Leveraging Students' Handwritten Notes to link to watched Instructional Videos

by

Ranjitha Jaddigadde Srinivasa

B.E., Sapthagiri College of Engineering, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies
(Electrical and Computer Engineering)

The University of British Columbia
(Vancouver)

April 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Leveraging Students' Handwritten Notes to link to watched Instructional Videos

submitted by **Ranjitha Jaddigadde Srinivasa** in partial fulfillment of the requirements for the degree of **Master of Applied Science** in **Electrical and Computer Engineering**.

Examining Committee:

Dr. Sidney Fels, Department of Electrical and Computer Engineering
Supervisor

Dr. Dongwook Yoon, Department of Computer Science
Supervisory Committee member

Dr. Julia Rubin, Department of Electrical and Computer Engineering
Supervisory Committee member

Abstract

Handwritten note-taking with pen and paper is still the preferred medium to achieve information seeking and comprehension from diverse learning objects. But students, especially in video-based learning settings, exercise laborious practices to re-find the corresponding video context when reviewing notes. We propose the orchestration of students' handwritten notebook content as interoperable links to retrieve previously watched instructional videos. This work articulates the research objectives in two phases. In phase 1, we analyzed the characteristic features of notebook content of watched videos. And, in phase 2, we investigated student expectations and requirements of the proposed video retrieval system.

Analysis of quality handwritten notebook samples and the related video materials in a lab study with ten engineering students revealed distinctive characteristic representations of note content such as text, formula, figures, and chiefly a hybrid of all the 3. A box plot interpretation of notes and the watched video content confirmed that at least 75% of the identified note samples demonstrated a verbatim overlap of 50% or more with the related video content, hinting at its potential use as a query artifact. Additionally, the video references to collected note samples exhibited referencing at three temporal levels: point, interval, and whole video.

A 12-student lab study indicated higher satisfaction for video matches returned at the 'interval' level and showcased students' existing workarounds for linking back to videos. Overall, students rated a positive Mean score for the system's usability to re-find note-specific video context. A medium-fidelity prototype was built, using off-the-shelf computer vision algorithms, to deduce technology requirements associated with the proposed approach. When tested on the 181 identified note samples, the prototype

system matched 77.5% of the samples to corresponding watched videos. The proposed method worked exceptionally well to find suitable videos for textual notes — yielding a 98% accuracy. The note content overlap with the video results further highlights the fragmented nature of the evaluated accuracy across all three temporal levels. Overall, the presented work ascertains the prospect of augmenting prevalent Personalized learning (PL) strategies, such as handwriting notes for future reference, to easily re-find and connect to the watched videos.

Lay Summary

When learning from instructional videos, students often take handwritten notes to improve recall and comprehension. When reviewing their notes, it can be difficult for students to return to the corresponding part of the video. In this thesis, we present the idea of using content available in student notebooks to connect to relevant video context in a video archive. A set of notebook samples were analyzed to explore distinct types of content and references made to videos. A prototype was designed to test the resources and requirements to develop a fully working system. Experimental studies conducted to validate the proposed system demonstrated that it is feasible for implementation. The findings indicated that the students find the proposed approach usable.

Preface

The data collection section of the experimental study discussed in Chapter 3 of this work was approved by UBC Behavioural Research Ethics Board(BREB), Study Number: H13-01589-A022.

I was the primary investigator for the work contained within this thesis. The idea of addressing interoperability between student notebooks and video medium was presented to me by my supervisor Dr. Sidney Fels at the beginning of my degree. Further iterations were my designs. All the ideation, design, implementation, and testing for this project was carried out at the Human Communication Technologies (HCT) Laboratory in the Electrical and Computer Engineering department at the University of British Columbia (UBC) (Vancouver campus).

A condensed version of Chapter 2, 3, 4, and 5 is submitted to a conference, titled “NoteLink: A Point-and-Shoot Linking Interface between Students’ Handwritten Notebooks and Instructional Videos”. I was the lead investigator for this work and was solely responsible for the ideation, software development, conduction of user studies, data analysis, and manuscript composition. Samuel Dodson(School of Information, UBC) and Dr. Kyoungwon Seo(Electrical and Computer Engineering Department) were co-investigators. Dr. Sidney Fels (Electrical and Computer Engineering Department) and Dr. Dongwook Yoon (Department of Computer Science) provided valuable feedback towards the experimental design. Samuel Dodson contributed to the presentation of background work and data collection. Dr. Kyoungwon Seo contributed with results articulation.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

API Application Programming Interface

CAD Computer-Aided Design

CC Connected Component

ECAR Educause Center for Applied Research

FLANN Fast Library for Approximate Nearest Neighbors

GAN Generative Adversarial Networks

HCT Human Communication Technologies

HWR Handwriting Word Recognition

ICAP Interactive Constructive Active and Passive

ISEE Interactive Shared Education Environment

LBP Local Binary Pattern

OCR Optical Character Recognition

OpenCV Open Computer Vision

PIM Personal Information Management (PIM)

PL Personalized Learning

ROI Region of Interest

SD Standard Deviation

SIFT Scale Invariant Feature Transform

SSIM Structural Similarity

TRECVid TREC Video Retrieval Evaluation

UBC University of British Columbia

URL Uniform Resource Allocator

Acknowledgements

I would like to thank my supervisor, Dr. Sidney Fels, for giving me the opportunity to work on this project. It was an exciting idea from the start, and I thoroughly enjoyed the ideation process to shape my work to what it looks like today. I continue to remain excited about the future of this work. I am incredibly grateful for the guidance and unwavering support, even when things did not seem to go my way. Thank you for being so patient and encouraging throughout my degree.

I want to thank Dr. Dongwook Yoon for his insights in the course of building my work. His views on the user study methods, procedures have greatly helped the work address crucial questions. A big thank you is extended to other members of my team—Samuel Dodson and Dr. Kyoungwon Seo—for allowing me to constantly pick at their brains whenever I was stuck and needed help. Samuel deserves a special mention for all his assistance with the thesis organization. I would also like to specifically thank Sameer Sunani for helping me with the initial configurations in the prototype development.

This work was supported by the University of British Columbia Teaching Learning Enhancement Fund (TLEF), the Natural Sciences and Engineering Research Council of Canada (NSERC), and Microsoft. I also thank ViDeX team for their feedback on the work, particularly Gokberk Ozsoy for his contributions in note recognition, especially for formulas. [1]

Lastly, I owe everything to my parents and my husband for being the most significant supporters they are. My parents have always given me the best of opportunities, but this time they encouraged me to move all the way to Canada. Words cannot express my gratitude to my husband, Kiran Kadaba, for all he had done to support me, even when

[1] https://matheqrecognition.blogspot.com/

things were not going my way. I would not be half the person I am without the two support systems.

CHAPTER 1

Introduction

Students frequently engage in note-taking to improve recall and comprehension [40] [85]. With paper-based notebooks, in particular, students maintain the notes in one place and leverage a versatile platform for information gathering practices [63][77]. Pen and paper medium enables students to handwrite flexibly and review notes in many learning situations, such as during lectures, labs, or tutorials. However, it cannot be easy to return to a specific source of the notes, for example, a related chapter in a specific textbook when reviewing for an upcoming exam. This is fundamentally an issue of re-retrieval in pen-paper note-taking medium when information spans across a diverse collection of media services. Re-finding of information, a salient learning activity under the umbrella term of *personalized learning*, can be challenging, especially as students learn from instructional videos.

The adoption of personalized learning approaches has increased significantly in recent years [18]. Personalized learning facilitates an efficient way of connecting instruction to students' preferences, interests, and needs [35] [75]. For instance, when a student takes notes watching an instructional video, he is enabling himself to prepare in case he forgets essential parts of the lecture. He makes a note of equations, diagrams, etc. that are relevant to the discussion, such that when he returns to the part of the discussion later, he can comprehend with relative ease. His learning activity is a conscious or an unconscious effort to be equipped with the required knowledge in future need. This is a classic illustration of personalized learning, which aligns well with the findings from prior research, that emphasizes the intentions of note-taking in video-based learning as predominantly to review the notes afterward [47][18].

Thus, there is an imperative need to better understand the critical attributes of reviewing handwritten notes captured when learning from instructional videos and exploring design opportunities to facilitate effective linking to video information without hindering students' natural note-taking practices.

1.1 Thesis Objective

When learning from notebooks, students often operate re-finding of the source with the help of navigational cues [79][97]. With physical information objects, such as textbooks, navigational cues include properties of the artifact (e.g., its size, color, thickness), paratext (e.g., page numbers and headings), and the relative positions of the information sought (e.g., "about half-way through the textbook on the top of a left-hand page"). In e-books, while some previously-mentioned contextual cues are lost, students commonly use text-search, highlighting, and scrolling to find information [56]. Bookmarks and annotations are other widely used cues to identify important parts of the book [68] [79]. With audio learning materials, indexing the transcripts converted from the speech is relatively easy, similar to e-documents. These types of contextual cues can be less salient in the video.

In the case of instructional videos, as the video content is visually changing with time, students need to navigate the video to find the relevant content, which can be difficult and time-consuming. Fewer and less rich cues are extracted from the visual, auditory, textual, and temporal information such as slide transitions, talking head instructor. Current video interfaces leverage these cues to facilitate navigation and re-finding of video information, through supporting time-linked video-based annotation [44], [62] or table of contents [101], for example. However, re-finding relevant information in a collection of videos whose cues are to be identified from handwritten paper-based notes is an under-explored area of research.

In handwritten notes of instructional video content, students can write down time-codes and slide titles as links/cues to associate to a video [23]. Unfortunately, when these cues are present in a notebook, it can be difficult to look up the related video using a timestamp or slide title when learning from an extensive video collection. Moreover, these links are not always recorded by students because they are time-consuming to create and can become a distraction to learning.

When learning from videos, students often transcribe or record verbatim content in their notebooks rather than paraphrasing [8]. Comprehending, summarizing key points, paraphrasing, and noting them down simultaneously might be too challenging, given the time and efforts in scrolling back and forth. This is especially evident in courses involving technical learning such as mathematics, engineering as pointed by researchers on note-taking strategies [85]. Facts, definitions, and graphical representations are recorded when students believe copying word to word will help recognize source material. This salient feature of notes structure, often termed as *Verbatim transcriptions* by researchers [63] [41] is leveraged in the current work to augment reviewing in students' learning process. Hence, the focus of the thesis is to evaluate the effectiveness of exercising content of notes that are verbatim reference(s) to a specific part(s) of watched videos as query elements to aid content-based video retrieval.

1.2 Research questions

The primary research objective of the current work is to glean the distinctive features and requirements of linking notes to watched videos. First, we investigate the distinguishing traits in the notebook content, such as the types of note representations and the extent of verbatim occurrence of video content in the notebooks. Second, we showcase the requirements of rendering the found matching videos. Students' attitudes towards the ways of rendering videos, along with their expectations, are captured. The following research questions address the above-defined objectives:

1. What are the characteristics of the note content captured when learning from instructional videos?

2. What are the requirements of a watched video rendering system to meet students' reviewing expectations?

Throughout this work, *watched video notes* are the notes captured when learning from instructional videos, and a *video timepoint* as the topically related timestamp in an instructional video.

1.2.1 Characteristics of Watched Video Notes

To answer research question 1, we investigated the distinct types of representations in watched video note content, along with the nature of the reference made, if done manually, to a certain time in the video source. To this end, we conducted a lab study with ten undergraduate students. Examining a total of 181 identified note samples from the data collected led to the following findings:

- Students use four distinctive representations of content types in engineering notes: text, formula, figure, and the hybrid types.

- Most (atleast 75%) of the notes demonstrated a verbatim overlap of some amount with the watched videos.

- The note reference to video is often made to distinct temporal levels in a video.

The recorded observations yielded subsequent exploration of the student expectations and requirements of watched video re-finding mechanism that follows in the next section.

1.2.2 Student Expectations, Requirements of a Watched Video Re-finding System

Two research elements can reflect the ability of the proposed mechanism to re-find watched instructional videos from the already captured notes. First, students' expectations of a prototype system that transfers their existing re-finding strategies. Second, implementation requirements to utilize notebook content as a linking artifact. We briefly outline the incorporated methodology to deduce insights on both the expectations and requirements below:

1. A 12-student study elicits the existing workarounds employed by students to link notes to corresponding video sources. Students' expectations towards a) video timepoint retrieval in the conceptualized temporal levels, b) to-be-rendered video surrogates, the elements of a video to help quick recollection of the included content, were recorded with the help of survey-based questions. Usability evaluation followed by the qualitative analysis of the interview transcripts reports the usefulness and ability of the proposed approach to link note-book content to relevant video materials.

2. Design and development of a a mid-fidelity prototype present the implementation challenges and requirements of using Computer Vision to link handwritten notes to video content. The implementation process employed readily available off-the-shelf Optical Character Recognition (OCR) techniques and image matching algorithms. Testing the prototype against a set of 181 identified note samples revealed salient takeaways of the prototype's linking feasibility.

1.3 Contributions

Firstly, the proposed research work provides a new design solution to link physical notes with corresponding digital video sources. The novelty of the proposed approach is in employing note content instead of special markers or pre-recorded indices to aid video retrieval, which is the primary contribution of this work. The empirical pieces of evidence from controlled lab studies further support the contributions of the entailed work.

1. Firstly, this work recognizes, applies, and articulates the idea of video retrieval using note content. Identifying an application domain, that is, technical notes captured in selected engineering courses, and testing out a design solution to extract design recommendations for that domain is the core contribution of this work. The concept also aims to seamlessly bridge the gap between paper and digital learning materials to better user experience in linking different learning materials. The design application of a medium-fidelity prototype combines paper and mobile device merits where students can comfortably scan and access digital video sources related to the note document. Testing the prototype against a set of note samples also demonstrated the ability of the proposed approach by retrieving the correct video matches for 77.5% of the total note samples. In particular, 97.5% of the text-only notes matched to their corresponding videos also highlighted the advancement in handwritten recognition of textual elements. The design challenges and implications showcase the feasibility of bringing the proposed approach to use. The fact that early users reported that they find the system *usable* shows promise.

2. Content analysis of watched video notes, an important processing step to extensively explore the recognition capacity, is a significant empirical contribution of the work. A controlled lab study investigating ten students' previously captured notebooks reported students' increased use of textual descriptions in notebooks over the

non-textual representations such as figures and formulas. A special case of hybrid type addresses the complexities of recognizing closely connected text and non-text components. Additionally, the percentage of content overlap between the captured notes and the watched videos showed that most (at least 50%) of the note samples comprising of formulas and figures showed 100% overlap. In contrast, textual notes exhibited a more considerable variance of paraphrasing or noting additional information not available in the video materials. Interestingly, box plot distribution indicated that at least 75% of the observed notes with both textual and non-textual content types comprised some amount of content overlap. The findings also suggest that the notes with as low as 25% overlap with the video content also retrieve expected videos. Thus corroborating the design choice of this work, that is to utilize the notebook content to re-find watched videos

3. Outlines inferences on the timepoint in the video rendering. This work specifically addresses the difference between the expected video timestamp and the retrieved video timestamp by leveraging its temporal nature in point, interval, and whole video. Experimental results from a 12-student lab study substantiated the preference for video result playing from any point in time inside the section/interval where the note content occurs. A significantly lower satisfaction score was recorded for a timepoint anywhere in the video than the timepoint in the relevant section/interval in a Wilcoxon signed-rank test ($Z = -2.240$; $p = 0.011$). While the apparent expectation of a video retrieval system is to render the exact matching slide/point as expected, qualitative analysis of the interview transcripts confirmed a higher temporal preference for interval level with 0 negative codes. The overall choice for video rendering in the interval temporal level paves the way for a whole new exploration in the video context to be rendered.

In sum, note-taking is an important learning activity. This paper discusses the utilization of technology support to prevailing note-taking habits employed by students. Specifically, it describes the characteristics and requirements of a technique that lets students link to a video source from which notes are captured. The goal is to allow students to continue to produce notes in natural ways without heeding to efforts that go to re-finding or reviewing later, so having the potential to serve interoperability between handwritten notes and video materials for review purposes.

1.4 Thesis Overview

The thesis is organized as follows:

Chapter 1 Introduction: Introduces the impediments associated with linking of notes to watched video as well as the motivation and objectives of the work contained in this thesis.

Chapter 2 Background: Describes past work done in learning ecologies, video hyperlinks and utilizing handwritten notes.

Chapter 3 Characteristics of Watched Video Note Content: Describes the experimental user study conducted to explore the attributes of content types in notes and how a video timepoint is employed in the linking process.

Chapter 4 Rendering Watched Video Matches: Includes study methods and findings related to preferences for video rendering in temporal levels and related surrogates. Articulates findings on the existing workarounds to link to videos, usability evaluation of the proposed approach.

Chapter 5 Pragmatics of Linking Notes to Videos: Introduces the design and implementation of note recognition, matching, and rendering of video, explains the resources necessary for its implementation. Also, synopses of technology requirements associated with implementation are covered. Accuracy analysis of matching notes to videos is demonstrated.

Chapter 6 Conclusions and Future Work: Presents the concluding remarks for this thesis highlighting a summary of contributions and limitations and discusses possible future directions for this work.

CHAPTER 2

Background

This section covers some of the foundational studies related to linking paper-based handwritten notes with digital materials such as e-documents, video and audio sources. Learning occurs in coordination with a collection of media types. Dodson et al. found that undergraduates in flipped classrooms interact with heterogeneous information ecologies, composed of learning materials that span text, video, and audio [23]. These learning materials are accessed through a variety of platforms, such as video players, learning management systems, and communication backchannels. A challenge for learners is the limited interoperability within information ecologies, which can result in a kind of information archipelagos [43]. The focus of this work is to build on the previous work that highlights fundamental mechanisms to augment cross-linking between notes and video materials in specific. An overview of handwritten notes' characteristics when learning from digital media and the current state of handwritten content recognition is also contained in this section.

2.1 In-house Video Functionalities for Note-taking

Nowadays, increasingly, instructional videos are being embedded in traditional courses. Also, video materials are an essential means of information delivery in massive online open courses (MOOCs) and flipped classrooms [21]. Students often overt distinct learning behaviors in video-based learning settings, such as bursts of asynchronous engagement across various information source materials and(or) selective sampling of video content for future review. Such characteristic video learning behaviors are classified by the Interactive, Constructive, Active, and Passive (ICAP) framework [13]. Behaviors that include physical information manipulation for effective seeking or generation of additional information for comprehension fall under the *active learning* category. The affordances available in the video learning objects, such as browsing, highlighting, annotating or note-taking, are administered to control active learning.

Studies that operationalize the ICAP framework in relation to video-based learning have emphasized positive learning outcomes when students engage cognitively and meaningfully with the materials rather than passively viewing video content. Mitrovic et al. [60] incorporated ICAP framework behaviours in an active viewing system to learn soft skills. The focus was to evaluate conceptual understanding and reflection of video materials by leveraging commenting/rating on videos over passive viewing. Although the findings indicated higher conceptual understanding, the participants found commenting on the videos and rating comments cognitively demanding.

Dodson et al. [22] described students' behaviors in video-based learning using a ViDeX framework that is grounded in the model of active viewing. Common behaviors identified in the previous work and their observations were categorized and assigned to distinct groups of ICAP framework as seen in the figure 2.1. They discuss various affordances exercised in active and constructive behaviors, such as highlighting transcripts (constructive) and seeking subsequent highlighted instances (active).

A growing number of video platforms have included active-viewing affordances to facilitate video-based learning. Pausing, browsing, replay, and archive mechanisms are now standard in-built features. As video players are revamping active viewing, new affordances for constructive behaviors are now a priority.

In the learning activities pertaining to constructive behavior, learners construct their meaning to expand and extend beyond the video content itself [22]. Video platforms

Interactive Behaviors

Communicating with others.
Cooperating with others.
Collaborating with others.

Constructive Behaviors

Making connections between learning objects.
Taking notes to record sense making.
Highlighting video content for future use.
Tagging video content for future use.

Active Behaviors

Browsing for general information.
Searching for specific information.
Triaging between learning objects, such as a video and a textbook.
Re-watching specific video content.
Pausing video content to reflect on the video content or engage in another viewing behavior.

Passive Behaviors

Playing video content.

Figure 2.1: Behaviors of active and passive viewing, categorized using the ICAP framework [22].

administer supporting mechanisms both in-house and external information mediums to facilitate constructive actions. In-house mechanisms support annotation activities such as adding notes to specific points in the video, highlight/underline a video frame or transcript, or marking/tagging a video scene. Overall, they focus on elevating the synchronization between the notes and the video context for easy review. Marshall [58] articulated various functions of annotations from paper to digital note-taking forms, which can transfer to relevant in-house video actions.

First, highlighting or underlining is performed to signal note-taker for future attention on the specific part of the notes. Fong et al., in a video platform ViDeX, demonstrated text-book style highlighting of video content [27]. The subsequent findings corroborated that video transcripts will allow users to highlight, search, and review the video more easily. Second, tagging acts as placeholders to personalize the reflection of an important point in the video. TagLecture [30] introduced a game to motivate the need to increase the number of tags associated with video lectures. Consequent investigation revealed that youth and almost 50% of the adults did not prefer playing the tagging game due to the type of labels: “they think they are too personal and subjective to tag video lectures properly”.

Next, in-video note-taking allows recording reflections or interpretations of the information present. The interfaces today support in-house note-taking that is timestamped to the current video time for efficient review. A tool called Interactive Shared Education Environment (ISEE) [62] explored the issues in video annotation. It generates hyperlinked timestamps, Smartlinks, to associate the notes with video contents. Similarly, Dorn et al., examined the use of spatiotemporal anchored collaboration affordances to enable collaborative annotation and discussion of video content as a first-order learning activity [24]. Results indicated that students take advantage of the system’s affordances to interact in meaningful ways, though overall student annotation authoring is restricted. The limitations of in-video note-taking mechanisms are the linear-only capability, meaning there is less room for physical non-linear formatting like adding doodlings/markings of reflections around the notes taken.

NoteStruct [52], in-video note-taking system, prompted learners to perform a series of note-taking activities. First, learners highlight, comment and add questions while watching the video. Second, learners walk through every highlight they made and elaborate or merge key points. In the final phase, learners review all the notes generated in steps 1 and 2 in a free-form text editor and are free to edit parts of the notes. The free-form editor partly allocates non-linear formatting but fails to transfer the free-form formatting affordances found in the paper medium. Also, the 3-step note-taking is both time-consuming and demands additional effort when learning from videos.

Furthermore, Note-taking with a video annotation system was compared with handwritten note-taking when learning from videos by Bargeron et al., [2]. Results indicated that note-taking with MRAS, a video annotation system, consumed a lot more time than traditional note-taking, even though the number of words captured was the same on average. However, it is important to note that participants preferred MRAS over handwritten notes because MRAS was able to contextualize the notes with the corresponding video timestamp. This very observation indicates that retrieving the context of captured notes is an important activity to support students' constructive behaviors.

The preceding discussion pointed to several video platforms with features for textbook-style highlighting, note-taking, and tagging, along with their contributions to transfer physical annotating practices. Nevertheless, previous work suggests that students often take notes with paper-based notebooks, even when provided with video annotation tools [36] [94]. As Sellen and Harper [77] points out, the paperless office is a myth. Sam et al. found that notebooks are used to synthesize, organize and orchestrate important information from multiple class materials in one place for use [23]. The authors also emphasize that students prefer interoperability over medium-specific tools. Thus, in the following section, significant work of cross-referencing video materials from paper notes is covered.

2.2 Cross-Media Interaction with Paper Notes

The key characteristic of students learning from various information objects is to acquire and keep items of value, as defined in an umbrella term Personal Information Management (PIM) [92] [4]. The onus is often on the students to manage their information in the right place, in the suitable form to meet their learning needs. Microsoft's OneNote, for example, provides interactive features for note-taking but also constrains the use of a tabbed platform to aid the organization of notes and references to other information sources. Several researchers have highlighted the general lack of progress within HCI towards closing the gap between the affordances of paper and the benefits of digital media [5] [4].

Hypertext is often thought of digital environments' content, particularly the Web; however, hypertext can also exist in non-digital environments [50]. For example, Marshall's studies of students' textbook annotations suggest that annotations are conceptual hyperlinks within and between Lexia [58]. Literary theorists have argued that hyperlinks can have various extents or scopes [50], allowing links to reference objects at different levels of granularity. For example, a student can create a 'word level' link by defining an unfamiliar word from her textbook or a 'collection level' link by writing a note that summarises a chapter. Other scholars have explored the linking between media. Augmenting paper-based documents within digital information [34], in particular. For example, a lepidopterist's paper-based fieldnotes can be linked to photographs of butterflies [99] [54].

Embedded Media Markers (EMM) are indicators, frequently employed in terms of glyph codes, barcodes, or transparent marks, signifying the availability of additional media associated with that part of the document. Post-captured in-video notes are printed on paper to include place markings or marginal markings that promote incidental reflections of material circumstances. Lynn D et al. developed a system for video access from notes on a paper medium [93]. The invention includes a note-taking system that allows a user to select keyframes or make annotations during a note-taking session (meeting, presentation, or other activity); the captured annotation indexes a position of a video. The captured keyframes and annotations are printed on a paper sheet with glyph encoding. The printed sheet is used to access the relevant parts of the video by scanning the glyph codes associated with the target keyframe or annotation. The printed sheet can then be annotated or elaborated in free-form, allowing non-linear formatting of the captured notes.

PaperChains [69] connects paper-based material with digital audio using low-tech, easily accessible equipment. The authors make use of standard camera phone handsets to photograph physical content and then allow the photographed item to be augmented with digital audio in multiple locations via interaction with the image. The approach uses two precisely placed QR (quick response) codes printed on paper, allowing the PaperChains system to detect the item, its orientation, and dimensions, with any camera-enabled mobile handset. This method requires no additional specialist hardware (such as a dock or camera-pen) – users interact directly with the photograph using the phone as a proxy. When reviewing the audio files, touching anywhere on the photograph plays the audio for that location. Although time-consuming, utilization of EMM-type indicators synchronizes notes contextually with the corresponding video material. But the limitation here is the onus in maintaining media-only special paper notes and also accomodating the associated technology requirements.

More recent work in this area has adopted technology such as Anoto, which uses a special dot-marked paper and a camera to recognize document areas while writing. ChronoViz [28] integrates researcher's paper notes into the composite time-coded data set of video files. It exploits Anoto digital pen technology to support the integration of paper-based digital notes. The pen includes an onboard infrared camera and tracks its position on the paper in real-time by reading the dot pattern and, in turn, making it possible to navigate data in more flexible and powerful ways.

Today, this work continues with research exploring how novel tools can support paper-based note-taking practices for active, constructive learning when interacting with the video. We address the imperative need to nurture interoperability between students' paper-based notebooks, central to information weaving of videos and other media, and the corresponding context. The integration of interactive navigation techniques with existing note-taking practices enables students to navigate to moments noted during initial learning activity quickly.

2.3 Searchable-Video Data

Current work on instructional video indexing and matching gathers the visual, textual, and auditory streams of data from videos, opening up the content of instructional video for topic matching through traditional information retrieval approaches [45] [83] [87] [97] [96]. TRECVid [1], a workshop series, has been pioneering content-based video analysis, retrieval, and detection since 2001. Of special importance is their semantic indexing task, assigning semantic tags to video samples, that aims at evaluating methods and systems for detecting visual, auditory, or multi-modal concepts in video shots. The methods follow a "bag of visual words" matching approach or more elaborate aggregation methods like generating Fisher Vectors or SuperVectors. In VCenter [37], for example, Hsiao and Wang segment video into a series of frames from which only the most representative frames are used for indexing. iVIEW [53] is another system that supports full-content searching of multilingual text and audio extracted from the video.

