.

Hier wird eine Gesamt-Erkennungsquote von ca. 71,7 % erreicht.

## Bilder mit Brille

Auch bei Bildern, worauf die abgebildete Person eine Brille trägt, wurden die entsprechenden Helligkeitseinstellungen in Bezug auf die Erkennung des Gesichts überprüft. Die erreichten Ergebnisse bei erhöhter Helligkeit sind in Tabelle 9 und in entsprechender Abbildung in Anlagen, Teil 8 dargestellt. Tabelle 10 und entsprechender Abbildung in Anlagen, Teil 8 können die Ergebnisse der verminderten Helligkeitseinstellungen entnommen werden.

**Tabelle 9: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Erhöhung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 1,76	50	100 %
Helligkeit = 1,86	50	100 %
Helligkeit = 1,87	50	100 %
Helligkeit = 1,88	49	98 %
Helligkeit = 1,89	49	98 %
Helligkeit = 1,90	48	96 %
Helligkeit = 1,91	46	92 %
Helligkeit = 1,92	43	86 %
Helligkeit = 1,93	29	58 %
Helligkeit = 1,94	23	46 %
Helligkeit = 1,95	8	16 %
Helligkeit = 1,96	0	0 %
Durchschnitt	37,083 $\approx$ 37,1	74,166 $\approx$ 74,2 %

Face Mesh ist ebenfalls in der Lage die Mehrheit der vorliegenden Gesichter mit Brille zu detektieren, bei denen die Helligkeit erhöht worden ist. Erneut fällt auf, dass die Erkennungsquote bei den 1,76 bis 1,86-fachen Helligkeitswerten bei 100 % gehalten werden kann. Die entscheidenden bzw. am meisten aussagenden Ergebniswerte liegen hier bei 1,92- bis 1,96-facher Helligkeit. Am meisten aussagend meint, dass sich diese Werte signifikant von denen der Gesichter ohne Brille unterscheiden. Die durchschnittliche Erkennungsquote beträgt ca. 74,2 % und liegt damit nur knapp unter der der frontal aufgenommenen Gesichter ohne Brille.

**Tabelle 10: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verminderung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 0,1	50	100 %
Helligkeit = 0,06	50	100 %
Helligkeit = 0,05	50	100 %
Helligkeit = 0,04	48	96 %
Helligkeit = 0,03	37	74 %
Helligkeit = 0,02	10	20 %
Helligkeit = 0,008	0	0 %
Durchschnitt	35	70 %

Bei den verminderten Helligkeitswerten können sowohl bei einem 0,1-fachen als auch bei einem 0,06-fachen Helligkeitswert 100 % der dargestellten Gesichter erkannt werden. Bei einem 0,02-fachen Helligkeitswert werden im Gegensatz zu der Kategorie „frontal“ zwei Gesichter weniger erkannt. Gleichzeitig liegen die Ergebniswerte bei 0,05- bis 0,03-fache Helligkeit insgesamt ebenfalls etwas unter den Ergebnissen der Kategorie „frontal“, jedoch stets bei über 70 %. Bei einem 0,05-fachen Helligkeitswert kann das Ergebnis der frontal positionierten Gesichter ohne Brille sogar um 2 % übertroffen werden, sodass alle dargestellten Gesichter erkannt werden. Dabei liegt die durchschnittliche Erkennungsquote bei 70 % und damit 1,7 % unter der der frontalen Bilder ohne Brille.

## Bilder mit Sonnenbrille

Weiter wurden die gleichen Helligkeitseinstellungen bei Bildern mit Sonnenbrille angewendet. Die dabei erreichten Ergebnisse sind für die erhöhte Helligkeit Tabelle 11 sowie entsprechender Abbildung in Anlagen, Teil 8 und für die verminderte Helligkeit Tabelle 12 sowie entsprechender Abbildung in Anlagen, Teil 8 zu entnehmen.

**Tabelle 11: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Erhöhung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 1,76	50	100 %
Helligkeit = 1,86	49	98 %
Helligkeit = 1,87	49	98 %
Helligkeit = 1,88	48	96 %
Helligkeit = 1,89	46	92 %
Helligkeit = 1,90	45	90 %
Helligkeit = 1,91	45	90 %
Helligkeit = 1,92	35	70 %
Helligkeit = 1,93	20	40 %
Helligkeit = 1,94	11	22 %
Helligkeit = 1,95	0	0 %
Helligkeit = 1,96	0	0 %
Durchschnitt	33,166 $\approx$ 33,2	66,333 $\approx$ 66,3 %

Bei Erhöhung der Helligkeit in der Kategorie „mit Sonnenbrille“ können insgesamt gute Ergebnisse im Rahmen von einer durchschnittlich hohen Erkennungsquote erreicht werden. Bis zu einem 1,86-fachen Helligkeitswert kann eine Erkennungsquote von 100 % und bis zu der 1,91-fachen Helligkeit von min. 90 % erreicht werden. Auffallend ist jedoch, dass bereits bei 1,95-facher Helligkeit die untere Grenze von null erkannten Gesichtern erzielt wird. Die durchschnittliche Erkennungsquote beträgt etwa 66,3 % und liegt damit unter der der frontalen Bilder mit und ohne Brille.

**Tabelle 12: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verminderung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 0,1	50	100 %
Helligkeit = 0,06	50	100 %
Helligkeit = 0,05	50	100 %
Helligkeit = 0,04	48	96 %
Helligkeit = 0,03	42	84 %
Helligkeit = 0,02	5	10 %
Helligkeit = 0,008	0	0 %
Durchschnitt	35	70 %

Bei Verminderung der Helligkeit von Bildern aus der Kategorie „mit Sonnenbrille“ wird bis zu der 0,05-fachen Helligkeit eine Erkennungsquote von 100 % erreicht. Besonders auffällig ist auch hier die große Differenz zwischen der 0,03- und 0,02-fachen Helligkeit in Bezug auf die Anzahl der erkannten Gesichter. Die untere Grenze wird bei 0,008-facher Helligkeit erreicht.

Die durchschnittliche Erkennungsquote beträgt 70 %.



## Links laterale Bilder

Wiederholend wurden links und rechts laterale Bilder bzw. Gesichter separat voneinander untersucht. In Tabelle 13 sowie entsprechender Abbildung in Anlagen, Teil 8 sind die Ergebnisse für die links lateralen Gesichter bei erhöhter Helligkeit dargestellt. Die Ergebnisse der links lateralen Bilder bei verminderter Helligkeit sind Tabelle 14 und entsprechender Abbildung in Anlagen, Teil 8 zu entnehmen. Tabelle 15 und die entsprechende Abbildung in Anlagen, Teil 8 zeigen die Ergebnisse der erhöhten Helligkeit bei Bildern, auf denen das Gesicht jeweils rechts lateral zur Kamera positioniert ist. In Tabelle 16 sowie in entsprechender Abbildung in Anlagen, Teil 8 sind die Ergebnisse bezüglich verminderter Helligkeit rechts lateraler Gesichter dargestellt.

**Tabelle 13: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Erhöhung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 1,76	50	100 %
Helligkeit = 1,86	50	100 %
Helligkeit = 1,87	50	100 %
Helligkeit = 1,88	49	98 %
Helligkeit = 1,89	48	96 %
Helligkeit = 1,90	48	96 %
Helligkeit = 1,91	45	90 %
Helligkeit = 1,92	34	68 %
Helligkeit = 1,93	22	44 %
Helligkeit = 1,94	14	28 %
Helligkeit = 1,95	1	2 %
Helligkeit = 1,96	0	0 %
Durchschnitt	30,5	61 %

Die in Tabelle 13 dargestellten Ergebnisse zeigen bis zu einer 1,87-fachen Helligkeit eine Erkennungsquote von 100 %. Im Gegensatz zu den frontalen Gesichtern wird bei 1,88-facher Helligkeit keine Erkennungsquote von 100 % erreicht. Darüber hinaus fällt auf, dass die Anzahl der erkannten Gesichter bei einem 1,92-fachen Helligkeitswert nur 34 und damit

68 % beträgt. Auch die Ergebnisse bei 1,93- bis 1,95-facher Helligkeit fallen um einiges schlechter aus als bei den frontalen positionierten Gesichtern.

Die durchschnittliche Erkennungsquote beträgt 61 %.

**Tabelle 14: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verminderung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 0,2	50	100 %
Helligkeit = 0,18	49	98 %
Helligkeit = 0,16	49	98 %
Helligkeit = 0,14	49	98 %
Helligkeit = 0,12	49	98 %
Helligkeit = 0,1	49	98 %
Helligkeit = 0,06	49	98 %
Helligkeit = 0,05	49	98 %
Helligkeit = 0,04	42	84 %
Helligkeit = 0,03	29	58 %
Helligkeit = 0,02	1	2 %
Helligkeit = 0,008	0	0 %
Durchschnitt	38,75	77,5 %

Im Vergleich zu den anderen Kategorien werden in Tabelle 14 mehr Abstufungen vorgenommen, da bei dem sonst gewählten Anfangswert, entsprechend dem 0,1-fachen des Default-Helligkeitswerts, bereits nicht auf allen Bildern ein Gesicht detektiert werden kann. Bei einer 0,2-fachen Helligkeit wird eine Erkennungsquote von 100 % erreicht. Von dieser Besonderheit abgesehen, werden Ergebnisse mit den erwarteten Tendenzen erzielt.

Die durchschnittliche Erkennungsquote beträgt 77,5 %.

## Rechts laterale Bilder

**Tabelle 15: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Erhöhung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 1,76	50	100 %
Helligkeit = 1,86	48	96 %
Helligkeit = 1,87	48	96 %
Helligkeit = 1,88	48	96 %
Helligkeit = 1,89	47	94 %
Helligkeit = 1,90	47	94 %
Helligkeit = 1,91	47	94 %
Helligkeit = 1,92	45	90 %
Helligkeit = 1,93	30	60 %
Helligkeit = 1,94	20	40 %
Helligkeit = 1,95	5	10 %
Helligkeit = 1,96	1	2 %
Helligkeit = 1,98	0	0 %
Durchschnitt	33,538 $\approx$ 33,5	67,077 $\approx$ 67,1 %

Zu den vorliegenden Ergebnissen in Bezug auf die Gesichtserkennung lässt sich bei dieser Kategorie festhalten, dass bereits ab einer 1,86-fachen Helligkeit nur noch 48 und damit 96 % der Gesichter detektiert werden können. Dieser Aspekt steht im Gegensatz zu den bereits vorgestellten Kategorien „frontal“ und „mit Brille“. Allerdings ist ebenso festzuhalten, dass dennoch erst ab der 1,93-fachen Helligkeit die Erkennungsquote signifikant mit immer höheren Differenzen sinkt. Um auch hier die untere Grenze der Erkennungsquote von 0 % zu erreichen, wurde die Helligkeit ebenfalls auf das 1,98-Fache erhöht.

Die durchschnittliche Erkennungsquote beträgt etwa 67,1 %.

**Tabelle 16: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verminderung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 0,12	50	100 %
Helligkeit = 0,1	49	98 %
Helligkeit = 0,06	49	98 %
Helligkeit = 0,05	48	96 %
Helligkeit = 0,04	45	90 %
Helligkeit = 0,03	29	58 %
Helligkeit = 0,02	1	2 %
Helligkeit = 0,008	0	0 %
Durchschnitt	33,875 $\approx$ 33,9	67,75 %

Im Vergleich zu den ersten zwei präsentierten Kategorien zeigen die Werte eine leicht verschlechterte Erkennungsquote auf. So muss ebenfalls die 0,12-fache Helligkeit überprüft werden, um die obere Grenze von 50 erkannten Gesichtern zu erreichen. Außerdem liegt der Ergebniswert bei 0,03-facher Helligkeit mit 58 % bedeutend unter den Ergebniswerten der vorangegangenen Kategorien – mit Ausnahme der links lateralen Bilder.

Insgesamt wird eine durchschnittliche Erkennungsquote von 67,75 % erreicht.

## Bilder mit teilweiser Verdeckung

Des Weiteren wurden abschließend in diesem Unterkapitel die Bilder mit teilweiser Verdeckung des Gesichts in Bezug auf die Helligkeit verändert und mittels der Gesichtserkennung überprüft, ob ein Gesicht erkannt werden kann. Die Werte bei Erhöhung der Helligkeit sind Tabelle 17 und entsprechender Abbildung in Anlagen, Teil 8 zu entnehmen, die bei Verminderung der Helligkeit Tabelle 18 sowie entsprechender Abbildung in Anlagen, Teil 8.

**Tabelle 17: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen erhöhten Helligkeitswerten von Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Erhöhung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 1,76	50	100 %
Helligkeit = 1,86	50	100 %
Helligkeit = 1,87	50	100 %
Helligkeit = 1,88	50	100 %
Helligkeit = 1,89	50	100 %
Helligkeit = 1,90	48	96 %
Helligkeit = 1,91	43	86 %
Helligkeit = 1,92	34	68 %
Helligkeit = 1,93	25	50 %
Helligkeit = 1,94	15	30 %
Helligkeit = 1,95	3	6 %
Helligkeit = 1,96	3	6 %
Helligkeit = 1,98	0	0 %
Durchschnitt	32,384 $\approx$ 32,4	64,769 $\approx$ 64,8%

Auch die Gesichter, welche teilweise durch einen Gegenstand o. ä. verdeckt sind, können bei veränderten Helligkeitswerten, besonders bei steigender Helligkeit, noch gut detektiert werden. Im direkten Vergleich zu beispielsweise Tabelle 7 fällt allerdings auf, dass bereits bei 1,92-facher Helligkeit ein größerer Abfall der Erkennungsquote zu beobachten ist. Erst an dieser Stelle sinkt die Erkennungsquote, zunächst in immer größer werdenden Differenzen, deutlich. Dennoch kann erst bei 1,98-facher Helligkeit die untere Grenze von 0 % erkannten Gesichtern erreicht werden. Die durchschnittliche Erkennungsquote beträgt etwa 64,8 % und liegt damit etwa 10,3 % unter der der frontalen Bilder.

**Tabelle 18: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen verminderten Helligkeitswerten von Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verminderung auf das x-fache)	Anzahl erkannter Gesichter	Prozentualer Anteil erkannte Gesichter
Helligkeit = 0,1	50	100 %
Helligkeit = 0,06	50	100 %
Helligkeit = 0,05	49	98 %
Helligkeit = 0,04	48	96 %
Helligkeit = 0,03	43	86 %
Helligkeit = 0,02	6	12 %
Helligkeit = 0,008	0	0 %
Durchschnitt	35,143 $\approx$ 35,1	70,286 $\approx$ 70,3 %

Bei Verminderung der Helligkeit können vergleichsweise ähnliche Ergebnisse zu den anderen Bildkategorien erreicht werden. So beträgt die Erkennungsquote bei 0,1- und 0,06-facher Helligkeit 100 % und bei 0,008-facher 0 %. Auch der Ergebniswert von 12 % bei 0,02-facher sowie die Werte bei 0,05-0,03-facher Helligkeit entsprechen etwa der erwarteten Tendenz. Erneut fällt dabei auf, dass, trotz der geringen Differenz von 0,01 der Helligkeitseinstellung zwischen 0,03- und 0,02-facher Helligkeit, der Sprung bei diesen Werten in Bezug auf den Abfall der Erkennungsquote vergleichsweise groß ist.

Die durchschnittliche Erkennungsquote beträgt hier etwa 70,3 %.

### **5.3 Ergebnisse der Gesichtserkennung bei Anwendung Gauß'scher Unschärfe auf das Ausgangsmaterial in Bezug auf die verschiedenen Bildkategorien**

#### **Frontale Bilder**

Zunächst werden auch in dieser Kategorie der Informationsreduktion die frontal positionierten Gesichter untersucht. Die Ergebnisse dafür sind Tabelle 19 und entsprechender Abbildung in Anlagen, Teil 9 zu entnehmen.

**Tabelle 19: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von frontalen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil der erkannten Gesichter
Radius = 5	50	100 %
Radius = 6	50	100 %
Radius = 7	50	100 %
Radius = 8	50	100 %
Radius = 9	50	100 %
Radius = 10	50	100 %
Radius = 12	50	100 %
Radius = 13	49	98 %
Radius = 14	48	96 %
Radius = 15	44	88 %
Radius = 16	41	82 %
Radius = 17	39	78 %
Radius = 18	34	68 %
Radius = 19	31	62 %
Radius = 20	28	56 %
Radius = 25	7	14 %
Radius = 30	2	4 %
Radius = 35	0	0 %
Durchschnitt	37,389 $\approx$ 37,4	74,778 $\approx$ 74,8 %

Die in Tabelle 19 dargestellten Ergebnisse zeigen, dass bei einer Radius-Einstellung von fünf bis 15 mindestens eine Erkennungsquote von 88 % erreicht wird. Konkret bleibt die Erkennungsquote der Einstellungswerte fünf bis zwölf, d. h. über der Verdopplung des Radius, bei 100 %. Auch die weiteren Ergebniswerte zeigen eine langsam sinkende Erkennungsquote. Ab einem Wert von 20 sinkt diese allerdings deutlich rapider, was sich unter anderem durch die größeren Unterschiede in Bezug auf die Einstellungswerte begründen lässt. Der größte Unterschied mit einer Differenz von 42 % liegt zwischen den Werten 20

und 25. Hier wird die untere Grenze von einer Erkennungsquote von 0 % bei der Einstellung 35 erreicht.

Die durchschnittliche Erkennungsquote beträgt 74,8 %.



## Bilder mit Brille

Weiterhin wurden die Bilder mit Gauß'scher Unschärfe verändert, auf denen die abgebildete Person eine Brille trägt. Die entsprechenden Ergebnisse sind in Tabelle 20 und in entsprechender Abbildung in Anlagen, Teil 9 dargestellt.

**Tabelle 20: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von Bildern mit Brille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil der erkannten Gesichter
Radius = 5	50	100 %
Radius = 6	50	100 %
Radius = 7	50	100 %
Radius = 8	49	98 %
Radius = 9	49	98 %
Radius = 10	48	96 %
Radius = 12	43	86 %
Radius = 13	36	72 %
Radius = 14	27	54 %
Radius = 15	23	46 %
Radius = 16	18	36 %
Radius = 17	16	32 %
Radius = 18	13	26 %
Radius = 19	11	22 %
Radius = 20	11	22 %
Radius = 25	2	4 %
Radius = 30	0	0 %
Durchschnitt	29,176 $\approx$ 29,2	58,353 $\approx$ 58,4 %

Die hier dargestellten Ergebnisse zeigen mehrere Auffälligkeiten bei dem Vergleich mit den Ergebnissen aus Tabelle 19. Zum einen werden bereits bei einer Radius-Einstellung von acht nicht mehr 100 % der Gesichter erkannt. Zum anderen ist die Entwicklung der

Erkennungsquote ins Negative deutlich stärker, beispielsweise bei den Einstellungen zwischen zwölf und 15 als bei frontal positionierten Gesichtern ohne Brille. Die Anzahl der erkannten Gesichter halbiert sich dabei fast. Weiterhin wird bereits bei einem Radius von 30 die untere Grenze von null erkannten Gesichtern erreicht.

Im Durchschnitt liegt die Erkennungsquote für die vorliegenden Ergebnisse bei etwa 58,4 %.

## Bilder mit Sonnenbrille

Fortführend wurden die Ergebnisse der Bilder mit Sonnenbrille, siehe Tabelle 21 und entsprechende Abbildung in Anlagen, Teil 9, in Bezug auf die Veränderung mittels Gauß'scher Unschärfe überprüft.

**Tabelle 21: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von Bildern mit Sonnenbrille in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil der erkannten Gesichter
Radius = 5	50	100 %
Radius = 6	50	100 %
Radius = 7	50	100 %
Radius = 8	49	98 %
Radius = 9	49	98 %
Radius = 10	49	98 %
Radius = 12	43	86 %
Radius = 13	40	80 %
Radius = 14	39	78 %
Radius = 15	35	70 %
Radius = 16	33	66 %
Radius = 17	30	60 %
Radius = 18	30	60 %
Radius = 19	28	56 %
Radius = 20	25	50 %
Radius = 25	15	30 %
Radius = 30	7	14 %
Radius = 35	2	4 %
Radius = 40	0	0 %
Durchschnitt	32,842 $\approx$ 32,8	65,684 $\approx$ 65,7 %

Besonders auffällig ist bei diesen dargestellten Ergebnissen, dass nicht nur eine, sondern zwei weitere Abstufungen notwendig waren, um die untere Grenze von null erkannten Gesichtern zu erreichen. Bei einer Einstellung mit dem Wert 30 ist darüber hinaus ein deutlich höherer Erkennungswert als null oder auch zwei, wie es beispielsweise bei den frontalen Bildern der Fall war, zu erkennen. Im weiteren Vergleich zu frontalen Bildern mit und ohne Brille ist zu beobachten, dass bei einer Einstellung von 15 in etwa ein Mittelwert dieser Bildkategorien in Bezug auf die Anzahl der erkannten Gesichter erreicht wird. Damit liegt dieser Ergebniswert also auch deutlich über dem der Bildkategorie „mit Brille“.

Die durchschnittliche Erkennungsquote der Ergebnisse aus Tabelle 21 beläuft sich auf ca. 65,7 %.

## Links laterale Bilder

Sowohl links als auch rechts laterale Bilder bzw. Gesichter wurden mittels Gauß'scher Unschärfe verschlechtert. Die Ergebnisse der links lateralen Gesichter sind in Tabelle 22 und in entsprechender Abbildung in Anlagen, Teil 9 dargestellt, die Ergebnisse der rechts lateralen Bilder bzw. Gesichter in Tabelle 23 sowie in entsprechender Abbildung in Anlagen, Teil 9.

**Tabelle 22: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von links lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil der erkannten Gesichter
Radius = 5	50	100 %
Radius = 6	50	100 %
Radius = 7	49	98 %
Radius = 8	47	94 %
Radius = 9	47	94 %
Radius = 10	47	94 %
Radius = 12	40	80 %
Radius = 13	34	68 %
Radius = 14	29	58 %
Radius = 15	24	48 %
Radius = 16	18	36 %
Radius = 17	15	30 %
Radius = 18	13	26 %
Radius = 19	11	22 %
Radius = 20	12	24 %
Radius = 25	5	10 %
Radius = 30	0	0 %
Durchschnitt	28,882 $\approx$ 28,9	57,765 $\approx$ 57,8 %

Bei links lateralen Bildern wird in den ersten zwei Abstufungen jeweils eine Erkennungsrate von 100 % erreicht. Allerdings liegt bereits bei einem Einstellungswert von 8 nur noch eine Erkennungsquote von 94 % vor. Die weiteren Ergebnisse zeigen eine nicht bedeutend schnell sinkende Erkennungsquote. Auffällig sind die Ergebnisse bei den Werten 19 und 20, da bei einem Wert von 20 im Vergleich zu 19 ein Gesicht mehr erkannt wird. Die untere Grenze wird bei einer Einstellung von 30 erreicht.

Die durchschnittliche Erkennungsquote liegt bei 57,8 %.

## Rechts laterale Bilder

**Tabelle 23: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von rechts lateralen Bildern in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil der erkannten Gesichter
Radius = 5	50	100 %
Radius = 6	50	100 %
Radius = 7	50	100 %
Radius = 8	49	98 %
Radius = 9	48	96 %
Radius = 10	47	94 %
Radius = 12	44	88 %
Radius = 13	41	82 %
Radius = 14	39	78 %
Radius = 15	33	66 %
Radius = 16	27	54 %
Radius = 17	18	36 %
Radius = 18	12	24 %
Radius = 19	10	20 %
Radius = 20	9	18 %
Radius = 25	3	6 %
Radius = 30	1	2 %
Radius = 35	0	0 %
Durchschnitt	29,5	59 %

Grundsätzlich weisen die Ergebnisse aus Tabelle 23 in etwa die erwarteten Tendenzen auf. Trotzdem fallen zwei Aspekte auf: zum einen, dass erstmals bei einem Radius von zehn eine Anzahl von unter 48 erkannten Gesichtern vorliegt. Zum anderen fällt auf, dass auch hier eine weitere Abstufung überprüft werden musste, um die untere Grenze zu

erreichen. Der Abfall von 20 auf 25 fällt mit 6 % Unterschied in der Anzahl der erkannten Gesichter wesentlich geringer aus als bei den vorher vorgestellten Kategorien.

Durchschnittlich beträgt die Erkennungsquote 59 %.



## Bilder mit teilweiser Verdeckung

Abschließend wird die Gauß'sche Unschärfe ebenfalls auf Bilder bzw. Gesichter mit teilweiser Verdeckung angewendet. Tabelle 24 und entsprechende Abbildung in Anlagen, Teil 9 zeigen die Ergebnisse.

