

Titel der großen Studienarbeit

Projektarbeit

des Studienganges **Angewandte Informatik / Betriebliches Informationsmanagement**  
an der Dualen Hochschule Baden-Württemberg Mannheim

von

**Maximilian Ludwig, Kevin Wrona, Fabian Brandmüller**

10. Februar 2020

Bearbeitungszeitraum  
Betreuer der DHBW

**23.09.2019 - 20.04.2019**  
**Eckhard Kruse**

## Erklärung

Ich versichere hiermit, dass ich meine Projektarbeit mit dem Thema: "Titel der großen Studienarbeit" selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

---

Ort

Datum

Unterschrift

## Abstract

# Inhaltsverzeichnis

<b>I. Einleitung</b>	<b>I</b>
I.1. Aufgabenstellung . . . . .	III
<b>II. Stand der Technik</b>	<b>III</b>
II.1. Gesichtserkennung vs. Emotionserkennung . . . . .	III
II.1.1. Gesichtserkennung . . . . .	III
II.1.2. Emotionserkennung . . . . .	V
II.1.2.1. Emotionen . . . . .	V
II.1.2.2. Abgrenzung zur Gesichtserkennung . . . . .	VI
II.2. Emotionserkennung mithilfe von Deep Learning . . . . .	VI
II.2.1. OpenCV vs. Tensorflow . . . . .	VI
II.2.2. Supervised vs. Unsupervised Learning . . . . .	VI
II.3. Alternative Ansätze zur Umsetzung von Emotionserkennung . . . . .	VI
<b>III. Ergebnis</b>	<b>VIII</b>
III.1. Erstellung eines Modells . . . . .	VIII
III.2. Testen des Modells . . . . .	VIII
<b>IV. Diskussion</b>	<b>VIII</b>
IV.1. Reflexion der Ergebnisse . . . . .	VIII
IV.1.1. Alternativen . . . . .	VIII
IV.2. Reflexion Vorgehen . . . . .	VIII
IV.3. Reflexion der Literatur . . . . .	VIII
IV.4. Offene Implikationen . . . . .	IX

# Abbildungsverzeichnis

II.1. Phasen der Gesichtserkennung	
Quelle: <a href="https://alitarhini.files.wordpress.com/2010/12/untitled1.png">https://alitarhini.files.wordpress.com/2010/12/untitled1.png</a> . . . . .	IV

## Abkürzungen

# I. Einleitung

”Several researchers have stated that facial expression recognition appears to play one of the most important roles in human communication”<sup>1</sup> Dieses Zitat von Katherine B. Leeland gibt einen Einblick in die Relevanz der Emotionserkennung für den Menschen. Fragen zu dieser Thematik stellen sich allerdings nicht erst seit Beginn der Digitalisierung. Bereits Darwin fragte sich, ob von den Gesichtsausdrücken einer Person nicht auch der Emotionale Zustand abgeleitet werden kann.<sup>2</sup> Einen solchen Zustand von einem Mitmenschen abzulesen ist jedoch nicht leicht zu realisieren. Bereits durch kleine Änderungen in der Mimik werden verschiedene Emotionen ausgedrückt. Zum Beispiel indem eine Person die Lippen zusammen presst und die Augen zusammen kneift bei Wut, oder die Mundwinkel nach unten gezogen werden bei Trauer.<sup>3</sup> Durch derartige Ausdrücke können Emotionen wie Wut oder Trauer Ausdruck gewinnen. Emotionserkennungssoftware gibt es bereits und wird auch in der Wirtschaft eingesetzt. Die Anwendungsgebiete reichen dabei von Jobinterviews, in denen analysiert wird in wie weit die Bewerber für den jeweiligen Job geeignet sind,<sup>4</sup> bis hin zur Automobilindustrie. Dort wird mittels geeigneter Sensorik versucht die Emotion und somit der physiologische Zustand des Autofahrers zu analysieren.<sup>5</sup> Diese Daten legen den Grundbaustein für Warnsysteme, welche den Fahrer darauf hinweisen können, dass sein Zustand ungeeignet zum Betrieb eines Kraftfahrzeugs ist. Jedoch werden solche Einsatzszenarien auch durchaus kontrovers diskutiert. Auf Kritik stößt unter anderem dass die sogenannten ”Basisemotionen”, die verwendet werden um den KIs Emotionserkennung beizubringen, selbst umstritten sind.<sup>6</sup> Aber auch ethische Bedenken werden zunehmend geäußert, vor allem bezüglich der Anwendungsgebiete. Denn je nach Emotion die erkannt werden soll, liegt die Fehlerrate sehr hoch. So hat das Fraunhofer Institut, welches an Einsatzgebieten von Emotionserkennung in Fahrzeugen arbeitet, festgestellt, dass eine Emotionserkennung je nach Zielemotion eine Vorhersagekraft zwischen 6 und 95% haben kann.<sup>7</sup> Diese negativen Aspekte treffen jedoch nur teilwei-

---

<sup>1</sup>Vgl. Leeland, *Face Recognition: New Research*, S. 1.