Temporal linking appears to be a widespread approach in video platforms for entertainment too. In a study of YouTube comments, Yarmand et al. found that comments reference a variety of temporal aspects of video: from single points to intervals to whole videos [98]. Our approach utilizes all the visual, textual, and auditory data available from the video materials to deduce matched results of videos from their corresponding notes. This work also explores the effectiveness of linking note-based video content to temporal aspects, as suggested in the previous studies.

2.4 Note Content Characteristics

Notebooks fulfill a vital function for students because free-form writing, in particular, allows them to leverage a greater breadth of expressions [77]; especially when, in our case, writing notes on the content that appear in the videos, which include equations, figures, and tables. Real-time notes are unwieldy: most contain mixed cursive and graphical writing, often written at angles, in various sizes, and at multiple locations all over the page. These constraints do not change the problem statement but require a more profound investigation into the content of the notes.

Analyses of the lexical structure of notes by Piolet et al., [72] showed three characteristics of notes. First, abbreviating procedures, for instance writing down 'poss' for 'possibility'. Second, symbols substituting for syntax such as equals, star, arrows. Third, the physical formatting of the notes in a non-linear way. As it is possible that each individual uses different abbreviations/formatting techniques, or in some cases the same individual can use different abbreviations/substitutes for the same word in various parts of his notes, the current work does not observe the stated characteristics. Rather, the focus of the work is to examine the features of notes in relation to the video materials.

In the context of recall performance evaluation, several characteristics of notebook content have been identified in empirical studies, including note quantity [63], note quality [70], and verbatim overlap [63]. *Note quantity* is the number of words while *note quality* is measured by counting the number of factual statements in the notes compared to the lecture video, which can be judged as either right or wrong, present in the notes. *Verbatim overlap* stands for word-for-word overlap between the video content and the content of the notes. In this work, we focus on the content of notes that can be overlapped and matched with the video content.

2.5 Handwriting Recognition

The notebook content, written in relevance to instructional videos, must be translated to machine-readable information to assist the semantic lookup of the video source. Recognizing handwriting content remains an active area of research [73] [86]. A major area of focus has been on handwriting recognition through the analysis of stroke data, such as the direction and order letters are drawn [81] [19]. This information is readily captured by devices that support digital ink, such as mobile phones and tablets. However, students continue to use pen and paper for taking notes [36], meaning stroke data is not always available. Applications in off-line handwriting recognition through various domains include recognizing postal addresses [80], bank checks [29], writer identification [65], historical document recognition [65], and calligraphy imitation [102]. Various studies comparing on-line versus off-line handwritten recognition [73] [86] have articulated the underlying methods and gaps in the respective categories.

Yuan and Seales [100] proposed a semi-automatic solution to reading manuscripts that involves a document analysis (DA) module and a graphical user interface. The DA module detects and ranks regions of interest in an image using which users can manually configure the parameters of the DA module, thereby ranking candidate regions for linking. Our work, on the contrary, proposes a fully automated solution that identifies regions of interest for each image in every video from the library and delivers a ranked list of videos without user intervention.

Alternative approaches to handwriting word recognition (HWR) have demonstrated substantial improvements in recognition accuracy by employing a combination of convolutional and recurrent neural networks [12] [71] [74] [102]. Deep-learning approaches require a large amount of training data to discern the most important features for character recognition. Collecting and annotating a sufficiently large dataset of handwritten notes with connected components in different settings remains expensive and laborious.

To our knowledge, no such dataset is publicly available. Therefore, in this work, the medium-fidelity prototype is built upon readily available off-the-shelf Optical Character Recognition (OCR) engines whose application can be scalable to different note types.

CHAPTER 3

Characteristics of Watched Video Notes

This chapter aims to investigate *how* and *what* types of notebook content link to *what* timepoint in a video. The diversity of the captured note content and the reference point in the video material are surveyed and discussed. Identifying students' notes' distinct characteristics is especially trivial to inform the selection of effective recognition techniques. We also deduce a taxonomy of handwritten expressions and the temporal levels of referencing the corresponding video timepoint. Additionally, a dataset of watched video notes is created, which is later used for testing the performance of an automated linking prototype.

3.1 Notes Collection

In this section, previously taken student notes on instructional videos are analyzed for any information that could connect to the video. We inspect the identified information to uncover the attributes of 1) the content type of the notes and 2) the corresponding destination within the video, allowing us to compare the students' notes to the correct point in the video.

3.1.1 Participants

We recruited ten university students — four from an electromagnetics course, two from machine learning, three from a data analysis course, and one from a crash course on software engineering. The study targeted students from engineering courses specifically to increase the likelihood of finding diverse types of technical notes. 4 of the 12 students were undergraduates and 6 graduates, all of them between 18 to 28 years old. Table 3.1 shows the demographic data of the ten students, their level of education at the time of the study, and the course from which the notes. The table also includes the number of videos and the average length of the videos in each course. The inclusion criteria were that each unit of sample possessed pre-written notes recorded on paper notebooks watching instructional videos. Before the study, each participant confirmed they had at least eight pages of handwritten notes they had taken while learning with video that ensured a sufficient amount of data in each session.

3.1.2 Procedure

Students located pages in their notebooks containing handwritten notes related to a point in an instructional video. Each instance of a video-related note was recorded as a link to populate a dataset of notebook-video mappings. A flatbed scanner recorded note pages to maintain the resolution, luminance, and color quality of the data collected. The notes collected were full-page copies of students' notebooks. Given that we made copies of students' actual notes, they might have included personally identifiable information or content that they would like to exclude from the database. Therefore, each participant was provided with a plain piece of paper to conceal all the content they were not comfortable sharing.

Ground truthing was a crucial step for understanding the linking patterns and validating the core idea of note-content-based video access. We collected student-labeled ground truth observations along with the actual copies of video-related notebook content. The course videos that students learned from were available in a video learning platform ViDeX [27] prior to the study. We used a laptop to view the scanned notebook pages and the videos to which the notes belonged. Once the copy of relevant notebook pages

Table 3.1: Demographic data of Participants P1 to P10

Participant	Gender	Level	Course	No. of Videos/-Course	Average video length (hh:mm:ss)
P1	Female	2nd year, Undergraduate	Engineering Electromagnetics (ELEC211)	15	0:07:56
P2	Female	2nd year, Undergraduate	Circuits & Electromagnetics (BMEG220)	16	0:08:03
P3	Male	1st year, MASc	Machine Learning (CPSC340)	12	0:50:43
P4	Male	2nd year, Undergraduate	Circuits & Electromagnetics (BMEG220)	16	0:08:03
P5	Female	2nd year, Undergraduate	Engineering Electromagnetics (ELEC211)	15	0:07:56
P6	Male	1st year, MASc	React, JS	1	2:25:26
P7	Male	1st year, MASc	Machine Learning (CPSC340)	12	0:50:43
P8	Male	2nd year, Meng	Data Analysis (EOSC510)	13	0:18:41
P9	Male	1st year, MASc	Data Analysis (EOSC510)	13	0:18:41
P10	Female	2nd year, MASc	Data Analysis (EOSC510)	13	0:18:41

was taken, participants marked a rectangular box around the content that can be linked to the video on each page using editing software like paint, PowerPoint slides. Each marking included the following details as shown in 3.1.

1. Name/title of the video.

2. Timestamp of the referenced video material that links to the note content.

3. A confidence interval representing their level of confidence, measured using the

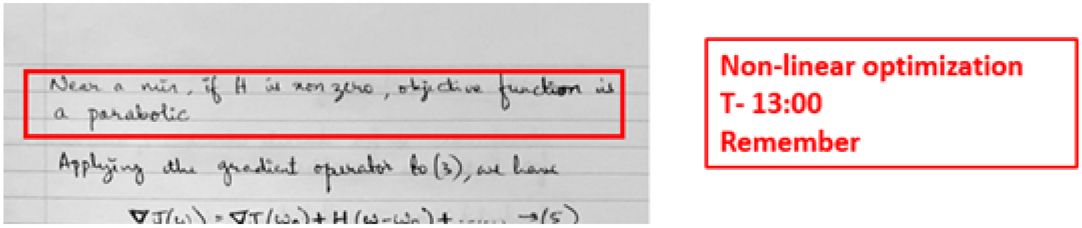

Figure 3.1: An example of a student-labelled ground truth of a scanned note, recording 1) the name of the matching video, 2) the timestamp within the matching video, and 3) the students' Remember, Know, Guess confidence interval.

Remember, Know, Guess paradigm [25], of their responses. Participants reported if they remembered the video occurrence or knew about the occurrence or guessed if they took some time to go through the videos and guessed the occurrence.

We returned students' notebooks after completing the data collection session, which took approximately 45 minutes, during which we gathered as many links as possible. Pilot testing with four students prior to the data collection recorded any impediments through the process that needs attention.

3.1.3 Description of the Captured Notes

Five to ten pages of notes per participant were recorded, comprising a minimum of 10 to a maximum of 35 video links. An important aspect to note is that each student took considerable time to scroll through the videos and the video timeline to record the relevant timestamp. However, the time taken to identify links did not indicate a general pattern. Also, we did not see any consistency in the number of notes per student and their capacity of recalling the related video context as Remember, Know, or Guess. For instance, two students from the same course had contrasting patterns of link identification. P1 marked 9 note links to 9 different videos, whereas P5 marked 20 links to 6 different videos, with at least 2 links pointing to each video.

Students marked most video links with a confidence level of Remember when asked to indicate the basis on which they judged that they had studied the video previously. The apparent reason could be as students had the videos included as part of their ongoing curriculum. They marked two links as Know and four as Guess. One student tagged a part of the note page as video-related content but did not remember the video and marked the confidence level as 'Unsure'. In total, we collected 181 video links which can be seen in detail in A.1 and analyzed for various characteristics so that we can apply appropriate recognition analysis in the later phase. It is particularly imperative to identify the types of note content that exist to aid the recognition of handwritten content and thereby elicit design ideas to support the robust linking of notes to videos.

3.2 Results and Discussion

The identified set of notes and their corresponding video references set the stage for three main subtasks:

1. Understanding the similarities and differences between the various content types of the notes.
2. Developing a taxonomy of similar notes.
3. Understanding the temporal nature of the video references.

3.2.1 Content-based Categorization of Note Content

We observed the links for collective identities whose characteristics were derived based on the subject matter in the note content. Previous work has demonstrated the distinction between textual and non-textual content types in handwritten notes [20] [38]. They emphasize the need for assigning a specific content type to corresponding specialized recognition systems for high-quality accuracy. Accordingly, the following four main units categorize the links based on the content type:

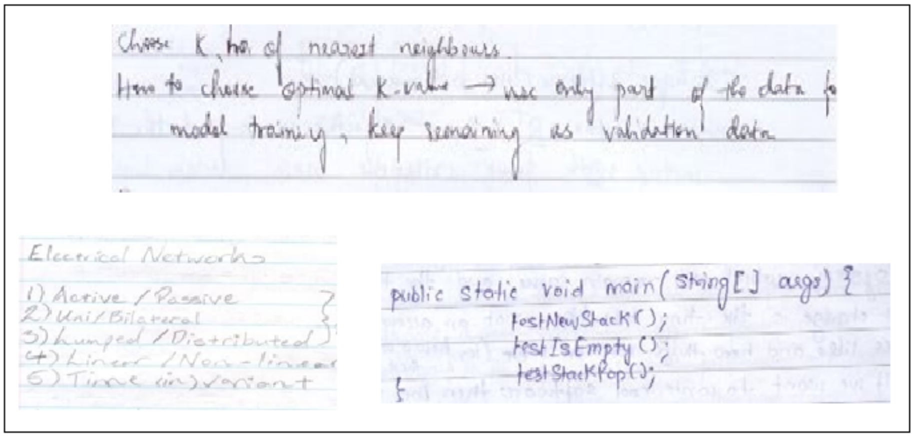

Figure 3.2: Notes with text-only content

1. Text: A link that has plain textual content. This type includes groups that have textual descriptions such as comparison table, a summary of a concept, definitions, bulletin of important pointers, freestyle flowcharts, piece of code. An example for each of these groups is shown in 3.2.

2. Formula: A link that includes mathematical symbols, equations as shown in figure 3.3. One line equation or a block of derivation were both regarded as formulas.

3. Figure: A link that comprises handwritten illustrations of graphical elements. This includes pictorial representations of electromagnetic elements, graphs, circuits whose examples can be seen in figure 3.4.

4. Hybrid: A link that uses a combination of the above three types in close connection as shown in figure 3.5. An example could be a figure with one or more lines of textual details of the figure.

The categorization was devised based on the examples described in the hierarchical annotation of online handwritten documents [38]. Mapping video-related note content

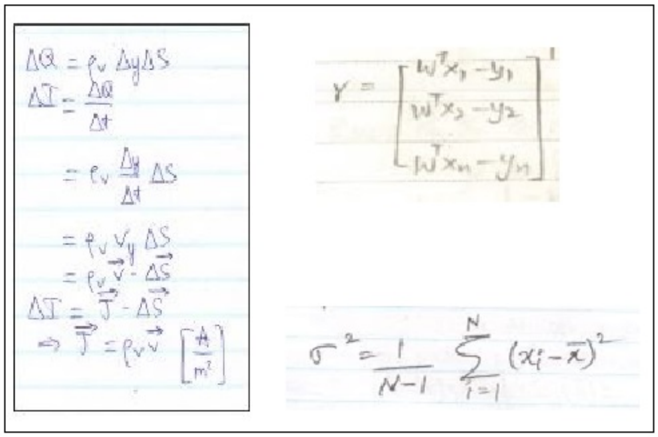

Figure 3.3: Notes with formula-only content

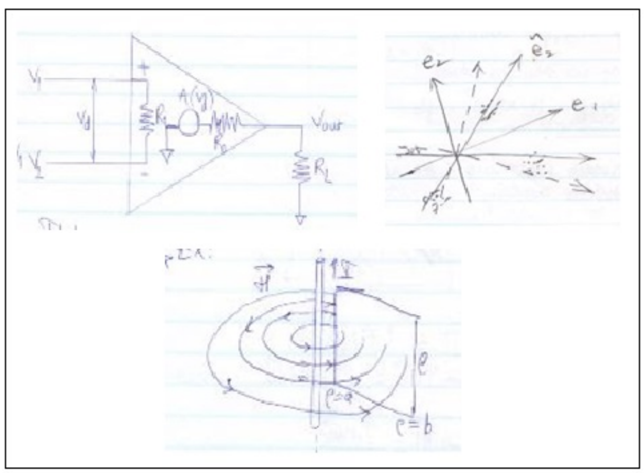

Figure 3.4: Notes with figure-only content

to the categories mentioned above showed that most links were text-based. 39.22% of the references (n = 71) were written letters or words, 34.25% (n = 62) were formulas, and 4.97% (n = 9) were drawings of figures/drawings/circuits, not surprising given that our participants are student engineers. We also found 21.54% (n = 39) of hybrid cases representing a combination of textual and non-textual content. It is important to note that textual content type accounted for the highest usage of the video links data collected, indicating its broader use over other content types to link to videos. This is consistent

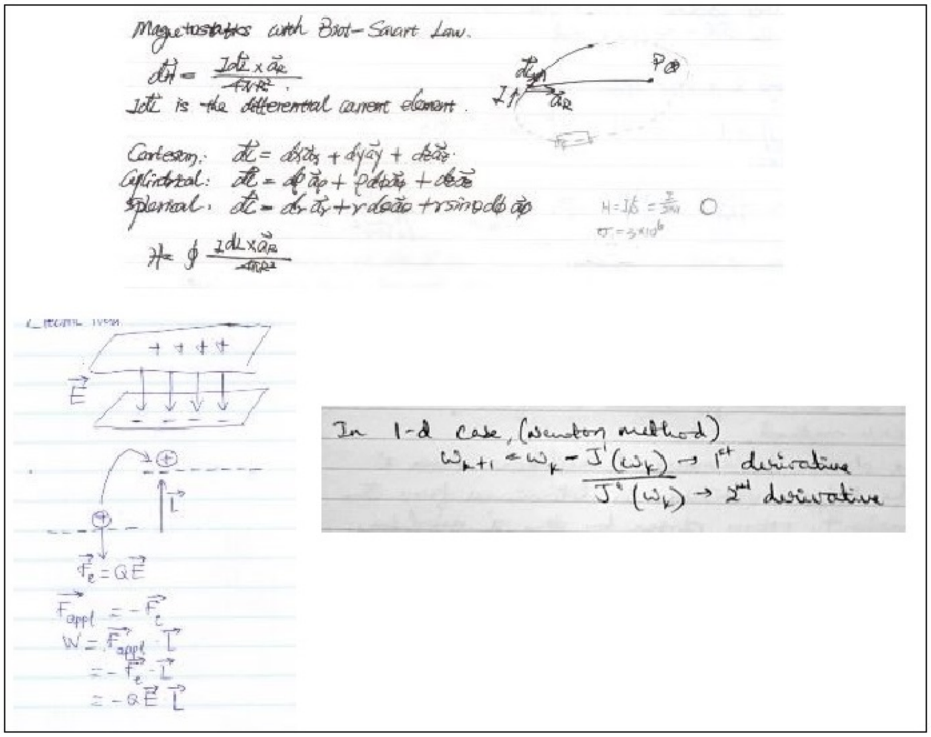

Figure 3.5: Notes with a combination of textual and non-textual content

with the findings from previous work [38] that reports text as the frequently used content type in online handwritten documents.

3.2.2 Unfolding the Hybrid Type

In hybrid content types, textual and non-textual contents haphazardly amalgamate to structure a reference to some point in a video. This poses a challenge in assigning a recognition technique to hybrid types found in notes. Further scrutiny into the samples of this type displayed that one of the 2 or 3 types dominate the other in obvious ways, as listed below. We leveraged this factor of the predominance of one type over the others to unfold and assign the hybrid type to text, formula, or figure.

1. The content load of one type exceeds the other types. This can be seen in the sample shown in figure 3.6. The link comprises all three types; text, formula, and

figure. However, it has more textual characters and letters when compared to mathematical symbols and pictorial representations. So, the content type chosen for this case is text.

2. The topical similarity between one content type and the video far exceeding the other(s) distinguishes the hybrid types. Figure 3.7 shows one such example. Although both content types have about an equal number of characters, the text 'sample mean' here is not unique and can be semantically related to several videos. However, the equation can be a finer query that could be semantically close to the expected video. Hence, the formula was the choice in this case.

3. Sometimes, the significance of a content type is manually highlighted in hybrid notes. In the sample shown in figure 3.8, equations are highlighted in textual boxes to indicate its significance. Hence the accentuated type was chosen, in this case, formula.

   We applied the above-discussed directions to all the hybrid cases to assign each case to one of the other three content types. In the end, 98 of the total records were tagged as text type, 65 as formula, and 18 as figure.

3.2.3 Verbatim Overlap

The Verbatim overlap is a measure of the word-for-word overlap between the lecture content and the content of the notes. Mueller and Oppenheimer [63] explored the effects of verbatim overlap in three studies (n = 65, n = 149, and n = 109). In all three studies, all participants took notes, either by hand or with a computer. The authors scored these notes on their verbatim overlap. They compared a fixed chunk of text in the notes with a chunk of text in the lecture transcript and reported a percentage of match for each.

We followed a similar procedure to calculate the percentage of matching for each

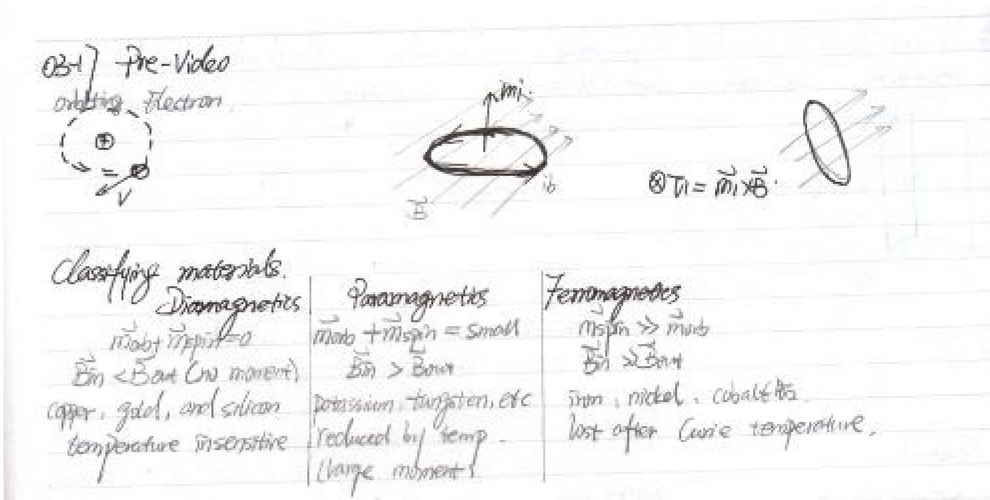

Figure 3.6: Hybrid link consisting of connected text, non-text components

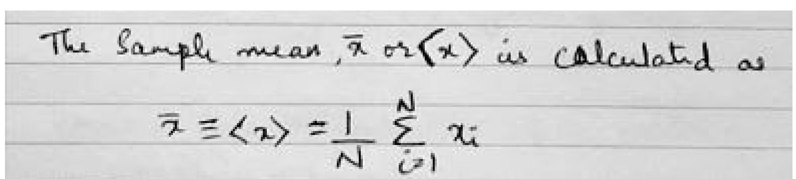

Figure 3.7: Hybrid link with semantically broad text content

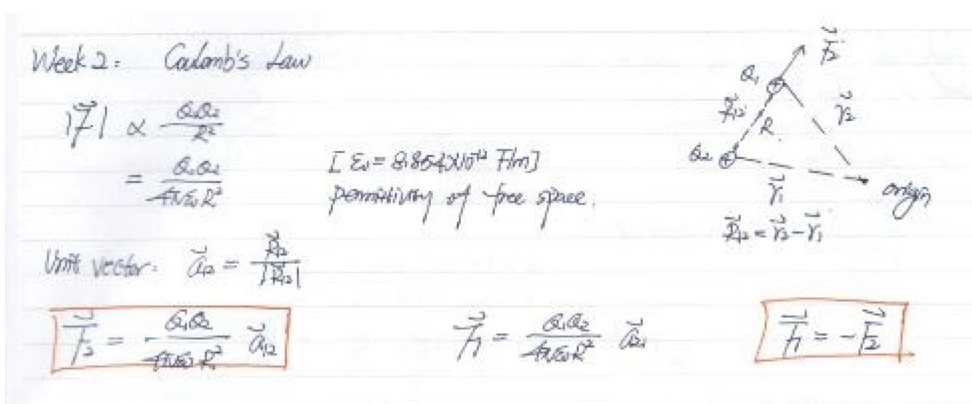

Figure 3.8: Hybrid link with highlighted content

recorded link in the notes. As the content of the notes included both textual and non-textual components, each chunk of the components in the notes was systematically compared with the matching chunk in the video with the following guidelines.

- Firstly, for text, each word that is a noun or a term that has lexical significance and is not a stop word was considered as a chunk of data. Stop words are commonly used words, such as “how”, “what”, “or”, that are excluded from searches to parse information objects faster. We identified the chunks discretely irrespective of the sequence of their appearance.

- For formulas, defining a chunk was not straightforward as the length of equations could vary from a simple two-term equation to a number of lines. Therefore, an equation connected by less than or equal to 3 terms on either side of the equals sign was regarded as a chunk. If an equation exceeded more than 3 terms, we assigned it to the next chunk.

- In figures, a graphical representation that has connecting nodes from one end to another was considered one chunk. As we looked into freehand paper drawings, if the figure was structurally similar to the one in the video, the overlap was true otherwise false.

- For the hybrid type, we identified the chunks according to the type assigned. For example, if the hybrid representation has a higher number of graphical components, we followed the figures’ procedure to identify and match the chunks.

Therefore, we computed the percentage of content overlap as:

$$\%\ of\ overlap = \frac{\#\ of\ matching\ chunks}{total\ \#\ of\ chunks\ in\ the\ notes} * 100 \tag{3.1}$$

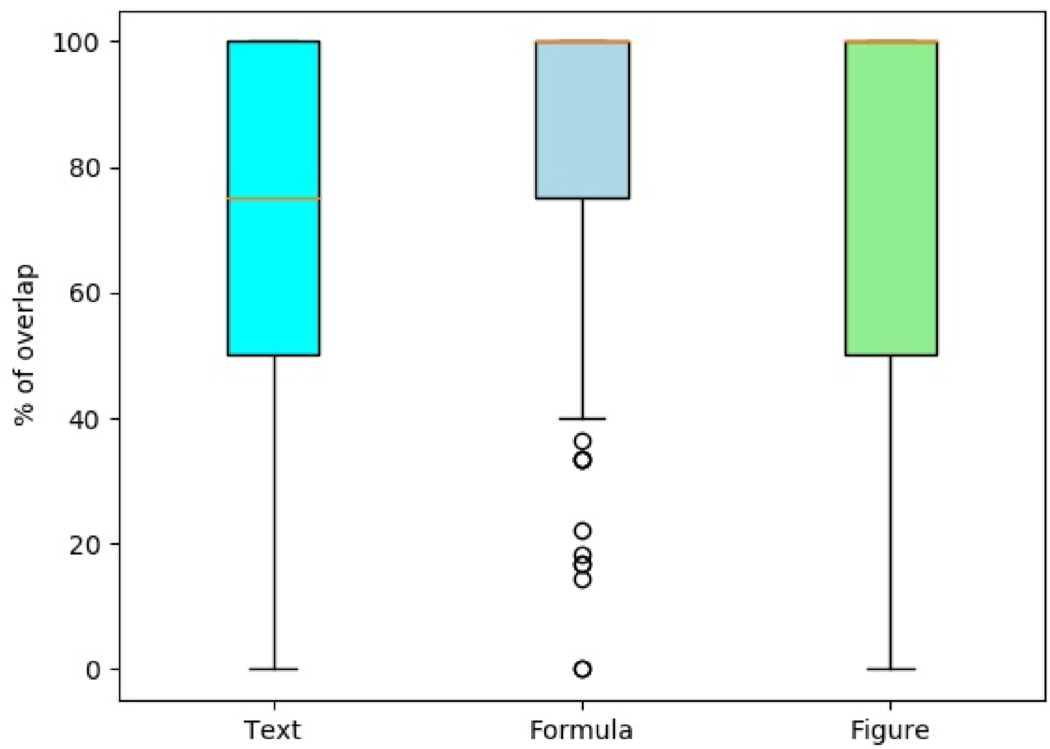

Figure 3.9: Box plot comparison of content overlap in the text, formula and figure representations respectively.

The box plot in figure 3.9 visualizes the overall distribution of the content overlap across text, formula, and figure representations in the watched video note content.

A glance at the plots suggests that the Inter Quartile Regions (IQRs) of all the three boxes are above the 50% overlap. This very fact implies that at least half of the note links identified in each type, that is, text, formula, and figure, matched half or more of the content chunk for chunk as found in the videos. We outline some of the observations specific to each type below:

- In the case of texts and figures, at least 25% of the note samples showed an overlap of less than 50% giving rise to two possible interpretations. One, students completely paraphrased the content in the notes in one's own words or two, the matching video timestamp as indicated in the ground truth did not exactly match the context of the notes. The IQR for the text type also exhibits a larger variance in the 50-100% window of overlap compared to the other two types. This indicates that with the text type, there is more paraphrasing when compared to formula and figures.

- The median is the lowest for the text type pointing to 75% of overlap and is at 100% for the formula and figure. Formulas in specific manifest 100% overlap in the case of at least half of the formula-based notes. This implies that when learning from videos, formulas are transferred literally to notes for reviewing purposes.

- Furthermore, for both text and figure, there are cases when there is no overlap at all, with a minimum overlap of 0%. But, it is extreme in the case of formula, indicating some amount of obvious overlap for most cases.

