

Digitalisierung in der Amateurmusik

Anforderungen an ein digitales Notenarchiv für einen Musikverein

Stefan Balke,¹ Adrian Maiworm²

Abstract: Die Rolle der Notenverantwortlichen ist in einem Musikverein von großer Bedeutung. Neben der Bereitstellung der Noten für die regelmäßige Probenarbeit sind sie ebenfalls für die Archivierung des Notenbestandes verantwortlich. Ein wachsender Anteil dieses Bestandes wird heutzutage in digitaler Form erworben bzw. bei Eingang durch Einscannen digitalisiert. Die Ablage dieser digitalen Artefakte stellt viele Vereine vor eine besondere und oftmals ungewohnte Herausforderung. In diesem Beitrag wollen wir einen exemplarischen Digitalisierungsprozess im Amateurbereich am Beispiel eines Sinfonischen Blasorchesters beschreiben. Ferner legen wir typische Anforderungen aus dem Orchester an die Notenverantwortlichen dar und explorieren in einem ersten Anwendungsfall, wie diese durch algorithmische Unterstützung automatisiert werden können.

Keywords: Computergestützte Musikverarbeitung, Multimedia, Datenarchivierung

1 Einleitung

Die digitale Transformation von vormaligen analogen Prozessen und Abläufen schreitet in allen gesellschaftlichen Bereichen voran, so auch in der Amateurmusik. Viele Musikverlage und Komponist*innen bieten ihre Werke heutzutage in digitaler Version an: Angefangen von den Notentexten als Scans³ bis hin zu Demoaufnahmen auf großen Streamingplattformen. Was sich im Endeffekt als eine Erleichterung für die Notenverantwortlichen herausstellen soll, erfordert vorab Hilfsmittel, die eine strukturierte Ablage der digitalen Artefakte (Notentexte, Audioaufnahmen, Metadaten) gewährleisten können.

Der Status Quo in den Musikvereinen lässt sich als eine Kombination von gedruckten und eingescannten Notentexten beschreiben. Im Normalfall werden lediglich Kopien an die Musizierenden herausgegeben, die Originale verbleiben im Archivschrank. Durch die lange Tradition mancher Musikvereine stammen manche Noten allerdings aus den 50er und 60er Jahren des letzten Jahrhunderts. Damals waren Fotokopierer noch nicht im heutigen Maßstab verbreitet. Aus diesem Grund wurden zwangsläufig die Originalnoten herausgegeben, was die heutige Vollständigkeit mancher Notensätze signifikant schmälert und die Noten im schlimmsten Fall für den Verein unbrauchbar bzw. für die Notenverantwortlichen extrem

¹ Hochschule Weserbergland, Am Stockhof 2, 31785 Hameln, Deutschland balke@hsw-hamelnde

² Musikzug Ennest, 57439 Attendorn, Deutschland

³ Exporte in digitalen Notentextformaten (z.B. MusicXML) sind heute noch die absolute Ausnahme.

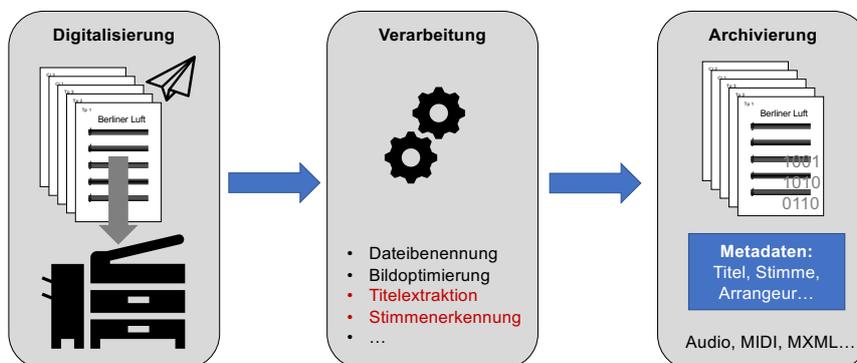


Abb. 1: Übersichtsdarstellung eines Digitalisierungsprozesses. Im ersten Schritt werden die Notentexte gescannt. In der Verarbeitung können neben der Optimierung der gescannten Bilder weitere Algorithmen, bspw. zur Metadatenextraktion angewendet werden. Die Archivierung erfolgt auf einem Speichermedium, das sowohl strukturierte Metadaten, als auch unstrukturierte Daten in Form von Scans oder Audioaufnahmen aufnehmen kann.

mühsam in der Wiederbeschaffung machen. Die vollständige Archivierung dieser existierenden Notenbestände, in Kombination mit neu erworbenen Werken, ist für viele Vereine ein Mammutprojekt und stellt die Notenverantwortlichen vor ungewohnte Herausforderungen, insbesondere wenn der Datenbestand eine gewisse Größe erreicht.