<sup>2</sup>Vgl. Leeland, *Face Recognition: New Research*, S. 2.

<sup>3</sup>Vgl. Li, *Handbook of Face Recognition*, S. 249.

<sup>4</sup>Vgl. Schreiner, *Künstliche Intelligenz: Emotionserkennung will gelernt sein*.

<sup>5</sup>Vgl. Diederichs, *KI-gestützte Emotionserkennung im Fahrzeug aus physiologischen Daten*, Herausforderung.

<sup>6</sup>Vgl. Kuhn, *Emotionen sind schwer definierbar*.

<sup>7</sup>Vgl. Diederichs, *KI-gestützte Emotionserkennung im Fahrzeug aus physiologischen Daten*, Ergebnis.

se auf das hier behandelte Forschungsprojekt zu, wie im Folgenden dargelegt werden soll: Es werden in dieser Arbeit die Basisemotionen als Referenz herangezogen. Dies ist notwendig, da die Alternativen Verfahrensweisen zur Ermittlung der Emotion eines Individuums nicht praktikabel für den Use Case dieser Arbeit wären. Eine erweiterte Betrachtung der Alternativen findet sich in Kapitel II. Des weiteren ist diese Studienarbeit kein wissenschaftliches Werk über die psychologischen Emotionen die hinter verschiedenen Gesichtsausdrücken stehen, auch wenn sie dies indirekt thematisiert. In dieser Arbeit geht es darum die These zu verifizieren oder zu widerlegen, ob es möglich ist mit einer künstlichen Intelligenz zu erkennen, in wie weit ein Pokerface einer Person anhand dessen Gesichtsausdruck abgelesen werden kann. Diese Aufgabe hat also nicht den Anspruch eine Emotion korrekt vorherzusagen, sondern vorherzusagen wann keine Emotion vorliegt. Denn Ein Pokerface wird im Allgemeinen als ein emotionsloser Gesichtsausdruck definiert. Dies impliziert, dass eine Person mit Pokerface keine Emotion zu erkennen gibt. Daher könnte bei einer nicht messbaren, oder neutralen Emotion ein Pokerface vorliegen. Dies mittels KI zu testen und einen sogenannten "Pokerface Detektor" zu entwickeln ist folglich Ziel dieser Arbeit. Mit diesem Pokerface Detektor sind verschiedenste Einsatzmöglichkeiten in der Praxis denkbar. Diese haben eine gewisse Schnittmenge mit denen von "normaler" Emotionserkennungssoftware, jedoch gibt es auch einige weitere. Im Folgenden werden einige mögliche Szenarien expliziert:

- Polizeiverhöre

Es ist denkbar, dass der zu entwickelnde Detektor bei Polizeiverhören eingesetzt werden könnte. Hierdurch könnten Beamte die Verwendung eines Pokerfaces durch den Beschuldigten erkennen, welches auf eine Lüge hinweisen könnte. Neben der gebräuchlichen Verwendung eines Lügendetektors wäre der Einsatz der in dieser Arbeit erstellten Lösung eine kostengünstige Variante.

- Gerichtsverhandlungen

Das zweite Einsatzgebiet ist ähnlich zu dem ersten. Bei Gerichtsverhandlungen gelten die gleichen Voraussetzungen wie bei einem Verhör der Polizei. Zwar müssen die Vorgeklagten eine eidesstattliche Erklärung abgeben nur die Wahrheit zu sagen, jedoch ist zu bezweifeln ob dies auch der Fall ist. Nun soll nicht der Eindruck entstehen dass das hier gebaute Werkzeug ein Lügendetektor ist. Es ist ebenfalls nicht möglich, dass von einem Pokerface immer auf eine Lüge geschlossen werden kann. Jedoch ist ein Pokerface ein Zeichen dafür, dass sich diese Person ihren emotionalen Zustand nicht anmerken lassen möchte. Und dies wiederum deutet eher daraufhin dass die Person nicht die Wahrheit sagt oder nur teilweise. Abgesehen von Einsatzgebieten die zur Entlarvung von Lügen führen, kann auch ein klassisches Szenario verwendet werden:

- Pokerspiel

Es ist anzunehmen, dass der erste Begriff der mit dem Wort Pokerface in Verbindung gebracht wird, das Pokerspiel selber ist. Und auch in diesem kann ein Pokerface Detektor nützlich sein. So kann ein Mitspieler zum Beispiel mittels einer Kamera das Gesicht des Gegenübers scannen und analysieren ob ein Pokerface vorliegt oder nicht, und dementsprechend agieren. Nachdem verschiedene Anwendungsszenarien beleuchtet wurden wird im Folgenden die konkrete Forschungsfrage beleuchtet.

## **I.1. Aufgabenstellung**

Das Projekt selber wird an der DHBW in Mannheim durchgeführt und von Prof. Dr. Erckhard Kruse betreut. Wie eingangs erwähnt soll mittels künstlicher Intelligenz ein Pokerface erkannt werden. Zur Umsetzung dieser Aufgabe wird eine Bilderkennungssoftware angefertigt, welches ein übermitteltes Bild nach vorhandenen Emotionen analysiert. Sollte keine Emotion durch die Software erkannt werden, wird dem Anwender das Vorhandensein eines Pokerfaces zurück gegeben. Ein konkretes Einsatzgebiet nach Abschluss der Entwicklung ist nicht vorgesehen, da es sich um ein Forschungsprojekt handelt. Jedoch sind wie bereits in der Einleitung beschrieben einige verschiedene Einsatzmöglichkeiten denkbar, an die das Werkzeug leicht angepasst werden kann.

# **II. Stand der Technik**

Technologien todo:

- Poker erklären (Wenn wir Poker als testusecase benutzen)

In diesem Abschnitt soll der aktuelle Forschungs- und Entwicklungsstand im Bereich Emotionserkennung thematisiert werden. Jedoch muss dafür erst einmal eine Unterscheidung der Begrifflichkeiten Emotionserkennung und Gesichtserkennung erfolgen, da beide gebiete Überschneidungen haben, jedoch inhaltlich und von ihren Zielen verschieden sind.

