

**BACHELORARBEIT**

# **Digitales Fruchtlexikon: Augmented Reality und KI-gestützte Objekterkennung als interaktiver Ansatz zur Wissensvermittlung im Alltag**

---

vorgelegt am 23. September 2024  
Michelle Lamowski, 2577377

Erstprüferin: Prof. Dr. Eike Langbehn  
Zweitprüfer: Prof. Dr.-Ing. Sabine Schumann

---

**HOCHSCHULE FÜR ANGEWANDTE  
WISSENSCHAFTEN HAMBURG**  
Department Medientechnik  
Finkenau 35  
22081 Hamburg

# Zusammenfassung

Die vorliegende Bachelorarbeit untersucht den Einsatz von Augmented Reality (AR) und Künstliche Intelligenz (KI)-gestützter Objekterkennung zur interaktiven Wissensvermittlung im Alltag. Während AR bereits in Spielen erfolgreich angewendet wird, bleibt ihr Potenzial für andere Bereiche, wie die Wissensvermittlung, weitgehend ungenutzt, obwohl sie aufgrund zahlreicher pädagogischer Vorteile als eine wichtige Technologie für die Bildung gilt. Zur Analyse des Potenzials von AR in diesem Kontext werden zunächst bestehende Arbeiten herangezogen, die sich auf die Nutzung von AR im schulischen Bildungswesen konzentrieren. Um die Anwendungsmöglichkeiten im Alltag zu evaluieren, wird eine eigene AR-basierte App entwickelt, die mithilfe eines maschinellen Lernmodells Früchte erkennt und in Echtzeit Nährstoffinformationen über der Kamera eines mobilen Geräts anzeigt. Die App bietet den Nutzern zudem die Möglichkeit, die identifizierten Früchte zu speichern und zu sammeln, wodurch spielerisches Lernen gefördert wird. In einer abschließenden Nutzerstudie wird die Anwendung im Vergleich zu traditionellen Lernmethoden hinsichtlich Effizienz und Benutzerfreundlichkeit getestet. Die Ergebnisse zeigen, dass AR-basierte Lernmethoden das Verständnis und die Motivation der Lernenden signifikant steigern können.

## Abstract

This bachelor's thesis examines the use of augmented reality and AI-supported object recognition for interactive knowledge dissemination in everyday life. While AR has already been successfully applied in gaming, its potential for other areas, such as knowledge transfer, remains largely untapped, despite being considered a key technology for education due to its numerous educational advantages. To assess AR's potential in this context, existing studies focusing on its use in school education are analyzed. In order to evaluate its applications in daily life, a custom AR-based app is developed that uses a machine learning model to recognize fruits and display real-time nutritional information via a mobile device's camera. Additionally, the app allows users to save and collect identified fruits, promoting playful learning. A final user study evaluates the app's efficiency and usability compared to traditional learning methods. The results indicate that AR-based learning methods can significantly enhance both learners' understanding and motivation.

# Inhaltsverzeichnis

|   |           |
|---|-----------|
| <b>Abbildungsverzeichnis</b>  | <b>IV</b> |
| <b>Tabellenverzeichnis</b>  | <b>V</b>  |
| <b>1 Einleitung</b>   | <b>1</b>  |
| 1.1 Vorwort . . . . .   | 1         |
| 1.2 Motivation . . . . .  | 1         |
| 1.3 Zielsetzung und Hypothesen . . . . .  | 2         |
| 1.4 Aufbau der Arbeit . . . . .   | 3         |
| <b>2 Grundlagen</b>   | <b>4</b>  |
| 2.1 Lerntheorien . . . . .  | 4         |
| 2.1.1 Theorie des situierten Lernens . . . . .  | 4         |
| 2.1.2 Theorie des konstruktivistischen Lernens . . . . .  | 5         |
| 2.1.3 Weitere Lernmethoden . . . . .  | 6         |
| 2.2 Webentwicklung . . . . .  | 6         |
| 2.2.1 Einführung in die Webentwicklung . . . . .  | 6         |
| 2.2.2 Asynchrone Programmierung . . . . .   | 7         |
| 2.2.3 Progressive Web App . . . . .   | 8         |
| 2.3 Mixed Reality . . . . .   | 8         |
| 2.3.1 Virtual Reality . . . . .   | 8         |
| 2.3.2 Augmented Reality . . . . .   | 9         |
| 2.3.3 Taxonomien von Augmented Reality . . . . .  | 9         |
| 2.3.4 3D-Nutzerschnittstellen . . . . .   | 10        |
| 2.4 Künstliche Intelligenz . . . . .  | 11        |
| 2.4.1 Maschinelles Lernen . . . . .   | 11        |
| 2.4.2 Deep Learning . . . . .   | 13        |
| 2.4.3 Transfer Learning . . . . .   | 14        |
| 2.4.4 Anwendungsbeispiele . . . . .   | 15        |
| 2.5 Vergleichbare Arbeiten . . . . .  | 15        |
| 2.5.1 Augmented Reality Teaching and Learning . . . . .   | 15        |
| 2.5.2 Current status, opportunities and challenges of Augmented Reality<br>in education . . . . . | 17        |

|          |   |           |
|----------|---|-----------|
| 2.5.3    | EcoMOBILE: Integrating Augmented Reality and probeware with environmental education field trips . . . . . | 18        |
| <b>3</b> | <b>Konzept</b>  | <b>20</b> |
| 3.1      | Kernkonzept . . . . .   | 20        |
| 3.2      | Designkonzept . . . . .   | 21        |
| 3.3      | Entwicklungsumgebungen . . . . .  | 25        |
| 3.3.1    | IntelliJ IDEA Ultimate . . . . .  | 25        |
| 3.3.2    | Google Colab . . . . .  | 26        |
| 3.4      | Frameworks . . . . .  | 26        |
| 3.4.1    | Vue . . . . .   | 27        |
| 3.4.2    | Ionic . . . . .   | 28        |
| 3.4.3    | Tailwind CSS . . . . .  | 29        |
| 3.4.4    | Media Pipe: Model Maker . . . . .   | 29        |
| 3.4.5    | Media Pipe: Object Detector . . . . .   | 31        |
| 3.5      | Entwicklungsplattformen . . . . .   | 32        |
| 3.5.1    | Datenannotation: Label Studio . . . . .   | 32        |
| 3.5.2    | Deployment: Firebase . . . . .  | 32        |
| <b>4</b> | <b>Umsetzung</b>  | <b>34</b> |
| 4.1      | Entwicklung des Modells . . . . .   | 34        |
| 4.1.1    | Datensammlung . . . . .   | 34        |
| 4.1.2    | Datenaufbereitung . . . . .   | 36        |
| 4.1.3    | Modelltraining . . . . .  | 36        |
| 4.1.4    | Modellvalidierung . . . . .   | 38        |
| 4.1.5    | Modellbereitstellung . . . . .  | 39        |
| 4.2      | Entwicklung der Progressive Web App . . . . .   | 39        |
| 4.2.1    | Benutzeroberfläche . . . . .  | 39        |
| 4.2.2    | Composables . . . . .   | 45        |
| 4.2.3    | Deployment . . . . .  | 47        |
| <b>5</b> | <b>Evaluation</b>   | <b>48</b> |
| 5.1      | Durchführung . . . . .  | 48        |
| 5.2      | Ergebnisse . . . . .  | 49        |
| 5.3      | Auswertung . . . . .  | 51        |
| 5.3.1    | Vergleich der Mittelwerte . . . . .   | 51        |
| 5.3.2    | Inhaltsanalyse des mündlichen Feedbacks . . . . .   | 56        |
| 5.3.3    | Inhaltsanalyse der Beobachtungen . . . . .  | 56        |
| 5.4      | Limitationen . . . . .  | 57        |
| <b>6</b> | <b>Fazit</b>  | <b>58</b> |

**Literatur**

**60**

**Anhang**

**67**

# Abbildungsverzeichnis

|     |  |    |
|-----|--|----|
| 2.1 | Abgrenzung von Augmented Reality und Augmented Virtuality . . . . .  | 10 |
| 2.2 | Teilbereiche der Künstlichen Intelligenz . . . . .   | 11 |
| 2.3 | Abbildung eines Deep Learning-Modells . . . . .  | 13 |
| 2.4 | Unterschied zwischen traditionellen maschinellen Lernmethoden und modernen Transferlern-Strategien . . . . . | 14 |
| 3.1 | Designkonzept der CameraPage . . . . .   | 21 |
| 3.2 | Designkonzept der GalleryPage . . . . .  | 22 |
| 3.3 | Designkonzept der FullScreenImagePage . . . . .  | 23 |
| 3.4 | Designkonzept der KnowHowPage . . . . .  | 24 |
| 3.5 | Designkonzept der SettingPage . . . . .  | 25 |
| 3.6 | Ersetzen der Klassifikationsebene durch spezialisierte Ebenen mithilfe des Model Maker-Tools . . . . .       | 30 |
| 4.1 | Beispiele für problematische Datensätze . . . . .  | 35 |
| 4.2 | Beispiele für eigene Datensätze . . . . .  | 36 |
| 4.3 | Annotation eines Apfels in Label Studio . . . . .  | 37 |
| 4.4 | Hauptbereiche der Anwendung . . . . .  | 40 |
| 4.5 | Funktionalitäten der CameraPage . . . . .  | 41 |
| 4.6 | Abbildung einer farblich markierten Bounding Box . . . . .   | 44 |
| 4.7 | Abbildung der Liste von Nährstoffen einer Zitrone . . . . .  | 45 |
| 4.8 | Abbildung der gespeicherten Fruchtobjekte auf der GalleryPage . . . . .                                      | 46 |

# Tabellenverzeichnis

|      |   |    |
|------|---|----|
| 5.1  | Reihenfolge der Methoden für jede Gruppe . . . . .  | 49 |
| 5.2  | Anzahl der richtigen Antworten im ersten und zweiten Durchlauf aller Gruppen                                | 50 |
| 5.3  | Anzahl der richtigen Antworten im ersten und zweiten Durchlauf aller Gruppen sortiert nach Frucht . . . . . | 51 |
| 5.4  | Durchschnittliche Anzahl der richtigen Antworten aller Gruppen . . . . .                                    | 52 |
| 5.5  | Durchschnittliche Anzahl der richtigen Antworten für den Apfel . . . . .                                    | 52 |
| 5.6  | Durchschnittliche Anzahl der richtigen Antworten für die Orange . . . . .                                   | 53 |
| 5.7  | Durchschnittliche Anzahl der richtigen Antworten für die Birne . . . . .                                    | 53 |
| 5.8  | Durchschnittliche Anzahl der richtigen Antworten für die Pflaume . . . . .                                  | 53 |
| 5.9  | Durchschnittliche Anzahl der richtigen Antworten sortiert nach Methode . .                                  | 54 |
| 5.10 | Ergebnisse der richtigen Antworten sortiert nach angewandeter Reihenfolge für die Gruppen A und B . . . . . | 55 |
| 5.11 | Ergebnisse der richtigen Antworten sortiert nach angewandeter Reihenfolge für die Gruppen C und D . . . . . | 55 |

# 1 Einleitung

## 1.1 Vorwort

In dieser Arbeit wird aus Gründen der Lesbarkeit auf die explizite Nennung mehrerer Geschlechtsformen verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

## 1.2 Motivation

In den letzten Jahren hat sich das Smartphone immer weiterentwickelt und an nahezu unverzichtbaren Features dazu gewonnen. Dazu gehören eine eingebaute Kamera, GPS für die Navigation und eine Internetverbindung (Chou & ChanLin, 2012). Dies führte dazu, dass sich auch das Aufkommen der Augmented Reality (AR)-Technologie begünstigt hat (Zsila et al., 2017).

Insbesondere Spiele, die AR einsetzen, sind weltweit populär geworden (Zsila et al., 2017). Dies beweist Pokémon Go, welches im Jahr 2016 als beliebtestes Spiel in der Geschichte der Smartphone-Spiele galt (Zsila et al., 2017). Mehr als 100 Millionen Nutzer aus 30 Ländern haben Pokémon Go innerhalb weniger Wochen heruntergeladen und spielen durchschnittlich 26 Minuten am Tag (Smith, 2024).

Allerdings besitzt AR weit mehr Potenzial als lediglich im Bereich der Unterhaltung. Ein Beispiel für einen sinnvollen Anwendungsbereich ist die Wissensvermittlung im Alltag. Noch widmen sich jedoch nur wenige Forschungs- und Entwicklungsteams der Anwendung von AR in diesem Bereich, obwohl AR aufgrund der pädagogischen Vorteile als eine der wichtigsten aufkommenden Technologien für die Bildung angesehen wird (Dunleavy & Dede, 2014; L. Johnson et al., 2010; L. F. Johnson et al., 2010; Martin et al., 2011).

Die immersive Natur von AR ermöglicht es, die physische Umgebung durch digitale Informationen zu erweitern und diese interaktiv darzustellen. Diese Kombination schafft eine innovative Form des Lernens, die konventionelle Ansätze übersteigt (Facer et al., 2014; Klopfer, 2008; Klopfer & Squire, 2008; Schmalstieg & Wagner, 2007; K. D. Squire et al., 2007).

Darüber hinaus hat die Anwendung von AR das Potenzial, die Motivation der Lernenden erheblich zu steigern. Eine gezielte Verbindung von Freude, Herausforderung und Neugier trägt dazu bei, das Engagement der Nutzer zu erhöhen. Dies ermöglicht es den Nutzern, die Anwendung eigenständig zu erkunden, was das Lernen intuitiv und ansprechend gestaltet (Sotiriou & Bogner, 2008).

Insbesondere die Nutzung mobiler AR-Anwendungen bietet den Vorteil, dass sie flexibel und ortsunabhängig einsetzbar sind. Da moderne Smartphones und Tablets mit AR-fähigen Technologien ausgestattet sind, ist es möglich, Lerninhalte überall und jederzeit zu vermitteln, wodurch die Wissensvermittlung nicht länger an feste Orte oder Zeiten gebunden ist (Klopfer & Sheldon, 2010). Diese Flexibilität erlaubt es, Lernprozesse stärker in den Alltag zu integrieren und Lerninhalte in realen Situationen anzuwenden. So können beispielsweise AR-Anwendungen an unkonventionellen Orten, wie in Supermärkten, genutzt werden.

Um eine authentische Umgebung zu gewährleisten spielt die Verarbeitung von Echtzeitdaten der realen Welt eine zentrale Rolle (Rosenbaum et al., 2007). Durch fortschrittliche Methoden der Objekterkennung können AR-Anwendungen relevante Informationen in Echtzeit bereitstellen, ohne dass der Nutzer manuelle Eingaben vornehmen muss. Dies trägt entscheidend zu einer benutzerfreundlichen Interaktion bei, da die Bedienung vereinfacht und der Aufwand für den Nutzer reduziert wird.

AR besitzt somit das Potenzial, das Lernen in alltäglichen Kontexten zu erleichtern, indem es kontextbezogene Informationen zugänglicher macht und den Wissenserwerb in den Alltag integriert.

### **1.3 Zielsetzung und Hypothesen**

Um den Einsatz von AR im Wissensbereich und die Wirksamkeit dieser Technologie für die Wissensaufnahme zu untersuchen, wird eine AR Anwendung entwickelt: Ein digitales Fruchtlexikon, welches Informationen über Früchte mithilfe von AR darstellt. Die Anwendung nutzt die Kamera eines Mobilgeräts um Nährstoffe anzuzeigen, sobald die Kamera auf eine Frucht gerichtet wird. Für die Erkennung der anvisierten Früchte wird ein Objekterkennungsmodell trainiert. Im Rahmen der Evaluation wird untersucht, wie Nutzer die AR-Technologie im Vergleich zu herkömmlichen Lernmethoden wahrnehmen und welche Vor- oder Nachteile sich aus der Anwendung dieser Technologie in der Wissensvermittlung ergeben.

Für diesen Kontext werden folgende Thesen formuliert:

- H1: AR-basierte Lernanwendungen verbessern das Verständnis und Erinnerungsvermögen von Lerninhalten.
- H2: AR-Anwendungen erhöhen die Motivation und das Engagement der Lernenden.

- H3: Mobile AR-Anwendungen sind in alltäglichen Kontexten effektiv anwendbar.

## **1.4 Aufbau der Arbeit**

Zu Beginn werden im Kapitel 2 wichtige Grundlagen zu den Themen Lerntheorien, Webentwicklung, Mixed Reality (MR) und Künstliche Intelligenz (KI) erläutert. Außerdem wird ein Überblick über vergleichbare Arbeiten gegeben, um grundlegende Ansätze und Methoden zu identifizieren und deren Relevanz für die eigene Untersuchung zu verdeutlichen. Im folgenden Kapitel 3 wird das Konzept beschrieben, welches die Basis für die Entwicklung der Anwendung bilden soll. Des Weiteren wird der Ansatz des Modelltrainings in diesem Kapitel erklärt. Das Kapitel 4 behandelt die Umsetzung des vorgestellten Konzeptes und die Implementierung der Anwendung. In Kapitel 5 wird die Nutzerstudie aufgestellt, umgesetzt und evaluiert. Abschließend wird im Kapitel 6 ein Fazit gezogen.

## 2 Grundlagen

Zur Schaffung einer fundierten theoretischen Basis für diese Arbeit werden zunächst zentrale Konzepte eingehend erläutert. Dazu gehören insbesondere ausgewählte Lerntheorien, die die wissenschaftliche Grundlage für AR-basierte Lernplattformen darstellen. Des Weiteren werden in diesem Kapitel grundlegende Prinzipien der Webentwicklung vorgestellt, auf die in Kapitel 3 weiter eingegangen wird. Die Basis für das zu erforschende Prinzip der Wissensvermittlung bildet jedoch die AR. Zu diesem Zweck werden zunächst die Begrifflichkeiten aus dem Bereich MR im Abschnitt 2.3 erläutert. Ein unverzichtbarer Bestandteil der Arbeit ist das Erkennen von Objekten, das mithilfe von Objekterkennungstechnologien realisiert wird. Daher werden abschließend relevante Themengebiete der KI im Abschnitt 2.4 erklärt. Zuletzt werden vergleichbare Arbeiten herangezogen, um die Kernkonzepte AR-basierter Lernplattformen zu analysieren und einzuordnen.

### 2.1 Lerntheorien

Es gibt eine Vielzahl von Lerntheorien, die unterschiedliche Ansätze und Perspektiven zum Lernen bieten. Besonders relevant für diese Arbeit sind die Theorie des situierten Lernens und die Theorie des konstruktivistischen Lernens. Zusätzlich werden weitere Lernmethoden vorgestellt, die für diese Arbeit bedeutend sind.

#### 2.1.1 Theorie des situierten Lernens

Laut Brown et al. (1989) basiert die Theorie des situierten Lernens auf der Annahme, dass alles Lernen in einem bestimmten Kontext stattfindet. Die Qualität des Lernens ist daher ein Ergebnis der Interaktionen zwischen den Menschen, Orten, Objekten, Prozessen und der Kultur innerhalb und in Bezug auf den jeweiligen Kontext. Lernen wird in diesen Kontexten als ein gemeinschaftlicher und partizipativer Prozess betrachtet. Dabei verändern sich die Lernenden durch ihre Handlungen und Beziehungen zur Welt (Driscoll, 1994). Das Niveau des Lernens wird maßgeblich von der Qualität der sozialen Interaktionen im Lernkontext beeinflusst. Diese Theorie wird durch die Arbeiten von Bandura (1977) und Vygotsky (1978) gestützt.

Für die Theorie des situierten Lernens ist der Transfer von besonderer Bedeutung, da dadurch gezeigt wird, wie Wissen und Fähigkeiten aus einem spezifischen Kontext auf neue, verwandte Kontexte angewendet werden können (Dunleavy & Dede, 2014). Dabei beschreibt Transfer die Fähigkeit, Wissen aus einer Situation erfolgreich auf eine andere anzuwenden. Dies wird sichtbar, wenn die Anleitung zu einer Lernaufgabe zu einer besseren Leistung in einer Transferaufgabe führt, idealerweise unter realen Bedingungen (Mestre, 2002).

Bei der Bewertung des Lernens als Vorbereitung auf zukünftiges Lernen fokussieren sich Forscher auf den Transfer, indem sie erweiterte Leistungen messen. Hierbei erwerben die Lernenden Kompetenzen in einer vielfältigen Lernumgebung. Dieses Wissen wird anschließend von den Lernenden zur Problemlösung in realen Kontexten verwendet (Dunleavy & Dede, 2014). Im konventionellen Unterricht verlangt die Vorbereitung auf zukünftiges Lernen einen Ferntransfer. Dabei wird Wissen aus einer Situation auf einen völlig anderen Kontext angewendet, dessen zugrunde liegende Semantik zwar mit dem Lernstoff verwandt, aber dennoch unterschiedlich ist (Perkins & Salomon, 1992).

### **2.1.2 Theorie des konstruktivistischen Lernens**

Die Theorie des konstruktivistischen Lernens, wie sie von Bruner (1974) und Vygotsky (1978) vertreten wird, beschreibt, dass Menschen neues Wissen und Verständnis auf Basis ihrer vorhandenen Kenntnisse und Überzeugungen entwickeln. Dieser Lernprozess wird durch ihren Entwicklungsstand, frühere Erfahrungen und ihren soziokulturellen Kontext geprägt.

Wissen ist eng mit dem Umfeld verknüpft, in dem es angewendet wird. Lernen umfasst daher die Bewältigung authentischer Aufgaben innerhalb sinnvoller und realistischer Situationen (Lave & Wenger, 1991). Lernende entwickeln persönliche Interpretationen der Realität, die auf ihren Erfahrungen und Interaktionen mit anderen basieren. Dadurch schaffen sie neuartige, situationsspezifische Verständnisse (Dunleavy & Dede, 2014).

Bruner (1974), Cunningham (1992), Driscoll (1994), Piaget (1970) und Vygotsky (1978) fassen auf Basis der konstruktivistischen Lerntheorie fünf Bedingungen zusammen, die das Lernen am effektivsten fördern. Dazu gehört die Einbettung des Lernens in relevante Umgebungen, die Einbeziehung von sozialem Austausch in die Lernerfahrung, sowie die Bereitstellung multipler Perspektiven und Darstellungsformen. Zudem wird die Bedeutung selbstgesteuerten und aktiven Lernens hervorgehoben, ebenso wie die Unterstützung und Erleichterung metakognitiver Strategien innerhalb der Lernumgebung.

### **2.1.3 Weitere Lernmethoden**

Weitere relevante Lernmethoden umfassen das spielbasierte Lernen, bei dem Spielelemente genutzt werden, um den Lernprozess spielerisch und motivierend zu gestalten (Rosenbaum et al., 2007; K. Squire & Klopfer, 2007; K. D. Squire & Jan, 2007). Ortsbasiertes Lernen integriert den physischen Standort als Lernkontext, um Wissen in realen Umgebungen zu vermitteln (Klopfer, 2008; Mathews, 2010). Eine weitere Methode ist die Anwendung partizipativer Simulationen, bei denen die Lernenden aktiv an simulierten Szenarien teilnehmen, um komplexe Systeme besser zu verstehen (Klopfer & Sheldon, 2010; Rosenbaum et al., 2007; K. Squire & Klopfer, 2007). Ebenfalls wichtig für diese Arbeit ist das Problembasierte Lernen, das die selbstständige Wissensaneignung durch die Lösung realer Probleme fördert (Liu et al., 2009; K. Squire & Klopfer, 2007). Daneben ermöglichen es Rollenspiele verschiedene Perspektiven und soziale Interaktionen zu erleben (Rosenbaum et al., 2007). Studio-basierte Pädagogik setzt auf praxisorientiertes und kollaboratives Arbeiten in Projekten (Mathews, 2010). Schließlich teilt die Jigsaw-Methode Lerninhalte in Teilbereiche auf, die von Gruppen bearbeitet und anschließend zusammengeführt werden (Dunleavy et al., 2009).

## **2.2 Webentwicklung**

Die in dieser Arbeit entwickelte Anwendung basiert auf Webtechnologien. Um ein fundiertes Verständnis dieser Technologien zu ermöglichen, wird zunächst eine Einführung in die Webentwicklung gegeben. Im Anschluss wird auf die asynchrone Programmierung eingegangen, da diese im späteren Verlauf der Arbeit von Bedeutung ist. Da die Anwendung als Progressive Web App (PWA) veröffentlicht wird, werden die spezifischen Anforderungen und Vorteile dieser Technologie abschließend erläutert.

### **2.2.1 Einführung in die Webentwicklung**

Webentwicklung umfasst die Erstellung und Gestaltung von Webseiten und Webanwendungen. Dabei bilden die drei Haupttechnologien Hypertext Markup Language (HTML), Cascading Style Sheets (CSS) und JavaScript das Fundament moderner Webentwicklung (MDN Web Docs, 2024c).

HTML5 ist eine Auszeichnungssprache, die die grundlegende Struktur einer Webseite definiert. Webseiten bestehen in der Regel aus einer Vielzahl von HTML-Elementen, die zusammen den Inhalt der Seite bilden. Diese Elemente werden hierarchisch organisiert und ermöglichen die Erstellung von komplexen und dynamischen Webinhalten (MDN Web Docs, 2024d). Die hierarchische Struktur wird durch das Document Object Model (DOM) abgebildet, das

eine Schnittstelle zur Manipulation der Seiteninhalte bietet (MDN Web Docs, 2024b). Um die HTML-Elemente zu gestalten, wird wiederum die Stylesheet-Sprache CSS verwendet. Dabei wird die Gestaltung der HTML-Elemente in einem CSS-Regelwerk festgehalten. Hier hinterlegte Selektoren ermöglichen das Selektieren und damit das Gestalten gewünschter Elemente (MDN Web Docs, 2024a).

Neben diesen strukturgebenden und gestaltenden Technologien können die Webseiten mithilfe von JavaScript um interaktive Funktionen ergänzt werden. So ist es möglich, dynamische Inhalte zu erstellen, zu denen unter anderem interaktive Spiele, animierte 2D- und 3D-Grafiken sowie komplexe datenbankgestützte Anwendungen gehören. Die JavaScript-Kernsprache ermöglicht den Zugriff auf eine Vielzahl von Tools und Technologien, die über die grundlegenden Funktionen hinausgehen. Insbesondere bietet JavaScript umfassende Unterstützung durch Browser Application Programming Interfaces (API), die in modernen Webbrowsern integriert sind. Diese APIs ermöglichen beispielsweise die Erfassung und Bearbeitung von Videostreams von Webcams oder die Erzeugung von 3D-Grafiken. Darüber hinaus können Entwickler auch auf APIs von Drittanbietern zugreifen, um zusätzliche Funktionen zu integrieren und spezialisierte Dienste zu nutzen. Weiterhin verwendet JavaScript, ähnlich wie andere Programmiersprachen, Variablen und Funktionen, um logische Abläufe und Datenmanipulationen zu ermöglichen (MDN Web Docs, 2024f).