In diesem Beitrag beschreiben wir zunächst in Abschnitt 2 aus technischer Sicht ein digitales Notenarchiv eines Blasorchesters, als beispielhafter Repräsentant eines Musikvereins. Neben dem initialen Digitalisierungsprozess ziehen wir in diesem Abschnitt einen Vergleich zur existierenden Literatur, insbesondere den Forschungsarbeiten aus dem Bereich der „Digitalen Bibliotheken.“ Abgeschlossen wird dieser Abschnitt mit einer kurzen Beschreibung der typischen Anforderungen an ein digitales Notenarchiv eines Blasorchesters. Dies stellt den Rahmen für erste Aufgabenstellungen hinsichtlich algorithmischer Unterstützung in dieser Art von Archiven. Als exemplarische, erste Aufgabenstellung widmet sich Abschnitt 3 der automatisierten Extraktion von Metadaten wie dem Werktitel mittels Optical Character Recognition (OCR). Abschnitt 4 fasst den Beitrag zusammen und beschreibt die weiteren Schritte (z.B. Architekturauswahl, Datenstrukturierung etc.) zum Aufbau eines digitalen Notenarchives.

2 Digitales Notenarchiv eines Blasorchesters

Wie eingangs erwähnt stellt die wachsende Digitalisierung der Notenarchive und auch die steigenden Anforderungen der Musiker*innen, Notenverantwortliche in Musikvereinen vor mitunter neue Herausforderungen. Im Folgenden beschreiben wir einen Digitalisierungsprozess, den die Autoren in der Praxis mehrfach beobachtet haben bzw. selbst praktizieren.

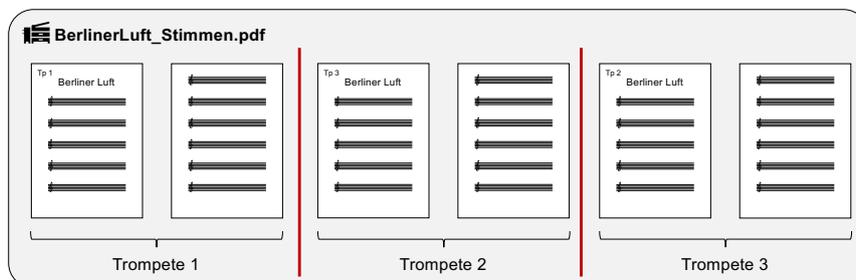


Abb. 2: Darstellung eines typischen Archivdokuments. Alle Stimmen sind aneinandergehängt. Der Beginn einer neuen Stimmen wird durch ein Notenblatt mit Titelinformation begonnen und endet vor dem nächsten Titelblatt, dargestellt durch die roten Trennlinien.

2.1 Digitalisierungsprozess

Abb. 1 stellt einen typischen Digitalisierungsprozess eines in Papierform vorliegenden Notensatzes schematisch dar (in unserem Beispiel den Notensatz für das Stück „Berliner Luft“).⁴ Im ersten Schritt der *Digitalisierung* wird der Notentext konventionell eingescannt. Dies erfolgt im Normalfall mit Flachbettscannern oder im Falle der Partitur mit speziellen Buchscannern (falls vorhanden). In vielen Musikvereinen erfolgt das Scannen über ein bestehendes Kopiersystem. Die Scans werden hierbei direkt auf ein Speichermedium (z.B. USB-Stick) gespeichert. In diesem Fall werden die Stimmen nacheinander eingescannt und in einer zusammenhängenden Datei abgelegt. Für die entstandene Datei vergibt das Kopiersystem mit Abschluss des Scanvorgangs einen eindeutigen Namen (meistens eine Kombination aus Datum und Uhrzeit).