## **II.1. Gesichtserkennung vs. Emotionserkennung**

### **II.1.1. Gesichtserkennung**

- use cases von Gesichtserkennung



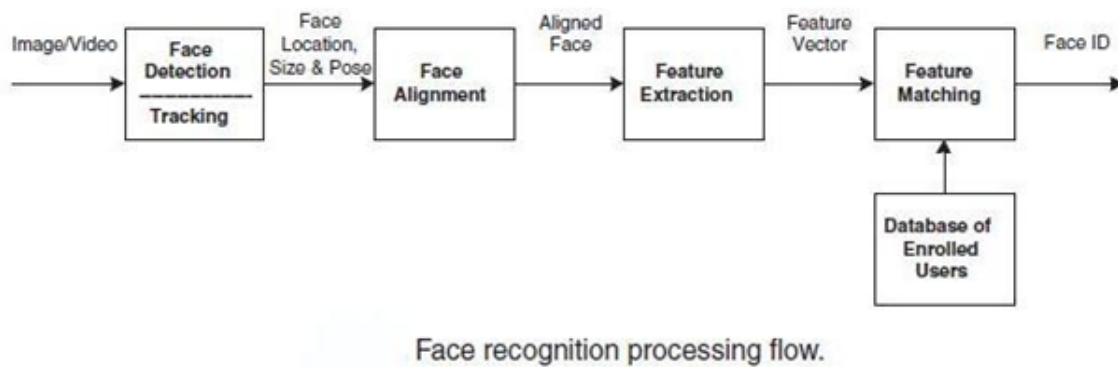


Abbildung II.1.: Phasen der Gesichtserkennung

Quelle: <https://alitarhini.files.wordpress.com/2010/12/untitled1.png>

- Grenzen der Gesichtserkennung
- Ausblick in der Gesichtserkennung

Gesichtserkennung ist eine Disziplin der Informatik in der es darum geht Gesichter wieder zu erkennen, und gegebenenfalls verschiedenen Personen zuzuordnen. Dabei lässt sich der Prozess der Gesichtserkennung in vier Phasen einteilen, Face "detection", "alignment", "feature extraction" und "matching".<sup>1</sup> Die Detection Phase ist dafür verantwortlich um zu erkennen ob Gesichter vorhanden sind in einem Bild, oder aber Video.<sup>2</sup> In der darauffolgenden Alignment Phase hingegen wird die Lokalisierung der Gesichter genauer, indem Gesichtskomponenten wie Augen, Augenbrauen, oder die Nase genauer lokalisiert werden. Dabei wird das Bild oder Video ebenfalls normalisiert, indem z.B. die Bildbeleuchtung angepasst wird.<sup>3</sup> In der Feature extraction hingegen werden die verschiedenen Gesichtskomponenten wie Augen, Nase, Mund, dem Bild oder Video entnommen. Dies ist ein wichtiger Schritt für weitere Prozesse wie Eye Tracking oder Face Tracking. Alternativ kann sogar eine bestimmte Person anhand der extrahierten Merkmale erkannt werden.<sup>4</sup> In der letzten Phase, dem Matching, geht es darum die gewonnen Daten mit den in der Datenbank vorhandenen Gesichtern abzugleichen. Wenn eine genügende Übereinstimmung gefunden wurde, wird ein Match mit einer Person ausgegeben.<sup>5</sup> Die Anwendungsgebiete von Software die Gesichtserkennung ermöglicht ist mannigfaltig. Sie reicht von Applikationen die ein Gerät wie ein Smartphone entsperren, wenn das Gesicht des Besitzers als Match ausgegeben wurde, bis hin zur Anwendung in Verbrechensbekämpfung. In jedem dieser Szenarien wird dabei der

<sup>1</sup>Vgl. Li, *Handbook of Face Recognition*, S. 2.

<sup>2</sup>Vgl. Li, *Handbook of Face Recognition*, S. 2.

<sup>3</sup>Vgl. Li, *Handbook of Face Recognition*, S. 2.

<sup>4</sup>Vgl. Saranya und Kumar, *Geometric shaped facial feature extraction for face recognition*, Abstract.

<sup>5</sup>Vgl. Li, *Handbook of Face Recognition*, S. 3.

oben beschriebene Ablauf durchgegangen, und abhängig vom zu liefernden Ergebnis eine Abschlussaktion vorgenommen.

## **II.1.2. Emotionserkennung**

In diesem Unterkapitel nun sollen Emotionen an sich thematisiert werden, da diese maßgeblich sind für das zu entwickelnde Tool. Eine Definition von Emotionserkennung ist per se nicht schwer zu geben. Prinzipiell beschäftigt sich Emotionserkennung mit der Analyse von Gesichtern und den Emotionen die diese Gesichter darstellen. Jedoch ist der Begriff der Emotionen nicht ganz so einfach zu definieren, wie im folgenden erläutert wird:

### **II.1.2.1. Emotionen**

- Def. von Emotionen

Grundsätzlich gibt es verschieden Ansätze Emotionen zu definieren und einzuteilen. Eine Variante ist dabei die eingangs erwähnte, nicht ganz unumstrittene Einteilung in Basisemotionen. Eine gängige Einteilung ist dabei die verschiedenen Emotionen in acht Bereiche einzuteilen. Diese Einteilung wurden 1984 von Plutchik postuliert und beinhaltet die Emotionskategorien Angst, Wut, Freude, Trauer, Akzeptanz, Ekel, Erwartung und Überraschung.<sup>6</sup> Jedoch ist dies nicht die einzige mögliche Einteilung. Als weiteres Beispiel teilte MacLean die Emotionen in lediglich sechs Kategorien ein, welche da wären: Verlangen, Wut, Angst, Niedergeschlagenheit, Freude und Zuneigung.<sup>7</sup> Wie sich bereits an den beiden Beispielen zeigt, geht die Meinungen der Forscher dabei stark auseinander, welche und wie viele Emotionen zu den sogenannten "Basis Emotionen" gehören. In dieser Arbeit werden die Emotionen in sechs Kategorien eingeteilt, in Wut, Trauer, Freude, Ekel, Überraschung und Neutral. Diese Einteilung entspricht an sich keiner gängigen Einteilung, jedoch wurde diese aus den folgenden Gründen gewählt: Die hier genannten Emotionen lassen sich gut anhand von Bildern erlernen, da diese zum Teil komplementär und somit eindeutig sind. Es ist aber auch einfacher Testdatensätze zu bekommen für ein freudiges Gesicht, oder ein überraschtes, als ein Gesicht mit dem emotionalen Ausdruck Akzeptanz. Des Weiteren wurde der Ausdruck "Neutral" hinzugefügt. Neutral repräsentiert ein emotionsloses Gesicht, und somit nach Definition einem Pokerface. Zudem sind die gewählten Emotionen häufig bei dem Test Usecase dieser Arbeit anzutreffen, dem Texas Holdem Poker.