## 2.2.2 Asynchrone Programmierung

Die asynchrone Programmierung ist eine Technik, die es Anwendungen ermöglicht, eine potenziell lang andauernde Aufgabe zu starten, während das Programm gleichzeitig weiterhin auf andere Ereignisse reagieren kann. Sobald die Aufgabe abgeschlossen ist, wird das Ergebnis dem Programm zur Verfügung gestellt. Dies ist insbesondere in der Webentwicklung von Bedeutung, da viele Funktionen, die von Browsern bereitgestellt werden, potenziell zeitintensiv sind und daher asynchron ausgeführt werden müssen. Beispiele hierfür sind das Senden von HTTP-Anfragen mit `fetch` oder der Zugriff auf die Kamera oder das Mikrofon eines Nutzers mittels `getUserMedia` (MDN Web Docs, 2024e).

Ein zentrales Konzept der asynchronen Programmierung sind Promises. Ein Promise-Objekt fungiert als Platzhalter für einen Wert, der zum Zeitpunkt der Erstellung des Promises noch nicht bekannt ist. Anstatt sofort einen endgültigen Wert zu liefern, gibt die asynchrone Methode ein „Versprechen“ zurück, den Wert zu einem späteren Zeitpunkt zu liefern. Diese Struktur ermöglicht es asynchronen Methoden, Werte wie synchrone Methoden zurückzugeben (MDN Web Docs, 2024h). Mithilfe der asynchronen Funktionen `async` und `await` wird der Umgang mit Promises vereinfacht. Eine Funktion, die als `async` deklariert ist, gibt automatisch ein Promise zurück. Innerhalb dieser Funktion kann das Schlüsselwort `await` verwendet werden,

um die Ausführung der Funktion so lange anzuhalten, bis das angegebene Promise erfüllt oder abgelehnt wird (MDN Web Docs, 2024e).

### **2.2.3 Progressive Web App**

PWAs werden mit Web-Technologien erstellt, bieten jedoch ein Benutzererlebnis, das dem von plattformspezifischen Anwendungen ähnelt. Sie kombinieren die besten Eigenschaften von Webseiten und nativen Apps. Wie Webseiten können sie auf mehreren Plattformen und Geräten mit einer einzigen Codebasis betrieben werden. Gleichzeitig ermöglichen sie die Installation auf dem Gerät und bieten Funktionen wie die Offline-Fähigkeit, Push-Benachrichtigungen und Caching-Funktionalitäten (MDN Web Docs, 2024g).

Die beiden wesentlichen Anforderungen an eine PWA sind ein Service Worker und ein Web App Manifest (Ionic Web Docs, 2024b). Service Worker sind spezialisierte JavaScript-Assets, die als Proxys zwischen Webbrowsern und Webservern agieren. Sie verbessern die Zuverlässigkeit und Leistungsfähigkeit von Webseiten, indem sie Offline-Zugriff ermöglichen. Dies wird durch die Nutzung der Cache-Schnittstelle erreicht, die es ermöglicht, Inhalte zwischenspeichern und bei Bedarf bereitzustellen (Chrome Developers, 2021).

Das Web App Manifest ist eine JSON-Datei, die Metadaten über eine Webanwendung bereitstellt. Es umfasst mindestens den Namen der Anwendung, die verwendeten Symbole und die URL, die beim Start der Anwendung geöffnet werden soll. Durch die Definition entsprechender Manifestattribute ermöglicht es Browsern das Verhalten von PWAs zu steuern und die Anwendung korrekt zu installieren (Google Developers, 2022).

## **2.3 Mixed Reality**

MR ist als eine immersive Technologie bekannt, in der die Benutzer in eine virtuelle Umgebung eintauchen (Klinker et al., 2017). Immersion beschreibt dabei den subjektiven Eindruck, an einer umfassenden, realistischen Erfahrung teilzunehmen (Dede, 2009). Die Technologie lässt sich in AR und Virtual Reality (VR) unterteilen. Im Folgenden werden diese beiden Technologien voneinander abgegrenzt und erläutert.

### **2.3.1 Virtual Reality**

VR ermöglicht die Erstellung computergenerierter Welten, in die die Nutzer eintauchen können (Przybilla et al., 2018). Dabei wird die wirkliche Welt vollständig durch eine virtuelle Umgebung, bestehend aus virtuellen Objekten, ersetzt. Dies wird durch eine Kombination von Hardware, wie VR-Headset, und Software ermöglicht (Klinker et al., 2020).

### 2.3.2 Augmented Reality

AR ermöglicht es, virtuelle Objekte in die reale Umgebung der Nutzer zu integrieren (Przybilla et al., 2018). Diese Darstellung erfolgt in Echtzeit und kann durch verschiedene Geräte, wie Smartphones, Tablets oder spezielle AR-Brillen, erlebt werden. Im Gegensatz zur VR bleibt bei AR die physische Welt der primäre Bezugspunkt. Virtuelle Objekte erweitern die Wahrnehmung dieser Welt lediglich (Klinker et al., 2020).

In der wissenschaftlichen Literatur werden verschiedene Formen von AR differenziert. Dazu zählen Hand-Held-AR, die auf mobilen Geräten wie Smartphones und Tablets basieren, sowie Head-Worn-AR, die durch Head-Mounted Displays realisiert werden (Billinghurst et al., 2015). AR-Anwendungen lassen sich zudem in Optical-See-Through (OST) und Video-See-Through (VST) einteilen (Rolland & Fuchs, 2001). VST-AR, wie es bei mobilen Geräten genutzt wird, zeigt die reale Welt über einen Livestream, in den virtuelle Elemente eingefügt werden. Im Gegensatz dazu fügen OST-Geräte, wie die Microsoft HoloLens, die virtuellen Objekte direkt in die Sicht der Nutzer ein, ohne dass ein Livestream nötig ist (Klinker et al., 2020).

AR-Erfahrungen können entweder ortsabhängig oder bildbasiert sein, wobei beide Ansätze unterschiedliche Funktionen von Smartphones nutzen, um immersive Erlebnisse zu schaffen (Azuma et al., 2001; Dede, 2009; L. Johnson et al., 2010). Ortsabhängige AR verwendet GPS-Technologie, um digitale Medien bereitzustellen, während sich die Nutzer durch einen physischen Raum bewegen. Die Erlebnisse sind dabei auf diesen Raum ausgerichtet und nutzen die Geschichte, Geografie und physische Struktur dieser Umgebung. Die Umgebung wird dabei durch Medien wie Text, Grafiken, Audio, Video oder 3D-Modelle ergänzt, um relevante Erzählungen, Navigations- oder wissenschaftliche Informationen bereitzustellen. Im Gegensatz dazu nutzt bildbasierte AR die Kamera eines Mobilgeräts, um digitale Inhalte anzuzeigen, sobald die Kamera auf ein spezifisches Objekt, wie einen QR-Code, ein 2D- oder 3D-Ziel, gerichtet wird (Dunleavy & Dede, 2014).

Jedoch ist die Skalierbarkeit ortsabhängiger AR-Erfahrungen erheblich eingeschränkt, da sie stark auf bestimmte Umgebungen zugeschnitten und dadurch weniger leicht auf andere Orte übertragbar ist. Bildbasierte AR ist im Unterschied dazu aufgrund ihrer hohen Portabilität flexibler einsetzbar (Klopfer & Sheldon, 2010).

### 2.3.3 Taxonomien von Augmented Reality

Die Autoren Milgram und Kishino (1994) beschreiben in ihrer Publikation ein MR Kontinuum. Dabei wird das Spektrum zwischen realer Welt und virtueller Welt beschrieben. Abbildung 2.1 stellt ein MR Kontinuum dar, in dem AR und Augmented Virtuality (AV) eingeordnet werden.

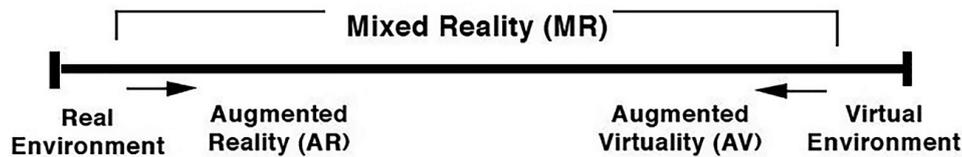


Abbildung 2.1: Abgrenzung von Augmented Reality und Augmented Virtuality (Quelle: Wüller & Koppenburger, 2021, S. 128)

AR wird demnach näher an der realen Welt eingeordnet, da diese Technologie eine größere Anzahl realer Elemente als virtueller Komponenten integriert. Im Gegensatz dazu wird AV näher an der virtuellen Umgebung verortet, da sie überwiegend aus virtuellen Elementen besteht. Hierbei werden reale Objekte in eine vollständig virtuelle Umgebung projiziert, wodurch die virtuelle Welt durch reale Elemente ergänzt wird (Wu et al., 2013). Die Autoren Wüller und Koppenburger (2021) fassen dies wie folgt zusammen: „Anwendungen, die mehr Elemente der realen als einer virtuellen Welt enthalten, gelten als AR, dominieren jedoch virtuelle Elemente, spricht man von AV“ (Wüller & Koppenburger, 2021).

Des Weiteren kann eine Unterscheidung zwischen AR und AV anhand des Ausmaßes der Augmentation vorgenommen werden. Dies wird durch das Gewicht der virtuellen Informationen bestimmt, die den Nutzern bereitgestellt werden. Diese Unterscheidung deutet auf die spezifische Anwendung der Technologie hin. Eine stark augmentierte Welt, in der der Anteil der virtuellen Informationen überwiegt, wird typischerweise durch den Einsatz von Head-Mounted Displays ermöglicht. Im Gegensatz dazu ist eine leicht augmentierte Realität, bei der der Anteil der virtuellen Informationen geringer ist, oft durch die Nutzung mobiler Geräte gekennzeichnet (Klopfer, 2008).

### 2.3.4 3D-Nutzerschnittstellen

Innerhalb einer virtuellen Umgebung werden Interaktionen durch Sensoren, wie etwa Kameras, erfasst. Diese Eingaben werden interpretiert und als Input für die AR-Anwendung verwendet (Wüller & Koppenburger, 2021). Da AR- und VR-Systeme in mobilen und dreidimensionalen Umgebungen eingesetzt werden, sind traditionelle Eingabemethoden wie Tastatur und Maus in diesem Kontext nicht geeignet (Jacob et al., 2008).

Studien haben gezeigt, dass direkte Interaktionen mit 3D-Objekten in virtuellen Umgebungen durch spezialisierte 3D-Nutzerschnittstellen und Interaktionsmethoden ermöglicht werden. Da AR eine stärkere Verbindung zwischen realen und virtuellen Objekten herstellt, ist es erforderlich Benutzereingaben zu berücksichtigen, die sich auf physische Objekte beziehen. Diese Eingaben müssen in eine computererzeugte virtuelle Darstellung umgesetzt werden (Billinghurst et al., 2015).

## 2.4 Künstliche Intelligenz

KI beschäftigt sich mit der Aufgabe, menschliche Wahrnehmungs- und Denkprozesse zu erfassen und zu systematisieren. Ziel ist es, diese Prozesse durch technische Systeme, insbesondere durch informationsverarbeitende Technologien, nachzubilden und nutzbar zu machen (Görz & Schneeberger, 2003). Dies wird durch die Definition von Winston (1992) unterstützt. Dieser formuliert: „Künstliche Intelligenz ist die Lehre der Berechnungen, die es ermöglichen, wahrzunehmen, zu denken und zu handeln“ (Winston, 1992).

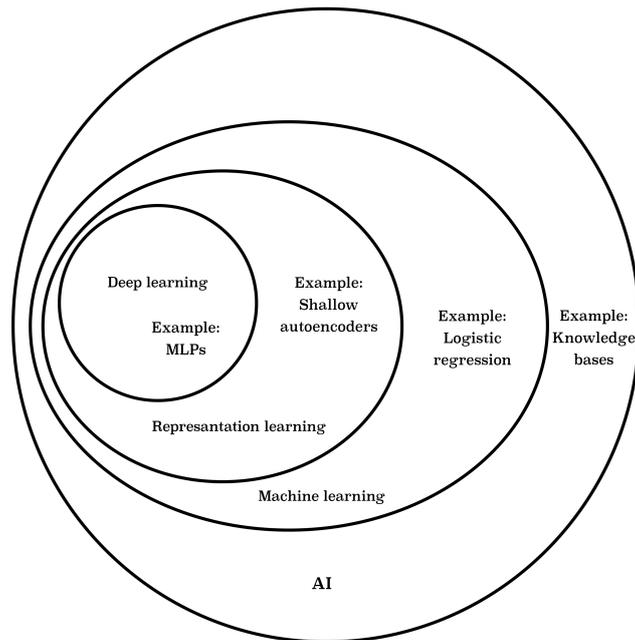


Abbildung 2.2: Teilbereiche der Künstlichen Intelligenz (Quelle: Goodfellow et al., 2016, S. 9)

KI ist ein weitläufiges und komplexes Forschungsfeld, das zahlreiche Teilbereiche umfasst. Abbildung 2.2 verdeutlicht diese Struktur, indem sie zeigt, wie Deep Learning (DL) als spezialisierter Bereich des Maschinellen Lernens (ML) in das umfassendere Feld der KI eingebettet ist.

### 2.4.1 Maschinelles Lernen

Ein wichtiges Teilgebiet der KI ist das ML. Beim ML wird dem Programm eine Lernmethode vorgegeben, auf deren Grundlage es dann mit bereitgestellten Daten trainiert wird. Dabei sucht es beispielsweise nach Mustern oder analysiert die Zusammenhänge zwischen Attributen und möglichen Verhaltensweisen. Mit jedem neuen Datensatz erweitert der Computer so kontinuierlich sein Wissen (Heesen, 2021). Dies befähigt das Programm Probleme zu lösen,

die ein Verständnis der realen Welt erfordern, und subjektiv wirkende Entscheidungen zu treffen (Goodfellow et al., 2016). Dadurch entfällt der Bedarf an spezifischer, zweckgebundener Programmierung (Goodfellow et al., 2016; Samuel, 1959). Dabei werden verschiedene Ansätze unterschieden, darunter Supervised Learning, Unsupervised Learning sowie Reinforcement Learning. Im Folgenden wird das Unsupervised Learning und das Reinforcement Learning kurz vorgestellt. Das Supervised Learning wird ausführlicher behandelt, da das in der Anwendung verwendete Modell mit dieser Methode trainiert wurde.

Das Unsupervised Learning arbeitet mit Daten, die weder klassifiziert noch gelabelt sind. Daher liefert die Maschine Ergebnisse ohne jegliche Beobachtungen (Singh et al., 2021). Die Hauptaufgabe des Unsupervised Learning besteht darin, Informationen nach bestimmten Eigenschaften wie zum Beispiel Form, Größe, Farbe oder Beschaffenheit zu gruppieren (Hastie et al., 2009).

Eine weitere Lernmethode ist das Reinforcement Learning, welches in Spielen und in der Robotik verwendet wird. Dabei lernt das Modell durch Interaktionen mit seiner Umgebung, bei denen es Belohnungen oder Strafen erhält, um seine Handlungen entsprechend anzupassen (Hinton & Sejnowski, 1999). Das Ziel dieses Ansatzes besteht nicht darin, die Maschine auf eine bestimmte Aktion zu trainieren. Vielmehr soll das Modell in die Lage versetzt werden, auf Basis der vorliegenden Daten eigenständig die erforderlichen Schritte zu identifizieren, um eine Aufgabe erfolgreich zu bewältigen (Singh et al., 2021).

Die dritte Lernmethode ist das Supervised Learning, bei dem das Modell mit gelabelten Daten trainiert wird. Bei diesen Daten liegt für jede Eingabe bereits die korrekte Ausgabe vor. Diese Eingabe-Ausgabe-Paare dienen als Referenz, um dem Modell zu ermöglichen, Muster zu erkennen und Vorhersagen unbekannter Eingaben zu treffen (Singh et al., 2021).

Ein Beispiel für das Supervised Learning ist das Klassifizieren von Bildern. Hierbei werden Bilder von verschiedenen Objekten, wie etwa Hunden, Katzen, Autos oder Fahrrädern, mit entsprechenden Labels versehen. Diese gekennzeichneten Bilder werden dann verwendet, um den Algorithmus zu trainieren. Die Aufgabe des Modells besteht darin, die Bilder korrekt den jeweiligen Kategorien zuzuordnen, indem es aus den Trainingsdaten Muster und Merkmale erkennt.

Beispielhafte Klassifizierungsalgorithmen für das Trainieren eines Modells mithilfe des Supervised Learning sind Entscheidungsbäume, Naive Bayes und k-Nearest Neighbors (KNN) (Singh et al., 2021). Diese Algorithmen eignen sich jedoch nur für flache Modelle und kleinere Datensätze. Moderne Aufgaben wie die Objekterkennung werden hingegen am effektivsten mit DL gelöst (Goodfellow et al., 2016).

## 2.4.2 Deep Learning

DL ist ein Teilbereich des ML, der durch den Einsatz neuronaler Netze komplexe Datenverarbeitung ermöglicht. Neuronale Netze bestehen aus einer Vielzahl von Neuronen, die in mehreren Schichten organisiert sind. Neuronen empfangen Signale von der vorherigen Schicht und verarbeiten diese Informationen mithilfe von Gewichten und Aktivierungsfunktionen. Jede Schicht baut demnach auf der vorherigen auf und setzt Schritt für Schritt aus einfacheren Teilen komplexere Konzepte zusammen (Goodfellow et al., 2016).

Abbildung 2.3 zeigt die Funktionsweise eines DL-Modells anhand der schrittweisen Analyse eines Bildes. Die Verarbeitung erfolgt in mehreren aufeinanderfolgenden Schichten, um komplexe Sensordaten eines Bildes effizient zu interpretieren. Zunächst werden die Eingabedaten in der sichtbaren Schicht abgebildet, welche die Rohdaten des Bildes enthält. Dazu zählen Pixelwerte, die die grundlegenden visuellen Informationen wie Farb- und Helligkeitswerte repräsentieren. Diese Pixelwerte werden anschließend durch eine Reihe verborgener Schichten verarbeitet, die zunehmend abstrakte Merkmale wie Kanten, Ecken und Konturen extrahieren. Durch die Erkennung mehrerer abstrakter Merkmale können beispielsweise ganze Teile spezifischer Objekte identifiziert werden. Die extrahierten Merkmale werden schließlich kombiniert, um das Objekt als Ganzes zu erkennen. Dieser mehrstufige Ansatz der Abstraktion und Verarbeitung ermöglicht es dem Modell, komplexe Muster präzise zu erkennen und zu interpretieren (Goodfellow et al., 2016).

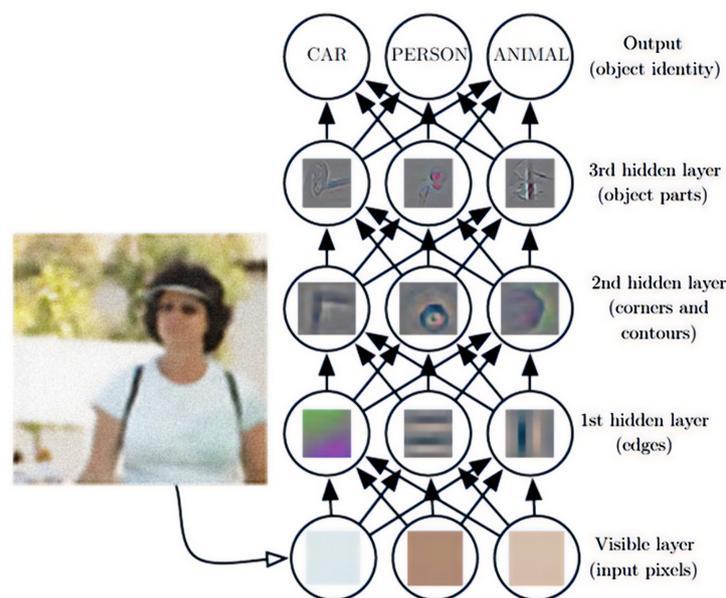


Abbildung 2.3: Abbildung eines Deep Learning-Modells (Quelle: Goodfellow et al., 2016, S. 6)

DL umfasst Optimierungsalgorithmen, die dafür sorgen, dass das neuronale Netz durch Anpassung seiner Parameter stetig verbessert wird. Diese Algorithmen minimieren Fehler

und optimieren die Leistung des Modells, indem sie die Gewichte der Verbindungen zwischen den Neuronen anpassen (Goodfellow et al., 2016).

### 2.4.3 Transfer Learning

Das Training von neuronalen Netzen kann äußerst ressourcenintensiv und zeitaufwendig sein. Darüber hinaus kann die Beschaffung geeigneter Daten eine erhebliche Herausforderung darstellen, da Daten in einigen Fällen selten oder kostspielig zu erwerben sind. Dies kann dazu führen, dass es in einigen Fällen nicht praktikabel ist ein neues Modell zu trainieren. In solchen Situationen kann Transfer Learning (TL) als Alternative eingesetzt werden (Hosna et al., 2022; Torrey & Shavlik, 2010).

TL ist ein Teilbereich des ML, bei dem Wissen aus früheren Aufgaben genutzt wird, um die Effizienz bei neuen Aufgaben zu steigern. Abbildung 2.4 zeigt die Unterschiede zwischen traditionellen ML- und modernen TL-Strategien. Dabei wird sichtbar, dass sich TL von herkömmlichen maschinellen Lernansätzen unterscheidet, bei denen jedes Modell von Grund auf neu trainiert wird (Seldon Technologies, 2021). Stattdessen wird beim TL Wissen aus zuvor bearbeiteten Aufgaben genutzt und auf ein ähnliches Anwendungsgebiet übertragen. Dies verbessert das Lernen, indem es Verknüpfungen zwischen früherem Wissen und der gegenwärtigen Aufgabe herstellt, wodurch es zu schnelleren, logischeren und dadurch besseren Lösungen kommt (Hosna et al., 2022; Pan & Yang, 2009).

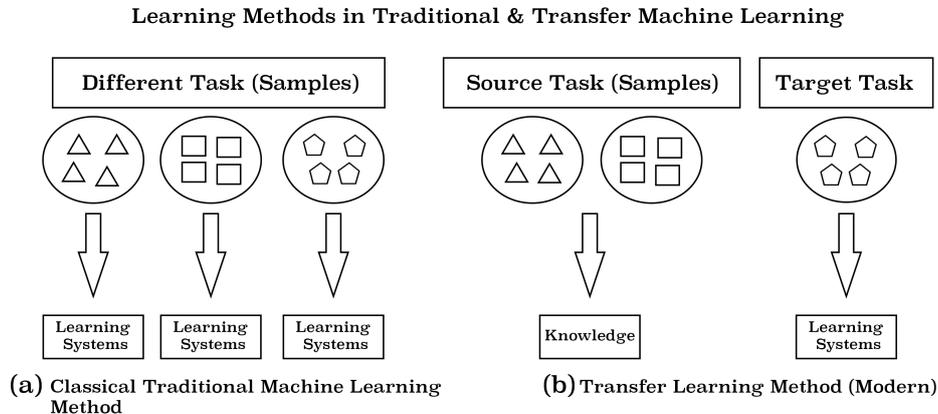


Abbildung 2.4: Unterschied zwischen traditionellen maschinellen Lernmethoden und modernen Transferlern-Strategien (Quelle: Hosna et al., 2022, S. 2)

In der Praxis kommen häufig vortrainierte Modelle, insbesondere DL-Modelle, zum Einsatz, die auf umfangreichen Datensätzen trainiert wurden. Beispiele hierfür sind ImageNet für Bildverarbeitungsaufgaben oder große Textkorpora für Aufgaben im Bereich des Natural Language Processing (NLP). Beim TL werden die oberen Schichten des neuronalen Netzes

entfernt und durch neu konfigurierte Schichten ersetzt. Dieser Ansatz ermöglicht es, die in den tieferen Schichten des Netzes bereits erlernten Merkmale beizubehalten, während die obersten Schichten an die spezifischen Anforderungen der neuen Zielaufgabe angepasst werden (Hosna et al., 2022).

#### **2.4.4 Anwendungsbeispiele**

Heutzutage bildet KI das Kernstück vieler in der Industrie eingesetzter Technologien. Besonders ausgeprägt ist der Einsatz in den Bereichen Sprachverarbeitung und Computer Vision. Die Sprachverarbeitung umfasst unter anderem intelligente Geräte und Sprachassistenten wie Alexa, Siri und Cortana (Heesen, 2021). Diese Technologien dienen der Automatisierung von Aufgaben, der Beschleunigung von Entscheidungsprozessen und der Verbesserung der Kundenkommunikation durch Chatbots (Columbia University, 2024).

Auch der Bereich der Computer Vision ist von zentraler Bedeutung. Hierzu zählen insbesondere Bildklassifizierung und Objekterkennung (Hosna et al., 2022). Besonders hervorzuheben ist die Objekterkennung, die in dieser Arbeit eine zentrale Rolle spielt. Objekterkennung bezieht sich auf die Fähigkeit eines KI-Agenten, spezifische Objekte in Bildern oder Videos zu identifizieren und deren Positionen zu bestimmen. Dabei wird nicht nur das Vorhandensein von Objekten erkannt, sondern auch deren räumliche Lage erfasst (Felzenszwalb et al., 2009; Zhao et al., 2019). Die Bilderkennung, die häufig mit der Objekterkennung kombiniert wird, umfasst darüber hinaus das Erkennen von Personen, Texturen oder komplexen Szenen. In beiden Bereichen kommen häufig Convolutional Neural Networks (CNNs) zum Einsatz, die speziell für die effiziente Verarbeitung visueller Daten entwickelt wurden (Krizhevsky et al., 2012). Diese Netzwerke sind in der Lage, relevante Merkmale in Bildern zu identifizieren, die für die Klassifizierung oder Lokalisierung von Objekten entscheidend sind (Zhao et al., 2019).