**Tabelle 24: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen Gauß'scher Unschärfe von Bildern mit teilweiser Verdeckung in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung	Anzahl erkannter Gesichter	Prozentualer Anteil der erkannten Gesichter
Radius = 5	50	100 %
Radius = 6	48	96 %
Radius = 7	48	96 %
Radius = 8	46	92 %
Radius = 9	40	80 %
Radius = 10	40	80 %
Radius = 12	34	68 %
Radius = 13	29	58 %
Radius = 14	25	50 %
Radius = 15	22	44 %
Radius = 16	16	32 %
Radius = 17	13	26 %
Radius = 18	10	20 %
Radius = 19	8	16 %
Radius = 20	7	14 %
Radius = 25	1	2 %
Radius = 30	1	2 %
Radius = 35	0	0 %
Durchschnitt	24,333 $\approx$ 24,3	48,667 $\approx$ 48,7 %

In dieser Kategorie fällt der Wert bei einer Einstellung von neun auf ein neues Minimum mit einer Erkennungsquote von 80 %. Bereits ab einer Einstellung von sechs wird keine Erkennungsquote von 100 % mehr erreicht. Weiterhin fällt auf, dass trotz der insgesamt schlechteren Ergebnisse eine weitere Abstufung von 35 notwendig ist, um die untere Grenze zu erreichen. Grundsätzlich weist die Erkennungsquote hier allerdings keine besonders großen Differenzen zwischen den einzelnen Abstufungen auf.

Trotz dieser auffallenden Werte wird dennoch eine weitere Abstufung benötigt, um an die untere Grenze zu gelangen.

Die durchschnittliche Erkennungsquote beträgt etwa 48,7 %.

## 5.4 Ergebnisse der Gesichtserkennung bei Anwendung von Interpolation auf das Ausgangsmaterial in Bezug auf die verschiedenen Bildkategorien

### Frontale Bilder

Beginnend in dieser Kategorie werden die Ergebnisse der frontal ausgerichteten Gesichter in Tabelle 25 und in entsprechender Abbildung in Anlagen, Teil 10 dargestellt.

**Tabelle 25: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf frontalen Bildern angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verkleinerung auf und Vergrößerung auf)	Anzahl erkannte Gesichter	Prozentualer Anteil der erkannten Gesichter
128x128 auf 512x512 Pixel	50	100 %
64x64 auf 512x512 Pixel	50	100 %
32x32 auf 512x512 Pixel	50	100 %
30x30 auf 512x512 Pixel	50	100 %
21x21 auf 512x512 Pixel	42	84 %
19x19 auf 512x512 Pixel	32	64 %
18x18 auf 512x512 Pixel	43	86 %
17x17 auf 512x512 Pixel	23	46 %
16x16 auf 512x512 Pixel	39	78 %
15x15 auf 512x512 Pixel	18	36 %
14x14 auf 512x512 Pixel	36	72 %
13x13 auf 512x512 Pixel	11	22 %
8x8 auf 512x512 Pixel	0	0 %
Durchschnitt	34,154 $\approx$ 34,2	68,308 $\approx$ 68,3 %

Bei den in Tabelle 25 dargestellten Ergebnissen ist zunächst festzustellen, dass bis zu einer Verkleinerung auf 30x30 Pixel und anschließender Vergrößerung stets eine Erkennungsquote von 100 % erreicht wird. In den nächsten zwei Abstufungen sinkt die Erkennungsquote zunächst weiter, jedoch schwankt sie danach stark bis schließlich die untere Grenze von 0 % bei Verkleinerung auf 8x8 Pixel und anschließender Vergrößerung erreicht wird. Die durchschnittliche Erkennungsquote beträgt etwa 68,3 %.

## Bilder mit Brille

Als nächstes werden die Ergebnisse der verkleinerten und wieder vergrößerten Bilder mit Brille in Tabelle 26 und in entsprechender Abbildung in Anlagen, Teil 10 dargestellt.

**Tabelle 26: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf Bildern mit Brille angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verkleinerung auf und Vergrößerung auf)	Anzahl erkannte Gesichter	Prozentualer Anteil der erkannten Gesichter
128x128 auf 512x512 Pixel	50	100 %
64x64 auf 512x512 Pixel	50	100 %
32x32 auf 512x512 Pixel	48	96 %
30x30 auf 512x512 Pixel	47	94 %
21x21 auf 512x512 Pixel	35	70 %
19x19 auf 512x512 Pixel	25	50 %
18x18 auf 512x512 Pixel	25	50 %
17x17 auf 512x512 Pixel	20	40 %
16x16 auf 512x512 Pixel	21	42 %
15x15 auf 512x512 Pixel	11	22 %
14x14 auf 512x512 Pixel	14	28 %
13x13 auf 512x512 Pixel	9	18 %
8x8 auf 512x512 Pixel	3	6 %
4x4 auf 512x512 Pixel	0	0 %
Durchschnitt	25,571 $\approx$ 25,6	51,143 $\approx$ 51,1 %

Bei diesen Ergebnissen fällt im Vergleich zu den frontalen Bildern ohne Brille unmittelbar auf, dass die Erkennungsquote in einem schnelleren Tempo sinkt. So kann beispielsweise bereits bei einer Verkleinerung auf 32x32 Pixeln und anschließender Vergrößerung keine Erkennungsquote mehr von 100 %, sondern nur von 96 % erreicht werden. Auch bei einer Verkleinerung auf 16x16 Pixel ist ein bedeutender Unterschied zu beobachten, da hier nur 42 % der Gesichter erkannt werden. Trotz der insgesamt schlechteren Ergebnisse werden bei der vermeintlich unteren Grenze noch 6 % der Gesichter erkannt. Um die untere Grenze zu erreichen, wurde eine weitere Abstufung vorgenommen. Dabei wird eine Erkennungs-

quote von 0 % erreicht. Grundsätzlich schwankt bzw. steigt die Erkennungsquote hier nur bei Verkleinerung auf 14x14 Pixel und anschließender Vergrößerung.

Die durchschnittliche Erkennungsquote beträgt etwa 51,1 %.

### Bilder mit Sonnenbrille

Im nächsten Schritt werden die Ergebnisse der verkleinerten und anschließend vergrößerten Bilder mit Sonnenbrille, siehe Tabelle 27 und entsprechende Abbildung in Anlagen, Teil 10, präsentiert.

**Tabelle 27: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf Bildern mit Sonnenbrille angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verkleinerung auf und Vergrößerung auf)	Anzahl erkannte Gesichter	Prozentualer Anteil der erkannten Gesichter
128x128 auf 512x512 Pixel	50	100 %
64x64 auf 512x512 Pixel	50	100 %
32x32 auf 512x512 Pixel	48	96 %
30x30 auf 512x512 Pixel	48	96 %
21x21 auf 512x512 Pixel	36	72 %
19x19 auf 512x512 Pixel	33	66 %
18x18 auf 512x512 Pixel	22	44 %
17x17 auf 512x512 Pixel	23	46 %
16x16 auf 512x512 Pixel	15	30 %
15x15 auf 512x512 Pixel	19	38 %
14x14 auf 512x512 Pixel	10	20 %
13x13 auf 512x512 Pixel	15	30 %
8x8 auf 512x512 Pixel	2	4 %
4x4 auf 512x512 Pixel	0	0 %
Durchschnitt	26,5	53 %

Tabelle 27 beinhaltet ähnliche Ergebnisse wie Tabelle 26. Dies war zu erwarten, da beide Kategorien Landmarken der Augen beeinflussen können. Die Ergebnisse der ersten drei Abstufungen gleichen denen der Ergebnisse der Bilder mit Brille. Allerdings fallen bei der Kategorie „mit Sonnenbrille“ mehrere Sprünge in der Erkennungsquote auf. Außerdem

werden bei einer Verkleinerung auf 8x8 Pixel zwei Gesichter mittels Face Mesh erkannt. Aus diesem Grund wurde eine weitere Abstufung mit einer Verkleinerung auf 4x4 Pixel vorgenommen.

Die durchschnittliche Erkennungsquote beträgt 53 %.

### Links laterale Bilder

Auch die lateral positionierten Gesichter wurden in Bezug auf die vorliegende Verschlechterung untersucht. Die Ergebnisse für die links lateralen Gesichter sind Tabelle 28 und entsprechender Abbildung in Anlagen, Teil 10 zu entnehmen. In Tabelle 29 sowie in entsprechender Abbildung in Anlagen, Teil 10 sind die Ergebnisse der rechts lateralen Gesichter einsehbar.

**Tabelle 28: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf links lateralen Bildern angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verkleinerung auf und Vergrößerung auf)	Anzahl erkannte Gesichter	Prozentualer Anteil der erkannten Gesichter
128x128 auf 512x512 Pixel	50	100 %
64x64 auf 512x512 Pixel	50	100 %
32x32 auf 512x512 Pixel	49	98 %
30x30 auf 512x512 Pixel	48	96 %
21x21 auf 512x512 Pixel	38	76 %
19x19 auf 512x512 Pixel	24	48 %
18x18 auf 512x512 Pixel	36	72 %
17x17 auf 512x512 Pixel	21	42 %
16x16 auf 512x512 Pixel	32	64 %
15x15 auf 512x512 Pixel	10	20 %
14x14 auf 512x512 Pixel	20	40 %
13x13 auf 512x512 Pixel	2	4 %
8x8 auf 512x512 Pixel	0	0 %
Durchschnitt	29,231 $\approx$ 29,2	58,462 $\approx$ 58,5 %

Im Vergleich zu der Kategorie „mit Brille“ scheint es keine zum Großteil einheitliche Tendenz in Bezug auf die Erkennungsquoten zu geben. Zudem sind auch hier mehrere

Sprünge zu beobachten. So sinkt z. B. bei einer Verkleinerung auf 19x19 Pixel und anschließender Vergrößerung die Erkennungsquote bereits auf 48 % und steigt bei der nächsten Abstufung auf 72 %. Schließlich kann bei einer Verkleinerung auf 8x8 Pixel die untere Grenze erreicht werden.

Die durchschnittliche Erkennungsquote beträgt etwa 58,5 %.

## Rechts laterale Bilder

**Tabelle 29: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf rechts lateralen Bildern angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verkleinerung auf und Vergrößerung auf)	Anzahl erkannte Gesichter	Prozentualer Anteil der erkannten Gesichter
128x128 auf 512x512 Pixel	50	100 %
64x64 auf 512x512 Pixel	50	100 %
32x32 auf 512x512 Pixel	50	100 %
30x30 auf 512x512 Pixel	48	86 %
21x21 auf 512x512 Pixel	38	76 %
19x19 auf 512x512 Pixel	30	60 %
18x18 auf 512x512 Pixel	35	70 %
17x17 auf 512x512 Pixel	22	44 %
16x16 auf 512x512 Pixel	28	56 %
15x15 auf 512x512 Pixel	12	24 %
14x14 auf 512x512 Pixel	23	46 %
13x13 auf 512x512 Pixel	3	6 %
8x8 auf 512x512 Pixel	1	2 %
4x4 auf 512x512 Pixel	0	0 %
Durchschnitt	27,857 $\approx$ 27,9	55,714 $\approx$ 55,7 %

Auffällig ist bei den vorliegenden Ergebnissen, dass bei den ersten drei Abstufungen eine Erkennungsquote von 100 % erreicht wird. Allerdings beträgt die Differenz zwischen der nächsten und übernächsten Abstufung 10 %. Darüber hinaus sind auch für diese Kategorie teilweise Sprünge in der Erkennungsquote zu verzeichnen. Zudem musste erneut die

Abstufung auf 4x4 Pixel vorgenommen werden, um die untere Grenze von 0 % zu erreichen. Die durchschnittliche Erkennungsquote beträgt etwa 55,7 %.

### Bilder mit teilweiser Verdeckung

Zuletzt werden die Ergebnisse der auf die Gesichter mit teilweiser Verdeckung angewendeten Interpolation in Tabelle 30 und in entsprechender Abbildung in Anlagen, Teil 10 dargestellt.

**Tabelle 30: Darstellung der Ergebnisse der Gesichtserkennung bei in verschiedenen Abstufungen auf Bildern mit teilweiser Verdeckung angewendeter Interpolation in Bezug auf die absolute und relative Anzahl erkannter Gesichter**

Einstellung (Verkleinerung auf und Vergrößerung auf)	Anzahl erkannte Gesichter	Prozentualer Anteil der erkannten Gesichter
128x128 auf 512x512 Pixel	50	100 %
64x64 auf 512x512 Pixel	50	100 %
32x32 auf 512x512 Pixel	45	90 %
30x30 auf 512x512 Pixel	40	80 %
21x21 auf 512x512 Pixel	22	44 %
19x19 auf 512x512 Pixel	12	24 %
18x18 auf 512x512 Pixel	18	36 %
17x17 auf 512x512 Pixel	13	26 %
16x16 auf 512x512 Pixel	16	32 %
15x15 auf 512x512 Pixel	8	16 %
14x14 auf 512x512 Pixel	9	18 %
13x13 auf 512x512 Pixel	3	6 %
8x8 auf 512x512 Pixel	1	2 %
4x4 auf 512x512 Pixel	0	0 %
Durchschnitt	20,5	41 %

Tabelle 30 legt nahe, dass die Ergebnisse dieser Kategorie mit am schlechtesten sind. Zum einen liegt bei der dritten Abstufung nur noch eine Erkennungsquote von 90 % vor. Zum anderen zeigen die Abstufungen einer Verkleinerung auf 15x15 und 14x14 Pixel jeweils eine geringere Anzahl an erkannten Gesichtern. Ebenfalls sind zwei Sprünge der Erkennungsquote bei Verkleinerung auf 18x18 und auf 16x16 Pixel zu erkennen. Trotz dieser im



Vergleich schlechten Werte wurde die Abstufung 4x4 Pixel überprüft, da bei 8x8 Pixeln vermeintlich immer noch ein Gesicht erkannt wird.

Die durchschnittliche Erkennungsquote beträgt 41 %.

## 6 Diskussion der erzielten Ergebnisse

Im folgenden Kapitel werden die in Kapitel 5 dargelegten Ergebnisse der Gesichtserkennung in Bezug auf verschiedene Informationsreduktions-Methoden und Bildkategorien im Rahmen einer Diskussion ausgewertet. Des Weiteren werden diese Ergebnisse, sofern möglich, in einen Kontext mit vergleichbaren Arbeiten gestellt.

Auch in diesem Kapitel betragen die in den Abbildungen veranschaulichten Bilder zum Zweck der besseren Lesbarkeit etwa ein Drittel ihrer eigentlichen Größe.

### 6.1 Auswertung der Ergebnisse bei Verpixelung

Vor der Auswertung der einzelnen Ergebnisse jeder Bildkategorie ist anzumerken, dass es zunächst den Anschein hat, dass Zheng et al. [ZhRa2010] deutlich bessere Erkennungsraten bei vergleichbarer Pixelanzahl erreichen konnten. Demnach wäre MediaPipe Face Mesh bei einer veränderte Pixelanzahl der Bilder nicht so robust wie beispielsweise die Variation des Viola-Jones Detectors unter Zunahme von Color Census Transform [ZhRa2010]. Dabei ist allerdings zu beachten, dass unterschiedliche Datensätze ([Ro2020] in der vorliegenden Arbeit und [GTFD2022] bei Zheng et al. [ZhRa2010]) verwendet wurden und deswegen ein Vergleich der Ergebnisse nur eingeschränkt möglich ist. Zu bedenken ist auch, dass Zheng et al. [ZhRa2010] sich mit ihrer Pixelzahlangabe lediglich auf die das Gesicht repräsentierenden Pixel und nicht, wie in der vorliegenden Arbeit, auf die Gesamtpixelzahl des Bildes beziehen. Auch aus diesem Grund ist ein direkter Vergleich nicht möglich. Allerdings kann aufgrund der Tendenzen vermutet werden, dass der verwendete Algorithmus bei Zheng et al. [ZhRa2010] bei frontalen Bildern ähnliche Ergebnisse wie Face Mesh erzielt. Diese Annahme beruht u. a. darauf, dass bei einer manuell grob geschätzten Anzahl der gesichtsrepräsentierenden Pixel von 16x16 aus der Kategorie „frontal“ (mit Reduzierung auf insgesamt 32x32 Pixel) 98 % der Gesichter erkannt werden konnten, was keinen signifikanten Unterschied zu der Detektionsrate von 97,2 % bei Zheng et al. [ZhRa2010] bedeutet.

Dem gegenüber wurden bei lateralen Bildern bei dieser Pixelanzahl von Face Mesh nur 84 % bzw. 82 % der dargestellten Gesichter erkannt, womit die Detektionsraten deutlich unter denen aus Zheng et al. [ZhRa2010] liegt. Andere Kategorien sind für einen Vergleich allerdings ungeeignet, da bei Zheng et al. [ZhRa2010] nur frontal und lateral positionierte Gesichter als Ausgangsmaterial dienen.

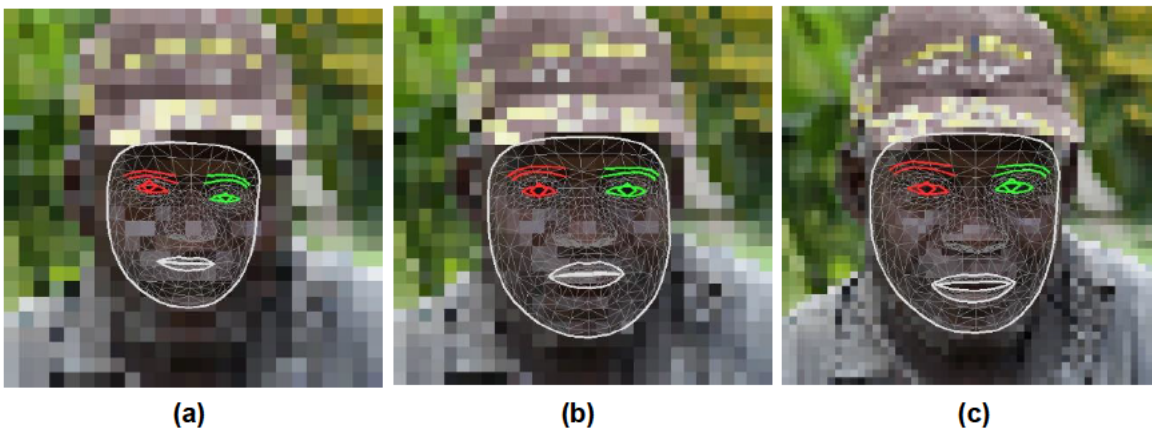
#### Frontale Bilder

Angesichts einer Erkennungsrate von mindestens 94 % bis zu einer Pixelanzahl von 25x25 entsteht zunächst der Eindruck, dass die Gesichtserkennung bei frontalen Bildern robust

ist. Es müsste allerdings ergänzend getestet werden, wie gut sich die erkannten Landmarken zur forensischen Personenidentifikation eignen. Für einen derartigen Test müsste ein Algorithmus zur Gesichtsidifikation angewendet werden, der anschließend die Zuverlässigkeit der Ergebnisse aufzeigen soll.

Weiterhin ist zu beachten, dass es sich bei dem erkannten Bild bei den Abstufungen 14x14 bis 16x16 Pixel jeweils um ein anderes handelt. Aus diesem Grund wurden die vorliegenden Ergebnisbilder manuell betrachtet. Dabei konnte kein offensichtlicher Fehler erkannt werden, der u. a. die Erkennung bei einer so geringen Pixelanzahl wie 14x14 erklären würde. Daher hat es den Anschein, dass Face Mesh auf diesen Bildern das Gesicht korrekt erkennt. Für eine genaue Überprüfung wäre allerdings eine objektive Bestimmung der Korrektheit der Gesichtsnetze erforderlich. Eine mögliche Methode dafür wird im Ausblick vorgestellt.

Zur Auswertung der Beobachtung, dass die Erkennungsquote zwischen einer Pixelanzahl von 32x32 und 30x30 steigt, wird im Folgenden das nicht erkannte Bild bzw. das jeweilige Gesichtsnetz in mehreren Abstufungen, siehe Abbildung 4, im Vergleich untersucht. Eine eindeutige Ursache konnte dabei jedoch nicht identifiziert werden.



**Abbildung 4:** Darstellung der Ergebnisbilder mit bei einer Pixelanzahl von (a) 25x25 (b) 30x30 und (c) 50x50 erkannten Gesichtsnetzen der Kategorie „frontal“. Bild (a) entspricht dem Eingabebild 35\_frontal\_pixel\_25\_25.png, Bild (b) dem Eingabebild 35\_frontal\_pixel\_30\_30.png und Bild (c) dem Eingabebild 35\_frontal\_pixel\_50\_50.png

Allerdings kann festgestellt werden, dass das erkannte Gesichtsnetz bei 25x25 Pixeln (a) deutlich kleiner ist als die anderen erkannten Gesichtsnetze. Außerdem ist das Gesichtsnetz an dieser Stelle mehr nach rechts geneigt als beispielsweise das auf Bild (b). Auf dem Bild (b) ist das Gesichtsnetz wiederum weiter nach links geneigt als das auf Bild (c). Weiterhin hat es den Anschein, dass dieses Gesichtsnetz zwar größer ist als das Gesichtsnetz auf Bild (a), dennoch minimal kleiner ist als das auf Bild (c).

Abschließend lässt sich sagen, dass die Ergebnisse von Face Mesh bei dem vorliegenden Bild grundsätzlich starke Schwankungen bezüglich des Gesichtsnetzes aufweisen, was auch eine mögliche Ursache für die Nicht-Erkennung bei 32x32 Pixel darstellt. Auch ist zu vermuten, dass die erkannten Gesichtsnetze, hauptsächlich bei 25x25 und 32x32 Pixeln, nicht das eigentliche Gesicht repräsentieren, sondern eine andere Pixelkombination von

Face Mesh als Gesicht erkannt wird. Durch die veränderte Pixelanzahl mittels Interpolation werden die Pixel so unterschiedlich kombiniert, dass diese Pixelkombination nicht als Gesicht interpretiert wird.

### Bilder mit Brille

Bei Überprüfung der erkannten Gesichtsnetze bei einer Abstufung von 16x16 Pixeln ist zunächst anzunehmen, dass die Gesichter nicht fälschlicherweise (falsch positiv) erkannt wurden. Begründen lässt sich diese Annahme damit, dass die Gesichtsnetze, siehe Abbildung 6, jeweils an der Stelle auf dem Bild liegen, wo ein menschlicher Betrachter in etwa das Gesicht lokalisieren würde. Im direkten Vergleich zu den Gesichtsnetzen auf den Ausgangsbildern, siehe Abbildung 5, ist festzuhalten, dass die Gesichtsnetze zwar voneinander abweichen, aber die Positionen der Gesichtsnetze auf den verpixelten Bildern noch zum eigentlichen Gesicht zu passen scheinen, die Differenz also nicht signifikant hoch ist. Zugleich entsteht aber auch hier der Eindruck, dass von Face Mesh Pixelkombinationen als Gesicht erkannt werden, die nicht das eigentliche Gesicht repräsentieren. Demnach ist bereits an dieser Stelle festzuhalten, dass die in Kapitel 5 dargestellten Erkennungsquoten nicht akkurat wiedergeben, wie effizient bzw. korrekt Face Mesh die abgebildeten Gesichter wirklich erkennt.

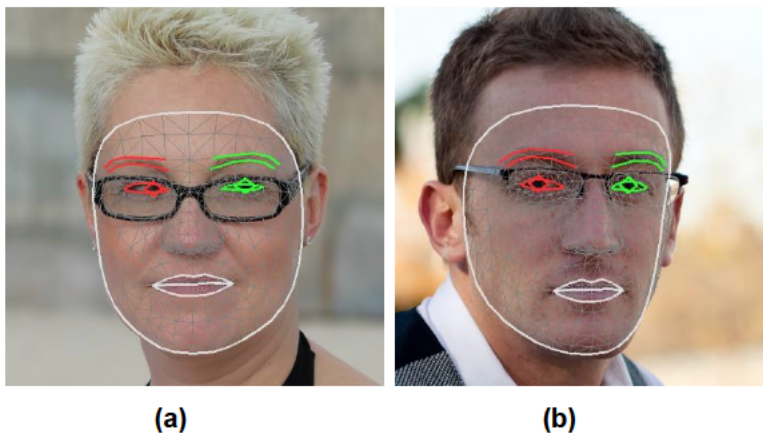
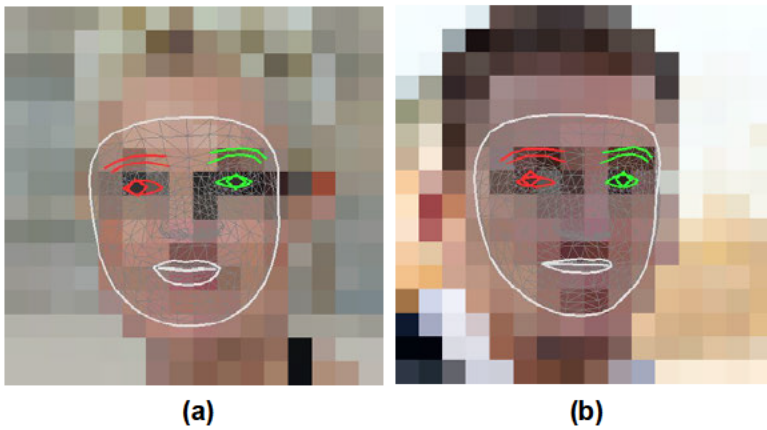


Abbildung 5: Darstellung der Ergebnisbilder mit erkannten Gesichtsnetzen der Kategorie „mit Brille“. Bild (a) entspricht dem Eingabebild 8\_brille.png, Bild (b) dem Eingabebild 26\_brille.png



**Abbildung 6:** Darstellung der in der Pixelanzahl veränderten Ergebnisbilder bei 16x16 Pixeln mit erkannten Gesichtsnetzen der Kategorie „mit Brille“. Bild (a) entspricht dem Eingabebild 8\_brille\_pixel\_16\_16.png, Bild (b) dem Eingabebild 26\_brille\_pixel\_16\_16.png

### **Bilder mit Sonnenbrille**

Die initial aufgestellte Vermutung, dass eine Sonnenbrille einen größeren Einfluss auf die Gesichtserkennung hat als eine Brille, lässt sich größtenteils bestätigen. Unter anderem lässt sich dies auf die Tatsache stützen, dass bereits bei einer Pixelanzahl von 30x30 eine Erkennungsquote von nur knapp unter 60 % erreicht wird. Die gleiche Quote von 58 % wird beispielsweise bei Bildern mit Brille, siehe Tabelle 2, erst bei 25x25 Pixeln erreicht.

