

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

6. 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

I. Einleitung	I
I.1. Aufgabenstellung	III
II. Stand der Technik	III
II.1. Gesichtserkennung	III
II.2. Emotionserkennung	V
II.2.1. Emotionen	V
II.3. Ansätze zur Umsetzung von Emotionserkennung	VI
II.3.1. Emotionserkennung mittels Deep Learning	VII
II.3.1.1. Tensorflow	VII
III. Methode	VIII
IV. Ergebnis	VIII
V. Diskussion	IX
V.1. Reflexion der Ergebnisse	IX
V.1.1. Alternativen	IX
V.2. Reflexion Vorgehen	IX
V.3. Reflexion der Literatur	IX
V.4. Offene Implikationen	IX

Abbildungsverzeichnis

II.1. Phasen der Gesichtserkennung	
Quelle: https://alitarhini.files.wordpress.com/2010/12/untitled1.png .	IV

Abkürzungen

I. Einleitung

”Several researchers have stated that facial expression recognition appears to play one of the most important roles in human communication”¹ Dieses Zitat von Katherine B. Leeland gibt einen Einblick in die Relevanz der Emotionserkennung für den Menschen. Dabei ist diese Relevanz nicht erst in jüngerer Zeit entstanden. Bereits Darwin stellte die Frage, ob von den Gesichtsausdrücken einer Person nicht auch der emotionale Zustand abgeleitet werden kann.² Einen solchen Zustand von einem Mitmenschen abzulesen ist jedoch nicht trivial. Dies liegt auch daran, dass die Übergänge von verschiedenen Emotionen durch leichte Veränderungen der Mimik ausgedrückt werden. 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.³ Durch solche kleinen Änderungen kann ein sonst neutral wirkendes Gesicht zu einem wütenden oder traurigen werden. 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 zu dem jeweiligen Job passen⁴ bis hin zur Automobilindustrie. Dort wird mittels geeigneter Sensorik versucht die Emotion und somit der physiologische Zustand des Autofahrers zu analysieren.⁵ Mit diesen Daten können dann wiederum Funktionen eingebaut werden wie Warnsysteme, die den Fahrer daraufhinweisen, 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, selber umstritten sind.⁶ 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.⁷ Diese ne-

¹Vgl. Leeland, *Face Recognition: New Research*, S. 1.

²Vgl. Leeland, *Face Recognition: New Research*, S. 2.

³Vgl. Li, *Handbook of Face Recognition*, S. 249.

⁴Vgl. Schreiner, *Künstliche Intelligenz: Emotionserkennung will gelernt sein*.

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

⁶Vgl. Kuhn, *Emotionen sind schwer definierbar*.

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

gativen Aspekte treffen jedoch nur teilweise auf das hier behandelte Forschungsprojekt zu, wie im Folgenden dargelegt werden soll: Es werden in dieser Arbeit wie im folgenden Kapitel beschrieben werden wird 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. Auf der anderen Seite ist diese Studienarbeit kein wissenschaftliches Werk über die psychologischen Emotionen die hinter verschiedenen Gesichtsausdrücken stehen, auch wenn sie dies indirekt thematisiert. Es geht hier 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 eine 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 daher Ziel dieser Arbeit. Mit diesem Pokerface Detektor sind verschiedenste Einsatzmöglichkeiten in der Praxis denkbar. Diese haben eine gewisse Schnittmenge mit denen von "normaler" Emotionserkennungs software, jedoch gibt es auch einige weitere. Im Folgenden werden einige mögliche Szenarien expliziert:

- Polizeiverhöre

Es ist denkbar, dass ein Detektor wie der, der in dieser Arbeit entwickelt werden wird bei Polizeiverhören eingesetzt wird. Für die Beamten kann es nicht immer direkt ersichtlich sein, ob der verhörte mit einem Pokerface lügt, oder die Wahrheit sagt. Da die Möglichkeiten dies zu prüfen ebenfalls nicht zahlreich oder einfach verfügbar sind, wäre es eine Vereinfachung für die Beamten wenn eine einfache Webcam zusammen mit einem Computer reichen würde um einen Lügner zu entlarven.