## **2.5 Vergleichbare Arbeiten**

Im folgenden Abschnitt wird der Einsatz von AR im Bildungsbereich näher beleuchtet. Dabei werden sowohl die pädagogischen Chancen als auch die Herausforderungen von AR im Kontext des Lehrens und Lernens untersucht.

### **2.5.1 Augmented Reality Teaching and Learning**

In dem Buch „Augmented reality teaching and learning“ von Dunleavy und Dede (2014) werden Chancen und Herausforderungen von AR für Lehren, Lernen und Unterrichtsgestaltung untersucht. Diese Forschung stützt sich auf zwei zentrale Theorien, nämlich die Theorie

des situierten Lernens und die konstruktivistische Lerntheorie, die bereits im Abschnitt 2.1 umfassend behandelt wurden.

Laut der Theorie des situierten und konstruktivistischen Lernens eignet sich AR als pädagogischer Ansatz, da es Lernende in einen realen physischen und sozialen Kontext versetzt. AR fördert und erleichtert partizipative und metakognitive Lernprozesse, darunter Untersuchungen, gezielte Beobachtungen, Peer-Coaching, wechselseitiges Lehren und eine aktive Einbindung der Lernenden (Dunleavy et al., 2009; Klopfer & Sheldon, 2010; Palincsar, 1998; K. Squire, 2010). Außerdem steht das Lernen mit AR im Einklang mit den in Abschnitt 2.1.2 genannten Bedingungen der konstruktivistischen Lerntheorie, die als förderlich für effektives Lernen gelten. Besonders deutlich wird dies durch die Einbettung von AR in relevante, authentische Umgebungen, die den Lernprozess für die Lernenden bedeutsam machen. Darüber hinaus fördert AR sozialen Austausch und kollaborative Lernprozesse, indem es Interaktionen zwischen den Lernenden in die Lernerfahrung integriert. Schließlich ermöglicht AR eigenverantwortliches Lernen, indem es den Lernenden die Möglichkeit gibt, ihre Lernaktivitäten selbst zu steuern und zu gestalten.

Diese Wirkungen sind durch eine Reihe von designbasierten Forschungen (DBR) bewiesen, welche die Machbarkeit und Praktikabilität des Einsatzes von AR in der K-12-Umgebung zum Lehren und Lernen untersuchen. Die Studien zeigen, dass AR die physische Umgebung mit digitalen Informationen ergänzt und somit vielfältige Lernmöglichkeiten schafft, die anderweitig nicht realisierbar wären (Facer et al., 2014; Klopfer, 2008; Klopfer & Squire, 2008; Schmalstieg & Wagner, 2007; K. D. Squire et al., 2007).

AR ermöglicht den Einsatz externer Ressourcen und zusätzlicher Software zur effektiveren Lösung spezifischer Probleme. Darüber hinaus lassen sich die Technologien auf unvorhergesehene, aber vorteilhafte Weise verwenden. So kann beispielsweise die Internet- oder Videoaufzeichnungsfunktion eines Smartphones genutzt werden, wodurch handschriftliche Notizen überflüssig werden (Perry et al., 2008). Zudem wird deutlich, dass AR die Motivation der Schüler erheblich steigern kann (Dunleavy & Dede, 2014).

Die Forschung zeigt jedoch auch Herausforderungen auf, die mit dem Einsatz von Technologie im Lernkontext verbunden sind. Dazu zählen insbesondere die potenzielle kognitive Überlastung, die durch die Komplexität der Aktivitäten, sowie die Schwierigkeiten bei der Entscheidungsfindung im Team bedingt ist (Dunleavy et al., 2009; Klopfer & Squire, 2008; Perry et al., 2008). Diese Herausforderungen können jedoch durch ein verbessertes Design und fortschreitende technologische Entwicklungen reduziert werden (O'Shea et al., 2009; Perry et al., 2008).

Im Abschnitt 2.3.2 werden zwei Haupttypen von AR-Erfahrungen beschrieben: ortsabhängige AR und bildbasierte AR. Ortsabhängige AR eignet sich besonders gut für authentische Umweltbeobachtungen, da sie die Lernerfahrung vertieft, indem sie den physischen Raum

mit relevanten Themen verknüpft. Die Auswahl umstrittener oder narrativ ansprechender Räume ermöglicht es, den Lernenden durch thematisch relevante und vertraute Umgebungen ein höheres Maß an Kontrolle und Bedeutung zu bieten (Klopfer & Sheldon, 2010; Perry et al., 2008; K. D. Squire et al., 2007). Bildbasierte AR ist allerdings flexibler einsetzbar, da sie unabhängig vom physischen Standort funktioniert und vielfältige visuelle Inhalte direkt auf Objekte projizieren kann (Klopfer & Sheldon, 2010). Insbesondere eine Kombination beider Ansätze kann zu pädagogisch wertvollen und immersiven Lernerfahrungen führen, da sie sowohl die räumliche Relevanz als auch die visuelle Informationsvielfalt integriert (Klopfer & Sheldon, 2010).

## **2.5.2 Current status, opportunities and challenges of Augmented Reality in education**

Der Artikel von Wu et al. (2013) untersucht die Anwendung von AR im Bildungsbereich und zeigt auf, wie AR als innovatives Werkzeug zur Verbesserung von Lernerfahrungen beitragen kann. Die Studie analysiert aktuelle Entwicklungen, Möglichkeiten und Herausforderungen, die mit der Integration von AR in Bildungskontexten verbunden sind.

Die Autoren definieren AR und nicht als spezifische Technologie. Diese Sichtweise ermöglicht es, den pädagogischen Wert von AR und das Potenzial zur Unterstützung von Lernprozessen stärker hervorzuheben (Wu et al., 2013). AR wird aufgrund der pädagogischen Vorteile als eine der wichtigsten aufkommenden Technologien für die Bildung angesehen (L. Johnson et al., 2010; L. F. Johnson et al., 2010; Martin et al., 2011). Es wird Lernenden ermöglicht, Inhalte aus verschiedenen Perspektiven zu betrachten, wodurch das Verständnis komplexer Themen verbessert wird (Chen et al., 2011). Durch die Einbindung von mobilen Geräten, drahtloser Verbindung und ortsbezogener Technologie fördert AR kollaboratives und situiertes Lernen, das es den Lernenden erlaubt, in realen Kontexten zu interagieren (Broll et al., 2008; Dunleavy et al., 2009). Die immersive Natur von AR schafft zudem ein Gefühl der Präsenz, das das Bewusstsein der Lernenden für Gemeinschaften stärken kann (K. D. Squire & Jan, 2007). Dadurch wird auch die soziale Interaktivität gefördert, indem Lernende durch direkte Interaktionen miteinander enger zusammenarbeiten können (Birchfield & Megowan-Romanowicz, 2009). Darüber hinaus ermöglicht AR die Visualisierung unsichtbarer Konzepte oder Ereignisse, indem virtuelle Objekte oder Informationen auf physische Objekte oder Umgebungen projiziert werden. Dies erleichtert das Verständnis komplexer räumlicher Beziehungen und abstrakter Konzepte (Arvanitis et al., 2009; Dunleavy et al., 2009). Schließlich hat AR das Potenzial, die Kluft zwischen Lernen in formellen und informellen Kontexten zu überbrücken, indem es das wissenschaftliche Lernen in der Schule mit realen Erfahrungen verbindet (Wu et al., 2013).

Der Erfolg von AR begründet sich auf die im Abschnitt 2.1.3 erwähnten Lernansätze, die sich in drei Hauptkategorien zusammenfassen lassen: ortsbasiertes Lernen, spielerisches Lernen und problemorientiertes Lernen. Beim ortsbasierten Lernen liegt der Schwerpunkt auf der Interaktion der Lernenden mit ihrer physischen Umgebung, was durch AR unterstützt wird und ein Gefühl von Authentizität schafft, das beim Lernen hilfreich ist (Rosenbaum et al., 2007). Spielerisches Lernen, das durch die Kombination von Spaß, Herausforderung und Neugier gekennzeichnet ist, kann durch AR zusätzlich angeregt werden, da es die Motivation und das Interesse der Lernenden erhöht (Sotiriou & Bogner, 2008). Beim problemorientierten Lernen wird die Entwicklung von selbstgesteuertem Lernen, Selbstmotivation, Problemlösungsfähigkeiten und Wissensanwendungsfähigkeiten gefördert. AR unterstützt diese Aspekte, indem es realistische Probleme und Szenarien bietet, die die Lernenden aktiv lösen müssen (Liu, 2009).

Wu et al. heben abschließend hervor, dass AR zwar ein großes Potenzial für den Bildungsbereich besitzt, die Forschung jedoch noch in einem frühen Entwicklungsstadium steckt und zahlreiche Studien methodische Begrenzungen aufweisen. Daher besteht ein Bedarf an weiteren umfassenden und methodisch soliden Studien, um die tatsächlichen Auswirkungen von AR auf das Lehren und Lernen besser zu verstehen und zu bewerten.

### **2.5.3 EcoMOBILE: Integrating Augmented Reality and probeware with environmental education field trips**

Kamarainen et al. (2013) analysieren die wesentlichen Merkmale des Lernens und Lehrens im Rahmen des EcoMOBILE-Projekts (Ecosystems Mobile Outdoor Blended Immersive Learning Environment) (Harvard Graduate School of Education, 2024). In dieser Studie werden die Einstellungen der Schüler, ihre Lernfortschritte sowie die Rückmeldungen der Lehrkräfte systematisch erhoben, um die Effektivität des Unterrichts mit dieser Anwendung umfassend zu bewerten.

Das EcoMOBILE-Projekt kombiniert den Einsatz von AR mit Werkzeugen zur Wasseranalyse. Mobile, drahtlose Geräte, die mit der AR-Anwendung FreshAiR ausgestattet sind, ermöglichen den Schülern die Navigation in einer Teichumgebung, in der virtuelle Medien und Informationen über die physische Umgebung projiziert werden. An ausgewählten AR-Hotspots führen die Schüler Messungen der Wasserqualität durch. Zur Evaluierung werden Umfragen durchgeführt, bei denen die Schüler vor und nach dem EcoMOBILE-Pilotcurriculum befragt wurden. Diese Befragungen konzentrierten sich auf affektive Aspekte sowie auf das inhaltliche Verständnis. Darüber hinaus erfolgte eine Nachbefragung, um die Schülermeinungen hinsichtlich ihrer Zufriedenheit mit dem Exkursionserlebnis zu erfassen. Lehrkräfte wurden ebenfalls befragt, um deren Einschätzungen zur Motivation der Schüler zu dokumentieren. Die Ergebnisse zur Selbstwirksamkeit der Schüler zeigten eine positive Entwicklung ihres

Wissens und ihrer Fähigkeiten im Bereich der Ökosystemwissenschaften. Insbesondere ihre Fähigkeit zentrale Themen zu verstehen und wissenschaftliche Methoden anzuwenden hat sich verbessert. Die Analyse des inhaltlichen Verständnisses erfolgte durch Fragen zur Inhaltsbewertung, wobei die Antworten als richtig oder falsch klassifiziert wurden. Die Vorher-Nachher-Befragungen ergaben signifikante Lernfortschritte mit einer durchschnittlichen Steigerung der Inhaltsbewertung um 19 %. In der Umfrage zur Bewertung des Exkursionserlebnisses, die auf einer Skala von 1 bis 7 durchgeführt wurde, lag die durchschnittliche Bewertung bei 5,4.

Die positiven Ergebnisse lassen sich durch die Erkenntnisse verschiedener Forschender begründen. Laut diesen fördert die Durchführung der Lernaktivitäten in einer realen Umgebung das Verständnis von Ökosystemen und beeinflusst die Einstellung der Schüler zur Natur positiv (Ballantyne & Packer, 2002; Bitgood, 1989; Bogner, 1998; Garner & Gallo, 2005; Gottfried, 1980; Knapp & Barrie, 2001; Manzanal et al., 1999). Gleichwohl könnte die Neuheit des sozialen und physischen Kontexts potenziell zu Ablenkungen führen und die Konzentration auf relevante Lernaufgaben beeinträchtigen (Falk, 1983; Orion & Hofstein, 1994). Diese Befürchtung wurde jedoch durch die Befragungsergebnisse entkräftet. Diese zeigen, dass die Technologie keine Ablenkung von der realen Umgebung darstellte. Die Nachbefragungen und Rückmeldungen der Lehrkräfte belegten vielmehr positive Auswirkungen auf die Motivation und das Engagement der Schüler.

Darüber hinaus unterstützt die Verknüpfung von Wissen und wissenschaftlicher Praxis die Theorie des situierten Lernens. Dies wird dadurch belegt, dass sie Verbindungen zu eigenen Erfahrungen schafft und die Kognition sowohl im physischen als auch im psychosozialen Kontext verankert (Council et al., 2012; Greeno, 1998). Des Weiteren belegen empirische Studien, dass der Einsatz von mobilen Geräten verschiedene Aspekte des Lehrens und Lernens fördert (Metcalf & Tinker, 2004; Nachmias & Linn, 1987; Nicolaou et al., 2007; Zucker et al., 2008). Die Analyse der Inhaltsbefragung zeigt zudem eine Förderung kognitiver Fähigkeiten wie kritisches Denken und Problemlösung sowie die Entwicklung eines tieferen Verständnisses (McCombs & Whisler, 1997; Perkins & Salomon, 1992).

Abschließend betont der Autor, dass weitere Studien erforderlich sind, um die Auswirkungen unterschiedlicher Einsätze sowie die langfristige Wirksamkeit der Lernfortschritte der Schüler zu untersuchen und den Einsatz der Technologie zu bewerten.

## 3 Konzept

In diesem Kapitel wird das grundlegende Konzept der entwickelten Anwendung vorgestellt, das sich durch den Einsatz moderner Technologien wie ML und PWA auszeichnet. Das Kernkonzept wird im Abschnitt 3.1 detailliert dargestellt, während das Design im Abschnitt 3.2 erläutert wird. Technische Frameworks und weitere Tools, die für die Projektdurchführung erforderlich sind, werden in den Abschnitten 3.4 und 3.5 beschrieben. Diese Abschnitte bieten eine umfassende Übersicht über die verwendeten Technologien und deren Anwendung im Projekt.

### 3.1 Kernkonzept

Das Kernkonzept der entwickelten PWA basiert auf der Nutzung mobiler Endgeräte zur Identifikation von Früchten mithilfe der integrierten Kamera. Der Schwerpunkt liegt auf der Verwendung von Smartphones, da diese Geräte aufgrund ihrer ständigen Verfügbarkeit und Handlichkeit besonders gut für das Prinzip der Anwendung geeignet sind.

Auf der Startseite der App wird ein kontinuierlicher Video-Stream angezeigt, über den die Umgebung in Echtzeit erfasst wird. Die App ermöglicht es den Nutzern durch diesen Video-Stream Früchte zu entdecken und zu identifizieren. Durch dieses Verhalten werden Barrieren überwunden, da der Nutzer die Früchte nicht manuell suchen oder identifizieren muss.

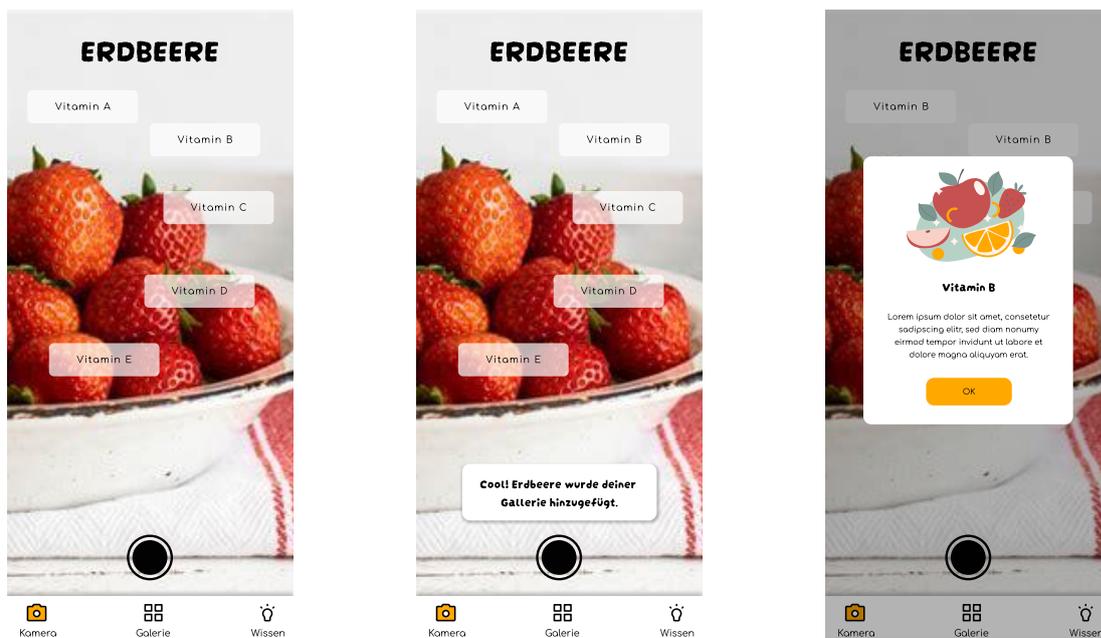
Das zugrunde liegende maschinelle Lernmodell der App wurde speziell auf die Erkennung von sechs unterschiedlichen Früchten trainiert: Apfel, Banane, Birne, Zitrone, Pflaume und Orange. Diese Auswahl basiert auf den deutlichen Unterschieden in Bezug auf die Form und Farbe der Früchte, was eine verbesserte Erkennung durch das Objekterkennungsmodell begünstigt. Sobald eine der Früchte im Video-Stream detektiert wird, werden der Name sowie die relevanten Nährstoffinformationen als Overlay über der Kameradarstellung eingeblendet. Darüber hinaus können Nutzer durch Interaktion mit den dargestellten Nährstoffen detailliertere Informationen abrufen. Zusätzlich bietet die App die Möglichkeit, ein Foto der erkannten Frucht aufzunehmen. Dieses Bild wird dann in einer Galerie innerhalb der App gespeichert, sodass die Nutzer die entdeckte Frucht zu einem späteren Zeitpunkt erneut betrachten und die entsprechenden Informationen abrufen können.

Ein zentraler Bestandteil der App ist das Sammelkonzept, bei dem die Nutzer ihren Fortschritt in Form einer Fortschrittsanzeige angezeigt bekommen. Diese zeigt an, wie viele der verfügbaren Früchte bereits entdeckt und gesammelt wurden, was den Anreiz zur weiteren Nutzung der App steigern soll.

### 3.2 Designkonzept

Das Design wurde mit dem Programm Figma entwickelt. Figma bietet eine intuitive Umgebung, um Benutzeroberflächen zu gestalten und interaktive Prototypen zu erstellen. Im Gestaltungsprozess lag der Schwerpunkt auf der mobilen Optimierung, um eine optimale Benutzererfahrung auf Smartphones sicherzustellen.

Der Startbildschirm der App verfügt über eine Tab-Navigation am unteren Bildschirmrand, die den Zugriff auf die CameraPage, die GalleryPage und die KnowHowPage ermöglicht. Auf der CameraPage wird ein Video-Stream bereitgestellt, der die Umgebung kontinuierlich auf das Vorkommen von Früchten analysiert.



(a) Erkennung einer Frucht

(b) Benachrichtigungsmodal

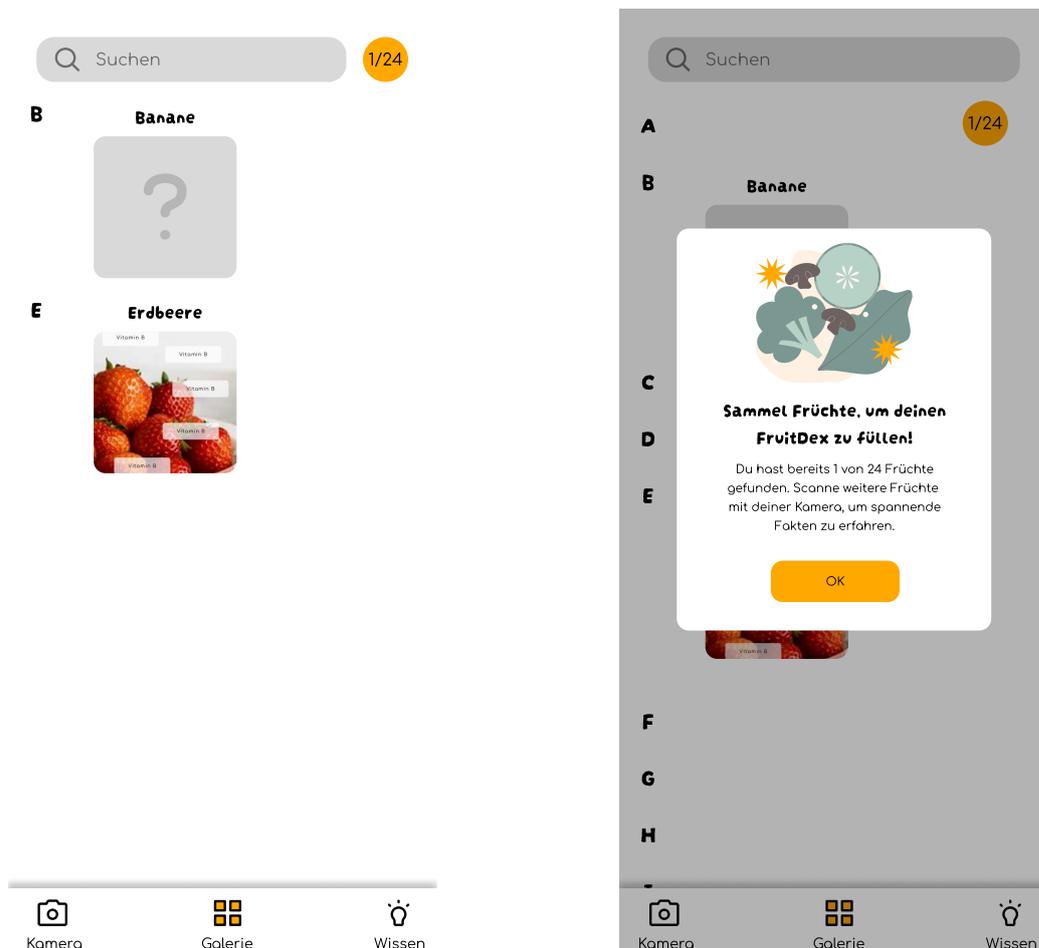
(c) Informationsmodal

Abbildung 3.1: Designkonzept der CameraPage (Quelle: Eigene Darstellung)

Sobald eine Frucht erkannt wird, erfolgt ein akustisches Signal, begleitet von einer Einblendung der wesentlichen Nährstoffinformationen. Durch Interaktion mit den angezeigten Informationen wird ein modales Fenster geöffnet, das weiterführende Details zu den Nährstoffen bereitstellt. Der Benutzer kann ein Foto der erkannten Frucht aufnehmen, wobei

dieser Vorgang durch ein Benachrichtigungsmodal bestätigt wird. Die Abbildungen 3.1 a bis c veranschaulichen diese Funktionalitäten am Beispiel einer Erdbeere.

Die GalleryPage bietet eine Suchleiste am oberen Bildschirmrand, um eine schnelle Auffindbarkeit der gesammelten Früchte zu ermöglichen. Unterhalb der Suchleiste befindet sich eine interaktive Fortschrittsanzeige, die den aktuellen Fortschritt des Benutzers anzeigt. Zusätzlich ermöglicht die in den Abbildungen 3.2 a und b visualisierte Fortschrittsanzeige dem Benutzer weitere Interaktionen.



(a) Ansicht der Galerie mit Fortschrittsanzeige

(b) Informationsmodal der Fortschrittsanzeige

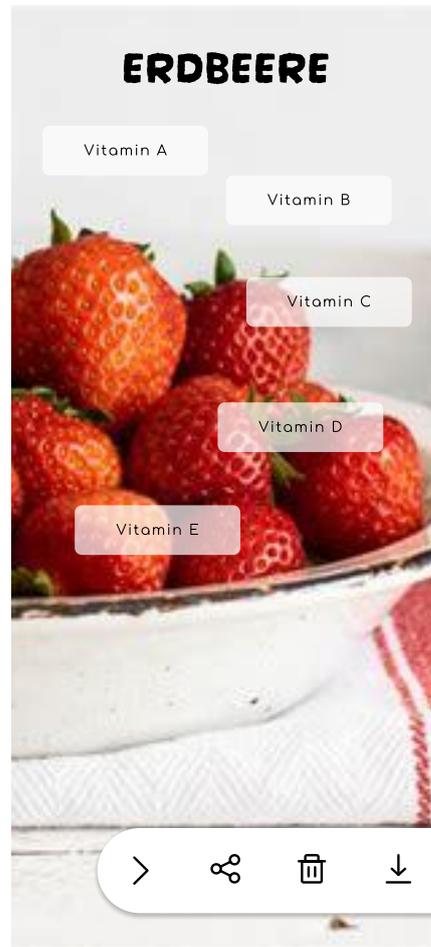
Abbildung 3.2: Designkonzept der GalleryPage (Quelle: Eigene Darstellung)

Die gefundenen Früchte werden alphabetisch aufgelistet und können durch das Auswählen der jeweiligen Vorschau in den Vollbildmodus versetzt werden. Früchte, die lediglich gesehen oder gelöscht wurden, bleiben in der Galerie bestehen, werden jedoch nicht als Fortschritt gezählt. Die Abbildungen 3.3 a und b stellen den Vollbildmodus dar. In diesem sind weitere Interaktionsmöglichkeiten verfügbar, wie das Teilen der Fruchtinformationen, das Löschen der Frucht aus der Galerie oder das Herunterladen des Bildes. Zusätzlich bieten die

Fruchtinformationen in diesem Bereich erweiterte Interaktionsmöglichkeiten und detaillierte Informationen.



