
PRAKTIKUMSBERICHT

Herr
Hannes Steiner

Entwicklung einer Dokumenten-Scanner-App

**Digitalisierung der Verwaltung an der Hochschule
Mittweida**

2020

Fakultät **Angewandte Computer- und
Biowissenschaften**

PRAKTIKUMSBERICHT

Entwicklung einer Dokumenten-Scanner-App

**Digitalisierung der Verwaltung an der Hochschule
Mittweida**

Autor:
Hannes Steiner

Studiengang:
Softwareentwicklung

Seminargruppe:
IF17wS-B

Matrikelnummer:
46540

Erstprüfer:
Prof. Dr. Marc Ritter

Zweitprüfer:
N.N.

Mittweida, März 2020

Bibliografische Angaben

Steiner, Hannes: Entwicklung einer Dokumenten-Scanner-App, Digitalisierung der Verwaltung an der Hochschule Mittweida, 46 Seiten, 19 Abbildungen, Hochschule Mittweida, University of Applied Sciences, Fakultät Angewandte Computer- und Biowissenschaften

Praktikumsbericht, 2020

Dieses Werk ist urheberrechtlich geschützt.

Referat

Im Rahmen eines zwölfwöchigen Forschungspraktikums an der Hochschule Mittweida arbeiteten der Student Tobias Kallauke und der Verfasser des Berichts gemeinsam an einem Projekt mit dem Hintergrund der Digitalisierung der Verwaltung von Lehr- und Forschungseinrichtung. Dabei entwickelten sie ein prototypisches Software-System, mit dessen Hilfe kontrollierte Klawsuren eingescannt und digitalisiert werden können.

I. Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis	II
Abkürzungsverzeichnis	III
1 Einleitung	1
2 Hochschule Mittweida	2
3 Problemstellung	3
3.1 Digitalisieren der Klausur-Daten	3
3.2 Klausuren-Vorlage	3
3.3 Klausuren-Vorlage verbessern	4
4 Anforderungen	5
5 Konzept der Dokumenten-Scanner-App	7
5.1 Scan-Vorlage erstellen und speichern	7
5.2 Scan-Vorlage verwenden	9
5.3 Erweiterung des Konzepts	11
6 Entwicklung des ersten Prototyps	12
6.1 Anforderungsplanung	12
6.2 Analyse und Definition	12
6.3 Grundlagen	13
6.4 Entwurf und Design	15
6.5 Implementierung	17
6.5.1 Registrieren und Anmelden	18
6.5.2 Scan-Vorlagen erstellen und speichern	19
6.5.3 Scan-Vorlage verwenden	22
7 Fazit	29
7.1 Stand des Prototyps	29
7.2 Probleme und Grenzen der App	29
7.2.1 Probleme beim Erkennen von Dokumenten	29
7.2.2 Probleme der Klausur-Vorlage beim Scannen	30

7.2.3 Weitere Probleme der App	31
7.3 Verbesserung der Klausuren-Vorlage	31
8 Ausblick	33
A Wasserfall-Modell	36
B Workflow	37
C Zusätzliche Funktionalität	40
D Installation der iOS App	41
E Tätigkeitsbericht	42
Literaturverzeichnis	45

II. Abbildungsverzeichnis

3.1 Umsetzung der Klausuren-Vorlage der Fakultät CB	4
5.1 Schematisches Schaubild zum Verständnis des Schablonen-Konzepts	7
5.2 Kompaktes Flussdiagramm zum Erstellen einer Scan-Vorlage	7
5.3 Schematische Darstellung des Vorgangs Scan-Vorlage erstellen und auf dem Server speichern	8
5.4 Schematischer Ablauf beim Verwenden einer Scan-Vorlage	9
5.5 Kompaktes Flussdiagramm zum Verwenden einer Scan-Vorlage	10
5.6 Schematische Darstellung des Vorgangs beim Verwenden einer Scan-Vorlage	10
6.1 Schematische Darstellung des Model-View-ViewModel Konzepts	14
6.2 Architektur Schemata	16
6.3 Willkommen- Registrier- und Anmelde-View	19
6.4 Die ersten Views zur Erstellung einer Scan-Vorlage	20
6.5 Views zur Erstellung von Regionen	21
6.6 Seiten-Vorschau-View	22
6.7 Views zur Erstellung von Kontrollmechanismen	23
6.8 Listen- und Detail-Ansicht von Scan-Vorlagen	24
6.9 Vorlagen-Ansicht vor und nach der Texterkennung	25
A.1 Das Wasserfall-Model nach Winston W. Royce	36
B.1 Gekürztes Flussdiagramm zum Erstellen einer Scan-Vorlage	38
B.2 Gekürztes Flussdiagramm zum Verwenden einer Scan-Vorlage	39

III. Abkürzungsverzeichnis

API	application programming interface, deutsch: Programmierschnittstelle, Seite 17
CRUD	Das Akronym CRUD steht für Create/Erstellen, Read/Lesen, Update/Aktualisieren und Delete/Löschen und umfasst die vier grundlegenden Operationen persistenter Speicher, Seite 5
Fakultät CB	Fakultät für angewandte Computer- und Biowissenschaften, Seite 2
HSMW	Hochschule Mittweida - university of applied science, Seite 2
ID	Steht für einen einzigartigen Identifikator, Seite 21
IDE	integrated development environment, deutsch: Integrierte Entwicklungsumgebung, Seite 17
MVVM	Model-View-ViewModel, Seite 14
OCR	optical character recognition, deutsch: optische Zeichenerkennung und im deutschen Synonym für Texterkennung, Seite 9

1 Einleitung

Digitalisierung wird gängig als Integration von digitaler Technologie in den Alltag verstanden und soll helfen Zeit einzusparen. Mit diesem Gedanken initiierten die Mitarbeiter Holger Langner und Falk Schmidsberger der Hochschule Mittweida das Projekt Memo Space. Im Zuge dessen sollen kleinere Forschungsergebnisse entstehen, die richtungsweisend für die Digitalisierung der Verwaltung von Lehr- und Forschungseinrichtung sind.

Im Rahmen eines zwölfwöchigen Forschungspraktikums an der Hochschule Mittweida arbeiteten der Student Tobias Kallauke und der Verfasser des Berichts gemeinsam an einem Forschungsprojekt von Memo Space. Dabei entwickelten sie ein prototypisches Software-System zum Einscannen und Digitalisieren von kontrollierten Klausuren.

2 Hochschule Mittweida

Die Hochschule Mittweida - university of applied science (HSMW) wurde vor über 150 Jahren gegründet. Heute lehrt und forscht sie mit ca. 6000 Studenten in fünf Fakultäten und vier Forschungsschwerpunkten [5]. Eines der Schwerpunkte ist die Angewandte Informatik, in dem Memo Space angesiedelt ist.

Nach eigener Einschätzung schreibt jeder Student der HSMW ca. 5 Prüfungen pro Semester. Das bedeutet, dass im Jahr um die 60.000 Klausuren kontrolliert werden. Dazu kommt, dass die Zensuren in das Verwaltungs-Intranet der HSMW eingetragen werden müssen. Dieser Vorgang ist ebenfalls aufwendig, repetitiv und deshalb kognitiv belastend.

Da die Mitarbeiter der Fakultät Angewandte Computer- und Biowissenschaften (Fakultät CB) Holger Langner und Falk Schmidsberger selbst Klausuren kontrollieren und die Problematik genau kennen, entstand hier eine der ersten Ideen für Memo Space.

3 Problemstellung

An der Kontrolle von Klausuren arbeiten zum Ende eines Semesters Hochschulmitarbeiter über Tage. Diese Aufgabe muss stets mit hoher Konzentration erledigt werden. Auch bei der anschließenden Eingabe der Noten in ein Webformular des Verwaltungs-Intranets der HSMW ist ein kognitiv belastender Navigationsaufwand notwendig. Denn um die Zensuren einzutragen, muss für jeden Studenten zunächst die jeweilige Seminargruppe und anschließend der Studenten-Name herausgesucht werden. Die unübersichtlichen Listen und die hohe Anzahl an Klicks sind für die Hochschulmitarbeiter belastend. Zusätzlich wird die Suche nach der Seminargruppe durch die ähnlich lautenden Bezeichnungen erschwert. Das führt beim Danebenklicken in der Seminargruppen-Liste dazu, dass man sich eventuell in der Navigation verirrt und von vorne beginnen muss.

3.1 Digitalisieren der Klausur-Daten

Für genau diesen Vorgang wird nach einer effizienteren Lösung gesucht. Die Prüfer sollen effektiv und möglichst zeitsparend diese Aufgabe verrichten, ohne die Überlastung ihrer Aufmerksamkeitsspanne. Zudem ist angedacht, alle relevanten Klausuren-Daten zu digitalisieren und in ein geeignetes Format zu bringen, um sie automatisiert in das Verwaltungs-Intranet überführen zu können. Darüber hinaus empfiehlt es sich, digitale Kopien der Klausuren abzuspeichern, um sie nicht nur analog zu archivieren.

3.2 Klausuren-Vorlage

Eine Klausuren-Vorlage bzw. ein Gestaltungsleitfaden für Klausuren der Fakultät Angewandte Computer- und Biowissenschaften bietet außerdem die Möglichkeit der Kontrolle des Prüfers. Durch das vorgegebene Layout der Klausur ist es möglich, Fehler des Prüfers zu erkennen. Die Klausur in Abbildung 3.1 ist nach dieser Klausuren-Vorlage angefertigt worden. Auf dem Deckblatt (3.1a) der Prüfung ist eine Tabelle mit drei Zeilen und jeweils einer Spalte pro Klausur-Aufgabe. In der ersten Zeile befinden sich die Nummern der Aufgaben, in der zweiten die zu erreichenden Punkte der Aufgabe. In der dritten Zeile trägt der Prüfer die erbrachten Punkte des Studenten ein. Unter der Tabelle befindet sich ein Feld für die erreichte Gesamtpunktzahl sowie ein weiteres für die aus den Punkten resultierende Note. Es soll nun überprüft werden, ob die Summe der erreichten Punkte mit der Gesamtpunktzahl übereinstimmt. Zudem wird auch die Note errechnet und mit dem Ergebnis des Prüfers verglichen werden. Ein weiteres Merkmal der Klausuren-Vorlage ist ein Feld, für die vom Studenten erreichten Punkte über jeder Aufgabenstellung auf den nachfolgenden Seiten. Zu sehen sind solche Felder in Abbildung 3.1b am rechten Seitenrand. Die dortige Angabe sollte mit der Angabe in der

Tabelle auf dem Deckblatt (3.1a) übereinstimmen und bietet somit noch eine weitere Möglichkeit der Kontrolle.

**Klausur
Grundlagen
Informations-
technologie**

Wintersemester
2019/2020

Professur
Medieninformatik
Prof. Dr. rer. nat.
Marc Ritter
Dipl.-Inf.
Falk Schmidberger
Dipl.-Inf.
Robert Manthey

Name: _____
Matrikel #: _____
Seminargruppe: _____

Aufgabe	1	2	3	4	5	6	7	8	9	10	11	12
Punkte	10	5	5	9	3	7	5	9	8	9	2	9
Score												

Punkte: / 81 Note: _____

Allgemeine Hinweise

Schreiben Sie Ihren Namen, Vornamen und Ihre Matrikelnummer leserlich auf alle Blätter, die Sie für Ihre Lösung verwenden - bevor Sie mit der Bearbeitung beginnen! Blätter ohne diese Angaben können nicht gewertet werden!

Geöffnete blaue Unterlagen (Spirale, Vorlesungsmitschriften, etc.) zugelassen. Sie werden als Tauschgegenmittel gewertet. Dessen abweichend wird bei Ihnen jeweils ein A4-Blatt mit ausschließlich handschriftlichen Notizen zugelassen!

Der Versuch der Täuschung hat die Wertung der Klausur mit "nicht bestanden" zur Folge.

Benutzen Sie keine Bleistiele, keine rotschreibenden Stifte und kein Tipp-Ex (oder ähnliche Produkte).

Zum Bestehen der Klausur mindestens die Hälfte der erreichbaren Punkte notwendig.

Bei mehreren Aufgabenbearbeitungen müssen Sie sich Hinweise und Hilfestellungen werden, falls erforderlich, offiziell für den zuständigen Hörsaal durchsetzen.

Geben Sie keine mehrdeutigen (oder mehrere) Lösungen an. In solchen Fällen wird statt die Lösung mit der geringeren Punktzahl gewertet (Eine richtige und eine falsche Lösung zu einer Aufgabe ergeben also 0 Punkte).

NUTZEN SIE DIE SATZFORM BEI DEN BEGRIFFEN ERLÄUTERN UND BESCHREIBEN!

NUTZEN SIE FÜR IHRE LÖSUNGEN AUCH DIE RÜCKSEITEN. GEBEN SIE STETS DIE ZUGEHÖRIGE AUFGABENNUMMER AN.

Rechnerarchitektur

Aufgabe 1
Vervollständigen Sie die Begriffe in der folgenden Skizze der von Neumann-Architektur eines Personal Computers. (/10 Punkte)

Betriebssysteme

Aufgabe 2
Welche 5 Hauptaufgaben hat ein Betriebssystem? (/5 Punkte)

Verschlüsselung

Aufgabe 3
Erläutern Sie das Verfahren der symmetrischen kryptografischen Verschlüsselung und nennen Sie Vor- und Nachteil. (/5 Punkte)

(a) Deckblatt der Klausur

(b) Erste Seite der Klausur mit drei Aufgaben

Abbildung 3.1: Umsetzung der Klausuren-Vorlage der Fakultät CB

Die Bilder sind beim Erstellen einer Scan-Vorlage mit der App entstanden.

3.3 Klausuren-Vorlage verbessern

Nachdem ein erster Prototyp zum Digitalisieren der Klausur-Daten entwickelt wurde, sollen außerdem Klausuren-Vorlagen entstehen, die für die Digitalisierung optimiert sind. Dabei sollten Probleme, die beim Einscannen der aktuellen Vorlage erkannt wurden, behoben werden.

4 Anforderungen

In diesem Kapitel werden die Anforderungen an das Software-System dargestellt.

Es soll eine App für das Betriebssystem iOS entstehen, mit der Klausuren digitalisiert werden können. Dabei müssen die, für die Notenfreigabe relevanten Daten der Klausur, in ein tabellarisches Format gebracht werden. Ferner muss die iOS-Anwendung Bilder von ausgefüllten Prüfungen aufnehmen können und den Inhalt mithilfe von Texterkennung digitalisieren.

Auch bei unterschiedlicher Seitengestaltung der Klausuren, soll die App in der Lage sein, alle relevanten Informationen, wie z. B. Name, Matrikelnummer und Note richtig zu lokalisieren und anschließend zu digitalisieren. Dafür muss der Benutzer digitale Vorlagen, sogenannte Scan-Vorlagen erstellen. Mit deren Hilfe können, die Informationen sofort gefunden und schneller digitalisiert werden, ohne die gesamte Klausuren-Seite zu verarbeiten.

Zusätzlich soll es dem Benutzer möglich sein, Kontrollmechanismen festzulegen. Wie im Abschnitt 3.2 beschrieben, kann überprüft werden, ob die Punkte auf dem Deckblatt mit den Punkten auf den nachfolgenden Seiten übereinstimmen. Des Weiteren wird kontrolliert, ob die Gesamtpunktzahl korrekt summiert sowie die daraus resultierende Note richtig errechnet wurde. Ein Kontrollmechanismus soll so gestaltet sein, dass diese Überprüfungsmöglichkeiten vom Benutzer bestimmt werden können. Bei Nichtübereinstimmung der genannten Kriterien erhält der Benutzer eine entsprechende Fehlermeldung.

Um alle relevanten Daten zu erfassen werden die kontrollierten Klausuren zunächst fotografiert. Bei der anschließenden Digitalisierung aller Schriftzeichen durch die App findet optische Zeichen- bzw. Texterkennung Anwendung. Für die Texterkennung auf externen Servern müssen die entsprechenden Server-Schnittstellen zur Kommunikation implementiert werden.