3.2.4 Watched Video Timepoint

Investigating how students marked the expected timestamp of video to each part of notes was crucial in determining the predicted outcome of searching through the videos. We examined the timestamp — i.e., the link destination — recorded as the ground truth for each collected note page and registered the following observations:

1. The note content did not always point to the video slide as indicated in the data collected.

2. Expected timestamp did not always indicate the start of an interval. The video content of the notes spread across a stretch in the video. When a software code is explained, for instance.

The recorded timestamp with the video title failed to indicate if a student is trying to re-find the whole video, an interval of similar content in the video, or a particular slide. Thus, there is added ambiguity in the video timestamp to be rendered that corresponds to expected. It is crucial to rationalize the video timepoint rendering evaluation to articulate the ability of an automated linking approach.

To this end, we conceptualized the timestamp to be rendered on the temporal attributes in the video content. Previous work has emphasized the provision of temporal context within search results in video [9] [90]. Yarmand et al., [98] presented a taxonomy for classifying video references in YouTube comments by temporal specificity. The authors identified three temporal characteristics of video referencing as *Point* that references distinct moment, *Interval* that references a span/section of time, and *Whole Video* that references the entire video. Accordingly, we define the retrieved video timepoint as a timepoint in the three temporal levels:

1. Point: when a video timestamp is pointing to specific content in a video,
2. Interval: when a video timestamp is pointing to an interval of the related video section,
3. Video: when a video timestamp is anywhere in the video

3.3 Summary

We performed a needs assessment by collecting previously taken notes of watched videos through a lab study with ten students to crucially analyze the types of notes and their attributes to being used as query elements. First, we articulated the outlook of identifying various information types: text, formula, drawing and hybrid, in the watched video notes. It is important to note that textual content type accounted for about half of the video link data collected, indicating its broader use over other content types to link to videos.

It is also important to note that the extension of hybrid type to one of the other three types also lead to a majority of text-dominant links; 26 out of 38 hybrid notes, to be precise. This is consistent with the findings from previous work [38] that reports text as the frequently used content type in handwritten documents.

In addition to the identified types of notes, investigation of the verbatim overlap factor also yielded interesting insights in relation to each type. The box plot interpretations implied that students tend to copy important formulas and figures as it is from the video slides, but paraphrase or summarize the textual explanations into their notes. Overall, at least 75% of the notes in each type showed some amount of verbatim overlap, confirming the ability of watched video notes as efficient query elements to retrieve related videos. While the 100% overlap of content between notes and corresponding video in at least half of the identified formula and figure note samples, and in 25% of the text samples, is a promising finding to aid the proposed approach, 0% overlap for text and figure samples poses a trivial challenge.

We conceptualized the timestamp in the resulting video on the temporal scale to a timepoint in point, section, or whole video. We believe the problem of accounting for a difference between the retrieved to expected video timestamp has been addressed with the temporal attributes through this approach. However, student reviews on the video retrieval in the 3 levels and other requirements in the video viewing end are important to determine the comprehensive features in the future systems. On the whole, the established characteristics of both note content and the corresponding watched video address the first research question of the work. In the following chapters, the emphasis is on exploring how to link back to the video source material in one of the three temporal levels. Also, we address the optimal temporal level among the three groups in the retrieved video.

CHAPTER 4

Rendering Watched Video Matches

Students' attitude on employing a linking device as an educational tool is vital to design, evaluate and roll out the right technology. As discussed in chapter 3, the timestamp rendered in one of the temporal levels can influence the likelihood of determining the video result as a potential match or not. Therefore, we evaluate the acceptable temporal difference between the retrieved video and the expected in this chapter. In turn, the chapter also draws on the inclusion of other video-related objects, along with the video result playing from an expected timestamp, to conveniently establish the match. Drawing inferences on these points are imperative in determining the optimization parameters for the system's ability to predict video matches. Therefore, the objective of this chapter is three-fold:

1. To learn the *temporal specificity*, the preferred temporal level among point, interval, and whole video levels where the temporal difference between retrieved to expected is acceptable.

2. To examine the preferences for *video surrogates* along with the video when rendered on search.

3. To evaluate the *usability* of the proposed approach, that is the idea of an automated linking interface over the existing linking approaches.

4.1 Experiment Design and Procedure

We conducted a user study with 12 Engineering students to demonstrate the preference for temporal specificity in point, interval, and whole video, along with the expectations for video surrogates in search results. With approval from the University of British Columbia (UBC) Behavioural Research Ethics Board, we obtained informed consent was obtained from all the participants. A pilot study with 2 participants confirmed any setbacks in the procedure of the study. The study began with only 4 participants to determine the total sample size. An a-priori power analysis was performed with the data received from the 4 participants using the G*Power application [26]. The chosen test family was the t-test, and the chosen statistical test was the 'Means: Difference between two dependent means (matched pairs)' test. We calculated the effect size based on the Likert-scale data for a question comparing the information effectiveness among the three temporal levels 'The information is effective in helping me complete the tasks and scenarios for learning'. The point level condition was the control group compared with the interval and whole video levels as the treatment groups. The calculated effect size was $dz = 1.03$ for point-interval group and $\alpha = 1.59$ for the point-whole video using the difference of means and standard deviations. The computed sample size was 6, 12 choosing an input error probability of 0:05 and desired power of 0:95. We finalized the sample size as 12 and recruited students accordingly.

4.1.1 Participants

With a sample comprising 12 university students, the controlled lab-based study included students from a range of engineering disciplines and degree levels. Table 4.1 outlines the students' basic information. We reached undergraduate and graduate engineering students with convenience sampling. Inclusion criteria were that students were over eighteen years of age and had experience learning with video. We scheduled half-hour appointments with each participant. The entire session was audio-recorded, which enabled us to focus on the verbal prompts. A verbatim transcript of the interview was later generated.

Table 4.1: Demographic data of the 12 students

Participants	**Gender**	**Level**
P1	Female	MASC
P2	Female	PhD
P3	Female	MASC
P4	Male	PhD
P5	Male	MASC
P6	Male	MASC
P7	Male	MASC
P8	Male	MASC
P9	Male	2nd year, Undergraduate
P10	Male	2nd year, Undergraduate
P11	Female	2nd year, Undergraduate
P12	Female	4th year, Undergraduate

4.1.2 Study Procedure

The conducted study was a within-subject study in which each subject performed the same tasks in all the conditions. Before beginning with the key tasks of the study, each participant went through an introductory session where they shared their existing workarounds when learning from the videos. The session was followed by a mock video-learning task to help students familiarize themselves with the video-learning settings and also generate sufficient real-time notes to guide the subsequent tasks. The remainder of the procedure was designed to address the key objectives of this chapter individually.

The study began with a few experience-based questions to learn about students' current video learning, reviewing, and note-taking habits. The questions followed a semi-structured interview pattern to allow students to express their reviews related to learning from videos. We performed a 2-student pilot testing before the experiment to formulate the questions that sensitively addressed video-based learning reviews. Although the session included exploratory questions, key questions asked were as follows:

1. How often do you learn from videos? What kinds?
2. What measures do you take to help you remember the things that you have learned for reviewing later?
3. How do these measures help you re-find the source efficiently?

As a next step, we gave participants a realistic task scenario: to learn from an instructional video and complete a comprehension quiz to improve ecological validity, in line with Borlund's work on simulated work tasks in lab studies [6]. We modeled the task to generate real-time notes whose purpose was to deduce inferences on the experience and traits involved in the process of video-based learning. Additionally, participants used the notes to reflect on the various conditions in the later part of the study.

Initially, we selected one 5-minute video to learn. The pilot study showed that students have varied interests and might get frustrated to learn unrelated content. Hence, we selected 3 videos of different genres to accommodate the curriculum-based choice of preferred video. Three instructional videos that spanned approximately 5 minutes each were chosen beforehand. Each student was free to choose any video from the three. The topic of the videos were: 1) Introduction to capacitors, 2) Nervous system, and 3) Reinforcement learning. The motive of the study was not revealed to the students at this point to avoid leading participants to perform a certain way. We told them that it is necessary to remember the video content after the viewing session and connect back to the video during the study. Participants were given blank sheets of paper to take notes while they watched the video. Students were asked to identify parts of notes that pointed to video-related content after completing the video learning session. Their general learning and linking habits were discussed, which later followed to a discussion on the conditions that supported the objective of the study. After gleaning on various existing linking strategies carried out, we explained the intent of our proposed linking approach to the students.

4.1.2.1 Temporal Specificity

The first objective of this study was to investigate the acceptable temporal difference in the video rendered to that of expected. We considered three conditions of delivering a timepoint in point level, interval level, and whole video level. The conditions were demonstrated relative to the participant's notes captured in the previous step and the related video. For instance, participants noted a video timestamp that connects to part of their notes as *mm:ss* in video time. In the case of point condition, the timestamp to be retrieved was demonstrated in two to three timepoints, say *mm:ss+0:02* or *mm:ss-0:02* in the same frame that pointed to the noted timestamp. If a frame spanned for about 30 seconds and the noted timestamp is inside the frame, the retrieved timestamp is either the start, middle, or end of the 30 seconds window. Thus, the temporal difference is now in this window condition. We repeated the procedure for the other two temporal conditions, that is, various timepoints in the corresponding section and the whole video, the order counterbalanced. Students discussed the pros and cons of all three conditions in detail to observe the acceptable difference. Additionally, they provided the preference for each condition based on the following after-scenario questionnaire on a Likert scale of 1 (strongly disagree) to 5 (strongly agree):

1. It is easy to find the information I need from this point.
2. The information is effective in helping me complete the tasks and scenarios for learning.
3. I'm satisfied with the retrieved video time point.

4.1.2.2 Re-scan, list of search results

Submitting a meaningful query is essential for video search effectiveness from a large video archive [91]. Taking the example of the Google search engine, a user often reformulates or rephrases the topical content of queries to retrieve the desired information. Similarly, students expressed their reviews on performing re-search operations with other relevant video-related queries from their notes. The discussion also covered traditional ranked-list presentation of more than one to two video results in the mobile scenario.

Along with the video and the timepoint at the acceptable temporal level, the video surrogates' rendering is significant in the video retrieval process. The next step was to glean on the video surrogate's expectations when rendering the predicted video.

4.1.2.3 Preference for Video Surrogates

Popular search engines such as Google provide a set of keywords along with the heading, URL in search results to highlight query relevant pages. This page-related information presented are surrogates for a web platform that assists searchers in deciding whether to examine the full web page. Similarly, surrogates for video are a condensed representation of the video as a whole while summarizing key features to cue a quick sense of the expected video. They facilitate rapid sense-making. Textual and non-textual video surrogates were used in Marchionini's work [57] to aid rapid gist determinations from the list of results to select which video to watch. While there are numerous types, four fundamental and frequently exercised video surrogates were selected for the study: 1) Title of the video, 2) Thumbnail, a small image that exemplifies video content, 3) Keywords to indicate significant video substance and, 4) A summary of the video content. We demonstrated three sample designs with the four surrogates to the students:

1. Video with title and thumbnail

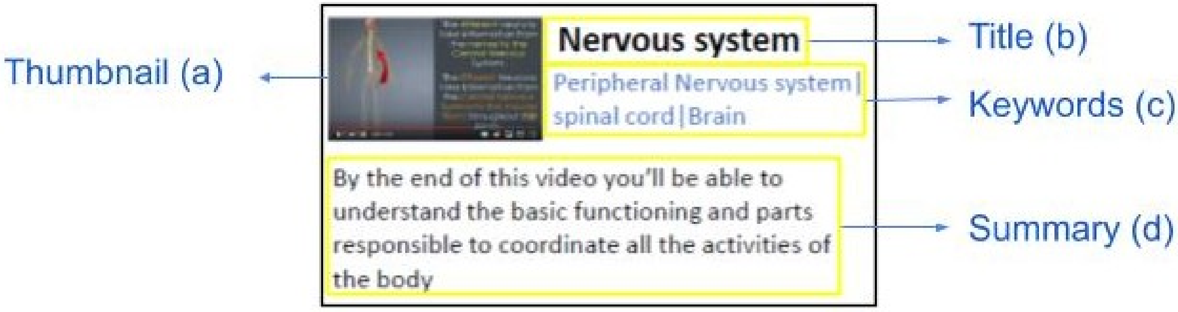

Figure 4.1: An example video surrogate, comprised of a thumbnail (a), title (b), keywords (c), and summary (d).

2. Video with title, thumbnail, and keywords

3. Video with title, thumbnail, keywords, and summary.

Figure 4.1 shows the design sample for the third condition. Students' preferences for each sample were discussed and rated based on a set of Likert-scale questions with a scale of 1 (strongly disagree) to 5 (strongly agree) in three scenarios:

1. It is easy to find the information I need from this point.

2. This design has all the functions and capabilities I expect it to have.

3. The organization of information on the layout screen is clear.

4.1.2.4 Usability

Each participant talked about the usability of the proposed system as a working application, recorded in terms of three key elements; efficiency, effectiveness, and satisfaction [9]. The ratings for the three elements were captured through the following set of Likert-scale questions (1; strongly disagree to 5; strongly agree).

1. Effectiveness is indicated by the quality of solution, accuracy, and completeness with which users complete specific tasks. Following questions were used to report on the effectiveness of our approach:

   (a) I think the system is easy to use.

   (b) I found the various functions in this system were well integrated.

   (c) I think that I would need the support of a technical person to be able to use this system.

2. Efficiency is shown by task completion time or learning time which was captured using the following questions in our study:

   (a) I believe I could become productive quickly using this system.

   (b) I would imagine that most people would learn to use this system very quickly.

   (c) I needed to learn a lot of things before I could get going with this system.

3. Satisfaction is confirmed by users' comfort and positive attitudes towards using the system. The study reported on users' satisfaction rates based on the below questions:

   (a) I think that I would like to use this system frequently.

   (b) I found the system unnecessarily complex.

   (c) Overall, I am satisfied with this system.

The items for each usability element were phrased with two positive and one negative question to avoid extreme response bias and acquiescent bias. Interview data was transcribed and de-identified before analysis.

4.2 Study Findings

In this section, we refer to the findings from the study discussed in the previous section, particularly with respect to preferences in terms of video retrieval and attitudes towards using the proposed system.

4.2.1 Video in "Interval" Temporal Level

We compared the difference in preference for the temporal levels of video timepoints (i.e., retrieved video timepoint in point, interval, or full video) based on the data from the three Likert-scale questions. Since the data collected was ordinal, it failed to meet the assumptions for parametric tests. Thus, we performed the Friedman test, a parametric alternative, that confirms whether there are overall differences between the groups of consideration. But it does not pinpoint which groups, in particular, differ from each other. To achieve this, one needs to run a post-hoc analysis. We conducted Wilcoxon signed-rank test with a Bonferroni correction applied, resulting in a significance level set at $\alpha < 0.017$. Here we used an exact p-value considering the small sample size ($n = 12$).

We analyzed the three questions, respectively.

1. First, participants' perceived ease in finding information from the retrieved time point was statistically significantly different among the three timepoint groups ($\chi^2(2) = 7.4, p = 0.021$). However, the post-hoc test did not locate significant differences between the combinations of the three groups despite an overall reduction in the whole video group compared to the point and interval groups.

2. Participants' perceived effectiveness in completing tasks for learning was statistically and significantly different among the groups ($\chi^2(2) = 6.897, p = 0.029$). But no statistical difference was observed in the combinations.

3. Lastly, participants' satisfaction with the retrieved timepoint was significantly different in the groups ($\chi^2(2) = 9.190, p = 0.008$). Additionally, the post-hoc test showed significant difference in the video versus interval groups ($Z = -2.240; p = 0.011$). Participants felt higher satisfaction when the retrieved video timepoint was in the interval when compared to anywhere in the whole video. The Ranks table provided some interesting data on the comparison of participants' satisfaction for the two groups. Ten participants had a higher satisfaction score for video playing in the interval of the expected content occurrence over a timepoint anywhere in the whole video. However, only 1 participant had a higher score for the whole video condition, and one participant saw no change in terms of satisfaction.

To learn more about the preference for the video timepoint retrieval in one of the three temporal groups, interview transcripts were examined for explanatory quotes. The quotes from the 12 participants were coded with five different preference levels, *very good, good, neutral, bad, very bad* as seen in figure 4.2. For example, P13 had the following reviews for each category:

- Whole video; *That can be **really inconvenient** because some videos are like two hours* → coded to **very bad**
- Interval; *it is somewhat helpful I'll probably be **pretty satisfied*** → coded to **good**
- Point; *Well, then I'll just be **pretty happy**. Yeah. Gets me to where I want to go* → coded to **very good**.

Quotes were coded as **neutral** if preference changed with contexts. For example, P1 reported that the length of the video influences preference when timepoint retrieved is anywhere in the video; '*If it's a 15-minute video. Yeah, I would be OK with it because it's just 15 minutes you would want to devote like a solid understanding and if it's a day*

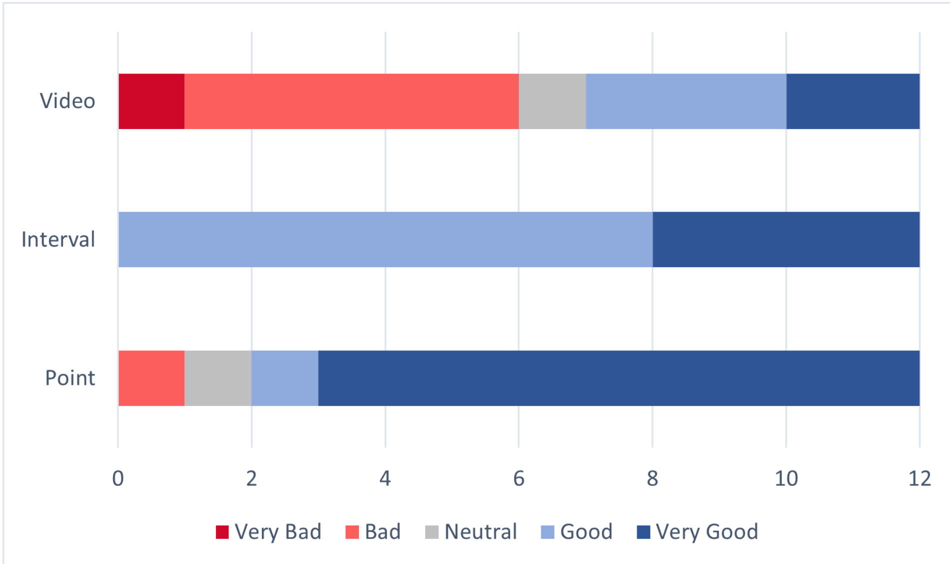

Figure 4.2: Coded mapping of preference levels (i.e., very good, good, neutral, bad, and very bad) to each level of temporal specificity (i.e., point, interval, and whole video).

before an exam. Yeah, but if it's an hour or two hours then it would become slight bit of an inconvenience to find it.'

Following the coded mapping of preference to each temporal group across the 12 participants, the point-level temporal group presented 9 - very good, 1 - good, and 1 - negative code tagged to bad. The interval-level temporal group displayed no negative codes, 8 - good and 4 - very good codes. Six participants exhibited negative codes, bad, in the case of the video-level temporal group with five participants still indicating positive codes, 2 - very good and 3 - good.

4.2.2 Selection-based Use of Video Surrogates

To examine the video metadata cues to be rendered along with the video, data from Likert-scale questions was analyzed to compare the three conditions: 1) Title + Thumbnail, 2) Title + Thumbnail + Keywords, and 3) Title + Thumbnail + Keywords + Summary. Once again, we performed the Freidman test along with the Wilcoxon signed-rank test to investigate the difference between the three conditions.

Participants' perceived ease in finding information from the retrieved timepoint showed no statistically significant difference between the conditions ($\chi^2(2) = 2.889, p = 0.240$). Participants' expectations of the design (i.e., design has all the expected functions) also showed no statistically significant difference between the conditions ($\chi^2(2) = 2.8, p = 0.272$). Participants' perceived information clarity was statistically significantly different among the groups ($\chi^2(2) = 8.914, p = 0.008$). Post hoc test indicated a significant difference in the clarity of information organization among the summary and keyword surrogates ($Z = -2.762; p = 0.004$). Ranks showed that 9 participants had a higher clarity score for keywords than the summary. No participant preferred summary over keywords. However, three participants saw no change in the organization with the two groups.

Since there was no significant difference recorded in 2 out of the 3 likert-scale questions, we could not deduce preference for each of the selected surrogates. Although there was the influence of the video surrogate combination in the organization of information, it was not clear about the ease of finding information. Interview excerpts further supported this evidence. Participant P3 preferred keywords; '*Yeah. I would prefer the keywords most because that one is short. For the abstract one, if it's one or two sentences, I think that will be better.*' Participant P10 pointed to the need for summary; '*I think more the information more it will be easy for me. So ok from here it is like more information so I can just read through and recall if that's the video I'm looking for*'. Therefore, Select or Deselect action items can be included in an expandable list of video surrogate choices. When a student first views the video result on search, the default value, title + thumbnail, is the choice that appears in the video control and selects the other surrogate options that they want.

The inclusion of transcripts also seemed to be influenced by the author's accent, video content type, compact video viewing in mobile applications. An additional observation was that all the participants except one (P3) preferred seeing a thumbnail that showed the video content that matches the note content. For example, P9 informed that '*if that slide would be in the thumbnail, Yeah that would be like the most optimal scenario*'. However, P3 reported that it might not be as useful: '*I think it's because, with the typical video thumbnail, I think for a video they always have a thumbnail selected for that video in the system and I remember that thumbnail. And if you change the thumbnail, during the search, I'll probably get confused.*'

4.2.3 Usability reports

We analyzed the final survey data on usability. The figures 4.3, 4.4, and 4.5 below show the stacked bar charts constructed from Likert scale data for 3 different usability elements: Effectiveness, Efficiency and Satisfaction. The counts of participants on each row who agree with the statement are shown to the right in dark blue color; the counts who disagree are shown to the left in dark red color. Participants who neither agree nor disagree are shown in a neutral color.

Participants thought that NoteLink was easy to use ($M = 4.42, SD = 0.49$) and were able to become productive quickly using NoteLink ($M = 4.33, SD = 0.62$). Overall, participants were satisfied with our NoteLink ($M = 4.33, SD = 0.47$). However, the reports did not indicate complete agreement in terms of the system's functioning. 2 participants marked 'neutral' when asked about the system's integration of functions. 2 Participants also agreed that there is lot to learn. All the 12 participants asserted that the system would be very useful while learning from videos if it were a real application and if it could scope well to their expectations. P1 said, '*If it had things I want, yes. Very useful. Saves so much time, and maybe I'll watch more videos then.*' P2 agreed: '*Definitely, for videos that have to be bookmarked. This is very useful.*'

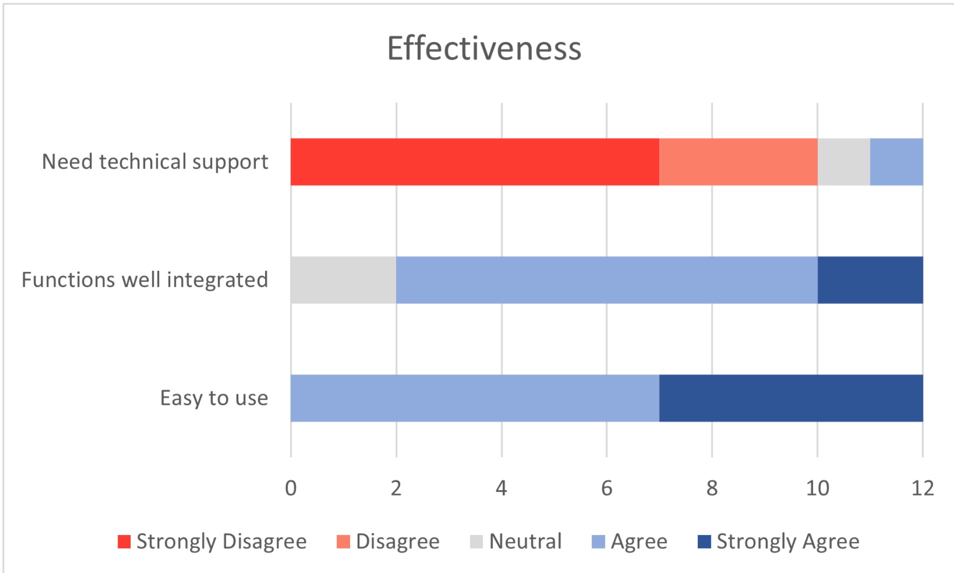

Figure 4.3: Stacked bar chart showing likert-scale observations for Effectiveness.

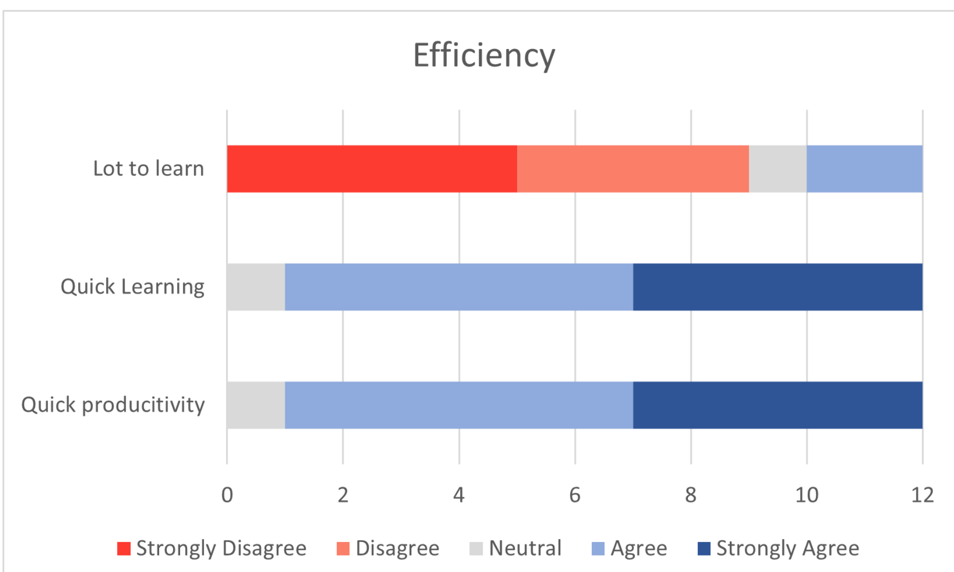

Figure 4.4: Stacked bar chart showing likert-scale observations for Efficiency.

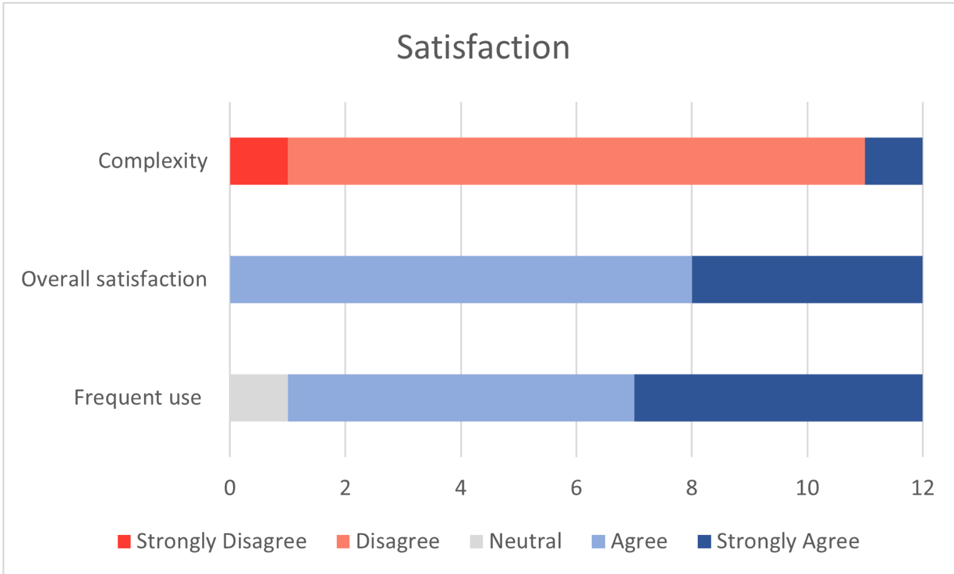

Figure 4.5: Stacked bar graph showing likert-scale observations for Satisfaction.