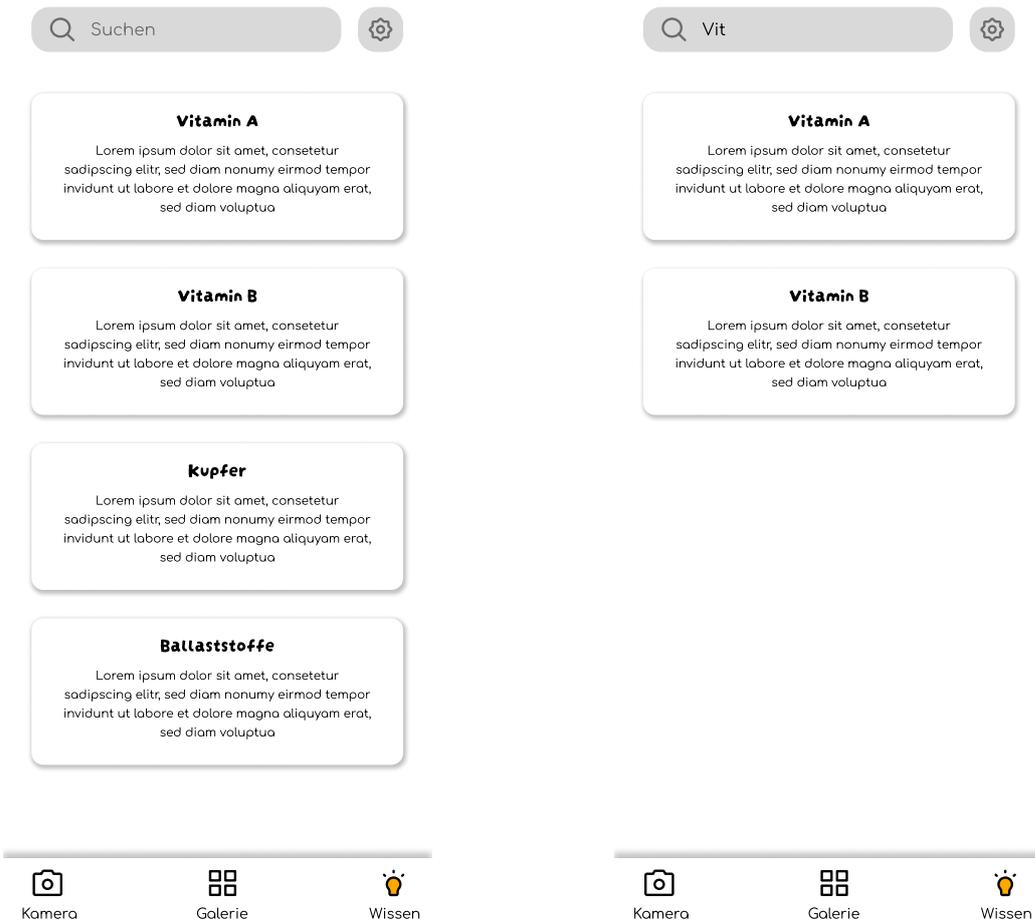
(a) Vollbildmodus



(b) Vollbildmodus mit Interaktionsbar

Abbildung 3.3: Designkonzept der FullScreenImagePage (Quelle: Eigene Darstellung)

Die Abbildungen 3.4 a und b repräsentieren die KnowHowPage. Diese liefert eine umfassende Übersicht über sämtliche in den Früchten enthaltenen Nährstoffe. Am oberen Bildschirmrand befindet sich eine Suchleiste, die den Benutzern ermöglicht, gezielt nach bestimmten Nährstoffen zu suchen. Neben der Suchleiste befindet sich ein Einstellungssymbol, das den Zugriff auf weitere Anpassungsoptionen und Konfigurationen der App ermöglicht.

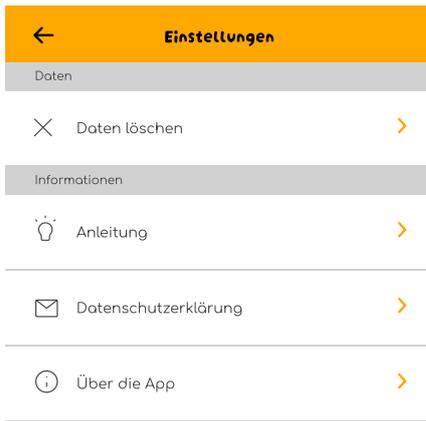


(a) Liste aller Nährstoffe

(b) Gefilterte Liste aller Nährstoffe

Abbildung 3.4: Designkonzept der KnowHowPage (Quelle: Eigene Darstellung)

Die SettingPage wird in den Abbildungen 3.5 a und b dargestellt und ist in *Daten* und *Informationen* gegliedert, um eine klare und benutzerfreundliche Navigation zu ermöglichen. Im Bereich *Daten* haben Benutzer die Möglichkeit, persönliche Daten zu verwalten. Mithilfe der Option *Daten löschen* können die gespeicherten Daten vollständig entfernt werden. Diese Funktion bietet eine wichtige Möglichkeit zur Wahrung der Privatsphäre und zur Bereinigung von Daten. Der Abschnitt *Informationen* bietet umfassende Details zur App. Er enthält mehrere Unterpunkte, die den Nutzern wichtige Informationen bereitstellen. Die Anleitung bietet eine detaillierte Erklärung zur Nutzung der App, einschließlich ihrer Funktionen und Bedienungshinweise. Die Datenschutzerklärung informiert darüber, wie Benutzerdaten gesammelt, verwendet und geschützt werden. Abschließend bietet der Punkt *Über die App* einen umfassenden Überblick über die Anwendung, einschließlich Informationen über die Entwicklerin, die Verwendungszwecke und weitere relevante Details.



(a) Liste der Einstellungen



(b) Beispielhafte Inhaltsseite

Abbildung 3.5: Designkonzept der SettingPage (Quelle: Eigene Darstellung)

## 3.3 Entwicklungsumgebungen

Die Wahl der geeigneten Tools ist entscheidend, um sowohl eine effiziente Entwicklung als auch eine reibungslose Integration und Ausführung sicherzustellen. In diesem Abschnitt werden die genutzten Entwicklungsumgebungen vorgestellt und deren Einsatzgebiete erläutert.

### 3.3.1 IntelliJ IDEA Ultimate

IntelliJ IDEA Ultimate ist eine integrierte Entwicklungsumgebung von JetBrains. Diese Umgebung ist für eine Vielzahl von Programmiersprachen optimiert, darunter JavaScript. Dies wird durch die JavaScript-fähige Codierungshilfe deutlich, welche die Vervollständigung von Schlüsselwörtern, Bezeichnungen, Variablen, Parametern und Funktionen, Fehler- und

Syntaxhervorhebung, Formatierung, Code-Inspektionen und Quick-Fixes sowie allgemeines und JavaScript-spezifisches Refactoring umfasst (Jetbrains, 2024). Des Weiteren unterstützt IntelliJ eine Vielzahl von populären Frameworks. Durch die Integration mit verschiedenen Tools für die Webentwicklung eignet sich die Umgebung, um moderne Web-, Mobil- und Desktop-Anwendungen mit JavaScript zu entwickeln (Jetbrains, 2024).

In dieser Arbeit wurde IntelliJ IDEA Ultimate in der Version 2023.1.2 verwendet. Die Entwicklungsumgebung wurde vor allem wegen ihrer leistungsstarken Funktionen zur Code-Navigation, automatischen Vervollständigung und Fehlererkennung eingesetzt.

### 3.3.2 Google Colab

Google Colaboratory, allgemein bekannt als Google Colab, ist eine Entwicklungsumgebung, die besonders in den Bereichen Data Science und ML beliebt ist. Basierend auf Jupyter-Notebooks ermöglicht Colab das Schreiben und Ausführen von Python-Code direkt im Browser, ohne dass komplexe Einrichtungsvorgänge erforderlich sind. Ein besonderer Vorteil von Colab ist der kostenlose Zugang zu leistungsstarken Rechenressourcen wie Graphics Processing Unit (GPU) und Tensor Processing Units (TPU), was die Entwicklung und das Training von rechenintensiven Modellen erheblich beschleunigt. Außerdem ist die Umgebung cloudbasiert, sodass das Projekt von jedem Gerät aus zugänglich ist. Dies ist besonders für eine gemeinsame Nutzung hilfreich.

In der Google Colab Umgebung sind populärere Python-Bibliotheken wie TensorFlow, PyTorch und Keras bereits vorinstalliert. Dadurch können Nutzer sofort mit dem Programmieren beginnen, ohne Zeit für die Installation und Aktualisierung dieser Bibliotheken aufwenden zu müssen (Levy, 2023).

In dieser Arbeit wurde Google Colab in der Version 2024-08-20 verwendet. Google Colab diente primär der Implementierung und dem Training eines maschinellen Lernmodells.

## 3.4 Frameworks

Frameworks werden in der Softwareentwicklung häufig eingesetzt, um die Effizienz zu steigern, bewährte Architekturen zu nutzen und die Entwicklungszeit zu verkürzen. Sie bieten eine strukturierte Grundlage, auf der komplexe Anwendungen erstellt und gewartet werden können. In diesem Abschnitt werden die in der Anwendung verwendeten Frameworks vorgestellt und ihre jeweiligen Vorteile sowie Einsatzmöglichkeiten erläutert.

### 3.4.1 Vue

Vue.js ist ein progressives JavaScript-Framework, das speziell für die Erstellung interaktiver Benutzeroberflächen und Single-Page-Anwendungen (SPA) entwickelt wurde. Es basiert auf standardisierten Webtechnologien wie HTML, CSS und JavaScript, erweitert diese jedoch um zusätzliche Kernkonzepte, die die Entwicklung moderner Webanwendungen erheblich erleichtern.

Eines dieser Kernkonzepte ist das deklarative Rendering. Vue.js verwendet eine HTML-basierte Template-Syntax, die es ermöglicht das gerenderte DOM deklarativ an die Daten der zugrunde liegenden Komponenteninstanz zu binden. Alle Vue-Templates sind syntaktisch gültiges HTML und können daher von standardkonformen Browsern und HTML-Parsern problemlos interpretiert werden.

Ein weiteres zentrales Merkmal von Vue.js ist die Reaktivität, durch die das DOM automatisch aktualisiert wird, sobald Änderungen im JavaScript-Zustand auftreten. Diese Reaktivität wird durch ein System zur Abhängigkeitsverfolgung ermöglicht. Beim erstmaligen Rendern einer Komponente verfolgt Vue.js alle Referenzen, die während des Renderprozesses verwendet werden. Wenn eine dieser Referenzen später geändert wird, löst Vue.js automatisch ein erneutes Rendern der betroffenen Komponenten aus.

Vue-Komponenten durchlaufen bei ihrer Erstellung eine Reihe von Initialisierungsschritten. Diese umfassen unter anderem die Einrichtung der Datenbeobachtung, die Kompilierung des Templates, das Einfügen der Instanz in das DOM und das Aktualisieren des DOM bei Änderungen der Daten. Während dieser Schritte werden Lifecycle Hooks ausgeführt, die es Entwicklern ermöglichen, eigenen Code zu spezifischen Zeitpunkten im Lebenszyklus der Komponente auszuführen. Ein wichtiger Hook in diesem Zusammenhang ist `onMounted`, der unmittelbar nach dem Einfügen der Komponente in das DOM ausgeführt wird. Die komponentenbasierte Architektur von Vue.js ermöglicht nicht nur eine effiziente Verwaltung und Aktualisierung von User Interface (UI)-Elementen, sondern fördert auch die Wiederverwendbarkeit und verbessert die Wartbarkeit des Codes, was den Entwicklungsprozess insgesamt vereinfacht. Dabei spielen Properties eine entscheidende Rolle. Properties sind explizit deklarierte Eigenschaften, über die Komponenten externe Daten entgegennehmen können. Dies ermöglicht eine klare Trennung von Daten und Präsentation und stellt sicher, dass die Kommunikation zwischen den Komponenten auf strukturierte Weise erfolgt.

Vue.js zeichnet sich zudem durch seine Flexibilität aus und kann auf verschiedene Arten genutzt werden. Dazu gehören die Erweiterung von statischem HTML, die Einbettung als Webkomponente auf einer beliebigen Seite sowie die Entwicklung von SPA. Außerdem unterstützt Vue.js Fullstack-Ansätze wie Serverseitiges Rendering (SSR) und Statische Seitengenerierung (SSG), sowie die Ausrichtung auf unterschiedliche Plattformen wie Desktop, Mobile, WebGL und das Terminal (Vue Web Docs, 2024).

Vue.js in der Version 3.3.0 wird in dieser Arbeit hauptsächlich aufgrund seiner komponentenbasierten Struktur verwendet. Durch die damit geschaffene Wiederverwendbarkeit wird der Entwicklungsprozess drastisch vereinfacht und die Wartbarkeit des Codes erheblich verbessert.

### 3.4.2 Ionic

Ionic ist ein Open-Source-Framework für die Entwicklung mobiler Anwendungen, welches mit standardisierten Webtechnologien wie HTML, CSS und JavaScript arbeitet. Außerdem verwendet Ionic moderne Web-APIs wie Custom Elements und Shadow DOM und lässt sich zudem nahtlos in Angular, React oder Vue integrieren.

Mithilfe von Ionic wird nur eine Codebasis benötigt, um plattformübergreifende Anwendungen zu erstellen, die sowohl auf iOS als auch auf Android sowie im Web und auf Desktop-Geräten ausgeführt werden können. Zusätzlich sorgen Adaptive Stylings dafür, dass sich Ionic-Apps an die jeweilige Plattform anpassen und für jedes Gerät eine native Benutzererfahrung geboten wird. Dies wird durch native Funktionalitäten wie modale Fenster, Navigation, Listen und Formulare deutlich, welche mit der Verwendung von nativen SDKs und der Emulation von UI-Richtlinien für native Anwendungen erreicht werden. Darüber hinaus ermöglicht Ionic den Zugriff auf native APIs durch Cordova oder Capacitor, sodass Funktionen wie Kamera, GPS oder Push-Benachrichtigungen nahtlos in Web-Apps integriert werden können.

Ein weiterer zentraler Bestandteil von Ionic ist das Konzept der UI-Komponenten. Das Framework stellt eine umfassende Bibliothek wiederverwendbarer, anpassbarer und leistungsstarker Komponenten zur Verfügung, die speziell für mobile Interfaces optimiert sind (Ionic Web Docs, 2024a).

Ionic ermöglicht zudem die Erstellung von PWAs durch die Unterstützung des Vite PWA Plugins (Fu, 2024). Dieses Plugin automatisiert die Generierung des Service Workers und des Web Application Manifests, die für die PWA-Funktionalitäten unerlässlich sind (Ionic Web Docs, 2024b).

Das Framework wird in der Version 8.0.0 verwendet, insbesondere wegen seiner nahtlosen Integration mit Vue, der Unterstützung für PWAs und der umfassenden plattformübergreifenden Unterstützung sowie leistungsfähigen UI-Komponenten. Diese Eigenschaften beschleunigen die Entwicklung und gewährleisten eine konsistente Benutzeroberfläche über verschiedene Plattformen hinweg.

### 3.4.3 Tailwind CSS

Tailwind ist ein utility-first CSS-Framework, das Entwicklern ermöglicht, vordefinierte Klassen direkt im HTML-Code zu verwenden. Diese Klassen repräsentieren jeweils eine spezifische CSS-Eigenschaft oder Regel. Durch das Anwenden dieser auf HTML-Elemente wird eine visuelle Gestaltung von Anwendungen ermöglicht. So entspricht beispielsweise die Klasse `text-center` der CSS-Eigenschaft zur Zentrierung von Text, während `bg-blue-500` eine bestimmte Hintergrundfarbe festlegt.

Darüber hinaus bietet Tailwind die Möglichkeit einer einfachen Konfiguration und Erweiterung. Mit der Konfigurationsdatei können Entwickler benutzerdefinierte Themen, Farben, Abstände und weitere Designvariablen festlegen, was die konsistente Gestaltung über das gesamte Projekt hinweg erleichtert (Tailwind Web Docs, 2024b).

In diesem Projekt wird Tailwind in Verbindung mit PostCSS in der Version 3.4.4 eingesetzt. Tailwind CSS durchsucht alle HTML-Dateien und JavaScript-Komponenten, generiert die entsprechenden Stile und schreibt diese in eine statische CSS-Datei. Dies erfolgt durch einen Watcher, der zur Laufzeit Änderungen im Projekt erkennt und diese unmittelbar umsetzt (Tailwind Web Docs, 2024a).

### 3.4.4 Media Pipe: Model Maker

Das Model Maker-Tool ist ein Teilbereich der Media Pipe Solutions, eine Sammlung von Open-Source-Bibliotheken und Tools, die von Google entwickelt wurden. Dieses Tool wird zur Anpassung von vortrainierten maschinellen Lernmodellen durch TL verwendet und ermöglicht es, bereits existierende Modelle effizient an neue Aufgaben anzupassen. Dieser Vorgang lässt sich in Abbildung 3.6 erkennen. Die Klassifikationsebene eines existierenden maschinellen Lernmodells wird durch spezialisierte Ebenen ersetzt. Diese Ebenen werden mit den bereitgestellten Daten neu trainiert, um das Modell auf die spezifischen Anforderungen der Anwendung abzustimmen (Google AI, 2024b). Da ein erheblicher Teil der vorhandenen Modelllogik wiederverwendet wird und mit weniger Daten durchgeführt werden kann, nimmt das Training weniger Zeit in Anspruch als das Training eines neuen Modells (Google AI, 2024c).

Für das Trainieren des Modells wird ein Datensatz benötigt, der die zu erkennenden Objekte enthält. Dieser Datensatz kann deutlich kleiner sein als der, der für das Training eines neuen Modells erforderlich wäre (Google AI, 2024b). Die Dokumentation empfiehlt die Verwendung von etwa 100 Datenpunkten für jede zu trainierende Klasse (Google AI, 2024c). Dabei wird in der Praxis häufig eine 80/10/10-Aufteilung angewendet, bei der 80 % der Daten zum Trainieren des Modells, 10 % zur Validierung und 10 % zum Testen genutzt werden (Raheja, 2024).

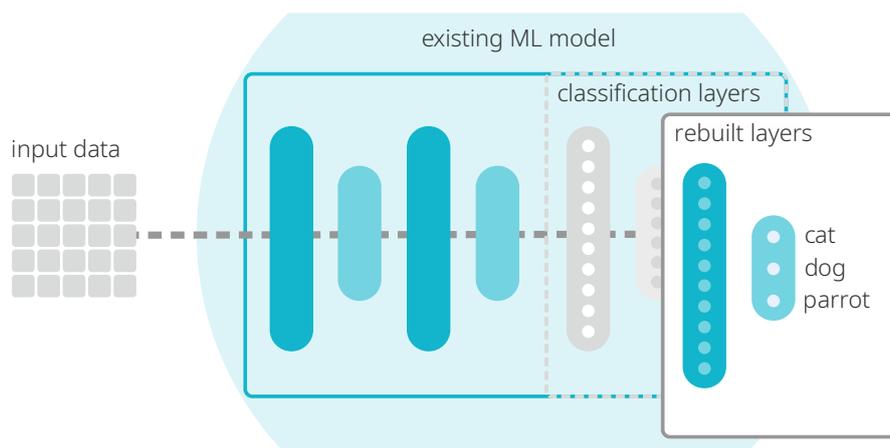


Abbildung 3.6: Ersetzen der Klassifikationsebene durch spezialisierte Ebenen mithilfe des Model Maker-Tools (Quelle: Google AI, 2024c)

MediaPipe Model Maker unterstützt das Lesen von Datensatzformaten wie dem COCO-Format und dem PASCAL-VOC-Format. Im COCO-Format werden alle Bilder in einem Datenverzeichnis gespeichert, während die Anmerkungen in einer einzigen JSON-Datei zusammengefasst sind. Im PASCAL-VOC-Format werden die Bilder ebenfalls in einem Datenverzeichnis abgelegt, jedoch sind die Anmerkungen für jedes Bild in separaten XML-Dateien im Verzeichnis Annotations organisiert.

Der Nutzer kann verschiedene Trainingsparameter festlegen, einschließlich der Modellarchitektur. Die MobileNetV2-Architektur dient als Grundlage für die verfügbaren Varianten, die jeweils auf unterschiedliche Leistungsanforderungen und Anwendungsgebiete abgestimmt sind. Die Wahl der Variante basiert auf der Eingangsbildgröße und der verwendeten Methode zur Merkmalsextraktion. Weitere anpassbare Trainingsparameter umfassen die Lernrate, die Batch-Größe und die Anzahl der Epochen (Google AI, 2024b).

Zur Evaluierung der Leistung von Objekterkennungsmodellen werden COCO-Metriken verwendet, da sie eine umfassende Analyse ermöglichen. Eine zentrale Metrik ist die Average Precision, die die Präzision des Modells bei verschiedenen Konfidenzschwellen bewertet. Sie misst, wie genau die Vorhersagen des Modells sind, indem sie den Anteil der korrekten Erkennungen im Verhältnis zu allen getätigten Vorhersagen betrachtet. Eine spezifische Variante des Präzisionswertes ist AP50, die angibt, wie gut die Vorhersagen des Modells mit der tatsächlichen Position eines Objekts übereinstimmen müssen. Hierbei muss die Übereinstimmung zwischen Vorhersage und tatsächlicher Position des Objekts mindestens 50 % betragen, um als korrekt zu gelten.

Ein weiteres bedeutendes Kriterium ist der Average Recall, der misst, wie viele der tatsächlich vorhandenen Objekte vom Modell korrekt erkannt wurden. Der Average Recall gibt Aufschluss darüber, wie vollständig das Modell arbeitet und ob es alle relevanten Objekte im Bild identifizieren kann (COCO Consortium, 2024).

Abschließend ist der Validierungsverlust eine wesentliche Metrik, da er die Generalisierungsfähigkeit eines Modells auf bisher nicht gesehenen Daten bewertet. Ein kontinuierlicher Rückgang des Validierungsverlusts deutet darauf hin, dass das Modell effektiv auf neue Daten anwendbar ist, ohne signifikante Überanpassung zu zeigen (vom Lehn, 2023).

Das Model Maker-Tool in der Version 0.10.14 wird in dieser Arbeit eingesetzt, um ein vortrainiertes Modell an einen spezifischen Anwendungsfall anzupassen.

### **3.4.5 Media Pipe: Object Detector**

Die Object Detector-Bibliothek ist ebenfalls ein Bestandteil der Media Pipe Solutions. Diese Bibliothek ermöglicht es, verschiedene Arten von Objekten in Bildern oder Videos zu erkennen und ihre genaue Position zu bestimmen. Dafür wird ein maschinelles Lernmodell verwendet, das sowohl Einzelbilder als auch kontinuierliche Videostreams als Eingaben akzeptiert. Die Ausgabe des Objekterkennungsmodells besteht aus einer Liste von Erkennungsergebnissen, wobei jedes Ergebnis ein erkanntes Objekt im Bild oder Video repräsentiert. Für jedes identifizierte Objekt liefert die API des Objekterkennungsmodells bestimmte Informationen. Hierzu zählen die Objektkategorie, die die Klassifizierung des erkannten Objekts angibt, der Wahrscheinlichkeitswert, der die Sicherheit der Zuordnung des Objekts zu einer bestimmten Kategorie darstellt, sowie die Koordinaten des Begrenzungsrahmens, auch als Bounding Box bezeichnet. Die Bounding Box wird durch die Werte `x-Origin` und `y-Origin` definiert, die die Position der oberen linken Ecke des Rechtecks im Bild angeben, sowie durch `Width` und `Height`, die die Größe des Rechtecks bestimmen.

Nutzer können entweder eigene Modelle verwenden oder auf vortrainierte Modelle von Media Pipe zurückgreifen. Ein Beispiel für ein vortrainiertes Modell ist das `EfficientDet-Lite0` Modell, das mit dem COCO-Datensatz trainiert wurde und ein ausgewogenes Verhältnis zwischen Latenz und Genauigkeit bietet. Eine weitere Option ist das `SSD MobileNetV2` Modell, das für seine Geschwindigkeit und Leichtigkeit bekannt ist, jedoch eine allgemein geringere Genauigkeit als die `EfficientDet`-Modelle bietet. Die genannten Modelle sind in verschiedenen Varianten verfügbar, um unterschiedlichen Anforderungen gerecht zu werden (Google AI, 2024a).

Die Bibliothek Object Detector in der Version 0.10.14 wird in dieser Arbeit eingesetzt, um das selbst trainierte Modell für die Nutzung in einer Webanwendung bereitzustellen.

## 3.5 Entwicklungsplattformen

In der Softwareentwicklung spielen spezialisierte Plattformen eine zentrale Rolle bei der Umsetzung verschiedener Entwicklungsaufgaben. Im Folgenden werden wichtige Werkzeuge zur Datenannotation und zum Deployment vorgestellt. Zunächst wird Label Studio im Abschnitt 3.5.1 behandelt, gefolgt von Firebase im Abschnitt 3.5.2.

### 3.5.1 Datenannotation: Label Studio

Label Studio ist eine Open-Source-Plattform, die speziell für die Datenannotation in verschiedenen Kontexten des maschinellen Lernens entwickelt wurde. Es bietet ein umfassendes Interface zur Unterstützung unterschiedlichster Trainingsmethoden. Dazu zählen unter anderem überwachte Lernverfahren, bei denen Modelle auf Basis gelabelter Daten trainiert werden. Die Plattform unterstützt die Annotation verschiedenster Datentypen, darunter Bilder, Audiodateien, Zeitreihen, Multidomain-Daten und Videos. Im Bereich der Computer Vision ermöglicht Label Studio unter anderem das Trainieren von Modellen für Bildklassifikationen, bei denen Bilder in vordefinierte Kategorien eingeordnet werden, sowie für Objekterkennung, bei der Objekte in Bildern identifiziert und durch Bounding Boxen markiert werden. Darüber hinaus unterstützt die Plattform die semantische Segmentierung, bei der Bilder in mehrere Segmente unterteilt werden (Label Studio Docs, 2024a). Die Plattform lässt sich nahtlos in bestehende Workflows einbinden und unterstützt die Integration in ML/KI-Pipelines über Webhooks, ein Python SDK und eine API. Zudem bietet Label Studio die Möglichkeit, ML-unterstütztes Labeling zu nutzen, bei dem Vorhersagen zur Unterstützung der Datenannotation verwendet werden können. Label Studio kann über die Containerisierungsplattform Docker installiert werden und bietet ein Interface, das über einen Webbrowser auf dem lokalen Rechner zugänglich ist (Label Studio Docs, 2024b).

In der vorliegenden Arbeit wurde Label-Studio in der Version 1.13.1 eingesetzt, um Bilder für die Objekterkennung zu annotieren.