Die durch die Texterkennung digitalisierten Daten, die Klausuren-Fotos und die digitalen Scan-Vorlagen sollen außerhalb der App auf einem Server gespeichert werden. Somit können alle Benutzer der App die bereits erstellten Scan-Vorlagen verwenden. Wie bereits erwähnt, werden die digitalisierten Daten zudem in ein geeignetes Format gebracht, um sie, wie angedacht, automatisiert in das Verwaltungs-Intranet zu überführen.

Damit Synchronität der Daten zwischen allen Apps und dem Server gewährleistet werden kann, sind auf beiden Seiten Schnittstellen notwendig. Diese sollten den standardmäßigen CRUD-Operationen entsprechen. Für genauere Details zu den Sicherheits-

mechanismen und den CRUD-Operationen zum Server siehe im Praktikumsbericht von Tobias Kallauke.

5 Konzept der Dokumenten-Scanner-App

Dieses Kapitel beschreibt, wie das Softwaresystem die Anforderungen umsetzt.

5.1 Scan-Vorlage erstellen und speichern

Im folgenden Abschnitt wird zuerst verkürzt und anschließend detailliert geschildert, wie eine Scan-Vorlage zu erstellen ist.

Um wirklich nur relevante Daten zu digitalisieren, benötigt es die sogenannten Scan-Vorlagen, die in der App durch den Benutzer angelegt werden können. Eine Scan-Vorlage beinhaltet mehrere Schablonen. Solch eine Schablone wird beispielsweise auf das Deckblatt gelegt, sodass nur noch die relevanten Informationen, die digitalisiert werden sollen, zu sehen sind. Alles andere, was nicht von Interesse ist, wird von der Schablone überdeckt. In der Abbildung 5.1 ist dieses Konzept schemenhaft zu sehen. Der Benutzer kann in zwei grobe Schritte, um eine digitale Schablone anzufertigen. Zuerst muss ein Foto von der Seite aufgenommen werden. Anschließend werden die Regionen auf diesem Foto manuell markiert, die digitalisiert werden sollen. Eine Scan-Vorlage ist somit eine Sammlung aller digitalen Schablonen zu einer Klausur, denn jede Seite benötigt eine eigene individuelle Schablone.

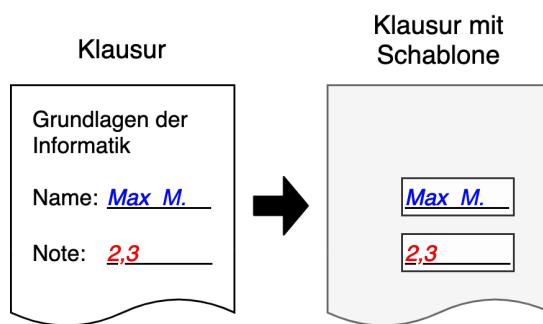


Abbildung 5.1: Schematisches Schaubild zum Verständnis des Schablonen-Konzepts

Nachfolgend wird detailliert das Erstellen einer Scan-Vorlage beschrieben. Der Vorgang ist in Abbildung 5.2 als kompaktes Flussdiagramm und in Abbildung 5.3 schematisch dargestellt.

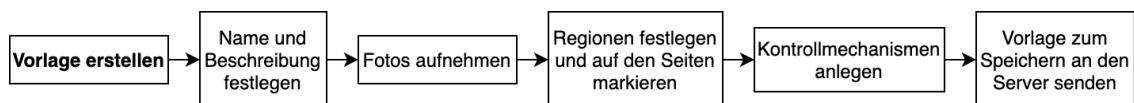


Abbildung 5.2: Kompaktes Flussdiagramm zum Erstellen einer Scan-Vorlage

1. Zu Beginn wird der Scan-Vorlage ein Name und eine Beschreibung gegeben. Anschließend muss jede Seite der Klausur fotografiert werden. Dabei wird in jedem

Bild das Dokument erkannt, vom Hintergrund getrennt bzw. ausgeschnitten und perfekt ausgerichtet. Zudem wird der Kontrast erhöht, sodass die Schrift leichter lesbar ist. Die Abbildung 3.1 resultiert aus dem Prozess (siehe auch Abbildung 6.4b).

2. Anschließend markiert der Benutzer diejenigen Regionen auf jeder Seite, die zu digitalisieren und/oder zu kontrollieren sind. In dem Schema 5.3 sind die Regionen blau hinterlegt (siehe auch Abbildung 6.6b). Zusätzlich muss jeder Region eine Bezeichnung und einer der folgenden Datentypen zugeordnet werden: Unbekannt, Vorname, Nachname, Matrikelnummer, Seminargruppe, Punkte und Note (siehe auch Abbildung 6.6a und 6.6b). Die Angabe des Datentyps ist wichtig, da dadurch die Texterkennung optimiert wird und die Ergebnisse der Texterkennung in das geeignete Format gebracht werden können.
3. Zuletzt wird die Vorlage zum Speichern an einen Server gesendet, wodurch alle Benutzer der App die bereits erstellten Scan-Vorlagen verwenden können.

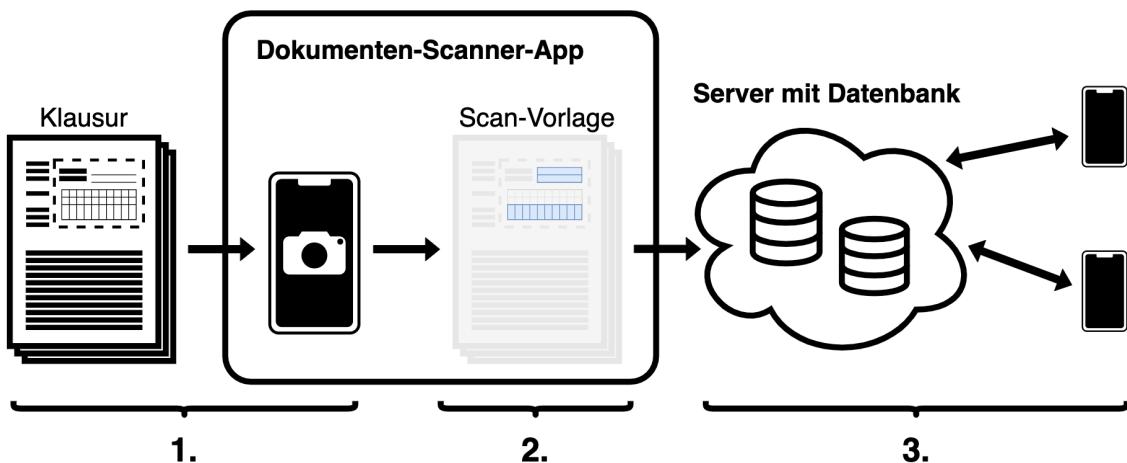


Abbildung 5.3: Schematische Darstellung des Vorgangs Scan-Vorlage erstellen und auf dem Server speichern

Zusätzlich können in der Phase der Erstellung einer Scan-Vorlage die in Kapitel 4 beschriebenen Kontrollmechanismen festgelegt werden. Dies geschieht nach Punkt 2, also nachdem alle Regionen angelegt wurden. Diese Kontrollmechanismen beruhen auf zwei simplen Konzepten.

Das erste Konzept ist der Vergleich. Zur Erinnerung: Es sollen die eingetragenen Punkte auf dem Deckblatt mit den Punkten neben der Aufgabenstellung auf den nachfolgenden Seiten auf Gleichheit überprüft werden. Das zweite Konzept baut auf dem ersten auf. Statt zwei Regionen bzw. Komponenten zu vergleichen, wird eine oder werden beide Komponenten zuvor berechnet. Beispielsweise muss die Summe der Punkte pro Aufgabe mit der Gesamtpunktzahl übereinstimmen. Das bedeutet, es wird erst die Summe berechnet und anschließend mit der vom Prüfer eingetragenen Punktzahl verglichen.

Oder aus der Gesamtpunktzahl wird die Note ermittelt und mit der vom Prüfer eingetragenen Note verglichen.

Beim Festlegen der Kontrollmechanismen wird wie folgt vorgegangen:

1. Der Benutzer wählt einen Kontrollmechanismus aus. Zur Auswahl stehen:
 - der Vergleich zweier Regionen,
 - die Gesamtpunktzahl berechnen und mit der eingetragenen Gesamtpunktzahl vergleichen,
 - oder die Note aus der Gesamtpunktzahl berechnen und mit der eingetragenen Zensur vergleichen.
2. Aus den zuvor angelegten Regionen müssen nun die passenden Regionen ausgewählt werden. Abhängig von dem gewählten Kontrollmechanismus wird dem Benutzer Auskunft darüber gegeben, welche und wie viele Regionen auszuwählen sind.
3. Der Kontrollmechanismus kann gespeichert werden, sobald in jeder zu vergleichenden Komponente die Mindestanzahl an auszuwählenden Regionen erreicht wurde. Diese werden in der Scan-Vorlage hinterlegt und mit an den Server zum Speichern gesendet. So besitzen alle Benutzer dieselben Kontrollmechanismen zu den jeweiligen Scan-Vorlagen.

5.2 Scan-Vorlage verwenden

Im folgenden Abschnitt wird zuerst verkürzt und anschließend detailliert die Verwendung und Funktionsweise einer erstellten Scan-Vorlage beschrieben.

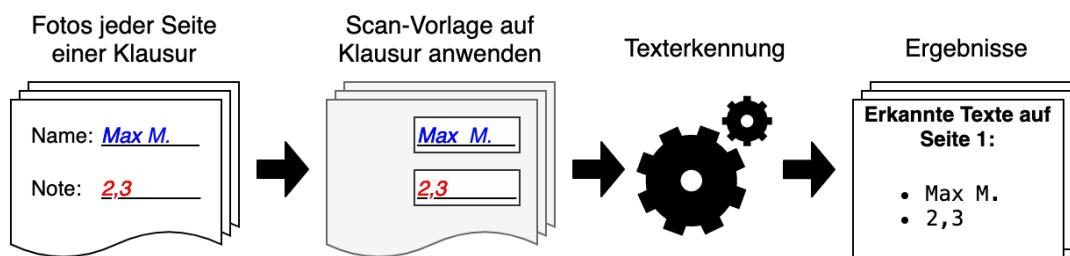


Abbildung 5.4: Schematischer Ablauf beim Verwenden einer Scan-Vorlage

Nach dem Erstellen einer Scan-Vorlage kann diese auf bewerteten Klausuren angewandt werden. In Abbildung 5.4 wird der Ablauf schematisch dargestellt. Zuerst werden die Seiten einer Klausur fotografiert. Anschließend stellen die digitalen Schablonen die relevanten Informationen jeder Seite für den nächsten Schritt zur Verfügung. Optical character recognition (OCR), auf deutsch optische Zeichenerkennung, wird dazu benutzt, die Daten aus den übrig gebliebenen Regionen zu digitalisieren.

Nachfolgend wird detailliert das Verwenden einer Scan-Vorlage beschrieben. Der Vorgang ist in Abbildung 5.5 als kompaktes Flussdiagramm und in Abbildung 5.6 schematisch dargestellt.

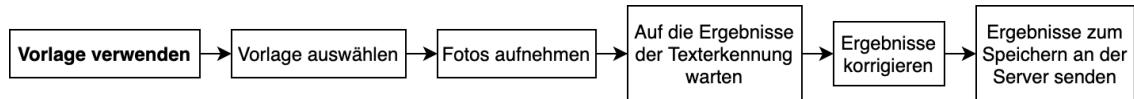


Abbildung 5.5: Kompaktes Flussdiagramm zum Verwenden einer Scan-Vorlage

1. Die Seiten der Klausur müssen dazu in derselben Reihenfolge, wie in der Scan-Vorlage fotografiert werden. Anschließend werden die Regionen der Scan-Vorlage auf das jeweilige Bild angewendet. Bei dem Prozess entstehen Bildausschnitte, in denen die relevanten Informationen enthalten sind.
2. Die entstandenen Bildausschnitte werden seitenweise in die Texterkennung überführt.
3. Die Inhalte der Ausschnitte werden mithilfe von OCR digitalisiert. Der Datentyp der zugehörigen Region aus der Scan-Vorlage nimmt nun Einfluss auf das Ergebnis. Durch ihn wird während der Worterkennungsphase in der Texterkennung eine Liste an vordefinierten möglichen Ergebnissen eingespeist. Diese Liste hat Vorrang vor dem sogenannten Standard-Wörterbuch und verbessert dadurch die Ergebnisse. Ein Beispiel für solch eine Liste können alle möglichen Noten sein. Es ist auch vorstellbar, dass alle Seminargruppen oder Namen von Personen, die an der Klausur teilgenommen haben, dort verwendet werden.
4. Die Ergebnisse werden im Anschluss an den Server gesendet. Zuvor kann der Benutzer jedoch die Ergebnisse anpassen bzw. korrigieren, falls es bei der Texterkennung zu Fehlern kam. Auch die Kontrollmechanismen finden vor dem Hochladen zum Server Anwendung und geben dem Nutzer durch Fehlermeldungen Hinweise.

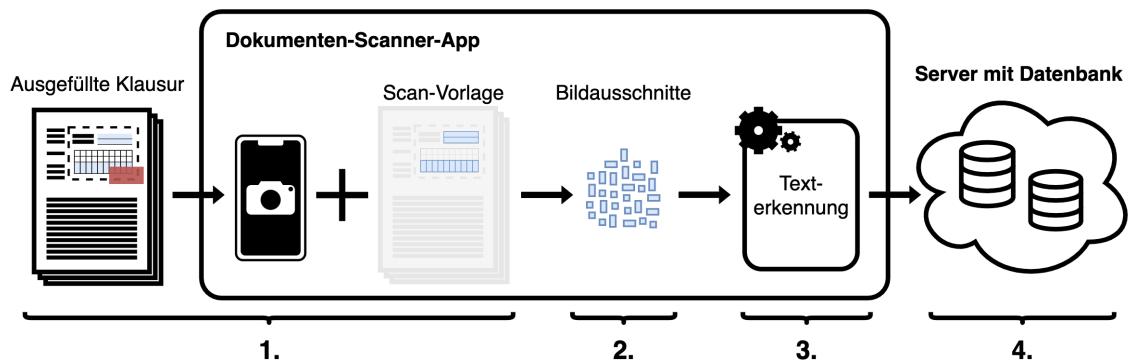


Abbildung 5.6: Schematische Darstellung des Vorgangs beim Verwenden einer Scan-Vorlage

5.3 Erweiterung des Konzepts

Weiter ist eine weitere Korrektur sinnvoll. Angenommen ein Student verschreibt sich bei seiner Seminargruppe und vergisst ein Zeichen. Die Texterkennung erkennt zwar jedes Zeichen, jedoch nicht die korrekte Bezeichnung der Seminargruppe. Deshalb sollte nach der Texterkennung die falsche Bezeichnung durch diejenige Seminargruppe ersetzt werden, die die größte Übereinstimmung mit der erkannten Bezeichnung hat. Bei diesem Vorgehen ist zu beachten, dass nur die Seminargruppen verwendet werden sollten, die tatsächlich die Klausur geschrieben haben. Ähnliches ist auch für die Namen oder Matrikelnummern der Studenten möglich.

Allerdings entstehen auch durch das OCR Fehler. Aus diesem Grund wird neben jedem Ergebnis die sogenannte confidence angezeigt. Diese sagt aus, wie sicher die Texterkennung des Ergebnisses ist. Jedoch kann es auch hier falsche Positiv-Ergebnisse geben. Deshalb ist auch die confidence kein perfekter Indikator für die Richtigkeit des Ergebnisses. Deswegen gibt es, wie in Punkt 4 beschrieben, die Möglichkeit, die Ergebnisse noch einmal zu überprüfen und zu korrigieren, bevor sie an den Server zum Abspeichern gesendet werden. Auch die Kontrollmechanismen reagieren auf die Änderungen der Ergebnisse, sodass Fehlermeldungen behoben werden können.