Lastly, qualitative analysis of the interview transcripts regarding the traditional list of search results showed that 11 out of 12 participants preferred having a list of probable video matches (in particular, three videos on average). For instance, P1 preferred having at most three options; '*Options are definitely good. But how many options are there? I don't want something like google where it has pages of matches. Maybe three, not more than that*'. Therefore, we decided to go with three options with an option to view more when required.

4.3 Discussion

The video timestamp to be retrieved was conceptualized and evaluated on a temporal scale to support semantic evaluation of the system's retrieval prediction. We established an overall difference in the preference for the predicted video timepoint in the three conditions; point/interval/Whole-video. The interval group, in particular, showed a significant increase in satisfaction compared to anywhere in the whole video. The reason is likely contextual, involving several aspects of the viewer's information need at the time that might influence the preferred temporal level. For instance, P9 had his preference change with the length of the video and complexity of the topic: '*If it's a short video, I'd be okay with just anywhere in the video. If it's a long video, I wouldn't be okay with just the video. If it's a very complex topic, I would definitely not prefer being taken to the point. But, if it's a rather simple topic, of course, I'd be okay with just, you know, like being taken to the point. …The sweet spot would be to use the middle one — interval.*' The coded preference mapping in figure 4.2 also showed that interval level match point is the most opted with 0 negative codes. The point-based temporal specificity also showed the most positive codes (10 out of 12). However, 6 out of 12 participants shared negative codes for the whole video level, indicating lower preference compared to the other two levels.

Interview transcripts revealed students' preference to view a list of probable video matches seen in the traditional queries to digital libraries. Christel [14] discussed two major challenges in locating needed information from the list of results: 1) Information returned can be too much, and 2) Information needs are different for different users. Similar trends were observed in the preference for video surrogates: title, thumbnail, keywords, and summary. Therefore, the enable/disable feature-based selection of video surrogates from the ranked list of video matches is suggested in this work. Furthermore, emulating design recommendations from literature [10], [82] on effective mobile information retrieval can boost the performance of our system in the human-computer interactive perspective.

Analysis of interview excerpts along with the notes captured during the study led to interesting themes towards linking and note-taking habits.

4.3.1 Link to Note vs Note of Links – Linking Strategies

We analyzed the current conventional workarounds that students perform to link their notes to videos in the interview excerpts. The findings led to a list of linking strategies that can compile identical methods employed in digital and paper-based physical mediums. This is a unique observation over the traditional categorization of digital versus physical linking techniques. We classified the collection of linking methods into two main groups whose elements are listed in the table 4.2:

1. One category comprises of links or video cues such as title, timestamp, content headings, and many more adjacent to the relevant part of the notes, that is termed as *link to Notes.*

2. Another category includes creating a document of links such as word or excel document of pointers to videos that is named as *Note of Links.*

Each participant expressed their views of an existing approach versus the proposed approach. P7 reported that '*I never thought it would be easier or difficult. If I had another choice like another process that makes things easier then definitely I would have felt that this would be difficult.*' The participant also found the system's approach to be easy. '*I would definitely like it. Instead of searching myself every time I go to YouTube and saving a link somewhere and forgetting things. Yeah, scanning and it gets to the right video, Yeah. That's really good*'. Thus, the proposed system hints at a great potential in supporting automatic linking to videos over the existing manual inclusion of links.

4.3.2 Impact on the Existing Note-taking Practices

Technologies that create new environments for note-taking have shown to bring change in the way students take notes. For instance, a study on the annotation of digital documents

Table 4.2: Identified current linking techniques

	Link to Notes	Note of Links
1	Title, interval/content headings, in-video text, author name	Excel sheet of video to note pointers
2	Timestamp, video link on digital notes	Word document of slide screenshots
3	Highlight, color code, 'x' identification marks to notes	Video Bookmarks

[59] demonstrated significant changes in the note content and anchoring style. On the same lines, students discussed the possible changes/influence on their existing note-taking habits with the introduction of the proposed linking system. Interview quotes indicated the proposed approach might not change students' note-taking habits; '*My note taking wouldn't change but would be far more helpful*' said P13. The possibility of extending the application of our system on digital notes also played a role in indicating no influence; '*...because I write electronic notes I can imagine that could also be something that I can use this snipping tool to take pictures and do that.*' Overall the ability of the system to retrieve video without adding links or pointers at the time of capturing notes assures preserving students' current note-taking style. '*I don't think my note taking process would change as much. No, I don't. I would still write my notes like this. Because it works even now, if I scan this word*' said participant P12.

4.4 Summary

This chapter conducted an experimental study to draw on the students' expectations of mechanizing their existing workarounds in linking notes to videos to make informed and inspired design solutions. The qualitative analysis of interview transcripts bolstered the evidence demonstrated by the qualitative data from Likert-scale questions. Results indicated that students prefer video rendered with a timepoint in the interval level, while point level matches also showed higher satisfaction. Rendering video with a timepoint anywhere in the related interval/section provides a considerably large window of time difference between the expected and the retrieved video timepoint. This is an important finding as it aids future video retrieval systems to conceptualize the timepoint difference in determining the right video context.

The results also showed that no participants had higher priority for summary than keywords, although they showed no significant priority for thumbnail. A deeper investigation into the effects of video surrogates on video retrieval systems is necessary for this work in the future. However, students indicated that the search results could go up to 3 possible matches. Ultimately, the usability evaluation revealed that students would find the proposed linking approach helpful when learning from instructional videos. The students' positive attitude towards the proposed method is paramount to advance into the design and development of watched video retrieval systems.

Overall, results established the requirements of the video rendering in terms of temporal specificity and the types of surrogates. The next chapter moves discusses the pragmatics of linking handwritten notes to corresponding videos with the help of a medium-fidelity prototype design. We can perform an informed flow of recognition analysis with the established content types used in watched video notes. Also, the preferences for temporal levels can enable the prototype design to find the right video matches. The prototype development and testing outline the technical requirements in the next chapter, which, with the findings in this chapter, comprehensively answers the second research objective of this work.

CHAPTER 5

Pragmatics of Linking Notes to Video

From the previous chapters, three conclusions are apparent: 1) students make links to video content in their notebooks using a variety of content representations (text, formulas, and figures), 2) in all the identified content types of watched video notes, there is some amount of verbatim overlap, and 3) expected timestamp in a video can be scaled to three temporal levels in a video; timestamp in a point, interval, and whole video. A medium-fidelity prototype was devised on top of the above hit inferences to evaluate the effectiveness of exercising watched video notes to link to corresponding videos that:

1. Recognizes the three handwritten content types,
2. Matches these to a collection of videos and,
3. Presents three best matching videos with timepoint in one of the three temporal levels.

5.1 Note Recognition Analysis

To limit the implementation efforts and focus on the implementation gaps associated with the proposed approach, we explored readily available off-the-shelf Optical Character Recognition (OCR) technologies to recognize handwritten images. The process of selecting the best OCR model directed us to employ different recognition models for different handwritten content types. The goal here was not to compare or evaluate the recognition accuracy or capacity of various models.

OCR Application Programming Interfaces (API) from PixLab[1], Google cloud vision[2], and Microsoft's computer vision read handwritten text API[3] were used to compare and identify the optimal recognition API. A random number of text-based samples from the collected video notes in phase I were sampled and then compared for the word matching count. The words matched did not include stop-words such as 'the', 'at', 'is', and so on. With nearly similar results, Microsoft's API worked slightly better with rotated characters and produced relatively fewer false positives when compared to the other APIs. However, it did not reliably recognize the special characters and symbols often used in formulas. It also could not interpret lines in figures that did not contain any meaningful text. Thus, we used Microsoft API to read text-only notes, and explored other OCR APIs for non-textual components.

Our choice to extract special symbols like scientific notations from a given image, without stroke data, was Mathpix OCR API[4]. The API's performance was evaluated with a few randomly picked formulas from the pilot study dataset and recorded the number of matched characters from the original equation. The result object returned with a confidence field which demonstrated a probability of more than 70% for all the tested cases. Hence, Mathpix was used for reading mathematical content, that is, formulas.

A two-step recognition process facilitated reading lines in pictorial representations. First, we used Scale Invariant Feature Transform (SIFT)[5] algorithm from Open Source Computer Vision Library (OpenCV) to extract feature-based keypoints and descriptors from a scaled and slightly rotated image. Second, we studied structural information that predicts the perceived quality of an image as Structural Similarity (SSIM) index[89].

Next, the Matcher block incorporates finding the corresponding video from the col-

[1] https://pixlab.io/api
[2] https://cloud.google.com/vision/docs/reference/rest/
[3] https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/concept-recognizing-text
[4] https://mathpix.com/ocr
[5] https://docs.opencv.org/3.4/da/df5/tutorial_py_sift_intro.html

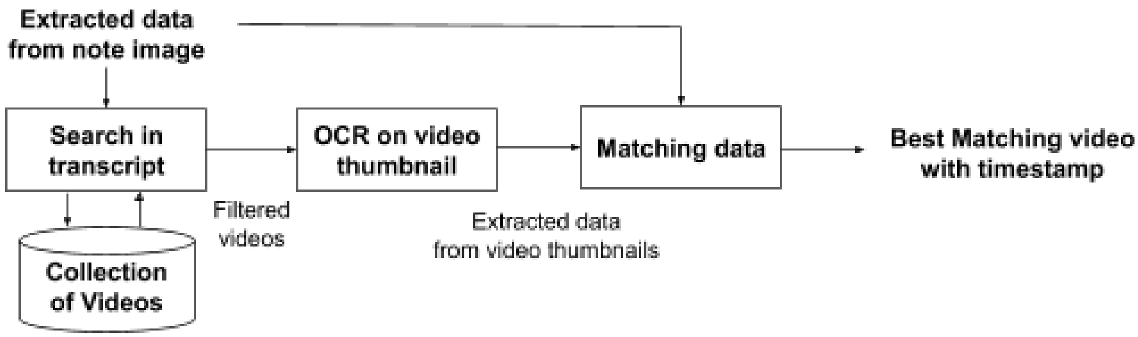

Figure 5.1: Matching flow for Text and Formula content.

lection of videos.

5.2 Finding Video Matches

Visual information from slides and the transcribed auditory text was compared with the recognized handwritten content. Figure 5.1 illustrates the approach used to pick the matching video for the handwritten units' text and formula. Extracted data is the output from OCRing text and formula handwritten input. Stop-words were filtered out before feeding into the matching block.

A quick search of data occurrence in the transcripts of videos provided a list of possible matches. We carefully examined the video slides of these matches to find the best matching video timepoint. Finding similar textual components was straightforward as the entire slide can be OCRed to extract all the text-only data. However, instructional slides are often content-heavy and include various types of information like equations, figures, and more content types in a slide. The non-textual objects, such as, formulas have to be identified and highlighted before reading them. Region of Interest (ROI) facilitates filtering or operating on unique portions of an image. Using OpenCV, ROIs for each video slide equipped the highlighting and extracting borders of mathematical content.

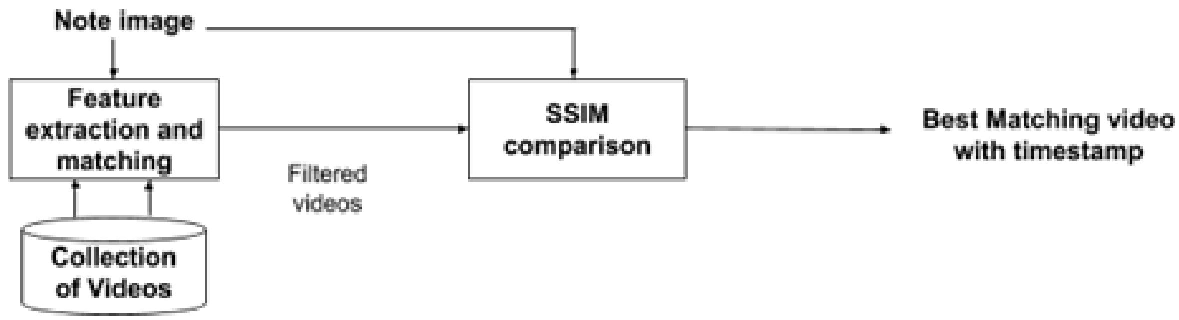

Figure 5.2: Matching pictorial content for figures

The next step directly applied the chosen OCR techniques on the slides for text type and the extracted ROIs for the formula type. The resulting data from this function was compared with the input handwritten data to arrive at the best matching ROI, which gives the best matching video with the timestamp where the data occurs.

The process flow for pictorial data is different from the text and formula as there is no quantifiable information that can easily compare in this case. A flowchart can be seen in the 5.2. Here, again, ROIs were extracted from each video slide. Feature descriptors were detected using SIFT from ROIs and matched with input image descriptors. For feature matching OpenCV's Fast Library for Approximate Nearest Neighbors (FLANN)[6] algorithm was used. Here, to eliminate possible false matches, a ratio test was performed that selects only the most promising matches for further screening. Following this approach through the collection of videos, a list of possible matches was captured. Later, the SSIM of the note image was compared with the list of ROIs from the filtered videos to get the final matching video with a timestamp where the pictorial information exists.

In each case, the matcher block picked 3 best possible video matches, in line with the inferences drawn from the students in the previous chapter. In the subsequent section, a mobile interface application that provides a platform to submit the watched video notes and play the retrieved videos is discussed.

[6]https://docs.opencv.org/3.4/dc/dc3/tutorial_py_matcher.html

5.3 Smartphone as a Pedagogical Interface

Enabling an interaction platform to feed the note content and consume related videos is an important step in designing the prototype. The goal here is to employ an interface that expects minimum effort from a user to provide the desired input and delivers video results that convey the information necessary to choose a match. Mobile devices allocate new pedagogical affordances to students in varied temporal conditions and learning settings [16]. They have proven to possess great potential in engaging students with effortless access to needed content and delivering information 'just-in-time' [95], [32], [42], [48]. The annual survey by the Educause Center for Applied Research [ECAR] on Mobile IT in higher education has tracked mobile technology usage since 2012 [31], [17]. The reports demonstrate a positive correlation among student use, the importance placed on these technologies, and students' academic success. Additionally, considering students' robust use of paper notebooks for their coursework, it is better to design a mechanism in the mobile world that adapts to the attributes of paper. The attempt here is to aid the ability to capture heterogeneous handwritten information in student notebooks and quickly retrieve required video sources in their own devices.

To this end, we leveraged the ability to use a camera to take pictures and video viewing in smartphones to design a mobile interface in this work. This interface employs a mobile camera to point to video-related notes and re-find corresponding video in a mobile-based video platform. Figure 5.3 illustrates the use of the interface in an example case, where a student points to a textual description in notes to retrieve the corresponding video. First, a student opens the application and is presented with the two options to feed the formula, take a picture, or upload an image(top). For either option, the user can crop/adjust the image containing the equation to delimit the note content that can be linked to the video. The corresponding video, that is, the video watched while noting down the content, is presented to the user, played back within the application (down).

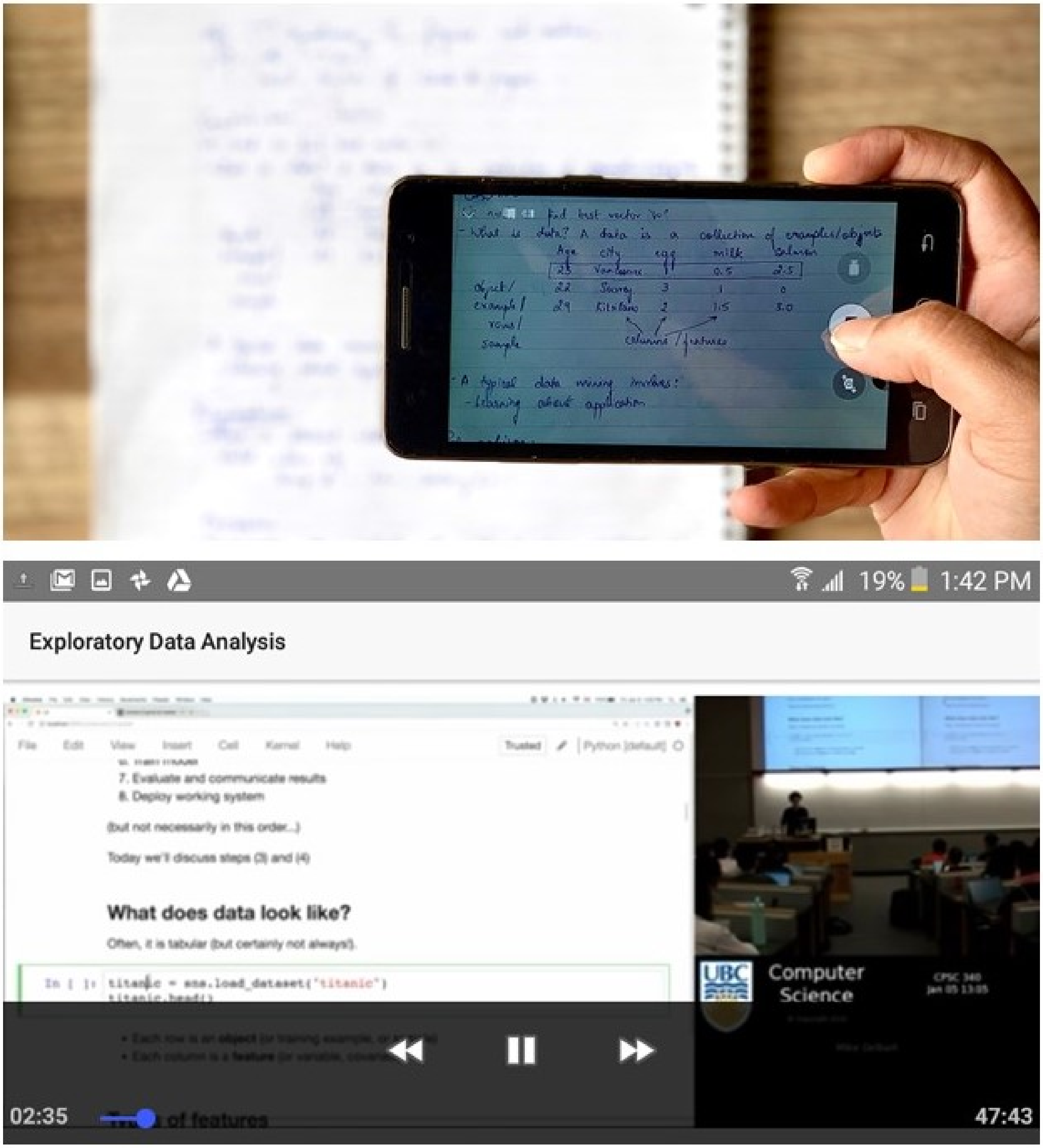

Figure 5.3: Mobile interface, is presented with the option to take a picture of students' handwritten note content (top) to find the corresponding video with topically similar content, read to be played back within the app (bottom)

The mobile application was built using Nativescript plugin API[7] for using device camera[8], background-http plugin[9] for enabling http methods and video player plugin[10] that uses the native video players to play remote. At the back end, the image input is fed to Recognizer and Matcher blocks to find the corresponding video from a collection of videos.

5.4 Accuracy of matching videos to notes

We assessed the ability of the prototype system to retrieve watched videos on the notebook data sample identified in chapter 3. Additionally, instructional videos that students learned to make notes were also collected to compare the video result with the ground truth video marked with the note data. We transcribed the spoken content of the videos to facilitate the matching between the note content and the video. The prototype tested a total of 181 watched video notes. The evaluation metric used was accuracy which is a fraction of the number of matched note samples to the total number of note samples:

$$\%\ of\ Accuracy = \frac{\#\ of\ matched\ notes}{total\ \#\ of\ notes} * 100 \tag{5.1}$$

Following guidelines were used to choose a video result as matched or not-matched:

1. As the search result gives only three different videos relevant to the note content, we treated any video result of the three as a match irrespective of their ranks.

2. Timestamp retrieved was compared with the observed timestamp collected as ground truth. In both cases, the video content was checked and marked if they belonged to

[7]https://www.nsplugins.com/

[8]https://market.nativescript.org/plugins/nativescript-camera/

[9]https://market.nativescript.org/plugins/nativescript-background-http/

[10]https://market.nativescript.org/plugins/nativescript-videoplayer/

the same point, the same interval where the video content marked as ground truth or as the whole video if they were from different intervals.

3. If any of the three matches did not match the expected video, it was marked as *No-Match* which is 0 relevance.

4. If the recognizer block threw an error or returned no information, it was marked as *No-Data*.

After analyzing all the video notes, 38 of the 181 notes were regarded as no-data and excluded from the set. Of the remaining 143 notes, one was marked as a special case as it did not include ground truth, that is, timestamp and the video. Therefore, the data from the Matcher block that was available for comparison was 142. About 77.5% of the total notes (n = 142) returned video(s) that matched the expected videos that students reported as ground truth. The pie chart in figure 5.4 illustrates the accuracy breakdown across all the note samples. 29.58% of the notes samples were matched to the exact point in video as that of the expected. 27.46% of the samples returned match results playing in the same interval or section where the students' note content is present. 20.42% of the samples produced the right videos but did not play in the same point or interval as was expected.

Figure 5.5 demonstrates the accuracy percentage for each of the types breaking down the accuracy of finding the video matches across the three different content types, that is, text, formula, and figure. The system matched textual content with an accuracy of 97.5%, formulas with 56.8%, and drawings with 38.9%. Table 5.1 shows the exact number of matches and accuracy predicted for the three content types on the temporal scale of point, interval, and video.

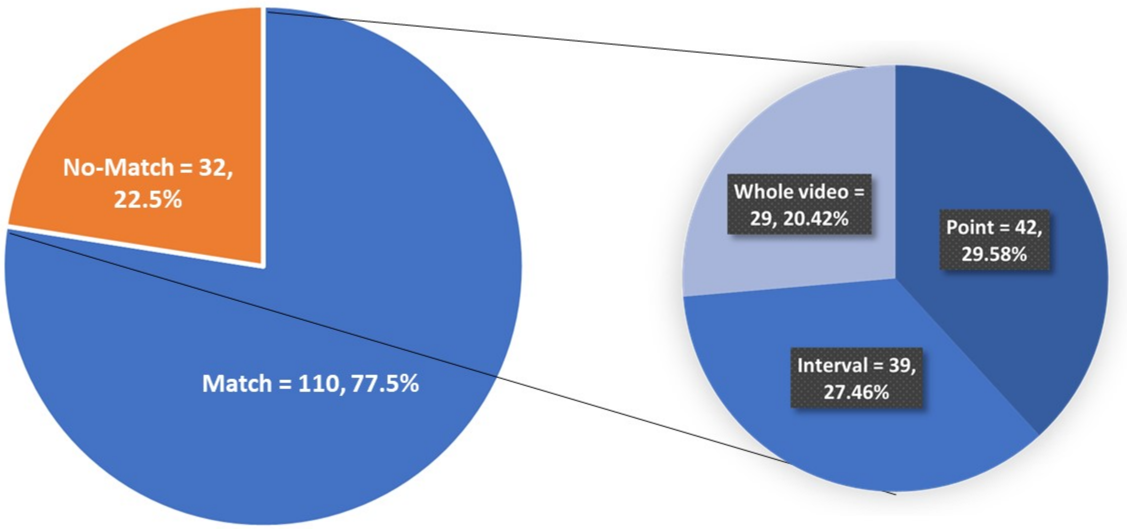

Figure 5.4: The pie charts showing the total accuracy of finding the matched and the breakdown across point, interval and whole video temporal levels

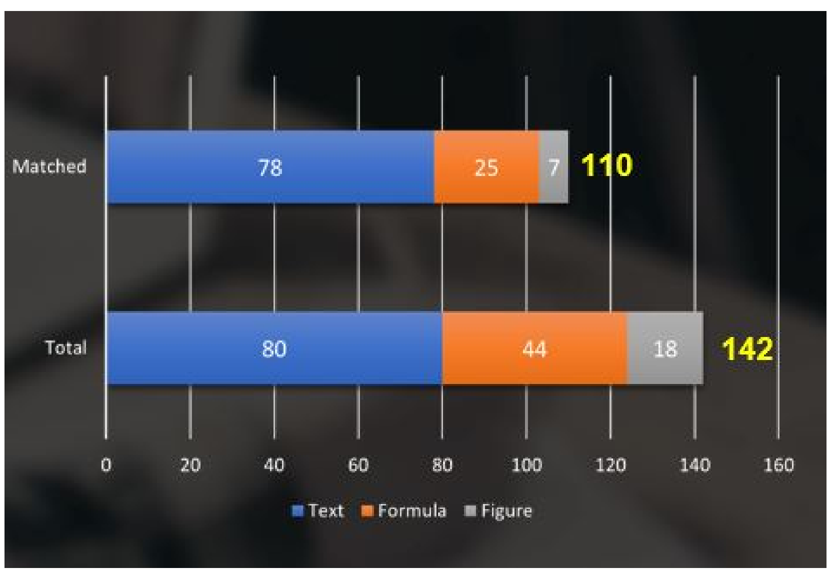

Figure 5.5: The bar chart showing the total number of matched note samples of the 3 content types; Text, Formula, and figure

Table 5.1: The number of count (n) of note-video matches, and accuracy of retrieving matching video is reported for each content type (i.e., text, formula, and drawing) overall and at the three levels of temporal specificity (i.e., point, interval, and whole video).

	Temporal Level					
Note type	Point	Interval	Whole video	Overall		
	Count (n)	Count (n)	Count (n)	Count (n)	Notes (N)	Accuracy($n/N\%$)
Text	36	27	15	78	80	97.5
Formula	5	10	10	25	44	56.8
Drawing	1	2	4	7	18	38.9
Total	42	39	29	110	142	77.5

5.5 Content Overlap

We computed the percentage of content overlap between the note content and the video slide deduced from the results of our system. Doing this will give us more insights into how the system performs concerning student's manual process of linking their previously written notes to corresponding video context. The proposed approach matched most of the textual components in student notes related to watched videos with a 97.5% accuracy. Thus, we first considered the general characteristics of the overlap results of the matched video for the text-only content and made the following observations from the boxplot shown in figure 5.6.

1. Atleast 50% of the video results that played from the exact video slide/point matched the note content word to word with more than 90% overlap. Minimum overlap is at 25%, and the IQR spans between 65-100%, indicating that video results playing from the same point as expected had an overlap greater than 65% for most of the note samples.