### 3.5.2 Deployment: Firebase

Deployment bezeichnet den Prozess, durch den eine Anwendung von der Entwicklungsumgebung in eine Produktionsumgebung überführt wird, um sie für Endbenutzer zugänglich zu machen. Für die vorliegende Anwendung erfolgt die Veröffentlichung als PWA, wofür ein Service Worker und ein Web Application Manifest erforderlich sind. Diese werden mit dem Vite PWA Plugin automatisch während des Build-Prozesses generiert.

Im Kontext von Ionic gibt es verschiedene Optionen für das Deployment einer PWA, wobei Firebase eine der bevorzugten Plattformen darstellt (Ionic Web Docs, 2024b). Firebase ist eine umfassende Entwicklungsplattform von Google, die eine Vielzahl von Diensten für die Entwicklung und Bereitstellung von Web- und Mobilanwendungen bietet. Das Firebase-Hosting zeichnet sich durch Vorteile wie schnelle Antwortzeiten durch Content Delivery Networks (CDNs), standardmäßig aktivem HTTPS und Unterstützung für HTTP2-Push aus (Ionic Web Docs, 2024b). Neben dem Hosting bietet Firebase auch weitere Funktionen wie Datenbankverwaltung, Authentifizierung, Cloud-Funktionen und Analytik (Firebase Developers, 2024).

# 4 Umsetzung

Im folgenden Kapitel wird die Umsetzung der App basierend auf dem vorgestellten Konzept in Kapitel 3 schrittweise erläutert. Dabei werden zunächst die Entwicklung des Objekterkennungsmodells unter Verwendung von MediaPipe Model Maker beschrieben und anschließend die Implementierung der PWA unter Einsatz der Frameworks Ionic und Vue behandelt.

## 4.1 Entwicklung des Modells

In der App findet ein selbst trainiertes Objekterkennungsmodell Anwendung. Die Funktionalität dieses Modells ist essenziell, da es die Grundlage für eine präzise und zuverlässige Erkennung der verschiedenen Fruchtsorten bildet. Um sicherzustellen, dass die Anwendung in realen Anwendungsszenarien optimal funktioniert, ist eine sorgfältige und systematische Vorgehensweise zur Entwicklung des Modells von entscheidender Bedeutung. Daher wird in diesem Kapitel der gesamte Entwicklungsprozess des Modells detailliert behandelt, angefangen bei der Datensammlung und -aufbereitung. Abschließend wird das Training und die Feinabstimmung des Modells umfassend erläutert.

### 4.1.1 Datensammlung

Der erste Schritt im Prozess der Modellentwicklung besteht in der Datensammlung. Da bei der Verwendung des MediaPipe Model Makers TL zum Einsatz kommt, ist für jede Klasse nur eine begrenzte Anzahl von Bildern erforderlich. In diesem Fall reichen etwa 100 Bilder pro Klasse aus.

Zunächst wurde in Erwägung gezogen, vorhandene Datenbanken wie Kaggle zu nutzen, die eine große Auswahl an Bildmaterialien bieten (Kaggle, 2024). Allerdings unterliegen die meisten verfügbaren Bilder Einschränkungen, die ihre Eignung für das Projekt beeinträchtigen. Viele der Bilder weisen einen weißen Hintergrund auf, der für die Erkennung in realen Anwendungsszenarien ungeeignet ist, wie in Abbildung 4.1a dargestellt. Einige Bilder sind mit Wasserzeichen versehen, was ihre Nutzung für Trainingszwecke einschränkt. In Abbildung 4.1b werden mehrere Äpfel abgebildet, die nicht vollständig auf dem Bild zu

sehen sind. Solche Sonderfälle können ebenfalls zu Fehlern bei der Objekterkennung führen und sind daher ungeeignet für das Training eines Modells. Zudem zeigen einige Bilder die Objekte in ungeeigneten Umgebungen, wie zum Beispiel im Wald, oder bei unzureichender Beleuchtung, was nicht zu den realistischen Anwendungsszenarien der Anwendung passt. Dies wird in Abbildung 4.1c verdeutlicht. Schließlich enthalten einige der Bilder Objekte, die in verschiedenen Unterkategorien dargestellt werden, wie etwa Äpfel in unterschiedlichen Farben. Diese Differenzierung kann dazu führen, dass der Agent diese Unterkategorien als separate Klassen interpretiert, was die Genauigkeit der Klassifizierung beeinträchtigen und zu einer erhöhten Fehleranfälligkeit in der Objekterkennung führen kann.



(a) Bild eines Apfels mit weißem Hintergrund



(b) Bild mehrerer angeschnittener Äpfel



(c) Bild eines Apfels mit unzureichender Beleuchtung

Abbildung 4.1: Beispiele für problematische Datensätze (Quelle: Minut, 2024)

Aufgrund dieser Einschränkungen wurde entschieden, eigene Bilder zu sammeln, die gezielt für den spezifischen Anwendungsfall optimiert sind. Diese maßgeschneiderte Datensammlung konzentriert sich auf mehrere wesentliche Aspekte, um die Relevanz und Präzision der Daten zu gewährleisten.

Die Abbildungen 4.2 a bis c zeigen Bilder, die in verschiedenen alltäglichen Umgebungen aufgenommen wurden. Dies dient dazu, eine breite Variabilität an Hintergründen zu bieten. Dazu gehören ein Holzfußboden, ein weißer Tisch und eine Couch. Diese unterschiedlichen Hintergründe sollen sicherstellen, dass das Modell in der Lage ist, die Früchte unter verschiedenen Bedingungen zu erkennen und korrekt zu klassifizieren. Ein weiterer wichtiger Aspekt ist die Einschränkung der Unterkategorien. Um die Komplexität der Klassifikation zu reduzieren und die Genauigkeit zu erhöhen, wurden bei den Äpfeln zum Beispiel nur rot-gelbe Äpfel der Sorte Pink Lady berücksichtigt. Eine weitere Differenzierung in Unterkategorien wurde vermieden. Außerdem wurden die Bilder mit der Kamera des Handymodells OnePlus Nord AC2003 im Hochformat aufgenommen. Dies ist besonders von Vorteil, da die Anwendung für

den Einsatz mit einer Handykamera konzipiert ist und somit die Trainingsdaten die tatsächlichen Einsatzbedingungen besser widerspiegeln. Diese Entscheidungen gewährleisten, dass die Bilder den realen Bedingungen entsprechen und die Klassifikation der Früchte präzise und zuverlässig durchgeführt wird.



(a) Bild eines Apfels auf einem weißen Tisch



(b) Bild eines Apfels auf einer Couch



(c) Bild eines Apfels auf einem Holzfußboden

Abbildung 4.2: Beispiele für eigene Datensätze (Quelle: Eigene Darstellung)

### 4.1.2 Datenaufbereitung

Nach der Erfassung der Rohdaten beginnt die Phase der Datenaufbereitung, die mehrere wesentliche Schritte umfasst. Zunächst werden die gesammelten Bilder sortiert, indem sie nach spezifischen Kategorien geordnet werden. Dies erleichtert die anschließende Annotation und gewährleistet, dass das Training auf einer sauberen und strukturierten Datenbasis erfolgt. Im Anschluss an die Sortierung werden die Bilder überprüft und aussortiert, um nur qualitativ hochwertige und relevante Bilder für das Training zu verwenden. Bilder, die beispielsweise unscharf, falsch belichtet oder nicht repräsentativ für die jeweilige Kategorie sind, werden aus dem Datensatz entfernt. Abbildung 4.3 stellt die manuelle Datenannotation mit dem Tool Label Studio dar, bei der die Bilder mit den entsprechenden Labels versehen werden. Neben den korrekten Labels müssen für jedes Bild Begrenzungsrahmen um die relevanten Objekte gezogen werden. Im nächsten Schritt erfolgt der Export des gelabelten Datensatzes im COCO-Format. Dieser Export bringt das Dataset in eine strukturierte Form, die unmittelbar für das Training des Modells verwendet werden kann.

### 4.1.3 Modelltraining

Nachdem alle notwendigen Bibliotheken und erforderlichen Python-Klassen in die Google Colab Umgebung importiert wurden, wird das vorbereitete Dataset in die Umgebung geladen.



Abbildung 4.3: Annotation eines Apfels in Label Studio (Quelle: Eigene Darstellung)

Das geladene Dataset wird in Trainings-, Validierungs- und Testdaten aufgeteilt. 80 % der Daten werden für das Training des Modells verwendet, 10 % werden für die Validierung reserviert, und die verbleibenden 10 % werden als Testdaten genutzt. Diese Aufteilung stellt sicher, dass das Modell auf einem separaten Teil der Daten validiert und getestet wird, um seine Leistung unabhängig vom Training zu überprüfen.

Vor der Durchführung des eigentlichen Trainingsprozesses mit dem Model Maker-Tool müssen die Retraining-Parameter konfiguriert werden. Dies wird im Codeblock 4.1 veranschaulicht, wobei das MobileNetV2-Modell als zugrunde liegende Architektur verwendet wird. Gleichzeitig erfolgt eine Anpassung der Hyperparameter, darunter die Lernrate, die Batch-Größe und die Anzahl der Epochen. Im vorliegenden Beispiel sind diese Parameter auf eine Lernrate von 0,3, eine Batch-Größe von 8 und 30 Epochen festgelegt.

#### Codeblock 4.1: Retrainingsoptionen

```
1 spec = object_detector.SupportedModels.MOBILENET_MULTI_AVG_I384
2 hparams = object_detector.HParams(export_dir='exported_model',
3     learning_rate=0.3, batch_size=8, epochs=30)
4 options = object_detector.ObjectDetectorOptions(
5     supported_model=spec,
6     hparams=hparams
7 )
```

Nachdem die Trainingsparameter festgelegt wurden, beginnt das Retraining des Modells. Während dieses Prozesses wird das Modell iterativ angepasst, um eine optimale Genauigkeit zu erreichen. Dieser Schritt ist entscheidend für die finale Leistungsfähigkeit des Modells und erfordert unter Umständen mehrere Durchläufe mit unterschiedlichen Parametereinstellungen, um das bestmögliche Ergebnis zu erzielen.

#### 4.1.4 Modellvalidierung

Nach Abschluss des Trainings wird das Modell anhand des separaten Testdatensatzes validiert. Der Codeblock 4.2 veranschaulicht, wie die Leistung des trainierten Modells mit der Funktion `evaluate` überprüft werden kann. Dabei liegt der Fokus auf den Metriken für den Validierungsverlust (Validation Loss) sowie den COCO-Metriken (Validation COCO Metrics).

Codeblock 4.2: Evaluierung

```
1     loss, coco_metrics = model.evaluate(test_data, batch_size=4)
2     print(f"Validation Loss: {loss}")
3     print(f"Validation COCO Metrics: {coco_metrics}")
```

Die Validierungsergebnisse zeigen einen niedrigen Validierungsverlust von 0,01. Zudem wurden hohe COCO-Metriken erreicht, darunter ein durchschnittlicher Präzisionswert von 0,91, eine durchschnittliche Präzision bei 50 % Intersection over Union (AP50) von 0,99 und ein Average Recall von 0,94. Der niedrige Validierungsverlust weist auf eine ausgezeichnete Generalisierungsfähigkeit des Modells hin, da es auf neue Daten ohne signifikante Überanpassung effektiv anwendbar ist. Außerdem belegt der hohe durchschnittliche Präzisionswert die hohe Genauigkeit der Modellvorhersagen, während der AP50-Wert zeigt, dass die Vorhersagen des Modells zu

98,74 % mit den tatsächlichen Objektpositionen übereinstimmen. Darüber hinaus zeigt der Average Recall, dass das Modell 94,20 % der tatsächlich vorhandenen Objekte korrekt erkennt, was auf eine umfassende Erkennungseffizienz hinweist.

Im Verlauf des Trainingsprozesses wurden verschiedene Parameterkonfigurationen getestet, um die Leistung des Objekterkennungsmodells weiter zu optimieren und die Auswirkungen unterschiedlicher Einstellungen zu evaluieren. Beispielsweise führte eine Batch-Größe von 32 zu einem hohen Validierungsverlust mit Werten von bis zu 13,79 und suboptimalen COCO-Metriken, darunter ein durchschnittlicher Präzisionswert von 0,01. Im Gegensatz dazu erzielte eine Lernrate von 0,01 zusammen mit einer Batch-Größe von 32 verbesserte Ergebnisse, einschließlich eines reduzierten Validierungsverlusts und COCO-Metriken wie einem durchschnittlichen Präzisionswert von 0,22. Besonders effektiv war jedoch eine Lernrate von 0,01 in Kombination mit einer Batch-Größe von 8, die zu einem Validierungsverlust von

0,42 und einem durchschnittlichen Präzisionswert von 0,85 führte. Allerdings erwies sich die im Abschnitt 4.1.3 erwähnte Konfiguration als die leistungsfähigste, da sie zu den besten Ergebnissen in Bezug auf Genauigkeit und Generalisierungsfähigkeit führte.

### **4.1.5 Modellbereitstellung**

Nach der Erstellung des Modells wird es in das TensorFlow Lite (TFLite)-Format konvertiert und exportiert, um die spätere Integration in die Anwendung zu ermöglichen. Das TFLite-Format ist besonders für den Einsatz in Webanwendungen optimiert, da es eine effiziente Ausführung auf mobilen Geräten durch seine reduzierte Modellgröße ermöglicht. Der Export umfasst zudem die Metadaten des Modells, einschließlich der Zuordnung der Labels, die für die Klassifizierung der Ergebnisse erforderlich sind.

## **4.2 Entwicklung der Progressive Web App**

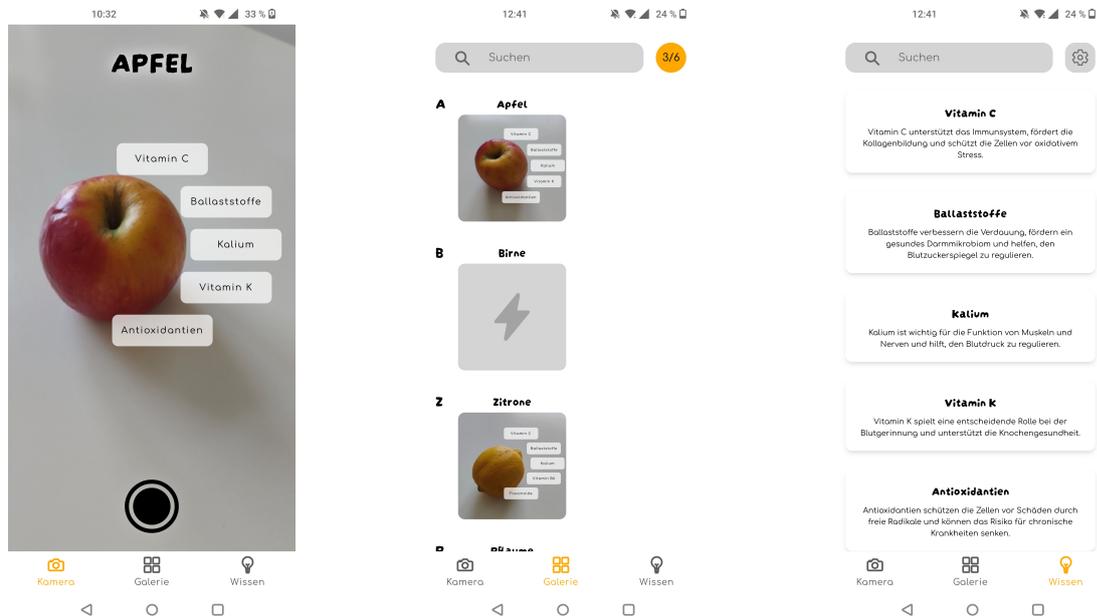
Die Anwendung wird als PWA bereitgestellt. Das in Abschnitt 3.2 vorgestellte Design wird mithilfe der Frameworks Ionic und Vue in IntelliJ IDEA Ultimate umgesetzt. Im Abschnitt 4.2.1 wird zunächst die Benutzeroberfläche der App detailliert beschrieben. Anschließend erfolgt im Abschnitt 4.2.2 eine eingehende Betrachtung der verwendeten Composables, welche einen wesentlichen Teil der Anwendungslogik bereitstellen. Schließlich wird im Abschnitt 4.2.3 der Prozess des Deployments erläutert.

### **4.2.1 Benutzeroberfläche**

In der Hauptnavigation der Anwendung wurde eine Tab-Navigation implementiert, die durch eine zentrale Tab-Komponente realisiert wird. Diese Navigation umfasst drei Registerkarten, welche dem zuvor definierten Design in Abschnitt 3.2 entsprechen. Diese Komponente dient in der PWA als zentrales Navigationselement, welches den Nutzern ermöglicht, intuitiv zwischen verschiedenen Hauptbereichen der Anwendung zu navigieren. Jede Registerkarte repräsentiert dabei einen eigenständigen Anwendungsbereich, der durch ein Symbol und eine Beschriftung visualisiert wird. Um eine funktionale Navigation innerhalb der Anwendung zu gewährleisten, ist die Definition eines Routers erforderlich. Dieser Router legt die notwendigen Routen fest, die die Zuordnung zwischen den vom Benutzer gewählten Pfaden und den entsprechenden Seiten der Anwendung bestimmen. Der Router wird in der gesamten Anwendung eingesetzt, um eine konsistente Navigation zwischen den verschiedenen Bereichen sicherzustellen. In Kombination mit dem Router wird gewährleistet, dass die entsprechenden Seiten geladen werden, sobald ein Benutzer eine Registerkarte auswählt. Der Router steuert

die Navigation zu den spezifischen Routen, die den einzelnen Tabs zugeordnet sind, und stellt sicher, dass die korrekte Ansicht im Inhaltsbereich der Anwendung angezeigt wird.

Abbildung 4.4 zeigt die drei Hauptbereiche der Anwendung. Diese umfassen die Seiten CameraPage, GalleryPage und KnowHowPage. Diese Seiten sind über die zuvor beschriebenen Tab-Komponenten zugänglich und kapseln spezifische Funktionalitäten innerhalb der PWA.



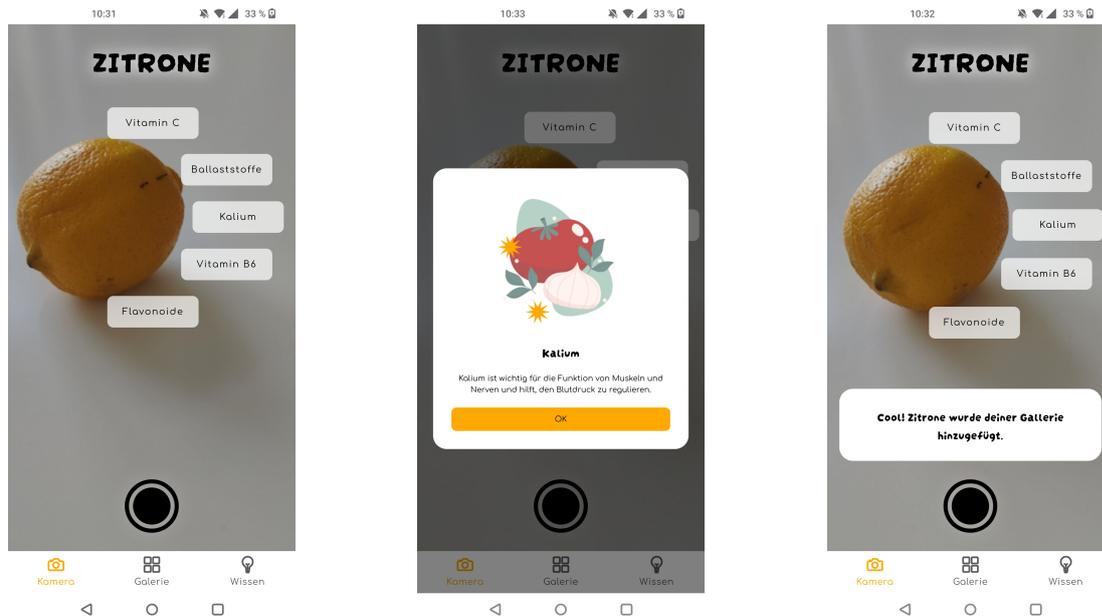
(a) Ansicht der CameraPage

(b) Ansicht der GalleryPage

(c) Ansicht der KnowHowPage

Abbildung 4.4: Hauptbereiche der Anwendung (Quelle: Eigene Darstellung)

Im Folgenden wird die Funktionalität der CameraPage chronologisch erläutert, da diese eine zentrale Rolle innerhalb der Anwendung einnimmt. Abbildung 4.5 veranschaulicht die wichtigsten Funktionalitäten. Die Verwendung des Medienstreams der Kamera wird in Abbildung 4.5a deutlich. Dieser wird in diesem Beispiel zur Erkennung einer Zitrone und zur Echtzeitanzeige relevanter Nährstoffe über den Video-Stream genutzt. Interagiert der Nutzer mit einem der dargestellten Nährstoffe, wird ein Informationsmodal angezeigt. Dies wird in Abbildung 4.5b visualisiert. Zuletzt lässt sich der aktuelle Frame des Medienstreams als Bild speichern. Dafür wird der Auslöser im unteren Bereich des Bildes betätigt. Abbildung 4.5c zeigt eine Benachrichtigung für den Nutzer nach einer erfolgreichen Interaktion.



(a) Erkennung einer Zitrone mit der Kamera      (b) Anzeige des Informationsmodal      (c) Anzeige des Benachrichtigungsmodals

Abbildung 4.5: Funktionalitäten der CameraPage (Quelle: Eigene Darstellung)

Das Template der CameraPage besteht aus einer strukturierten Anordnung von HTML-Elementen und Komponenten, die innerhalb der PWA implementiert sind. Die äußeren Ionic-Komponenten, wie `ion-page` und `ion-content`, sind für das Routing sowie die Integration der PWA-Features verantwortlich. Außerdem enthält dieses Template mehrere modale Fenster, darunter `informationModal`, `failureModal`, `loadingModal` und `sliderModal`. Diese Modale werden erwähnt, ohne dass eine detaillierte Betrachtung ihrer Funktionalität erfolgt, da sie nicht wesentlich zur Kernfunktionalität der Seite beitragen. Ein zentrales Element im Template ist das `video` HTML-Element, das mit den Tailwind-Utility-Klassen `h-full`, `w-full` und `object-cover` versehen ist. Diese Klassen stellen sicher, dass das Video im Vollbildmodus dargestellt wird, den gesamten verfügbaren Platz ausnutzt und die Ansicht ohne Verzerrungen vollständig abdeckt.

Innerhalb der PWA werden zahlreiche wiederverwendbare Komponenten integriert, darunter die `OverlayComponent`. Diese Komponente findet sowohl auf der `GalleryPage` als auch auf der `FullScreenImagePage` Anwendung. Sie zeichnet sich durch ihre Flexibilität aus und kann in verschiedenen Varianten konfiguriert werden. Die Varianten werden durch das Property `variant` gesteuert, das die Werte `camera`, `full` oder `preview` annehmen kann. Je nach übergebenem Wert wird das Markup der Komponente unterschiedlich gerendert, um die jeweils gewünschte Darstellungsform zu realisieren. Im Codeblock 4.3 wird dargestellt, wie die Variante `camera` als Property übergeben wird. Dadurch wird die Komponente als Overlay

in absoluter Positionierung über dem Video gerendert.

Codeblock 4.3: Anwendung der Komponente `OverlayComponent` in der Variante `camera`

```
1     <OverlayComponent
2         variant="camera"
3         <!-- Weitere Properties und Events -->
4     </OverlayComponent>
```

Nachdem die Komponente vollständig in den DOM eingefügt und gerendert wurde, wird der `onMounted`-Hook ausgeführt. Diese Ausführung erfolgt asynchron, um sicherzustellen, dass asynchrone Operationen effizient abgewickelt werden, ohne die Benutzeroberfläche zu blockieren. Im Rahmen dieser asynchronen Ausführung wird die Funktion `initializeObjectDetector` aufgerufen, die den Objektdetektor konfiguriert und initialisiert. Diese Funktion ist Teil des Composables `useMediaPipe`, die eine nahtlose Integration der `MediaPipe`-Technologie in die Anwendung ermöglicht. Eine detaillierte Erläuterung der Vorteile und Anwendungen von Composables erfolgt im Abschnitt 4.2.2. Zusätzlich wird die Methode `getUserMedia` der Schnittstelle `MediaDevices` aufgerufen. Die Ausführung der Funktion wird im Codeblock 4.4 dargestellt. Diese Methode ermöglicht den Zugriff auf die Kamera des Endgeräts.

Codeblock 4.4: Aufruf der `getUserMedia`-Funktion

```
1     await navigator.mediaDevices
2     .getUserMedia({
3         video: {
4             <!-- Anforderungen an den Medientyp video -->
5         }
6     })
7     <!-- Weitere Funktionalitäten -->
```

Sobald der Benutzer seine Einwilligung erteilt hat, gibt die Methode ein Promise zurück. Im Codeblock 4.5 wird der `then`-Befehl verwendet, um auf die erfolgreiche Auflösung des Promises zu reagieren. Dieser Befehl ermöglicht es, das `MediaStream`-Objekt dem `srcObject` des `video`-Elements zuzuweisen, wodurch der Live-Stream der Kamera auf der Seite angezeigt wird.

Codeblock 4.5: Zuweisung des `MediaStream`-Objekts zum `video`-Element

```
1     <!-- Weitere Funktionalitäten -->
2     .then(async (localMediaStream) => {
3         <!-- Funktionalitäten der modalen Fenster -->
4         video.value.srcObject = localMediaStream;
```

```

5         <!-- Funktionalitäten nach geladenen Metadaten -->
6     })
7     <!-- Weitere Funktionalitäten -->

```

Sollte der Zugriff auf die Kamera verweigert werden oder ein anderer Fehler auftreten, wird die `catch`-Methode ausgeführt, um auf das Fehlschlagen des Promises zu reagieren. Wie im Codeblock 4.6 dargestellt, wird in diesem Fall ein Modal mit der entsprechenden Fehlermeldung angezeigt.