6 Entwicklung des ersten Prototyps

Dieses Kapitel beschreibt einzelne Phasen der Entwicklung der iOS-App in einer ähnlichen Reihenfolge, wie im Wasserfall-Model zur Verwaltung der Entwicklung großer Softwaresysteme nach Winston W. Royce [7] (siehe Abbildung A.1).

6.1 Anforderungsplanung

Vor Beginn des Praktikums wurde diskutiert und kalkuliert, welches Thema aus dem Projekt Memo Space für ein zwölfwöchiges Praktikum geeignet ist. Dabei entstand eine Art Durchführbarkeits- / Machbarkeitsstudie, welche zur Entscheidung führte, das Dokumenten-Scanner-Softwaresystem zu entwickeln. Tobias Kallauke soll aufgrund seiner Kenntnisse und Erfahrungen ein Backend mit entsprechenden Schnittstellen und eine Android-App erstellen, während der Verfasser eine iOS-App programmiert. Die iOS-Anwendung soll den Anforderungen, die im gleichnamigen Kapitel 4 zu finden sind, erfüllen. Für weitere Details über den Server, dessen Architektur und die Android-App siehe im Praktikumsbericht von Tobias Kallauke.

6.2 Analyse und Definition

Die Aufgaben bzw. Ziele dieser Phase sind:

1. die Auseinandersetzung mit der Problemstellung und den Anforderungen,
2. die Durchführung einer Problemanalyse,
3. die Entwicklung von Ideen und eines genauen Konzepts der iOS-App sowie
4. das Entwickeln eines ersten minimalen Prototyps als Machbarkeitsnachweis.

Bei der Entstehung des Konzepts, welches im Kapitel 5 zu finden ist, spielen die Dokumentationen der Frameworks Vision¹, VisionKit² und PhotoKit³ von Apple eine entscheidende Rolle. Aus ihnen geht hervor, welche Funktionen der App noch zu entwickeln und welche in Frameworks schon vorhanden sind. Z. B. ist das Erkennen und das Geradeziehen von Dokumenten in Echtzeit sowie die Texterkennung auf Bildern in Vision und VisionKit enthalten. Bei der Entwicklung des ersten minimalen Prototyps half eine Beispiel-App von Apple [1], die das Erkennen von Objekten in Standbildern mithilfe der genannten Frameworks umsetzt.

¹ Dokumentation von Vision - <https://developer.apple.com/documentation/vision>

² Dokumentation von VisionKit - <https://developer.apple.com/documentation/visionkit>

³ Dokumentation von PhotoKit - <https://developer.apple.com/documentation/photokit>

Durch die Auseinandersetzung mit den Bibliotheken konnte festgestellt werden, dass die zu entwickelnde App nur Geräte ab der Betriebssystem-Version iOS 13.0 unterstützen wird. Grund dafür sind die Apple Frameworks SwiftUI und VisionKit.

SwiftUI bietet die Möglichkeit, Benutzeroberflächen für alle Apple-Plattformen in der Programmiersprache Swift zu erstellen. Die deklarative Swift-Syntax ist einfach zu lesen und schnell zu schreiben, sodass es möglich ist, die App in wenigen Wochen für iPhone und iPad zu entwickeln. Als Alternative gibt es UIKit⁴, welches unter allen iOS Versionen funktioniert. Dieses Framework ist allerdings nicht deklarativ, sodass Views sowohl im Code, als auch in Interface-Dateien getrennt voneinander erstellt und konfiguriert werden müssen [8]. Dadurch dauert die Entwicklung einer App im Gegensatz zu SwiftUI deutlich länger, wie man auch in dem Video⁵ von Paul Hudson, einem in der Swift-Community sehr bekannten Programmierer und Autor, sieht. Sehr ähnliche Erfahrung hat der Verfasser vor dem Praktikum in seiner Freizeit mit den beiden Frameworks gesammelt.

VisionKit dagegen ist das Framework zum Scannen der Dokumente. Auch hierfür existiert eine Alternative, welche auch auf älteren iOS Versionen funktioniert. Allerdings unterstützt das Framework WeScan⁶ noch kein stapelweises Scannen. Das bedeutet, dass man immer nur ein Foto machen kann, welches erst abgespeichert werden muss, bevor man das nächste aufnehmen kann. Das ist für den Benutzer umständlich und zeitintensiv. Zu Informationen anderer verwendeter Frameworks siehe im Kapitel 6 Abschnitt 6.5 und Abschnitt 6.5.

Zur Projektplanung und zum Projektmanagement wurde das Scrum-Konzept⁷, welches sich für agile Softwareentwicklung anbietet, verwendet. Als Versionsverwaltung der App-Software wurde Git⁸ in Kombination mit GitHub Issues⁹ und GitHub Project Boards als Planungstools benutzt. Dadurch konnte der Verfasser jedem sogenannten Sprint selbstständig Aufgaben zuordnen und den Fortschritt nachvollziehen. Informationen über den Server und der Android-App zu diesem Thema sind im Praktikumsbericht von Tobias Kallauke zu finden.

6.3 Grundlagen

In den folgenden zwei Absätzen werden wichtige Konzepte erläutert und Grundwissen vermittelt, die im nachfolgendem Abschnitt 6.4 nochmals thematisiert werden.

⁴ Dokumentation von UIKit - <https://developer.apple.com/documentation/uikit>

⁵ SwiftUI vs UIKit – Comparison of building the same app in each framework - <https://www.youtube.com/watch?v=qk2y-TiLDZo>

⁶ WeScan GitHub Repository - <https://github.com/WeTransfer/WeScan>

⁷ Der Scrum Guide™ - <https://www.scrumguides.org/scrum-guide.html>

⁸ Git Internetseite - <https://git-scm.com/>

⁹ Mastering Issues - <https://guides.github.com/features/issues/>

Model-View-ViewModel Model-View-ViewModel (MVVM) ist ein Software-Architekturmuster. Unter einem Software-Architekturmuster oder auch Entwurfsmuster ist eine bewährte Lösungsvorlage für wiederkehrende Entwurfsprobleme zu verstehen. MVVM erleichtert die Trennung der Entwicklung der grafischen Benutzeroberfläche (der View in Abbildung 6.1) von der Entwicklung der Geschäftslogik oder der Back-End-Logik (dem Model). So ist die View nicht von einer bestimmten Model-Plattform abhängig. MVVM verwendet das Konzept eines Schichtmodells¹⁰ und ist eine abstrakte Darstellung einer Benutzeroberfläche in Form einer Datenstruktur. Diese Struktur enthält Daten, die auf der Benutzeroberfläche angezeigt werden sollen und Anweisungen, die auf der Benutzeroberfläche aufgerufen werden können. Dieses sogenannte ViewModel besitzt keine direkte Verbindung zur View, wie es sonst bei anderen Entwurfsmustern üblich ist, um Daten auf der Benutzeroberfläche anzuzeigen. Stattdessen verwendet eine MVVM-View eine Bindungsfunktion (data binging) zur bidirektionalen Zuordnung von Daten aus dem ViewModel zu den jeweiligen Objekten auf der View. Z. B. bildet eine Liste von Zahlen im ViewModel die Einträge in einem Dropdown-Menü auf der View ab. Aber auch das Binden von Daten aus dem Model an Benutzereingaben durch Maus, Tastatur oder Touch-Screens ist möglich. Beispielsweise kann ein Mausklick eine Anweisung in dem ViewModel auslösen. Die Anweisung besorgt Daten aus dem Model, wodurch die View durch data binding aktualisiert wird. [6] [4] [3]

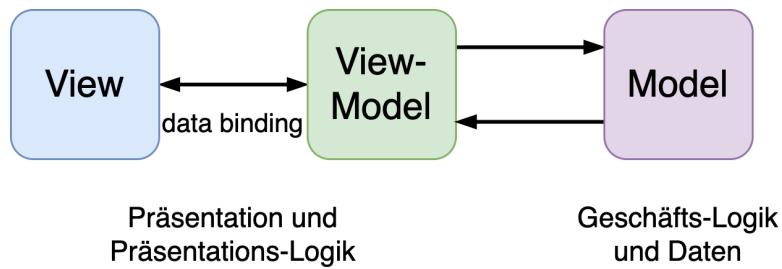


Abbildung 6.1: Schematische Darstellung des Model-View-ViewModel Konzepts

Redux.js Redux.js ist eine Bibliothek¹¹ für die Programmiersprache JavaScript und stellt einen sogenannten vorhersehbaren Zustandscontainer¹² für JavaScript-Anwendungen bereit. In diesem Container wird der Zustand der gesamten Anwendung in einem sogenannten Objektbaum gespeichert. Dieser Baum ist mit der Ordner-Struktur von Windows vergleichbar. Es gibt einen Wurzel-Ordner, in dem wiederum Ordner oder aber auch Dateien enthalten sind. Im Zustandscontainer bezeichnet man die Ordner als States bzw. auf deutsch Zustände. Die Dateien sind einfache Daten, die unter anderem auf der Benutzeroberfläche angezeigt werden sollen. Diese Aufteilung in die States ist dazu gedacht, jeder View nur die wirklich benötigten Daten bereitzustellen. Diese Struktur erleichtert das Testen oder Untersuchen der Anwendung und ermöglicht es, durch

¹⁰ Geschichtete Anwendung - [https://docs.microsoft.com/en-us/previous-versions/msp-n-p/ff650258\(v=3dpanp.10\)](https://docs.microsoft.com/en-us/previous-versions/msp-n-p/ff650258(v=3dpanp.10))

¹¹ Redux.js Github Repository - <https://github.com/reduxjs/redux>

¹² Redux Core Concepts - <https://redux.js.org/introduction/core-concepts>

Hinzufügen eines neuen States den aktuellen Entwicklungsstand der Anwendung beizubehalten und dadurch den Entwicklungsprozess zu beschleunigen.

Ein weiteres wichtiges Prinzip¹³ von Redux ist, dass die Daten in den States schreibgeschützt sind. Die einzige Möglichkeit den Zustand bzw. die Daten zu ändern, besteht darin, eine Aktion auszulösen. Die Aktion beschreibt, wie der State und dessen Daten sich ändert. Dadurch wird sichergestellt, dass weder die Views noch Netzwerk-Rückrufe jemals direkt den Zustand ändern können. Stattdessen bringen sie nur die Absicht zum Ausdruck, den State zu verändern und lösen eine Aktion aus, die die Änderung für sie vornimmt. Da alle Aktionen zentralisiert sind und eine nach der anderen in einer strikten Reihenfolge erfolgen, gibt es weniger Programmfehlerquellen.

6.4 Entwurf und Design

In der Entwurfsphase wird normalerweise über die Datenhaltung, die Verteilung des Software-Systems im Netz und die Benutzeroberfläche entschieden¹⁴. Jedoch standen diese Grundsatzentscheidungen schon zu Beginn des Praktikums durch die Kenntnisse von Tobias Kallauke und dem Verfasser fest. Zur Datenspeicherung verwendet der Server PostgreSQL¹⁵, eine relationale Datenbank. Zum Speichern der Bilder wird die Server-Verzeichnisstruktur genutzt. In der App hingegen werden keine Daten persistent gespeichert. Durch die Aufteilung von App und Server handelt es sich um eine sogenannte Client-Server-Architektur und für die Benutzeroberfläche der iOS-App wird, wie schon erwähnt, das deklarative SwiftUI verwendet. Weitere Informationen zum Entwurf der Server-Architektur und der Android-App befinden sich im Praktikumsbericht von Tobias Kallauke.

Zur Definition einzelner Teil-Workflows, die zusammenhängende Aufgaben umfassen, wurden Flussdiagramme und schriftliche Zustandsautomaten angefertigt (siehe Anhang B). Diese Modelle halfen dabei Views zu entwickeln und deren Design festzulegen. Das Aussehen der App sollten sich jedoch im Laufe der Zeit noch ändern, da erst einmal die Umsetzung des Konzepts im Vordergrund stand.

Ausgehend vom Design und den Teil-Workflows entstand ein grobes Konzept zur Handhabung der Daten innerhalb der App. Da SwiftUI das Entwurfsmuster MVVM umsetzt, wird eine Struktur für das ViewModel und für den Datenfluss der asynchronen Server-Rückrufe benötigt. Die Struktur-Entwicklung erwies sich bereits während des ersten Prototyps als problematisch. Denn auch bei steigender Komplexität sollte das ViewModel noch übersichtlich bleiben, um eine schnelle Weiterentwicklung zu gewährleisten. Bei der Suche nach einer besseren Lösung wurde die JavaScript-Bibliothek Redux.js zur

¹³ Redux Three Principles - <https://redux.js.org/introduction/three-principles>

¹⁴ Vorlesung 9, Folie 27 aus Softwaretechnik Grundlagen von Prof. Dr.-Ing. Wilfried Schubert (2019) - https://www.staff.hs-mittweida.de/~wschub/intranet/ss19/Fach_SWT/Fach_SWT_Zeitplan.htm

¹⁵ PostgreSQL Internetseite - <https://www.postgresql.org/>

Verwaltung von Zustandsinformationen in Webanwendungen gefunden.

Das Konzept von Redux hilft komplexe Views mit vielen Daten schnell und einfach zu implementieren. Somit werden Server-Antworten und zwischengespeicherte Daten sowie lokal erstellte Daten, die noch nicht auf dem Server gespeichert wurden, strukturiert und zentral abgelegt. Das erleichtert nicht nur das Wiederverwenden von Daten, sondern reduziert die Wiederholung von Server-Aufrufen. Zudem ermöglicht Redux durch die modulare Aufteilung des State-Containers eine schnelle Weiterentwicklung der App, trotz steigender Komplexität. Zudem stellten die Vorteile hinsichtlich des Testens und Untersuchens der App, das schnelle Auffinden von Fehlern in Aussicht. Es erschien nun möglich, trotz begrenzter Praktikumszeit, den Prototyp möglichst fehlerfrei und weit zu entwickeln.

Daher lag es nah die wesentlichen Konzepte von Redux umzusetzen und einen Redux ähnlichen State Container, als ViewModel zu implementieren. Aus der Definition von Teil-Workflows sollte der Container oder auch Store genannt mindestens 5 States haben:

- für Authentifizierung sowie Registrierung,
- für das Anlegen von Vorlagen, um Zwischenergebnisse zu speichern,
- für das Ausführen von Server-Aufrufen zum Speichern und Abrufen von Vorlagen,
- für die Steuerung des Workflows bzw. den stellvertretenden Views sowie
- für sonstige Daten, die sehr häufig verwendet werden.