Zu den jeweils bei 14x14, 15x15 und 16x16 Pixeln erkannten Gesichtern lässt sich sagen, dass es sich für 15x15 und 16x16 Pixeln um dasselbe Bild handelt, auf dem ein Gesicht erkannt wird. Ebenfalls wird dieses in den nachfolgenden Kategorien erkannt, woraus sich schließen lässt, dass diese Erkennung sich gezielt auf das abgebildete Gesicht bezieht. Bei der Gesichtserkennung bei 14x14 Pixeln ist dies allerdings nicht der Fall. Auf diesem Bild wird nur an dieser Stelle und bei einer Anzahl von 21x21 Pixeln ein Gesicht erkannt. Deswegen ist wiederum zu vermuten, dass es sich bei 14x14 Pixeln um ein erkanntes Muster handelt, welches nicht auf das eigentliche Gesicht zurückzuführen ist.

### **Links laterale Bilder**

Wie bereits erwähnt scheint die Erkennungsquote bei einer Pixelanzahl von 32x32 bis 29x29 Pixel zu stagnieren. Überprüft man allerdings, um welche Bilder es sich dabei jeweils handelt, fällt auf, dass diese bei keiner der Abstufungen vollständig übereinstimmen. Das bedeutet, dass zwar dieselbe Erkennungsquote erreicht, jedoch stets eine andere Kombination aus Bildern bzw. Gesichtern erkannt wird. Auch eine solche Entwicklung bestätigt die Vermutung, dass ausschlaggebende Pixel sich grundsätzlich mit jeder Abstufung verändern, was wiederum die Erkennung des Gesichts erheblich beeinflusst. Eine derartige Veränderung der Pixel ist dabei auf die Interpolation zurückzuführen, durch welche Pixel zusammengefasst werden. Damit setzen sich beispielsweise auch die Farben eines neuen Pixels zusammen.

Die Beobachtung, dass das in Abbildung 2 und Abbildung 3 dargestellte Bild bzw. Gesicht erst bei einer Pixelanzahl von 80x80 erkannt wird, ließe sich beispielsweise mit der vollständig lateralen Position des Gesichts erklären. Dabei scheint eine höhere Qualität des Bildes notwendig zu sein.

### **Rechts laterale Bilder**

Auch wenn die Erkennungsquote hier zwischen 19x19 und 18x18 Pixeln nicht steigt, weicht diese Stufe nicht vollständig von den anderen Kategorien ab. Damit ist gemeint, dass bei Überprüfung der detektierten Gesichter bzw. Gesichtsnetze auffällt, dass keines der Gesichter, welches bei 18x18 Pixeln erkannt worden ist, bei 19x19 Pixeln ebenfalls erkannt wurde. Dieser Aspekt führt zu der Schlussfolgerung, dass zwar mehr „neue“ Gesichter erkannt wurden, gleichzeitig aber der Schritt zwischen den zwei Abstufungen ein Abfall der Erkennungsquote mit sich bringt.

Ein, wenn auch minimaler, Abfall bei einer Pixelanzahl von 32x32 ist besonders auffällig, da dies auch bei frontalen Bildern der Fall war. Betrachtet man an dieser Stelle auf welchen Bildern ein Gesicht erkannt worden ist und auf welchen nicht, fällt auf, dass nur vier der jeweils acht (bei 30x30 Pixeln) bzw. neun (bei 32x32 Pixel) nicht erkannten Gesichter übereinstimmen. Dies führt zu der Annahme, dass bei den anderen neun Gesichtern sich besonders die die Augen- und Mundpartie darstellenden Pixel verändern bzw. ausschlaggebend sind.

### **Bilder mit teilweiser Verdeckung**

Insgesamt entsteht hier der Eindruck, dass die Gesichtserkennung bei den vorliegenden Bildern schlechter performed als bei den vorausgegangenen Versuchen, mit Ausnahme der Kategorie „mit Sonnenbrille“. Dieser Eindruck wird durch eine durchschnittliche Erkennungsquote von etwa 46,5 % bestätigt.

Durch diesen Gesamteindruck kann die Erkennungsquote von 0 % bei 19x19 Pixeln teilweise erklärt werden, da dieser den Anschein erweckt, dass eine höhere Qualität der Bilder vorausgesetzt wird, um ähnlich hohe Ergebniswerte wie in den anderen Kategorien zu erreichen.

Als Ursache für den wiederholten Abfall der Erkennungsquote bei 32x32 Pixeln können die zuvor dargelegten Begründungen bei anderen Kategorien angeführt werden.

## 6.2 Auswertung der Ergebnisse bei veränderter Helligkeit

Vor der Auswertung der Ergebnisse bei veränderter Helligkeit der einzelnen Bildkategorien wird ein kategorienübergreifender Vergleich mit den Ergebnissen von Zhou et al. [ZhLi2018] versucht. Dabei ist allerdings zu beachten, dass bei Zhou et al. [ZhLi2018] an dieser Stelle nicht nur die Helligkeit, sondern gleichzeitig auch der Kontrast eines Bildes verringert worden ist, indem die Pixelwerte in bestimmten Verhältnissen (0,8, 0,5 und 0,2) neuskaliert wurden. [ZhLi2018] Ein direkter Vergleich mit erhöhter Helligkeit in Bezug auf die Robustheit bzw. der Limitierungen von Face Mesh ist nicht möglich.

Ein Vergleich bei verringerten Helligkeitswerten kann nur abgeschätzt werden. Festzuhalten ist, dass Face Mesh auch bei größeren Verhältnisunterschieden zwischen ursprünglichem und neuem Helligkeitswert eine höhere Erkennungsquote erreicht. Verallgemeinert lässt sich also sagen, dass ausgehend von den jeweils dargestellten Ergebnissen eine höhere Robustheit als beispielsweise bei den Algorithmen Haar AdaBoost (Viola-Jones Detector) und HoG SVM vorzuliegen scheint. Bei ersterem wird u. a. bei einem Verhältnis von 0,2 eine Erkennungsquote von nur knapp 30 %, bei HoG SVM eine Erkennungsquote von etwa 65 % erreicht. Face Mesh kann z. B. bei links lateralen Bildern bei auf das 0,2-Fache reduzierter Helligkeit alle abgebildeten Gesichter detektieren. Insgesamt sind die Ergebnisse bei den beiden genannten Algorithmen so signifikant schlechter, dass davon auszugehen ist, dass sie allgemein eine geringere Robustheit bei veränderter Helligkeit des Materials als Face Mesh aufweisen. S<sup>3</sup>FD hingegen kann bei einem Verhältnis von 0,2 unter Berücksichtigung des veränderten Kontrasts mit ca. 95 % eine ähnliche hohe Erkennungsquote wie Face Mesh erreichen. Eine Erkennungsquote von etwa 90 % bei selbem Verhältnis legt nahe, dass der Faster R-CNN Algorithmus insgesamt weniger robust als Face Mesh ist. [ZhLi2018]

Vorab ist noch anzumerken, dass ein Pixel im RGB-Farbraum nur einen maximalen bzw. minimalen Wert annehmen kann. Dabei bedeutet eine Erreichung des Maximums, dass ein Pixel weiß und eine Erreichung des Minimums, dass ein Pixel schwarz ist. [KeJe2022] Bei konstanter Erhöhung bzw. Minderung der Werte entsprechend der erhöhten bzw. verminderten Helligkeit werden demnach an einer gewissen Stelle diese Grenzen erreicht, wodurch Bereiche aus nur weißen und schwarzen Pixeln entstehen. Solche Bereiche reduzieren den Informationsgehalt der Bilder bedeutend, was wiederum die Gesichtserkennung deutlich erschwert.

### Frontale Bilder

Die Ergebnisse, wie u. a. ein Abfall der Erkennungsquote erst bei 1,89-facher Helligkeit, legen nahe, dass frontal aufgenommene Gesichter erst bei gravierender Veränderung der Helligkeit von Face Mesh nicht mehr erkannt werden können. Um zu überprüfen, ob Face Mesh nur bei vermeintlich „idealen“ Bedingungen (frontale Aufnahme) zu diesen Ergebnissen in der Lage ist, wird dieser Aspekt bei weiteren Kategorien noch einmal aufgegriffen.

Im Gegensatz zu den verpixelten Bildern ist eine stetig fallende Erkennungsquote zu beobachten. Auch ist festzustellen, dass die durchschnittliche Erkennungsquote sowohl bei Erhöhung als auch bei Verminderung der Helligkeit jeweils höher ist als beim Verpixeln der frontalen Bilder. Diese zwei Aspekte führen zu der Annahme, dass die Erkennungsquote parallel zu dem Grad der Informationsreduktion sinkt und somit keiner besonderen Erklärung bedarf.

### **Bilder mit Brille**

Durch die beiden durchschnittlichen Erkennungsquoten - 74,2 % bei erhöhter und 70 % bei verminderter Helligkeit - entsteht der Eindruck, dass bei dieser Form von Informationsreduktion eine Brille zu keinem ausschlaggebenden Unterschied führt. Auch hier werden in den ersten Abstufungen hohe Erkennungsquoten erreicht, was zu der Annahme führt, dass Face Mesh eine höhere Robustheit bei erhöhter Helligkeit aufweist als beispielsweise bei Verpixelung der Bilder.

Die insgesamt leicht schlechtere Erkennungsquote der Bilder mit Brille bei verminderter Helligkeit im Vergleich zu frontalen Bildern ohne Brille beruht auf dem Ergebnis bei 0,02-facher Helligkeit. Dabei zeichnet sich bereits ab, dass diese Stufe die Limitierung von Face Mesh bei verminderter Helligkeit aufzeigt, da die dort erreichte Anzahl erkannter Gesichter sich bedeutend von der bei der Kategorie „frontal“ unterscheidet.

Es darf allerdings nicht außer Acht gelassen werden, dass die Grundhelligkeit mancher Bilder auch innerhalb einer Kategorie voneinander abweichen kann. Durch diesen Aspekt verursachte Auswirkungen auf die Gesichtserkennung müssten zusätzlich, z. B. durch Normalisierung der Bilder, untersucht werden. Im Ausblick wird dieser Aspekt erneut aufgegriffen.

### **Bilder mit Sonnenbrille**

In den ersten Abstufungen werden auch hier hohe Erkennungsquoten erreicht. Im Vergleich zum Verpixeln der Bilder fällt diese Kategorie allerdings besonders auf. Im vorangegangenen Kapitel konnte die Vermutung gestützt werden, dass zumindest minimal schlechtere Erkennungsquoten als bei der Kategorie „mit Brille“ aufgrund der Verdeckung der Augen durch die Sonnenbrille erreicht werden. Die hier bei verminderter Helligkeit erreichten Ergebnisse widerlegen diese Vermutung (u. a. durch die Gesamt-Erkennungsquote von 70 %) an entsprechender Stelle. Davon ausgehend lässt sich vermuten, dass die durch die Sonnenbrille verdeckten Gesichtslanmarken bei verminderter Helligkeit von Face Mesh ausreichend geschätzt werden können, sofern die Pixelanzahl nicht verändert wird. Ausschlaggebend könnte dabei auch die frontale Positionierung der Gesichter im Verhältnis zur Kamera sein. Allerdings hätte dies ebenfalls in Kapitel 5.1 der Fall sein müssen. Im Gegensatz dazu kann die eingangs aufgestellte Vermutung durch die Gesamterkennungsquote von 66,3 % gegenüber der 74, 2 % der Kategorie „mit Brille“ gestützt werden.

Ein weiterer bedeutender Aspekt ist, dass für jede der vorliegenden Bildkategorien andere Gesichter auf den Bildern abgebildet sind. Gleichzeitig wurden diese Gesichter unter



unterschiedlichen Bedingungen, wie z. B. Hintergrund und Belichtung, aufgenommen, welches ebenfalls Einfluss auf die Gesichtserkennung nach Informationsreduktion haben kann. Dieser Aspekt wird noch einmal im Ausblick thematisiert.

### **Links laterale Bilder**

Die im Vergleich zu den vorherigen Kategorien schlechteren Ergebniswerte bei erhöhter Helligkeit (u. a. eine Gesamterkennungsquote von nur 61 %) deuten darauf hin, dass Face Mesh links laterale Gesichter bei erhöhter Helligkeit schlechter detektieren kann. Die allgemein von Bild zu Bild variierenden Unterschiede, wie Hintergrund, generelle Helligkeit oder Kontrast zum Hintergrund, können dabei ebenfalls das Ergebnis beeinflussen. Um diesen Aspekt zu untersuchen, müssten die Bilder auf einen möglichst gleichen Ausgangszustand gebracht werden. Dafür könnten die Gesichter beispielsweise jeweils vor einem einfarbigen Hintergrund platziert werden.

Darüber hinaus variieren die genannten Aspekte nicht nur zwischen den Bildern innerhalb einer Kategorie, sondern auch zwischen Bildern verschiedener Bildkategorien. Ein maximal akkurater Vergleich könnte nur erfolgen, wenn an dieser Stelle für alle Kategorien das gleiche Ausgangsmaterial, angepasst an die jeweilige Kategorie durch z. B. Einfügen einer Brille, verwendet wird. Da ein solches Ausgangsmaterial hier nicht vorliegt, lassen sich nur Tendenzen ablesen, die für einen Vergleich verwendet werden können. Auch hier deutet die ab 1,92-facher Helligkeit stärker sinkende Tendenz der Erkennungsquote darauf hin, dass etwa an dieser Stelle die Grenze von Face Mesh liegt.

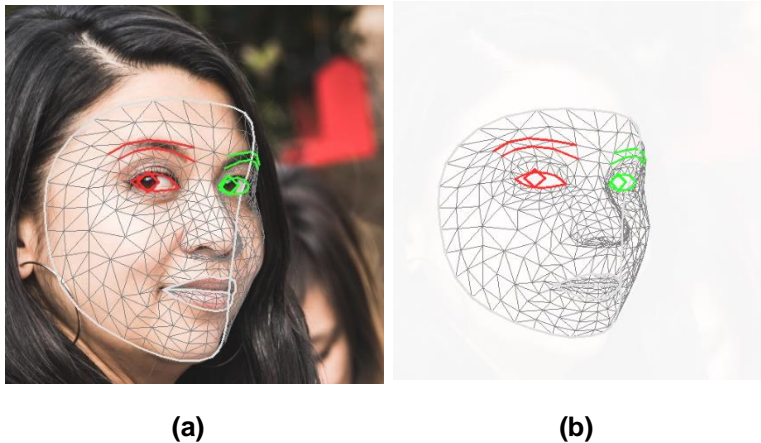
Bezüglich der Überprüfung des nicht erkannten links lateralen Bildes bei 0,18- bis 0,05-facher Helligkeit ist zu erwähnen, dass es sich stets um dasselbe Bild (19\_links\_heller\_x.png) handelt, auf dem kein Gesicht erkannt wird. (Das x stellt einen Platzhalter für den jeweiligen Helligkeitswert dar.) Dieses Bild ist bereits bei der Informationsreduktion durch Verpixelung aufgefallen. Ausgehend davon lässt sich vermuten, dass aufgrund der vollständig lateralen Position der Person zur Kamera die Qualität des Bildes verhältnismäßig hoch sein muss, damit Face Mesh in der Lage ist, ein Gesicht auf Basis verschiedener Gesichtslanmarken zu detektieren. Die vermeintlich höhere Gesamt-Erkennungsquote von 77,5 % ist durch die weiter vorgenommenen Abstufungen zu erklären, da durch diese vermehrt hohe Erkennungsquoten in die Berechnung einfließen.

### **Rechts laterale Bilder**

In dieser Bildkategorie sind die Ergebniswerte bei erhöhter Helligkeit bis zu einem 1,92-fachen Wert unauffällig. Die langsam sinkende Erkennungsquote ist ein Indiz dafür, dass rechts laterale Gesichter bei diesen Einstellungen weiterhin zuverlässig mittels Face Mesh erkannt werden können. Die Grenze, ab der große Differenzen zwischen den Anteilen der erkannten Gesichter auftreten, scheint hier bei einer 1,93-fachen Helligkeit zu liegen. Diese Aussage beruht darauf, dass verhältnismäßig kleine Abstufungen ausgewählt wurden und sich ab dieser Einstellung die Werte deutlich von Abstufung zu Abstufung verändern. Die Anzahl von nur 48 erkannten Gesichtern bei einem 1,86-fachen Helligkeitswert lässt

vermuten, dass rechts laterale Gesichter weniger robust von Face Mesh detektiert werden können als beispielsweise frontale Bilder. Daraus lässt sich wiederum schließen, dass die Position bzw. der Winkel des abgebildeten Gesichts durchaus bedeutender Faktor in der Gesichtserkennung ist.

Dies wird durch die insgesamt geringere durchschnittliche Erkennungsquote als bei frontalen Gesichtern von 67,1 % bestätigt. Allerdings wurde auch bei rechts lateralen Gesichtern bei 1,96-facher Helligkeit auf einem Bild ein Gesichtsnetz erstellt. Abbildung 7 zeigt sowohl dieses Gesichtsnetz als auch das dazugehörige Referenz-Gesichtsnetz, abgebildet auf dem Originalbild.



**Abbildung 7: Darstellung der Ergebnisbilder mit bei (a) 1- und bei (b) 1,96-facher Helligkeit erkannten Gesichtsnetzen der Kategorie „links lateral“. Bild (a) entspricht dem Eingabebild 3\_links.png, Bild (b) dem Eingabebild 3\_links\_heller\_500.png**

Die dargestellten Gesichtsmeshes zeigen sowohl in der Form als auch in der Größe nicht außer Acht zu lassende Unterschiede auf. Das Gesichtsmesh auf Bild (b) scheint insgesamt kleiner und auch kürzer zu sein. Aufgrund dessen ist davon auszugehen, dass von Face Mesh jedenfalls teilweise das Gesicht wohl in anderen Bildelementen erkannt wurde als in denen, die tatsächlich das Gesicht repräsentieren. Gestützt wird diese Annahme dadurch, dass ein menschlicher Betrachter auf Bild (b) nur minimale Konturen auf dem Bild erkennen kann.

Bei verminderten Helligkeitswerten ist die geringere Differenz der Erkennungsquoten zwischen der 0,03- und 0,02-fachen Helligkeit darauf zurückzuführen, dass bereits bei 0,03-facher Helligkeit nur noch 58 % der Gesichter erkannt werden. Dieser Aspekt kann damit erklärt werden, dass bei rechts lateralen Gesichtern augenscheinlich früher die Grenze erreicht wird, ab der erheblich weniger Gesichter zuverlässig detektiert werden können.

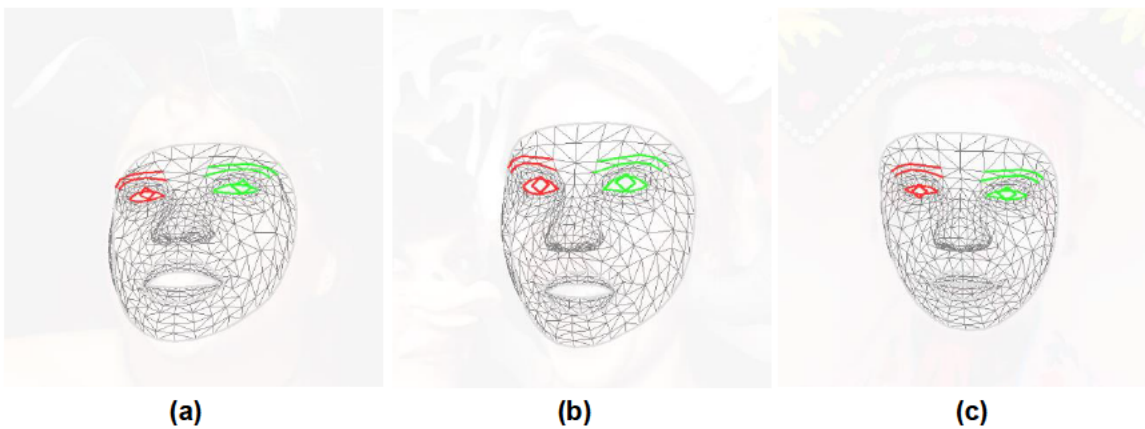
Demnach ist festzuhalten, dass eine laterale Position des Gesichts Auswirkungen auf die Gesichtserkennung mittels Face Mesh hat. Diese Behauptung wird auch weiter dadurch gestützt, dass die obere Grenze von 100 % an erkannten Gesichtern erst bei einer weiteren Abstufung erreicht wurde.

Grundsätzlich können die an dieser Stelle insgesamt leicht besseren Ergebnisse im Vergleich zu den links lateral gerichteten Gesichtern darauf hindeuten, dass Face Mesh rechts laterale Gesichter besser erkennt als links laterale Gesichter, also Face Mesh Landmarken auf der rechten Gesichtshälfte besser schätzen kann.

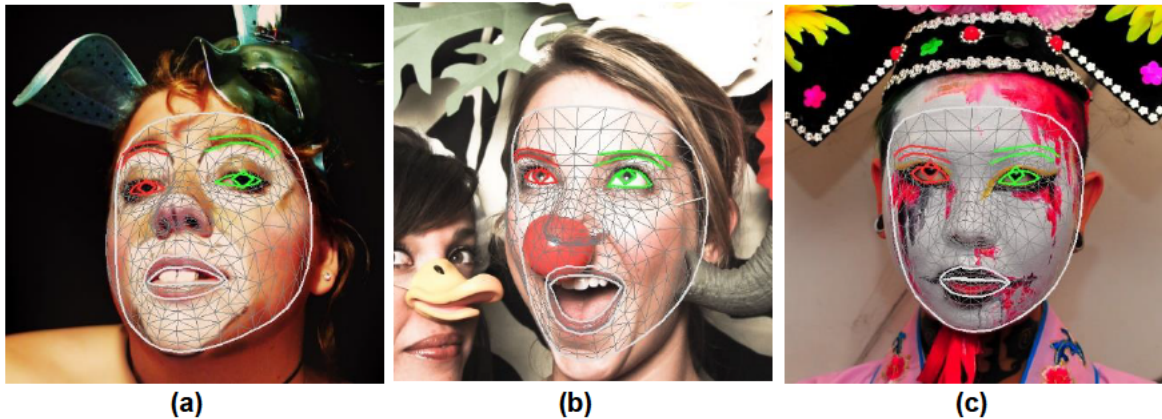
### **Bilder mit teilweiser Verdeckung**

Die bei 1,76- bis 1,89-fachem Helligkeitswert nicht sinkende Erkennungsquote deutet darauf hin, dass bei zunehmender Helligkeit die Detektion der Gesichter mit teilweiser Verdeckung durch Face Mesh relativ robust ist. Bei näherer Betrachtung der entsprechenden Gesichtsnetze sind keine Auffälligkeiten zu verzeichnen. Damit kann die Robustheit von Face Mesh bei dem hier vorliegenden Material vorerst bestätigt werden. Die Beobachtung, dass die Erkennungsquote ab 1,92-facher Helligkeit verhältnismäßig schnell sinkt sowie die ähnlichen Beobachtungen bei den anderen Bildkategorien sprechen erneut dafür, dass an dieser Stelle die Grenze von Face Mesh erreicht wurde.

Auffällig sind in dieser Kategorie die drei vermeintlich erkannten Gesichter bei 1,96-facher Helligkeit, da in den vorherigen Kategorien, wenn überhaupt, an dieser Stelle zumeist nur noch ein Gesicht erkannt worden ist. Die entsprechenden Bilder mit Gesichtsnetzen werden in den folgenden Abbildungen dargestellt. Abbildung 8 zeigt die erkannten Gesichtsnetze bei 1,96-facher Helligkeit und Abbildung 9 die entsprechenden Gesichtsnetze auf den gleichen Bildern mit ihrem ursprünglichen Helligkeitswert.