2. IQR of the matched content in the case of interval condition also manifests a considerable amount of overlap, that is 55-80%, with at least 75% of the note samples

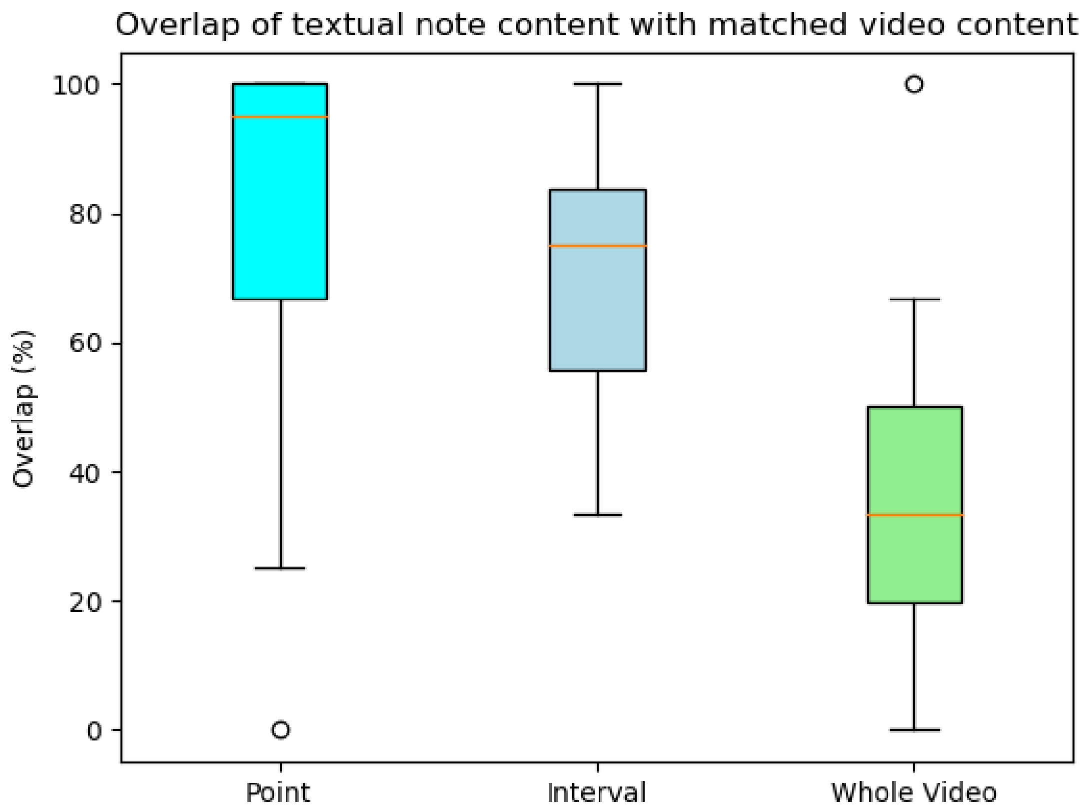

Figure 5.6: Results showing overlap between video results and text-only content.

matched in this condition showing an overlap greater than 50%. This hints that the note content in the case of textual components summarize an interval of a video and can point to multiple frames in a video interval

3. In the case of the whole video condition, the matched results still show some overlap. However, it offers a decreased percentage of overlap. An interesting factor to note is, more than 75% of the notes showed overlap between 20-40% indicating that the note content might re-appear in various parts of the video.

Similarly, figure 5.7 also demonstrates the percentage of overlap for the formula-based content.

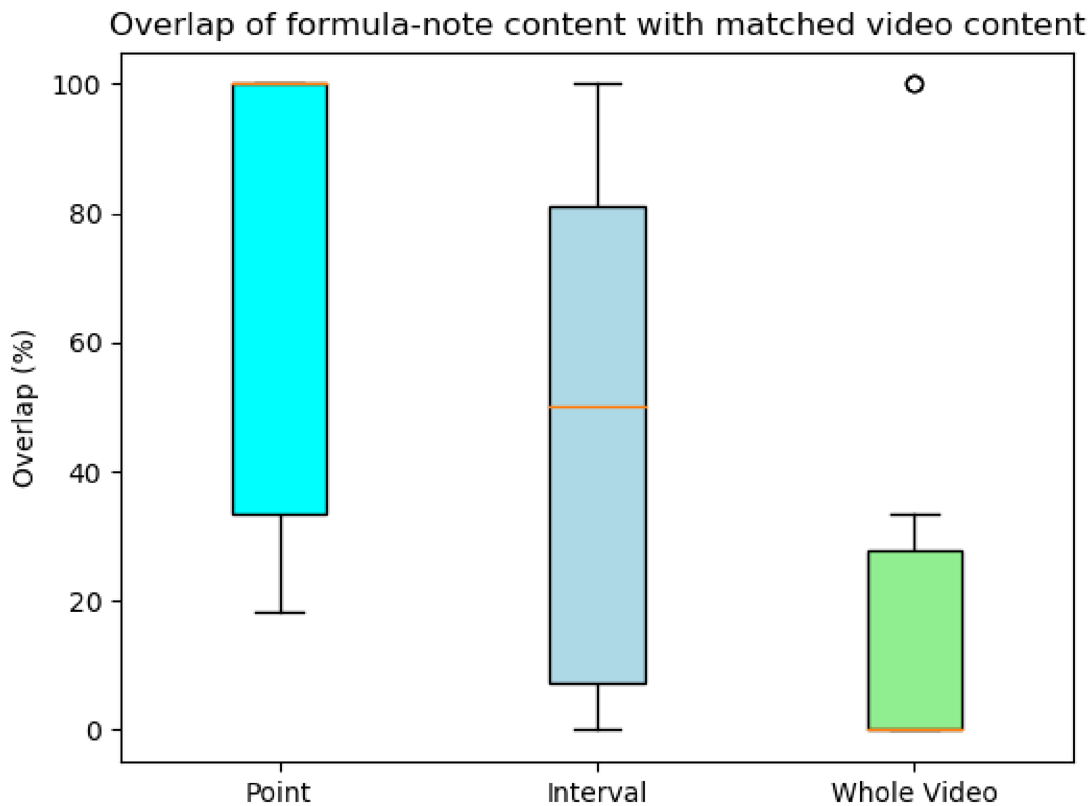

Figure 5.7: Boxplot results showing overlap between video results and formula-only content.

1. Similar to the case of textual note samples, the point condition demonstrated an overlap greater than 65% for at least 75% of the note samples.

2. Interval condition confirmed a wide range of overlap in the IQR region. However, it also had matches with 0% overlap.

3. In the whole video condition, most cases showed an overlap of less than 30%, hinting that the video results started playing randomly.

In the case of figure-type notes, the accuracy was predominantly low and recorded only seven matches. Hence, we did not include the box plot representation of this type.

5.6 Discussion

We showcased the accuracy of retrieving watched videos on both temporal levels and content types. The system matched approximately 78% of the note data to expected videos on one of the temporal levels. More than half of the returned videos exhibited a timepoint in point level or interval temporal level as preferred among the participants. The system did exceptionally well in matching textual note content to videos as finer legibility in textual content and matured attempts in word-spotting exists [88]. Analysis of content overlap between notes and the video results further highlighted interesting findings respective to retrieval in the three temporal levels. Close to half of the text-only samples were matched with timepoint in the point level and more importantly returned a relevant context with a good amount of overlap for timepoint in the interval level. Although a video point that is semantically closer to the note content exists, students' choice of topical relevance was at another timepoint in the same section or interval where the note content is present. The presented work has indeed conceptualized the difference of timestamp between the result to that of the expected video. This is an important finding in aiding the future video retrieval systems to allow a wider window of difference in the retrieved video timepoint.

However, in terms of formulas and drawings, the system matched less than half of the total links. The prototype system returned a few videos with point level matches that had a content overlap of 100%. Additionally, more than half of the equations and graphs that yielded no matching videos had an overlap of 100% in the actual observations. One of the reasons for low accuracy in the two non-textual content types could be the gap in generating non-textual Region of Interest (ROIs) in the Matcher block. The start and end of outlines for ROI bounding boxes are not obvious in equations and figures, which influences the matching effectiveness. ROI segmentation must be improved with training models prior to matching the note image to address this gap. Moreover,

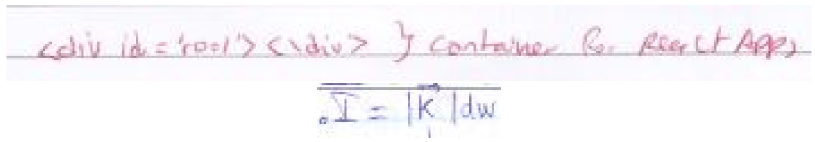

Figure 5.8: Link samples with low legibility

recognition of handwritten equations does not always yield success. Recent reports from Competitions on Recognition of Online Handwritten Mathematical Expressions (ICDAR 2019 CROHME) [55] have shown an accuracy of 77% for image-based recognition of equations. Similarly, advancements in feature matching approaches must find sound correspondences, especially with a higher proportion of false-positive matches in the case of free-form drawings.

We shouldnote that 38 video notes were excluded, marked as No-data, from the accuracy calculation. This was because the recognizer failed to detect the content before the note was passed to the matcher block due to legibility concerns, such as bad image quality or the information in the link might have been too less to detect content meaningfully. It can be evident in the sample data shown in Figure 5.8 where the legibility is compromised or the data to be recognized is too narrow for meaningful interpretation.

Additionally, in the proposed approach, the content types are physically categorized in notebooks for the hybrid type before feeding into the recognition module. The distinction between textual and non-textual content types is not always obvious. Of the 39 hybrid notes assigned to one of the three content types, one was dropped as a special case with no ground truth. Thus, 30 out of 38 hybrid notes were matched to the right videos, leading to an accuracy of 78.94%. This corroborates the right choice of content type assignment for improved accuracy in retrieving the video match.

5.7 Implementation requirements

By designing the medium-fidelity prototype, a number of technology impediments were identified that hinders the ability to recognize and match notes to videos effectively. The following sections outline a few essential requirements.

5.7.1 Automated Content Type Segmentation

The primary step of the proposed system is to interpret the text and non-text elements separately to carry out the subsequent recognition and matching processes. We must assign the mobile-scanned note content to the appropriate content labels with the identified types: text, formula, and figure before the corresponding recognition module assignment. Several researchers have proposed heuristic methods to separate text/non-text segments in online handwritten documents that utilize a combination of spatial and temporal information available from stroke level data [19], [20].

In classroom notebooks, however, this process is an arduous task without stroke-level information. Furthermore, the content of paper-based handwritten notes varies in size, shape, and orientation with different writing styles [76]. Most often, all the content types are closely connected and merge haphazardly, which is a significant issue in distinguishing them. Additional challenges include arbitrary layout, inconsistency in the note quality as they can be written on a variety of paper and can be old.

Previous work has introduced techniques that allow content segmentation, and layout analysis of a page in three main methods: top-down, bottom-up, and hybrid techniques [66]. The top-down techniques first detect the highest level of structures and then proceed to the bottom layer by training a classifier model that successively splits text, and non-text components [7], [38]. The top-down approach has shown a good (about 95%) prediction rate for identifying textual information, but not when non-textual elements like formulas and graphs are present. On the contrary, bottom-up methods begin with primary components and merge them into a segment [19]. These methods mostly focus on online handwritten documents and do not fit our case as the knowledge of digital strokes or individual pixel data are not available in paper-based documents. Hybrid methods combine the two approaches [78]. Florent et al. [61] proposed a hierarchical combination of Conditional Random Fields(CRFs) to extract the document layout of handwritten letters. Sarkar et al. [76] presented a typical connected components(CCs) classification-based method to separate text from non-text components. A similar method is proposed in [3] where the LBP operator is used to classify the CCs. Kundu et al. [49] recently proposed a method that applies Generative Adversarial Networks (GANs) where they considered 'text line extraction' in handwritten documents as an image-to-image translation task. Many more methods have considered the complex type of documents but have not addressed the most common issue of handwritten documents, i.e., overlapping components.

Recent advancement has introduced a few commercial/non-commercial modules that support text/non-text segmentation through pre-trained classifiers. Azure Custom vision[11] is one such service that facilitates the use of a machine-learning algorithm to train and classify the images. During the process of prototype implementation, we built a training model with hundreds of handwritten note samples by tagging labels pointing to the three content types. However, the test run with the data collected yielded us an utmost accuracy of 60%. The performance of the classifier significantly dropped for the images that had overlapping components of different types, that is, the hybrid type.

In the case of Hybrid images that consist of closely connected text and non-text components, assigning a label involves more work than classifying an image into an appropriate content type. Consider the example shown in Figure 3.7. Here, although the

[11] https://docs.microsoft.com/en-us/azure/cognitive-services/Custom-Vision-Service/overview

text and non-textual data elements are arbitrarily constructed, the textual information conveys semantically rich interpretations that can aid the accurate topical search of the related video. The pictorial data, in this case, might not be as helpful. Therefore, the classifier must point the following case to a text label instead of pointing to 3 independent types, which otherwise becomes taxing to process both in time and space. This poses an imperial need to possess prior knowledge of the content in the classifier.

Along with the content segmentation at the scanning end of the mobile interface, video slides also contain both text and non-text components. As seen in the implementation section, the Region of Interests (ROIs) extracted per image must be accurately mapped to non-textual components for efficient matching. Thus, there is an imperative need to develop an advanced content classification system that efficiently identifies different content types in both handwritten paper notebooks and the video slides as a mandatory pre-processing step.

5.7.2 Recognition Accuracy

Generally factors which affect accuracy of OCR can be disparate language classes [103], character merging and fragmentation [11], resolution (e.g. [84], [64]), illumination, skew and noise [46]. Handling each of these issues is significant in improving the overall recognition accuracy. As we collected the data was collected under controlled conditions, skew, blurriness and illumination are less likely to contribute towards OCR inaccuracy.

Character segmentation for multi-lingual handwritten recognition is now a significant area of research. The proposed system has been developed with recognizers to accommodate English, mathematical characters, and pictorial representations. However, the recognizer should be able to scale to notes of multiple languages. For instance, for notes with a mix of English and Korean/Chinese language, the OCR accuracy might significantly drop. Tesseract has proven to show very high accuracy for Latin characters. Park et al. [67] proposed a framework to tackle multi-lingual scripts. The words are split into multiple character images fed to a set of pre-trained classifier models. Unlike English, languages like Korean, Chinese have thousands of character classes. Thus, independent recognizers with explicit character segmentation and individual character recognition are important to map videos of interest.

Poor resolution of the notebook content scanned can greatly influence the OCR accuracy. Notes can be old or ill-maintained, which degrades the quality of content present in the notebook. Image interpolation as a pre-processing step in the Recognition pipeline can lessen the overall accuracy. Several researchers have proposed to tackle resolution degradation in printed document images [39], [15] [33]. Ankit et al. [51] showed an accuracy improvement of up to 21% in OCRing document images by employing a super-resolution-based pre-processing step.

5.7.3 Computing Time

Research suggests that even slightly higher retrieval latency by search engines can significantly decline in users' perceptions of result quality and engagement with the search results. The time between query submission to search response should be reduced as much as possible. In the case of the proposed system for video search, the processing time is dependent on the number of videos to be looked up.

In textual links, the processing time is directly proportional to the number of videos as the content is searched in the video transcript directly. However, the time taken depends on the number of ROIs extracted per video slide for formulas and figures. As the ROI extraction becomes more precise, the time factor increases with a potential improvement in the matching accuracy. A more grouped ROI can produce good speed, which might affect the accuracy of finding the right video match. Hence deeper work to address this trade-off is necessary to aid the system's ability to retrieve videos of interest efficiently.

CHAPTER 6

Conclusions

The work in this thesis investigated the use of engineering students' handwritten note content, captured when learning from educational videos, as a linking artifact to connect back to originally watched video context when reviewing. We observed distinct characteristics of notes, references to related videos, and verbatim overlap, a pivotal element to find relevant video content, in 10 engineering students' video notes. Results indicated overlap of some amount in at least 75% of the notes in all the identified types of note content. The retrieved video timepoint was conceptualized on the temporal scale, a timepoint in the same point, interval, or the whole video as that of the expected. We believe this approach accounts for the difference between the retrieved to expected video timestamp.

The findings from the 12-student lab study showed an overall preference for video timepoint retrieval in the 'interval' level. Also, a medium-fidelity prototype informed insights regarding the existing technology in designing a fully working application. Evaluating the prototype system against a set of 181 identified note samples, 77.5% of the samples were matched with the right videos as expected, with the video result playing in one of the temporal levels. Also, 97.5% of the total textual notes were matched to the videos observed in the ground truth. This is especially interesting given that more than half of the links identified were textual content indicating its broader use over other content types. The thematic analysis of interview transcripts in the study revealed that the proposed system of linking notebook content with corresponding videos was perceived as usable, effective, and easy to use. and can improve both the learning experience and learning outcomes. The implementation gaps given the existing technology such as automated content separation, recognition inaccuracy and computation time were derived in prototype development.

Along with the established results on the temporal specificity and video surrogates associated with matching video rendition, it was interesting to note how it affects users' linking behaviors. Students' response to the system's usability showcased its potential use as a supplement to bring context back to their notes rather than modeling notebook as a platform to include video-related annotations. Also, the existing linking practices as seen in the study transcripts in Table 4.2 are mostly manual approaches that confirm the proposed approach's novelty.

6.1 Limitations and Future Work

Although the results of our initial evaluation are positive, we recognized limitations within the study. With regard to the notes collection elaborated in chapter 3, the focus was only on technical notes to generate varied types of content types. The observations were based on a small corpus of class notes gathered from Engineering students at the University of British Columbia. The intention was to consider diverse content types accessible in engineering notes as confirmed during the pilot study. It is important to note that there is a great deal of individual variation in note-taking practices. So the characteristics of video-based notes described in this work are applicable only generally, not universally.

Additionally, the participants in the latter study, that reported on the systems' usability and requirements were recruited based on convenience sampling. All the students were part of varied degree levels pursuing different majors. Also, they were in no way connected and randomly picked. While an ample amount of qualitative data was collected that led to interesting inferences, the study might be vulnerable to selection bias and other influences that go beyond the control of the study.

Next, the collected notes benefited us with distinctive content types, data from other domains fall out of scope in this work, music-based courses or notes from a chemistry course, for example. The extension of this work in the future must address notes of other genre containing unique characters and different languages.

Along with the diverse genres, lecture characteristics such as the modality and the lecture structure is another critical factor that needs deeper investigation concerning the note-taking styles. Lectures can be seen to be presented with a classroom style written content or spoken content with talking heads. Previous work, however, shows no indications of lecture modality influencing the amount of lecture material students record or remember. Furthermore, the structure of lectures in a specific modality can have closely connected or disconnected/discontinuous segments stitched together depending on the complexity of the lecture topic. The temporal specificity of such lecture materials dramatically varies and is an important factor to consider in rendering the matching videos.

While the current work has confirmed students' positive attitude towards the instrumentation of notebook content-based video retrieval, the current work falls short of evidence in various learning settings, as shown in the examples below:

1. First, how a linking interface functions when working on assessments and preparing for exams. A detailed review of the expectations of a linking system's use when approaching class tests or an examination can contradict some of the observations here, chiefly pertaining to temporal specificity and surrogates of the retrieved video.

2. Second, how does the system perform on old notes, say older than a year or more. An interesting factor would be evaluating the validity of the matched video results.

Although the prototype designed for this work was of sufficient fidelity to draw inferences on the requirements, incorporating additional image correction methods can significantly improve performance. For instance, 38 records were dropped as the recognizer failed to translate the note content. Since the intention was to employ off-the-shelf recognition algorithms to implement a prototype of sufficient fidelity, a pre-processing step to improve the resolution was not a part of this work. Particularly for handwritten drawings, image warping as a mandatory pre-processing step can significantly aid the matching accuracy. Future work needs to address the optimization of matching non-textual data and more comparative experimental designs with more extensive data collections.

Lastly, the focus of this work emphasizes the interoperable nature between notebooks and video materials. However, nowadays, most paper documents have an electronic counterpart in a variety of media that is accessible on the World Wide Web or from some other online database or document corpus. While the video plays a central role in flipped classrooms and other video-based pedagogies, video is not the only media type students engage with when learning with video. Students make use of a diverse collection of information objects from within their information ecology. A typical engineering student, for example, will use her notebook during class, in labs, when doing homework, and tutorials, to name a few everyday academic tasks. Thus, as part of future work, the aim is to evaluate how notebooks can become more interoperable with other media types.

6.2 Extension of Application Interactions

We have demonstrated a set of primary user interactions with the presented prototype to associate a piece of watched-video note in a paper document with a particular video. The future can build on numerous extensions on that association, whose illustrations we briefly outline in this section.

At the most superficial level, when a user scans a part of notes, the mobile app could provide more options over the regular edit/crop features. Suppose the user feels that the information in the scanned part of notes could be insufficient in fetching the videos of interest. In that case, an option to add more information by typing/drawing free-hand can strengthen the context. This way, the actual notes can remain unchanged. Advanced collaborative actions such as inviting friends, colleagues, or peers to add on the scanned note can help locate the video counterpart better. Audio-augmented action items may also be designed on the edit page of scanned notes to provide bonus information.

The strength of the scanned query in fetching suitable videos may be improved, taking into account the types of errors likely to occur in the particular captured note data. One example of this is an indication of suspected errors in recognizing specific characters such as a mathematical symbol or highlighting a word as important; in this instance, a search engine may assign them a lower priority or treat highlighted characters as wildcards. Additionally, the same video(s) can often be searched using multiple scans of notes, pages of notes that entail a mathematical derivation, for example. The mobile app can implement these and many other extensions of "paper/digital integration" without requiring changes to the current writing processes, editing and sharing documents, giving such conventional paper documents a whole new layer of digital functionality.

The ability and time to recognize the scanned notes can also be improved by keeping frequently occurring or important records of capturing-retrieving video results. One can keep track of a map of a key portion of notes such as words, characters, or descriptors to the corresponding videos so that the occurrence of the same corpus can be used to enhance the recognition process, eventually increasing the likelihood of bringing up the same video results. To aid indexing of capture-retrieve process maps, students may choose markers such as emoticons, stars, or the capture time. Doing this, the future capture-retrieve actions on the same set of notes can be faster and may not require additional cueing to fetch the video match.

User feedback can be trivial in identifying a set of one or more candidate matching videos. Candidate videos can be weighted according to their probable relevance (for example, based on the number of other users who have scanned to fetch these videos or their popularity on the Internet). These weights can be applied in the iterative matching process. If there are likely to be delays or cost associated with processing a note query or receiving the results, this data can improve the performance of the local device, reduce processing costs, and provide helpful and timely user feedback.

Adding overlays or markups on the retrieved videos can also supplement the feedback process and, in turn, future lookup of the videos. For example, a student can attach text or an audio recording of his/her thoughts about a particular video segment for later retrieval as annotations. As another example, a user may also attach a picture(s) of notes used to retrieve the video that might assist other users when shared with the class group. Another possible functionality that may benefit users is to enable the lookup and subsequent inclusion of other related electronic counterparts of a specific video section. Say, a text document that details the functioning of a model in a PDF document or a podcast introduced in the video section. Thus, we believe that the retrieved video and its digital counterparts as markups can bolster the prospect of added note-related information, eventually creating a rich, interactive platform to connect to a wide range of information sources in a single scan.

Lastly, user-specific actions and history may also enhance many aspects of the system operation. Say, if the previous capture was within the last few minutes, it is very likely to be from the same video or area of study. Similarly, it is more likely that a note is being recorded in start-to-finish order. Or a student might frequently capture text-only or mathematics notes, as he/she finds it important for future review. Such user-specific factors can help the system establish the location of relevant videos in cases of ambiguity and also reduce the effort in looking up through a large video archive.

In summary, the proposed approach has established a new baseline for linking paper-based notes to instructional videos with room for future improvement. The design, implementation, and evaluation of the system have led to several interesting implications and extensions, including note-type scalability, recognition optimization, and improving interaction with the system. In turn, they are paving the way for future trials to investigate similar approaches in augmenting interoperability between notebook content and corresponding reference media.

References

[1] Awad George, Cees GM Snoek, Alan F. Smeaton and G. Quénot. Trecvid semantic indexing of video: A 6-year retrospective. *ITE Transactions on Media Technology and Applications*, pages 187–208, 2016.

[2] D. Bargeron, A. Gupta, J. Grudin, and E. Sanocki. Annotations for streaming video on the web: system design and usage studies. *Computer Networks*, 31(11-16):1139–1153, 1999.

[3] S. Bhowmik, R. Sarkar, and M. Nasipuri. Text and non-text separation in handwritten document images using local binary pattern operator. In *Proceedings of the First International Conference on Intelligent Computing and Communication*, pages 507–515. Springer, 2017.

[4] R. Boardman and M. A. Sasse. " stuff goes into the computer and doesn't come out" a cross-tool study of personal information management. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 583–590, 2004.

[5] R. Boardman, R. Spence, and M. A. Sasse. Too many hierarchies? the daily struggle for control of the workspace. In *Proc. HCI International 2003*, volume 1, pages 616–620. CRC Press: Taylor & Francis Group, 2003.

[6] P. Borlund. The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Information Research*, 8(3), 2003.

[7] T. M. Breuel. Two geometric algorithms for layout analysis. In *International workshop on document analysis systems*, pages 188–199. Springer, 2002.

[8] D. C. Bui, J. Myerson, and S. Hale. Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology*, 105(2):299, 2013.

[9] D. Byrne, P. Wilkins, G. J. F. Jones, A. F. Smeaton, and N. E. O'Connor. Measuring the Impact of Temporal Context on Video Retrieval. In *Proceedings of the 2008 International Conference on Content-Based Image and Video Retrieval*, pages 299–308, New York, {NY}, 2008. ACM.

[10] C. Carpineto, S. Mizzaro, G. Romano, and M. Snidero. Mobile Information Retrieval With Search Results Clustering: Prototypes and Evaluations. *Journal of the American Society for Information Science and Technology*, 60(5):877–895, 2009.

[11] R. G. Casey and E. Lecolinet. A survey of methods and strategies in character segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 18(7): 690–706, 1996.

[12] D. Castro, B. L. D. Bezerra, and M. Valença. Boosting the Deep Multidimensional

Long-Short-Term Memory Network for Handwritten Recognition Systems. In *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 127–132, Piscataway, {NJ}, 2018. IEEE.

[13] M. T. Chi. Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in cognitive science*, 1(1):73–105, 2009.

[14] M. G. Christel. Evaluation and User Studies With Respect to Video Summarization and Browsing. In *Multimedia Content Analysis, Management, and Retrieval 2006*, page 60730M, Bellingham, {WA}, 2006. International Society for Optics and Photonics, SPIE.

[15] F. Cloppet, N. Vincent, and Others. Adaptive fuzzy model for blur estimation on document images. *Pattern Recognition Letters*, 86:42–48, 2017.

[16] H. Crompton, D. Burke, and K. H. Gregory. The use of mobile learning in PK-12 education: A systematic review. *Computers & Education*, 110:51–63, 2017.

[17] E. Dahlstrom, J. D. Walker, and C. Dziuban. ECAR study of undergraduate students and information technology. Technical report, 2013, 2013.

[18] S. Deb and P. Bhattacharya. Enhancing personalized learning with interactive note taking on video lectures–an analysis of effective hci design on learning outcome. In *International Conference on Next Generation Computing Technologies*, pages 558–567. Springer, 2017.

[19] A. Delaye and C.-L. Liu. Text/non-Text Classification in Online Handwritten Documents With Conditional Random Fields. In *Chinese Conference on Pattern Recognition*, pages 514–521, Berlin, Germany, 2012. Springer.

[20] A. Delaye and C.-L. Liu. Contextual Text/non-Text Stroke Classification in Online Handwritten Notes With Conditional Random Fields. *Pattern Recognition*, 47(3): 959–968, 2014.

[21] S. J. DeLozier and M. G. Rhodes. Flipped classrooms: a review of key ideas and recommendations for practice. *Educational psychology review*, 29(1):141–151, 2017.

[22] S. Dodson, I. Roll, M. Fong, D. Yoon, N. M. Harandi, and S. Fels. An active viewing framework for video-based learning. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, pages 1–4, 2018.

[23] S. Dodson, I. Roll, N. M. Harandi, S. Fels, and D. Yoon. Weaving Together Media, Technologies, and People: Students' Information Practices in Flipped Classrooms. *Information and Learning Sciences*, 120(7/8):519–540, 2019.

[24] B. Dorn, L. B. Schroeder, and A. Stankiewicz. Piloting {TrACE}: Exploring Spatiotemporal Anchored Collaboration in Asynchronous Learning. In *Proceedings of the 18th {ACM} Conference on Computer Supported Cooperative Work & Social*

Computing, pages 393–403, New York, {NY}, 2015. ACM.