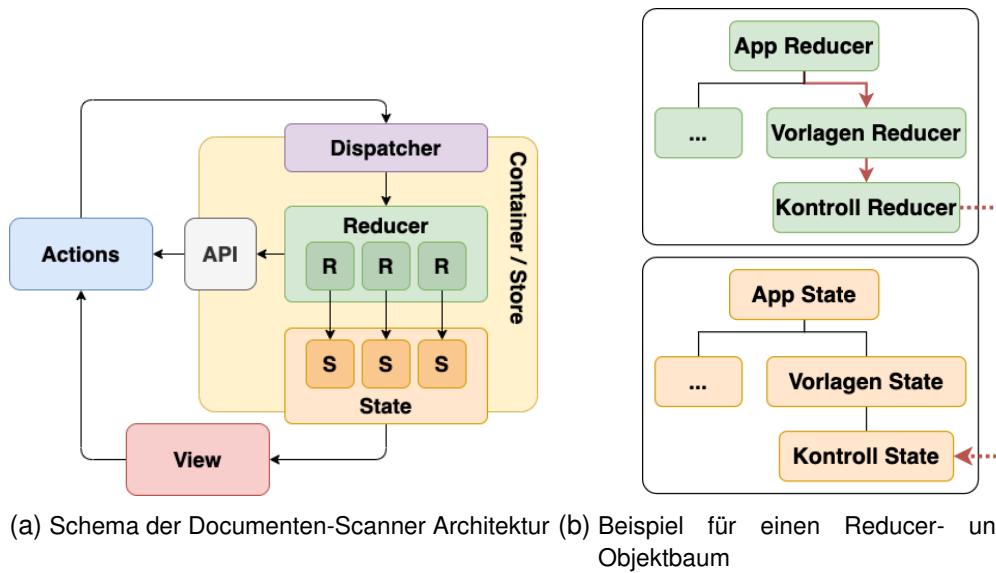


Abbildung 6.2: Architektur Schemata

Das Schema 6.2a zeigt, die in der App umgesetzte Redux ähnliche Struktur. Über eine View (rot) können Aktionen (blau) aufgerufen werden. Diese gelangt zuerst in einen sogenannten Dispatcher (lila), welcher als Verteiler dient. Anhand der Art einer Aktion,

wird der entsprechende Reducer (grün) vom Dispatcher beauftragt, die Aktion auszuführen. Ein Reducer ist für genau eine Art von Aktionen zuständig. Allerdings ist, wie in der Abbildung 6.2b zu sehen ist, eine Schachtelung von Reducern (grün) möglich. Dies bedeutet auch, dass Aktions-Arten ineinander geschachtelt werden können. Das hat den Hintergrund, dass die States wie ein Objektbaum aufgebaut sind. So hat jeder State seinen eigenen Reducer, was für die oben erwähnte Modularität sorgt. Möchte man beispielsweise eine Änderung im Kontroll-Mechanismen State (siehe Kontroll-State in 6.2b) vornehmen, muss die Haupt-Aktion in einer Kontroll-Mechanismen Reducer Aktion gekapselt werden und diese wiederum in einer Vorlagen Reducer Aktion. Zum Ausführen der Aktion, werden dann die entsprechenden Reducer die Aktions-Kapselung von außen nach innen auflösen. Nachdem die Aktion ausgeführt und eine State-Änderung herbeigeführt wurde, aktualisiert sich durch das data binding von MVVM die View. Genauer werden alle Views, die eine Bindung zu dem jeweiligen Datum haben, über die Änderung benachrichtigt und daraufhin aktualisieren diese sich.

Ein weiterer wichtiger Punkt sind die Server-Aufrufe. Für diese gibt es einen eigenen Reducer und eine besondere Schnittstelle. Diese Schnittstelle ist in der Abbildung 6.2a grau markiert und mit API beschriftet. Diese hat die Besonderheit selbst Aktionen an den Container zu senden. Beispielsweise löst ein Knopfdruck in einer View die Aktion aus, um alle Vorlagen vom Server zu laden. Diese Aktion gelangt über den vorgesehenen Reducer zu der API-Schnittstelle. Dort wird ein entsprechender Server-Aufruf gestartet und auf den Server-Rückruf gewartet. Sobald die Antwort des Servers angekommen ist, wird diese mit Hilfe einer Aktion zurück zum Container gesendet. Dort wird sich wieder ein entsprechender Reducer um eventuelle Fehlerbehandlung oder das Ablegen der heruntergeladenen Vorlagen kümmern.

6.5 Implementierung

Zur Realisierung der entworfenen Systemkomponenten wurde ausschließlich, die von Apple entwickelte IDE Xcode¹⁶, verwendet. Diese stellt Geräte-Simulatoren zur Verfügung, auf denen die App getestet werden kann. Die Simulatoren bieten unter anderem auch die Möglichkeit an, die Anwendung schnell auf verschiedenen iOS bzw. iPadOS Versionen zu testen. Auch ist das Testen der App auf echten Geräten durch Xcode möglich, was bei der Entwicklung unumgänglich war. Grund dafür ist die Benutzung der Kamera, die bei den Simulatoren zu gewollten Abstürzen führt, da diese keinen Zugriff auf eine Kamerasystem besitzen.

Die App wurde ausschließlich in der Programmiersprache Swift geschrieben. Zusätzlich wurden die Frameworks SwiftUI, Vision und VisionKit von Apple sowie das Framework Kingfisher¹⁷, zum Downloaden und Cachen von Bildern verwendet. Das Vision-

¹⁶ Xcode Internetseite - <https://developer.apple.com/xcode/>

¹⁷ Kingfisher GitHub Repository - <https://github.com/onevcat/Kingfisher>

Framework führt Erkennung von Gesichts- und Gesichtsmarkierungen, Texterkennung, Barcode-Erkennung, Bildregistrierung und allgemeine Merkmalsverfolgung durch. Jedoch wurde in der App nur die integrierte Texterkennung benutzt. Es ist aber anzunehmen, das Algorithmen aus VisionKit von Vision zur Dokumenten-Erkennung verwendet. Für mehr Informationen über die Frameworks SwiftUI und VisionKit siehe in Abschnitt 6.2.

Zum Einrichten und Testen des Backends wurde außerdem Visual Studio Community bzw. Visual Studio Code¹⁸ mit dem REST Client Plugin¹⁹ verwendet. Das Backend konnte als lokaler Server auf dem Entwicklungskomputer gestartet und über die IP-Adresse 0.0.0.0 im lokalen Netzwerk aufgerufen werden. Für die Verwaltung der PostgreSQL-Datenbank genügte die Anwendung PgAdmin4²⁰. Damit ist es möglich einzelne Einträge oder auch ganze Tabellen der Datenbank zu bearbeiten.

6.5.1 Registrieren und Anmelden

Die technische Implementierung der Registrierung und Anmeldung beinhaltet typische Charakteristiken von Anmelde- und Registrier-Formularen. Die E-Mail-Textfelder besitzen, wie auch das Passwort-Feld der Anmeldung eine AutoFill-Funktion. Dadurch kann das iOS-Gerät Registrier- und Anmelde-Daten vorschlagen und automatisch in die entsprechenden Felder einsetzen. Dies sieht man im dritten Bild 6.3c anhand des „Passwörter“-Knopfes über der Tastatur. Weiter sind alle Passwort-Felder gesichert. Das bedeutet, dass der Inhalt zum Schutz der Privatsphäre standardmäßig ausgeblendet und mit Punkten ersetzt wird. Durch das drücken des Auges rechts des Textfeldes kann der Inhalt jedoch eingesehen werden. Im mittlerem Bild 6.3b sind wegen der Sicherheitsstandards von iOS, bei einer Aufnahme des Bildschirms nicht einmal die Ersatz-Punkte des ersten Passwort-Textfelds zu sehen. Zusätzlich validieren alle Text-Felder ihren Inhalt mithilfe von regulären Ausdrücken. Beispielsweise muss ein Passwort aus 8 Zeichen bestehen und zwei von den drei folgenden Eigenschaften erfüllen:

1. Es ist mindestens ein Sonderzeichen enthalten.
2. Es ist mindestens ein Großbuchstabe enthalten.
3. Es ist mindestens eine Zahl enthalten.

Die Validierung ist jedoch nicht standardmäßig, sondern wurde selbst entwickelt.

¹⁸ Visual Studio Internetseite - <https://visualstudio.microsoft.com/de/>

¹⁹ REST Client Repository - <https://github.com/Huachao/vscode-restclient>

²⁰ PgAdmin Internetseite - <https://www.pgadmin.org/>

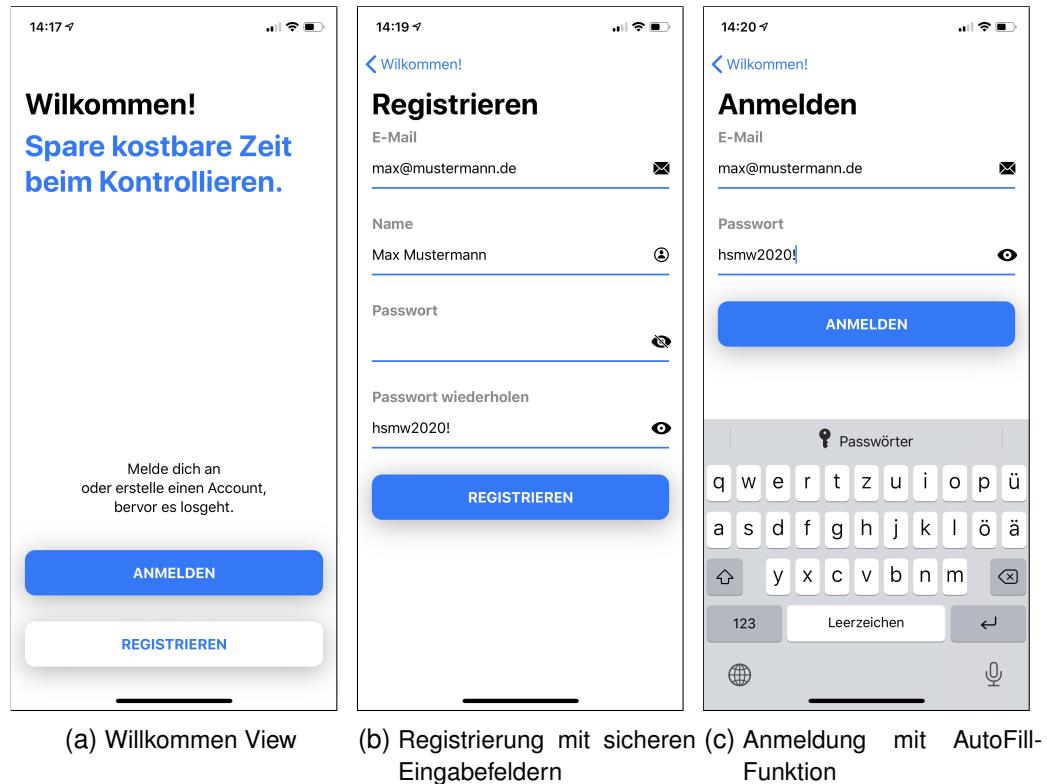


Abbildung 6.3: Willkommen-, Registrier- und Anmelde-View

6.5.2 Scan-Vorlagen erstellen und speichern

Bei der Implementierung des ersten Arbeitsschritts der Scan-Vorlagen wurde das Flussdiagramm B.1, welches im Anhang B zu finden sind, benutzt.

Scan-Vorlage erstellen

Im ersten Schritt sind ein Name und weitere Informationen, zu der Vorlage an zugegeben (siehe 6.4a), um im Anschluss die Fotos aufzunehmen. Bei der Kamera-View (siehe 6.4b) handelt es sich um die Scan-View des Frameworks VisionKit. Mithilfe von Kantenerkennung und anderen Algorithmen, die Tobias Kallauke in seinem Bericht beschreibt, kann dass Dokument sobald es erkannt ist, automatisch fotografiert werden. In Echtzeit wird das erkannte Dokument aus dem Bild ausgeschnitten und gerade gezogen. Ein manuelles Auslösen des Fotos und Anpassen der Dokumenten-Kanten im Bild, ist ebenfalls möglich. Des Weiterem werden alle erstellten Bilder zu einer Gruppe gesammelt. Diese können vor dem Abspeichern angesehen und nochmal bearbeitet werden. Die Abbildung 3.1 entstand durch diesen Prozess.

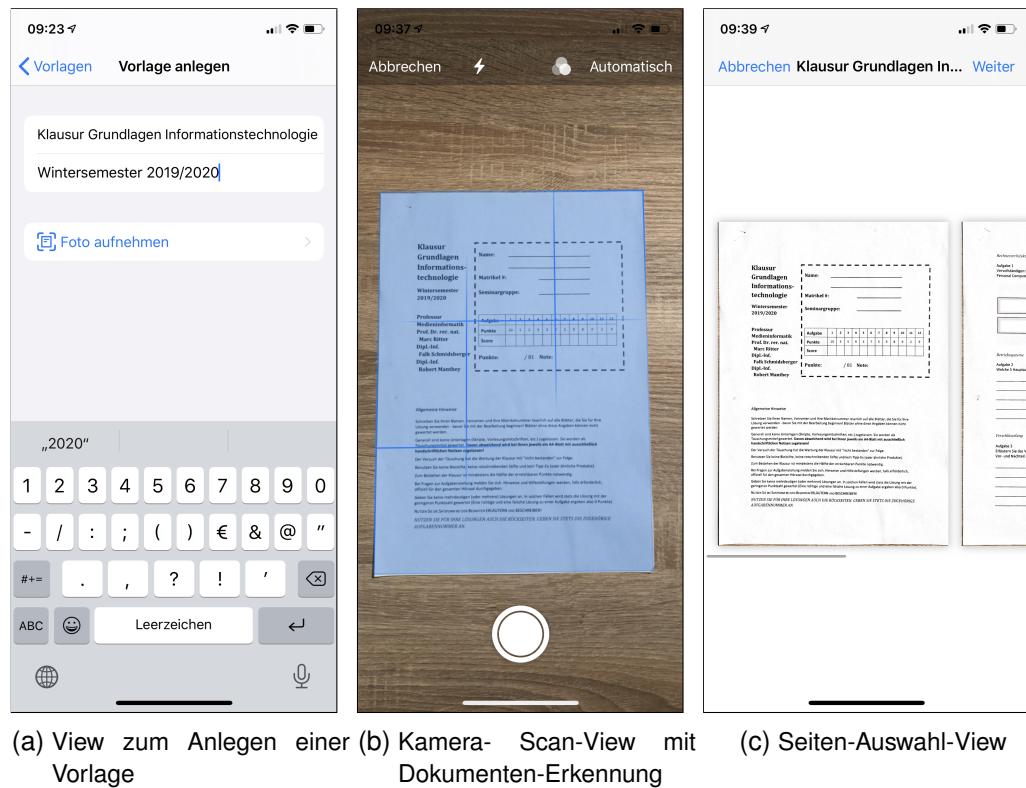


Abbildung 6.4: Die ersten Views zur Erstellung einer Scan-Vorlage

Regionen erstellen

Im nächsten Schritt sind die Regionen auf den Dokumenten-Seiten zu markieren, deren Inhalt beim Einscannen digitalisiert werden soll. Zu erst wird die gewünschte Seite aus einer Übersicht (siehe 6.4c) ausgewählt. Anschließend ist eine Vorschau der Seite mit alle eingetragenen Regionen zu sehen (siehe 6.6a bzw. 6.6b). Über einen Button können weitere Regionen hinzugefügt werden. Dazu legt man einen Namen (siehe 6.5a) und einen Datentyp (siehe 6.5b) fest. Die Datentypen sind für die Texterkennung und Erstellung der Tabelle für die Notenfreigabe wichtig. Dazu später noch mehr. Wenn die Eigenschaften der Region festgelegt sind, muss diese noch auf dem Bild markiert werden. Dazu zieht man mit einem Finger ein Rechteck in einer beliebigen Größe ein (siehe 6.5c). Die markierte Region kann anschließend noch bewegt oder neu gemacht werden. Um kleinere Regionen präzise zu markieren kann mit einer Zwei-Finger-Geste auch an die Seite heran bzw. auch heraus gezoomt werden. Dieses Vorgehen muss dann für alle nötigen Regionen auf den jeweiligen Seiten wiederholt werden.