Im zweiten Schritt werden die Dateien vom Speichermedium an einen Computer zur weiteren Verarbeitung übertragen. Darauf folgt die Benennung der Datei. Abb. 2 zeigt den Aufbau der Datei *BerlinerLuft_Stimmen.pdf*. Wie eingangs beschrieben sind in dieser Datei alle Stimmen hintereinander abgelegt. Eine Stimme startet mit einer Titelseite und endet vor der nächsten Titelseite (Anfang der nächsten Stimme). Weitere Operationen zur Bildoptimierung können optional folgen, bspw. das Drehen der Noten oder das Beschneiden der Ränder.

Im finalen Schritt der Archivierung werden die Scans auf einem geeigneten Speichermedium archiviert. Viele Vereine verwenden dazu Cloudanbieter, damit der Zugriff später für alle Musiker*innen einfach gewährleistet werden kann. Im Sinne der Archivierung und Übersicht werden für alle gescannten Werke weitere Metadaten in strukturierter Form erfasst (z.B. in einem Tabellenkalkulationsprogramm). Zu den Metadaten gehören der Stücktitel, Komponist*in bzw. Arrangeur*in, ggf. das Genre sowie weitere archivarische Anmerkungen (z.B. ob Stimmen fehlen). Neben den eigentlichen Scans und den dazugehörigen Metadaten

⁴ Bei digital vorliegenden Notentexten entfällt der erste Schritt des Scannens sowie Teile des zweiten Schrittes.

werden hier unter Umständen weitere Daten erfasst. Dazu zählen Demoaufnahmen der Stücke, Konzertmitschnitte, falls vorhanden MIDI-Dateien oder digitale Notentexte (z.B. MusicXML).

2.2 Literatureinordnung

Das Thema der digitalen Notenarchive ist mitnichten neu und wurde in der Vergangenheit bereits in vielen wissenschaftlichen Arbeiten und Projekten dokumentiert. Der nun folgende Literaturvergleich stellt eher einen Auszug dar und erhebt keinen Anspruch auf Vollständigkeit. Im Bereich der musikwissenschaftlich motivierten Editionsforchung sind in den letzten Jahren bereits viele Arbeiten entstanden, die sich (nicht ausschließlich) mit der Erstellung von kritischen Editionen beschäftigt. Beispielhaft sind hier die Projekte *Edirom* (<https://www.edirom.de/edirom-projekt>), *Beethovens Werkstatt* (<https://beethovens-werkstatt.de>) oder *FreischützDigital* [Rö15] genannt. Bibliotheken digitalisieren ihre Datenbestände und machen diese für die Öffentlichkeit zugänglich [Di12; Ku15; PKK18]. Von äußerst hoher Relevanz sind die Ergebnisse und Ausführungen aus dem *PROBADO* Projekt, welches sich zum Ziel gesetzt hatte, eine komplette Systemlandschaft, von der Digitalisierung, bis hin zu multimodalen Benutzerschnittstellen zu konzeptionieren und prototypisch zu entwickeln [Da12]. Weitere Literatur ist beispielsweise in den Konferenzbänden der seit 2014 jährlich stattfindenden International Conference on Digital Libraries for Musicology (DLfM) zu finden.⁵

Dieser Beitrag siedelt sich am ehesten an den Ausführungen des *PROBADO* Projektes an. Der größte Unterschied ist, dass die Perspektive eine andere ist: Wendet sich die Ausführungen des *PROBADO* Projektes (und auch der genannten Editionsprojekte) eher an Musikinteressierte oder Musikwissenschaftler*innen, so richten sich die hier präsentierten Darstellungen eher an die Notenverantwortlichen sowie die Musiker*innen und deren Anforderungen.

2.3 Anforderungen an ein digitales Notenarchiv

Die Anforderungen an ein digitales Notenarchiv in einem Musikverein sind vielfältig und wir betrachten im Folgenden nur die gängigsten. Eine der größten Herausforderungen für die Notenverantwortlichen ist es, die Musiker*innen von dem unregulierten Zugriff auf die Originalnoten abzuhalten. Ein Grund dafür ist, dass die Stimmen in Notensätzen häufig in einer bestimmten Reihenfolge vorsortiert sind (z.B. Partitur, Holz, hohes Blech, tiefes Blech, Schlagwerk, etc.). Diese Vorsortierung/Konvention ermöglicht ein schnelles Auffinden einzelner Register und Stimmen. Ist der Zugang zu den Originalnoten für alle frei, lässt sich diese Sortierung nicht lange halten, da Noten aus der Erfahrung heraus nicht

⁵ Link zu den gesammelten Proceedings: https://dl.acm.org/conference/dl_fm/proceedings.

wieder korrekt einsortiert werden nach dem Gebrauch. Im schlimmsten Fall werden für die Probe Originale entwendet und nicht wieder zurückgelegt. Die Notensätze zu erhalten ist eine Grundvoraussetzung für die Orchesterarbeit. Nicht immer sind alle Werke im aktiven Repertoire des Orchesters, werden aber zu bestimmten Anlässen immer wieder benötigt (z.B. Weihnachtslieder). Insbesondere dann erzeugen unvollständige Sätze weitere Arbeit für die Notenverantwortlichen.