[25] J. C. Dunn. Remember-Know: A Matter of Confidence. *Psychological Review*, 111(2): 524–542, 2004.

[26] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner. G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2):175–191, 2007.

[27] M. Fong, S. Dodson, X. Zhang, I. Roll, and S. Fels. {ViDeX}: A Platform for Personalizing Educational Videos. In *Proceedings of the 18th {ACM/IEEE} Joint Conference on Digital Libraries*, pages 331–332, New York, {NY}, 2018. ACM.

[28] A. Fouse, N. Weibel, E. Hutchins, and J. D. Hollan. {ChronoViz}: A System for Supporting Navigation of Time-Coded Data. In *Extended Abstracts on Human Factors in Computing Systems*, pages 299–304. ACM, New York, {NY}, 2011.

[29] K. Franke and M. Köppen. Towards an universal approach to background removal in images of bankchecks. In *in Proceedings 6th International Workshop on Frontiers in Handwriting Recognition (IWFHR), Tajon, Korea*. Citeseer, 1998.

[30] M. Furini, S. Mirri, and M. Montangero. {TagLecture}: The Gamification of Video Lecture Indexing Through Quality-Based Tags. In *2017 {IEEE} Symposium on Computers and Communications ({ISCC})*, pages 122–127, Piscataway, {NJ}, 2017. IEEE.

[31] J. Galanek, D. Gierdowski, and D. Christopher Brooks. ECAR study of undergraduate students and information technology. Technical report, 2018, 2018.

[32] R. Gangaiamaran and M. Pasupathi. Review on Use of Mobile Apps for Language Learning. *International Journal of Applied Engineering Research*, 12(21):11242–11251, 2017.

[33] U. Garain, A. Jain, A. Maity, and B. Chanda. Machine reading of camera-held low quality text images: an ICA-based image enhancement approach for improving OCR accuracy. In *2008 19th International Conference on Pattern Recognition*, pages 1–4. IEEE, 2008.

[34] F. Guimbretière. Paper Augmented Digital Documents. In *Proceedings of the 16th Annual {ACM} Symposium on User Interface Software and Technology*, pages 51–60, New York, {NY}, 2003. ACM.

[35] K. Gynther. Design Framework for an Adaptive MOOC Enhanced by Blended Learning: Supplementary Training and Personalized Learning for Teacher Professional Development. *Electronic Journal of e-Learning*, 14(1):15–30, 2016.

[36] N. M. Harandi, F. Agharebparast, L. Linares, S. Dodson, I. Roll, M. Fong, D. Yoon, and S. Fels. Student Video-Usage in Introductory Engineering Courses. *Proceedings of*

the Canadian Engineering Education Association (CEEA), pages 1–8, 2018.

[37] J.-H. Hsiao and Y.-Z. Wang. Vcenter: A digital video management system with mobile search service. In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*, page 508, New York, NY, 2007. Association for Computing Machinery.

[38] E. Indermühle, H. Bunke, F. Shafait, and T. Breuel. Text Versus Non-Text Distinction in Online Handwritten Documents. In *Proceedings of the 2010 {ACM} Symposium on Applied Computing*, pages 3–7, New York, {NY}, 2010. ACM.

[39] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graphical models and image processing*, 53(3):231–239, 1991.

[40] R. S. Jansen, D. Lakens, and W. A. IJsselsteijn. An integrative review of the cognitive costs and benefits of note-taking. *Educational Research Review*, 22:223–233, 2017.

[41] R. S. Jansen, D. Lakens, and W. A. IJsselsteijn. An integrative review of the cognitive costs and benefits of note-taking. *Educational Research Review*, 22:223–233, 2017.

[42] S. Järvelä, P. Näykki, J. Laru, and T. Luokkanen. Structuring and Regulating Collaborative Learning in Higher Education With Wireless Networks and Mobile Tools. *Journal of Educational Technology & Society*, 10(4):71–79, 2007.

[43] W. Jones. Personal Information Management. *Annual Review of Information Science and Technology*, 41(1):453–504, 2007.

[44] M. Kam, J. Wang, A. Iles, E. Tse, J. Chiu, D. Glaser, O. Tarshish, and J. Canny. Livenotes: A System for Cooperative and Augmented Note-Taking in Lectures. In *Proceedings of the {SIGCHI} Conference on Human Factors in Computing Systems*, pages 531–540, New York, {NY}, 2005. ACM.

[45] L. S. Kate, M. M. Waghmare, and A. Priyadarshi. An approach for automated video indexing and video search in large lecture video archives. In *2015 international conference on pervasive computing (ICPC)*, pages 1–5. IEEE, 2015.

[46] A. Khotanzad and Y. H. Hong. Invariant image recognition by Zernike moments. *IEEE Transactions on pattern analysis and machine intelligence*, 12(5):489–497, 1990.

[47] K. Kobayashi. Combined effects of note-taking/-reviewing on learning and the enhancement through interventions: A meta-analytic review. *Educational Psychology*, 26(3):459–477, 2006.

[48] A. Kukulska-Hulme and L. Shield. An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(3): 271–289, 2008.

[49] S. Kundu, S. Paul, S. K. Bera, A. Abraham, and R. Sarkar. Text-line extraction from handwritten document images using GAN. *Expert Systems with Applications*, 140: 112916, 2020.

[50] G. Landow. *Hypertext 3.0: Critical Theory and New Media in an Era of Globalization.* Johns Hopkins University Press, Baltimore, {MD}, 2006.

[51] A. Lat and C. V. Jawahar. Enhancing OCR Accuracy with Super Resolution. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 3162–3167, 2018.

[52] C. Liu, C.-L. Yang, J. J. Williams, and H.-C. Wang. Notestruct: Scaffolding note-taking while learning from online videos. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.

[53] M. R. Lyu, E. Yau, and S. Sze. A multilingual, multimodal digital video library system. In *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 145–153, New York, NY, 2002. Association for Computing Machinery.

[54] W. E. Mackay, G. Pothier, C. Letondal, K. Bøegh, and H. E. Sørensen. The missing link: augmenting biology laboratory notebooks. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 41–50, 2002.

[55] M. Mahdavi, R. Zanibbi, H. Mouchère, C. Viard-Gaudin, and U. Garain. {ICDAR} 2019 {CROHME}+ {TFD}: Competition on Recognition of Handwritten Mathematical Expressions and Typeset Formula Detection. In *15th {IAPR} International Conference on Document Analysis and Recognition*, pages 1533–1538, Piscataway, {NJ}, 2019. IEEE.

[56] A. Mangen. Hypertext Fiction Reading: Haptics and Immersion. *Journal of Research in Reading*, 31(4):404–419, 2008.

[57] G. Marchionini, Y. Song, and R. Farrell. Multimedia Surrogates for Video Gisting: Toward Combining Spoken Words and Imagery. *Information Processing & Management*, 45(6):615–630, 2009.

[58] C. Marshall. Toward an Ecology of Hypertext Annotation. In *Proceedings of the Ninth {ACM} Conference on Hypertext and Hypermedia*, pages 40–49, New York, {NY}, 1998. ACM.

[59] C. Marshall and A. J. B. Brush. Exploring the Relationship Between Personal and Public Annotations. In *Proceedings of the 4th {ACM}/{IEEE-CS} Joint Conference on Digital Libraries*, pages 349–357, New York, {NY}, 2004. ACM.

[60] A. Mitrovic, V. Dimitrova, L. Lau, A. Weerasinghe, and M. Mathews. Supporting constructive video-based learning: requirements elicitation from exploratory studies. In *International Conference on Artificial Intelligence in Education*, pages 224–237. Springer, 2017.

[61] F. Montreuil, S. Nicolas, E. Grosicki, and L. Heutte. A new hierarchical handwritten document layout extraction based on conditional random field modeling. In *2010 12th International Conference on Frontiers in Handwriting Recognition*, pages 31–36. IEEE, 2010.

[62] X. Mu. Towards Effective Video Annotation: An Approach to Automatically Link Notes With Video Content. *Computers & Education*, 55(4):1752–1763, 2010.

[63] P. A. Mueller and D. M. Oppenheimer. The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. *Psychological Science*, 25(6): 1159–1168, 2014.

[64] A. Namane and M. A. Sid-Ahmed. Character scaling by contour method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(6):600–606, 1990.

[65] A. J. Newell and L. D. Griffin. Writer identification using oriented basic image features and the delta encoding. *Pattern Recognition*, 47(6):2255–2265, 2014.

[66] O. Okun, D. Dœrmann, and M. Pietikainen. Page segmentation and zone classification: the state of the art. Technical report, OULU UNIV (FINLAND) DEPT OF ELECTRICAL ENGINEERING, 1999.

[67] J. Park, E. Lee, Y. Kim, I. Kang, H. I. Koo, and N. I. Cho. Multi-Lingual Optical Character Recognition System Using the Reinforcement Learning of Character Segmenter. *IEEE Access*, 8:174437–174448, 2020.

[68] J. Pearson, G. Buchanan, and H. Thimbleby. The Reading Desk: Applying Physical Interactions to Digital Documents. In *Proceedings of the {SIGCHI} Conference on Human Factors in Computing Systems*, pages 3199–3202, New York, {NY}, 2011. ACM.

[69] J. Pearson, S. Robinson, and M. Jones. Paperchains: Dynamic sketch+ voice annotations. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 383–392, 2015.

[70] S. T. Peverly, P. C. Vekaria, L. A. Reddington, J. F. Sumowski, K. R. Johnson, and C. M. Ramsay. The relationship of handwriting speed, working memory, language comprehension and outlines to lecture note-taking and test-taking among college students. *Applied Cognitive Psychology*, 27(1):115–126, 2013.

[71] V. Pham, T. Bluche, C. Kermorvant, and J. Louradour. Dropout Improves Recurrent Neural Networks for Handwriting Recognition. In *2014 14th International Conference on Frontiers in Handwriting Recognition*, pages 285–290, Piscataway, {NJ}, 2014. IEEE.

[72] A. Piolat, T. Olive, and R. T. Kellogg. Cognitive effort during note taking. *Applied cognitive psychology*, 19(3):291–312, 2005.

[73] R. Plamondon and S. N. Srihari. Online and Off-Line Handwriting Recognition: A Comprehensive Survey. {*IEEE*} *Transactions on Pattern Analysis and MacHine Intelligence*, 22(1):63–84, 2000.

[74] J. Puigcerver. Are multidimensional recurrent layers really necessary for handwritten text recognition? In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 67–72. IEEE, 2017.

[75] S. Redding. Through the Student's Eyes: A Perspective on Personalized Learning and Practice Guide for Teachers. *Center on Innovations in Learning, Temple University*, 2013.

[76] R. Sarkar, S. Moulik, N. Das, S. Basu, M. Nasipuri, and M. Kundu. Suppression of non-text components in handwritten document images. In *2011 International Conference on Image Information Processing*, pages 1–7. IEEE, 2011.

[77] A. Sellen and R. H. R. Harper. *The Myth of the Paperless Office*. {MIT} Press, Cambridge, {MA}, 2001.

[78] R. W. Smith. Hybrid page layout analysis via tab-stop detection. In *2009 10th International Conference on Document Analysis and Recognition*, pages 241–245. IEEE, 2009.

[79] Y.-C. Song, G.-D. Chen, and L.-Y. Li. Improving E-Book Reading With Information Cues: An User Investigation and Suggestion. In *2013 {IEEE} 13th International Conference on Advanced Learning Technologies*, pages 261–263, Piscataway, {NJ}, 2013. IEEE.

[80] S. N. Srihari. Recognition of handwritten and machine-printed text for postal address interpretation. *Pattern recognition letters*, 14(4):291–302, 1993.

[81] T. F. Stahovich and H. Lin. Enabling Data Mining of Handwritten Coursework. *Computers & Graphics*, 57:31–45, 2016.

[82] S. Sweeney and F. Crestani. Effective Search Results Summary Size and Device Screen Size: Is There a Relationship? *Information Processing & Management*, 42(4): 1056–1074, 2006.

[83] T. Tuna, J. Subhlok, L. Barker, V. Varghese, O. Johnson, and S. Shah. Development and evaluation of indexed captioned searchable videos for stem coursework. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education*, pages 129–134, 2012.

[84] R. A. Ulichney and D. E. Troxel. Scaling binary images with the telescoping template. *IEEE transactions on pattern analysis and machine intelligence*, (3):331–335, 1982.

[85] P. Van Meter, L. Yokoi, and M. Pressley. College Students' Theory of Note-Taking Derived From Their Perceptions of Note-Taking. *Journal of Educational Psychology*, 86 (3):323–338, 1994.

[86] P. C. Vashist, A. Pandey, and A. Tripathi. A Comparative Study of Handwriting Recognition Techniques. In *2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, pages 456–461, Piscataway, {NJ}, 2020. IEEE.

[87] F. Wang, C.-W. Ngo, and T.-C. Pong. Structuring low-quality videotaped lectures for cross-reference browsing by video text analysis. *Pattern Recognition*, 41(10):3257–3269,

2008.

[88] K. Wang, B. Babenko, and S. Belongie. End-To-End Scene Text Recognition. In *2011 International Conference on Computer Vision*, pages 1457–1464, Piscataway, {NJ}, 2011. IEEE.

[89] Z. Wang, A. C. Bovik, and H. R. Sheikh. Structural similarity based image quality assessment. *Digital Video image quality and perceptual coding*, pages 225–241, 2005.

[90] M.-F. Weng and Y.-Y. Chuang. Multi-Cue Fusion for Semantic Video Indexing. In *Proceedings of the 16th ACM International Conference on Multimedia*, pages 71–80, New York, {NY}, 2008. ACM.

[91] R. W. White and D. Morris. Investigating the Querying and Browsing Behavior of Advanced Search Engine Users. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 255–262, New York, {NY}, 2007. ACM.

[92] S. Whittaker. Personal information management: from information consumption to curation. *Annual review of information science and technology*, 45(1):1, 2011.

[93] L. D. Wilcox, P. Chiu, M. Sasaoka, J. Miyazaki, D. L. Hecht, and L. N. Flores. System and method for video access from notes or summaries, Jan. 12 2010. US Patent 7,647,555.

[94] A. E. Witherby and S. K. Tauber. The Current Status of Students' Note-Taking: Why and How Do Students Take Notes? *Journal of Applied Research in Memory and Cognition*, 8(2):139–153, 2019.

[95] N. Yamamoto. An Interactive Learning System Using Smartphone: Improving Students' Learning Motivation and Self-Learning. In *2014 Ninth International Conference on Broadband and Wireless Computing, Communication and Applications*, pages 428–431. IEEE, 2014.

[96] H. Yang and C. Meinel. Content based lecture video retrieval using speech and video text information. *IEEE transactions on learning technologies*, 7(2):142–154, 2014.

[97] S.-J. Yang, G.-D. Chen, and L.-Y. Li. How Students Use Contextual Cues in Finding Information in Paper and Electronic Textbooks. In *2011 {IEEE} 11th International Conference on Advanced Learning Technologies*, pages 302–304, Piscataway, {NJ}, 2011. IEEE.

[98] M. Yarmand, D. Yoon, S. Dodson, I. Roll, and S. S. Fels. "Can you believe [1:21]?!": Content and time-based reference patterns in video comments. In *Proceedings of the 2019 {CHI} Conference on Human Factors in Computing Systems*, pages 489:1—-489:12, New York, {NY}, 2019. ACM.

[99] R. Yeh, C. Liao, S. Klemmer, F. Guimbretière, B. Lee, B. Kakaradov, J. Stamberger,

and A. Paepcke. {ButterflyNet}: A Mobile Capture and Access System for Field Biology Research. In *Proceedings of the {SIGCHI} Conference on Human Factors in Computing Systems*, pages 571–580, New York, {NY}, 2006. ACM.

[100] C. J. Yuan and W. B. Seales. Guided linking: Efficiently making image-to-transcript correspondence. In *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*, page 471, New York, NY, 2001. Association for Computing Machinery.

[101] B. Zhao, S. Lin, X. Luo, S. Xu, and R. Wang. A Novel System for Visual Navigation of Educational Videos Using Multimodal Cues. In *Proceedings of the 25th {ACM} International Conference on Multimedia*, pages 1680–1688, New York, {NY}, 2017. ACM.

[102] B. Zhao, J. Tao, M. Yang, Z. Tian, C. Fan, and Y. Bai. Deep imitator: handwriting calligraphy imitation via deep attention networks. *Pattern Recognition*, page 107080, 2020.

[103] H. Zheng, J. Wang, Z. Huang, Y. Yang, and R. Pan. Chinese/english mixed character segmentation as semantic segmentation. *arXiv preprint arXiv:1611.01982*, 2016.

Appendices

APPENDIX A

Participant data

A.1 Notebook Samples

The notebook samples are grouped under each participant number, example P1 and can be accessed here[1]. Further details about each of the notebook sample is detailed in the following table A.1.

Table A.1: Notebook Sample details

.	Notes	Actual Type	Grouped Type	Course	Video, timepoint	. Overlap	. No Overlap	% of Overlap
1	P1r1	hybrid	equation	ELEC211	Coloumb's law, 3.23	4	6	40
2	P1r2	hybrid	text	ELEC211	Electric flux, 5.11	3	3	50
3	P1r3	hybrid	equation	ELEC211	Electric Dipole, 3.48	2	9	18.18
4	P1r4	hybrid	text	ELEC211	Biot-Savart law, 3.05	7	2	77.78
5	P1r5	equation	equation	ELEC211	Magnetic Force, 1.06	4	7	36.36

[1] https://ubcca-my.sharepoint.com/:f:/g/personal/ranjs92_student_ubc_ca/EtvgSRjJKktPqKHcXUxUauIBynZU9R26ePuSuwC-5YvZpA?e=VN1Kei

6	P1r6	hybrid	text	ELEC211	Magnetic materials, 4.54	16	5	76.19
7	P1r7	equation	equation	ELEC211	Magnetic Flux, 1.41	3	6	33.33
8	P1r8	equation	equation	ELEC211	Current, 4.17	2	7	22.22
9	P1r9	hybrid	text	ELEC211	Unsure			
10	P2r1	equation	equation	BMEG220	Potential Energy, 1:43, 1:56, 2:09	2	4	33.33
11	P2r2	graph	graph	BMEG220	Electric flux, 2.10	3	1	75.00
12	P2r3	equation	equation	BMEG220	Current, 3.05	3	4	42.86
13	P2r4	graph	graph	BMEG220	Conducting materials, 1.12	1	1	50.00
14	P2r5	text	text	BMEG220	Classification of Electrical networks, 1.25	3	2	60.00
15	P2r6	graph	graph	BMEG220	Electric Dipole, 1.20	1	1	50.00
16	P2r7	equation	equation	BMEG220	Capacitance, 1.32	1	0	100.00
17	P2r8	graph	graph	BMEG220	Ideal op-amps, 7.40	0	1	0.00

18	P2r9	equation	equation	BMEG220	Biot-Savart law, 4.05	1	0	100.00
19	P2r10	graph	graph	BMEG220	Magnetic Flux, 2.01	0	1	0.00
20	P2r11	equation	equation	BMEG220	Magnetic Force, 5.01	3	0	100.00
21	P2r12	graph	graph	BMEG220	Solenoids & Toroids, 7.57	1	0	100.00
22	P2r13	hybrid	text	BMEG220	Magnetic materials, 4.51	15	0	100.00
23	P2r14	equation	equation	BMEG220	Faraday's law, 6.47	1	0	100.00
24	P2r15	hybrid	text	BMEG220	Coulomb's law, 1.51	4	0	100.00
25	P3r1	text	text	CPSC340	Neural Networks - Prediction, 1:27	3	5	37.50
26	P3r2	hybrid	text	CPSC340	Neural Networks - Prediction, 4.25	2	1	66.67

27	P3r3	text	text	CPSC340	Neural Networks - Prediction, 7.25	3	1	75.00
28	P3r4	text	text	CPSC340	Neural Networks - Training, 22.19	5	1	83.33
29	P3r5	text	text	CPSC340	Neural Networks - Training, 22.57	4	0	100.00
30	P3r6	text	text	CPSC340	Neural Networks - Training, 23.56	3	3	50.00
31	P3r7	text	text	CPSC340	Convolutional networks, 2.04	6	0	100.00
32	P3r8	equation	equation	CPSC340	Convolutional networks, 2.31	1	0	100.00
33	P3r9	text	text	CPSC340	K-Nearest Neighbours, 32.27	7	0	100.00

34	P3r10	text	text	CPSC340	K-Nearest Neighbours, 36.50	7	0	100.00
35	P3r11	text	text	CPSC340	K-Nearest Neighbours, 37.53	3	3	50.00
36	P3r12	text	text	CPSC340	K-Nearest Neighbours, 40.37	2	0	100.00
37	P3r13	hybrid	text	CPSC340	Ordinary Least Squares, 29.08	2	1	66.67
38	P3r14	text	text	CPSC340	Ordinary Least Squares, 33.08	2	0	100.00
39	P3r15	hybrid	text	CPSC340	Normal Equations, 3.31	4	0	100.00
40	P3r16	equation	equation	CPSC340	Normal Equations, 11.40	1	0	100.00

41	P3r17	text	text	CPSC340	Exploratory Data Analysis, 0.35	5	6	45.45
42	P3r18	text	text	CPSC340	Exploratory Data Analysis, 2.36	7	4	63.64
43	P3r19	text	text	CPSC340	Exploratory Data Analysis, 5.45	3	2	60.00
44	P4r1	hybrid	graph	BMEG220	Coulomb's Law, 0:14	1	2	33.33
45	P4r2	equation	equation	BMEG220	Coulomb's Law, 1:18	1	5	16.67
46	P4r3	hybrid	text	BMEG220	Potential Energy, 0:20	3	7	30.00
47	P4r4	equation	equation	BMEG220	Potential Energy, 1:35	1	6	14.29
48	P4r5	hybrid	text	BMEG220	Potential Energy, 2:19	3	1	75.00
49	P4r6	equation	equation	BMEG220	Potential Energy, 3:04	2	4	33.33
50	P4r7	text	text	BMEG220	Classification of Electrical Networks, 0:27	9	1	90.00

51	P4r8	text	text	BMEG220	Classification of Electrical Networks, 0:44	1	2	33.33
52	P4r9	text	text	BMEG220	Classification of Electrical Networks, 1:04	2	2	50.00
53	P4r10	text	text	BMEG220	Classification of Electrical Networks, 2:16	4	3	57.14
54	P4r11	text	text	BMEG220	Classification of Electrical Networks, 3:09	4	4	50.00
55	P4r12	text	text	BMEG220	Classification of Electrical Networks, 4:05	4	6	40.00
56	P4r13	graph	graph	BMEG220	Capacitance, 0:33	1	0	100.00
57	P4r14	equation	equation	BMEG220	Capacitance, 1:13	1	1	50.00

58	P4r15	text	text	BMEG220	Capacitance, 1:32	0	4	0.00
59	P4r16	text	text	BMEG220	Ideal OpAmps, 0:40	2	2	50.00
60	P4r35	text	text	BMEG220	Ideal OpAmps, 1:36	2	0	100.00
61	P4r17	hybrid	graph	BMEG220	Ideal OpAmps, 1:18	1	0	100.00
62	P4r18	text	text	BMEG220	Ideal OpAmps, 2:34	5	2	71.43
63	P4r19	hybrid	text	BMEG220	Ideal OpAmps, 5:01	1	2	33.33
64	P4r20	hybrid	text	BMEG220	Ideal OpAmps, 7:25	7	3	70.00
65	P4r21	hybrid	text	BMEG220	Magnetic Flux, 0:15	5	1	83.33
66	P4r22	hybrid	text	BMEG220	Magnetic Flux, 0:39	2	6	25.00

67	P4r23	equation	equation	BMEG220	Magnetic Flux, 1:32	2	0	100.00
68	P4r24	equation	equation	BMEG220	Magnetic Flux, 2:00	1	5	16.67
69	P4r25	hybrid	text	BMEG220	Magnetic Flux, 3:06	3	0	100.00
70	P4r26	hybrid	text	BMEG220	Magnetic Flux, 3:16	1	2	33.33
71	P4r27	hybrid	text	BMEG220	Magnetic Flux, 4:35	1	3	25.00
72	P4r28	hybrid	text	BMEG220	Solenoids and Toroids, 0:26	1	3	25.00
73	P4r29	hybrid	text	BMEG220	Solenoids and Toroids, 3:03	1	5	16.67
74	P4r30	text	text	BMEG220	Solenoids and Toroids, 6:14	0	2	0.00
75	P4r31	hybrid	graph	BMEG220	Solenoids and Toroids, 6:45	1	1	50.00
76	P4r32	hybrid	text	BMEG220	Solenoids and Toroids, 8:30	0	2	0.00
77	P4r33	hybrid	graph	BMEG220	Solenoids and Toroids, 10:21	2	0	100.00

78	P4r34	equation	equation	BMEG220	Solenoids and Toroids, 13:09	2	1	66.67
79	P5r1	equation	equation	ELEC211	Biot-Savart law, 2.20	1	0	100.00
80	P5r2	equation	equation	ELEC211	Biot-Savart law, 3.29	1	0	100.00
81	P5r3	equation	equation	ELEC211	Biot-Savart law, 5.12	1	0	100.00
82	P5r4	equation	equation	ELEC211	Biot-Savart law, 6.05	1	0	100.00
83	P5r5	equation	equation	ELEC211	Capacitance, 1.20	1	0	100.00
84	P5r6	equation	equation	ELEC211	Capacitance, 2.10	1	0	100.00
85	P5r7	equation	equation	ELEC211	Capacitance, 3.10	1	0	100.00
86	P5r8	equation	equation	ELEC211	Capacitance, 3.42	1	0	100.00
87	P5r9	equation	equation	ELEC211	Capacitance, 4.15	1	0	100.00
88	P5r10	equation	equation	ELEC211	Coulomb's law, 1.50	1	0	100.00
89	P5r11	equation	equation	ELEC211	Magnetic Force, 3.53	0	1	0.00

90	P5r12	equation	equation	ELEC211	Magnetic Flux, 1.35	1	0	100.00
91	P5r13	equation	equation	ELEC211	Magnetic Flux, 2.15	1	0	100.00
92	P5r14	equation	equation	ELEC211	Magnetic Flux, 2.51	3	0	100.00
93	P5r15	equation	equation	ELEC211	Magnetic Flux, 3.20	1	0	100.00
94	P5r16	equation	equation	ELEC211	Magnetic Force, 0.40	1	1	50.00
95	P5r17	equation	equation	ELEC211	Magnetic Force, 2.43	1	0	100.00
96	P5r18	equation	equation	ELEC211	Magnetic Force, 6.50	0	1	0.00
97	P5r19	equation	equation	ELEC211	Magnetic Materials, 7.14	1	0	100.00
98	P5r20	equation	equation	ELEC211	Magnetic Materials, 7.40	1	0	100.00
99	P6r1	text	text	React	0.1	3	0	100.00
100	P6r2	text	text	React	0.17	0	1	0.00
101	P6r3	text	text	React	0.47	5	0	100.00
102	P6r4	text	text	React	3.23	3	0	100.00
103	P6r5	hybrid	graph	React	3.29	2	2	50.00
104	p6r6	text	text	React	3.29	2	3	40.00
105	P6r7	hybrid	graph	React	5.07	2	0	100.00

106	P6r8	text	text	React	6.38	3	1	75.00
107	P6r9	text	text	React	10.07	2	6	25.00
108	P6r10	text	text	React	11.34	3	0	100.00
109	P6r11	text	text	React	13.43	1	4	20.00
110	P6r12	text	text	React	16.35	2	1	66.67
111	P6r13	text	text	React	28.31	4	4	50.00
112	P6r14	text	text	React	39.3	3	1	75.00
113	P7r1	text	text	CPSC340	CS3 Design in Computing, 12.09	6	3	66.67
114	P7r2	text	text	CPSC340	CS3 Design in Computing, 12.40	2	2	50.00
115	P7r3	text	text	CPSC340	Fundamentals of learning, 35.36	3	6	33.33
116	P7r4	text	text	CPSC340	Fundamentals of learning, 41.43	7	0	100.00
117	P7r5	text	text	CPSC340	Fundamentals of learning, 42.48	7	0	100.00
118	P7r6	text	text	CPSC340	k-nearest neighbors,6.37	3	1	75.00