#### Codeblock 4.6: Fehlerbehandlung der `getUserMedia`-Funktion

```

1     <!-- Weitere Funktionalitäten -->
2     .catch((error) => {
3         informationModal.value.showFailureModal(error);
4     });

```

Sobald die Metadaten des Videos vollständig geladen sind, wird die asynchrone Funktion `handlePrediction` des Composables `useMediaPipe` aufgerufen. Die Verwendung von `await` gewährleistet, dass die Ausführung des nachfolgenden Codes angehalten wird, bis die Funktion `handlePrediction` ihre Ausführung abgeschlossen hat. Diese Funktion erhält den aktuellen Frame des Medienstreams als Eingabeparameter und veranlasst die Objekterkennung auf diesem Frame. Als Rückgabewert wird eine neue Instanz der Klasse `Prediction` erstellt, die mit den Ergebnissen der Objekterkennung initialisiert wird. Die Klasse `Prediction` strukturiert die von `MediaPipe` gelieferten Daten und erleichtert die weitere Verarbeitung in der Anwendung. Diese Instanz enthält detaillierte Informationen über die Klassifikation des erkannten Objekts, die Bounding Box, eine Liste der Nährstoffe sowie die Wahrscheinlichkeit der Vorhersage.

Um eine kontinuierliche Aktualisierung der Vorhersagen zu gewährleisten, wird die Funktion `handlePrediction` mithilfe der Methode `requestAnimationFrame` rekursiv aufgerufen. Diese Methode ermöglicht eine regelmäßige und reibungslose Durchführung der Bildverarbeitung, indem sie die Funktion bei jedem Frame-Rendering des Bildschirms erneut ausführt. Codeblock 4.7 veranschaulicht diese Vorgehensweise.

#### Codeblock 4.7: Rekursives Aufrufen der `handlePrediction`-Funktion

```

1 const handlePrediction = async () => {
2     <!-- Weitere Funktionalitäten -->
3     requestAnimationFrame(() => {
4         handlePrediction();
5     });
6 };

```

Die reaktive Variable speichert die zuvor als Rückgabewert erstellte Instanz der Klasse Prediction und stellt deren Eigenschaften zur Verfügung. Diese Daten werden an die OverlayComponent übergeben, die für die Darstellung des Kamera-Overlays zuständig ist. Das Klassifikationslabel wird innerhalb der Vue-Komponente verwendet und durch Interpolation mittels der Mustache-Syntax `{{variable}}` im Template dynamisch in der Benutzeroberfläche dargestellt.

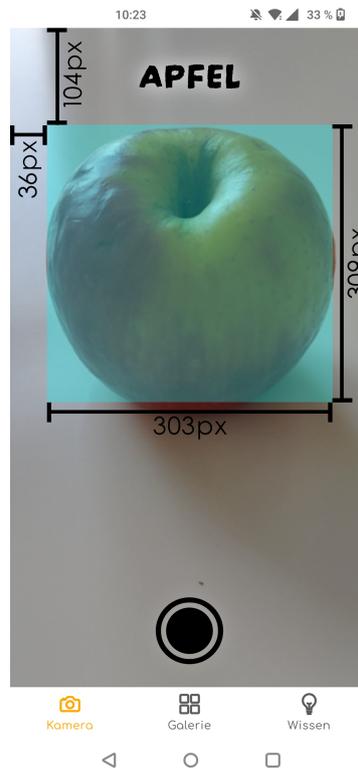


Abbildung 4.6: Abbildung einer farblich markierten Bounding Box  
(Quelle: Eigene Darstellung)

Die Bounding Box ist absolut positioniert, sodass die Position und Größe dieses Elementes mithilfe der Werte `left`, `top`, `width` und `height` definiert werden kann. Durch den Einsatz der Reaktivität werden die Werte der Bounding Box direkt aus der reaktiven Variable abgeleitet und zur Laufzeit basierend auf den Vorhersagen dynamisch aktualisiert. Abbildung 4.6 zeigt die farblich markierte Bounding Box eines erkannten Apfels. Dabei liegt der Wert der horizontalen Verschiebung bei 36px, der vertikalen Verschiebung bei 104px, während die Höhe und Breite 309px beziehungsweise 303px betragen.

Die Liste der erkannten Nährstoffe wird durch die Komponente `FactListItemComponent` verarbeitet und visualisiert. Dabei erfolgt die Darstellung der Liste durch eine iterative Verarbeitung der einzelnen Einträge. Die Positionierung der Fakten basiert auf geometrischen Transformationen, die auf elliptischen Berechnungen basieren. Jeder Fakt wird dabei mit individuellen

translateY- und translateX-Werten versehen, um eine kreisförmige Anordnung um die Bounding Box zu realisieren. Abbildung 4.7 zeigt die Transformation und visuelle Darstellung der Nährstoffe am Beispiel einer Zitrone. Da die Liste der Fakten als Kindelemente der Bounding Box behandelt wird, erfolgt ihre initiale Positionierung mittig in Bezug auf dieses Element. Für den Fakt Vitamin C wurden dabei eine vertikale Verschiebung von -119px und eine horizontale Verschiebung von 68px berechnet.

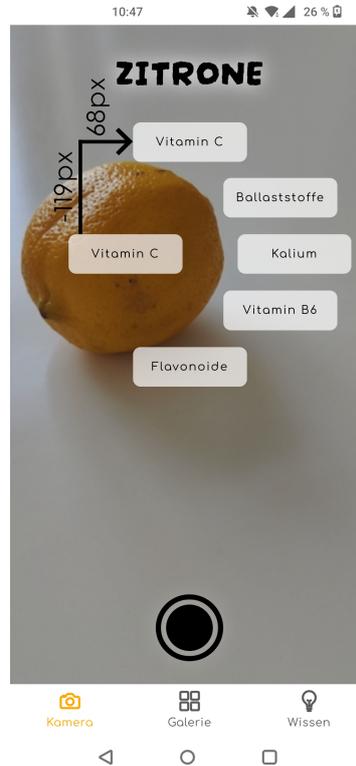


Abbildung 4.7: Abbildung der Liste von Nährstoffen einer Zitrone (Quelle: Eigene Darstellung)

## 4.2.2 Composables

Ein wesentlicher Teil der Anwendungslogik wird über sogenannte Composables bereitgestellt. Diese ermöglichen es, wiederverwendbare Funktionalitäten zu kapseln und stellen somit eine flexible und modulare Struktur für die Entwicklung der Anwendung sicher. Composables können unter anderem dazu verwendet werden, Shared References zu verwalten, spezifische Berechnungen durchzuführen oder externe Frameworks zu integrieren. Die wichtigsten Composables in diesem Projekt sind useFruits, useCalculations und useMediaPipe.

Das useFruits-Composable bietet mehrere Funktionen zur effizienten Verwaltung von Frucht-daten innerhalb der Anwendung. Die Funktion setFruitToLocalStorage ermöglicht es, Frucht-

objekte im Local Storage zu speichern, in dem relevante Attribute wie der Name der Frucht, das gespeicherte Bild, die Bounding Box und weitere Zusatzinformationen enthalten sind. Bei Änderungen im Local Storage oder beim Mounten der Komponente wird eine Shared Reference, die als sortedFruit bezeichnet wird, mit den aktualisierten Fruchtobjekten überschrieben. Dies führt zu einer Aktualisierung der Shared References, auf die mehrere Komponenten innerhalb der Anwendung zugreifen. Ein Beispiel für die zentrale Datenhaltung ist die GalleryPage. Abbildung 4.8 visualisiert, wie sortedFruit verwendet wird, um alle gespeicherten Fruchtobjekte anzuzeigen.

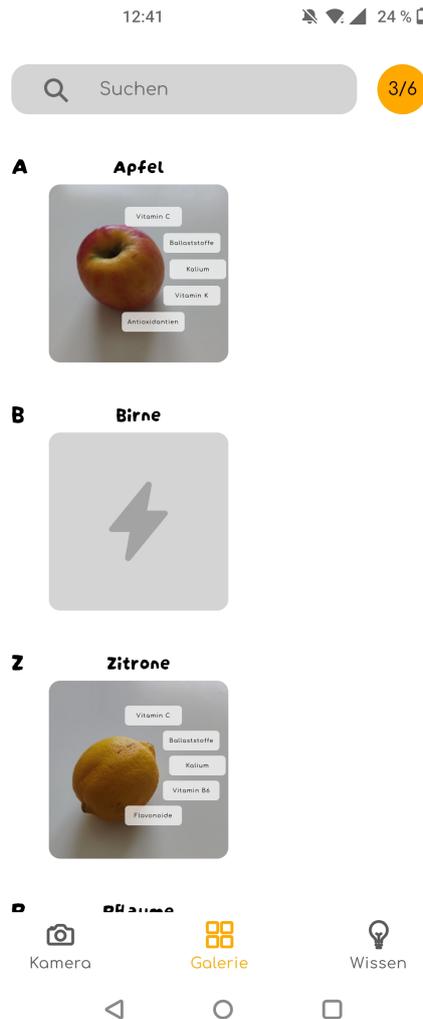


Abbildung 4.8: Abbildung der gespeicherten Fruchtobjekte auf der GalleryPage  
(Quelle: Eigene Darstellung)

Das useMediaPipe-Composable besteht aus zwei Hauptfunktionen. Die Funktion initializeObjectDetector initialisiert eine Objektdetektor-Instanz und ermöglicht die Konfiguration des Pfads zu einem TensorFlow Lite-Modell sowie der Verwendung der GPU für

mobile Geräte. Das in Abschnitt 4.1 entwickelte Modell wird dabei als spezifisches Modell verwendet. Die Funktion `detectImage` veranlasst die Erkennung von Objekten in einem Bild. Dabei wird eine Vorhersage erstellt, die verschiedene Eigenschaften wie das Klassifikationslabel, die Genauigkeit der Vorhersage und die Position des Objekts innerhalb des Bildes umfasst. Darüber hinaus werden die erkannten Objekte mit zusätzlichen Informationen angereichert, die aus einer externen JSON-Datei geladen werden, wie etwa die Liste der Nährstoffe.

Das `useCalculations-Composable` bietet eine Reihe von Funktionen zur Durchführung komplexer Berechnungen im Zusammenhang mit der Objekterkennung. Es umfasst beispielsweise Methoden zur Transformation von Videoframes auf ein Ziel-Canvas, zur Filterung von Bounding Boxen basierend auf einem Schwellenwert und zur Berechnung von Positionen für die Liste von Nährstoffen mittels geometrischer Transformationen. Da diese Funktionen an mehreren Stellen in der Anwendung benötigt werden, sind sie in diesem Composable zusammengefasst, um eine zentrale und wiederverwendbare Berechnungslogik zu gewährleisten.

### 4.2.3 Deployment

Für das Deployment der PWA wird zunächst die Firebase CLI installiert und das Firebase-Projekt im Anwendungsverzeichnis initialisiert. Dabei wird das öffentliche Verzeichnis, welches üblicherweise das Ergebnis des Build-Prozesses enthält, als `dist` festgelegt, wobei `dist` für `distributable` steht. Um eine reibungslose Navigation innerhalb der Anwendung zu gewährleisten, wird das Routing so konfiguriert, dass alle Anfragen im Sinne einer SPA auf die Hauptseite der PWA umgeleitet werden. Zusätzlich wurden die Cache-Control-Header in der Konfigurationsdatei angepasst, um durch optimiertes Caching die Performance zu verbessern. Das Projekt wird unter dem Namen `What The Fruits` in den Firebase-Einstellungen konfiguriert und ist über die Domain `fruit-dex.web.app` zugänglich. Der Deployment-Prozess wird schließlich mit dem Befehl `deploy` abgeschlossen, wodurch die Anwendung bereitgestellt wird.

# 5 Evaluation

Die Evaluation dieser Studie verfolgt das Ziel, die Wirksamkeit der in der Arbeit entwickelten AR-Anwendung im Vergleich zur traditionellen Wissensvermittlung durch gedruckte Texte zu bewerten. Durch eine Kombination von quantitativen und qualitativen Methoden soll der Einfluss der jeweiligen Lernmethode auf den Wissenszuwachs und die Zufriedenheit der Teilnehmer untersucht werden. Hierbei wird insbesondere Wert auf die statistische Auswertung der Wissenszuwächse sowie die subjektive Einschätzung der Teilnehmer hinsichtlich der Benutzerfreundlichkeit und des Lerneffekts gelegt.

## 5.1 Durchführung

Die Evaluation der entwickelten PWA wurde mit insgesamt 15 Teilnehmern im Alter von 23 bis 43 Jahren durchgeführt. Diese wurden in vier Gruppen mit jeweils drei bis vier Personen aufgeteilt. Um eine Beeinflussung der Ergebnisse durch Interaktionen der Teilnehmenden zu vermeiden, erfolgte die Datenerhebung sequenziell, wobei jede Gruppe separat getestet wurde. Die Dauer der Erhebung betrug insgesamt 30 Minuten pro Gruppe und fand in derselben Umgebung statt, in der auch die Datensätze für das Objekterkennungsmodell aufgenommen wurden. Diese konstanten Umgebungsbedingungen sollten die Erkennungsleistung des Modells begünstigen.

Zu Beginn jeder Sitzung wurde eine Wissensabfrage durchgeführt, um das Vorwissen der Teilnehmenden bezüglich der Nährstoffe ausgewählter Früchte zu erfassen. Diese Abfrage bestand aus einem dichotomen Test mit 20 Fragen, bei dem die Teilnehmenden entscheiden mussten, ob eine Aussage über eine bestimmte Frucht korrekt oder inkorrekt war. Der Test bestand aus Fragen zu insgesamt vier Früchten, wobei jeweils fünf pro Frucht abgefragt wurden. Dieser Test befindet sich im Anhang dieser Arbeit und diente als Grundlage für die spätere Analyse des Wissenszuwachses in Bezug auf die verschiedenen Früchte.

Im weiteren Verlauf mussten die Teilnehmenden zwei verschiedene Methoden zur Wissensaufnahme anwenden. Die erste Methode bestand in der Nutzung der AR-Anwendung, die im Rahmen dieser Arbeit entwickelt wurde. Die zweite Methode war traditioneller Natur und vermittelte dieselben Informationen in Form eines Fließtextes, welcher ebenfalls im Anhang zu finden ist.

Pro Methode sollten die Teilnehmenden jeweils zwei verschiedene Früchte untersuchen, sodass insgesamt Wissen über die vier vorkommenden Früchte vermittelt wurde. Für jede Methode wurden die entsprechenden Früchte zur Verfügung gestellt, damit die Teilnehmenden mit ihnen interagieren konnten. Dies sollte vergleichbare Bedingungen für die Wissensaufnahme sicherstellen.

Die Methoden wurden in unterschiedlicher Reihenfolge durchgeführt, um potenzielle Abfolgeeffekte zu vermeiden. Tabelle 5.1 veranschaulicht die Abfolge der Methoden. In Gruppe A erhielten die Teilnehmenden dabei zunächst Informationen zu Äpfeln und Orangen mittels der AR-Anwendung, gefolgt von Birnen und Pflaumen, die durch Fließtexte vermittelt wurden. In Gruppe B begann die Wissensvermittlung mit Birnen und Pflaumen über die AR-Anwendung, anschließend wurden Äpfel und Orangen durch Fließtexte vermittelt. Gruppe C startete hingegen mit der traditionellen Methode, wobei Äpfel und Orangen behandelt wurden. Danach nutzten sie die AR-Anwendung für Birnen und Pflaumen. In Gruppe D erfolgte die Wissensaufnahme ebenfalls zunächst über Fließtexte zu Birnen und Pflaumen, gefolgt von der AR-Anwendung für Äpfel und Orangen.

Tabelle 5.1: Reihenfolge der Methoden für jede Gruppe

| <b>Gruppe</b> | <b>Phase 1</b>                  | <b>Phase 2</b>                  |
|---------------|---------------------------------|---------------------------------|
| Gruppe A      | AR-Anwendung (Äpfel, Orangen)   | Fließtext (Birnen, Pflaumen)    |
| Gruppe B      | AR-Anwendung (Birnen, Pflaumen) | Fließtext (Äpfel, Orangen)      |
| Gruppe C      | Fließtext (Äpfel, Orangen)      | AR-Anwendung (Birnen, Pflaumen) |
| Gruppe D      | Fließtext (Birnen, Pflaumen)    | AR-Anwendung (Äpfel, Orangen)   |

Um die Vergleichbarkeit der Methoden sicherzustellen, erhielten die Teilnehmenden für jede Methode jeweils fünf Minuten Zeit. Im Anschluss an beide Methoden wurde eine zehnmütige Pause eingelegt, um kurzfristige Erinnerungseffekte zu reduzieren. Während dieser Pause gaben die Teilnehmenden Feedback und bewerteten dabei die Benutzerfreundlichkeit, den Spaßfaktor, den empfundenen Lernfaktor sowie die theoretische Anwendbarkeit der AR-Anwendung im Supermarkt. Nach der Pause erfolgte eine zweite Wissensabfrage, um den Lerneffekt durch die beiden Methoden zu messen. Die Erhebung der Daten erfolgte vollständig anonymisiert.

## 5.2 Ergebnisse

Tabelle 5.2 zeigt die Anzahl der richtigen Antworten der Gruppen A bis D im ersten und zweiten Durchlauf. Die maximale Anzahl richtiger Antworten pro Durchlauf beträgt dabei 20. Die Tabelle listet für jeden Teilnehmer der Gruppen die Ergebnisse beider Durchläufe auf und vergleicht die Anzahl der richtigen Antworten vor und nach der Wissensvermittlung.

Tabelle 5.2: Anzahl der richtigen Antworten im ersten und zweiten Durchlauf aller Gruppen

| <b>Gruppe</b> | <b>Teilnehmer</b> | <b>Erster Durchlauf</b> | <b>Zweiter Durchlauf</b> |
|---------------|-------------------|-------------------------|--------------------------|
| A             | Person 1          | 11                      | 12                       |
|               | Person 2          | 6                       | 9                        |
|               | Person 3          | 7                       | 8                        |
|               | Person 4          | 9                       | 10                       |
| B             | Person 1          | 11                      | 18                       |
|               | Person 2          | 10                      | 11                       |
|               | Person 3          | 10                      | 16                       |
|               | Person 4          | 11                      | 12                       |
| C             | Person 1          | 10                      | 14                       |
|               | Person 2          | 10                      | 14                       |
|               | Person 3          | 10                      | 15                       |
| D             | Person 1          | 10                      | 11                       |
|               | Person 2          | 9                       | 11                       |
|               | Person 3          | 10                      | 15                       |
|               | Person 4          | 12                      | 12                       |
| <b>Gesamt</b> |                   | <b>146</b>              | <b>187</b>               |

Die Tabelle 5.3 stellt die Anzahl der richtigen Antworten der Gruppen sortiert nach den abgefragten Früchten dar. Außerdem wird zwischen dem ersten und zweiten Durchlauf und den Wissensvermittlungsmethoden unterschieden. Anhand der Tabelle wird sichtbar, wie viele der maximal fünf möglichen korrekten Antworten pro Frucht von den Teilnehmern im ersten und im zweiten Durchlauf erreicht wurden.

Tabelle 5.3: Anzahl der richtigen Antworten im ersten und zweiten Durchlauf aller Gruppen sortiert nach Frucht

| Gruppe             | Apfel            |       | Orange           |       | Birne            |       | Pflaume          |       |
|--------------------|------------------|-------|------------------|-------|------------------|-------|------------------|-------|
|                    | 1. DL            | 2. DL |
| <b>A / Methode</b> | <i>AR</i>        |       | <i>AR</i>        |       | <i>Fließtext</i> |       | <i>Fließtext</i> |       |
| Person 1           | 2                | 5     | 1                | 1     | 5                | 3     | 3                | 3     |
| Person 2           | 3                | 3     | 2                | 2     | 1                | 2     | 0                | 2     |
| Person 3           | 3                | 3     | 2                | 1     | 1                | 1     | 1                | 3     |
| Person 4           | 3                | 3     | 3                | 2     | 1                | 2     | 2                | 3     |
| <b>B / Methode</b> | <i>Fließtext</i> |       | <i>Fließtext</i> |       | <i>AR</i>        |       | <i>AR</i>        |       |
| Person 1           | 4                | 5     | 3                | 3     | 1                | 5     | 3                | 5     |
| Person 2           | 3                | 1     | 3                | 3     | 2                | 2     | 2                | 5     |
| Person 3           | 2                | 5     | 2                | 2     | 3                | 5     | 3                | 4     |
| Person 4           | 3                | 3     | 2                | 3     | 3                | 2     | 3                | 4     |
| <b>C / Methode</b> | <i>Fließtext</i> |       | <i>Fließtext</i> |       | <i>AR</i>        |       | <i>AR</i>        |       |
| Person 1           | 2                | 3     | 4                | 3     | 1                | 4     | 3                | 4     |
| Person 2           | 4                | 5     | 3                | 3     | 2                | 4     | 1                | 2     |
| Person 3           | 3                | 4     | 3                | 3     | 1                | 4     | 3                | 4     |
| <b>D / Methode</b> | <i>AR</i>        |       | <i>AR</i>        |       | <i>Fließtext</i> |       | <i>Fließtext</i> |       |
| Person 1           | 4                | 4     | 4                | 2     | 1                | 2     | 1                | 3     |
| Person 2           | 4                | 4     | 1                | 2     | 2                | 3     | 2                | 2     |
| Person 3           | 3                | 5     | 2                | 5     | 2                | 2     | 3                | 3     |
| Person 4           | 3                | 4     | 4                | 1     | 3                | 3     | 2                | 4     |

## 5.3 Auswertung

### 5.3.1 Vergleich der Mittelwerte

Um den Wissenszuwachs der Teilnehmenden zu analysieren, wurde ein Vergleich der Mittelwerte zwischen den Ergebnissen vor und nach der Wissensvermittlung durchgeführt. Tabelle 5.4 zeigt die durchschnittliche Anzahl der richtigen Antworten der Gruppen im ersten und zweiten Durchlauf. Dabei lässt sich ein durchschnittlicher Anstieg von korrekten Antworten für alle Gruppen feststellen. Der Mittelwert stieg von 9,7 im ersten Durchlauf auf 12,5 im zweiten Durchlauf, was einem Anstieg von etwa 28,9 % entspricht. Dies deutet darauf hin, dass die Teilnehmer durch die Nutzung beider Lernmethoden ihr Wissen erweitern konnten.

Außerdem lag der Mittelwert der richtigen Antworten bei 9,7 von 20 maximal möglichen Punkten. Dies entspricht 48,5 % und deutet auf ein solides Grundwissen der Teilnehmer hin.

Tabelle 5.4: Durchschnittliche Anzahl der richtigen Antworten aller Gruppen

| <b>Gruppe</b>       | <b>Erster Durchlauf</b> | <b>Zweiter Durchlauf</b> | <b>Teilnehmer</b> |
|---------------------|-------------------------|--------------------------|-------------------|
| <b>Gruppe A</b>     | 33                      | 39                       | 4                 |
| <b>Gruppe B</b>     | 42                      | 57                       | 4                 |
| <b>Gruppe C</b>     | 30                      | 42                       | 3                 |
| <b>Gruppe D</b>     | 41                      | 49                       | 4                 |
| <b>Durchschnitt</b> | 9,7                     | 12,5                     | 15                |

Die Tabellen 5.5 bis 5.8 zeigen die durchschnittliche Anzahl der richtigen Antworten pro Frucht im ersten und zweiten Durchlauf. Dabei wird der Wissenszuwachs der Teilnehmer abhängig von der jeweiligen Frucht deutlich. Der Maximalwert liegt bei 5 richtigen Antworten. Der Mittelwert der richtigen Antworten bei dem Apfel stieg von 3,1 im ersten auf 3,8 im zweiten Durchlauf, was einem Wissenszuwachs von 22,6 % entspricht. Außerdem wird sichtbar, dass der Mittelwert bereits im ersten Durchlauf mit 3,1 von maximal 5 Punkten relativ hoch war. Dies deutet darauf hin, dass die Teilnehmer schon vor der Studie über ein solides Grundwissen in Bezug auf den Apfel verfügten.

Im Gegensatz dazu sank der Mittelwert bei der Orange leicht von 2,6 auf 2,4, was einem Rückgang von etwa 7,7 % entspricht. Dies könnte auf spezifische Schwierigkeiten bei der Vermittlung der Informationen zu dieser Frucht hindeuten.

Bei der Birne war eine deutliche Steigerung von 1,9 auf 2,9 richtige Antworten zu verzeichnen. Dies entspricht einem Wissenszuwachs von 52,6 % und legt nahe, dass die Informationen zur Birne besonders effektiv vermittelt wurden.

Ein ähnliches Muster zeigt sich bei der Pflaume, bei der der durchschnittliche Wissenszuwachs von 2,1 auf 3,4 richtige Antworten einen Anstieg von 61,9 % darstellt. Auch hier ist die Steigerung deutlich und spricht für eine erfolgreiche Wissensvermittlung.

Tabelle 5.5: Durchschnittliche Anzahl der richtigen Antworten für den Apfel

| <b>Frucht</b>       | <b>Erster Durchlauf</b> | <b>Zweiter Durchlauf</b> | <b>Teilnehmer</b> |
|---------------------|-------------------------|--------------------------|-------------------|
| <b>Apfel</b>        | 46                      | 57                       |                   |
| <b>Durchschnitt</b> | 3,1                     | 3,8                      | 15                |

Tabelle 5.6: Durchschnittliche Anzahl der richtigen Antworten für die Orange

| <b>Frucht</b>       | <b>Erster Durchlauf</b> | <b>Zweiter Durchlauf</b> | <b>Teilnehmer</b> |
|---------------------|-------------------------|--------------------------|-------------------|
| <b>Orange</b>       | 39                      | 36                       |                   |
| <b>Durchschnitt</b> | 2,6                     | 2,4                      | 15                |

Tabelle 5.7: Durchschnittliche Anzahl der richtigen Antworten für die Birne

| <b>Frucht</b>       | <b>Erster Durchlauf</b> | <b>Zweiter Durchlauf</b> | <b>Teilnehmer</b> |
|---------------------|-------------------------|--------------------------|-------------------|
| <b>Birne</b>        | 29                      | 44                       |                   |
| <b>Durchschnitt</b> | 1,9                     | 2,9                      | 15                |

Tabelle 5.8: Durchschnittliche Anzahl der richtigen Antworten für die Pflaume

| <b>Frucht</b>       | <b>Erster Durchlauf</b> | <b>Zweiter Durchlauf</b> | <b>Teilnehmer</b> |
|---------------------|-------------------------|--------------------------|-------------------|
| <b>Pflaume</b>      | 32                      | 51                       |                   |
| <b>Durchschnitt</b> | 2,1                     | 3,4                      | 15                |

Die Ergebnisse der Tabelle 5.9 zeigen die durchschnittliche Anzahl der richtigen Antworten sortiert nach verwendeter Methode. Im ersten Durchlauf erzielten die Teilnehmer bei der AR-Methode eine durchschnittlich Anzahl von 5,0 richtigen Antworten. Im zweiten Durchlauf stieg dieser Wert auf 6,7 richtige Antworten, was einer Steigerung auf 33,5 % entspricht. Im Vergleich dazu lag der Mittelwert bei der traditionellen Methode im ersten Durchlauf bei 4,7 richtigen Antworten, während er sich im zweiten Durchlauf auf 5,8 richtige Antworten erhöhte. Daraus lässt sich eine prozentuale Steigerung von 23,4 % errechnen. Diese Ergebnisse zeigen, dass die AR-Anwendung im Durchschnitt einen größeren Wissenszuwachs bewirkte als die traditionelle Methode.