Um den oben genannten unkontrollierten Zugriff auf die Originalnoten zu unterbinden, bedarf es einer einfach zu handhabenden Alternative. Der Zugang zu einem digitalen Notenarchiv muss demnach so simpel wie möglich gestaltet werden und von allen nutzbar sein. In der Praxis haben sich hierfür Cloudlösungen bewährt. Der Zugang zum Archiv kann mit Hilfe dieser Lösungen an das aktive Orchester erfolgen bzw. auch ausscheidenden Mitgliedern wieder entzogen werden.

Eine typische Probensituation verdeutlicht die nächste Anforderung: Ein/e Musiker*in ist erkrankt. Die Stimme soll aber in der Probe besetzt sein. Im Normalfall sind dann die Notenverantwortlichen gezwungen ins Archiv zu gehen, die Noten aufzusuchen und ggf. direkt eine Kopie anzufertigen. Dieser manuelle Prozess erzeugt Unruhe in der Probe und schmälert die effektive Zeit für die musikalische Probenarbeit. Insbesondere in der frühen Phase einer Konzertvorbereitung treten diese Situationen in den Proben mehrmals auf. Ein digitales Notenarchiv sollte in der Lage sein, Noten schnell auffindbar zu machen, entweder mit bereitgestellten Suchfunktionen oder einer intuitiven Ablagestruktur (gängig ist hier die Sortierung nach Werken und anschließend Stimmen).

Darüberhinaus gibt es für die musikalisch Verantwortlichen im Musikverein immer wieder Situationen, wo eine Gesamtsicht auf das Archiv von Nutzen sein kann (z.B. bei der Zusammenstellung einer Werkreihenfolge für ein Konzert). Hierzu sind weiterführende Metadaten wie Komponist*in, Arrangeur*in, Schwierigkeitsgrad, Genre oder spezielle Spielanweisungen oder benötigte Instrumente (oftmals in den Percussioninstrumenten zu finden) hilfreich in der Planung. Auch bereits gespielte Konzertreihenfolgen können Inspiration sein bzw. eine Hilfe sein, Wiederholungen zu vermeiden.

An dieser Stelle halten wir die grundlegenden Anforderungen fest:

- Komplette Notensätze vorhalten bzw. den Bestand daraufhin prüfen.
- Einfacher Zugang von allen Musiker*innen von möglichst überall.
- Schnelle Auffindbarkeit einzelner Stimmen um die Probenarbeit nicht unnötig zu verzögern.
- Alle Metadaten im Überblick bereitstellen, ggf. auch Informationen darüber, wann ein Stück gespielt wurde.

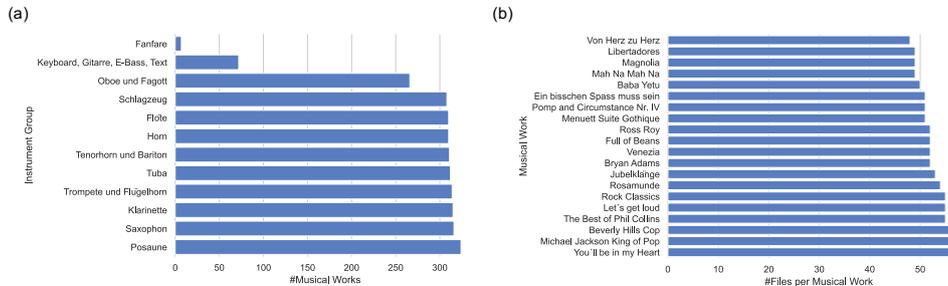


Abb. 3: Übersichtsstatistiken des verwendeten Datensatzes. (a) Zeigt die Anzahl der musikalischen Werke pro Register. (b) Stellt die Anzahl der Dateien/Stimmen pro musikalischem Werk dar.