119	P7r7	text	text	CPSC340	k-nearest neighbors, 22.14	7	2	77.78
120	P7r8	text	text	CPSC340	k-nearest neighbors, 29.49	7	0	100.00
121	P7r9	text	text	CPSC340	L2 Regularization, 12.52	7	0	100.00
122	P7r10	equation	equation	CPSC340	L2 Regularization, 13.30	2	0	100.00
123	P7r11	hybrid	graph	CPSC340	PCA Intuition, 4.13	1	0	100.00
124	P7r12	hybrid	text	CPSC340	PCA training, 27.22	11	3	78.57
125	P8r1	hybrid	text	EOSC510	Mean & Variance, 9.03	1	0	100.00
126	P8r2	text	text	EOSC510	Mean & Variance, 13.11	3	0	100.00
127	P8r3	text	text	EOSC510	Non-linear optimization, 0.02	2	0	100.00

128	P8r4	text	text	EOSC510	Non-linear optimization, 4.30	6	0	100.00
129	P8r5	text	text	EOSC510	Non-linear optimization, 13.00	4	2	66.67
130	P8r6	hybrid	equation	EOSC510	Non-linear optimization, 15.40	2	3	40.00
131	P8r7	hybrid	text	EOSC510	Eigen vector approach, 0.08	4	2	66.67
132	P8r8	text	text	EOSC510	Classification: k nearest neighbouring classifier , 0.05	5	0	100.00
133	P8r9	text	text	EOSC510	Classification: k nearest neighbouring classifier, 9.36	3	1	75.00

134	P8r10	text	text	EOSC510	Classification: k nearest neighbouring classifier, 13.02	6	0	100.00
135	P8r11	text	text	EOSC510	Mean & Variance, 4.17	2	0	100.00
136	P8r12	hybrid	equation	EOSC510	Mean & Variance, 6.28	1	0	100.00
137	P9r1	text	text	EOSC510	Correlation & Regression, 0.34	0	1	0.00
138	P9r2	text	text	EOSC510	Correlation & Regression, 0.56	0	1	0.00
139	P9r3	equation	equation	EOSC510	Mean & Variance, 2.15	1	0	100.00
140	P9r4	equation	equation	EOSC510	Mean & Variance, 4.26	1	0	100.00
141	P9r5	equation	equation	EOSC510	Mean & Variance, 4.44	1	0	100.00
142	P9r6	equation	equation	EOSC510	Mean & Variance, 6.40	2	0	100.00
143	P9r7	equation	equation	EOSC510	Mean & Variance, 7.28	1	0	100.00

144	P9r8	text	text	EOSC510	Mean & Variance, 8.39	5	0	100.00
145	P9r9	equation	equation	EOSC510	Mean & Variance, 9.15	4	0	100.00
146	P9r10	equation	equation	EOSC510	Mean & Variance, 11.06	1	0	100.00
147	P9r11	equation	equation	EOSC510	Mean & Variance, 11.43	1	0	100.00
148	P9r12	text	text	EOSC510	Mean & Variance, 13.11	4	0	100.00
149	P9r13	equation	equation	EOSC510	Mean and Variance, 14.28	1	0	100.00
150	P9r14	equation	equation	EOSC510	Mean & Variance, 16.15	1	0	100.00
151	P9r15	hybrid	text	EOSC510	Mean & Variance, 18.25	3	0	100.00
152	P9r16	equation	equation	EOSC510	Mean & Variance, 19.12	1	0	100.00
153	P9r17	text	text	EOSC510	Linear Regression, 0.08	4	0	100.00
154	P9r18	equation	equation	EOSC510	Linear Regression, 0.56	1	0	100.00

155	P9r19	equation	equation	EOSC510	Linear Regression, 2.20	1	0	100.00
156	P9r20	equation	equation	EOSC510	Linear Regression, 4.19	2	0	100.00
157	P9r21	equation	equation	EOSC510	Linear Regression, 5.05	2	0	100.00
158	P9r22	text	text	EOSC510	Geometric approach, 00:27	2	2	50.00
159	P9r23	text	text	EOSC510	Geometric approach, 6.41	7	0	100.00
160	P9r24	text	text	EOSC510	PCA applied to real data 00:22	1	0	100.00
161	P9r25	text	text	EOSC510	PCA applied to real data 00:26	1	0	100.00
162	P9r26	text	text	EOSC510	PCA applied to real data 00:54	4	0	100.00

163	P9r27	hybrid	graph	EOSC510	Rotated PCA 00:17	1	0	100.00
164	P9r28	hybrid	graph	EOSC510	Rotated PCA 00:29	1	0	100.00
165	P9r29	graph	graph	EOSC510	Rotated PCA 00:41	1	0	100.00
166	P9r30	graph	graph	EOSC510	Rotated PCA 2.44	1	0	100.00
167	P9r31	text	text	EOSC510	Rotated PCA 3.25	5	0	100.00
168	P10r1	equation	equation	EOSC510	Mean and Variance, 0.01	1	0	100.00
169	P10r2	equation	equation	EOSC510	Mean and Variance, 4.17	1	0	100.00
170	P10r3	equation	equation	EOSC510	Mean and Variance, 4.17	1	0	100.00
171	P10r4	equation	equation	EOSC510	Mean and Variance, 15.51	1	0	100.00
172	P10r5	equation	equation	EOSC510	MLP, 0.02	1	0	100.00
173	P10r6	text	text	EOSC510	MLP, 8.06	1	0	100.00

174	P10r7	text	text	EOSC510	PCA – Geometric approach 0:01	6	0	100.00
175	P10r8	text	text	EOSC510	PCA – Eigen vector approach 0:08	5	0	100.00
176	P10r9	equation	equation	EOSC510	PCA – Eigen vector approach 0:08	1	0	100.00
177	P10r10	equation	equation	EOSC510	Complex Data 7:45	1	0	100.00
178	P10r11	equation	equation	EOSC510	Complex Data 9:17	1	0	100.00
179	P10r12	hybrid	text	EOSC510	Scaling 9:35	2	0	100.00
180	P10r13	text	text	EOSC510	KNN classifier 2:28	2	2	50.00
181	P10r14	text	text	EOSC510	KNN classifier 4.36	3	3	50.00

APPENDIX B

Study Results

B.1 Accuracy-Overlap results

In this section, the results of the accuracy calculation are outlined in the table B.1. The columns show the note sample No. along with the type. 3 best matches are listed in the results column which can be compared with the ground truth observations. The level demonstrates the temporal level in which the results were retrieved. The percentage pf overlap for both the actual ground truth video and the results videos are listed.

Table B.1: Accuracy-Overlap Evaluation Results

Note	Note Type	Results	Ground Truth	Level	Actual Overlap (%)	Result Overlap (%)
P1r1	formula	{"name": "6- Conduction Materials", "timestamp": "00:03:48"}, {"name": "9- Magnetic Flux", "timestamp": "00:06:24"}, {"name": "11- Magnetic Materials", "timestamp": "00:05:22"}	Coloumb's law, 3.23	No match	40.00	NA
P1r2	text	{"name": "8- Biot-Savart Law", "timestamp": "00:04:07"}, {"name": "9- Magnetic Flux", "timestamp": "00:00:44"}, {"name": "3- Electric Flux", "timestamp": "00:03:28"}	Electric flux, 5.11	3rd, Section	50.00	33.33

P1r3	formula	{"name": "9- Magnetic Flux", "timestamp": "00:03:05"}, {"name": "4- Electric Dipole", "timestamp": "00:03:40"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}	Electric Dipole, 3.48	2nd, frame	18.18	18.18
P1r4	text	{"name": "8- Biot-Savart Law", "timestamp": "00:02:24"}, {"name": "10- Magnetic Force", "timestamp": "00:02:14"}, {"name": "12- Faraday's Law", "timestamp": "00:00:00"}	Biot-Savart law, 3.05	1st, Frame	77.78	77.78

P1r5	formula	{"name": "11-Magnetic Materials", "timestamp": "00:07:12"}, {"name": "9- Magnetic Flux", "timestamp": "00:00:20"}, {"name": "12- Faraday's Law", "timestamp": "00:05:45"}	Magnetic Force, 1.06	No match	36.36	NA
P1r6	text	{"name": "11-Magnetic Materials", "timestamp": "00:04:59"}, {"name": "12- Faraday's Law", "timestamp": "00:00:10"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:20"}	Magnetic materials, 4.54	1st, Frame	76.19	76.19

P1r7	formula	{"name": "9- Magnetic Flux", "timestamp": "00:02:55"}, {"name": "11-Magnetic Materials", "timestamp": "00:07:12"}, {"name": "8- Biot-Savart Law", "timestamp": "00:04:47"}	Magnetic Flux, 1.41	1st, section	33.33	66.67
P1r8	formula	{"name": "5- Current", "timestamp": "00:01:43"}, {"name": "7- Capacitance", "timestamp": "00:01:38"}, {"name": "10- Magnetic Force", "timestamp": "00:02:24"}	Current, 4.17	1st, video	22.22	11.11

P1r9	text	{"name": "10- Magnetic Force", "timestamp": "00:00:34"}, {"name": "2- Potential Energy", "timestamp": "00:00:50"}, {"name": "11- Magnetic Materials", "timestamp": "00:03:09"}	Unsure	Special case	.DIV/0!	.DIV/0!
P2r1	formula	{"name": "2- Potential Energy", "timestamp": "00:02:05"}, {"name": "13 - Magnetic Force", "timestamp": "00:00:20"}, {"name": "5- Conducting Materials", "timestamp": "00:02:05"}	Potential Energy, 1:43, 1:56, 2:09	1st, frame	33.33	33.33

P2r2	figure	{"name": "8- Electric Dipole", "timestamp": "01:10"}, {"name": "9- Capacitance", "timestamp": "00:35"}, {"name": "14 - Solenoids and Toroids", "timestamp": "09:30"}	Electric flux, 2.10	No match	75.00	NA
P2r3	formula	{"name": "4- Current", "timestamp": "00:01:43"}, {"name": "15 - Magnetic Materials", "timestamp": "00:03:09"}	Current, 3.05	1st, section	42.86	28.57

P2r4	figure	{"name": "2- Potential Energy", "timestamp": "01:45"}, {"name": "1 - Coulomb's Law", "timestamp": "02:20"}, {"name": "8- Biot-Savart Law", "timestamp": "00:50"}	Conducting materials, 1.12	No match	50.00	NA
P2r5	text	"name": "6- Classification of Electrical Networks", "timestamp": "00:01:20"}, {"name": "2- Potential Energy", "timestamp": "00:05:19"}, {"name": "9- Capacitance", "timestamp": "00:04:45"}	Classification of Electrical networks, 1.25	1st, frame	60.00	60.00

P2r6	figure	{"name": "14 - Solenoids and Toroids", "timestamp": "06:00"}, {"name": "1 - Coulomb's Law", "timestamp": "02:20"}, {"name": "9- Capacitance", "timestamp": "00:35"}	Electric Dipole, 1.20	No match	50.00	NA
P2r7	formula		Capacitance, 1.32	No data	100.00	NA
P2r8	figure	{"name": "2- Potential Energy", "timestamp": "01:45"}, {"name": "1 - Coulomb's Law", "timestamp": "02:15"}, {"name": "9- Capacitance", "timestamp": "00:35"}	Ideal op-amps, 7.40	No match	0.00	NA
P2r9	formula		Biot-Savart law, 4.05	No data	100.00	NA

P2r10	figure	{"name": "8- Biot-Savart Law", "timestamp": "01:50"}, {"name": "9- Capacitance", "timestamp": "00:35"}, {"name": "5- Conducting Materials", "timestamp": "04:15"}	Magnetic Flux, 2.01	No match	0.00	NA
P2r11	formula	{"name": "17 - Faraday's Law", "timestamp": "00:05:45"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:20"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:43"}	Magnetic Force, 5.01	No match	100.00	NA

P2r12	figure	{"name": "1 - Coulomb's Law", "timestamp": "02:25"}, {"name": "8- Biot-Savart Law", "timestamp": "02:15"}, {"name": "14 - Solenoids and Toroids", "timestamp": "06:00"}	Solenoids & Toroids, 7.57	3rd, section	100.00	0.00
P2r13	text	{"name": "15 - Magnetic Materials", "timestamp": "00:04:59"}, {"name": "14 - Solenoids and Toroids", "timestamp": "00:01:56" , {"name": "13 - Magnetic Force", "timestamp": "00:00:00"}	Magnetic materials, 4.51	1st, Frame	100.00	100.00

P2r14	formula	{"name": "17 - Faraday's Law", "timestamp": "00:07:35"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:20"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:43"}	Faraday's law, 6.47	1st, video	100.00	0.00
P2r15	text	{"name": "1 - Coulomb's Law", "timestamp": "00:02:02"}, {"name": "18 - Displacement Current \u00a9Carol Jaeger", "timestamp": "00:07:09"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:20"}	Coulomb's law, 1.51	1st, frame	100.00	100.00

P3r1	text	{"name": "Neural Networks - Prediction", "timestamp": "00:03:09"}, {"name": "K-Nearest Neighbors", "timestamp": "00:32:12"}, {"name": "PCA", "timestamp": "00:37:40"}	Neural Networks - Prediction, 1:27	1st, section	37.50	87.50
P3r2	text	{"name": "Convolutional Neural Networks", "timestamp": "02:40"}, {"name": "CS3: Design in Computing", "timestamp": "01:15"}, {"name": "Neural networks", "timestamp": "10:50"}	Neural Networks - Prediction, 4.25	No match	66.67	NA

P3r3	text	{"name": "Neural Networks - Prediction", "timestamp": "00:08:01"}, {"name": "PCA", "timestamp": "00:02:30"}, {"name": "Normal Equations", "timestamp": "00:28:40"}	Neural Networks - Prediction, 7.25	1st Section	75.00	75.00
P3r4	text	{"name": "Neural networks", "timestamp": "00:23:26"}, {"name": "K-Nearest Neighbors", "timestamp": "00:18:16"}, {"name": "CS3: Design in Computing", "timestamp": "00:38:54"}	Neural Networks - Training, 22.19	1st section	83.33	100.00

P3r5	text	{"name": "Fundamentals of Learning", "timestamp": "00:15:18"}, {"name": "Neural Networks - Prediction", "timestamp": "00:37:24"}, {"name": "K-Nearest Neighbors", "timestamp": "00:09:02"}	Neural Networks - Training, 22.57	No match	100.00	NA
P3r6	text	{"name": "Neural networks", "timestamp": "00:24:40"}, {"name": "Neural Networks - Prediction", "timestamp": "00:11:32"}, {"name": "Ordinary least squares", "timestamp": "00:20:27"}	Neural Networks - Training, 23.56	1st section	50.00	100.00

P3r7	text	{"name": "Convolutional Neural Networks", "timestamp": "00:02:28"}, {"name": "PCA - Training", "timestamp": "00:02:21"}, {"name": "Neural networks", "timestamp": "00:00:00"}	Convolutional networks, 2.04	1st Section	100.00	100.00
P3r8	formula		Convolutional networks, 2.31	No data	100.00	NA
P3r9	text	{"name": "K-Nearest Neighbors", "timestamp": "00:32:47"}, {"name": "Normal Equations", "timestamp": "00:08:17"}, {"name": "Ordinary least squares", "timestamp": "00:19:29"}	K-Nearest Neighbours, 32.27	1st frame	100.00	100.00

P3r10	text	{”name”: ”K-Nearest Neighbors”, ”timestamp”: ”00:37:32”}, {”name”: ”Exploratory Data Analysis”, ”timestamp”: ”00:38:39”}, {”name”: ”Normal Equations”, ”timestamp”: ”00:44:44”}	K-Nearest Neighbours, 36.50	1st frame	100.00	100.00
P3r11	text	{”name”: ”K-Nearest Neighbors”, ”timestamp”: ”00:39:06”}, {”name”: ”PCA - Training”, ”timestamp”: ”00:01:00”}, {”name”: ”Normal Equations”, ”timestamp”: ”00:35:58”}	K-Nearest Neighbours, 37.53	1st Section	50.00	100.00
P3r12	text		K-Nearest Neighbours, 40.37	No data	100.00	NA

P3r13	text	{"name": "Ordinary least squares", "timestamp": "00:00:00"}, {"name": "Normal Equations", "timestamp": "00:01:13"}, {"name": "K-Nearest Neighbors", "timestamp": "00:23:31"}	Ordinary Least Squares, 29.08	1st video	66.67	33.33
P3r14	text		Ordinary Least Squares, 33.08	No data	100.00	NA
P3r15	text	{"name": "K-Nearest Neighbors", "timestamp": "00:03:04"}, {"name": "Normal Equations", "timestamp": "00:20:55"}, {"name": "Neural networks", "timestamp": "00:06:30"}	Normal Equations, 3.31	1st video	100.00	25.00
P3r16	formula	[]	Normal Equations, 11.40	No data	100.00	NA

P3r17	text	{"name": "K-Nearest Neighbors", "timestamp": "00:09:02"}, {"name": "Exploratory Data Analysis", "timestamp": "00:22:03"}, {"name": "Fundamentals of Learning", "timestamp": "00:32:37"}	Exploratory Data Analysis, 0.35	2nd video	45.45	27.27
P3r18	text	{"name": "Exploratory Data Analysis", "timestamp": "00:05:16"}, {"name": "Regularization", "timestamp": "00:02:18"}, {"name": "Fundamentals of Learning", "timestamp": "00:47:27"}	Exploratory Data Analysis, 2.36	1st section	63.64	54.55

P3r19	text	{"name": "Ordinary least squares", "timestamp": "00:02:27"}, {"name": "Exploratory Data Analysis", "timestamp": "00:46:34"}, {"name": "K-Nearest Neighbors", "timestamp": "00:23:31"}	Exploratory Data Analysis, 5.45	2nd video	60.00	60.00
P4r1	figure	{"name": "17 - Faraday's Law", "timestamp": "00:00:00"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:50"}, {"name": "3- Electric Flux", "timestamp": "00:00:20"}	Coulomb's Law, 0:14	No match	33.33	NA

P4r2	formula	{"name": "1 - Coulomb's Law", "timestamp": "00:02:02"}, {"name": "5- Conducting Materials", "timestamp": "00:02:55"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:00"}	Coulomb's Law, 1:18	1st, section	16.67	50.00
P4r3	text	{"name": "2- Potential Energy", "timestamp": "00:00:50"}, {"name": "14 - Solenoids and Toroids", "timestamp": "00:01:56"}, {"name": "17 - Faraday's Law", "timestamp": "00:00:32"}	Potential Energy, 0:20	1st, section	30.00	40.00

P4r4	formula	{"name": "2- Potential Energy", "timestamp": "00:02:05"}, {"name": "15 - Magnetic Materials", "timestamp": "00:07:02"}, {"name": "17 - Faraday's Law", "timestamp": "00:06:05"}	Potential Energy, 1:35	1st, section	14.29	85.71
P4r5	text	{"name": "8- Biot-Savart Law", "timestamp": "00:02:43"}, {"name": "2- Potential Energy", "timestamp": "00:00:25"}, {"name": "13 - Magnetic Force", "timestamp": "00:00:00"}	Potential Energy, 2:19	2nd, video	75.00	50.00

P4r6	formula	{"name": "12- Magnetic Flux", "timestamp": "00:00:20"}, {"name": "4- Current", "timestamp": "00:01:43"}, {"name": "15 - Magnetic Materials", "timestamp": "00:03:19"}	Potential Energy, 3:04	No match	33.33	NA
P4r7	text	{"name": "6- Classification of Electrical Networks", "timestamp": "00:00:30"}, {"name": "17 - Faraday's Law", "timestamp": "00:04:06"}, {"name": "2- Potential Energy", "timestamp": "00:00:25"}	Classification of Electrical Networks, 0:27	1st, frame	90.00	90.00

P4r8	text	{”name”: ”8- Biot-Savart Law”, ”timestamp”: ”00:02:54”}, {”name”: ”6- Classification of Electrical Networks”, ”timestamp”: ”00:00:00”}, {”name”: ”14 - Solenoids and Toroids”, ”timestamp”: ”00:00:00”}	Classification of Electrical Networks, 0:44	2nd, video	33.33	0.00
P4r9	text	{”name”: ”6- Classification of Electrical Networks”, ”timestamp”: ”00:01:10”}, {”name”: ”8- Biot-Savart Law”, ”timestamp”: ”00:02:43”}, {”name”: ”13 - Magnetic Force”, ”timestamp”: ”00:00:00”}	Classification of Electrical Networks, 1:04	1st, Frame	50.00	50.00

P4r10	text	{"name": "6- Classification of Electrical Networks", "timestamp": "00:02:47"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:43"}, {"name": "3- Electric Flux", "timestamp": "00:05:58"}	Classification of Electrical Networks, 2:16	1st, Section	57.14	71.43
P4r11	text	{"name": "6- Classification of Electrical Networks", "timestamp": "00:03:17"}, {"name": "8- Biot-Savart Law", "timestamp": "00:06:38"}, {"name": "14 - Solenoids and Toroids", "timestamp": "00:00:00"}	Classification of Electrical Networks, 3:09	1st, Frame	50.00	50.00

P4r12	text	{"name": "6- Classification of Electrical Networks", "timestamp": "00:04:54"}, {"name": "10- Ideal OpAmps", "timestamp": "00:10:58"}, {"name": "15 - Magnetic Materials", "timestamp": "00:05:22"}	Classification of Electrical Networks, 4:05	1st, Section	40.00	80.00
P4r13	figure	{"name": "8- Biot-Savart Law", "timestamp": "02:15"}, {"name": "13 - Magnetic Force", "timestamp": "06:20"}, "name": "12- Magnetic Flux", "timestamp": "02:25"}	Capacitance, 0:33	No match	100.00	NA
P4r14	formula	[]	Capacitance, 1:13	No data	50.00	NA

P4r15	text	{"name": "18 - Displacement Current \u00a9Carol Jaeger", "timestamp": "00:09:20"}, {"name": "9- Capacitance", "timestamp": "00:02:45"}, {"name": "2- Potential Energy", "timestamp": "00:00:25"	Capacitance, 1:32	2nd, video	0.00	50.00
P4r16	text		Ideal OpAmps, 0:40	No data	50.00	NA
P4r35	text	{"name": "10- Ideal OpAmps", "timestamp": "00:01:40"}, {"name": "6- Classification of Electrical Networks", "timestamp": "00:01:37"}	Ideal OpAmps, 1:36	1st frame	100.00	100.00

P4r17	figure	{"name": "14 - Solenoids and Toroids", "timestamp": "06:00"}, {"name": "2- Potential Energy", "timestamp": "03:25"} {"name": "18 - Displacement Current \u00a9Carol Jaeger", "timestamp": "02:40"}	Ideal OpAmps, 1:18	No match	100.00	NA
P4r18	text	{"name": "10- Ideal OpAmps", "timestamp": "00:01:40"}, {"name": "17 - Faraday's Law", "timestamp": "00:00:00"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:00"}	Ideal OpAmps, 2:34	1st section	71.43	57.14

P4r19	text	{"name": "15 - Magnetic Materials", "timestamp": "00:03:09"}, {"name": "10- Ideal OpAmps", "timestamp": "00:03:36"}, {"name": "17 - Faraday's Law", "timestamp": "00:00:00"}	Ideal OpAmps, 5:01	2nd, video	33.33	66.67
P4r20	text	{"name": "10- Ideal OpAmps", "timestamp": "00:08:03"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:43"}, {"name": "14 - Solenoids and Toroids", "timestamp": "00:04:30"}	Ideal OpAmps, 7:25	1st section	70.00	90.00

P4r21	text	{"name": "12- Magnetic Flux", "timestamp": "00:00:20"}, {"name": "18 - Displacement Current \u00a9Carol Jaeger", "timestamp": "00:07:09"}, {"name": "15 - Magnetic Materials", "timestamp": "00:05:22"}	Magnetic Flux, 0:15	1st frame	83.33	83.33
P4r22	text	{"name": "12- Magnetic Flux", "timestamp": "00:01:00"}, {"name": "13 - Magnetic Force", "timestamp": "00:00:34"}, {"name": "17 - Faraday's Law", "timestamp": "00:00:32"}	Magnetic Flux, 0:39	1st section	25.00	75.00
P4r23	formula		Magnetic Flux, 1:32	No data	100.00	NA

P4r24	formula	{"name": "12- Magnetic Flux", "timestamp": "00:00:20"}	Magnetic Flux, 2:00	1st, video	16.67	33.33
P4r25	text	{"name": "12- Magnetic Flux", "timestamp": "00:03:45"}, {"name": "17 - Faraday's Law", "timestamp": "00:02:08"}, {"name": "3- Electric Flux", "timestamp": "00:01:48"}	Magnetic Flux, 3:06	1st Section	100.00	100.00
P4r26	text	{"name": "12- Magnetic Flux", "timestamp": "00:03:15"}, {"name": "17 - Faraday's Law", "timestamp": "00:03:38"}, {"name": "3- Electric Flux", "timestamp": "00:01:00"}	Magnetic Flux, 3:16	1st frame	33.33	33.33

P4r27	text	{"name": "18 - Displacement Current \u00a9Carol Jaeger", "timestamp": "00:06:28"}, {"name": "17 - Faraday's Law", "timestamp": "00:00:32"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:44"}	Magnetic Flux, 4:35	3rd, video	25.00	25.00
P4r28	text	{"name": "14 - Solenoids and Toroids", "timestamp": "00:00:00"}	Solenoids and Toroids, 0:26	1st frame	25.00	25.00

P4r29	text	{"name": "14 - Solenoids and Toroids", "timestamp": "00:05:30"}, {"name": "17 - Faraday's Law", "timestamp": "00:03:38"}, {"name": "15 - Magnetic Materials", "timestamp": "00:01:09"}	Solenoids and Toroids, 3:03	1st section	16.67	50.00
P4r30	text		Solenoids and Toroids, 6:14	No data	0.00	NA
P4r31	figure	{"name": "14 - Solenoids and Toroids", "timestamp": "06:00"}, {"name": "18 - Displacement Current \u00a9Carol Jaeger", "timestamp": "03:05"}, {"name": "5- Conducting Materials", "timestamp": "00:15"}	Solenoids and Toroids, 6:45	1st Section	50.00	0.00

P4r32	text	{"name": "14 - Solenoids and Toroids", "timestamp": "00:10:01"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:54"}, {"name": "17 - Faraday's Law", "timestamp": "00:04:06"}	Solenoids and Toroids, 8:30	1st section	0.00	50.00
P4r33	figure	{"name": "2- Potential Energy", "timestamp": "01:45"}, {"name": "8- Biot-Savart Law", "timestamp": "00:50"}, {"name": "14 - Solenoids and Toroids", "timestamp": "05:00"}	Solenoids and Toroids, 10:21	3rd video	100.00	50.00

P4r34	formula	{"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}, {"name": "12- Magnetic Flux", "timestamp": "00:00:20"}	Solenoids and Toroids, 13:09	No match	66.67	NA
P5r1	formula	{"name": "8- Biot-Savart Law", "timestamp": "00:06:07"}, {"name": "12- Faraday's Law", "timestamp": "00:07:30"}, {"name": "1 - Coulomb's Law", "timestamp": "00:01:52"}	Biot-Savart law, 2.20	1st video	100.00	100.00
P5r2	formula	{"name": "12- Faraday's Law", "timestamp": "00:04:06"}, {"name": "5- Current", "timestamp": "00:01:58"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}	Biot-Savart law, 3.29	3rd video	100.00	0.00