**Abbildung 8:** Darstellung der Ergebnisbilder mit bei jeweils 1,96-facher Helligkeit erkannten Gesichtsnetzen der Kategorie „teilweise verdeckt“. Bild (a) entspricht dem Eingabebild 10\_verdeckt\_heller\_500.png, Bild (b) dem Eingabebild 4\_verdeckt\_heller\_500.png und Bild (c) dem Eingabebild 43\_verdeckt\_heller\_500.png



**Abbildung 9: Darstellung der Ergebnisbilder mit bei jeweils 1-facher Helligkeit erkannten Gesichtsmesh-Netzwerken der Kategorie „teilweise verdeckt“. Bild (a) entspricht dem Eingabebild 10\_verdeckt.png, Bild (b) dem Eingabebild 4\_verdeckt.png und Bild (c) dem Eingabebild 43\_verdeckt.png**

Der größte Unterschied zwischen den zusammengehörenden Gesichtsmesh-Netzwerken ist bei Bild (b) in beiden Abbildungen zu erkennen. Das Gesichtsmesh-Netzwerk in Abbildung 8 (b) weist dabei einen deutlich kleineren detektierten Mund als das abgebildete Gesichtsmesh-Netzwerk in Abbildung 9 (b) auf und scheint insgesamt kürzer zu sein. In Abbildung 8 (a) scheint die Neigung des Gesichtsmesh-Netzwerkes größer zu sein als in Abbildung 9 (a). Für diese jeweils zwei erkannten Gesichtsmesh-Netzwerke lässt sich dementsprechend festhalten, dass Face Mesh aller Wahrscheinlichkeit nach die Gesichter nicht vollständig korrekt erkennt. Die erkannten Gesichtsmesh-Netzwerke weisen aber nicht außer Acht zulassende Ähnlichkeiten auf, woraus sich schließen lässt, dass von Face Mesh einzelne Teile der Gesichter immer noch richtig detektiert, sie aber mit falsch platzierten Landmarken verknüpft werden.

Das auf Bild (c) in Abbildung 8 erkannte Gesichtsmesh-Netzwerk scheint mit dem auf Bild (c) in Abbildung 9 größtenteils übereinzustimmen, weshalb sich hierbei vermuten lässt, dass dieses korrekt detektiert wird.

Insgesamt, unter der Berücksichtigung der möglicherweise falsch erkannten Gesichtsmesh-Netzwerke, würde sich dann eine Anzahl von einem erkannten Gesicht ergeben. Dies würde wiederum mit vorhergehenden Beobachtungen aus den anderen Bildkategorien übereinstimmen.

Der wiederholend große Abfall der Erkennungsquote bei 0,03- und 0,02-facher Helligkeit ist auch hier durch eine erreichte Grenze der Leistung von Face Mesh zu erklären.

### **6.3 Auswertung der Ergebnisse bei Anwendung Gauß'scher Unschärfe**

Im Folgenden sollen die in Kapitel 5.3 dargelegten Ergebnisse ausgewertet werden. Vorab lässt sich bereits sagen, dass im Vergleich zu Zhou et al. [ZhLi2018] eine höhere Anzahl an Einstellungen überprüft wurde und somit auch die Bildinformation jeweils erheblich mehr

beeinflusst wurde. Allgemein kann festgehalten werden, dass Face Mesh bei vergleichbaren Einstellungen (sigma bzw. radius gleich fünf und sechs) wesentlich bessere Ergebnisse erreicht hat als die vier anderen Gesichtserkennungsalgorithmen. So konnten bei den Abstufungen fünf und sechs bei fast allen Bildkategorien von Face Mesh alle dargestellten Gesichter erkannt werden. Die Ergebnisse von Zhou et al. [ZhLi2018] liegen bei diesen Abstufungen bezüglich der Erkennungsquote für den S<sup>3</sup>FD-Algorithmus bereits jeweils unter 90 %, für Faster R-CNN unter 80 % und für den Haar AdaBoost-Algorithmus bereits bei unter 60 %. Bei Verwendung des HoG SVM wird bei einer Einstellung von fünf eine Erkennungsrate von unter 50 % und bei einer Einstellung von sechs unter 40 % erreicht. [ZhLi2018]

Wenn auch ein vollständiger Vergleich aufgrund der unterschiedlichen Einstellungswerte nicht möglich ist, kann dennoch anhand der gemeinsamen Werte festgehalten werden, dass die Detektion von Gesichtern mittels Face Mesh bei Anwendung Gauß'scher Unschärfe u. a. mit einer Erkennungsquote von jeweils 100 % bei frontalen Bildern um einiges robuster ist als die Detektion der bei Zhou et al. [ZhLi2018] vorgestellten Algorithmen.

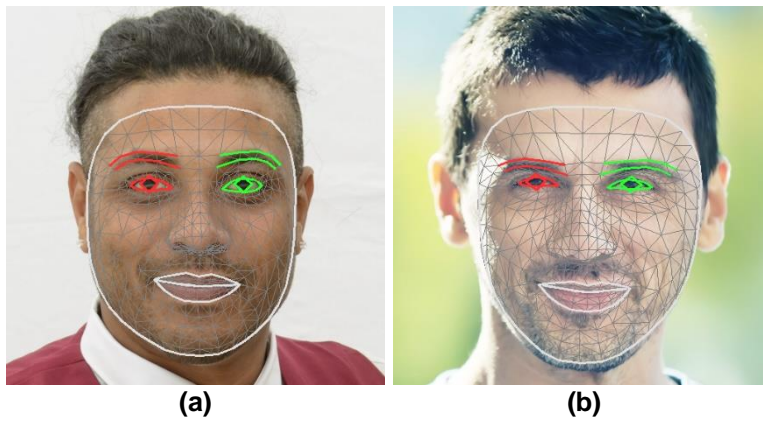
## Frontale Bilder

Die Tatsache, dass bei einer Radius-Einstellung von fünf bis zwölf stets eine Erkennungsquote von 100 % erreicht wird, lässt zunächst vermuten, dass diese Art von Informationsreduktion bei frontalen Bildern (bei verhältnismäßig geringem Ausmaß) keinen bedeutenden Einfluss auf die Gesichtserkennung mittels Face Mesh hat.

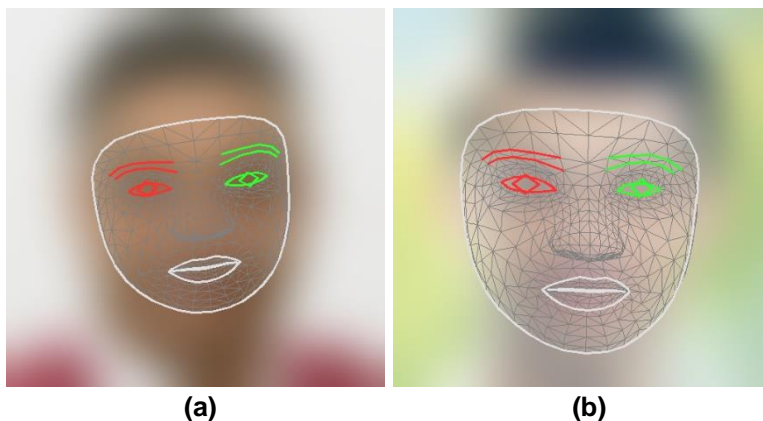
Bis zu einer Radius-Einstellung von einschließlich 20 sinkt die Erkennungsquote bei jeder Abstufung um einen verhältnismäßigen Anteil, woraus sich wiederum schließen lässt, dass dies im direkten Zusammenhang mit dem Grad der Informationsreduktion steht.

Zwischen einer Einstellung von 20 und 25 ist ein größerer Abfall der Erkennungsquote im Vergleich zu den vorangegangenen Abstufungen zu verzeichnen. Dies kann allerdings dadurch begründet werden, dass in diesem Fall eine größere Differenz zwischen den beiden Abstufungen besteht. Es könnten zusätzliche Abstufungen eingefügt werden. Der Fokus wurde allerdings primär daraufgelegt, an welcher Stelle die untere Grenze erreicht werden kann und nicht darauf mit welcher Geschwindigkeit die Erkennungsquote an dieser Stelle sinkt. Zudem lässt sich sagen, dass bei diesen weiteren Abstufungen eine ähnliche Geschwindigkeit zu erwarten wäre, da ausgehend von den vorangegangenen Differenzen (von etwa 6-10 %) zwischen den Stufen bei Radius-Einstellung 25 ein Wert von 14 % der allgemeinen Tendenz entspricht.

Im Vergleich dazu nimmt die Geschwindigkeit des Abfalls der Erkennungsquote bei den Einstellungen 25 und 30 ab. Zu erklären ist dies u. a. damit, dass bei einem solch hohen Grad der Informationsreduktion bei den zu verarbeitenden Bildern nur noch vereinzelt Unterschiede auftreten. Auf den folgenden Abbildungen sind die von Face Mesh detektierten Gesichtsnetze bei einer Einstellung von 30, siehe Abbildung 11, sowie auch der detektierten Gesichtsnetze auf den entsprechenden Referenzbildern, siehe Abbildung 10, dargestellt.



**Abbildung 10:** Darstellung der Ergebnisbilder mit erkannten Gesichtsnetzen der Kategorie „frontal“. Bild (a) entspricht dem Eingabebild 22\_frontal.png, Bild (b) dem Eingabebild 49\_frontal.png



**Abbildung 11:** Darstellung der Ergebnisbilder mit bei Anwendung Gauß'scher Unschärfe mit Radius 30 erkannten Gesichtsnetzen der Kategorie „frontal“. Bild (a) entspricht dem Eingabebild 22\_frontal\_gaussian\_30.png, Bild (b) dem Eingabebild 49\_frontal\_gaussian\_30.png

Vergleicht man die jeweils zueinander gehörigen Gesichtsnetze, so ist ein signifikanter Unterschied unmittelbar zu erkennen. Aufgrund dessen, dass Face Mesh in Abbildung 11 scheinbar nicht dem tatsächlichen Gesicht entsprechende Bildelemente als Gesichtsmuster erkennt, ist davon auszugehen, dass diese Gesichtsnetze fälschlicherweise erstellt worden sind. Theoretisch wird also bereits bei dieser oder bei Einstellung eines kleineren Radius die Grenze von Face Mesh erreicht. Betrachtet man in diesem Zusammenhang die erstellten Gesichtsnetze bei einer Radius-Einstellung von 25 fällt auch dort auf, dass nicht alle Gesichtsnetze korrekt zu sein scheinen. Deshalb ist bereits an dieser Stelle festzuhalten, dass die angegebene Anzahl der erkannten Gesichter nicht immer eine korrekte Auskunft darüber gibt, wie viele Gesichter tatsächlich noch korrekt erkannt werden können.

### **Bilder mit Brille**

Wie bereits in Kapitel 5.3 festgestellt werden konnte, ist bei der Kategorie „mit Brille“ in Bezug auf die Informationsreduktion mittels Gauß'scher Unschärfe ein Unterschied zu der Kategorie „frontal“ auffallend. Anders als bei der Kategorie „frontal“ sinkt die Erkennungsquote bereits ab einer Einstellung von acht. Dieser Aspekt ist ein Indiz dafür, dass das Tragen einer Brille durchaus bei Anwendung Gauß'scher Unschärfe einen Einfluss auf die

Gesichtserkennung mittels Face Mesh hat. Auch sinkt die Erkennungsquote insgesamt vermeintlich schneller, was ebenfalls mit dem genannten Aspekt zu begründen ist.

Weiterhin ist zwischen den Abstufungen von 20 und 25 eine erheblich geringere Differenz in der Erkennungsquote zu beobachten als bei der Kategorie „frontal“. Dies ist wiederum dadurch zu erklären, dass die Anzahl der erkannten Gesichter bereits bei einer Radius-Einstellung von 20 wesentlich geringer ist. Gemäß der Beobachtung bei der Kategorie „frontal“, dass eine bereits geringe Erkennungsquote langsamer sinkt, ist dieser Aspekt ebenfalls zu stützen.

Im Gegensatz zu der Kategorie „frontal“ kann die untere Grenze hier bereits bei einer Radius-Einstellung von 30 erreicht werden. Diese Tatsache verstärkt die zuvor implizierte Vermutung, dass das Tragen einer Brille die Ergebnisse mittels Face Mesh bei Informationsreduktion unter Anwendung Gauß'scher Unschärfe generell beeinflusst und dabei verschlechtert.

Jedoch sind bei Betrachtung der bei einer Radius-Einstellung von 25 erstellten Gesichternetze ähnliche Beobachtungen wie bei frontalen Gesichtern ohne Brille zu vermerken. Dadurch wird deutlich, dass besonders bei dieser Art von Informationsreduktion Face Mesh ab einem gewissen Grad nicht nur Schwierigkeiten hat Gesichter zu erkennen, sondern auch oftmals fälschlicherweise Gesichter in nicht das Gesicht repräsentierenden Bildelementen detektiert.

Verallgemeinert kann bereits an dieser Stelle festgehalten werden, dass die Limitierung der Leistung von Face Mesh in Bezug auf die ordnungsgemäße Erkennung von Gesichtern unter der liegt, welche aus den jeweiligen Erkennungsquoten hervorgeht. Es können dabei allerdings manuell nicht immer klare Grenzen gezogen werden, sodass eine spezifischere und objektive Analyse der Gesichternetze notwendig wäre. Dies stellt allerdings eine Herausforderung für zukünftige Arbeiten dar.

## **Bilder mit Sonnenbrille**

Ähnlich zu der Kategorie „mit Brille“ sinkt die Erkennungsquote der Gesichter in dieser Kategorie bereits ab einer Einstellung von acht, weshalb zunächst davon auszugehen ist, dass das Tragen einer Sonnenbrille hierbei im Vergleich zum Tragen einer Brille keinen signifikanten Unterschied bewirkt.

Allerdings, wie auch in den Ergebnisbeobachtungen festgehalten, wird die untere Grenze erst bei einer Radius-Einstellung von 40 erreicht. In diesem Zusammenhang ist bei einer Einstellung von 30 eine Erkennungsquote von 14 % zu verzeichnen. Da sich dies deutlich von den Ergebnissen der vorangegangenen Kategorien unterscheidet, ist davon auszugehen, dass es sich an dieser Stelle, zumindest bei der Mehrheit der erkannten Gesichter, um fälschlicherweise als Gesichtsmuster erkannte Bildelemente handelt. Ausgehend davon ist ebenfalls zu vermuten, dass das Tragen einer Sonnenbrille einen ähnlichen Einfluss auf die Gesichtserkennung mittels Face Mesh wie das Tragen einer Brille hat. Dabei hat es den Anschein, dass Bildelemente, die die Sonnenbrille repräsentieren, mehr falsch positiv

erkannte Gesichter hervorbringen. Hierbei ist zu vermuten, dass der Algorithmus die dunkleren Bildelemente der Sonnenbrille im Gesamtkontext mit höherer Wahrscheinlichkeit als Muster eines Gesichts klassifiziert.

### **Links laterale Bilder**

Die bedeutend hohe Abweichung der Ergebnisse der Gesichtserkennung im Vergleich zu frontal positionierten Gesichtern deutet auf zwei Aspekte hin. Zum einen können die Abweichungen mit den individuellen Unterschieden in Bezug auf Hintergrund, Kontrast oder Helligkeit etc. zu erklärt werden. Konkret würde das bedeuten, dass Face Mesh bei zunehmender Informationsreduktion Schwierigkeiten hat die Gesichtslandmarken und damit Gesichter beispielsweise vor einem bestimmten Hintergrund zu detektieren.

Zum anderen können die abweichenden Ergebnisse der Gesichtserkennung darauf hindeuten, dass links laterale Gesichter mittels Face Mesh nicht gleich gut erkannt werden. Dies würde u. a. damit zusammenhängen, wie gut das neuronale Netz auf links lateral gerichtete Gesichter trainiert ist.

Die Geschwindigkeit des Abfalls der Erkennungsquote passt grundsätzlich in das zu erwartende Schema. Die abnehmende Erkennungsquote ist in Relation mit dem Grad der Informationsreduktion zu erklären.

### **Rechts laterale Bilder**

Auch bei den rechts lateralen Gesichtern fällt auf, dass die Erkennungsquote in Bezug auf die Anzahl der erkannten Gesichter bereits bei einer Radius-Einstellung von acht sinkt. Dementsprechend ist davon auszugehen, dass eine laterale Positionierung des Gesichts bei Anwendung Gauß'scher Unschärfe allgemein Einfluss auf die Gesichtserkennung mittels Face Mesh hat.

Bei den vorliegenden Ergebnissen fällt auf, dass erstmalig bei einer Radius-Einstellung von zehn nur eine Anzahl der erkannten Gesichter unter 48 erreicht werden kann. Dies ist als Auswirkung der rechts lateralen Positionierung des Gesichts oder aber auch der individuellen Unterschiede der einzelnen Bilder, wie beispielsweise der Kontrast zum Hintergrund oder die generelle Belichtung, zu deuten.

Die geringere Differenz der Erkennungsquote zwischen den Abstufungen von einer Radius-Einstellung von 20 und 25 ist anhand der bereits genannten Beobachtungen im Kontext der anderen Kategorien damit zu erklären, dass bei bereits geringer Anzahl erkannter Gesichter die Erkennungsquote allgemein langsamer sinkt. Dabei ist davon auszugehen, dass die Detektion einige der Gesichter, die auch bei hoher Informationsreduktion noch erkannt werden, verhältnismäßig robust ist, weshalb diese auch bei steigender Informationsreduktion zunächst noch detektiert werden können.

Eine klare Differenzierung beider möglichen Ursachen ist jedoch unter den gegebenen Umständen nicht möglich, da nicht jede Entscheidung des neuronalen Netzes, auf dem Face



Mesh arbeitet, nachvollzogen werden kann. Die Versuchsreihen könnten jedoch beispielsweise erweitert bzw. spezifiziert werden, indem jedes Gesicht ausgeschnitten und vor einen einfarbigen Hintergrund gesetzt wird. Dieser Aspekt wird im Ausblick erneut aufgegriffen.

### **Bilder mit teilweiser Verdeckung**

Bei dieser Kategorie fällt unmittelbar auf, dass die Erkennungsquote bereits ab einer Radius-Einstellung von sechs sinkt. Daraus lässt sich der Schluss ziehen, dass die Detektion der Gesichter dieser Bildkategorie bei Informationsreduktion mittels Gauß'scher Unschärfe nicht besonders robust ist. Unter anderem lässt sich dies damit begründen, dass bei der Gauß'schen Unschärfe Bildelemente ineinander „verschwimmen“. Deshalb wird die Zuordnung einzelner Bildelemente zu einem Gesichtsmuster oder eben auch zu keinem erheblich erschwert. Kommen bei diesem Prozess beispielsweise Gegenstände unmittelbar über Teilen des Gesichts hinzu, so „verschwimmen“ auch deren repräsentierende Pixel mit denen des Gesichts. Folglich kann dabei ein Gesichtsmuster schwieriger erkannt werden. Eine Brille z. B. macht im Vergleich einen wesentlich geringeren Anteil aus.

Auch wenn eine Abstufung über eine Radius-Einstellung von 30 hinaus notwendig war, um die untere Grenze der Erkennungsquote zu erreichen, sind die Ergebnisse insgesamt nicht besser als die der Kategorien, bei denen ein solcher Schritt nicht notwendig gewesen ist. Dies zeigt auf, dass eine weitere notwendige Abstufung keinen direkten Aufschluss darüber geben kann, wie robust die Gesichtserkennung gegenüber entsprechender Informationsreduktion ist.

## 6.4 Auswertung der Ergebnisse bei Anwendung von Interpolation

Allgemein lässt sich zu den Ergebnissen der Gesichtserkennung in Bezug auf Bildmaterial, welches mittels Interpolation zunächst verkleinert und anschließend wieder vergrößert worden ist, sagen, dass MediaPipe Face Mesh bei einigen Versuchsreihen sehr große Schwankungen in der Anzahl der erkannten Gesichter hervorbringt. Diese Beobachtung lässt bereits darauf schließen, dass bei dieser Art von Informationsreduktion viele falsch erkannte Gesichtsnetze vorliegen. Besonders bei den Abstufungen, bei denen die Erkennungsquote im Vergleich zu der vorausgegangenen steigt statt sinkt, ist diese Schlussfolgerung mittels einer Betrachtung der Gesichtsnetze zu bestätigen. Dabei können die erkannten Gesichtsnetze anhand manueller Betrachtung als nicht korrekt bewertet werden. Beispielsweise liegen die vermeintlich erkannten Pupillen außerhalb der detektierten Augen oder der Umriss des Gesichts ist stark verformt und vergrößert.

Den genannten Beobachtungen nach ist MediaPipe Face Mesh bei durch Interpolation verkleinerten und anschließend vergrößerten Bildern bereits nach wenigen Abstufungen nicht robust genug, um korrekte Ergebnisse zu erzielen.

### Frontale Bilder

Die ersten vier Abstufungen zeigen keine Auffälligkeiten, da stets eine Erkennungsquote von 100 % erreicht werden kann. Allem Anschein nach kann Face Mesh hier also noch zuverlässige Ergebnisse erzielen. Erstmals zeigt sich einer der angesprochenen Schwankungen in der Erkennungsquote bei einer Verkleinerung auf 18x18 Pixel und anschließender Vergrößerung. Die Ursache dafür, und auch für die weiteren Schwankungen, ist der vorausgehenden Erläuterung zu entnehmen. Insgesamt machen es die falsch detektierten Gesichtsnetze schwierig zu bestimmen, wo die genaue Grenze der Leistung von Face Mesh unter den gegebenen Bedingungen liegt. Denn die Entscheidung, ob sie als korrekt oder inkorrekt zu werten sind, ist bei manchen Gesichtsnetzen äußerst subjektiv. Es müsste demnach eine Methode angewandt werden, die diese Bewertung anhand ausgewählter Eigenschaften objektiv übernehmen kann. Weitere Informationen dazu werden im Ausblick erläutert.

### Bilder mit Brille

In Bezug auf die hier vorliegenden Ergebnisse lässt sich sagen, dass die Gesichtserkennung mittels Face Mesh auf den ersten Blick robuster erscheint als bei frontalen Bildern ohne Brille. Ausgehend von den Ergebnissen der anderen Kategorien der Informationsreduktion ist allerdings festzuhalten, dass diese Beobachtung eine Ausnahme darstellt. Anhand subjektiver Bewertung der Gesichtsnetze ist bereits bei Verkleinerung auf 21x21 Pixel und anschließender Vergrößerung davon auszugehen, dass die Grenze der Leistung von Face Mesh fast erreicht ist, d. h., dass bereits an dieser Stelle kaum noch Gesichter korrekt

detektiert werden können. Die vermeintlich besseren Ergebnisse können demnach als nicht vollständig korrekt gewertet werden.

### **Bilder mit Sonnenbrille**

Wie auch bei den Bildern, auf denen die abgebildete Person eine Brille trägt, stellt das Tragen einer Sonnenbrille Face Mesh vor ähnliche Schwierigkeiten. Diese Schwierigkeiten sind bei dieser Art der Informationsreduktion jedoch fast vollständig auf diese selbst statt auf die Unterschiede zwischen den einzelnen Bildkategorien zurückzuführen. Aufgrund dessen, dass Face Mesh auch hier, nach Betrachtung der detektierten Gesichtsnetze, bereits ab einer Verkleinerung auf 21x21 Pixel und anschließender Vergrößerung nach einer manuellen Bewertung nur noch vereinzelt Gesichter korrekt detektiert, ist ebenfalls von einem deutlich niedrigeren Schwellwert in Bezug auf die Limitierung von Face Mesh auszugehen als von Verkleinerung auf eine Pixelanzahl von 4x4 und anschließender Vergrößerung auf 512x512 Pixel.

### **Links laterale Bilder**

Bei links lateralen Bildern sind erneut drei Sprünge in der Erkennungsquote in Bezug auf die Anzahl der erkannten Gesichter zu beobachten, welche ebenfalls mit der allgemeingefassten Erklärung zu begründen sind. Anhand der Betrachtung der detektierten Gesichtsnetze ab einer vorangegangenen Verkleinerung auf 21x21 Pixel ist auch hier die Limitierung von Face Mesh deutlich früher anzusetzen, als aus den Ergebniswerten hervorgeht. Generell ist davon auszugehen, dass die Grenze bereits bei Verkleinerung auf 30x30 Pixel fast erreicht wird. Es könnten weitere Abstufungen zwischen 30x30 und 21x21 Pixeln untersucht werden, allerdings müsste auch bei diesen stets eine Betrachtung der von Face Mesh detektierten Gesichtsnetze vorgenommen werden. Die zunächst tabellarisch dargestellten Ergebnisse haben gezeigt, dass weitere Abstufungen nicht unbedingt sinnvoll wären, da bereits zwischen den genannten Abstufungen kein deutlicher Abfall der Erkennungsquote zu beobachten ist. Solche Abstufungen könnten zwar nachgearbeitet werden, allerdings zeigen die detektierten Gesichtsnetze bereits ab einer Verkleinerung auf 30x30 Pixel fälschlicherweise Detektierte auf.

### **Rechts laterale Bilder**

Ähnliche Schlüsse wie die bei den bereits diskutierten Bildkategorien können auch bei rechts lateralen Gesichtern gezogen werden. So repräsentiert die Erkennungsquote nicht vollständig die tatsächlich korrekt erkannten Gesichter. Dementsprechend wäre auch hier eine manuelle Überprüfung der einzelnen detektierten Gesichtsnetze notwendig, um eindeutige Schlüsse ziehen zu können. Auch hier sei auf die im Ausblick erwähnte Methode verwiesen.

## **Bilder mit teilweiser Verdeckung**

Zuletzt sind auch bei den Gesichtern mit teilweiser Verdeckung ähnliche Abweichungen zwischen den Erkennungsquoten und den Beobachtungen zu korrekt detektierten Gesichternetzen zu erkennen. Dabei deutet der erhebliche Sprung der Erkennungsquote zwischen einer Verkleinerung auf 30x30 und 21x21 Pixeln darauf hin, dass die das Gesicht repräsentierenden Bildelemente bereits einen so geringen Anteil auf dem Bild aufweisen, dass an dieser Stelle deutlich weniger Gesichter detektiert werden. Aufgrund dessen kann ein Unterschied zu den anderen Bildkategorien festgestellt werden, wenn auch der Gesamtvergleich durch inkorrekte Gesichternetze erheblich beeinträchtigt wird.