3 Proof of Concept: Metadatenextraktion mittels OCR

Viele der oben genannten Anforderungen können durch ein geeignetes, relationales Datenbankschema und einer sortierten Ablage für die nicht-relationalen Daten realisiert werden. Problematisch ist, dass insbesondere die alten Notenbestände nur als Scans vorliegen und ggf. auch in völlig unsortierter Form. Im Folgenden beschreiben wir ein erstes Experiment zur Metadatenextraktion mittels Optical Character Recognition (OCR), das wir auf einem digitalen Notenarchiv eines Musikvereins durchgeführt haben. Es handelt sich bei dem Musikverein um ein Sinfonisches Blasorchester, das sowohl konzertant auftritt, als auch im Zuge der Brauchtumpflege (z.B. Volksfeste oder kirchliche Anlässe). Im Wesentlichen bedeutet das für den Datensatz, dass sowohl Noten im DIN A4, als auch im Format DIN A5 vorliegen. Die Forschungsfrage ist, inwiefern existierende, modellbasierte OCR-Programme in der Lage sind, die textuellen Daten in einem gemischten Dokument aus Text und Notentext, zu extrahieren. Als OCR-Programm verwenden wir für diese Experimente das Pythonpaket EasyOCR, da es sehr einfach in der Nutzung ist und in Vorexperimenten gute Ergebnisse erzielt hat.⁶ EasyOCR verwendet zwei hintereinander geschaltete neuronale Netze. Das erste Modell erkennt die relevanten, textuellen Felder im Bild (Detection). Das zweite Modell wandelt diese in Text um (Recognition). Wir nehmen an, dass der Notentext als solcher eher störend für das Modell wirken wird, da diese Art von Dokumenten typischerweise nicht im Training verwendet wird.⁷

Unser betrachteter Datensatz umfasst insgesamt 351 musikalische Werke, die sich auf 12.403 Einzelstimmen (z.B. 1. Trompete, 2. Klarinette, 4. Posaune, etc.) verteilen.⁸ Die Einzelstimmen bilden zusammengefügt ein Dokument mit 17.372 Seiten Notentext. Spezialstimmen wie Partituren oder Einzelstimmen mit mehreren Werken pro Seite wurden im Vorfeld bestmöglich aussortiert. In Abb. 3 sind zwei exemplarische Übersichtsstatistiken

⁶ <https://github.com/JaidedAI/EasyOCR>

⁷ Die Implementierung der Experimente kann unter https://github.com/stefan-balke/sheet-music-archive_GI eingesehen werden.

⁸ In diesem Datensatz sind die Einzelstimmen in separaten PDFs abgelegt. Die Anzahl der Einzelstimmen entspricht demnach der Anzahl der Dateien.

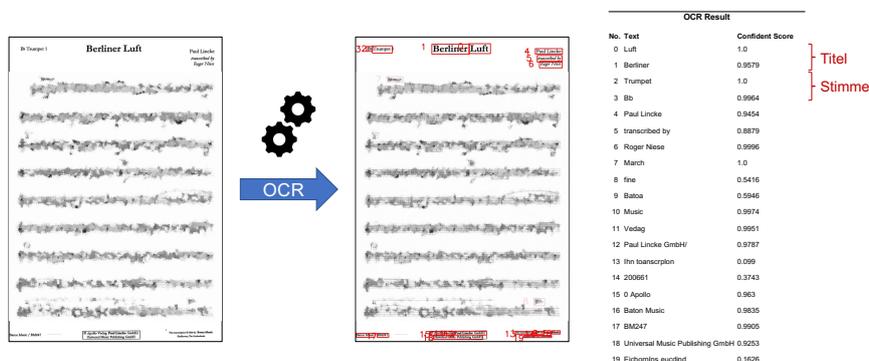


Abb. 4: Darstellung der Verarbeitungspipeline. Das Deckblatt der Einzelstimme wird geöffnet und mittels eines OCR-Programmes verarbeitet. Das Ergebnis des Detection Modells besteht aus Bounding Boxes (rote Rechtecke) und den dazugehörigen, extrahierten Texten als Ergebnis des Recognition Modells (Tabelle). (Aus lizenzrechtlichen Gründen ist der eigentliche Notentext für die Publikation unkenntlich gemacht worden.)