Kontrollmechanismus erstellen

Zum Schluss können noch die sogenannte Kontrollmechanismen erstellt werden. Diese Funktion ist allerdings noch nicht sehr weit fortgeschritten und beinhaltet aktuell nur den Kontrollmechanismen-Typ zum Vergleichen von Regionen, wie im Kapitel 5 beschrieben

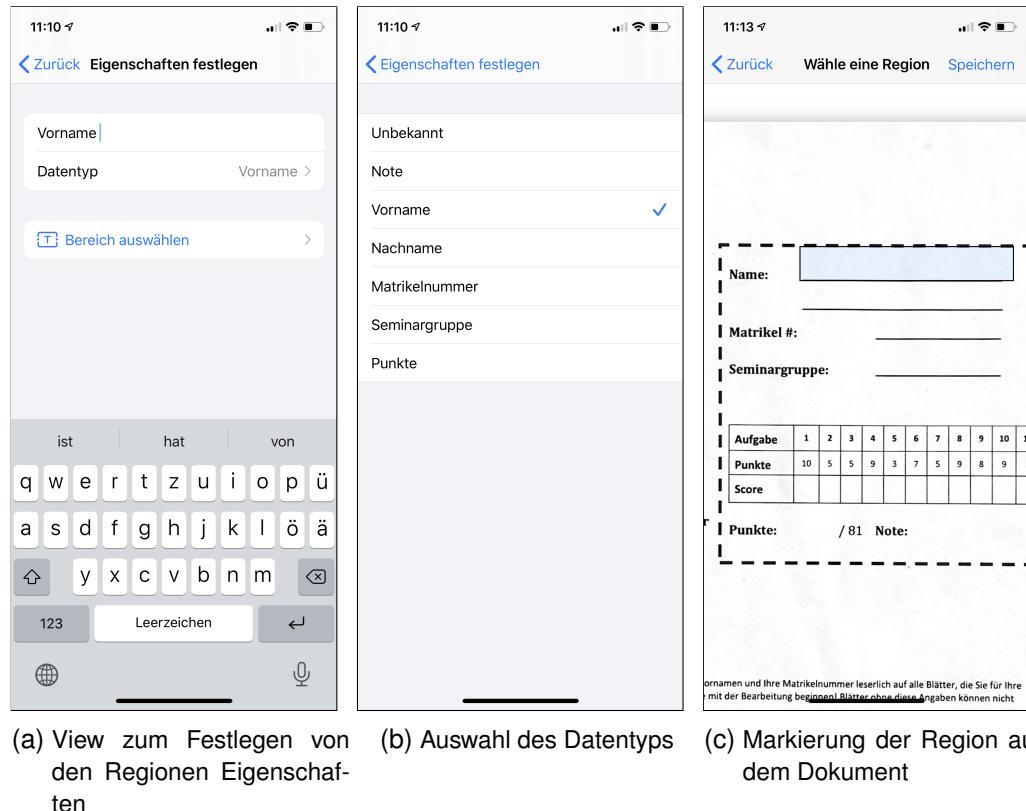


Abbildung 6.5: Views zur Erstellung von Regionen

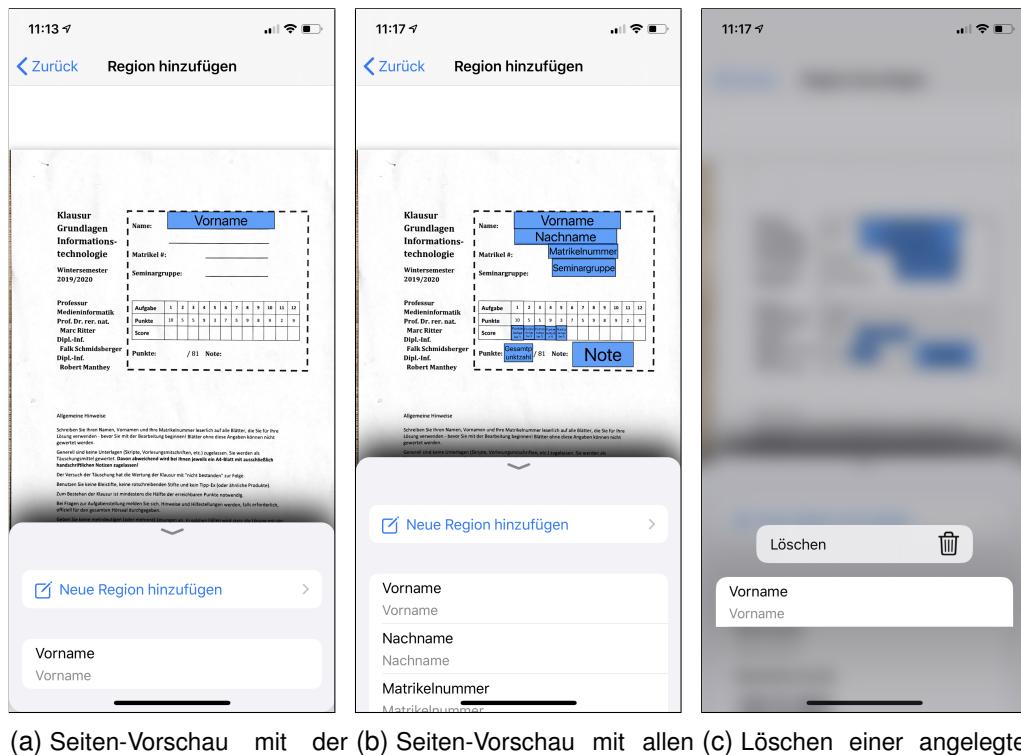
ist. Beim Erstellen einer Kontrolle dieses Typs wählt man zwei Regionen aus (siehe Abbildung 6.7b), deren Inhalt nach der Texterkennung auf Gleichheit überprüft wird. Zur Auswahl stehen alle vorher angelegten Regionen, wie in Abbildung 6.7c zu sehen ist. In Abbildung 6.7a ist eine Übersicht aller angelegten Kontrollen zu der Scan-Vorlage zu sehen sowie ein Button zum Speichern der Vorlage.

Scan-Vorlage speichern

Wenn eine Vorlage gespeichert werden soll, passiert das in mehreren Schritten. Dazu wird die entsprechenden Server-Schnittstellen aufgerufen und auf die Server-Rückrufe gewartet. Für genauere Details zur API des Servers siehe im Praktikumsbericht von Tobias Kallauke.

Im folgenden Abschnitt, ist beschrieben, in welcher Reihenfolge das Hochladen einer Vorlage zum Server geschieht. Dabei werden mögliche Fehler von Seiten des Server, der Internetverbindung und des Clients ignoriert. Jedoch ist im aktuellen Stand der App das Abfangen des Fehler schon integriert, die Fehlerbehandlung aber noch nicht.

1. Der Name und die Beschreibung der Vorlage sowie die Liste an Kontrollmechanismen wird gesendet. In der Antwort des Servers befindet sich die Vorlagen-ID , die in den nächsten Schritten benötigt wird.



(a) Seiten-Vorschau mit der ersten Region (b) Seiten-Vorschau mit allen Regionen (c) Löschen einer angelegten Region

Abbildung 6.6: Seiten-Vorschau-View

2. Die Bilder der Seiten werden gesendet. Als Antwort zu jedem Bild wird der Pfad zurück geschickt, wo das Bild vom Server gespeichert wurde. Dieser Pfad wird im nächsten Schritt benötigt.
3. Die Seiten mit einer Seiten-Nummer und dem Pfad zu dem Bild, wird mit der Vorlagen-ID gesendet. Als Antwort wird eine Seiten-ID zurückgegeben, die im nächsten Schritt verwendet wird.
4. Die Regionen der Seiten werden gesendet. Eine Region hat X- und Y-Koordinaten. Diese repräsentieren den Abstand vom Bildursprung, der sich in jedem Bild oben links befindet. Außerdem besitzt eine Region noch eine Höhe und Breite sowie einen Namen und einen Datentyp (6.5a und 6.5c). Die Koordinaten sowie die Höhe und Breite sind alle in Pixel angegeben. Zusammen mit der Seiten-ID wird das Attribute an den Server gesendet.

6.5.3 Scan-Vorlage verwenden

Bei der Implementierung dieses Arbeitsschritts wurde das Flussdiagramm B.2, welches im Anhang B zu finden sind, benutzt.

Um eine Klausur zu digitalisieren, muss die vorher erstellte Scan-Vorlage ausgewählt werden (6.8a). Danach bietet die Benutzeroberfläche über einen Button die Möglichkeit an, eine Klausur ein zu scannen (6.8b). Bevor die Bilder aufgenommen werden können,

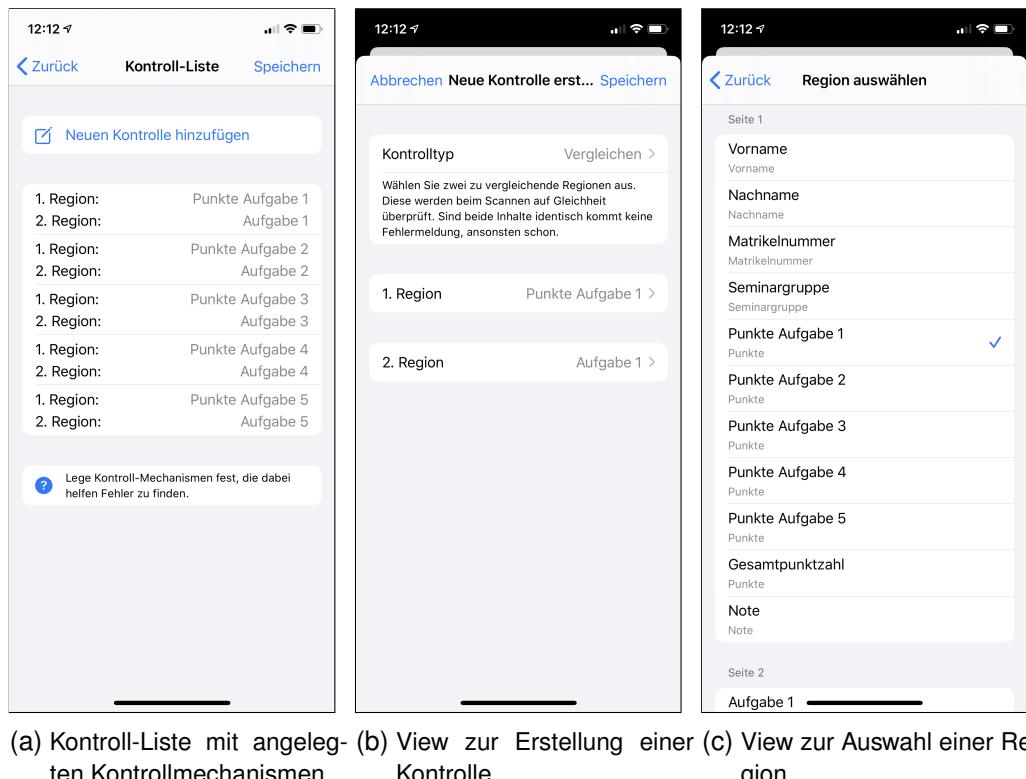


Abbildung 6.7: Views zur Erstellung von Kontrollmechanismen

erscheint ein Dialogfenster (6.9a), in dem der Benutzer sich entscheiden muss, welche OCR-Engine benutzt werden soll. Aktuell gibt es zwei Möglichkeiten.

- Die OCR-Engine des Vision Frameworks von Apple, welche direkt auf dem Gerät und ohne eine Internetverbindung die Texterkennung durchführt,
- oder die OCR-Engine Tesseract, die über eine Schnittstelle des Servers aufzurufen ist.

Nach der Auswahl öffnet sich die Kamera und die Bilder können aufgenommen werden. Wichtig dabei ist, dass die Seiten beim Fotografiert die selbe Reihenfolge, wie in der Scan-Vorlage haben. Auch muss die Anzahl der eingescannten Seiten mit denen in der Vorlage übereinstimmen. Bei der Implementierung der Scan-View wurde auf die Selbe zurückgegriffen, wie sie auch beim Vorlagen Erstellen zu finden ist (siehe 6.4b). Grundsätzlich konnten viele Komponenten der gesamten Benutzeroberfläche dank SwiftUI immer wieder verwendet werden. Durch die Scan-View wird die Seiten aus dem Bild ausgeschnitten, geglättet und der Kontrast des Bildes wird ebenfalls leicht angepasst, damit Buchstaben leichter zu erkennen sind und Schatten bzw. Falten verschwinden.

Nach dem Einscannen der Klausur-Seiten beginnt der Prozess der Texterkennung. Durch Aktivitätsanzeigen wird dem Benutzer mitgeteilt, dass die Anwendung gerade beschäftigt ist. Die Abläufe bei der online OCR-Engine sind jedoch leicht anders, als die, des Vision Frameworks.

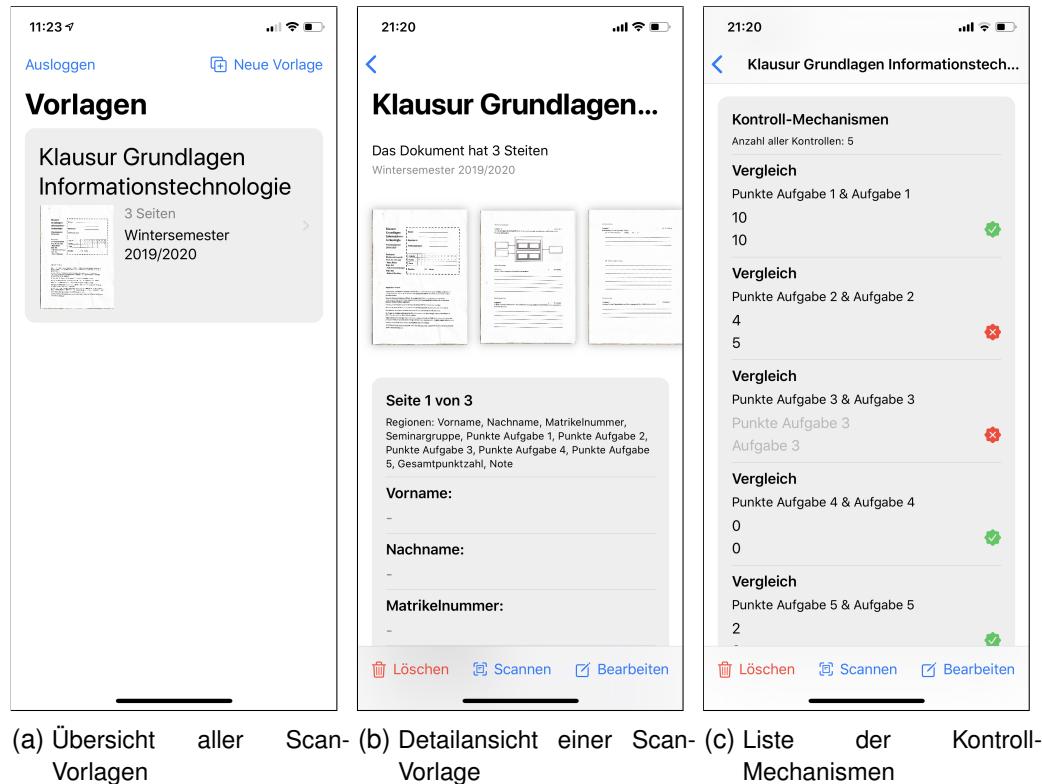


Abbildung 6.8: Listen- und Detail-Ansicht von Scan-Vorlagen

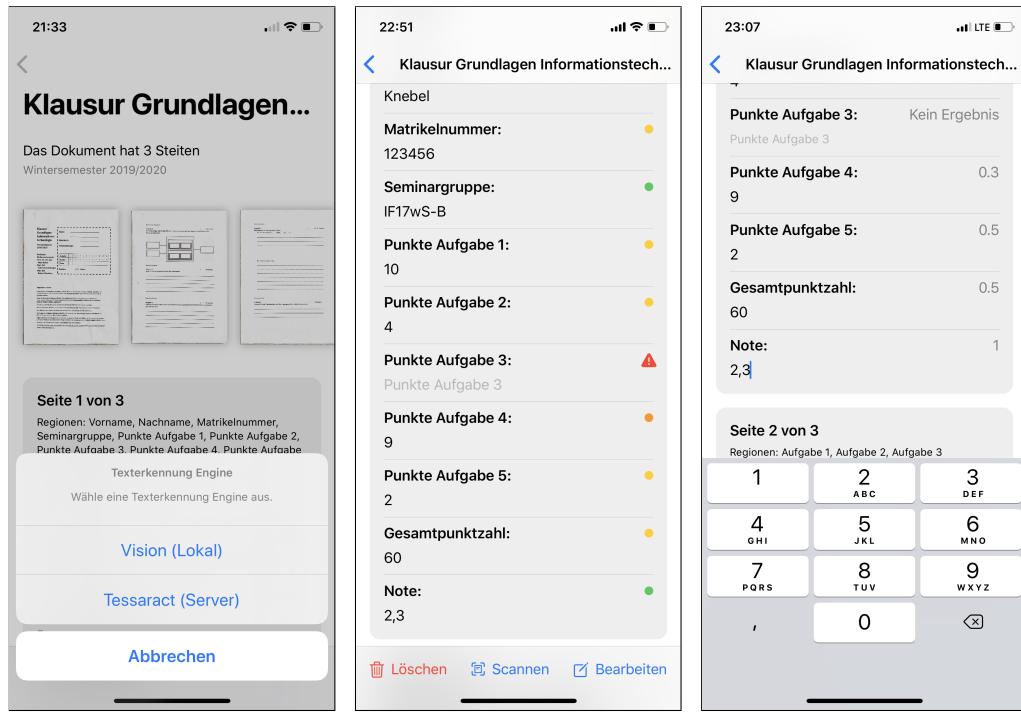
Ablauf mit Vision

Bei der Verwendung von Vision gibt es folgende grobe Schritte:

1. Die markierten Regionen in der Vorlage auf die neuen Bilder anwenden
2. Mit den Regionen jedes Bilds Bildausschnitten erstellen
3. Die Ausschnitte in die OCR-Engine geben
4. Texterkennungs-Ergebnisse darstellen

Zu 1.: Die markierten Regionen in der Vorlage auf die neuen Bilder anwenden
 Damit ist gemeint, dass nun das erste Bild des neuen Scans mit den Regionen der ersten Seite der Vorlage verarbeitet wird. Zur Erinnerung aus Kapitel 5: Das Ziel ist es, die markierten Regionen in der Vorlage auf die zu digitalisierende Seite anzuwenden und minimale Bildausschnitte zu erstellen, in denen der zu digitalisierende Text steht (siehe Abbildung 5.6). Diese Ausschnitte werden dann von der Texterkennung verarbeitet. Diesen Vorgang wird für jede Seite wiederholt. Jedoch ergibt sich hierbei ein Problem.