## 7 Ergebnisse und Ausblick

Im abschließenden Kapitel werden sowohl die Ergebnisse zusammengefasst als auch aufgetretene Schwierigkeiten dargelegt. Außerdem erfolgt eine Bewertung der Arbeit. Zum Schluss wird in einem Ausblick dargestellt, welche weiteren Versuche, Lösungen der genannten Schwierigkeiten und generelle Weiterentwicklungen bzw. weiterführende Arbeitsschritte möglich sind.

### 7.1 Ergebnisse und Problemstellungen

Verallgemeinert lässt sich bezüglich der erreichten Ergebnisse der Gesichtserkennung mittels MediaPipe Face Mesh auf informationsreduziertem Bildmaterial sagen, dass zum Großteil ähnliche Ergebnisse wie bei Untersuchung vergleichbarer Algorithmen erzielt werden konnten. An einigen Stellen übertrifft Face Mesh jedoch diese Algorithmen. So können zum einen ähnliche Werte in Bezug auf verpixelte Gesichter im Vergleich zu der Arbeit von Zheng et al. [ZhRa2010] erreicht werden. Dabei kann Face Mesh zwar beispielsweise bei etwa 16x16 gesichtsrepräsentierenden Pixeln bei frontalen Bildern mit einer Detektionsrate von 98 % die Detektionsrate unter Zunahme von Color Tensus Transform von 97,2 % übertreffen, die Differenz ist allerdings mit 0,8 % bedeutend gering.

Zum anderen kann im direkten Vergleich mit den Algorithmen aus Zhou et al. [ZhLi2018] festgehalten werden, dass Face Mesh bei gleicher Einstellung der Standardabweichung bei Anwendung der Gauß'schen Unschärfe Face Mesh signifikant bessere Erkennungsquoten erreichen kann. Diese betragen pro Bildkategorie bei einer Standardabweichung bzw. Radius-Einstellung von sechs mindestens 96 %. Im Gegensatz dazu kann bei den Algorithmen aus Zhou et al. [ZhLi2018] der S<sup>3</sup>FD-Algorithmus mit knapp über 80 % die höchste Erkennungsrate erzielen. Mit Faster R-CNN kann lediglich eine Quote von etwas über 60 % erreicht werden, mit Haar AdaBoost eine Rate von 50 % und bei HoG SVM liegt die Erkennungsrate bei unter 40 %. [ZhLi2018] Gegenüber diesen Algorithmen lässt sich somit sagen, dass Face Mesh robuster bei Anwendung Gauß'scher Unschärfe ist. Zu ergänzen ist dabei jedoch, dass Face Mesh bei weiteren Abstufungen Schwierigkeiten in Bezug auf die korrekte Detektion der Gesichter und Erstellung der Gesichtsnetze aufgezeigt hat. Beispielsweise lagen die erkannten Pupillen außerhalb der erkannten Augen oder die Umrandung des Gesichtsnetzes war deutlich deformiert.

Bezüglich der Informationsreduktion der Bilder unter Veränderung der Helligkeit lassen sich die von Zhou et al. [ZhLi2018] vorgestellten Algorithmen annähernd mit Face Mesh vergleichen. Annähernd zum einen, weil Zhou et al. [ZhLi2018] nur verminderte Helligkeit berücksichtigt, zum anderen, weil sie dabei gleichzeitig die Kontrasteinstellungen verändert

haben, was in der vorliegenden Arbeit nicht der Fall ist. Bei höchster Einstellung bzw. größtem Abstandverhältnis weisen alle vier vorgestellten Algorithmen eine niedrigere Erkennungsrate als Face Mesh auf, welches unter anderem durch gleichzeitige Verminderung des Kontrasts verursacht sein kann. Dennoch ist dieser Aspekt ein Indiz dafür, dass Face Mesh auch hier eine höhere Robustheit gegenüber dieser Informationsreduktion aufweist. Bei Erhöhung der Helligkeit ist aufgrund der vorliegenden Ergebnisse und der Tatsache, dass die Erkennungsquote in den meisten Bildkategorien erst nach einigen Abstufungen sinkt, ebenfalls von einer gewissen Robustheit seitens Face Mesh auszugehen.

Im Gegensatz dazu zeigen die zahlreichen Sprünge der Erkennungsquote bei Informationsreduktion mittels Interpolation auf, dass Face Mesh an dieser Stelle wenig robust ist. Zudem konnten diese Schwierigkeiten in der Diskussion anhand Betrachtung der von Face Mesh erkannten Gesichtsnetze bestätigt werden.

Allgemein ist ebenfalls anzumerken, dass die Erkennungsquoten und damit die Anzahl der erkannten Gesichter nicht immer die tatsächlichen Ergebnisse aufzeigen. Grund dafür sind zwar erstellte, aber nicht dem wirklich abgebildeten Gesicht entsprechenden Gesichtsnetze. In diesem Zusammenhang ist auf eine weitere Problemstellung hinzuweisen. Einige Gesichtsnetze können zwar objektiv als falsch bewertet werden (z. B. wenn die detektierten Pupillen außerhalb der detektierten Augen liegen), andere bedürfen bisher einer rein subjektiven Einschätzung. Dieses Problem gilt es in Zukunft zu lösen.

Des Weiteren werden individuelle Unterschiede der einzelnen Bilder nicht berücksichtigt. Beispielsweise wurde die Varietät der Hintergründe oder auch die Grundhelligkeit der Bilder bei der Gesichtserkennung mittels Face Mesh außer Acht gelassen, obwohl diese einen Einfluss auf die Ergebnisse haben können. Für einen akkuraten Vergleich der Ergebnisse zwischen den jeweiligen Bildkategorien müsste demnach eine Normalisierung durchgeführt werden. Weiter können die verschiedenen Gesichter selbst Unterschiede bzgl. der Ergebnisse verursachen, wenn Face Mesh beispielsweise schwieriger Gesichtslanmarken detektieren kann als bei anderen Gesichtern.

Abschließend lässt sich sagen, dass MediaPipe Face Mesh bei den Informationsreduktionskategorien, bei denen die Pixel in Anzahl und bzw. oder in ihrer Position verändert werden – bei Verpixelung und Interpolation – erhebliche Schwierigkeiten aufweist. Im Gegensatz dazu weisen die Erkennungsquoten beispielsweise bei erhöhter Helligkeit der einzelnen Pixel, welche an dieser Stelle somit in Anzahl und Position gleichbleiben, keine signifikanten Schwankungen auf.

## **7.2 Bewertung der Arbeit**

Im Rahmen der vorliegenden Bachelorarbeit konnten für vier vorgestellte Kategorien der Informationsreduktion von vorgegebenem Ausgangsmaterial die jeweiligen Schwellwerte der Gesichtserkennung mittels MediaPipe Face Mesh aufgezeigt werden. Diese

Schwellwerte spiegeln jeweils den Grad der Informationsreduktion wider, bei dem eine vollständige Nicht-Erkennung der Gesichter erreicht wird. Aufgetretene Schwierigkeiten bezüglich exakter Bestimmung dieser Werte wurden ebenfalls dargelegt. Damit wurde die Aufgabenstellung in ihren wesentlichen Bestandteilen erfüllt.

Des Weiteren wurden die zur Anwendung der Informationsreduktion erarbeiteten Python-Skripte in ihren grundlegenden Elementen erläutert. Gleichzeitig konnten darin verwendete Methoden, wie z. B. Interpolation, skizziert werden.

Bezüglich der konkreten Versuchsreihen ist anzumerken, dass an manchen Stellen weitere Abstufungen denkbar bzw. sinnvoll wären.

Zusammenfassend lässt sich sagen, dass die vorliegende Bachelorarbeit mit deren Ergebnissen als Grundlage für weitere Versuchsreihen zur Analyse der Robustheit von anderen Gesichtserkennungs-Algorithmen oder auch für im nächsten Schritt angewandte Super Resolution dienen kann. Letzteres wird im Ausblick weiter ausgeführt.

### **7.3 Ausblick**

Um einen genaueren Vergleich der Ergebnisse in Bezug auf die Gesichtserkennung mittels Face Mesh zwischen den einzelnen Bildkategorien gewährleisten zu können, müsste für jede Bildkategorie dieselbe Person jeweils in verschiedenen Situationen dargestellt werden. So würde eine Person unter anderem frontal, links lateral, rechts lateral, mit Sonnenbrille oder mit Brille abgebildet werden. Unter diesen Voraussetzungen wäre ein solcher Vergleich realisierbar. Um eine noch höhere Vergleichbarkeit zu erzielen, könnten die Bilder beispielsweise vor einem einfarbigen Hintergrund bei gleicher Belichtung aufgenommen werden. Für das vorliegende Material könnten alternativ die Gesichter mit Bildbearbeitungs-Tools ausgeschnitten und vor einen einheitlichen Hintergrund gesetzt werden.

Als Methode zur objektiven Bestimmung der Korrektheit erstellter Gesichtsnetze wäre es denkbar, die euklidische Distanz zwischen den einzelnen erkannten Gesichtslanmarken zu berechnen. Die euklidische Distanz zweier Punkte, in diesem Fall zweier Landmarken, in der Ebene oder im Raum ist „die zum Beispiel mit einem Lineal gemessene Länge einer Strecke, die diese zwei Punkte verbindet. Dieser Abstand ist invariant unter Bewegungen.“ [Hö2021] Dafür wurde im Rahmen dieser Bachelorarbeit bereits ein Python-Skript entwickelt. Um dieses Skript bzw. dessen Ergebnisse verwenden zu können, müsste allerdings zusätzlich eine Skala entwickelt werden, welche angibt ab wann die Distanz und damit der Unterschied zwischen zwei Gesichtsnetzen (zwischen dem auf dem Referenzbild und auf dem informationsreduzierten Bild erkannten Gesichtsnetz) als zu groß und damit als inkorrekte Detektion klassifiziert wird.

Allgemein könnten auch weitere Kategorien der Informationsreduktion, wie z. B. Veränderung der Kontrasteinstellungen, und deren Auswirkung auf die Gesichtserkennung mittels Face Mesh oder auch anderer Algorithmen überprüft und ausgewertet werden.

Im Anschluss an die Schwellwertanalyse der Gesichtserkennung mittels Face Mesh wäre die Anwendung von sogenannten Face-Recognition-Algorithmen und Evaluierung derer Ergebnisse denkbar. Algorithmen dieser Art sind sowohl in der Lage ein Gesicht generell zu erkennen als auch wiederzuerkennen, beispielsweise im Rahmen der forensischen Personenidentifikation [AO2022]. Vor Durchführung eines solchen Schrittes könnte eine sogenannte Super Resolution auf den Bildern angewendet werden. Super Resolution (hochauflösende) Techniken rekonstruieren ein Bild oder eine Sequenz mit höherer Auflösung aus den beobachteten Low Resolution (niedrig aufgelösten) Bildern [YuSh2016]. Ein solcher Prozess wäre also bei einigen Informationsreduktions-Kategorien, z. B. bei verpixeltem Material, denkbar anzuwenden. Im Anschluss daran könnte dann speziell überprüft werden, ob zuvor nicht mehr mittels der Gesichtserkennung detektierte Gesichter ausschlaggebend verbessert werden können, sodass nach Anwendung der Super Resolution das Gesicht von der Gesichtserkennung detektiert werden kann.

Abschließend lässt sich sagen, dass die vorliegenden Ergebnisse in Bezug auf die Gesichtserkennung mittels MediaPipe Face Mesh aufzeigen können, bis zu welchem Grad der Informationsreduktion die Detektion dargestellter Gesichter zuverlässig ist.



# Quellenverzeichnis

## Literatur:

- [AmSa2022] Aman; Sangal, A. L. (2022): Drowsy Alarm System Based on Face Landmarks Detection Using MediaPipe Face Mesh. In: Rawat, S., Kumar, A.; Kumar, P.; Anguera, J. (Hg.) Proceedings of First International Conference on Computational Electronics for Wireless Communications. Lecture Notes in Networks and Systems, vol. 329, S.363-375. Springer, Singapore. DOI: 10.1007/978-981-16-6246-1\_31.
- [BeKe2006] Berbar, Mohammed A.; Kelash, Hamdy M.; Kandeel, Amany A. (2006): Faces and Facial Features Detection in Color Images. In: Sarfraz (Hg.) 2006 – Geometric Modeling and Imaging--New Trends (GMAI'06), S. 209–214. DOI: 10.1109/GMAI.2006.18.
- [JaLe2010] Jain, Vidit; Learned-Miller, Erik (2010): FDDB: A benchmark for face detection in unconstrained settings. Tech. Rep. UM-CS-2010-009, University of Massachusetts, Amherst.
- [KIMu2020] Kleesiek, Jens; Murray, Jacob M.; Strack, Christian.; Kaissis, Georgios; Braren, Rickmer (2020): Wie funktioniert maschinelles Lernen? In: Der Radiologe, vol. 60, S. 24-31. DOI: 10.1007/s00117-019-00616-x.
- [PaS2018] Patil, M.; S. Malini, Manjunatha (2018): Interpolation Techniques in Image Resampling. In: International Journal of Engineering & Technology vol. 7, no. 3.34, S. 567–570. DOI: 10.14419/ijet.v7i3.34.19383.
- [ShMo2014] Sharifara, Ali; Mohd Rahim, Mohd S.; Anisi, Yasaman (2014): A general review of human face detection including a study of neural networks and Haar feature-based cascade classifier in face detection, 2014 International Symposium on Biometrics and Security Technologies (ISBAST), S. 73-78. DOI: 10.1109/ISBAST.2014.7013097.

- [YuSh2016] Yue, Linwei; Shen, Huanfeng; Li, Jie; Yuan, Qiangqiang; Zhang, Hongyan; Zhang, Liangpei (2016): Image super-resolution: The techniques, applications, and future. In: Signal Processing, vol. 128, S. 389–408. DOI: 10.1016/j.sigpro.2016.05.002.
- [ZhLi2018] Zhou, Yuqian; Liu, Ding; Huang, Thomas (2018): Survey of Face Detection on Low-Quality Images. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), S. 769–773. DOI: 10.1109/FG.2018.00121.
- [ZhRa2010] Zheng, Jun; Ramírez, Geovany A.; Fuentes, Olac (2010): Face Detection in Low-Resolution Color Images. In: Campilho, A., Kamel, M. (Hg.) Image Analysis and Recognition. ICIAR 2010. Lecture Notes in Computer Science, vol. 6111, S. 454–463. Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-642-13772-3\_46.

#### Onlinequellen:

- [Ad2022] Adobe (2022): So einfach funktioniert der Gaußsche Weichzeichner. Online verfügbar unter <https://www.adobe.com/de/creativecloud/photography/discover/gaussian-blur.html>, zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [AO2022] AO Kaspersky Lab (2022): What is Facial Recognition – Definition and Explanation. Online verfügbar unter <https://www.kaspersky.com/resource-center/definitions/what-is-facial-recognition>, zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [Ca2022] Cambridge in Colour (2022): Understanding Digital Image Interpolation. Online verfügbar unter <https://www.cambridgeincolour.com/tutorials/image-interpolation.htm>, zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [DaKI2020] Dannoritzer, Marlen; Klauth, Jan (2020): Eine deutsche Metropole gehört zu den am stärksten überwachten Städten. In: WELT, 03.08.2020. Online verfügbar unter <https://www.welt.de/wirtschaft/article212708481/Kontrolle-per-Kamera-Die-am-staerksten-ueberwachten-Staedte-der-Welt.html>, zuletzt geprüft am 09.09.2022.

- [De2021] Deepchecks (2021): What is False Positive Rate | Deepchecks. Online verfügbar unter <https://deepchecks.com/glossary/false-positive-rate/>, zuletzt aktualisiert am 05.08.2021, zuletzt geprüft am 31.08.2022.
- [Do2022a] Doxygen (2022): OpenCV: Introduction. Online verfügbar unter <https://docs.opencv.org/4.x/d1/dfb/intro.html>, zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [Do2022b] Doxygen (2022): OpenCV: Basic Operations on Images. Online verfügbar unter [https://docs.opencv.org/4.x/d3/df2/tutorial\\_py\\_basic\\_ops.html](https://docs.opencv.org/4.x/d3/df2/tutorial_py_basic_ops.html), zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [Do2022c] Doxygen (2022): OpenCV: Operations on arrays. Online verfügbar unter [https://docs.opencv.org/3.4/d2/de8/group\\_\\_core\\_\\_array.html#gafafb2513349db3bcff51f54ee5592a19](https://docs.opencv.org/3.4/d2/de8/group__core__array.html#gafafb2513349db3bcff51f54ee5592a19), zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [Do2022d] Doxygen (2022): OpenCV: Smoothing Images. Online verfügbar unter [https://docs.opencv.org/4.x/dc/dd3/tutorial\\_gaussian\\_median\\_blur\\_bilateral\\_filter.html](https://docs.opencv.org/4.x/dc/dd3/tutorial_gaussian_median_blur_bilateral_filter.html), zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [Go2020] Google LLC (2020): MediaPipe Face Mesh. Online verfügbar unter [https://google.github.io/mediapipe/solutions/face\\_mesh](https://google.github.io/mediapipe/solutions/face_mesh), zuletzt aktualisiert am 29.06.2022, zuletzt geprüft am 29.08.2022.
- [GTFD2022] Autor unbekannt: Georgia Tech Face Database. Online verfügbar unter [http://www.anefian.com/face\\_reco.htm](http://www.anefian.com/face_reco.htm), zuletzt geprüft am 05.09.2022.
- [Hö2021] Högel, Bianca (2021): Euklidischer Abstand. Online verfügbar unter [https://www.biancahoegel.de/geometrie/koerper/abstand\\_euklid.html](https://www.biancahoegel.de/geometrie/koerper/abstand_euklid.html), zuletzt aktualisiert am 08.05.2021, zuletzt geprüft am 31.08.2022.
- [IT2022] ITWissen.info (2022): Gammakorrektur. Online verfügbar unter <https://www.itwissen.info/Gammakorrektur-gamma-correction.html>, zuletzt aktualisiert am 26.08.2022, zuletzt geprüft am 26.08.2022.

- [KeJe2022] Kennedy, John; Jenks, Alma; White, Steven (2022): RGB-Farbräume, 11.05.2022. Online verfügbar unter <https://docs.microsoft.com/de-de/windows/win32/wcs/rgb-color-spaces>, zuletzt geprüft am 09.09.2022.
- [LuCl2022a] Lundh, Fredrik; Clark, Alex and Contributors (2022): Image Module. Online verfügbar unter <https://pillow.readthedocs.io/en/stable/reference/Image.html>, zuletzt aktualisiert am 01.07.2022, zuletzt geprüft am 31.08.2022.
- [LuCl2022b] Lundh, Fredrik; Clark, Alex and Contributors (2022): ImageFilter Module. Online verfügbar unter <https://pillow.readthedocs.io/en/stable/reference/ImageFilter.html>, zuletzt aktualisiert am 01.07.2022, zuletzt geprüft am 31.08.2022.
- [Pt2022] Ptc (2022): Lightweight models. Online verfügbar unter [https://support.ptc.com/help/creo/ced\\_modeling/r20.5.0.0/en/index.html#page/ced\\_modeling/OSDM\\_Main/LightweightModels.html](https://support.ptc.com/help/creo/ced_modeling/r20.5.0.0/en/index.html#page/ced_modeling/OSDM_Main/LightweightModels.html), zuletzt aktualisiert am 01.09.2022, zuletzt geprüft am 01.09.2022.
- [Ro2020] Rougetet, Arnaud (2020): Flickr-Faces-HQ Dataset (FFHQ). Online verfügbar unter <https://www.kaggle.com/datasets/arnaud58/flickrfaceshq-dataset-ffhq>, zuletzt aktualisiert am 09.03.2020, zuletzt geprüft am 31.08.2022.
- [St2022] Statista (2022): Anzahl der Überwachungskameras (CCTV) je 1.000 Einwohner in ausgewählten Großstädten weltweit. Online verfügbar unter <https://de.statista.com/statistik/daten/studie/1136552/umfrage/cctv-kameras-in-ausgewaehlten-grossstaedten/>, zuletzt aktualisiert am 02.09.2022, zuletzt geprüft am 02.09.2022.
- [TL2021] Team Learnopencv (2021): Image Resizing with OpenCV. In: LearnOpenCV, 26.05.2021. Online verfügbar unter <https://learnopencv.com/image-resizing-with-opencv/>, zuletzt geprüft am 31.08.2022.
- [ZVA2022] Zentralverband der Augenoptiker und Optometristen (2022): Brillenstudie. Online verfügbar unter <https://www.zva.de/brillenstudie>, zuletzt aktualisiert am 30.08.2022, zuletzt geprüft am 30.08.2022.

## Codequellen:

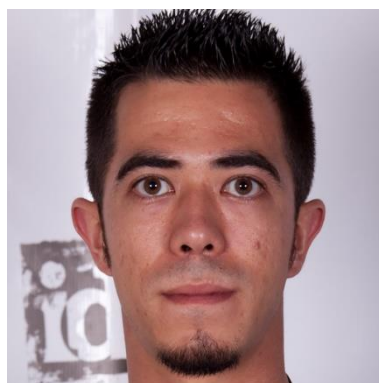
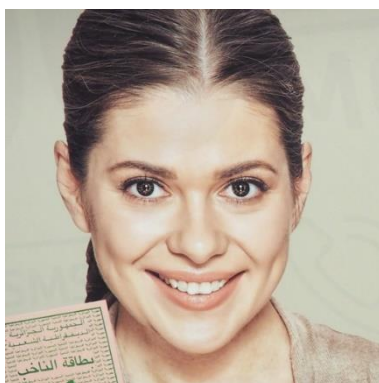
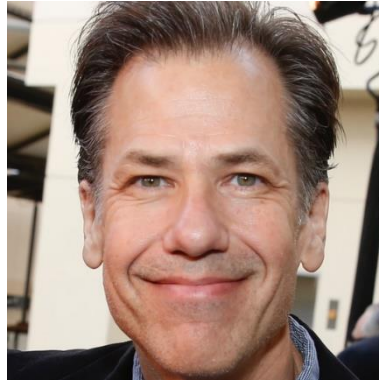
- [Ge2019] Adaptiert nach: GeeksforGeeks (2019): Python PIL | GaussianBlur method. Online verfügbar unter <https://www.geeksforgeeks.org/python-pil-gaussianblur-method/>, zuletzt aktualisiert am 14.07.2019, zuletzt geprüft am 31.08.2022.
- [Ge2020] Adaptiert nach: GeeksforGeeks (2020): Changing the contrast and brightness of an image using Python - OpenCV. Online verfügbar unter <https://www.geeksforgeeks.org/changing-the-contrast-and-brightness-of-an-image-using-python-opencv/>, zuletzt aktualisiert am 12.12.2021, zuletzt geprüft am 29.08.2022.
- [Se2022] Adaptiert nach: Seltmann, Stefan (2022): VERY Best Practice: Arbeiten in Python mit Pfaden - Teil 2. Online verfügbar unter <https://www.btelligent.com/blog/best-practice-arbeiten-in-python-mit-pfaden-teil-2/>, zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.
- [StOv2022] Adaptiert nach: Stack Overflow (2022): How to pixelate image using OpenCV in Python? Online verfügbar unter <https://stackoverflow.com/questions/55508615/how-to-pixelate-image-using-opencv-in-python>, zuletzt aktualisiert am 31.08.2022, zuletzt geprüft am 31.08.2022.

# Anlagen

Teil 1 .....	A-1
Teil 2 .....	A-2
Teil 3 .....	A-3
Teil 4 .....	A-4
Teil 5 .....	A-5
Teil 6 .....	A-6
Teil 7 .....	A-7
Teil 8 .....	A-10
Teil 9 .....	A-16
Teil 10 .....	A-19

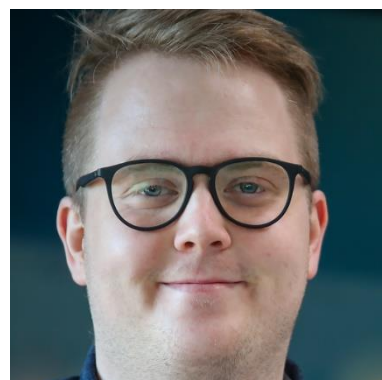
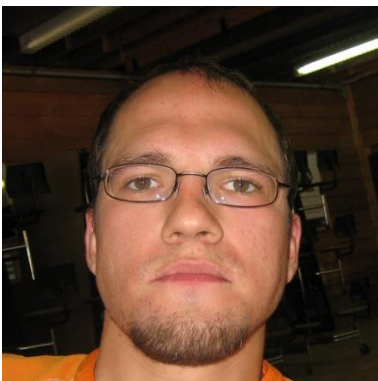
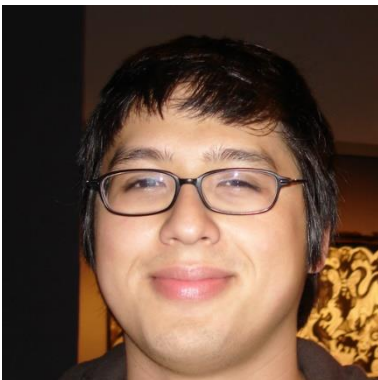
# Anlagen, Teil 1

Die folgenden Bilder entsprechen der Kategorie „frontal“, dargestellt in etwa einem Drittel ihrer eigentlichen Größe.