- Gerichtsverhandlungen

Das zweite Einsatzgebiet ist ähnlich zu dem ersten. Bei Gerichtsverhandlungen gelten die gleichen Vorraussetzungen wie bei einem Verhöf der Polizei. Zwar müssen die hier vorgeladenen eine eidestattliche Erklärung abgeben nur die Wahrheit zu sagen, jedoch ist zu bezweifeln ob dies auch jeder so handhabt. 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 klassischeres 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 betreut von Prof. Dr. Erckhard Kruse. Wie eingangs erwähnt soll mittels künstlicher Intelligenz ein Pokerface erkannt werden. Dafür soll wiederum eine Bilderkennungssoftware angefertigt werden, die Bilder oder Videos als Eingabe von einer Webcam oder anderen Kamera bekommt, und das Bild analysiert. Je nachdem welche Emotionen gezeigt werden wird dann von der Software ein Rückschluss gezogen auf ein eventuell vorhandenes Pokerface. Ein konkretes Einsatzgebiet nach Abschluss der Entwicklung ist nicht vorgesehen, da es sich um ein Forschungsprojekt handelt. Jedoch sind wie bereits 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 thematisiert werden im Bereich Emotionserkennung. Jedoch muss dafür zuerst eine Unterscheidung der Begrifflichkeiten Emotionserkennung und Gesichtserkennung erfolgen, da beide Gebiete Überschneidungen haben, jedoch inhaltlich und von ihren Zielen verschieden sind.

II.1. Gesichtserkennung

- use cases con Gesichtserkennung

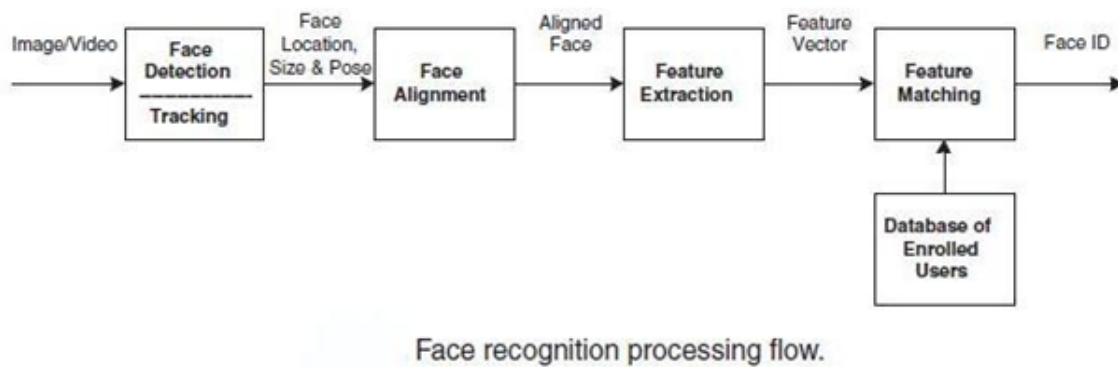


Abbildung II.1.: Phasen der Gesichtserkennung

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".¹ Die Detection Phase ist dafür verantwortlich um zu erkennen ob Gesichter vorhanden sind in einem Bild, oder aber Video.² 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.³ In der Feature extraction hingegen werden die verschiedenen Gesichtskomponenten wie Augen, Nase, Mund, dem Bild oder Video entnommen. Dies ist ein enorm wichtiger Schritt für weitere Prozesse wie Eye Tracking oder Face Tracking. Alternativ kann sogar eine bestimmte Person anhand der extrahierten Merkmale erkannt werden.⁴ In der letzten Phase, dem Matching, geht es darum die gewonnenen Daten mit den in der Datenbank vorhandenen Gesichtern abzugleichen. Wenn eine genügende Übereinstimmung gefunden wurde, wird ein Match mit einer Person ausgegeben.⁵ 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

¹Vgl. Li, *Handbook of Face Recognition*, S. 2.

²Vgl. Li, *Handbook of Face Recognition*, S. 2.

³Vgl. Li, *Handbook of Face Recognition*, S. 2.

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

⁵Vgl. Li, *Handbook of Face Recognition*, S. 3.

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

II.2. Emotionserkennung

In diesem Unterakptiel 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 wie der Name Verrät 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.2.1. Emotionen

- Def. von Emotionen

Grundsätzlich gibt es viele verschieden Ansätze Emotionen zu definieren, bzw. Emotionen 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.⁶ Jedoch ist dies bei weitem nicht die einzige mögliche Einteilung. Als weiteres Beispiel teilte MacLean die Emotionen in lediglich sechs Kategorien ein, Verlangen, Wut, Angst, Niedergeschlagenheit, Freude und Zuneigung.⁷ Wie sich bereits an den beiden Beispielen zeigt, geht die Meinungen der Forscher dabei sehr 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 ein Pokerface. Außerdem sind dieses Arten der Gesichtsausdrücke auch nicht selten anzutreffen bei dem Test Usecase dieser Arbeit, dem Texas Holdem Poker.

⁶Vgl. Leeland, *Face Recognition: New Research*, S. 3.

⁷Vgl. Leeland, *Face Recognition: New Research*, S. 3.

II.3. Emotionserkennung mittels 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, die Applikations Architektur im allgemeinen, und auch der Aufbau der Netzes selbst.