Tabelle 5.9: Durchschnittliche Anzahl der richtigen Antworten sortiert nach Methode

| Gruppe              | Wissensvermittlung AR |            | Wissensvermittlung Fließtext |            | Gruppe |
|---------------------|-----------------------|------------|------------------------------|------------|--------|
|                     | 1. DL                 | 2. DL      | 1. DL                        | 2. DL      |        |
| Gruppe A            | 19                    | 20         | 14                           | 19         | 4      |
| Gruppe B            | 20                    | 32         | 22                           | 25         | 4      |
| Gruppe C            | 11                    | 21         | 19                           | 21         | 3      |
| Gruppe D            | 25                    | 27         | 16                           | 22         | 4      |
| <b>Durchschnitt</b> | <b>5,0</b>            | <b>6,7</b> | <b>4,7</b>                   | <b>5,8</b> |        |

Die Tabellen 5.10 und 5.11 stellen die Ergebnisse der richtigen Antworten sortiert nach angewandeter Reihenfolge dar. Dabei zeigt die Tabelle 5.10 die Ergebnisse der Gruppen A und D, welche zu Beginn die AR-Anwendung verwendet haben und im Anschluss die Informationen mithilfe eines Fließtextes vermittelt bekamen. Dabei lässt sich erkennen, dass der Mittelwert von 4,9 auf 6,5 beim Verwenden der AR-Methode stieg, was einer Steigerung von etwa 32,7 % entspricht. Bei der Textmethode war eine Verbesserung von 4,5 auf 5,5 zu beobachten, was einer Steigerung von rund 22,2 % entspricht. Dies deutet darauf hin, dass bei beiden Methoden im Verlauf der Studie eine Verbesserung der Leistung festzustellen war, wobei der Anstieg bei der AR-Methode deutlich ausgeprägter ist.

Die Gruppen C und D dagegen führten die Studie in der umgekehrten Reihenfolge durch. Dies zeigt die Tabelle 5.11, in welcher eine ähnliche Tendenz zu erkennen ist. Hier stiegen die Ergebnisse bei der Textmethode von 5,0 auf 6,1, was einer Steigerung von etwa 22 % entspricht. Bei der AR-Methode verbesserte sich der Mittelwert von 5,1 auf 6,9, was einer Zunahme von etwa 35,3 % entspricht. Auch hier zeigt die AR-Methode einen etwas stärkeren Anstieg im Vergleich zur Textmethode.

Die Ergebnisse zeigen, dass die AR-Anwendung eine stärkere Lernwirkung erzielt, auch wenn der Einfluss der Reihenfolge auf die Leistung nicht signifikant erscheint. Es wurde zunächst angenommen, dass die Reihenfolge der Lerninhalte einen Einfluss auf die Ergebnisse haben könnte, da typischerweise das zuletzt Erlernte präziser im Gedächtnis bleibt. Jedoch könnte dieser Effekt durch die im Rahmen der Untersuchung durchgeführte Feedbackrunde abgeschwächt worden sein, da diese als Ablenkung vom zuvor Gelernten diente. Zusammengefasst deuten diese Ergebnisse darauf hin, dass die Wissensvermittlung mit der AR-Anwendung effektiver war.

Tabelle 5.10: Ergebnisse der richtigen Antworten sortiert nach angewendeter Reihenfolge für die Gruppen A und B

| Gruppe              | AR                       |            | Fließtext                |            |
|---------------------|--------------------------|------------|--------------------------|------------|
|                     | 1. DL                    | 2. DL      | 1. DL                    | 2. DL      |
| <b>A / Frucht</b>   | <i>Apfel und Orange</i>  |            | <i>Birne und Pflaume</i> |            |
| Person 1            | 3                        | 6          | 8                        | 6          |
| Person 2            | 5                        | 5          | 1                        | 4          |
| Person 3            | 5                        | 4          | 2                        | 4          |
| Person 4            | 6                        | 5          | 3                        | 5          |
| <b>B / Frucht</b>   | <i>Birne und Pflaume</i> |            | <i>Apfel und Orange</i>  |            |
| Person 1            | 4                        | 10         | 7                        | 8          |
| Person 2            | 4                        | 7          | 6                        | 4          |
| Person 3            | 6                        | 9          | 4                        | 7          |
| Person 4            | 6                        | 6          | 5                        | 6          |
| <b>Durchschnitt</b> | <b>4,9</b>               | <b>6,5</b> | <b>4,5</b>               | <b>5,5</b> |

Tabelle 5.11: Ergebnisse der richtigen Antworten sortiert nach angewendeter Reihenfolge für die Gruppen C und D

| Gruppe              | Fließtext                |            | AR                       |            |
|---------------------|--------------------------|------------|--------------------------|------------|
|                     | 1. DL                    | 2. DL      | 1. DL                    | 2. DL      |
| <b>C / Frucht</b>   | <i>Apfel und Orange</i>  |            | <i>Birne und Pflaume</i> |            |
| Person 1            | 6                        | 6          | 4                        | 8          |
| Person 2            | 7                        | 8          | 3                        | 6          |
| Person 3            | 6                        | 7          | 4                        | 7          |
| <b>D / Frucht</b>   | <i>Birne und Pflaume</i> |            | <i>Apfel und Orange</i>  |            |
| Person 1            | 2                        | 5          | 8                        | 6          |
| Person 2            | 4                        | 5          | 5                        | 6          |
| Person 3            | 5                        | 5          | 5                        | 10         |
| Person 4            | 5                        | 7          | 7                        | 5          |
| <b>Durchschnitt</b> | <b>5,0</b>               | <b>6,1</b> | <b>5,1</b>               | <b>6,9</b> |

### **5.3.2 Inhaltsanalyse des mündlichen Feedbacks**

Das Feedback der Nutzer zur Anwendung der App *What The Fruits* war überwiegend positiv, insbesondere hinsichtlich der Interaktivität und Benutzerfreundlichkeit. Viele Teilnehmer äußerten, dass sie das Fotografieren der Früchte innerhalb der Anwendung als besonders unterhaltsam empfanden. Diese Funktion trug dazu bei, den Lernprozess spielerisch zu gestalten und das Interesse an der App aufrechtzuerhalten. Außerdem konnten die Teilnehmer ihre gespeicherten Informationen jederzeit abrufen, um Wissensinhalte zu wiederholen.

Auch die Sammelfunktion wurde als besonders motivierend beschrieben. Diese weckte die Neugier der Teilnehmenden und regte dazu an, das Fruchtlexikon zu vervollständigen.

Die Teilnehmenden gaben zudem an, dass die Darstellung der Nährstoffe prägnant und übersichtlich war, wodurch diese schnell erfasst werden konnten.

Die Möglichkeit, die App zukünftig auch im Supermarkt zu verwenden, wurde ebenfalls positiv bewertet. Dies gilt unter der Prämisse, dass mehr als sechs Früchte integriert sind. Der potenzielle Nutzen der Anwendung im Alltag wird dadurch unterstrichen.

Nur eine geringe Anzahl der Teilnehmenden gab an, dass sie das Lernen durch klassische Methoden bevorzugen. Dazu zählen das Lesen von Fließtexten oder die Nutzung traditioneller Lehrmaterialien. Dies deutet darauf hin, dass die AR-Anwendung nicht für jeden Lernenden geeignet ist und weiterhin eine Präferenz für bewährte Lernmethoden besteht.

### **5.3.3 Inhaltsanalyse der Beobachtungen**

Die Beobachtungen während der Nutzung der App ergänzen das Feedback und verdeutlichen den positiven Gesamteindruck. Die Teilnehmer interagierten nicht nur mit der App, sondern auch verstärkt miteinander, was zu einer lebendigen und kommunikativen Atmosphäre führte. Sie unterstützten sich gegenseitig beim Umgang mit der Anwendung und lösten Probleme gemeinsam. Beim Nutzen der PWA wurde sichtbar mehr gelacht als beim Lesen des Fließtextes. Es wurde deutlich, dass die Nutzer erkennbar Spaß am gemeinsamen Erkunden der Anwendung hatten. Diese soziale Komponente des Lernens förderte eine spielerische Herangehensweise und trug zur positiven Wahrnehmung bei. Darüber hinaus wurde die App durchgehend für die gesamten fünf Minuten verwendet, während die meisten Teilnehmenden den Fließtext bereits nach etwa drei Minuten beiseitelegten.

Auf der anderen Seite zeigte sich, dass die anfängliche Wissensabfrage bei einigen Teilnehmern zu einem Gefühl von Prüfungsdruck führte. Dieser Druck wirkte sich insofern aus, dass die App nicht mit der gewünschten Leichtigkeit genutzt wurde.

Des Weiteren wurde beobachtet, dass die Nutzer durch zusätzliche Bereiche der App abgelenkt wurden. Sie navigierten durch die Anwendung und interagierten mit Inhalten, die für den Zweck dieser Studie nicht relevant waren. Dies könnte darauf zurückzuführen sein, dass

sie sich zunächst an die neue Technologie gewöhnen mussten und der Zeitraum von fünf Minuten zur Anwendung nicht ausreichend war.

## 5.4 Limitationen

Trotz der positiven Ergebnisse gibt es einige Limitationen, die in dieser Studie berücksichtigt werden müssen. Eine der wesentlichen Einschränkungen dieser Studie betrifft die geringe Stichprobengröße. Mit nur 15 Teilnehmenden ist die Aussagekraft hinsichtlich der Verallgemeinerung der Ergebnisse eingeschränkt. Eine größere Stichprobe könnte zu repräsentativeren Erkenntnissen beitragen und eine fundiertere Bewertung der Anwendbarkeit der AR-Technologie in unterschiedlichen Kontexten ermöglichen. Zudem ist der Umfang der trainierten Früchte ein weiterer limitierender Faktor. Das in der Anwendung verwendete Objekterkennungsmodell ist auf eine kleine Anzahl von Früchten beschränkt. Eine Erweiterung des Trainings auf eine größere Vielfalt von Früchten würde die Relevanz der Anwendung in realen Nutzungsszenarien signifikant erhöhen.

Die Studie wurde in einer kontrollierten Umgebung durchgeführt, um die Vergleichbarkeit der beiden Methoden sicherzustellen. Dadurch blieb jedoch die Untersuchung der Alltagstauglichkeit unberücksichtigt. In zukünftigen Studien sollte die Anwendung von den Teilnehmenden in realen Umgebungen, wie beispielsweise einem Supermarkt, getestet werden, um den Einsatz im Alltag besser evaluieren zu können.

Ein weiterer Aspekt betrifft den Prüfungsdruck, den einige Teilnehmende empfanden. Die anfängliche Wissensabfrage führte bei einigen zu Stress, was sich negativ auf ihre Nutzung der App auswirkte. Dieses Stressgefühl könnte die Ergebnisse beeinflusst haben und legt nahe, dass die Methodik zur Wissensüberprüfung angepasst werden sollte, um eine authentischere Lernerfahrung zu ermöglichen.

Auch die Technologie bringt gewisse Einschränkungen mit sich. Während mobile AR-Anwendungen auf Smartphones zugänglich und flexibel sind, ist die Nutzung dennoch von der Leistungsfähigkeit des Geräts und der Kameraqualität abhängig. Dies kann das Nutzererlebnis negativ beeinflussen und stellt eine Limitation dar, die in zukünftigen Anwendungen optimiert werden könnte.

Insgesamt zeigen diese Limitationen auf, dass trotz der vielversprechenden Ergebnisse weitere Untersuchungen und Optimierungen notwendig sind, um das volle Potenzial von AR-basierten Lernanwendungen im Alltag zu erfassen.

## 6 Fazit

Insgesamt ist festzustellen, dass die Anzahl der richtig beantworteten Fragen in allen Gruppen mithilfe der AR-Anwendung deutlich höhere Ergebnisse erzielte als bei der traditionellen Methode. Dies zeigt, dass die AR-Anwendung eine effektivere Lernmethode im Vergleich zur traditionellen Wissensvermittlung durch gedruckte Texte ist.

Dieser Befund lässt sich durch die Theorie des situierten Lernens erklären, die betont, dass Lernen am effektivsten in einem authentischen Kontext stattfindet. Die Teilnehmer der Studie konnten digitale Informationen direkt mit den physischen Objekten in ihrer Umgebung verknüpfen, was zu einem tieferen Verständnis und einer aktiveren Auseinandersetzung mit den Inhalten führte. Darüber hinaus werden durch AR-basierte Lernmethoden effektivere Lernbedingungen im Sinne der konstruktivistischen Lerntheorie geschaffen. Die Darbietung multipler Blickwinkel vermittelt den Nutzern ein übersichtliches Verständnis der Inhalte. Zudem wird selbstgesteuertes und aktives Lernen gefördert, indem Lernende die Möglichkeit haben, Informationen im eigenen Tempo zu erkunden und interaktiv zu verarbeiten. Diese Umsetzung wird durch bildbasierte-AR ermöglicht, welche die physische Umgebung mit digitalen Informationen ergänzt, sodass vielfältige Lernmöglichkeiten geschaffen werden. Außerdem werden die Lerninhalte visuell ansprechend dargestellt, sodass die Aufnahme und das Verständnis von Informationen erleichtert wird. Die bildbasierte-AR wird durch KI-gestützte Objekterkennung realisiert, die für eine präzise Erkennung der Früchte unerlässlich ist.

Allerdings wurde beobachtet, dass die Einführung der neuen Technologie auch zu Ablenkung führen kann. Die Teilnehmenden interagierten teilweise mit zusätzlichen Funktionen der Anwendung, die für das eigentliche Lernziel nicht relevant waren. Diese Ablenkungen könnten auf die Notwendigkeit zurückzuführen sein, sich zunächst an die neuartige Technologie zu gewöhnen. Eine längere Testphase könnte möglicherweise helfen, diese Anpassungsprobleme zu überwinden und den Nutzern mehr Zeit geben, sich vollständig auf die Lerninhalte zu konzentrieren. Des Weiteren ist es wichtig, zwischen den unterschiedlichen Lerntypen zu differenzieren, da AR nicht für alle Nutzer gleichermaßen effektiv ist. Diese Erkenntnisse bestätigen die Hypothese H1 teilweise, wonach der Einsatz von AR das Verständnis und das Erinnerungsvermögen von Lerninhalten fördert.

Zusätzlich wurde die Qualität der sozialen Interaktionen im Lernkontext durch die intensive Kommunikation zwischen den Teilnehmenden während der Studie signifikant beeinflusst.

Die Teilnehmenden interagierten miteinander, lachten und kommunizierten verstärkt, was den Lernprozess angenehmer und gemeinschaftlicher gestaltete. Diese dynamischen Interaktionen fördern nicht nur eine höhere Motivation, sondern tragen auch zur Verbesserung des Lernniveaus bei.

Auch die Prinzipien des spielbasierten Lernens, die durch Herausforderungen und Neugier die Motivation und das Engagement der Lernenden steigern, sind in der AR-Anwendung integriert. Die Teilnehmenden konnten durch die bildbasierten AR-Elemente Informationen auf spielerische Weise entdecken und anwenden, was zu einer effizienteren Wissensaufnahme führte. Dies zeigt sich insbesondere in der positiven Reaktion auf die Sammelfunktion der App, die den Lernprozess durch spielerische Elemente motivierend gestaltete. Durch diese Gamification-Elemente wurde das Lernen nicht nur effektiver, sondern auch ansprechender und nachhaltiger für die Teilnehmenden. Diese Ergebnisse bestätigen Hypothese H2, wonach AR-Anwendungen eine motivierende Wirkung haben.

Der alltägliche Vorteil von AR-basierten Lernanwendungen wird durch den Einsatz mobiler AR besonders deutlich. Die Nutzer können die App überall einsetzen und eigene relevante Lernräume schaffen, ohne auf ortsabhängige AR angewiesen zu sein. Im Gegensatz zu Head-Worn-AR, das Nutzer von ihrer Umgebung isoliert, ermöglicht die Nutzung des Smartphones eine Verbindung zwischen der realen und der digitalen Welt. Dieses Feedback wird durch die Tatsache bestätigt, dass die Teilnehmenden angaben, die Anwendung im Supermarkt nutzen zu wollen, sobald eine größere Auswahl an Früchten integriert wäre. Leider konnte die Anwendung aus Gründen der Vergleichbarkeit nicht im Supermarkt getestet werden, weshalb Hypothese H3 nicht abschließend geklärt werden konnte.

Zusammenfassend deuten die Ergebnisse darauf hin, dass die AR-Anwendung nicht nur das Engagement der Lernenden steigert, sondern auch die Qualität des Lernens verbessert. Die spielerischen Elemente, die visuelle Darstellung der Informationen, sowie die erhöhte soziale Interaktion tragen maßgeblich zum Lernerfolg bei. Zukünftige Anwendungen könnten von der Integration solcher Technologien profitieren, um das Lernen in alltäglichen Kontexten weiter zu fördern. Das Prinzip der AR-basierten Lernanwendung lässt sich zudem auf weitere Bereiche übertragen, wie beispielsweise im Tourismus, in Tierparks oder bei der Erkundung von Blumen und Pflanzen. In diesem Zusammenhang könnte AR die Nutzererfahrung bereichern und die Interaktion mit der Umgebung intensivieren.

# Literatur

- Arvanitis, T. N., Petrou, A., Knight, J. F., Savas, S., Sotiriou, S., Gargalakos, M., & Gialouri, E. (2009). Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities. *Personal and ubiquitous computing*, 13, 243–250.
- Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE computer graphics and applications*, 21(6), 34–47.
- Ballantyne, R., & Packer, J. (2002). Nature-based excursions: School students' perceptions of learning in natural environments. *International research in geographical and environmental education*, 11(3), 218–236.
- Bandura, A. (1977). Social learning theory. *Englewood Cliffs*.
- Billinghurst, M., Clark, A., Lee, G., et al. (2015). A survey of augmented reality. *Foundations and Trends® in Human-Computer Interaction*, 8(2-3), 73–272.
- Birchfield, D., & Megowan-Romanowicz, C. (2009). Earth science learning in SMALLab: A design experiment for mixed reality. *International Journal of Computer-Supported Collaborative Learning*, 4, 403–421.
- Bitgood, S. (1989). School field trips: An overview. *Visitor behavior*, 4(2), 3–6.
- Bogner, F. X. (1998). The influence of short-term outdoor ecology education on long-term variables of environmental perspective. *The Journal of environmental education*, 29(4), 17–29.
- Broll, W., Lindt, I., Herbst, I., Ohlenburg, J., Braun, A.-K., & Wetzels, R. (2008). Toward next-gen mobile AR games. *IEEE Computer Graphics and Applications*, 28(4), 40–48.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. 1989, 18(1), 32–42.
- Bruner, J. (1974). *Toward a theory of instruction*. Harvard university press.
- Chen, Y.-C., Chi, H.-L., Hung, W.-H., & Kang, S.-C. (2011). Use of tangible and augmented reality models in engineering graphics courses. *Journal of Professional Issues in Engineering Education & Practice*, 137(4), 267–276.
- Chou, T.-L., & ChanLin, L.-J. (2012). Augmented Reality Smartphone Environment Orientation Application: A Case Study of the Fu-Jen University Mobile Campus Touring System.
- Chrome Developers. (2021). *Service worker overview*. Verfügbar 23. August 2024 unter <https://developer.chrome.com/docs/workbox/service-worker-overview>

- COCO Consortium. (2024). *Detection Evaluation*. Verfügbar 6. September 2024 unter <https://cocodataset.org/#detection-eval>
- Columbia University. (2024). *Artificial Intelligence (AI) vs. Machine Learning*. Verfügbar 27. August 2024 unter <https://ai.engineering.columbia.edu/ai-vs-machine-learning/>
- Council, N. R., of Behavioral, D., Sciences, S., on Science Education, B., & on a Conceptual Framework for New K-12 Science Education Standards, C. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.
- Cunningham, D. J. (1992). Beyond educational psychology: Steps toward an educational semiotic. *Educational psychology review*, 4, 165–194.
- Dede, C. (2009). Immersive interfaces for engagement and learning. *science*, 323(5910), 66–69.
- Driscoll, M. P. (1994). *Psychology of learning for instruction*. Allyn & Bacon.
- Dunleavy, M., & Dede, C. (2014). Augmented reality teaching and learning. In *Handbook of Research on Educational Communications and Technology* (S. 735–745). Springer New York.
- Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of science Education and Technology*, 18, 7–22.
- Facer, K., Joiner, R., Stanton, D., Reid, J., Hull, R., & Kirk, D. (2014). Savannah: mobile gaming and learning? In *Handbook of research on new literacies* (S. 1003–1020). Routledge.
- Falk, J. H. (1983). Field trips: A look at environmental effects on learning. *Journal of Biological Education*, 17(2), 137–142.
- Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2009). Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, 32(9), 1627–1645.
- Firestore Developers. (2024). *Weitere Informationen zu Build-Produkten*. Verfügbar 3. September 2024 unter <https://firebase.google.com/docs/build>
- Fu, A. (2024). *PWA Vite Plugin*. Verfügbar 9. September 2024 unter <https://vite-pwa-org.netlify.app>
- Garner, L. C., & Gallo, M. A. (2005). Field trips and their effect on student achievement and attitudes: A comparison of physical versus virtual field trips to the Indian River Lagoon. *Journal of College Science Teaching*, 34(5), 14.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Google AI. (2024a). *Aufgabenübersicht zur Objekterkennung*. Verfügbar 13. August 2024 unter [https://ai.google.dev/edge/mediapipe/solutions/vision/object\\_detector](https://ai.google.dev/edge/mediapipe/solutions/vision/object_detector)
- Google AI. (2024b). *MediaPipe Object Detector Customization*. Verfügbar 13. August 2024 unter [https://ai.google.dev/edge/mediapipe/solutions/customization/object\\_detector](https://ai.google.dev/edge/mediapipe/solutions/customization/object_detector)
- Google AI. (2024c). *MediaPipe-Modellerstellung*. Verfügbar 13. August 2024 unter [https://ai.google.dev/edge/mediapipe/solutions/model\\_maker](https://ai.google.dev/edge/mediapipe/solutions/model_maker)

- Google Developers. (2022). *Add a web app manifest*. Verfügbar 23. August 2024 unter <https://web.dev/articles/add-manifest>
- Görz, G., & Schneeberger, J. (Hrsg.). (2003). *Handbuch der Künstlichen Intelligenz*. Oldenbourg Wissenschaftsverlag.
- Gottfried, J. (1980). Do children learn on school field trips? *Curator: The Museum Journal*, 23(3), 165–174.
- Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American psychologist*, 53(1), 5.
- Harvard Graduate School of Education. (2024). EcoMOBILE (Ecosystems Mobile Outdoor Blended Immersive Learning Environment) Project. Verfügbar 30. August 2024 unter <http://ecomobile.gse.harvard.edu>
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Bd. 2). Springer.
- Heesen, B. (2021). Ausblick Machine Learning & Künstliche Intelligenz. In *Data Science und Statistik mit R: Anwendungslösungen für die Praxis* (S. 367–369). Springer Fachmedien Wiesbaden.
- Hinton, G., & Sejnowski, T. J. (1999). *Unsupervised learning: foundations of neural computation*. MIT press.
- Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: a friendly introduction. *Journal of Big Data*, 9(1), 102.
- Ionic Web Docs. (2024a). *Introduction to Ionic*. Verfügbar 13. August 2024 unter <https://ionicframework.com/docs/>
- Ionic Web Docs. (2024b). *Progressive Web Apps in Vue*. Verfügbar 23. August 2024 unter <https://ionicframework.com/docs/vue/pwa>
- Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 201–210.
- Jetbrains. (2024). *IntelliJ IDEA: JavaScript*. Verfügbar 13. August 2024 unter <https://www.jetbrains.com/help/idea/javascript-specific-guidelines.html>
- Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). *The 2010 Horizon Report*. ERIC.
- Johnson, L. F., Levine, A., Smith, R. S., & Haywood, K. (2010). Key emerging technologies for postsecondary education. *Education Digest*, 76(2), 34–38.
- Kaggle. (2024). *Datasets*. Verfügbar 3. September 2024 unter <https://www.kaggle.com/datasets>
- Kamarainen, A. M., Metcalf, S., Grotzer, T., Browne, A., Mazzuca, D., Tutwiler, M. S., & Dede, C. (2013). EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. *Computers Education*, 68, 545–556.
- Klinker, K., Fries, V., Wiesche, M., & Krcmar, H. (2017). CatCare: Designing a serious game to foster hand hygiene compliance in health care facilities.