## Anlagen, Teil 2

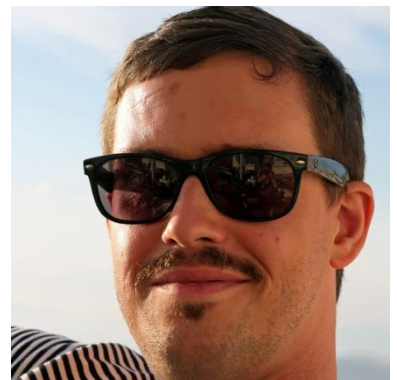
Die im Folgenden dargelegten Bilder, dargestellt in etwa einem Drittel ihrer eigentlichen Größe, sind der Kategorie „mit Brille“ zuzuordnen.





## Anlagen, Teil 3

Die folgenden Bilder, dargestellt in etwa einem Drittel ihrer eigentlichen Größe, repräsentieren die Bildkategorie „mit Sonnenbrille“.



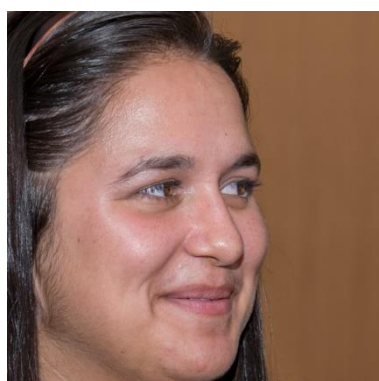
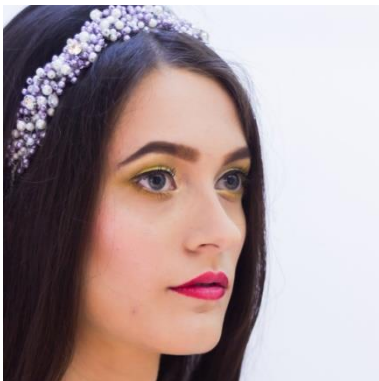
## Anlagen, Teil 4

Die folgenden Bilder, dargestellt in etwa einem Drittel ihrer eigentlichen Größe, stellen eine Auswahl des Ausgangsmaterials entsprechend der Kategorie „links lateral“ dar.



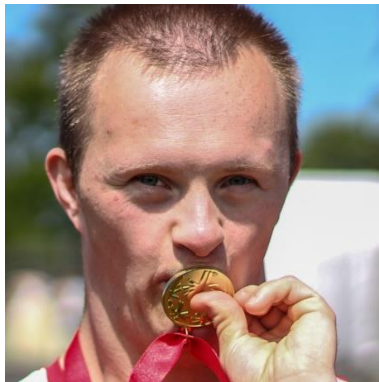
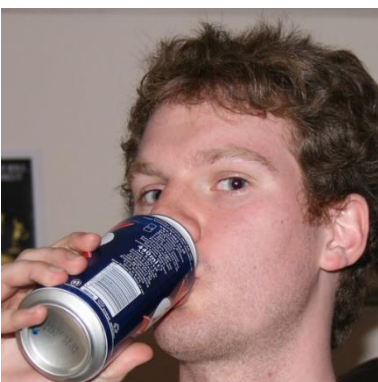
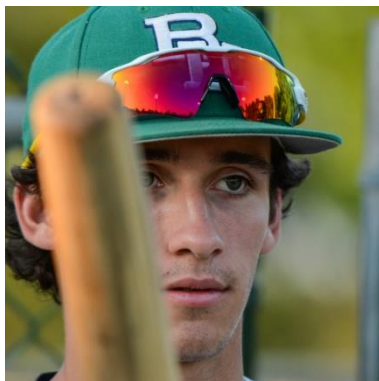
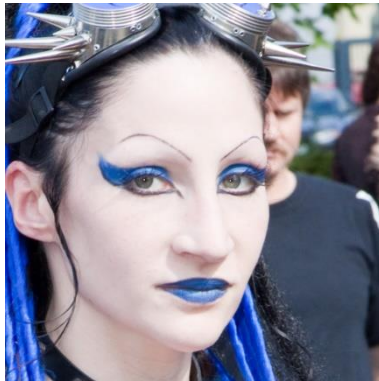
## Anlagen, Teil 5

Bei den folgenden Bildern, dargestellt in etwa einem Drittel ihrer eigentlichen Größe, handelt es sich um solche aus der Kategorie „rechts lateral“.



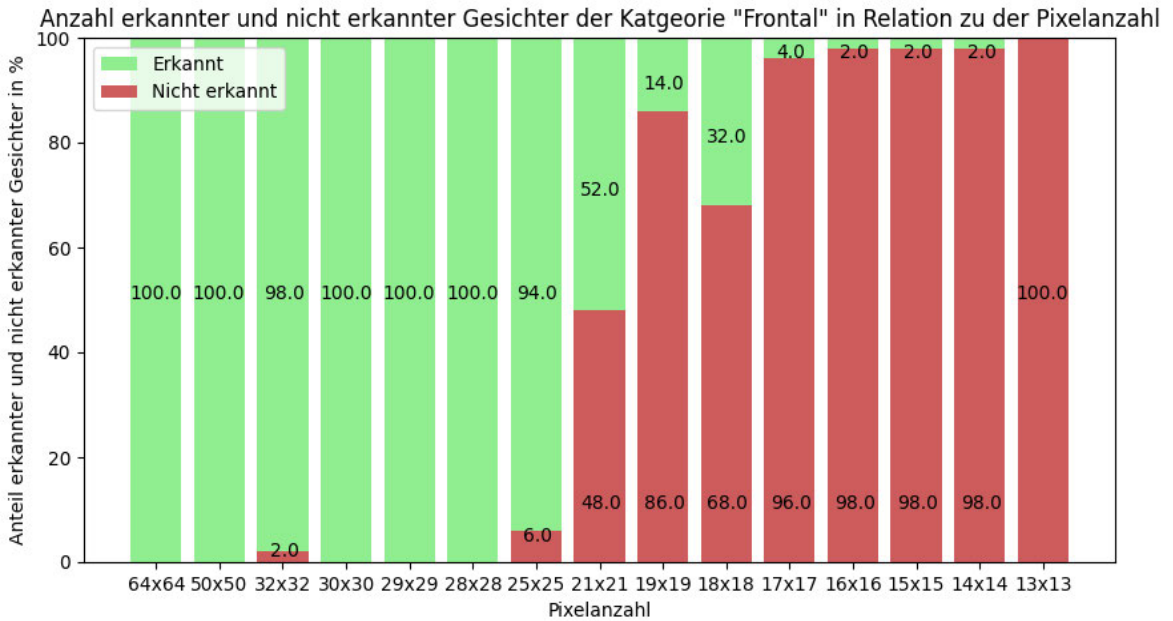
## Anlagen, Teil 6

Die folgenden Bilder, dargestellt in etwa einem Drittel ihrer eigentlichen Größe, entsprechen der Kategorie „teilweise verdeckt“.

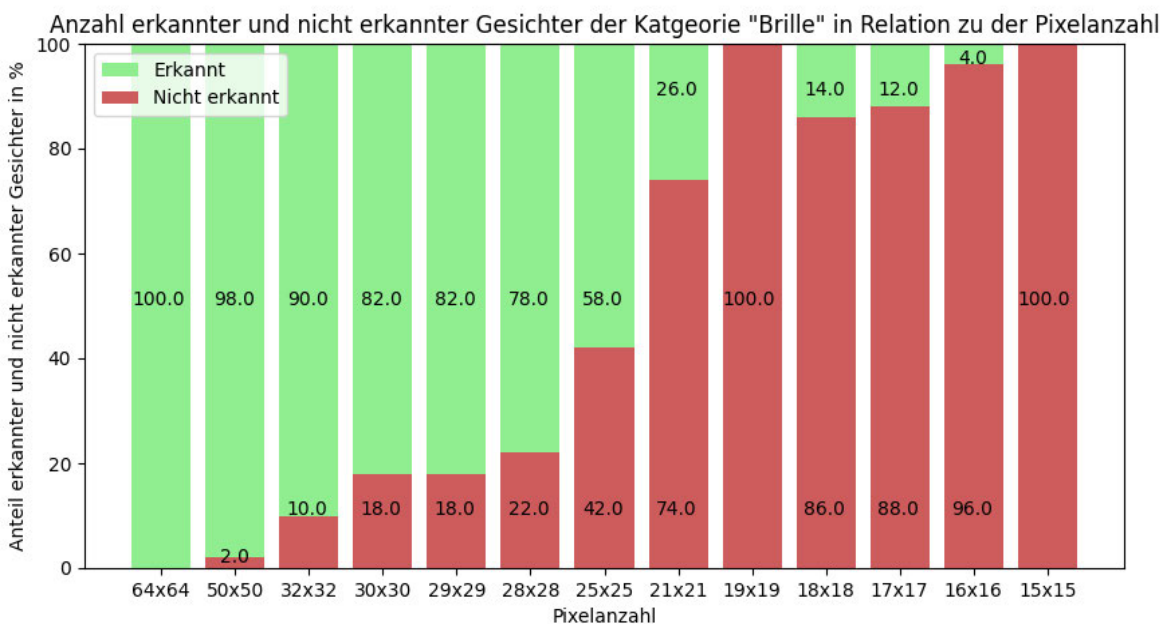


# Anlagen, Teil 7

Im Folgenden werden die Ergebniswerte der einzelnen Bildkategorien in Bezug auf die Informationsreduktion „Verpixeln“, dargestellt in Kapitel 5.1, jeweils in Form eines Diagramms visuell veranschaulicht. Die Reihenfolge orientiert sich an der zuvor festgelegten. Das folgende Diagramm zeigt die Ergebnisse aus Tabelle 1 in graphischer Form.

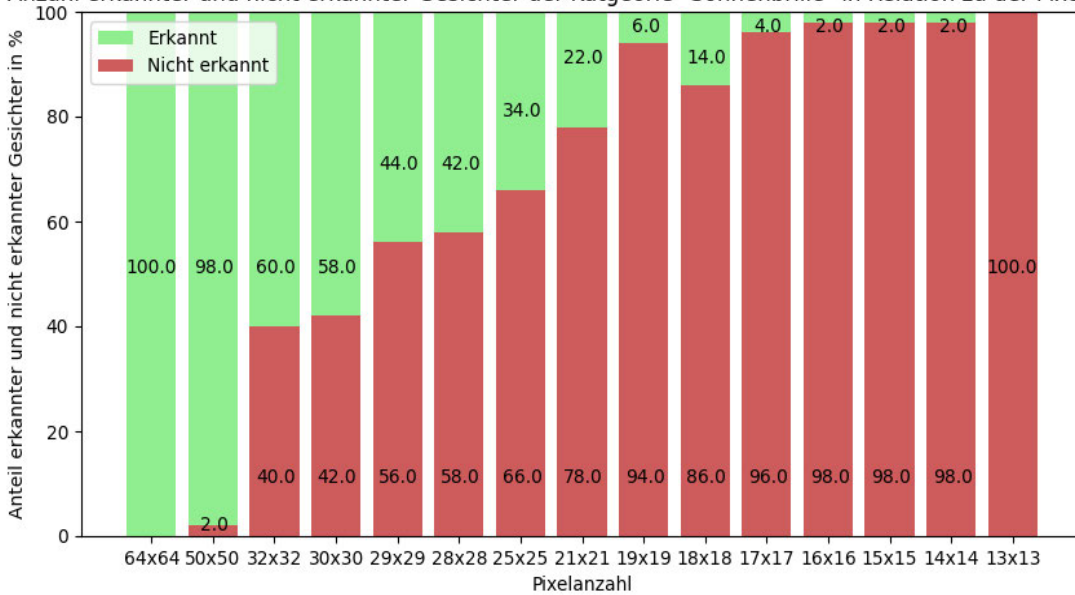


Folgendes Diagramm veranschaulicht die Ergebnisse aus Tabelle 2.



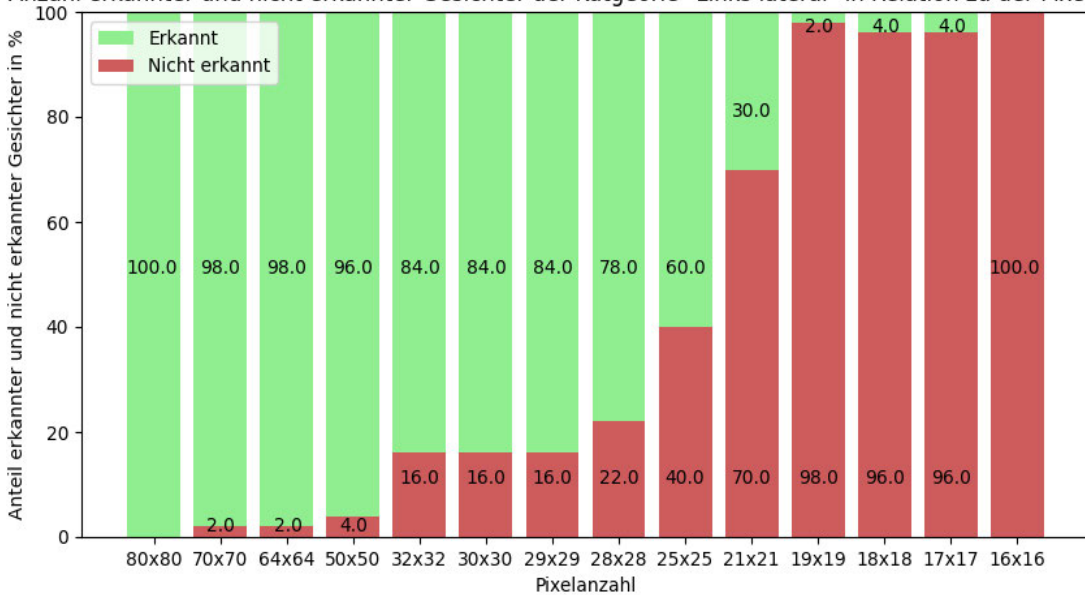
Das nachstehende Diagramm repräsentiert die in Tabelle 3 dargestellten Ergebnisse.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Sonnenbrille" in Relation zu der Pixelanzahl



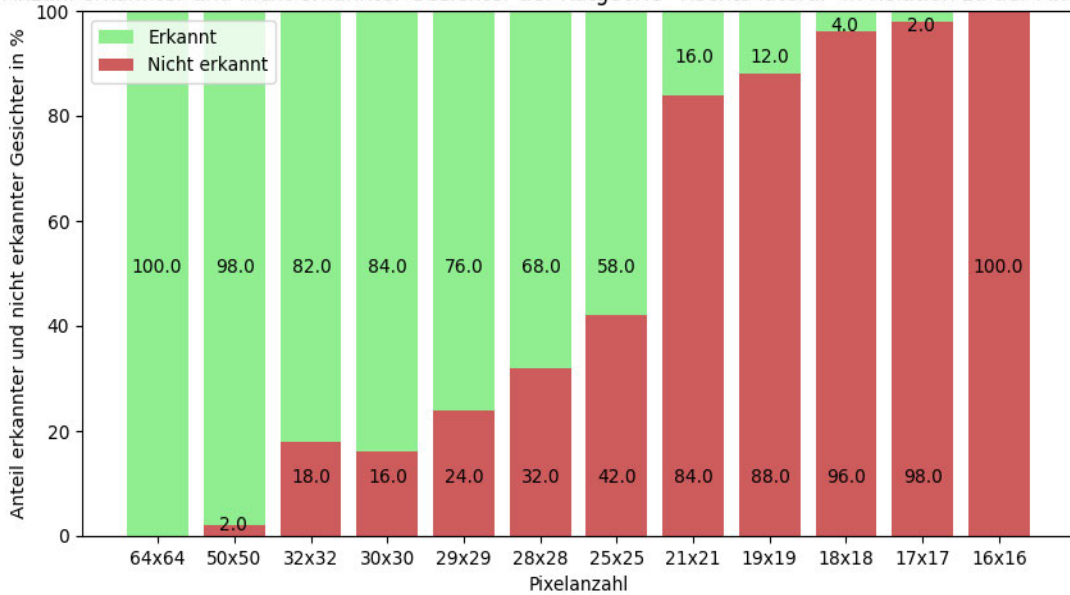
Im Anschluss können die Ergebnisse aus Tabelle 4 dem Diagramm entnommen werden.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Links lateral" in Relation zu der Pixelanzahl



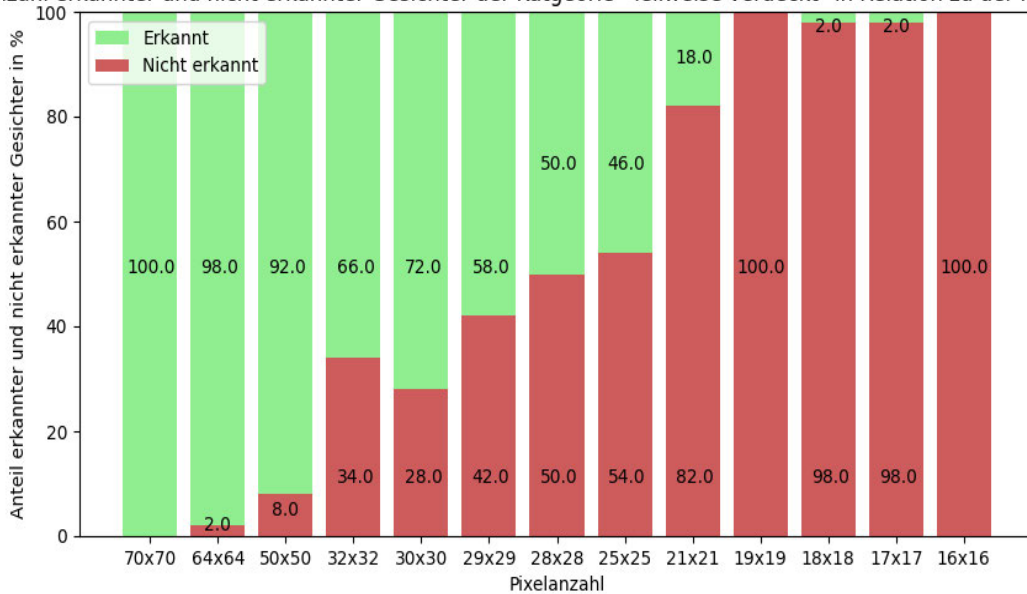
Dem folgenden Diagramm können die Ergebnisse aus Tabelle 5 entnommen werden.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Rechts lateral" in Relation zu der Pixelanzahl



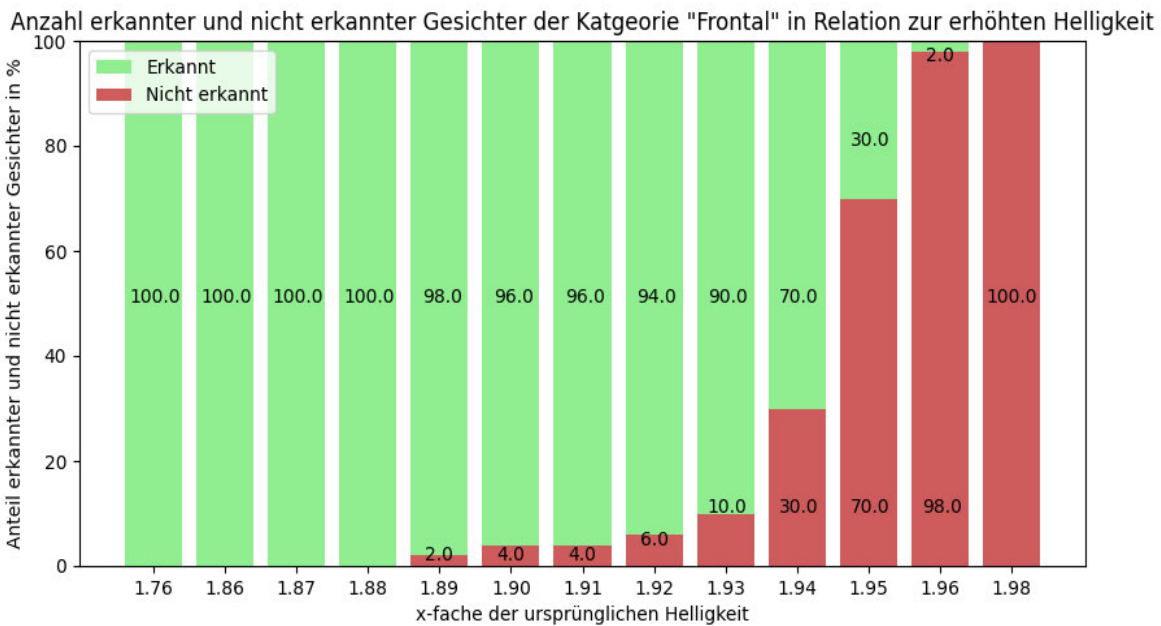
Das nachstehende Diagramm stellt die Ergebnisse aus Tabelle 6 in graphischer Form dar.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Teilweise verdeckt" in Relation zu der Pixelanzahl

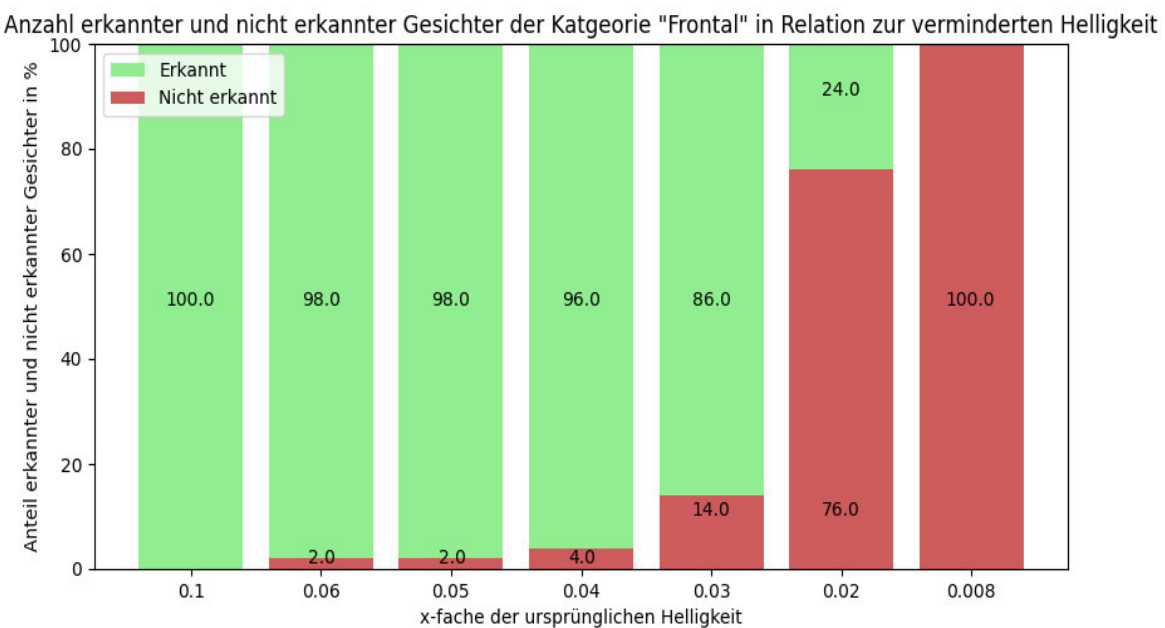


# Anlagen, Teil 8

Die nachfolgenden Abbildungen stellen die Ergebniswerte der Gesichtserkennung mittels MediaPipe Face Mesh in Bezug auf die Informationsreduktions-Kategorie „Veränderte Helligkeit“ pro Bildkategorie, siehe Kapitel 5.2, in Form eines Diagramms dar. Das nachstehende Diagramm veranschaulicht die Ergebnisse aus Tabelle 7 in graphischer Form.

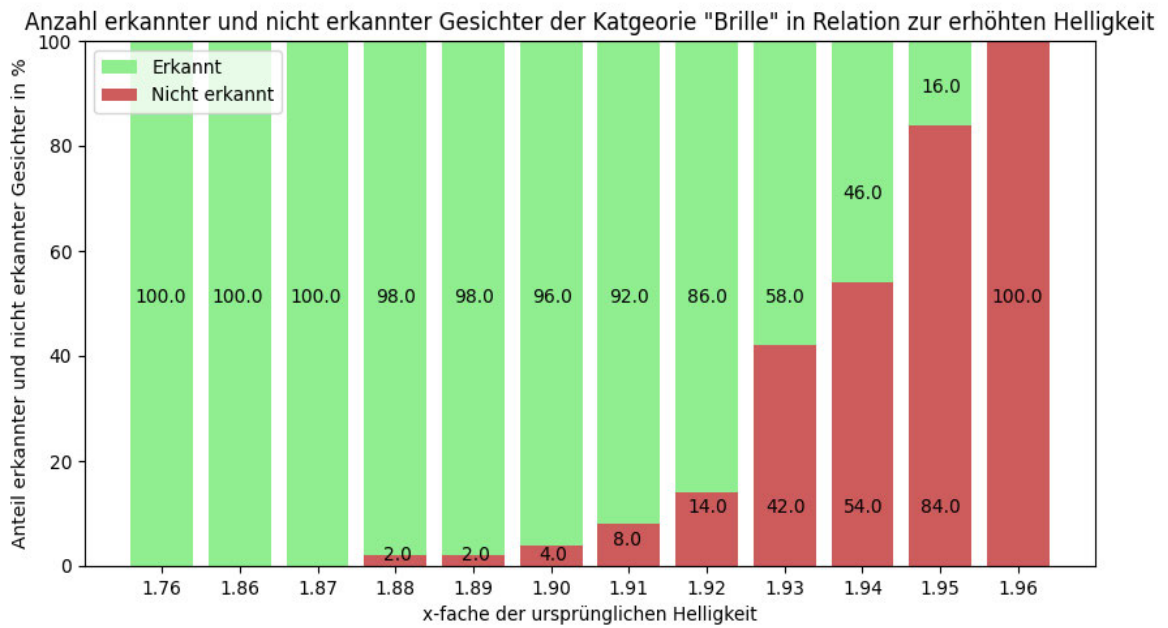


Folgende Graphik zeigt die Ergebnisse aus Tabelle 8 in Form eines Diagramms.