Wie man in den Abbildungen 3.1a und 3.1b sieht, haben beide, durch die App entstandenen Bilder, unterschiedliche Maße. Die Abbildung 3.1a ist größer und etwas breiter als die Abbildung 3.1b. Das liegt daran, dass sich der Winkel und der Abstand der Kamera zum Dokument bei den Aufnahme geändert hat, da das Gerät, während des Fotografie-



(a) Dialogfenster zur Auswahl der OCR-Engine (b) Übersicht nach der Texterkennung mit farblicher confidence (c) Änderung der Note mit spezieller Tastatur für Dezimalzahlen und confidence als Zahlenwert

Abbildung 6.9: Vorlagen-Ansicht vor und nach der Texterkennung

rens in den Händen gehalten wurde. Das bedeutet also, wenn die Bilder in der Vorlage nicht genau so groß sind, wie die Bilder des neuen Scans, können im schlechtesten Fall wichtige Informationen durch die absoluten Positionen der Regionen abgeschnitten werden. Würde man immer den selben Abstand und Winkel garantieren, wie das beispielsweise in einem richtigen Kopiergerät oder Scanner der Fall ist, könnte man nun die Berechnung der absoluten Positionen in relative Positionen der Regionen weglassen. Zusätzlich sind die relative Höhe und Breite der Regionen ebenfalls zu berechnen.

Zu 2.: Mit den Regionen jedes Bilds Bildausschnitten erstellen Nachdem die neue Position und Maße aller Regionen einer Seite bestimmt sind, entstehen daraus Bildausschnitte für die Texterkennung. Zur Erinnerung, die Regionen dienen als eine Art Schablone, so dass alles, was sich außerhalb dieser befindet, nicht verwendet wird (siehe Abbildung 5.1). Auch trotz der Umrechnung in relative Positionen kann es bei falscher Erstellung der Vorlagen oder bei falschem Einstellen der Klausur passieren, dass trotzdem Teile der Regionen wegfallen. Für mehr Details über falsches Erstellen oder Einstellen siehe im Unterabschnitt 7.2.1.

Zu 3.: Die Ausschnitte in die OCR-Engine geben Zu jedem Bildausschnitt wird eine sogenannte Texterkennungs-Anfrage konfiguriert. Auch hierfür wird immer noch die zugehörige Region des Ausschnitts benötigt. In der Konfiguration werden an Hand des Region-Datentyps (6.5b) unterschiedliche Einstellungen getroffen. Beispielsweise bekommt die Anfrage einer Region, vom Typ Note, eine Liste an möglichen Noten. Die Liste hat, wie schon im Kapitel 5 erwähnt, Vorrang vor dem Standard-Wörterbuch, in der Wörterkennungsphase und sorgt für bessere Ergebnisse²¹. Auch könnte hier die Sprache des zu erkennenden Texts angegeben oder die Texterkennungs-Stufe eingestellt werden. Diese bestimmt, welche Techniken bei der Texterkennung verwendet werden. Entweder man gibt der Geschwindigkeit Vorrang vor der Genauigkeit oder aber man setzt auf eine längere, rechenintensivere Erkennung. In der Anwendung ist die OCR-Genauigkeit entscheidend und auf die Einstellung der Sprache wurde verzichtet, da es sich in keinem Fall um normalen Text handelt, sondern um Eigennamen, Abkürzungen und Ziffern.

Standardmäßig werden bei einer Texterkennungs-Anfrage zunächst alle Glyphen oder Zeichen im Eingabebild lokalisiert und dann jede Zeichenfolge analysiert²². Anschließend wird eine Korrektur an den erkannten Wörter auf der Grundlage eines Wörterbuchs und anderer Wahrscheinlichkeits-Heuristiken für Zeichenpaare durchgeführt [2]. Der genaue Ablauf der Texterkennung oder die verwendeten Algorithmen von Vision sind allerdings nicht bekannt.

Anschließend an die Texterkennung durch Vision erfolgt in einigen Fällen noch eine selbst entwickelte Korrektur, wie sie in Kapitel 5 Abschnitt 5.3 beschrieben ist. Auch hier spielt der Regionen-Datentyp (6.5b) wieder eine Rolle. Beispielsweise wird der Text einer vermeidlichen Seminargruppe mit allen in der App hinterlegten Seminargruppen verglichen. Der Seminargruppen-Name mit der größten Übereinstimmung wird dann angezeigt. Die confidence wird ebenfalls angepasst. Zur Erinnerung, diese sagt aus, wie sicher sich die Texterkennung ist, dass das Ergebnis stimmt. Ähnliche Korrekturen gibt es für die Note und Punkte.

Danach werden die erkannten Texte bzw. Worte und deren confidences an den State-Container der App gesendet. Für die Ergebnisse der Texterkennung ist ein eigener State vorgesehen, in dem auch die Server-Ergebnisse von Tesseract gesichert werden. Somit ist das Anzeigen der Ergebnisse bei beiden OCR-Engines gleich.

Zu 4.: Texterkennungs-Ergebnisse darstellen Die Ergebnissen der Texterkennung und die jeweilige confidence werden auf der Benutzeroberfläche, nach Seiten sortiert,

²¹ Vision *costumWords* Dokumentation <https://developer.apple.com/documentation/vision/vnrecognizetextrequest/3152640-customwords>

²² Quelle: *VNRecognizeTextRequest* Dokumentation <https://developer.apple.com/documentation/vision/vnrecognizetextrequest>

angezeigt (6.9b). Alle OCR-Ergebnisse können zudem noch angepasst werden. Abhängig vom Datentyp der Region passt sich die Tastatur beim Ändern der Texte an (6.9c). Beispielsweise sind die Tastaturen beim Anpassen von Noten und Punkten auf Zahlen und Dezimalpunkte spezialisiert. Zudem kann die confidence farblich (siehe Abbildung 6.9b) aber auch als Zahlenwert (siehe Abbildung 6.9c) angezeigt werden. Wurde kein Text erkannt, ist dementsprechend ein Warnsymbol an der Stelle der confidence zu sehen.

Ablauf mit Tesseract

Etwas anders ist der Ablauf, wenn die Texterkennung Tesseract auf dem Server verwendet wird:

1. Bilder an den Server senden
2. Auftrag zur Texterkennung an den Server senden
3. Serverseitige Texterkennung
4. Texterkennungs-Ergebnisse darstellen

Zu 1.: Bilder an den Server senden Nachdem das Dokument fotografiert wurden, werden die Aufnahmen an den Server gesendet. Dieser gibt, wie beim Speichern der Scan-Vorlage, die Pfade der Bilder als Antwort zurück.

Zu 2.: Auftrag zur Texterkennung an den Server senden Anschließend wird für jeden Pfad eine Auftrag zur Texterkennung an die OCR-Schnittstelle des Servers gestellt. In der Anfrage wird der Pfad des Bildes und die passende Seiten-ID der ausgewählten Vorlage gesendet. Ab hier übernimmt der Server die Arbeit der Texterkennung.

Zu 3.: Serverseitige Texterkennung Mithilfe der Seiten-ID, die in der Datenbank des Servers hinterlegt ist, werden dann die Regionen der Seite auf das Bild, welches sich an dem gesendet Pfad befindet, angewendet. Für mehr Details siehe im Praktikumsbericht von Tobias Kallauke.

Zu 4.: Texterkennungs-Ergebnisse darstellen Nach der Texterkennung sendet der Server die Ergebnisse der Seiten als Antwort zurück. Diese werden dann im selben State des State-Containers gesichert, wie es auch bei Vision der Fall ist. Die Ergebnisse bestehen aus der Region-ID, zur Zuordnung, dem erkannten Text sowie der confidence des OCR, zu dem jeweiligem Wort. Auch ähnlich, wie es bei Vision der Fall ist.

Die Darstellung der Ergebnisse und das Verhalten der Benutzeroberfläche ist ebenfalls identisch.

7 Fazit

Dieses Kapitel reflektiert den aktuellen Stand der App und zieht einen Vergleich mit den Anforderungen aus Kapitel 4.

7.1 Stand des Prototyps

Während der Entwicklung konnten viele Anforderungen umgesetzt werden. Jedoch fehlt zum Zeitpunkt der Abgabe ein essentieller Bestandteil der App. Auf Grund mangelnder Zeit war es nicht mehr möglich die digitalisierten Daten aus der Texterkennung an den Server zu senden, wo sie gespeichert und in ein tabellarisches Format gebracht werden sollten.

Außerdem wurde nur einer der Kontrollmechanismen umgesetzt. Grund dafür ist, dass dafür zu wenig Zeit in der Planung investiert wurde. Grundsätzlich könnten die fehlenden Mechanismen recht schnell implementiert werden, allerdings ist der aktuelle Ansatz der dahinter liegenden Datenstruktur nicht zum Speichern in der Datenbank geeignet. Sobald Änderungen an der Datenstruktur vorgenommen werden könnten, kann es zu Fehlern kommen. Es existiert jedoch schon eine Idee zur Umsetzung einer besseren Datenstruktur. Dafür müssten allerdings neben der App auch der Server und die Datenbank angepasst werden. Des Weiteren wurden Fehlermeldungen und/oder -behandlungen bezüglich des Servers nur sporadisch implementiert. Aber auch hierfür sind die Grundlagen schon gelegt, da Fehler stets aufgefangen und zentral abgelegt werden.

7.2 Probleme und Grenzen der App

In diesem Abschnitt werden Probleme aufgeführt, die beim Verwenden der App beobachtet werden können. Auch wird hier über mögliche Lösungen zu einigen Problemen diskutiert.

7.2.1 Probleme beim Erkennen von Dokumenten

Bei der Erkennung von Dokumenten in einem Bild wird in der Bildverarbeitung auf Algorithmen der Segmentierung zurück gegriffen. Im Praktikumsbericht von Tobias Kallauke wird ein gängiger Ablauf zum Erkennen von Dokumenten dargestellt. Die genaueren Hintergrund und Schwächen der Algorithmen, welche die grundlegenden Probleme und Grenzen der App begründen, würden den Rahmen des Praktikumsberichts sprengen, weshalb darauf nicht weiter eingegangen wird. Allerdings muss klar sein, dass für die

Erkennung des Dokuments es wichtig ist, dass die Ecken sowie die Seiten sich gut vom Hintergrund abheben.

Das bedeutet im Umkehrschluss, dass wenn die Seiten und Ecken sich nicht gut abheben, kann das Dokument nicht oder nur schlecht erkannt werden. Allerdings gibt es für diese Probleme zwei effiziente Lösungen.

- Die erste Lösung besteht darin, einen Blitz zu verwenden. Meistens schaltet sich bei zu wenig Licht die zusätzliche Belichtung automatisch ein. Grundsätzlich empfiehlt es sich aber, immer eine Blitz zu verwenden, da dadurch das Dokument gleichmäßig belichtet wird. Außerdem werden Schatten vom Benutzer oder durch kleinere Falten beseitigt, welche durch eine Deckenbeleuchtung entstehen könnten. Bei einer zu starken Grundbelichtung verschlechtert ein Blitz jedoch nur das Ergebnis, da dadurch Text und Kanten unkenntlich gemacht werden könnten. Nach eigener Erfahrung ist es am geeignetsten, bei Tageslicht und mit Blitz die Dokumente aufzunehmen.
- Auch wenn bei guter Belichtung, die Wahrscheinlichkeit für falsche Kantenerkennung reduziert worden ist, kommt es bei schnellen Bewegungen der Kamera oder zu spitzem Winkel zum Dokument sowie ungeeignetem Hintergrund trotzdem dazu, dass ein falscher Bildausschnitt als Dokument gewählt wird. Deshalb ist es möglich das Bild noch einmal aufzunehmen, oder aber das Dokument im Bild selbst zu markieren. Dafür stellt VisionKit eine besondere View zur Verfügung, die es einem ermöglicht, die Ecken des Dokuments per Hand auszuwählen. Das hilft auch gegen die Problematik, wenn die Ecken des Papiers umgekippt oder abgerissen sind, da man selbst die Position der Ecke approximiert angeben kann.

7.2.2 Probleme der Klausur-Vorlage beim Scannen

Dieser Abschnitt bezieht sich ausschließlich auf entstandene Probleme mit der Klausur, welche in Abbildung 3.1 zu sehen ist. Diese stand während des Praktikums als Beispiel-Klausur zur Verfügung und ist stellvertretend für alle Klausuren-Vorlagen der Fakultät CB. Im Abschnitt 7.3, werden mögliche Änderungen und Lösungen zu den hier aufgeführten Problemen diskutiert.

Wie in der Abbildung 3.1a zu sehen ist, besitzt die Klausur ein Deckblatt mit Feldern, die der Student auszufüllen hat. Darunter Name, Artikelnummer und Seminargruppe. Diese Informationen sind essentiell bei der Digitalisierung und doch wird nicht klar zwischen einem Feld für den Vornamen und einem Feld für den Nachnamen differenziert. In den Abbildungen 6.5 und 6.6 wird diese Problematik nochmal dargestellt. Bei den Datentypen der Regionen (6.5b) wird zwischen Vornamen und Nachnamen unterschieden, auf der Klausur (6.5c) jedoch nicht. In der abgebildeten Vorlage (6.6b) wurde der Vorname auf den ersten und der Nachname auf dem zweiten Strich des Feldes Namen

gesetzt. Auch verleitet ein einzelner Strich dazu, über die Länge hinaus zu schreiben. Wenn die Regionen der Scan-Vorlage dann zu klein gewählt ist, können Informationen fehlen. Bei dem Feld für die Note wird dies noch extremer. Da nicht mal ein Markierung für die Note vorhanden ist, müsste man einen sehr viel größeren Bereich markieren (6.6b). Allerdings wird hinter der Note auch die Unterschrift des Prüfers gesetzt, welche bei der Digitalisierung nicht mit auftauchen sollte. Im Gegensatz dazu sind die Felder in der Tabelle für die Punkte zu jeder Aufgabe zu klein. Die geringe Größe verleitet dazu, dass Feld komplett aus zu nutzen und die Ziffer möglichst groß rein zu schreiben. Dazu kommt dass solch kleinen Regionen kaum Spielraum für Toleranz lassen und somit entstehen hier immer wieder Fehler. Entweder die Ziffer ist abgeschnitten, wenn man die Region zu klein wählt oder aber die Ränder des Feldes werden als eine Eins oder als der Buchstabe I erkannt, wenn man die Region zu groß wählt.