dargestellt. Abb. 3a zeigt die Anzahl der musikalischen Werke pro Register.⁹ Die typisch anzutreffenden Register eines Sinfonischen Bläserorchesters weisen alle eine Werkanzahl >300 auf. Bei dem Register der Doppelrohrbläser (Oboe und Fagott) sinkt die Zahl auf ca. 260 ab. Dies hängt damit zusammen, dass diese Instrumente insbesondere in der traditionellen Blasmusik nicht besetzt sind. Ähnlich verhält es sich mit dem Register um das Keyboard herum, auch diese Instrumente sind oftmals nur in konzertanten Stücken besetzt. Eine Sonderstellung haben die Fanfaren, diese sind lediglich im Zusammenhang eines Sinfonischen Bläserorchesters als Spezialinstrumente zu betrachten. Beispielsweise können die drei Posten des „Der große Zapfenstreichs“ auf diesen geblasen werden und ergänzen dann die Trompeten. Abb. 3b zeigt die Anzahl der Einzelstimmen pro musikalischem Werk (Darstellung wurde reduziert auf die Top 20 des Datensatzes). Bei den dargestellten Stücken handelt es sich fast ausschließlich um konzertante Werke. Wie eben erwähnt werden zum einen in diesen Werken Spezialinstrumente eingesetzt, zum anderen sind die Arrangements komplexer gestaltet, meistens mit 4-stimmigen Posaunen- und Trompetensätzen, 2 Oboen plus einer eigenen Stimme für Englischhorn etc.

Der eigentliche Ansatz zur Metadatenextraktion ist in Abb. 4 dargestellt. Für den ersten Schritt sind wir ausschließlich an dem Titel des Werkes interessiert (im Beispiel „Berliner Luft“). Mittels eines OCR-Programmes wird die erste Seite einer Einzelstimme (hier der ersten Trompete) analysiert und die Texte extrahiert. Als Ergebnis liefert EasyOCR die Bounding Boxes aus dem Detection Modell (rote Rechtecke) und die extrahierten Texte aus dem Recognition Modell, mitsamt einer Konfidenz (Tabelle). In unserem Beispiel hat EasyOCR alle relevanten Textfelder extrahiert, was in 19 Bounding Boxes resultiert. Nach qualitativer Begutachtung macht es dies auch für andere Beispiele sehr zuverlässig. Darüberhinaus

⁹ Ein sog. Register fasst bspw. alle Hörner oder alle Posaunen zu einer Gruppe zusammen.

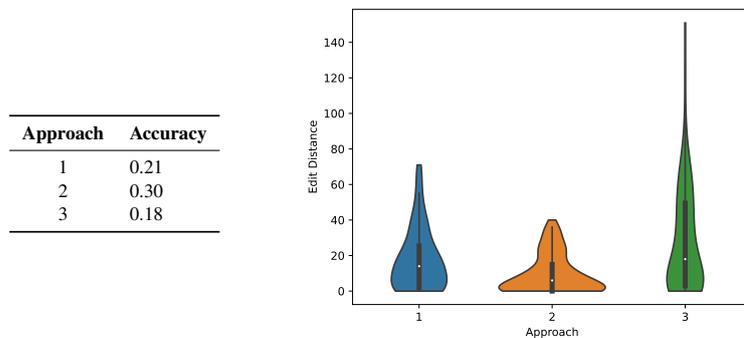


Abb. 5: Evaluationsergebnisse der drei Ansätze zur Titelerkennung. Die Tabelle zeigt die mittlere Accuracy auf den 99 Sample. Die Graphik stellt die Verteilung der Edit Distance der drei Ansätze als sog. Violin-Plots dar.

werden auch oftmals Dynamikzeichen (z.B. p, mf, etc.), sowie Spielanweisungen mit detektiert und extrahiert. Die Extraktion des Textes ist ebenfalls von sehr hoher Qualität, ebenso bei den anderen Beispielen. Beim dargestellten Beispiel zeigt sich, dass Texte, die eigentlich zusammenhängend sind (z.B. der Titel), in zwei Bounding Boxes aufgeteilt wird. Darüberhinaus ist auffällig, dass bei der Stimmenangabe die „1“ außerhalb der erkannten Bounding Box liegt, weshalb der extrahierte Text dies auch nicht anzeigt.