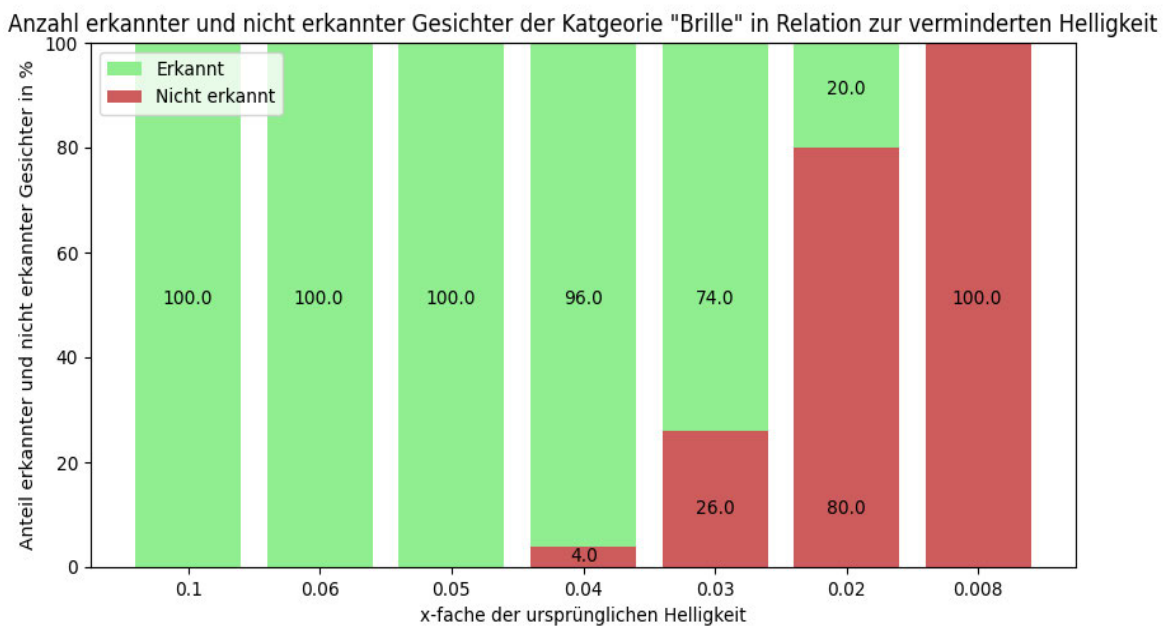




Die folgende Graphik repräsentiert die Ergebnisse aus Tabelle 9.

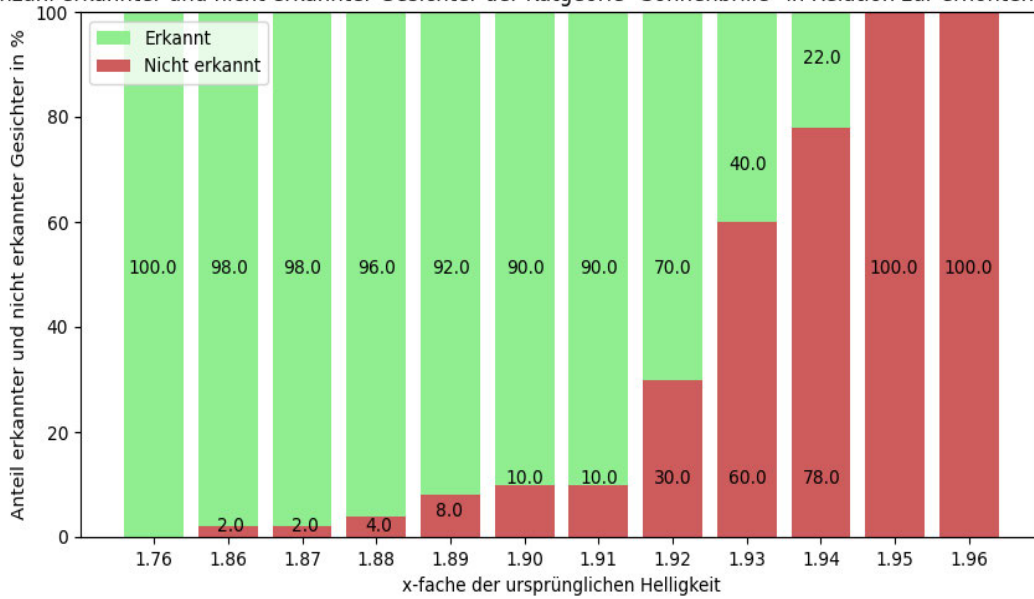


Nachstehende Graphik zeigt die Ergebnisse aus Tabelle 10.



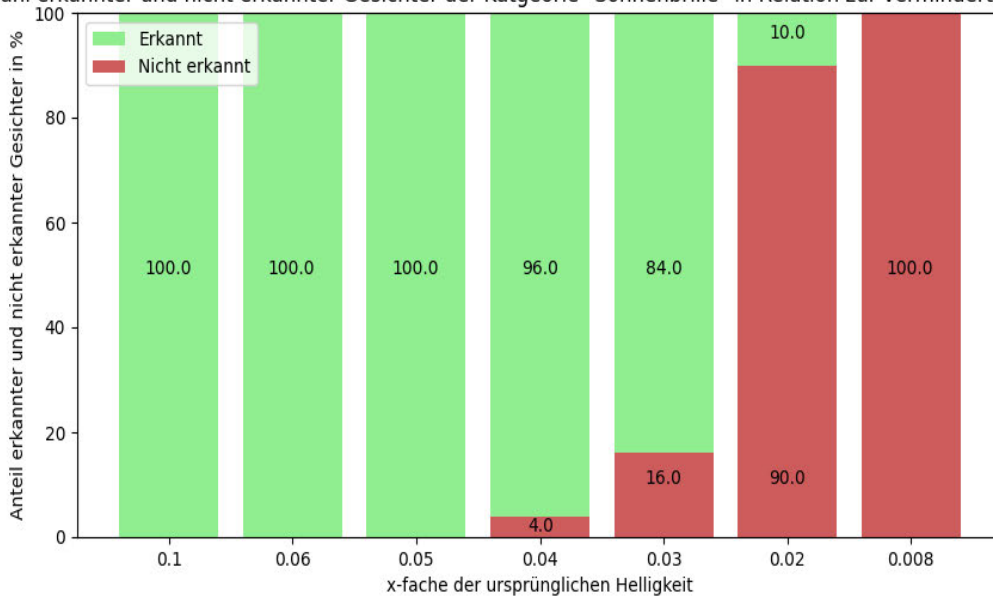
Im Anschluss können dem Diagramm die Ergebnisse aus Tabelle 11 entnommen werden.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Sonnenbrille" in Relation zur erhöhten Helligkeit



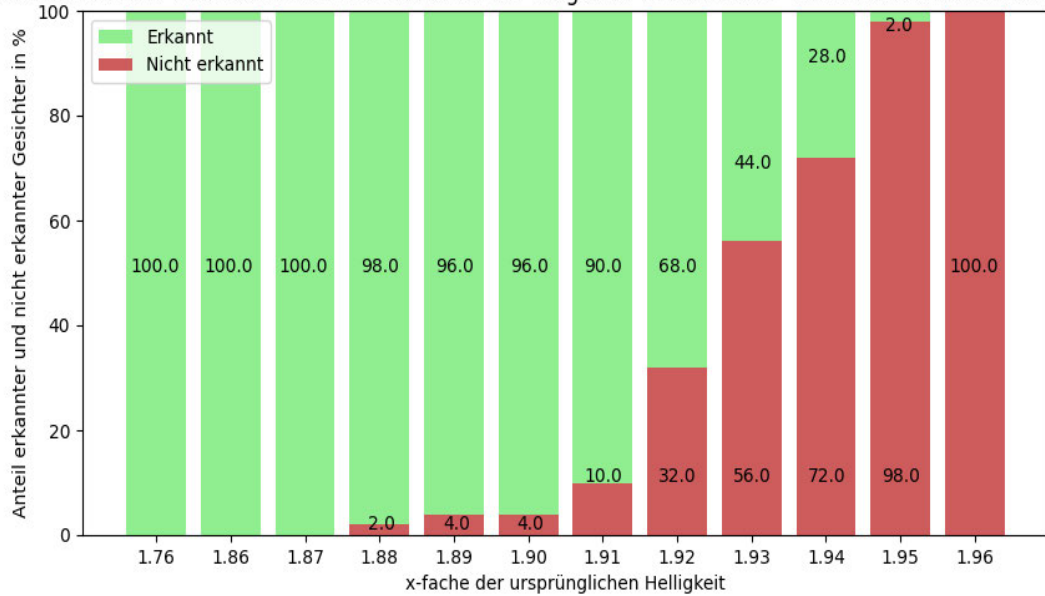
Folgendes Diagramm visualisiert die Ergebniswerte aus Tabelle 12.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Sonnenbrille" in Relation zur verminderten Helligkeit



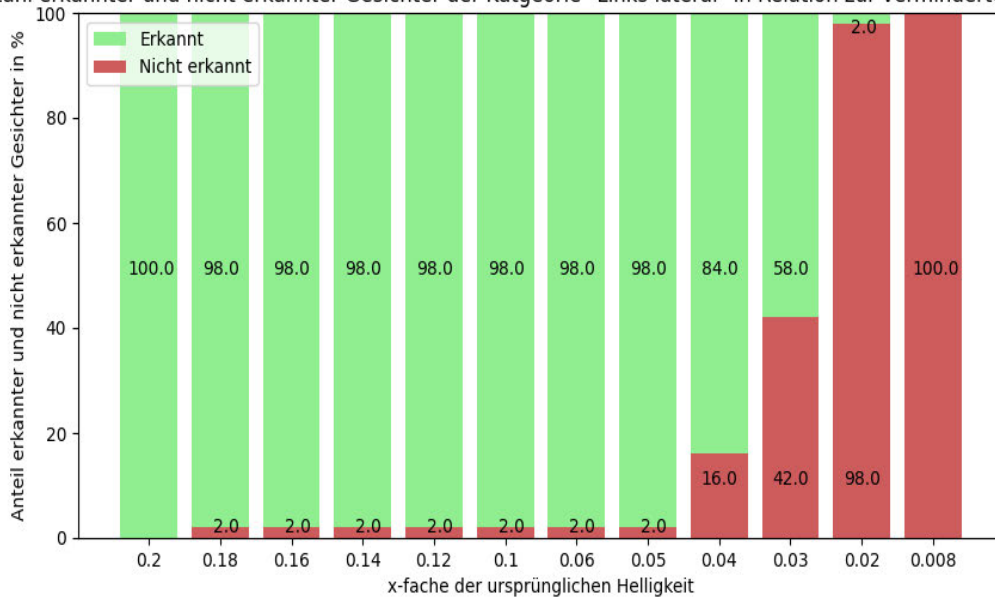
Dem anschließenden Diagramm können die Ergebnisse aus Tabelle 13 entnommen werden.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Links lateral" in Relation zur erhöhten Helligkeit



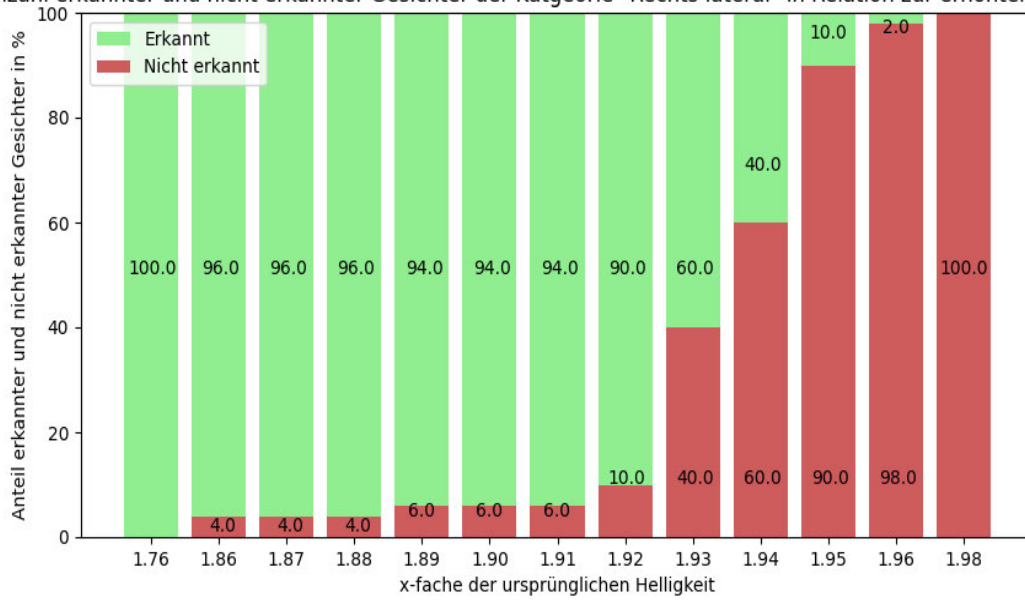
Die folgende Graphik zeigt die Ergebnisse aus Tabelle 14 in Form eines Diagramms.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Links lateral" in Relation zur verminderten Helligkeit



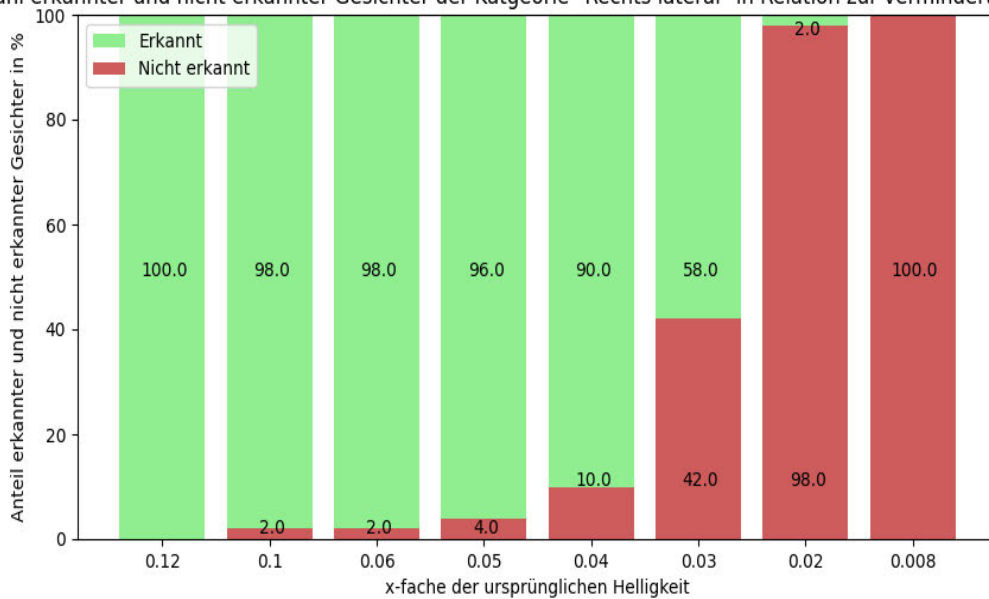
Folgende Graphik veranschaulicht die Ergebnisse aus Tabelle 15.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Rechts lateral" in Relation zur erhöhten Helligkeit



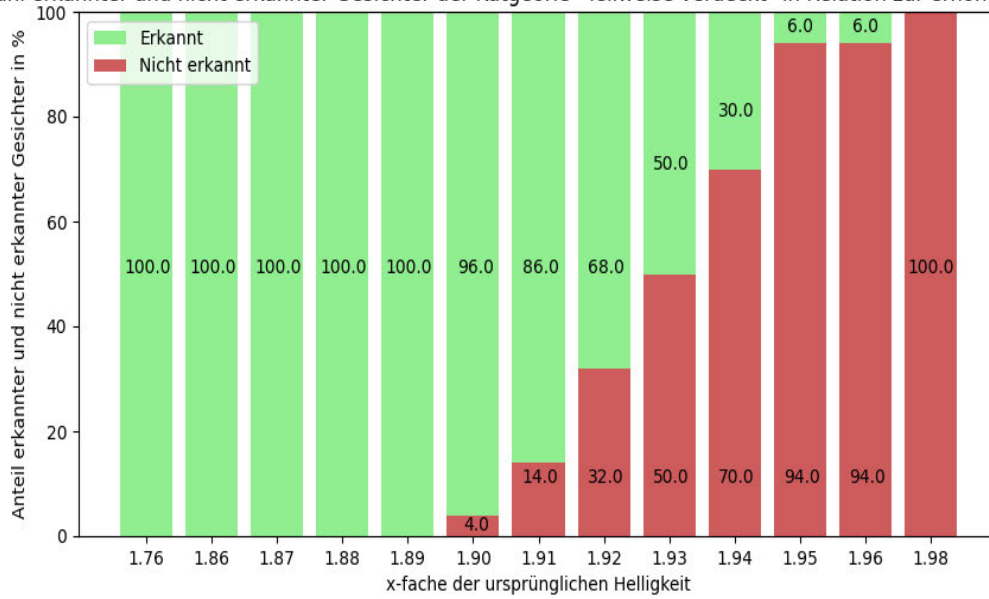
Die folgende Graphik repräsentiert die Ergebnisse aus Tabelle 16.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Rechts lateral" in Relation zur verminderten Helligkeit



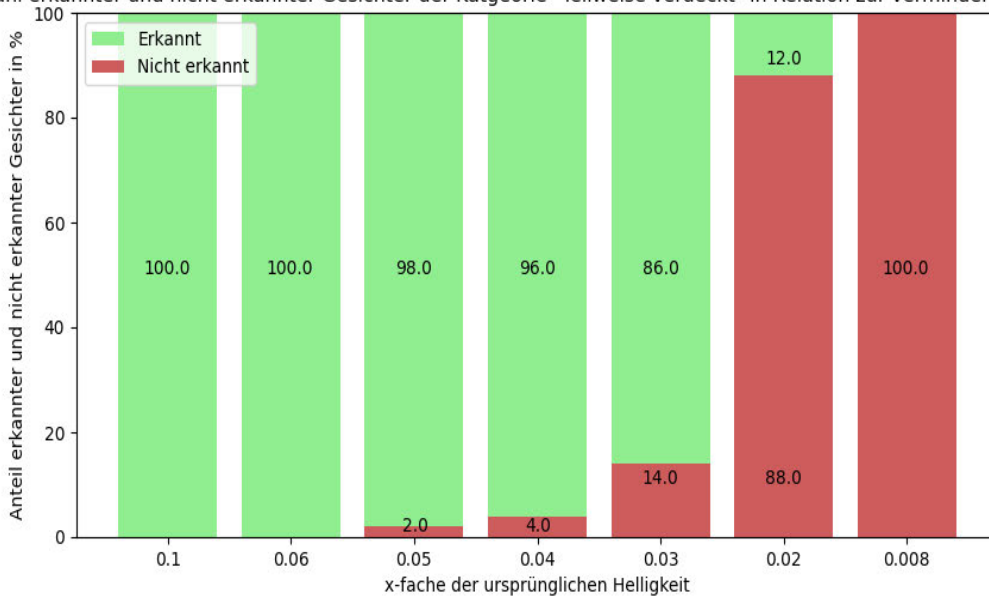
Dem folgenden Diagramm können die Ergebnisse aus Tabelle 17 entnommen werden.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Teilweise verdeckt" in Relation zur erhöhten Helligkeit



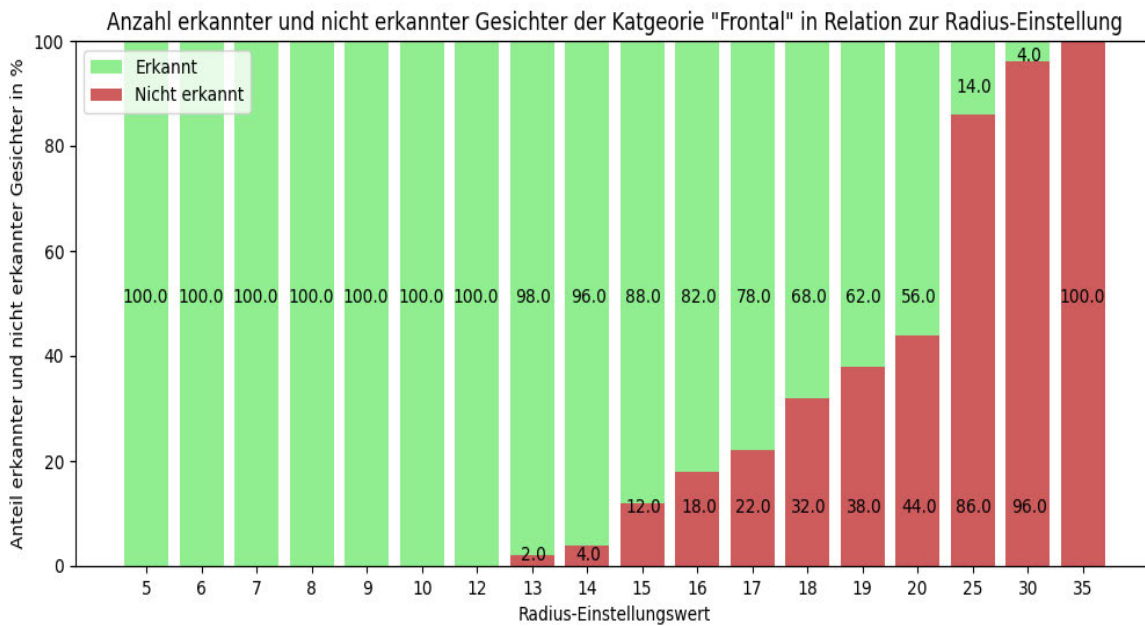
Nachfolgendes Diagramm veranschaulicht die Ergebnisse aus Tabelle 18.