---

<sup>6</sup>Vgl. Leeland, *Face Recognition: New Research*, S. 3.

<sup>7</sup>Vgl. Leeland, *Face Recognition: New Research*, S. 3.

#### II.1.2.2. Abgrenzung zur Gesichtserkennung

## II.2. Emotionserkennung mithilfe von Deep Learning

In dem nun folgenden Kapitel wird erörtert wie das Ziel der Arbeit, das Erkennen eines Pokerfaces, mittels eines neuronalen Netzes umgesetzt werden kann. Dabei wird weniger auf die generellen Eigenschaften von Neuronalen Netzen Bezug genommen, als auf die in dieser Arbeit spezifischen Aspekte. Diese sind unter anderem die verwendeten Komponenten, Applikationen und die Architektur im allgemeinen, sowie der Aufbau der Netzes selbst.

### II.2.1. OpenCV vs. Tensorflow

Libraries werden hier expliziert

### II.2.2. Supervised vs. Unsupervised Learning

Layer, Neuronen und Algorithmen explizieren

## II.3. Alternative Ansätze zur Umsetzung von Emotionserkennung

In diesem Abschnitt nun werden verschiedene alternative Ansätze dargestellt und expliziert, die dazu verwendet werden können um Emotionen zu erkennen. Dieses Unterkapitel beschäftigt sich mit alternativen Ansätzen zu den bereits explizierten Basisemotionen. Diese sind wie bereits erwähnt umstritten, was die Frage zulässt warum diese überhaupt verwendet werden sollten. Ein weiterer kreativer Ansatz zur Erkennung von Emotionen wäre die Analyse der Stimmlage. Dieser Ansatz beruft sich darauf, dass das Sprachzentrum eines Menschen einer der wichtigsten Aspekte der Kommunikation und somit auch der Preisgabe von Informationen über den emotionalen Zustand eines Individuums ist.<sup>8</sup> Dieser Ansatz ist jedoch nicht zielführend, da hier hauptsächlich die Stimme analysiert wird. Von einer Stimme kann nun auf eine Emotion geschlossen werden. Für den Usecase ist dieser Ansatz allerdings ungeeignet, aus folgenden Gründen: Es kann möglich sein eine Emotion anhand der Sprache zu erkennen. Das Äquivalent eines Pokerfaces wäre dementsprechend eine neutrale Stimmlage, welche keine Emotionen suggeriert. Nun kann aber keine Aussage getroffen werden aus welchen Gründen eine Person neutral spricht. Es könnte von einem Pokerface stammen, oder einer monotonen Sprechweise, oder einen gelangweilten Gemütszustand. Dies ist nicht eindeutig

---

<sup>8</sup>Vgl. Rao und Koolagudi, *Emotion Recognition using Speech Features*, Abstract.

identifizierbar. Gleiches könnte nicht über ein neutrales Gesicht gesagt werden, da dies gemeinhin als Pokerface bezeichnet wird. Ein weitere Ansatz wäre die Analyse der derzeit vernommenen Musik. Diese kann einem bestimmten Gemütszustand zugesprochen werden, welches auf eine aktuelle Emotion übertragbar ist.<sup>9</sup> Ziel dieses Forschungszweiges ist es daher die hinter Liedern oder Klängen stehenden Emotionen zu ermitteln und diese entsprechend zu kategorisieren. Dieser Ansatz erscheint zunächst durchaus interessant, hat jedoch genauso Nachteile wie die Analyse von Emotionen anhand von Bildern die Basisemotionen zeigen. Dieser liegt hier unter anderem in der Genauigkeit der Analysen. So z.B. lieferte ein Testprojekt an der Russischen HSE (Higher School of Economics) das Ergebnis von einer maximalen Genauigkeit von 71%.<sup>10</sup> In dem Versuchsaufbau wurden Spektrogramme von Klangfragmenten ausgewertet und versucht mittels Neuronalen Netzen eine Klassifikation der hinter dem Klang liegenden Emotion zu erreichen.<sup>11</sup> Der generelle Ansatz anhand von Musik die Emotion eines Individuums abzulesen ist zwar praktikabel und von dem Versuchsaufbau auch vergleichbar zu dem Ansatz bereits gelabelte Bilder zu verwenden. Jedoch lässt sich auf diese Weise aus zwei Gründen nicht die eigentliche Zielaufgabenstellung ableiten, das Erkennen eines Pokerfaces. Zum einen handelt es sich in dieser Arbeit um eine visuelle Problemstellung, in welcher das Erkennen des Gemütszustandes anhand des Gesichtsausdruckes erkannt werden soll, also einem vorhandenen bzw. nicht vorhandenen Pokerface. Zum anderen würde die Analyse von Musik einen Rückschluss auf den allgemeinen Gemütszustand des Betroffenen folgern und nicht eine kurzzeitige Stimmungsschwankung aufgrund beispielsweise eines schlechten Blattes, wie in diesem Usecase.