Im folgenden evaluieren wir drei regelbasierte Ansätze, um die Titelinformation zu extrahieren. Wir betrachten im Rahmen des Proof of Concepts einen zufällig gezogenen Unterdatenbestand von 99 Samples (bleibt für alle drei Ansätze identisch). Im ersten Ansatz nehmen wir die größte Bounding Box (Höhe x Breite in Pixeln) und nehmen den extrahierten Text als Titelinformation. Im zweiten Ansatz filtern wir die Bounding Boxes vor. Es werden nur die behalten, die sich in den oberen 20% des Dokuments befinden und wählen davon die größte als Titelnkandidat aus. Im dritten Ansatz filtern wir wie im zweiten Ansatz vor, zusätzlich führen wir die potenziellen Bounding Boxes noch algorithmisch zusammen. Das Zusammenführen der Bounding Boxes ist in EasyOCR vorimplementiert und führt im Wesentlichen eine Gruppierung auf Grundlage der Abstände der Boxes durch.¹⁰

Die Evaluation führen wir auf zwei Arten durch: Zum Einen mit einer Accuracy, die prüft, ob der extrahierte Text exakt der Ground Truth aus dem Datensatz entspricht. Durch OCR Fehler ist dies allerdings ein sehr hartes Maß, weshalb wir zusätzlich noch die Edit Distance berechnen, um ein Gefühl dafür zu bekommen, wie weit der extrahierte Text vom eigentlichen Text entfernt ist. Abb. 5 zeigt die Ergebnisse der drei Ansätze. Die Accuracy des ersten Ansatzes liegt bei 0.21. Das Filtern auf die obersten 20% des Dokuments hebt die Accuracy auf 0.30. Das zusätzliche Zusammenführen der Bounding Boxes ergibt eine Accuracy von 0.18. Qualitativ betrachtet waren die Ergebnisse des dritten

¹⁰ Für das Experiment verwenden wir folgenden Aufruf: `easyocr.utils.get_paragraph(input_boxes, x_ths=0.5, y_ths=0)`. Die beiden Parameter `x_ths` und `y_ths` wurden manuell ermittelt.

Ansatzes am vielversprechendsten, es passiert hier allerdings häufig, dass zu viele Boxen zusammengeführt werden und beispielsweise der Titel mit dem Arrangeur vermischt wird. Ebenfalls in Abb. 5 zeigt rechts die Verteilung der Edit Distances. Hier geht es nicht um absolute Zahlen, sondern rein um die relative Entwicklung (die Edit Distance variiert mit der Länge der zu vergleichenden Zeichenketten). Im relativen Vergleich der Ansätze wird deutlich, dass von Ansatz 1 zu 2 die Verteilung etwas schmaler wird, hingegen zum 3. Ansatz wieder sehr in die Breite geht. Letzteres ist vor allen dem Phänomen geschuldet, dass zu viele Boxen zusammengeführt werden, die Zeichenkette länger wird und somit die Edit Distance proportional anwächst.

4 Zusammenfassung und Ausblick

In dieser Arbeit wurden digitale Notenarchive in Musikvereinen betrachtet. Ausgehend vom Digitalisierungsprozess, über die Literatureinordnung bis hin zur Anforderungsanalyse wurden grundlegende Aspekte und Anforderungen beschrieben und erörtert. Mit einem ersten Proof of Concept wurde gezeigt, in wie fern sich OCR-Verfahren, die für Textdokumente ausgelegt und trainiert sind, eignen, um den Titel aus einem Notentext zu extrahieren. Die Ergebnisse der Experimente zeigen, dass dies als ein erster explorativer Schritt in diese Forschungsrichtung zu betrachten ist. Das verwendete Pythonpaket EasyOCR bietet die Möglichkeit, eigene Modelle zu trainieren, dazu muss allerdings der Datensatz mit weiteren Annotationen angereichert werden. Von der reinen Größe her hat dieser wahrscheinlich das Potenzial nach erfolgter Annotation, die Modelle mit ausreichend Daten zu versorgen.

Im Vergleich zu ähnlichen Experimenten, wie beispielsweise [BAM15], zeigt sich, dass insbesondere die Erkennung der relevanten Bounding Boxes (Detection) ein Hauptproblem für die existierenden OCR Modelle ist. Dieser Schritte wurde in den Experimenten von Balke et al. [BAM15] durch regelbasierte Verfahren übersprungen und es wurde lediglich die eigentliche Text Recognition durch ein OCR-Programm übernommen. Die damaligen Recognition Modellen hatten allerdings qualitativ betrachtet eine schlechtere Qualität.