Anzahl erkannter und nicht erkannter Gesichter der Kategorie "Teilweise verdeckt" in Relation zur verminderten Helligkeit

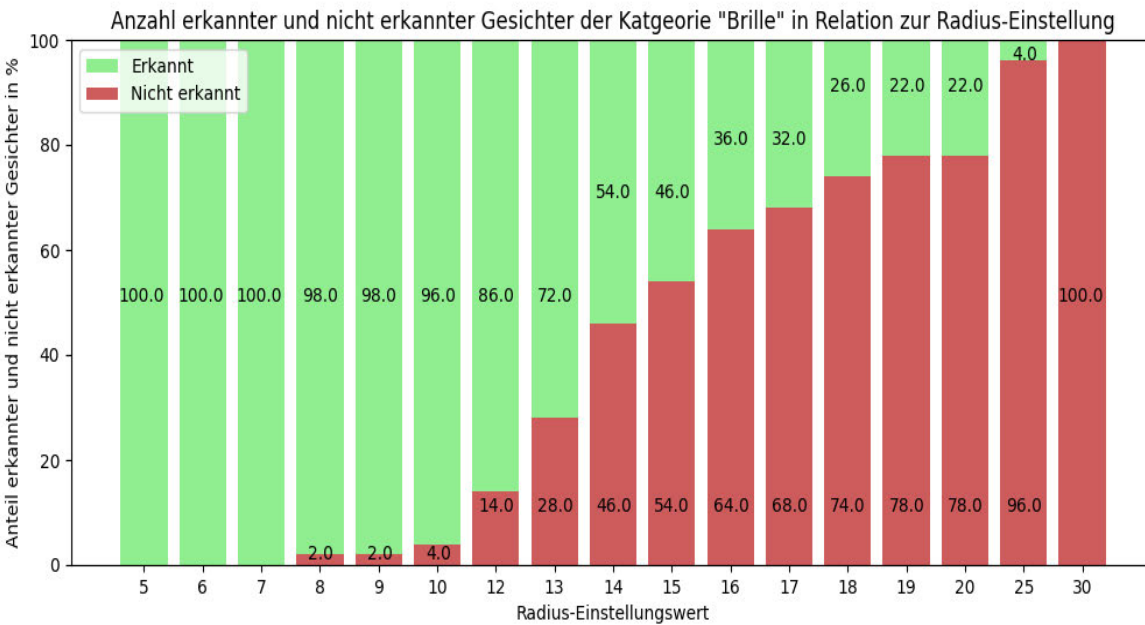


# Anlagen, Teil 9

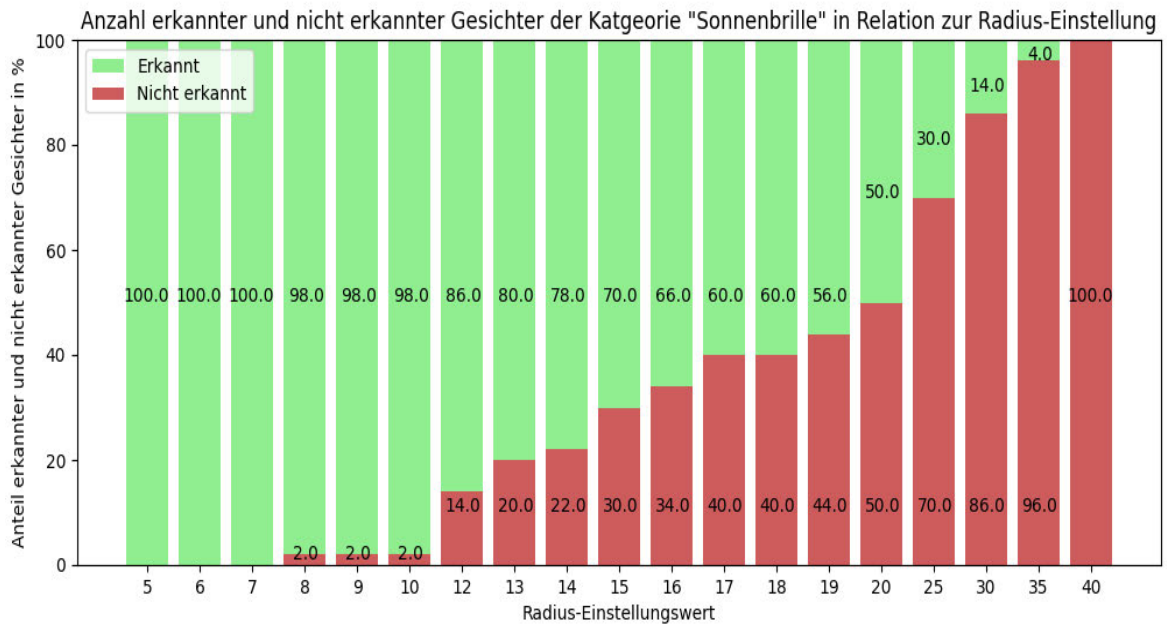
Im Folgenden werden die Ergebnisse bezüglich der Gesichtserkennung bei Informationsreduktion mit „Gauß’scher Unschärfe“ der einzelnen Bildkategorien, siehe Kapitel 5.3, jeweils in einem Diagramm veranschaulicht. Das nachstehende Diagramm zeigt die Ergebnisse aus Tabelle 19.



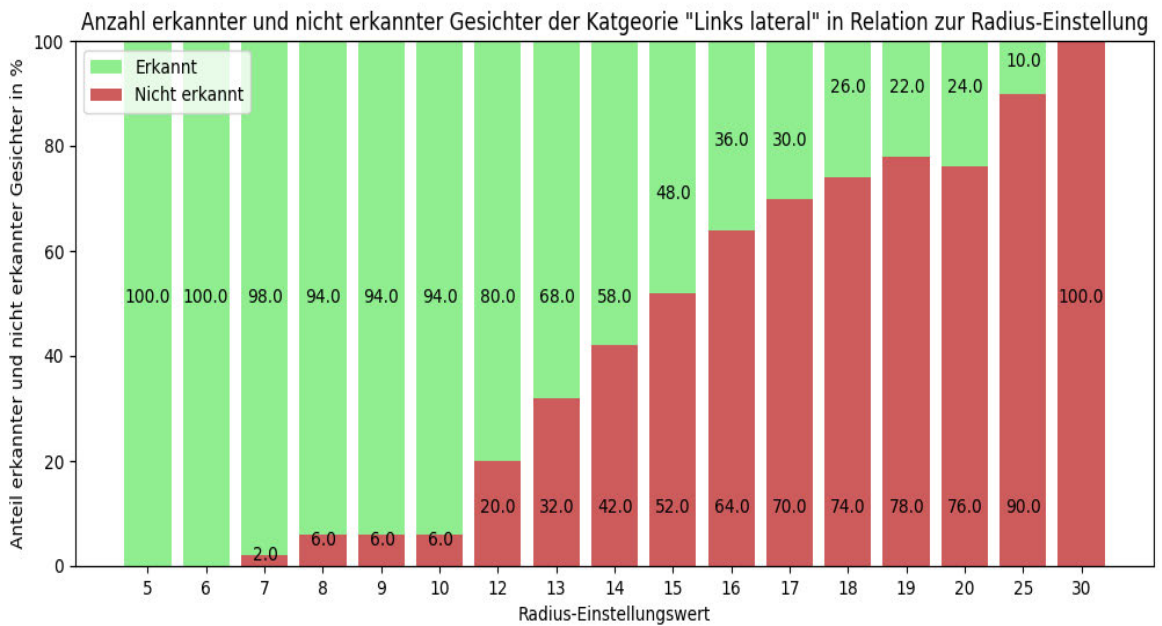
Die folgende Graphik veranschaulicht die Ergebnisse aus Tabelle 20.



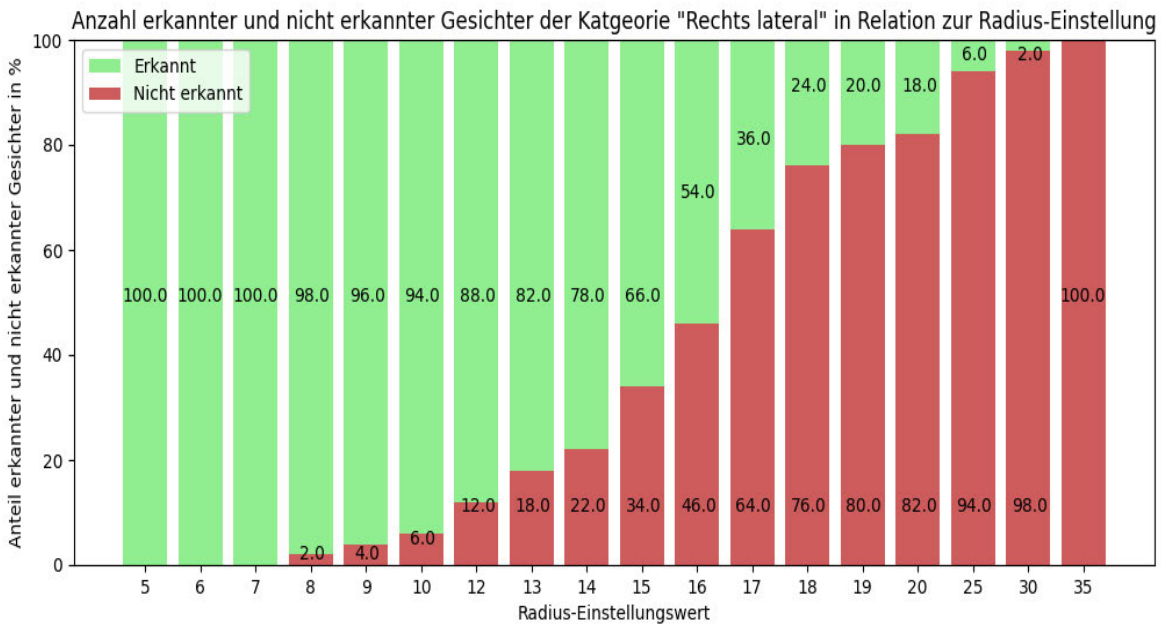
Im Folgenden sind die Ergebnisse aus Tabelle 21 in Form eines Diagramms dargestellt.



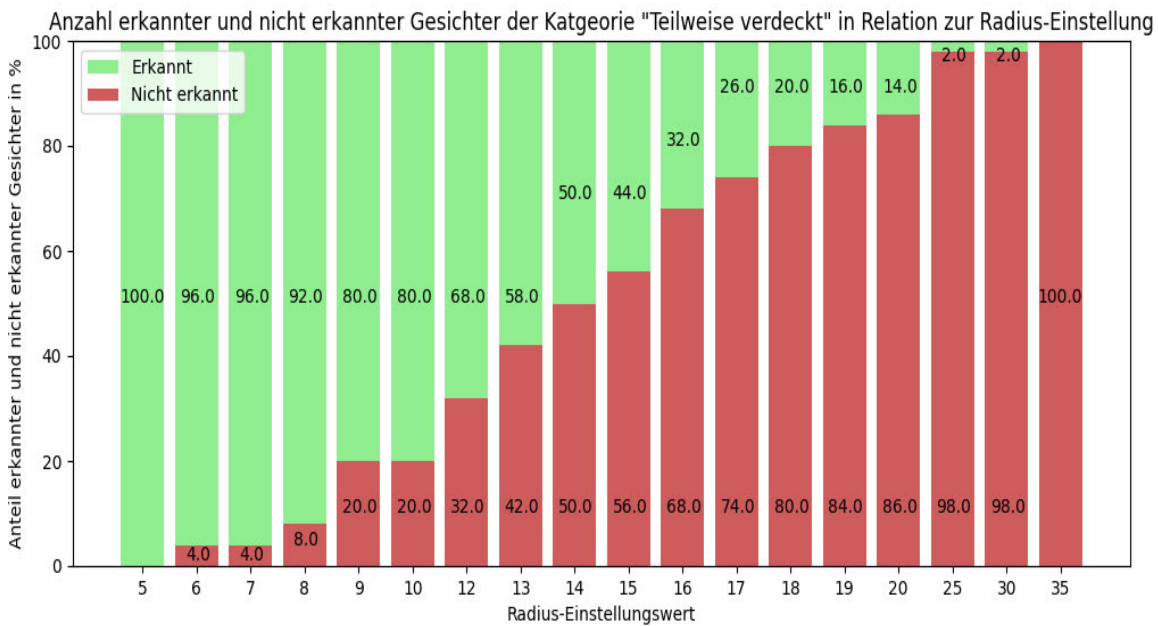
Die folgende Graphik veranschaulicht die Ergebnisse aus Tabelle 22.



Die folgende Graphik zeigt die Ergebnisse aus Tabelle 23 in Form eines Diagramms.



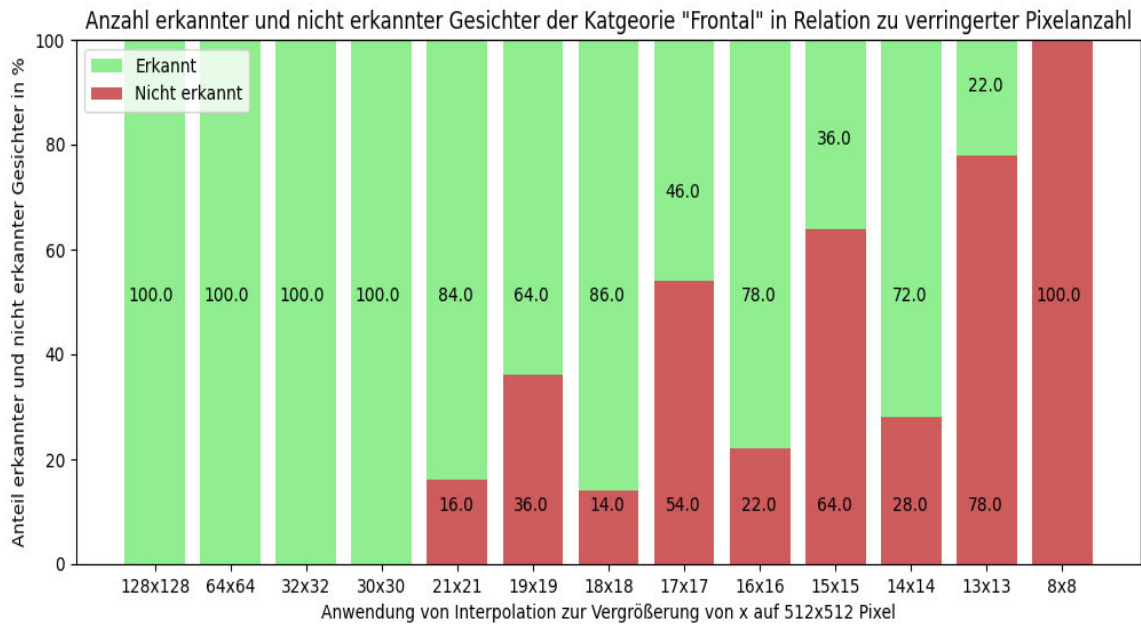
Nachstehendes Diagramm stellt die Ergebnisse aus Tabelle 24.



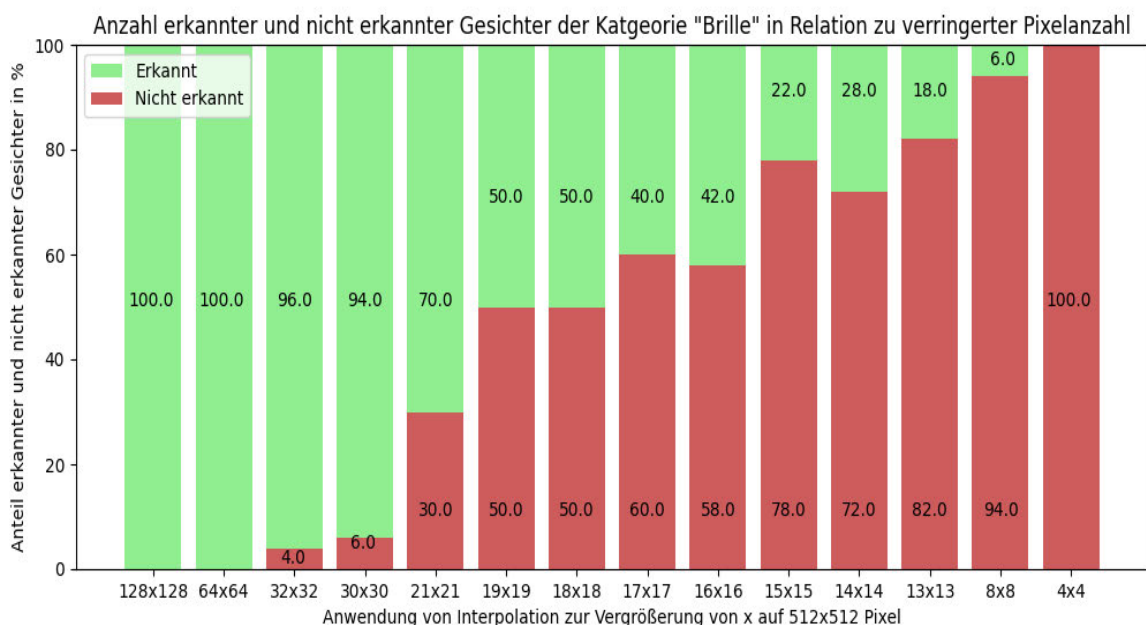


# Anlagen, Teil 10

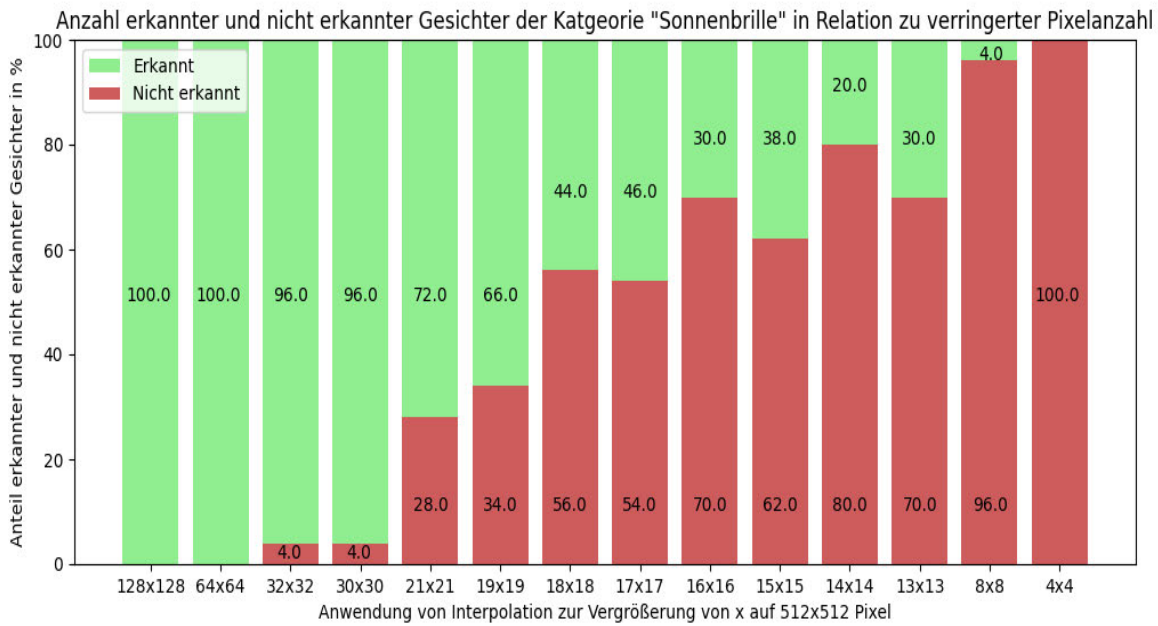
In diesem Abschnitt wird für jede Bildkategorie ein Diagramm dargestellt, welches die Ergebniswerte der Gesichtserkennung in Bezug auf zur Informationsreduktion angewendete Interpolation veranschaulicht. Die folgende Graphik veranschaulicht die Ergebnisse aus Tabelle 25.



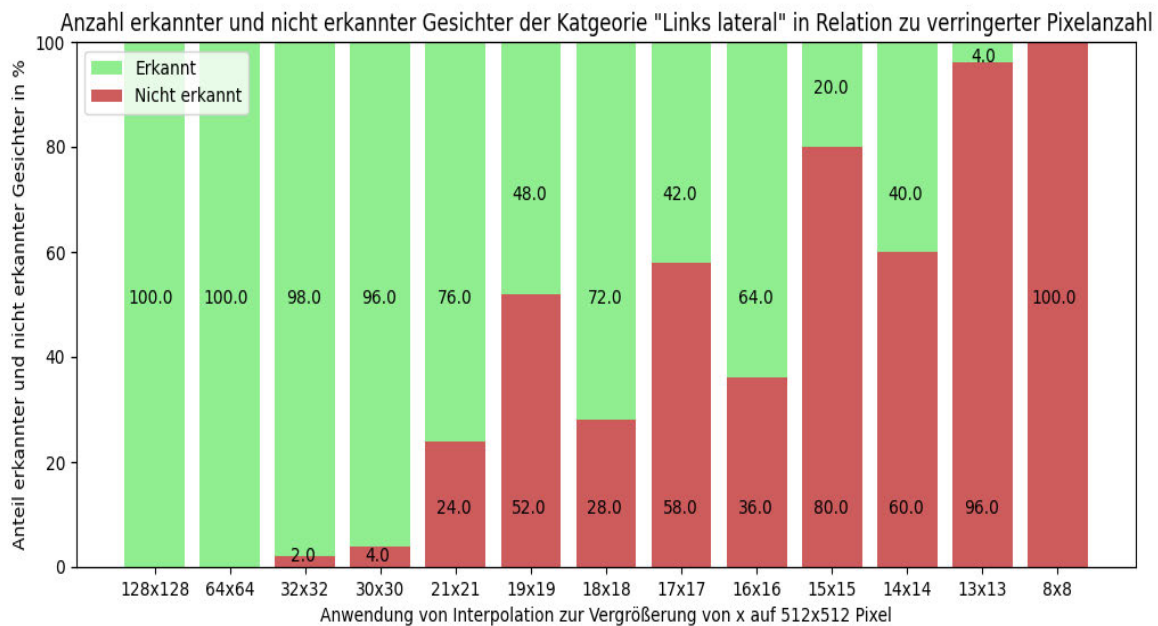
Die nachstehende Graphik stellt die Ergebnisse aus Tabelle 26 dar.



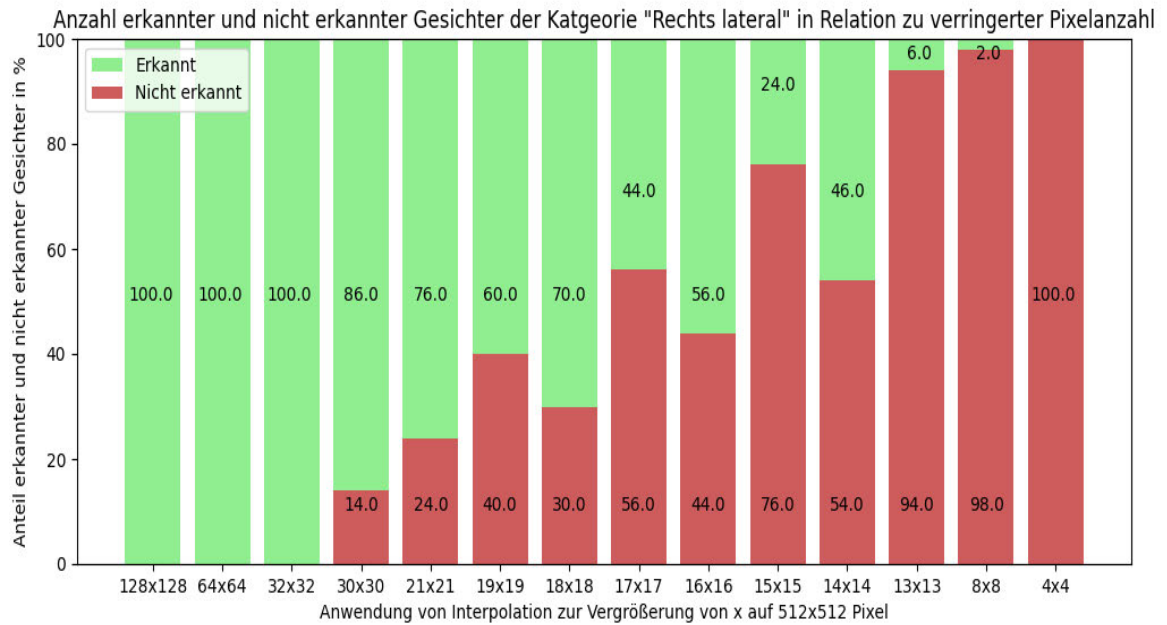
Im Folgenden sind die Ergebnisse aus Tabelle 27 in Form eines Diagramms dargestellt.



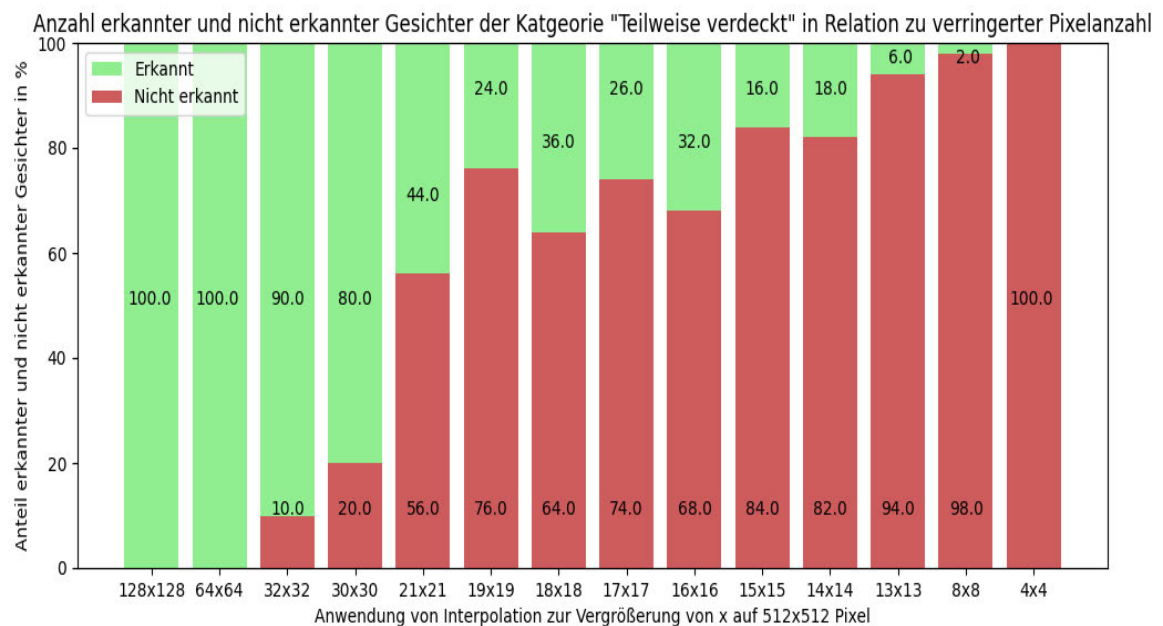
Nachstehendes Diagramm zeigt die Ergebnisse aus Tabelle 28.



Folgende Graphik veranschaulicht die Ergebnisse aus Tabelle 29 in Form eines Diagramms.



Die nachstehende Graphik stellt die Ergebnisse aus Tabelle 30 in Form eines Diagramms dar.



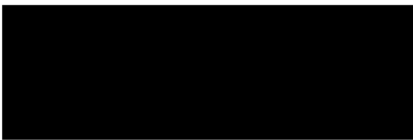
# Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Mittweida, den 13. September 2022



Jane Marie Wachter