7.2.3 Weitere Probleme der App

Obwohl SwiftUI die Entwicklung der App erst möglich gemacht hat, bringt das Framework von Juni 2019 einige Schwierigkeiten mit sich. Seit der Veröffentlichung gibt es nur eine sehr kleine Dokumentation und der Schnittstellen-Umfang, kann auch noch lange nicht mit dem des Frameworks von UIKit mithalten, welches seit Jahren in der iOS-Entwicklung als Standard benutzt wird. Grund für den geringen Umfang und schlechter Dokumentation ist vermutlich, die ständige Weiterentwicklung Seitens Apple. Das führt jedoch zu einigen provisorischen Lösungen und macht das Auffinden von Fehlern besonders schwierig. Deshalb kann es bei der Benutzung der App zu unerwarteten Abstürzen oder Aufhängern kommen. Des Weiteren sind die verwendeten Texterkennung-Engines nur für gedruckte Schrift ausgelegt. Handschriftliche Zeichen werden nur dann gut erkannt, wenn sie in Druckschrift geschrieben sind. Bei Ziffern ist es besonders auffällig, da z. B. eine gedruckte 4 oder 7 ganz andere Charakteristiken haben, als wenn sie geschrieben wurde.

7.3 Verbesserung der Klausuren-Vorlage

Dieser Abschnitt beschäftigt sich mit Verbesserungen und Änderungen an Hand der Klausur in Abbildung 3.1. Sie steht jedoch stellvertretend für die Klausuren-Vorlage der Fakultät CB. Alle folgenden Vorschläge haben den Hintergrund die App bei der Digitalisierung zu unterstützen und basieren auf der gesammelten Erfahrung mit der Anwendung, während der Entwicklung sowie den angegeben Problemen aus Abschnitt 7.2.

Der erste Vorschlag bezieht sich auf die Problematik, dass die Ecken und Kanten bei der Erkennung des Dokuments entscheidend sind. Sie sind die wichtigsten Referenz-Punkte bei der Objekt-Erkennung, können aber unter bestimmten Bedienungen, wie im Unterabschnitt 7.2.1 erklärt, nicht immer erkannt werden. Aus diesem Grund empfiehlt

es sich eigene Referenzpunkt auf dem Dokument anzubringen. Beispielsweise durch einen Rahmen, wie es schon auf der Klausur (3.1a) gemacht wurde, oder durch QR-Code ähnliche Muster. Der Vorteil bei eigenen Referenzpunkt ist, dass sie so gestaltet werden können, dass sie sich immer vom Blatt abheben. So mit ist die Dokumenten-Erkennung nur noch vom Licht abhängig. Ein weiterer Vorteil ist, dass dadurch die Kamera noch näher an den Text heran kann und die Auflösung noch besser wird. Der größte Nachteil daran ist allerdings, dass die Erkennung nicht mehr allein von Frameworks übernommen werden kann. Es müsste ein künstliches neuronales Netz trainiert werden, welche diese Referenzpunkte in Bildern erkennt.

Der zweite Vorschlag bezieht sich auf den Unterabschnitt 7.2.2. Einige dort angesprochenen Probleme bezogen sich auf die Ungewissheit, wie groß eine Region abzustecken ist. Die einfachste Lösung ist es, einen genauen Rahmen für alle Felder fest zu legen. Auch müssen alle Felder ausreichend beschriftet sein, so dass jedem klar ist, wo was hineinzutragen ist. Und um das Problem zu vermeiden, dass der Rahmen als ein Buchstabe oder eine Zahl erkannt wird, kann dieser nur schwach eingezogen werden. Mithilfe von Bildverarbeitung-Algorithmen könnten diese feinen Linien dann auch ohne Verlust der wichtigen Daten heraus gerechnet werden.

Abschließend ist zu erwähnen, dass laut der Problemstellung aus Kapitel 3 eine Klausuren-Vorlage entwickelt werden sollte. Auf Grund mangelnder Zeit war das jedoch nicht möglich.

8 Ausblick

Der aktuelle Stand der App bietet eine gute Grundlage für mögliche Weiterentwicklungen. Jedoch fehlen einige wichtige Elemente zur Inbetriebnahme des Software-Systems, wie schon im Kapitel 7 verdeutlicht. In den folgenden Abschnitten werden einige Ideen erläutert, wie die App und das gesamte Software-System weiter entwickelt und verbessert werden kann.

Verbesserung der Benutzerfreundlichkeit und des Workflows Gerade weil die App noch ein Prototyp ist, sollten bei einer Weiterentwicklung zu allererst Dinge implementiert werden, die die Benutzung verbessern. Beispielsweise werden Login-Daten und schon heruntergeladene Scan-Vorlagen nicht gespeichert. Auch muss zu jeder Region ein Name vergeben werden. Allerdings gibt auch der Datentyp (6.5b) der Region in den meisten Fällen darüber Auskunft, wie die Region heißen sollte (siehe dazu Abbildung 6.5a). Handelt es sich um den Vor- bzw. Nachnamen, die Matrikel-Nummer, Seminar-Gruppe oder die Note, taucht diese Region auch nur genau einmal in einer Scan-Vorlage auf. Ausschließlich der Datentyp Punkte sollte mehrmals vergeben werden können. Aus diesen Aspekten sollte der Workflow beim erstellen von Regionen angepasst werden.

Weitere Dokument- und Datentypen Auf dem Abschnitt *Verbesserung der Benutzerfreundlichkeit und des Workflows* aufbauend, ist folgende Idee entstanden. Durch die generischen Vorlagen ist es möglich neben Klausuren auch andere Dokumente nach diesem Schema zu digitalisieren. Beispielsweise Krankenscheinen, Urlaubsanträgen oder Arbeitsverträgen. Dafür könnten spezielle Workflows zum Erstellen einer Vorlage implementiert werden und weitere Regionen-Datentypen hinzukommen, wie z. B. Datum, IBAN-Nummern, E-Mail-Adressen und Telefonnummern.

Server-Schnittstellen Zum Zeitpunkt der Abgabe sind in der iOS-App nicht alle verfügbaren Server-Schnittstellen implementiert. Es fehlt noch das Ändern ganzer Vorlagen und deren Einzelteile, wie Seiten und Regionen. Die Views zum Ändern einer Scan-Vorlage sind schon implementiert, da die Views vom Erstellen einer Scan-Vorlage an der Stelle wieder verwendet werden konnte. Es fehlt lediglich die interne Logik sowie Schnittstellen zum Server.

Erweiterung der Texterkennung Wie im Abschnitt 7.2 erwähnt unterstützten Vision und Tesseract das Erkennen von Handschrift nicht. Jedoch wäre es möglich eigne

Neuronale-Netze zu trainieren und implementieren, die diese Aufgabe übernehmen. Beispielsweise existiert ein Datensatz an handgeschriebene Ziffern²³, mit dem solch ein einfaches Netz für das Erkennen von Ziffern, trainiert werden könnte. Außerdem bietet die Server Architektur die Möglichkeit, weitere oder mehr OCR-Engines zu implementieren. Für mehr Details siehe im Praktikumsbericht von Tobias Kallauke.

Klausuren-Einsicht Die archivierten Klausuren-Bilder könnten über ein Web-Portal dazu genutzt werden, eine online Einsicht der Prüfungen zu gestalten. Jedoch bringt solch eine Plattform auch ein Problem mit sich. Durch sie ist die Verbreitung von Klausuren noch einfacher als zuvor, wodurch die Professoren dazu angehalten werden, immer wieder neue Aufgaben auszudenken. Aber auch dafür könnte es in Zukunft eine Lösung geben.

Weitere Plattformen Neben der iOS-App sollte Tobias Kallauke die selbe für Android Geräte entwickeln. Jedoch existierten zum Zeitpunkt der Entwicklung keine Frameworks, die das Erkennen von Dokumenten in der Kamera übernimmt. Deshalb entwickelte Tobias ein Prototyp, der sich mit diesem Problem beschäftigt. Allerdings könnte das Software-System auf noch mehr Plattformen Anwendung finden. Neben einem Programm für Linux, MacOS und Windows ist auch einen Web-Plattform erdenklich. Alle könnten den selben Umfang besitzen nur das Erstellen der Scans erfolgt beispielsweise über Scan- oder Kopiergeräte.

Intelligente Referenzpunkte In Abschnitt 7.3 wurde erwähnt, dass eigene Referenzpunkte auf den Klausuren Vorteile mit sich bringen. Möglich wäre auch, dass die Referenzpunkte ähnlich wie QR-Codes aufgebaut sind. Diese enthalten die nötigen Informationen zu den Regionen auf der Seite und ermöglichen das Digitalisieren, ohne die Auswahl der richtigen Vorlage. Daran angelehnt könnten ein LATEX-Paket oder Word-PlugIn entwickelt werden, wodurch die QR-Codes automatisch für jede Seite und deren Regionen generiert werden.

Klausuren-Vorlage entwickeln Auf Abschnitt 7.3 und den *intelligenten Referenzpunkten* aufbauen könnten Klausuren-Vorlagen umgesetzte und getestet werden. Auch die App müsste angepasst werden, so dass die Referenzpunkte erkannt werden, wie im Abschnitt 7.3 beschrieben.

Benutzer Akzeptanz Studie Aktuell ist unklar, ob die App tatsächlich Zeit einspart. Aus diesem Grund bietet es sich an eine umfangreiche Benutzer Akzeptanz Studie an-

²³ MNIST-Datensatz - <http://yann.lecun.com/exdb/mnist/>

zusehen legen. Außerdem könnten so weitere Probleme entdeckt und Lösungen entwickelt werden, bevor die App benutzt werden kann.

Anhang A: Wasserfall-Modell

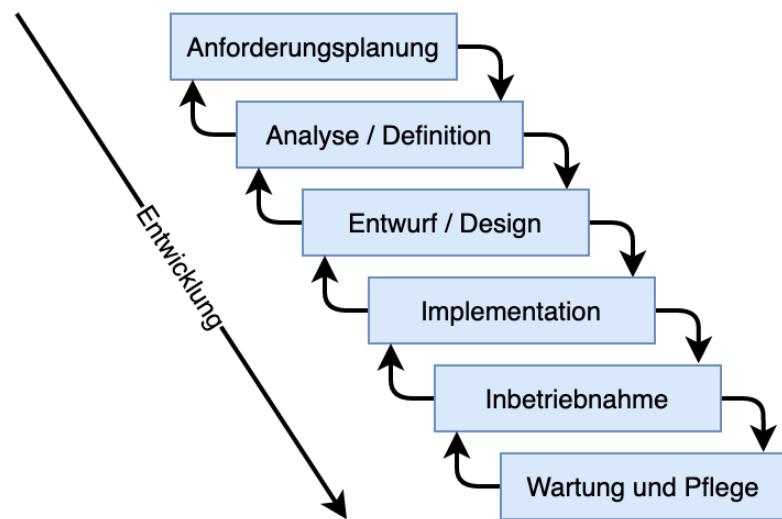


Abbildung A.1: Das Wasserfall-Modell nach Winston W. Royce

Anhang B: Workflow

Scan-Vorlage erstellen

1. "Neue Vorlage erstellen"
2. Foto machen
3. Frage: Ist Foto gut?
 - a) Ja: gehe zu 4.
 - b) Nein: gehe zu 2.
4. Neues Attribut hinzufügen
5. Bereich auf Bild auswählen
6. Frage: Ist Bereich gut?
 - a) Ja: gehe zu 7.
 - b) Nein: gehe zu 5.
7. Name für Attribut festlegen
8. Datentyp für Attribut festlegen (Name, Matrikelnummer, Note, ...)
9. Frage: Sind alle Attribute vorhanden?
 - a) Ja: gehe zu 10.
 - b) Nein: gehe zu 4.
10. Fertig
11. Vorlage an Server senden