---

<sup>9</sup>Vgl. Yang und Chen, *Music Emotion Recognition*, S. 1.

<sup>10</sup>Vgl. Popova, Rassadin und Ponomarenko, *Emotion Recognition in Sound*, Abstract.

<sup>11</sup>Vgl. Popova, Rassadin und Ponomarenko, *Emotion Recognition in Sound*, Abstract.

## **III. Ergebnis**

### **III.1. Erstellung eines Modells**

### **III.2. Testen des Modells**

## **IV. Diskussion**

Das nunmehr letzte Kapitel soll sich mit der kurzen Zusammenfassung der Ergebnisse des letztes Teils und deren Bewertung widmen. Des Weiteren sollen die angewandten Methoden reflektiert werden, offene Fragen beantwortet und auch weitere Punkte aufgezeigt werden die verbessert oder noch implementiert werden können. Dazu soll zunächst die Ergebnisse kurz zusammengefasst werden.

### **IV.1. Reflexion der Ergebnisse**

#### **IV.1.1. Alternativen**

### **IV.2. Reflexion Vorgehen**

Mehr darauf eingehen dass das Kontrovers ist und auch die Basisemotionen kontrovers sind – aber keine andere Möglichkeit vorhanden

### **IV.3. Reflexion der Literatur**

Bezüglich der Literatur ergeben sich nun einige Schwierigkeiten. Dies liegt unter anderem daran, dass das generelle Thema der Gesichts und Emotionserkennung immer noch vor allem aus psychologischer Sicht in der Literatur behandelt wurde. Zwar gibt

es Fachbücher auch aus informationstechnischer Sicht, welche ebenfalls in dieser Arbeit verwendet wurden.

## **IV.4. Offene Implikationen**

# Literatur

## Literaturquellen

- [Lee08] Katherine B. Leeland. *Face Recongnition: New Research*. 1. Aufl. Nova Science Publishers INC, 2008. ISBN: 978-16045646625.
- [Li05] Anil K. Li Stan Z.and Jain. *Handbook of Face Recognition*. Springer Science und Business Media, 2005. ISBN: 978-0387405957.
- [PRP17] Anastasiya S. Popova, Alexandr G. Rassadin und Alexander A. Ponomarenko. *Emotion Recognition in Sound*. Cham, Deutschland: Springer Verlag, 2017. ISBN: 978-3319666037.
- [RK12] K. Sreenivasa Rao und Shashidhar G. Koolagudi. *Emotion Recognition using Speech Features*. 1. Auflage. Berlin, Deutschland: Springer Science und Business Media, 2012. ISBN: 978-1461451433.
- [SK17] R Benedict Saranya und J. Satheesh Kumar. *Geometric shaped facial feature extraction for face recognition*. Coimbatore - India: IEEE, 2017. ISBN: 978-1509037698.
- [YC11] Yi-Hsuan Yang und Homer H. Chen. *Music Emotion Recognition*. Boca Raton Florida: CRC Press, 2011. ISBN: 978-1466508927.

## Sonstige Quellen

- [Die20] Dr. Frederik Diederichs. *KI-gestützte Emotionserkennung im Fahrzeug aus physiologischen Daten*. <https://www.hci.iao.fraunhofer.de/de/Human-Centered-AI/feinfuehlige-technik/KI-gestuetzte-Emotionserkennung.html>. Einsichtnahme:24.01.2020. 2020.
- [Kuh19] Johannes Kuhn. *Emotionen sind schwer definierbar*. <https://www.sueddeutsche.de/digital/software-emotionen-simulation-ki-1.4377004-2>. Einsichtnahme:24.01.2020. 2019.
- [Sch19] Maximilian Schreiner. *Künstliche Intelligenz: Emotionserkennung will gelernt sein*. <https://mixed.de/kuenstliche-intelligenz-emotionserkennung-will-gelernt-sein/>. Einsichtnahme:24.01.2020. 2019.