II.3.1. OpenCV vs Tensorflow

Libraries werden hier expliziert

II.3.2. Aufbau des neuronalen Netzes

Layer, Neuronen und Algorithmen explizieren

II.4. 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. Dabei beschäftigt sich dieses Unterkapitel mit alternativen Ansätzen um die bereits definierten Basisemotionen zu erkennen. Es gibt einige kreative alternative Ansätze zur Emotionserkennung. Zum Beispiel dass Emotionen anhand der Sprache eines Menschen ermittelt werden. 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.⁸ 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 use case ist dieser Ansatz allerfings ungeeignet, aus folgenden Gründen:

Es kann möglich sein eine Emotion anhand der Sprache zu erkennen. Das Äquivalent eines Pokerfaces wäre wiederum eine neutrale Stimmhaltung, da es Ziel ist möglichst emotionslos zu wirken bei einem Pokerface. Nun kann aber keine Aussage getroffen werden warum eine Person neutral spricht. Es könnte von einem Pokerface stammen, oder einer monotonen Sprechweise, oder einen gelangweilten Gemütszustand. Dies ist nicht eindeutig identifizierbar. Gleiches könnte nicht über ein neutrales Gesicht gesagt werden, da dies gemeinhin als Pokerface bezeichnet wird. Ein weitere Ansatz ist, dass

⁸Vgl. Rao und Koolagudi, *Emotion Recognition using Speech Features*, Abstract.

Musik eine gewisse Emotion zugeschrieben wird. hierbei geht es darum, dass dem Hörer bei gewissen Liedern eine gewisse Emotion zu teil wird.⁹ Ziel dieses Forschungszweiges ist es daher die hinter Liedern oder Klängen stehenden Emotionen zu ermitteln, um sie z.B. 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%.¹⁰ In dem Versuchsaufbau wurden Spektrogramme von Klangfragmenten ausgewertet und versucht mittels Neuronalen Netzen eine Klassifikation der hinter dem Klang liegenden Emotion zu erreichen.¹¹ Der generelle Ansatz anhand von Musik die Emotion eines Individuums abzulesen ist zwar praktikabel und vom Versuchsaufbau auch relativ ähnlich zu dem Ansatz bereits gelabelte Bilder zu verwenden. Jedoch lässt sich auf diese Weise nicht die eigentliche Zielaufgabenstellung ableiten, das Erkennen eines Pokerfaces, da dies eine rein visuelle Problemstellung ist.

⁹Vgl. Yang und Chen, *Music Emotion Recognition*, S. 1.

¹⁰Vgl. Popova, Rassadin und Ponomarenko, *Emotion Recognition in Sound*, Abstract.

¹¹Vgl. Popova, Rassadin und Ponomarenko, *Emotion Recognition in Sound*, Abstract.

III. Methode

Als vorgehen könnten wir sagen dass wir die neuen Algorithmen oder Daten beim Poker teste, das wäre Lustig und ein Use case der Sinn macht. Mehr als sich stupide vor die Kamera zu setzen.

Diese Arbeit soll methodisch mit der MoSCoW Priorisierung bearbeitet werden. Diese Art der Priorisierung teilt die zu bearbeitenden Anforderungen in vier Kategorien ein:¹

- Must - Core Anforderungen die unbedingt umgesetzt werden müssen
- Should - Anforderungen die ebenfalls umgesetzt werden müssen, jedoch um Nachhinein noch durch Change Request verändert werden können.
- Could - Anforderungen die nach den Must und Should Anforderungen umgesetzt werden sollen, sofern noch Ressourcen und Zeit vorhanden sind um diese zu bearbeiten
- Won't - Anforderungen die nicht in diesem Projekt bzw. Release erfolgen, jedoch in einer zukünftigen Version bearbeitet werden sollen.
- Must
 - placeholder
- Should
 - placeholder
- Could
 - placeholder
- Won't
 - placeholder

Dabei sollen die einzelnen Anforderungen entsprechend ihrer Priorität abgearbeitet werden. So kann am Ende der Erfolg der Arbeit deutlich besser eingeordnet werden

¹vgl. Roger, *Handbuch Projektmanagement: Agil – Klassisch – Hybrid*, S. 90.

IV. Ergebnis

V. Diskussion

Das nunmehr letzte Kapitel soll sich mit der kurzen Zusammenfassung der Ergebnisse des letztens 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.

V.1. Reflexion der Ergebnisse

V.1.1. Alternativen

V.2. Reflexion Vorgehen

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

V.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.

V.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.
- [Rog18] Schneider Patrickand Witschi Ursand Wüst Roger. *Handbuch Projektmanagement: Agil – Klassisch – Hybrid*. Auflage 4. Sebastopol: SSpringer Gabler”, 2018. ISBN: 978-3662578773.
- [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/sn-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.