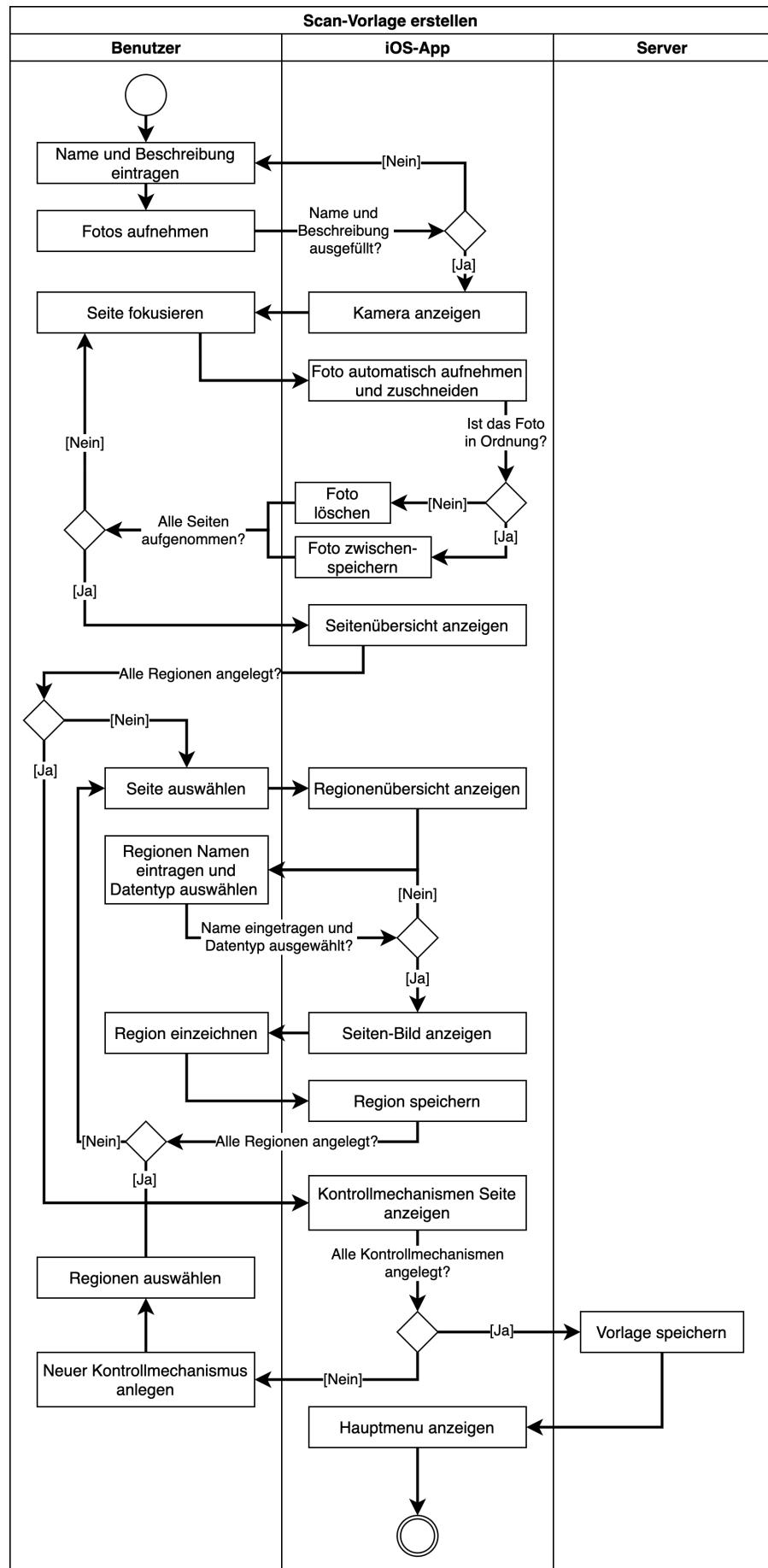


Abbildung B.1: Gekürztes Flussdiagramm zum Erstellen einer Scan-Vorlage

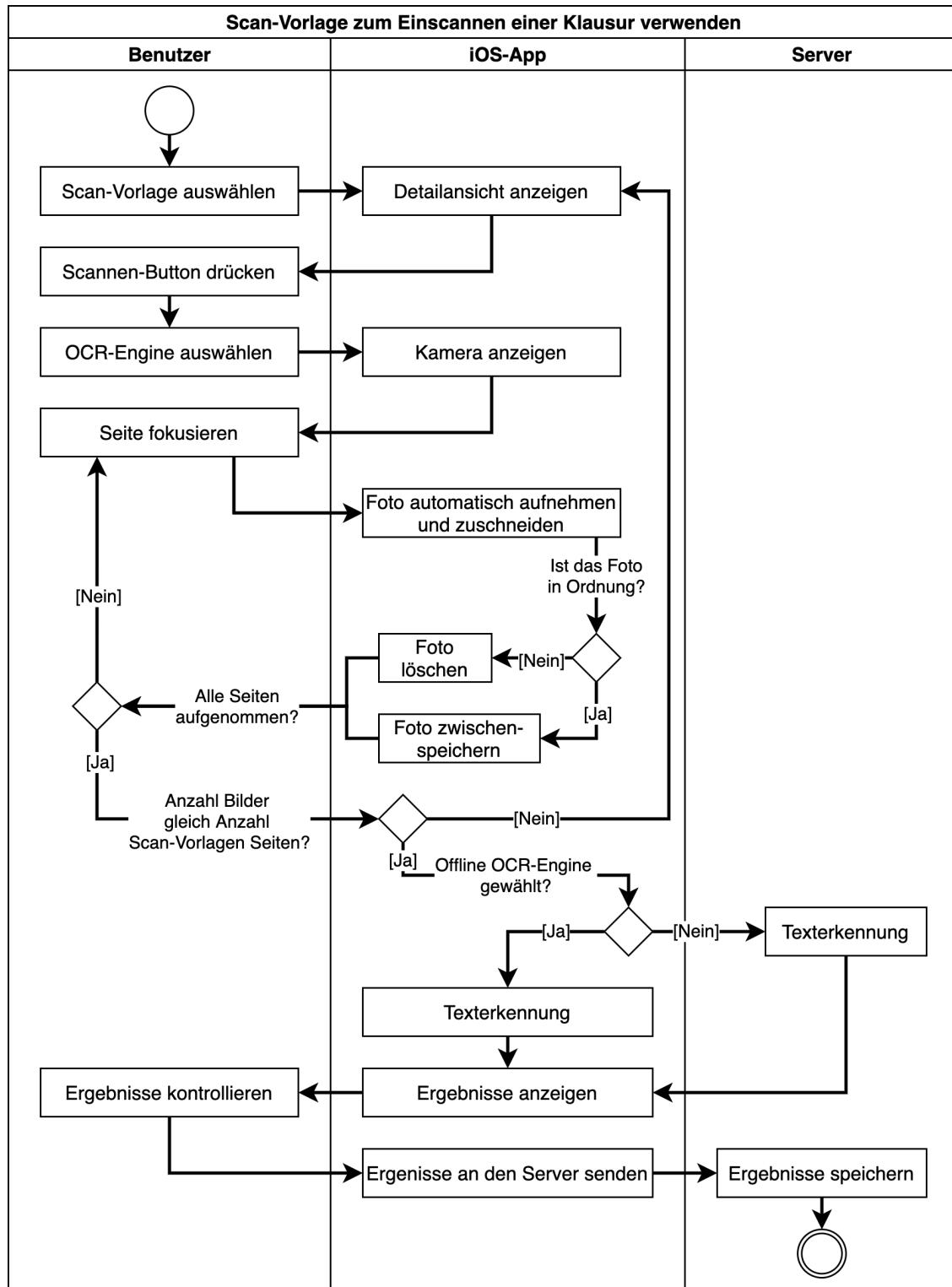


Abbildung B.2: Gekürztes Flussdiagramm zum Verwenden einer Scan-Vorlage

Anhang C: Zusätzliche Funktionalität

- Textgröße
- Dunkelmodus
- Überschrift nochmal ändern
- Kontrollmechanismus löschen siehe Bilder L5

Anhang D: Installation der iOS App

Im folgenden Kapitel ist beschrieben, wie die App auf einem iPhone oder iPad zu installieren ist.

Voraussetzungen:

- macOS 10.15.4 oder neuer
- Xcode 11.4.1 oder neuer installiert
- iPhone mit iOS 13.0 / iPad mit iPadOS 13.4 oder neuer
- Internetverbindung

Installation: Stellen sie sicher, dass Xcode installiert und das iPhone/iPad mit dem Mac über ein USB-Kabel verbunden sind.

1. Lade das Projekt herunter. (.zip anschließend entpacken)
2. Öffne die Datei *DocumentenScanner.xcodeproj*. Xcode öffnet sich.
3. Klicke in der Ordnerübersicht bzw. Navigator ganz oben auf das Xcode Projekt *DocumentenScanner*.
4. Wähle im Editor-Fenster nun unter *Targets* das Projekt *DocumentenScanner* aus. (Entweder links oben als Dropdown-Menu oder in dem geöffneten Menü links innerhalb des mittleren Fenster.)
5. Wähle den Reiter *Signing & Capabilities* aus. Es erscheint im mittleren Fenster ein Unterpunkt *Signing*.
6. Wähle ein passendes Team aus oder lege eins an.
7. Ändere den Bundle Identifier passend zu dem ausgewählten Team. Z. B. com.<TeamName>.DocumentenScanner. Es taucht ein Button auf, mit dem die App signiert werden kann. Falls kein Button erscheint mache mit den nächsten Schritten weiter und kommen bei einer auftretenden Fehlermeldung bezüglich der Signierung hier her zurück. Der Button ist dann zu sehen.
8. Wähle nun ein Simulator oder das angeschlossene Gerät in der Toolbar aus den *Active Devices* aus.
9. Drücke anschließend den Start bzw. Run Button der *Build Controls* oder die Tastenkombination *cmd + R*.
10. Folge den Anweisungen von Xcode und auf dem iPhone/iPad. Beim Verwenden eines Simulators tauchen keine weiteren Anweisungen auf. Die App startet automatisch.

Anhang E: Tätigkeitsbericht

24.02. - 01.03. Zunächst habe ich mich mit der Problemstellung auseinander gesetzt, Ideen gesammelt, Problemanalyse betrieben und einen kleinen Prototypen entwickelt. Dazu erstellte ich eine minimale Projektplanung, arbeitete mich in die Frameworks *Vision* und *VisionKit* ein und setzte eine Versionsverwaltung auf. Zusätzlich suchte ich nach einer passenden App-Architektur, die für das deklarative GUI-Framework SwiftUI sowie für asynchrone Aufgaben, wie z. B. API-Aufrufe geeignet ist. Dabei stieß ich auf *Cleancode Architecture* und *Redux*.

02.03. - 08.03. In dieser Woche habe ich die Texterkennung auf den berechneten Regionen eines neuen Fotos implementiert, den Workflow sowie viele andere Kleinigkeiten in der App verbessert und alle Fehler der letzten Woche behoben, sodass ich neue Dinge implementieren konnte. Zudem probierte ich CI sowie Lint für das Projekt aus. Da CI für eine iOS-App mit *Github Actions* schwer aufzusetzen war und ab April kostenpflichtig wurde, verwarf ich meine Pläne. Des Weiteren pflegte ich das Projekt Management durch *Issues* und *Project Boards* in GitHub. Anschließend programmierte ich den App-Workflow so um, dass nun mehr als eine Seite aufgenommen und analysiert werden konnte.

Abgesehen von neuen Quellcode begann ich mit dem Schreiben des Praktikumsberichts und arbeitete mich dazu in \LaTeX und die Bachelorarbeit-Vorlage für \LaTeX der Hochschule Mittweida ein.

09.03. - 15.03. Zu Beginn der dritten Woche schaute ich mir Möglichkeiten für serverseitiges OCR an. Dabei sammelte ich Informationen zu dem Framework Vapor und Swift unter Linux. Da die Frameworks *Vision* und *CoreML* von *Apple* unter Linux nicht funktionierten, stellte sich IronOCR als beste Option herausstellte. Mithilfe der in der App verwendeten Datentypen entwickelte ich ein Datenbankmodell und erstellte dazu noch eine JSON-Struktur, die später für die APIs verwendet werden könnte. Außerdem gab es ein Meeting, in dem Tobias Kallauke und ich unseren aktuellen Stand präsentierten, um weitere Schritte und Aufgaben zu planen. Bis zum Ende der Woche arbeitete ich fortlaufend an meinem Beleg und schrieb den Datenfluss in der App um. Nun ähnelte er sehr dem Redux-Model.

16.03. - 22.03. Anfangs schrieb ich meinen Praktikumsbericht weiter, bearbeitete alte Issues und fügte neue dem Project Board hinzu. Außerdem gepflegte ich die Dokumentation, um anschließend Kontrollmechanismen hinzuzufügen. Dabei entstanden neue Views. Der Redux-Store musste dadurch angepasst werden. Es kam eine Erweiterung für die Texterkennung hinzu, so dass man durch die Auswahl eines Datentyps, das Resultat der Erkennung verbessern konnte. Des Weiteren habe ich bis zum Ende der Woche den Kontroll-Typ *Vergleich* vollständig implementiert und die App auf Fehler und

Abstürze kontrolliert sowie den Beleg um einige Kapitel erweitert.

23.03. - 29.03. Ich begann den Workflow und die Navigation in der App zu verbessern und vereinfachen. Dabei beseitigte ich einigen Quellcode des Prototyps, erweiterte die Dokumentation und behob einige Fehler. Anschließend überarbeitete ich einige Views, sodass sie übersichtlicher und einfacher zu benutzen sind. Nach dem iOS 13.4 Update Mitte der Woche funktionierte ein Teil der App nicht, da sich das Verhalten von Views geändert hatte. Ich behob die Fehler, testete ausgiebig die App und fügte iPad Unterstützung hinzu. Des Weiteren entstand eine neue verbesserte permanente Scan-Vorlage und ich schrieb einen großen Teil des Berichts.

30.03. - 05.04. Das Backend für die App war soweit, dass ich es aufsetzen und die API-Schnittstellen implementieren konnte. Dazu erstellte ich Views für Registrieren und Anmelden, die mithilfe von regulären Ausdrücken, die Eingaben überprüfen. Außerdem entwickelte ich für die anfallenden asynchronen Aufgaben einen Schicht im App-Store. Dabei las ich mich in das Framework Combine ein und überlegte mir einen geeigneten Aufbau. Da der Ansatz von Combine sehr neu für mich war, dauerte es zwei Tage, bis ein erster API-Service mit Fehler-Handling funktionierte. Zum Ende der Woche waren alle der Create-Schnittstellen implementiert, getestet und dokumentiert. Nebenbei erstellte ein paar Issues für das Backend und sprach mich mit Tobias über OCR auf dem Server ab.

06.04. - 12.04. Diese Woche startete mit dem Umschreiben der Kontroll-Mechanismen und deren Analyse. Anschließend integrierte ich die Neuerungen vom Backend und erstellte die API für den Upload von Bildern. Danach fügte ich die APIs zusammen, um Vorlagen vollständig auf dem Server zu speichern und ab zu rufen. Dazu schrieb ich eine eigene JSON-Decoder-Funktion, um die App internen Datentypen zu unterstützen. Zusätzlich wurde die App etwas benutzerfreundlicher und ein Problem mit dem Start der iPad Version wurde behoben. Nach einem Meeting folgten noch weitere Absprachen mit Tobias und ich arbeitete weiter an dem Bericht.

14.04. - 19.04. Ein kritisches Problem mit den Sessions der vorherigen Woche konnte in dieser endlich gelöst werden. Dazu konnte ich die Anwendung etwas optimieren und einen sehr großen Teil des Beleges fertig stellen. Dabei half auch die Beantwortung vieler Fragen während eines Meetings mit dem Betreuer. Allerdings entstanden Probleme mit der Datenbank, die die Funktionalität der App einschränkten.

20.04. - 26.04. Ziel dieser Woche war es, so viel wie möglich des Belegs zu schreiben. Dafür fertigte ich einige Schemata und Bilder an. Zusätzlich konnte ich einige Fehler der App analysieren und beheben, sodass die Performance verbessert und Server-Aufrufe eingespart wurden. Außerdem kam die Online-Texterkennung über den Server hinzu sowie die zweite Korrektur in der Texterkennung mit Vision. Jedoch konnte die Online-Texterkennung auf Grund von Server-Problemen nicht getestet werden. Auch in

dieser Woche konnten einige Fragen zum Aufbau des Berichts in einem Meeting geklärt werden.

27.04. - 03.05. Im Mittelpunkt der Woche stand der Bericht, an dem ich täglich arbeitete. Ich erstellte einige Bilder für ihn und den Vortrag. Auch gab es diese Woche wieder ein Meeting.

04.05. - 10.05. Auch diese Woche arbeitete ich hauptsächlich am Beleg und könnte ihn so weit fertig stellen, dass nur noch wenige Dinge zu erledigen sind. Zudem gelang es Tobias die Fehler der online OCR-Engine Tesseract zu beseitigen. Dadurch konnte ich die letzten fehlenden Schritte für den Umgang mit den OCR-Schnittstellen implementieren und testen.

11.05. - 15.05.

Literaturverzeichnis

- [1] APPLE. Detecting Objects in Still Images | Apple Developer Documentation, Internetseite (2019). Verfügbar unter: https://developer.apple.com/documentation/vision/detecting_objects_in_still_images.
- [2] APPLE. Text Recognition in Vision Framework - WWDC 2019 - Videos, Internetseite (2019). Verfügbar unter: <https://developer.apple.com/videos/play/wwdc2019/234/>.
- [3] BRAGGE, M. *Model-View-Controller architectural pattern and its evolution in graphical user interface frameworks*. Bachelorarbeit (2013). Verfügbar unter: <https://lutpub.lut.fi/handle/10024/92156>.
- [4] FREEMAN, A. *Pro ASP.NET Core MVC 2*. Expert's Voice in .NET. Apress, Buch, 2. (30. Juni 2010) Edition (2017). ISBN 978-1-4842-3150-0. Verfügbar unter: <https://www.apress.com/gp/book/9781484231494>.
- [5] MITTWEIDA, H. Hochschule Mittweida: Portrait, Internetseite. Verfügbar unter: <https://www.hs-mittweida.de/hochschule/portrait.html>.
- [6] PAPA, J. Fundamental MVVM, Internetseite (2011). Verfügbar unter: <https://visualstudiomagazine.com/articles/2011/08/15/fundamental-mvvm.aspx>.
- [7] ROYCE, D. W. W. MANAGING THE DEVELOPMENT OF LARGE SOFTWARE SYSTEMS. Artikel, (1970), 11. Verfügbar unter: <http://www-scf.usc.edu/~csci201/lectures/Lecture11/royce1970.pdf>.
- [8] SILLMANN, T. Einstieg in SwiftUI, Internetseite (2019). Verfügbar unter: <https://www.heise.de/developer/artikel/Einstieg-in-SwiftUI-4594018.html>.

Erklärung

Hiermit erkläre ich, dass ich meine Arbeit selbstständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die Arbeit noch nicht anderweitig für Prüfungszwecke vorgelegt habe.

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

Mittweida, 13.03.2020