Die präsentierten Arbeiten eröffnen eine Reihe weiterer Fragestellungen. Die Qualität der Titalextraktionen noch nicht in der Praxis einsetzbar und bedarf daher weiterer Anstrengungen. Insbesondere der Detection-Schritt, also das Auffinden ggf. der relevanten Bounding Boxes weist noch erhebliches Verbesserungspotenzial auf. Zudem könnten weitere Informationsquellen eingebunden werden, um die OCR-Ergebnisse zu verbessern (z.B. externe Datenbank wie <https://de.musicainfo.net>). Ein weiterer Anwendungsfall, der oben kurz angerissen wurde, ist das automatische Trennen der Scans in die Einzelstimmen (vergleichbar zu einer Boundary Detection aus der computergestützten Musikverarbeitung). In einer weiteren Ausbaustufe wird auch die Aufgabenstellung der Optical Music Recognition (OMR) ein Thema werden [CJP20]. Der derzeitige Fokus von OMR liegt vor allen Dingen auf Klaviermusik, die durch ihre Systemanordnung und den regelmäßigen auftretenden Anomalien, ein sehr schwieriges Feld darstellt. Die monophone Einzelstimmen der Instrumente der Sinfonischen Blasmusik könnten hier ein interessantes Zwischenszenario bieten.

Aktuelle Forschung in der Sinfonischen Blasmusik zeigt (z.B. [Ba23]), dass die Musik als solche und die zugehörigen Orchesterstrukturen als interessantes Versuchsfeld im Bereich der computergestützten Musikverarbeitung dienen kann. Durch die ehrenamtlich getragenen Strukturen und vor allen Dingen durch die Möglichkeit kleine Laborumgebungen einzurichten, entstehen viele neue praxisbezogene Forschungsfragen, die allerdings schnell Anschluss zu den derzeitig weltweit diskutierten Fragestellungen des angewandten Maschinellen Lernen finden können.

Danksagung

Wir danken Boris Wortmann vom Musikzug Ennest, der in seiner unermüdlichen, ehrenamtlichen Arbeit als Notenwart den Datensatz für unsere Experimente ermöglicht hat. Diese Arbeit wurde durch die Deutsche Stiftung für Engagement und Ehrenamt (DSEE) unterstützt.

Literatur

- [Ba23] Balke, S.: Sinfonische Blasmusik als neuer Anwendungsfall in der automatisierten Musikverarbeitung. In: Proceedings der Deutschen Jahrestagung für Akustik (DAGA). Hamburg, Germany, S. 1366–1369, 2023.
- [BAM15] Balke, S.; Achankunju, S. P.; Müller, M.: Matching Musical Themes based on Noisy OCR and OMR Input. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Brisbane, Australia, S. 703–707, 2015.
- [CJP20] Calvo-Zaragoza, J.; Jr., J. H.; Pacha, A.: Understanding Optical Music Recognition. *ACM Computing Surveys* 53/4, 77:1–77:35, 2020, ISSN: 0360-0300, URL: <https://doi.org/10.1145/3397499>.
- [Da12] Damm, D.; Fremerey, C.; Thomas, V.; Clausen, M.; Kurth, F.; Müller, M.: A digital library framework for heterogeneous music collections: from document acquisition to cross-modal interaction. *International Journal on Digital Libraries: Special Issue on Music Digital Libraries* 12/2-3, S. 53–71, 2012.
- [Di12] Diet, J.: Digitalisierung und Bereitstellung von Musikdokumenten in der Bayerischen Staatsbibliothek. In: Proceedings der Jahrestagung der Gesellschaft für Musikforschung. Göttingen, Germany, S. 29–35, 2012.
- [Ku15] Kurth, S.: Das Schott-Archiv in der Bayerischen Staatsbibliothek. In: Proceedings des Bibliotheksforums Bayern. S. 297–300, 2015.
- [PKK18] Paakkonen, T.; Kervinen, J.; Kettunen, K.: Digitisation and Digital Library Presentation System – Sheet Music to the Mix. In: Proceedings of the Workshop on Music Reading Systems. Paris, S. 1–2, 2018.
- [Rö15] Röwenstrunk, D.; Prätzlich, T.; Betzwieser, T.; Müller, M.; Szwillus, G.; Veit, J.: Das Gesamtkunstwerk Oper aus Datensicht – Aspekte des Umgangs mit einer heterogenen Datenlage im BMBF-Projekt “Freischütz Digital”. *Datenbank-Spektrum* 15/1, S. 65–72, 2015.