- Klinker, K., Wiesche, M., & Krcmar, H. (2020). Digital transformation in health care: Augmented reality for hands-free service innovation. *Information Systems Frontiers*, 22(6), 1419–1431.
- Klopfer, E. (2008). Augmented learning.
- Klopfer, E., & Sheldon, J. (2010). Augmenting your own reality: Student authoring of science-based augmented reality games. *New directions for youth development*, 2010(128), 85–94.
- Klopfer, E., & Squire, K. (2008). Environmental Detectives—the development of an augmented reality platform for environmental simulations. *Educational technology research and development*, 56, 203–228.
- Knapp, D., & Barrie, E. (2001). Content evaluation of an environmental science field trip. *Journal of science Education and Technology*, 10, 351–357.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Label Studio Docs. (2024a). *Open Source Data Labeling Platform*. Verfügbar 3. September 2024 unter <https://labelstud.io>
- Label Studio Docs. (2024b). *Zero to One: Getting Started with Label Studio*. Verfügbar 3. September 2024 unter <https://labelstud.io/blog/zero-to-one-getting-started-with-label-studio>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.
- Levy, M. (2023). *Google Colab for Deep Learning: A Step-by-Step Guide*. Verfügbar 13. August 2024 unter <https://www.dataquest.io/blog/getting-started-with-google-colab-for-deep-learning/>
- Liu, T.-Y., Tan, T.-H., & Chu, Y.-L. (2009). Outdoor natural science learning with an RFID-supported immersive ubiquitous learning environment. *Journal of Educational Technology & Society*, 12(4), 161–175.
- Liu, T.-Y. (2009). A context-aware ubiquitous learning environment for language listening and speaking. *Journal of computer assisted Learning*, 25(6), 515–527.
- Manzanal, R. F., Rodríguez Barreiro, L. M., & Casal Jiménez, M. (1999). Relationship between ecology fieldwork and student attitudes toward environmental protection. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 36(4), 431–453.
- Martin, S., Diaz, G., Sancristobal, E., Gil, R., Castro, M., & Peire, J. (2011). New technology trends in education: Seven years of forecasts and convergence. *Computers & Education*, 57(3), 1893–1906.
- Mathews, J. M. (2010). Using a studio-based pedagogy to engage students in the design of mobile-based media. *English Teaching: Practice and Critique*, 9(1), 87–102.
- McCombs, B. L., & Whisler, J. S. (1997). *The Learner-Centered Classroom and School: Strategies for Increasing Student Motivation and Achievement*. The Jossey-Bass Education Series. ERIC.

- MDN Web Docs. (2024a). *CSS basics*. Verfügbar 23. August 2024 unter [https://developer.mozilla.org/en-US/docs/Learn/Getting\\_started\\_with\\_the\\_web/CSS\\_basics](https://developer.mozilla.org/en-US/docs/Learn/Getting_started_with_the_web/CSS_basics)
- MDN Web Docs. (2024b). *Document Object Model*. Verfügbar 13. September 2024 unter [https://developer.mozilla.org/en-US/docs/Web/API/Document\\_Object\\_Model](https://developer.mozilla.org/en-US/docs/Web/API/Document_Object_Model)
- MDN Web Docs. (2024c). *Getting Started with the Web*. Verfügbar 23. August 2024 unter [https://developer.mozilla.org/en-US/docs/Learn/Getting\\_started\\_with\\_the\\_web](https://developer.mozilla.org/en-US/docs/Learn/Getting_started_with_the_web)
- MDN Web Docs. (2024d). *HTML basics*. Verfügbar 23. August 2024 unter [https://developer.mozilla.org/en-US/docs/Learn/Getting\\_started\\_with\\_the\\_web/HTML\\_basics](https://developer.mozilla.org/en-US/docs/Learn/Getting_started_with_the_web/HTML_basics)
- MDN Web Docs. (2024e). *Introducing asynchronous JavaScript*. Verfügbar 23. August 2024 unter <https://developer.mozilla.org/en-US/docs/Learn/JavaScript/Asynchronous/Introducing>
- MDN Web Docs. (2024f). *JavaScript basics*. Verfügbar 23. August 2024 unter [https://developer.mozilla.org/en-US/docs/Learn/Getting\\_started\\_with\\_the\\_web/JavaScript\\_basics](https://developer.mozilla.org/en-US/docs/Learn/Getting_started_with_the_web/JavaScript_basics)
- MDN Web Docs. (2024g). *Progressive web apps*. Verfügbar 23. August 2024 unter [https://developer.mozilla.org/en-US/docs/Web/Progressive\\_web\\_apps](https://developer.mozilla.org/en-US/docs/Web/Progressive_web_apps)
- MDN Web Docs. (2024h). *Promise*. Verfügbar 23. August 2024 unter [https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global\\_Objects/Promise](https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/Promise)
- Mestre, J. (2002). *Transfer of learning: Issues and research agenda*. Arlington, VA: *National Science Foundation*.
- Metcalf, S. J., & Tinker, R. F. (2004). Probeware and handhelds in elementary and middle school science. *Journal of science education and technology*, 13, 43–49.
- Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12), 1321–1329.
- Minut, M. (2024). *Fruits-262*. Verfügbar 3. September 2024 unter <https://www.kaggle.com/datasets/aelchimminut/fruits262>
- Nachmias, R., & Linn, M. C. (1987). Evaluations of science laboratory data: The role of computer-presented information. *Journal of research in science teaching*, 24(5), 491–506.
- Nicolaou, C. T., Nicolaidou, I., Zacharia, Z., & Constantinou, C. P. (2007). Enhancing fourth graders' ability to interpret graphical representations through the use of microcomputer-based labs implemented within an inquiry-based activity sequence. *Journal of computers in Mathematics and Science Teaching*, 26(1), 75–99.
- Orion, N., & Hofstein, A. (1994). Factors that influence learning during a scientific field trip in a natural environment. *Journal of research in science teaching*, 31(10), 1097–1119.
- O'Shea, P., Mitchell, R., Johnston, C., & Dede, C. (2009). Lessons learned about designing augmented realities. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, 1(1), 1–15.
- Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual review of psychology*, 49(1), 345–375.

- Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345–1359.
- Perkins, D. N., & Salomon, G. (1992). Transfer of learning. International encyclopedia of education. Ed. Torsten Husten, 2.
- Perry, J., Klopfer, E., Norton, M., Sutch, D., Sandford, R., & Facer, K. (2008). AR gone wild: two approaches to using augmented reality learning games in Zoos.
- Piaget, J. (1970). Science of education and the psychology of the child. Trans. D. Coltman.
- Przybilla, L., Schrieck, M., Klinker, K., Pflügler, C., Wiesche, M., & Krcmar, H. (2018). Combining design thinking and agile development to master highly innovative IT projects. *Projektmanagement und Vorgehensmodelle 2018-Der Einfluss der Digitalisierung auf Projektmanagementmethoden und Entwicklungsprozesse*, 113–124.
- Raheja, S. (2024). *A Comprehensive Guide to Train-Test-Validation Split in 2024*. Verfügbar 6. September 2024 unter <https://www.analyticsvidhya.com/blog/2023/11/train-test-validation-split/>
- Rolland, J. P., & Fuchs, H. (2001). Optical versus video see-through head-mounted displays. In *Fundamentals of wearable computers and augmented reality* (S. 129–172). CRC Press.
- Rosenbaum, E., Klopfer, E., & Perry, J. (2007). On location learning: Authentic applied science with networked augmented realities. *Journal of Science Education and Technology*, 16, 31–45.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210–229.
- Schmalstieg, D., & Wagner, D. (2007). Experiences with handheld augmented reality. *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 3–18.
- Seldon Technologies. (2021). *Transfer Learning for Machine Learning*. Verfügbar 27. August 2024 unter <https://www.seldon.io/transfer-learning>
- Singh, S., Ramkumar, K. R., & Kukkar, A. (2021). Machine Learning Techniques and Implementation of Different ML Algorithms. *2021 2nd Global Conference for Advancement in Technology (GCAT)*, 1–6.
- Smith, C. (2024). *Fun Facts About Pokemon for 2024*. Verfügbar 7. August 2024 unter <https://expandedramblings.com/index.php/fun-facts-about-pokemon/>
- Sotiriou, S., & Bogner, F. X. (2008). Visualizing the invisible: augmented reality as an innovative science education scheme. *Advanced Science Letters*, 1(1), 114–122.
- Squire, K. (2010). From information to experience: Place-based augmented reality games as a model for learning in a globally networked society. *Teachers College Record*, 112(10), 2565–2602.
- Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *The journal of the learning sciences*, 16(3), 371–413.
- Squire, K. D., & Jan, M. (2007). Mad city mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of science education and technology*, 16, 5–29.

- Squire, K. D., Jan, M., Matthews, J., Wagler, M., Martin, J., DeVane, B., & Holden, C. (2007). Wherever you go, there you are: Place-based augmented reality games for learning. In *The design and use of simulation computer games in education* (S. 273–304). Brill.
- Tailwind Web Docs. (2024a). *Get started with Tailwind CSS*. Verfügbar 3. September 2024 unter <https://tailwindcss.com/docs/installation/using-postcss>
- Tailwind Web Docs. (2024b). *Utility-First Fundamentals*. Verfügbar 3. September 2024 unter <https://tailwindcss.com/docs/utility-first>
- Torrey, L., & Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (S. 242–264). IGI global.
- vom Lehn, F. (2023). *Interpreting Training/Validation Accuracy and Loss*. Verfügbar 6. September 2024 unter <https://medium.com/@frederik.vl/interpreting-training-validation-accuracy-and-loss-cf16f0d5329f>
- Vue Web Docs. (2024). *Introduction to Vue.js*. Verfügbar 13. August 2024 unter <https://vuejs.org/guide/introduction.html>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes* (E. Rice, Ed. & Trans.)
- Winston, P. H. (1992). *Artificial intelligence (3rd ed.)* Addison-Wesley Longman Publishing Co., Inc.
- Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers Education*, 62, 41–49.
- Wüller, H., & Koppenburger, A. (2021). Digitalisierung in der Pflege. In M. Wiesche, I. M. Welpel, H. Remmers & H. Krcmar (Hrsg.), *Systematische Entwicklung von Dienstleistungsinnovationen : Augmented Reality für Pflege und industrielle Wartung* (S. 127–133). Springer Fachmedien Wiesbaden.
- Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object Detection With Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232.
- Zsila, Á., Orosz, G., Böthe, B., Tóth-Király, I., Király, O., Griffiths, M., & Demetrovics, Z. (2017). *An empirical study on the motivations underlying augmented reality games: The case of Pokémon Go during and after Pokémon fever*. Verfügbar 5. August 2024 unter <https://www.sciencedirect.com/science/article/abs/pii/S0191886917304117>
- Zucker, A. A., Tinker, R., Staudt, C., Mansfield, A., & Metcalf, S. (2008). Learning science in grades 3–8 using probeware and computers: Findings from the TEEMSS II project. *Journal of Science Education and Technology*, 17, 42–48.

# **Anhang**

**Texte**

**Fragebogen**

# Der Apfel: Ein Alleskönner unter den Früchten

Äpfel sind nicht nur köstlich und vielseitig einsetzbar, sondern auch ein echtes Kraftpaket an Nährstoffen. Sie gehören zu den beliebtesten Früchten weltweit, und das nicht ohne Grund: Ein Apfel bietet eine Fülle an gesundheitlichen Vorteilen, die den gesamten Körper unterstützen. Die vielen enthaltenen Vitamine und Mineralstoffe wirken sich positiv auf das Immunsystem, die Verdauung und das Herz-Kreislauf-System aus.

Eine der herausragenden Eigenschaften des Apfels ist sein hoher Gehalt an Vitamin C, einem der wichtigsten Vitamine für den menschlichen Körper. Vitamin C spielt eine entscheidende Rolle bei der Stärkung des Immunsystems, indem es die Produktion weißer Blutkörperchen fördert, die den Körper vor Infektionen schützen. Zudem unterstützt es die Bildung von Kollagen, einem Protein, das für die Gesundheit der Haut, der Gelenke und des Bindegewebes essenziell ist. Aber das ist noch nicht alles: Vitamin C ist ein starkes Antioxidans, das die Zellen vor oxidativem Stress schützt, der durch schädliche Moleküle entsteht. Diese Moleküle können Zellen schädigen und das Risiko für chronische Erkrankungen wie Herzkrankheiten und Krebs erhöhen.

Ein weiterer wichtiger Bestandteil des Apfels sind Ballaststoffe, die für eine gesunde Verdauung unverzichtbar sind. Diese unlöslichen Fasern regen die Darmtätigkeit an, unterstützen eine gesunde Verdauung und können Verstopfung vorbeugen. Doch Ballaststoffe bieten noch mehr Vorteile: Sie tragen dazu bei, den Blutzuckerspiegel zu regulieren und das Risiko für Typ-2-Diabetes zu senken. Zudem fördern sie das Wachstum einer gesunden Darmflora, was für das allgemeine Wohlbefinden wichtig ist. Ein Apfel als täglicher Snack kann also dabei helfen, die Verdauung in Schwung zu halten und gleichzeitig das Risiko für chronische Erkrankungen zu verringern.

Kalium, ein weiterer wertvoller Nährstoff im Apfel, ist von großer Bedeutung für die Aufrechterhaltung einer normalen Muskelfunktion und eines stabilen Nervensystems. Kalium hilft, den Blutdruck zu regulieren, indem es überschüssiges Natrium aus dem Körper entfernt und so die Blutgefäße entspannt. Ein regelmäßiger Verzehr von kaliumreichen Lebensmitteln wie Äpfeln kann das Risiko von Bluthochdruck und Schlaganfällen verringern.

Auch Vitamin K spielt eine entscheidende Rolle in der Ernährung und ist besonders für die Knochengesundheit wichtig. Vitamin K fördert die Blutgerinnung, was für die Heilung von Wunden unerlässlich ist. Gleichzeitig unterstützt es die Mineralisierung der Knochen und trägt zur Prävention von Knochenschwund und Osteoporose bei.

Neben all diesen Vitaminen und Mineralstoffen enthält der Apfel auch wertvolle Antioxidantien, die eine Vielzahl von Vorteilen bieten. Antioxidantien schützen die Zellen vor Schäden durch freie Radikale und tragen dazu bei, das Risiko für chronische Krankheiten zu verringern. Zudem können sie entzündungshemmend wirken und helfen, das Immunsystem zu stärken.

Alles in allem ist der Apfel ein wahres Multitalent und sollte in keiner gesunden Ernährung fehlen. Ob roh als Snack, in Salaten oder gekocht in Gerichten – die vielseitige Frucht bietet unzählige Möglichkeiten und liefert dabei wertvolle Nährstoffe, die den Körper rundum unterstützen.

# Die Birne: Eine gesunde und schmackhafte Nährstoffquelle

Birnen sind für ihre süße, saftige Textur und ihren milden Geschmack bekannt, aber sie bieten weit mehr als nur Genuss. Als äußerst nährstoffreiche Frucht tragen sie auf vielfältige Weise zur Gesundheit bei. Die in Birnen enthaltenen Vitamine und Mineralstoffe unterstützen verschiedene Körperfunktionen und bieten zahlreiche Vorteile für die Verdauung, das Immunsystem und die Knochengesundheit.

Ein zentraler Nährstoff in Birnen sind Ballaststoffe, die eine Schlüsselrolle in der Verdauung spielen. Ballaststoffe fördern die Darmbewegungen und tragen dazu bei, Verstopfung vorzubeugen, indem sie die Nahrung durch das Verdauungssystem leiten. Gleichzeitig fördern sie ein gesundes Darmmikrobiom, das wichtig für die allgemeine Gesundheit ist. Ballaststoffe unterstützen auch die Regulierung des Blutzuckerspiegels, was besonders für Menschen von Vorteil ist, die auf eine ausgeglichene Energiezufuhr achten oder das Risiko von Typ-2-Diabetes reduzieren möchten. Der regelmäßige Verzehr von Birnen kann somit helfen, die Verdauung in Balance zu halten und das Sättigungsgefühl zu verlängern, was auch beim Gewichtsmanagement hilfreich sein kann.

Vitamin C, das in Birnen enthalten ist, bringt zahlreiche Vorteile für das Immunsystem mit sich. Dieses wichtige Antioxidans unterstützt die Abwehrkräfte des Körpers, indem es die Produktion von weißen Blutkörperchen anregt und so hilft, Infektionen abzuwehren. Zudem spielt Vitamin C eine wesentliche Rolle bei der Kollagenbildung, die für die Gesundheit von Haut, Haaren, Nägeln und Gelenken wichtig ist. Darüber hinaus schützt Vitamin C die Zellen vor Schäden durch schädliche Moleküle, die oxidativen Stress verursachen und das Risiko von Herzkrankheiten, Krebs und anderen chronischen Erkrankungen erhöhen können.

Kalium ist ein weiterer wichtiger Nährstoff, der in Birnen zu finden ist. Kalium ist entscheidend für die Regulierung des Blutdrucks und unterstützt die normale Funktion von Muskeln und Nerven. Ein ausreichender Kaliumspiegel kann dazu beitragen, Bluthochdruck vorzubeugen und das Risiko von Herz-Kreislauf-Erkrankungen zu senken.

Eine weitere wertvolle Komponente der Birne ist Vitamin K, das besonders für die Blutgerinnung und die Knochengesundheit von großer Bedeutung ist. Vitamin K unterstützt den Prozess der Blutgerinnung und hilft dem Körper, Verletzungen und Wunden effizient zu heilen. Gleichzeitig spielt es eine wichtige Rolle bei der Mineralisierung der Knochen und trägt dazu bei, die Knochendichte zu erhalten und Osteoporose vorzubeugen.

Kupfer ist ein weniger bekanntes, aber ebenso wichtiges Spurenelement, das in Birnen enthalten ist. Kupfer spielt eine entscheidende Rolle bei der Bildung roter Blutkörperchen und unterstützt die Eisenaufnahme im Körper. Es ist außerdem wichtig für die Funktion des Immunsystems und trägt zur allgemeinen Gesundheit bei.

Insgesamt ist die Birne eine hervorragende Frucht, die sich positiv auf verschiedene Aspekte der Gesundheit auswirkt. Ob als Snack zwischendurch, in Salaten oder als Zutat in süßen und herzhaften Gerichten – Birnen sind eine köstliche und nahrhafte Ergänzung zu jeder Ernährung.

# Die Pflaume: Eine vitaminreiche Frucht mit gesundheitlichem Nutzen

Pflaumen sind zwar klein, aber sie stecken voller Nährstoffe und bieten eine Vielzahl von gesundheitlichen Vorteilen. Diese süßen und saftigen Früchte sind eine hervorragende Quelle für Vitamine, die für die allgemeine Gesundheit wichtig sind, insbesondere für das Immunsystem, die Verdauung und die Haut.

Ein besonders wichtiger Bestandteil der Pflaume ist Vitamin C, das eine Vielzahl von Funktionen im Körper erfüllt. Vitamin C stärkt das Immunsystem, indem es die Produktion von Antikörpern und weißen Blutkörperchen unterstützt. Darüber hinaus fördert es die Kollagenbildung, ein Protein, das für die Gesundheit von Haut, Gelenken und Bindegewebe entscheidend ist. Menschen, die regelmäßig Vitamin C konsumieren, profitieren von einem stärkeren Immunsystem, einer schnelleren Wundheilung und einem besseren Schutz vor Infektionen.

Pflaumen sind auch reich an Ballaststoffen, die die Verdauung auf natürliche Weise regulieren. Ballaststoffe fördern die Darmtätigkeit und verhindern Verstopfung, indem sie den Stuhlgang auflockern und das Wachstum nützlicher Darmbakterien unterstützen. Eine ballaststoffreiche Ernährung kann zudem helfen, das Risiko von Darmkrebs und anderen Verdauungsproblemen zu verringern.

Ein weiterer wichtiger Nährstoff in Pflaumen ist Vitamin K, das für die Blutgerinnung und die Knochengesundheit unerlässlich ist. Vitamin K hilft dabei, Verletzungen schnell zu heilen und fördert die Festigkeit der Knochen. Eine ausreichende Zufuhr dieses Vitamins kann das Risiko von Knochenbrüchen und Osteoporose verringern.

Darüber hinaus enthalten Pflaumen Vitamin A, das besonders wichtig für die Augengesundheit ist. Vitamin A spielt eine zentrale Rolle bei der Aufrechterhaltung der normalen Sehkraft und schützt vor Nachtblindheit. Es trägt auch zur Gesundheit der Haut und des Immunsystems bei.

Pflaumen liefern zudem eine ausreichende Menge an Kupfer, einem Spurenelement, das an der Bildung roter Blutkörperchen beteiligt ist und die Eisenaufnahme unterstützt. Kupfer ist außerdem wichtig für die Aufrechterhaltung eines gesunden Immunsystems und die Vorbeugung von Entzündungen.

Insgesamt sind Pflaumen eine wahre Nährstoffbombe, die zahlreiche gesundheitliche Vorteile bietet. Sie können roh, getrocknet oder in einer Vielzahl von Gerichten genossen werden und sind eine großartige Ergänzung zu einer ausgewogenen Ernährung.

# Die Orange: Eine Zitrusfrucht mit vielen Vitaminen

Orangen gehören zu den beliebtesten Zitrusfrüchten und sind weltweit für ihren frischen, süßen Geschmack und ihren hohen Gehalt an Vitamin C bekannt. Doch Orangen bieten weit mehr als nur Erfrischung – sie sind vollgepackt mit wichtigen Nährstoffen, die den Körper in vielerlei Hinsicht unterstützen.

Der wohl bekannteste Nährstoff der Orange ist Vitamin C, das eine zentrale Rolle bei der Stärkung des Immunsystems spielt. Vitamin C fördert die Bildung weißer Blutkörperchen, die für die Abwehr von Infektionen verantwortlich sind. Darüber hinaus ist es ein starkes Antioxidans, das die Zellen vor Schäden durch schädliche Moleküle schützt. Diese Schutzfunktion ist besonders wichtig, um das Risiko für Herz-Kreislauf-Erkrankungen und Krebs zu reduzieren. Vitamin C unterstützt zudem die Kollagenbildung, die für die Gesundheit von Haut, Blutgefäßen und Knochen essenziell ist.

Neben Vitamin C sind Orangen auch reich an Ballaststoffen, die eine wichtige Rolle für die Verdauung spielen. Ballaststoffe fördern die Darmgesundheit, indem sie die Verdauung regulieren und das Risiko von Verstopfung verringern. Zudem tragen sie zur Senkung des Cholesterinspiegels bei und helfen, den Blutzuckerspiegel nach den Mahlzeiten stabil zu halten. Eine ballaststoffreiche Ernährung kann das Risiko für Herzkrankungen und Typ-2-Diabetes erheblich verringern.

Kalium ist ein weiterer wertvoller Nährstoff in Orangen, der eine wichtige Rolle bei der Regulierung des Blutdrucks spielt. Kalium unterstützt die Funktion von Muskeln und Nerven und trägt zur Aufrechterhaltung eines stabilen Flüssigkeitshaushalts im Körper bei. Ein hoher Kaliumspiegel kann dazu beitragen, das Risiko von Bluthochdruck und Schlaganfällen zu verringern.

Orangen enthalten zudem Folat, ein essentielles Vitamin für die Zellteilung und die DNA-Produktion. Besonders für schwangere Frauen ist Folat von großer Bedeutung, da es zur gesunden Entwicklung des Fötus beiträgt und das Risiko von Neuralrohrdefekten verringert. Aber auch allgemein ist Folat wichtig für die Bildung roter Blutkörperchen und die Zellregeneration.

Thiamin, auch bekannt als Vitamin B1, ist ein weiterer Nährstoff, der in Orangen zu finden ist. Thiamin spielt eine entscheidende Rolle im Energiestoffwechsel und unterstützt die Funktion von Nerven und Muskeln. Es hilft dabei, die Nahrung in Energie umzuwandeln, und sorgt so dafür, dass der Körper den ganzen Tag über leistungsfähig bleibt.

Orangen sind somit nicht nur eine erfrischende Zwischenmahlzeit, sondern auch eine Quelle zahlreicher Vitamine und Mineralstoffe, die den Körper umfassend unterstützen. Sie sollten in keiner gesunden Ernährung fehlen – ob als Snack, in Smoothies oder in herzhaften Gerichten.

# Fruit Quiz

Lass dir Zeit.

**Name \***

**Durchlauf \***

**Pflaumen enthalten Vitamin C, das die Kollagenbildung unterstützt. \***

Richtig

Falsch

**Orangen enthalten Vitamin K, das die Knochengesundheit fördert. \***

Richtig

Falsch

**Birnen sind reich an Vitamin A, das die Sehkraft unterstützt. \***

Richtig

Falsch

**Der Apfel enthält Vitamin C, das das Immunsystem stärkt. \***

Richtig

Falsch

**Birnen enthalten Eisen, das den Sauerstofftransport im Blut unterstützt. \***

Richtig

Falsch

**Pflaumen enthalten Kupfer, welches wichtig für das Immunsystem ist. \***

Richtig

Falsch

**Orangen enthalten Ballaststoffe, die helfen, den Blutzuckerspiegel zu regulieren. \***

Richtig

Falsch

**Ein Apfel enthält Magnesium, das für die Muskel- und Nervenfunktion wichtig ist. \***

Richtig

Falsch

**Pflaumen sind reich an Folat, das für die Zellteilung wichtig ist. \***

Richtig

Falsch

**Birnen enthalten Kalium, das wichtig für die Funktion von Muskeln und Nerven ist. \***

Richtig

Falsch

**Orangen sind eine gute Quelle für Folat, das wichtig für die Zellteilung ist. \***

Richtig

Falsch

**Der Apfel ist eine gute Quelle für Vitamin K, das bei der Blutgerinnung hilft. \***

Richtig

Falsch

**Die Pflaume ist eine gute Quelle für Vitamin A, das wichtig für die Sehkraft ist. \***

Richtig

Falsch

**Orangen enthalten Kalium, das hilft, den Blutdruck zu regulieren. \***

Richtig

Falsch

**Ein Apfel enthält Ballaststoffe, die die Verdauung fördern. \***

Richtig

Falsch

**Birnen enthalten Thiamin, das wichtig für den Energiestoffwechsel ist. \***

Richtig

Falsch

**Der Apfel enthält Vitamin B6, das wichtig für den Stoffwechsel und die Gehirnfunktion ist. \***

Richtig

Falsch

**Pflaumen enthalten Flavonoide, die entzündungshemmend wirken und die Herzgesundheit fördern. \***

Richtig

Falsch

**Birnen sind eine gute Quelle für Vitamin K, das die Knochengesundheit unterstützt. \***

Richtig

Falsch

**Orangen sind eine gute Quelle für Kupfer, das für die Bildung roter Blutkörperchen wichtig ist. \***

Richtig

Falsch

# Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit mit dem Titel

**Digitales Fruchtlexikon: Augmented Reality und KI-gestützte Objekterkennung als interaktiver Ansatz zur Wissensvermittlung im Alltag**

selbstständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Alle Passagen, die ich wörtlich aus der Literatur oder aus anderen Quellen wie z. B. Internetseiten übernommen habe, habe ich deutlich als Zitat mit Angabe der Quelle kenntlich gemacht.

*Michelle Lamowski*

Hamburg, 19. September 2024