P5r3	formula	{"name": "8- Biot-Savart Law", "timestamp": "00:06:07"}, {"name": "12- Faraday's Law", "timestamp": "00:07:30"}, {"name": "5- Current", "timestamp": "00:01:58"}	Biot-Savart law, 5.12	1st section	100.00	0.00
P5r4	formula	{"name": "12- Faraday's Law", "timestamp": "00:07:30"}, {"name": "5- Current", "timestamp": "00:01:58"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:58"}	Biot-Savart law, 6.05	3rd video	100.00	0.00
P5r5	formula	{"name": "7- Capacitance", "timestamp": "00:01:38"}	Capacitance, 1.20	1st section	100.00	100.00
P5r6	formula	[]	Capacitance, 2.10	No data	100.00	NA
P5r7	formula	[]	Capacitance, 3.10	No data	100.00	NA

P5r8	formula	{"name": "7- Capacitance", "timestamp": "00:03:31"}	Capacitance, 3.42	1st section	100.00	100.00
P5r9	formula	{"name": "7- Capacitance", "timestamp": "00:01:38"}	Capacitance, 4.15	1st video	100.00	0.00
P5r10	formula	{"name": "7- Capacitance", "timestamp": "00:03:23"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:59"}, {"name": "9- Magnetic Flux", "timestamp": "00:00:15"}	Coulomb's law, 1.50	No match	100.00	NA

P5r11	formula	{"name": "12- Faraday's Law","timestamp": "00:07:30"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:43"}, {"name": "11- Magnetic Materials", "timestamp": "00:07:07"}	Magnetic Force, 3.53	No match	0.00	NA
P5r12	formula	{"name": "12- Faraday's Law", "timestamp": "00:04:21"}, {"name": "9- Magnetic Flux", "timestamp": "00:05:00"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}	Magnetic Flux, 1.35	2nd video	100.00	100.00

P5r13	formula	{"name": "9- Magnetic Flux", "timestamp": "00:05:00"}, {"name": "12- Faraday's Law", "timestamp": "00:04:06"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}	Magnetic Flux, 2.15	1st video	100.00	0.00
P5r14	formula	{"name": "9- Magnetic Flux", "timestamp": "00:03:05"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}, {"name": "5- Current", "timestamp": "00:00:43"}	Magnetic Flux, 2.51	1st frame	100.00	100.00

P5r15	formula	{"name": "8- Biot-Savart Law", "timestamp": "00:04:28"}, {"name": "1 - Coulomb's Law", "timestamp": "00:01:37"}, {"name": "12- Faraday's Law", "timestamp": "00:05:45"}	Magnetic Flux, 3.20	No match	100.00	NA
P5r16	formula		Magnetic Force, 0.40	No data	50.00	NA
P5r17	formula	{"name": "12- Faraday's Law", "timestamp": "00:06:00"}, "name": "11- Magnetic Materials", "timestamp": "00:07:07"}, {"name": "5- Current", "timestamp": "00:01:48"}	Magnetic Force, 2.43	No match	100.00	NA

P5r18	formula	{"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}, {"name": "5- Current", "timestamp": "00:00:43"}, {"name": "9- Magnetic Flux", "timestamp": "00:00:15"	Magnetic Force, 6.50	No match	0.00	NA
P5r19	formula	{"name": "12- Faraday's Law", "timestamp": "00:07:30"}, {"name": "11- Magnetic Materials", "timestamp": "00:07:07"}, {"name": "8- Biot-Savart Law", "timestamp": "00:02:43"}	Magnetic Materials, 7.14	2nd frame	100.00	100.00

P5r20	formula	{"name": "11-Magnetic Materials", "timestamp": "00:07:07"}, {"name": "8- Biot-Savart Law", "timestamp": "00:00:30"}, {"name": "5- Current", "timestamp": "00:01:48"}	Magnetic Materials, 7.40	1st section	100.00	0.00
P6r1	text	{"name": "React", "timestamp": "00:00:00"}	0.1	1st Section	100.00	66.67
P6r2	text		0.17	No data	0.00	NA
P6r3	text	{"name": "React", "timestamp": "00:01:00"}	0.47	1st section	100.00	80.00
P6r4	text	{"name": "React", "timestamp": "00:03:09"}	3.23	1st section	100.00	66.67
P6r5	figure	{"name": "React", "timestamp": "19:30"}	3.29	video	50.00	0.00

p6r6	text	{"name": "React", "timestamp": "00:03:42"}, {"name": "React", "timestamp": "00:01:00"}	3.29	1st section	40.00	80.00
P6r7	figure	{"name": "React", "timestamp": "3:10"}	5.07	1st video	100.00	0.00
P6r8	text	[]	6.38	No data	75.00	NA
P6r9	text	{"name": "React", "timestamp": "00:09:57"}, {"name": "React", "timestamp": "00:03:42"}, {"name": "React", "timestamp": "00:00:00"}	10.07	1st section	25.00	75.00
P6r10	text		11.34	No data	100.00	NA

P6r11	text	{"name": "React", "timestamp": "00:13:16"}, {"name": "React", "timestamp": "00:10:17"}, {"name": "React", "timestamp": "00:03:42"}	13.43	1st video	20.00	0.00
P6r12	text	{"name": "React", "timestamp": "00:15:02"}, {"name": "React", "timestamp": "00:00:00"}, {"name": "React", "timestamp": "01:75:35"}	16.35	1st section	66.67	66.67

P6r13	text	{"name": "React", "timestamp": "00:27:57"}, {"name": "React", "timestamp": "00:01:00"}, {"name": "React", "timestamp": "00:01:09"}	28.31	1st section	50.00	50.00
P6r14	text	{"name": "React", "timestamp": "00:29:48"}, {"name": "React", "timestamp": "00:02:29"}, {"name": "React", "timestamp": "00:01:00"}	39.3	1st section	75.00	75.00

P7r1	text	{"name": "CS3: Design in Computing", "timestamp": "00:12:14"}, {"name": "Fundamentals of Learning", "timestamp": "00:26:28"}, {"name": "Neural networks", "timestamp": "00:26:22"}	CS3 Design in Computing, 12.09	1st frame	66.67	66.67
P7r2	text	{"name": "CS3: Design in Computing", "timestamp": "00:13:14"}, {"name": "Neural networks", "timestamp": "00:30:12"}, {"name": "Neural Networks - Prediction", "timestamp": "00:37:04"}	CS3 Design in Computing, 12.40	1st section	50.00	50.00

P7r3	text	{"name": "K-Nearest Neighbors", "timestamp": "00:09:02"}, {"name": "PCA - Training", "timestamp": "00:15:46"}, {"name": "Fundamentals of Learning", "timestamp": "00:47:27"}	Fundamentals of learning, 35.36	3rd video	33.33	11.11
P7r4	text	{"name": "K-Nearest Neighbors", "timestamp": "00:24:46"}, {"name": "Fundamentals of Learning", "timestamp": "00:27:13"}, {"name": "CS3: Design in Computing", "timestamp": "00:03:24"}	Fundamentals of learning, 41.43	1st video	100.00	14.29

P7r5	text	{"name": "Fundamentals of Learning", "timestamp": "00:43:37"}, {"name": "K-Nearest Neighbors", "timestamp": "00:04:07"}, {"name": "Regularization", "timestamp": "00:16:20"}	Fundamentals of learning, 42.48	1st frame	100.00	100.00
P7r6	text	{"name": "Fundamentals of Learning", "timestamp": "00:49:06"}, {"name": "Neural networks", "timestamp": "00:27:24"}, {"name": "K-Nearest Neighbors", "timestamp": "00:21:38"}	k-nearest neighbors,6.37	3rd video	75.00	50.00

P7r7	text	{"name": "K-Nearest Neighbors", "timestamp": "00:22:11"}, {"name": "Fundamentals of Learning", "timestamp": "00:45:47"}, {"name": "Regularization", "timestamp": "00:25:10"}	k-nearest neighbors, 22.14	1st frame	77.78	77.78
P7r8	text	{"name": "K-Nearest Neighbors", "timestamp": "00:30:04"}, {"name": "Normal Equations", "timestamp": "00:35:58"}, {"name": "Fundamentals of Learning", "timestamp": "00:01:40"}	k-nearest neighbors, 29.49	1st frame	100.00	100.00

P7r9	text	{"name": "Regularization", "timestamp": "00:12:47"}, {"name": "PCA", "timestamp": "00:02:33"}, {"name": "Neural networks", "timestamp": "00:21:59"}	L2 Regularization, 12.52	1st frame	100.00	100.00
P7r10	formula	{"name": "Fundamentals of Learning", "timestamp": "00:14:09"}	L2 Regularization, 13.30	No match	100.00	NA
P7r11	figure	{"name": "Normal Equations", "timestamp": "39:40"}, {"name": "PCA", "timestamp": "15:20"}, {"name": "K-Nearest Neighbors", "timestamp": "09:00"}	PCA Intuition, 4.13	2nd video	100.00	0.00

P7r12	text	{"name": "PCA - Training", "timestamp": "00:28:35"}, {"name": "Fundamentals of Learning", "timestamp": "00:28:34"}, {"name": "Convolutional Neural Networks", "timestamp": "00:19:26"}	PCA training, 27.22	1st frame	78.57	78.57
P8r1	text		Mean & Variance, 9.03	No data	100.00	NA
P8r2	text	{"name": "Mean and Variance", "timestamp": "00:14:45"}, {"name": "PCA - Geometric approach", "timestamp": "00:00:07"}, {"name": "Multiple Linear Regression", "timestamp": "00:13:01"}	Mean & Variance, 13.11	1st frame	100.00	0.00

P8r3	text		Non-linear optimization, 0.02	No data	100.00	NA
P8r4	text	{"name": "Non-linear optimization", "timestamp": "00:04:38"}, {"name": "Multiple Linear Regression", "timestamp": "00:12:00"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:00:36"}	Non-linear optimization, 4.30	1st frame	100.00	100.00
P8r5	text	{"name": "Non-linear optimization", "timestamp": "00:11:43"}, {"name": "k-means clustering", "timestamp": "00:02:50"}, {"name": "Multiple Linear Regression", "timestamp": "00:08:03"}	Non-linear optimization, 13.00	1st frame	66.67	66.67

P8r6	formula	{"name": "Linear Regression", "timestamp": "00:00:00"}, {"name": "PCA - Geometric approach", "timestamp": "00:08:00"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:00:00"}	Non-linear optimization, 15.40	No match	40.00	NA
P8r7	text	{"name": "PCA - Eigenvector approach", "timestamp": "00:00:20"}, {"name": "Linear Regression", "timestamp": "00:00:20"}, {"name": "Multiple Linear Regression", "timestamp": "00:04:40"}	Eigen vector approach, 0.08	1st frame	66.67	66.67

P8r8	text	{"name": "KNN classifier", "timestamp": "00:00:20"}, {"name": "k-means clustering", "timestamp": "00:02:14"}, {"name": "Mean and Variance", "timestamp": "00:00:20"}	Classification: k nearest neighbouring classifier , 0.05	1st frame	100.00	100.00
P8r9	text	{"name": "KNN classifier", "timestamp": "00:09:48"}, {"name": "MSE and max likelihood", "timestamp": "00:04:05"}, {"name": "Mean and Variance", "timestamp": "00:02:37"}	Classification: k nearest neighbouring classifier, 9.36	1st frame	75.00	75.00

P8r10	text	{"name": "KNN classifier", "timestamp": "00:13:03"}, {"name": "Multiple Linear Regression", "timestamp": "00:09:43"}, {"name": "PCA - Geometric approach", "timestamp": "00:00:20"}	Classification: k nearest neighbouring classifier, 13.02	1st frame	100.00	100.00
P8r11	text		Mean & Variance, 4.17	No data	100.00	NA
P8r12	formula	{"name": "Mean and Variance", "timestamp": "00:09:49"}, {"name": "Linear Regression", "timestamp": "00:01:45"}, {"name": "Multiple Linear Regression", "timestamp": "00:08:29"}	Mean & Variance, 6.28	1st section	100.00	0.00

P9r1	text		Correlation & Regression, 0.34	No data	0.00	NA
P9r2	text		Correlation & Regression, 0.56	No data	0.00	NA
P9r3	formula	{"name": "Linear Regression", "timestamp": "00:04:48"}	Mean & Variance, 2.15	No match	100.00	NA
P9r4	formula	[]	Mean & Variance, 4.26	No data	100.00	NA
P9r5	formula	[]	Mean & Variance, 4.44	No data	100.00	NA
P9r6	formula	{"name": "Mean and Variance", "timestamp": "00:04:08"}, {"name": "Linear Regression", "timestamp": "00:04:42"}	Mean & Variance, 6.40	1st section	100.00	50.00
P9r7	formula	{"name": "Linear Regression", "timestamp": "00:04:48"}	Mean & Variance, 7.28	No match	100.00	NA

P9r8	text	{"name": "Mean and Variance", "timestamp": "00:07:37"}, {"name": "Multiple Linear Regression","timestamp": "00:08:24"}, {"name": "Linear Regression", "timestamp": "00:09:32"}	Mean & Variance, 8.39	1st frame	100.00	100.00
P9r9	formula	[]	Mean & Variance, 9.15	No data	100.00	NA
P9r10	formula	[]	Mean & Variance, 11.06	No data	100.00	NA
P9r11	formula	{"name": "MSE and max likelihood", "timestamp": "00:05:31"}, {"name": "Linear Regression", "timestamp": "00:04:48"}	Mean & Variance, 11.43	No match	100.00	NA

P9r12	text	{"name": "Mean and Variance", "timestamp": "00:09:17"}, {"name": "Linear Regression", "timestamp": "00:12:05"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:05:47"}	Mean & Variance, 13.11	1st section	100.00	75.00
P9r13	formula	{"name": "Linear Regression", "timestamp": "00:10:30"}, {"name": "MSE and max likelihood", "timestamp": "00:05:31"}	Mean and Variance, 14.28	No match	100.00	NA
P9r14	formula	[]	Mean & Variance, 16.15	No data	100.00	NA

P9r15	text	{"name": "Mean and Variance", "timestamp": "00:17:15"}, {"name": "PCA - Geometric approach", "timestamp": "00:11:52"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:22:23"}	Mean & Variance, 18.25	1st frame	100.00	100.00
P9r16	formula	[]	Mean & Variance, 19.12	No data	100.00	NA
P9r17	text	{"name": "Linear Regression", "timestamp": "00:00:20"}, {"name": "Multiple Linear Regression", "timestamp": "00:00:20"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:08:33"}	Linear Regression, 0.08	1st frame	100.00	100.00

P9r18	formula	[]	Linear Regression, 0.56	No data	100.00	NA
P9r19	formula	[]	Linear Regression, 2.20	No data	100.00	NA
P9r20	formula	{"name": "Mean and Variance", "timestamp": "00:04:07"}, {"name": "PCA - Geometric approach", "timestamp": "00:00:15"}	Linear Regression, 4.19	No match	100.00	NA
P9r21	formula	{"name": "Mean and Variance", "timestamp": "00:04:07"}, {"name": "Linear Regression", "timestamp": "00:04:48"}	Linear Regression, 5.05	2nd frame	100.00	100.00

P9r22	text	{"name": "PCA - Geometric approach", "timestamp": "00:00:20"}, {"name": "KNN classifier", "timestamp": "00:16:33"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:05:55"}	Geometric approach, 00:27	1st frame	50.00	50.00
P9r23	text	{"name": "Linear Regression", "timestamp": "00:00:20"}, {"name": "MSE and max likelihood", "timestamp": "00:04:05"}, {"name": "PCA - Geometric approach", "timestamp": "00:00:07"}	Geometric approach, 6.41	3rd video	100.00	42.86

P9r24	text		PCA applied to real data 00:22	No data	100.00	NA
P9r25	text		PCA applied to real data 00:26	No data	100.00	NA
P9r26	text	{"name": "PCA on real data", "timestamp": "00:01:40"}, {"name": "PCA - Geometric approach", "timestamp": "00:00:07"}, {"name": "Multiple Linear Regression", "timestamp": "00:10:10"}	PCA applied to real data 00:54	1st frame	100.00	100.00

P9r27	figure	{"name": "MSE and max likelihood", "timestamp": "02:20"}, {"name": "Non-linear optimization", "timestamp": "18:20"}, {"name": "k-means clustering", "timestamp": "05:00"}	Rotated PCA 00:17	No match	100.00	NA
P9r28	figure	{"name": "Rotated PCA", "timestamp": "00:02"}, {"name": "MSE and max likelihood", "timestamp": "02:40"}, {"name": "Non-linear optimization", "timestamp": "13:20"},	Rotated PCA 00:29	1st frame	100.00	100.00

P9r29	figure	{"name": "MSE and max likelihood", "timestamp": "01:00"}, {"name": "Non-linear optimization", "timestamp": "06:20"}, {"name": "k-means clustering", "timestamp": "05:00"}	Rotated PCA 00:41	No match	100.00	NA
P9r30	figure	{"name": "MSE and max likelihood", "timestamp": "05:00"}, {"name": "Non-linear optimization", "timestamp": "06:20"}, {"name": "KNN classifier", "timestamp": "01:00"}	Rotated PCA 2.44	No match	100.00	NA

P9r31	text	{"name": "Rotated PCA", "timestamp": "00:04:09"}, {"name": "k-means clustering", "timestamp": "00:08:26"}, {"name": "PCA - Geometric approach", "timestamp": "00:06:29"}	Rotated PCA 3.25	1st frame	100.00	100.00
P10r1	formula	{"name": "Linear Regression", "timestamp": "00:04:48"}	Mean and Variance, 0.01	No match	100.00	NA
P10r2	formula	[]	Mean and Variance, 4.17	No data	100.00	NA
P10r3	formula	[]	Mean and Variance, 4.17	No data	100.00	NA

P10r4	formula	{"name": "Complex Data", "timestamp": "00:00:48"}, {"name": "Linear Regression", "timestamp": "00:04:48"}	Mean and Variance, 15.51	No match	100.00	NA
P10r5	formula	[]	MLP, 0.02	No data	100.00	NA
P10r6	text		MLP, 8.06	No data	100.00	NA
P10r7	text	{"name": "PCA - Geometric approach", "timestamp": "00:00:10"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:05:55"}, {"name": "KNN classifier", "timestamp": "00:16:33"}	PCA – Geometric approach 0:01	1st frame	100.00	100.00

P10r8	text	{”name”: ”PCA - Eigenvector approach”, ”timestamp”: ”00:00:10”}, {”name”: ”Non-linear optimization”, ”timestamp”: ”00:00:10”}, {”name”: ”Linear Regression”, ”timestamp”: ”00:00:10”}	PCA – Eigen vector approach 0:08	1st frame	100.00	100.00
P10r9	formula	[]	PCA – Eigen vector approach 0:08	No data	100.00	NA
P10r10	formula	{”name”: ”Linear Regression”, ”timestamp”: ”00:04:48”}, {”name”: ”Complex Data”, ”timestamp”: ”00:00:48”}	Complex Data 7:45	2nd video	100.00	0.00

P10r11	formula	[]	Complex Data 9:17	No data	100.00	NA
P10r12	text		Scaling 9:35	No data	100.00	NA
P10r13	text	{"name": "MSE and max likelihood", "timestamp": "00:03:28"}, {"name": "KNN classifier", "timestamp": "00:03:20"}, {"name": "Multiple Linear Regression", "timestamp": "00:10:10"}	KNN classifier 2:28	2nd frame	50.00	50.00

P10r14	text	{"name": "KNN classifier", "timestamp": "00:13:03"}, {"name": "Multiple Linear Regression", "timestamp": "00:08:24"}, {"name": "PCA - Eigenvector approach", "timestamp": "00:15:28"}	KNN classifier 4.36	1st video	50.00	100.00

B.2 Friedman & Wilcoxon test results

The friedman and wilcoxon test results for the preference levels in temporal conditions and video surrogates is shown in the figures B.1, B.2, B.3, B.4, B.5, and B.6.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
point	12	4.5000	.79772	3.00	5.00	4.0000	5.0000	5.0000
interval	12	4.0833	.99620	2.00	5.00	3.2500	4.0000	5.0000
video	12	3.5833	1.08362	1.00	5.00	3.0000	4.0000	4.0000

Ranks

		N	Mean Rank	Sum of Ranks
interval - point	Negative Ranks	5[a]	3.60	18.00
	Positive Ranks	1[b]	3.00	3.00
	Ties	6[c]		
	Total	12		
video - interval	Negative Ranks	6[d]	3.67	22.00
	Positive Ranks	1[e]	6.00	6.00
	Ties	5[f]		
	Total	12		
point - video	Negative Ranks	1[g]	3.00	3.00
	Positive Ranks	7[h]	4.71	33.00
	Ties	4[i]		
	Total	12		

a. interval < point
b. interval > point
c. interval = point
d. video < interval
e. video > interval
f. video = interval
g. point < video
h. point > video
i. point = video

Test Statistics[a]

N	12
Chi-Square	7.400
df	2
Asymp. Sig.	.025
Exact Sig.	.021
Point Probability	.001

a. Friedman Test

Test Statistics[a]

	interval - point	video - interval	point - video
Z	-1.667[b]	-1.387[b]	-2.157[c]
Asymp. Sig. (2-tailed)	.096	.165	.031
Exact Sig. (2-tailed)	.188	.219	.047
Exact Sig. (1-tailed)	.094	.109	.023
Point Probability	.078	.023	.020

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Figure B.1: Temporal preference: a) It is easy to find the information I need from this point

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	50th (Median)	75th
point	12	4.5833	.66856	3.00	5.00	4.0000	5.0000	5.0000
interval	12	4.1667	1.02986	2.00	5.00	3.2500	4.5000	5.0000
video	12	3.8333	.83485	2.00	5.00	3.2500	4.0000	4.0000

Ranks

		N	Mean Rank	Sum of Ranks
interval - point	Negative Ranks	5[a]	3.20	16.00
	Positive Ranks	1[b]	5.00	5.00
	Ties	6[c]		
	Total	12		
video - interval	Negative Ranks	5[d]	3.10	15.50
	Positive Ranks	1[e]	5.50	5.50
	Ties	6[f]		
	Total	12		
point - video	Negative Ranks	1[g]	3.50	3.50
	Positive Ranks	7[h]	4.64	32.50
	Ties	4[i]		
	Total	12		

a. interval < point

b. interval > point

c. interval = point

d. video < interval

e. video > interval

f. video = interval

g. point < video

h. point > video

i. point = video

Test Statistics[a]

N	12
Chi-Square	6.897
df	2
Asymp. Sig.	.032
Exact Sig.	.029
Point Probability	.001

a. Friedman Test

Test Statistics[a]

	interval - point	video - interval	point - video
Z	-1.179[b]	-1.081[b]	-2.124[c]
Asymp. Sig. (2-tailed)	.238	.279	.034
Exact Sig. (2-tailed)	.313	.406	.055
Exact Sig. (1-tailed)	.156	.203	.027
Point Probability	.047	.031	.023

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on negative ranks.

Figure B.2: Temporal preference: b) The information is effective in helping me complete the tasks and scenarios for learning

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	50th (Median)	75th
point	12	3.9167	1.08362	2.00	5.00	3.0000	4.0000	5.0000
interval	12	4.0833	.99620	2.00	5.00	3.2500	4.0000	5.0000
video	12	3.0000	.85280	2.00	4.00	2.0000	3.0000	4.0000

Ranks

		N	Mean Rank	Sum of Ranks
interval - point	Negative Ranks	4[a]	3.88	15.50
	Positive Ranks	4[b]	5.13	20.50
	Ties	4[c]		
	Total	12		
video - interval	Negative Ranks	10[d]	5.75	57.50
	Positive Ranks	1[e]	8.50	8.50
	Ties	1[f]		
	Total	12		
point - video	Negative Ranks	2[g]	6.25	12.50
	Positive Ranks	9[h]	5.94	53.50
	Ties	1[i]		
	Total	12		

a. interval < point
b. interval > point
c. interval = point
d. video < interval
e. video > interval
f. video = interval
g. point < video
h. point > video
i. point = video

Test Statistics[a]

N	12
Chi-Square	9.190
df	2
Asymp. Sig.	.010
Exact Sig.	.008
Point Probability	.001

a. Friedman Test

Test Statistics[a]

	interval - point	video - interval	point - video
Z	-.359[b]	-2.240[c]	-1.877[b]
Asymp. Sig. (2-tailed)	.719	.025	.060
Exact Sig. (2-tailed)	.805	.021	.080
Exact Sig. (1-tailed)	.402	.011	.040
Point Probability	.078	.003	.007

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.
c. Based on positive ranks.

Figure B.3: Temporal preference: c) I'm satisfied with the retrieved video time point

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Title+Thumbnail	12	3.4167	1.16450	1.00	5.00	2.2500	4.0000	4.0000
Keywords	12	4.3333	.65134	3.00	5.00	4.0000	4.0000	5.0000
Summary	12	4.0000	.95346	3.00	5.00	3.0000	4.0000	5.0000

Ranks

		N	Mean Rank	Sum of Ranks
Keywords - Title+Thumbnail	Negative Ranks	1[a]	2.00	2.00
	Positive Ranks	6[b]	4.33	26.00
	Ties	5[c]		
	Total	12		
Summary - Keywords	Negative Ranks	6[d]	5.25	31.50
	Positive Ranks	3[e]	4.50	13.50
	Ties	3[f]		
	Total	12		
Title+Thumbnail - Summary	Negative Ranks	5[g]	6.60	33.00
	Positive Ranks	4[h]	3.00	12.00
	Ties	3[i]		
	Total	12		

a. Keywords < Title+Thumbnail
b. Keywords > Title+Thumbnail
c. Keywords = Title+Thumbnail
d. Summary < Keywords
e. Summary > Keywords
f. Summary = Keywords
g. Title+Thumbnail < Summary
h. Title+Thumbnail > Summary
i. Title+Thumbnail = Summary

Test Statistics[a]

N	12
Chi-Square	2.889
df	2
Asymp. Sig.	.236
Exact Sig.	.244
Point Probability	.026

a. Friedman Test

Test Statistics[a]

	Keywords - Title+Thumbnail	Summary - Keywords	Title+Thumbnail - Summary
Z	-2.050[b]	-1.155[c]	-1.269[c]
Asymp. Sig. (2-tailed)	.040	.248	.205
Exact Sig. (2-tailed)	.063	.398	.215
Exact Sig. (1-tailed)	.031	.199	.107
Point Probability	.023	.125	.010

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.
c. Based on positive ranks.

Figure B.4: Temporal preference: a) It is easy to find the information I need from this point

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	50th (Median)	75th
Title+Thumbnail	12	3.7500	.96531	2.00	5.00	3.0000	4.0000	4.7500
Keywords	12	4.3333	.65134	3.00	5.00	4.0000	4.0000	5.0000
Summary	12	3.9167	.99620	3.00	5.00	3.0000	3.5000	5.0000

Ranks

		N	Mean Rank	Sum of Ranks
Keywords - Title+Thumbnail	Negative Ranks	2[a]	3.00	6.00
	Positive Ranks	6[b]	5.00	30.00
	Ties	4[c]		
	Total	12		
Summary - Keywords	Negative Ranks	6[d]	4.67	28.00
	Positive Ranks	2[e]	4.00	8.00
	Ties	4[f]		
	Total	12		
Title+Thumbnail - Summary	Negative Ranks	5[g]	5.20	26.00
	Positive Ranks	4[h]	4.75	19.00
	Ties	3[i]		
	Total	12		

a. Keywords < Title+Thumbnail
b. Keywords > Title+Thumbnail
c. Keywords = Title+Thumbnail
d. Summary < Keywords
e. Summary > Keywords
f. Summary = Keywords
g. Title+Thumbnail < Summary
h. Title+Thumbnail > Summary
i. Title+Thumbnail = Summary

Test Statistics[a]

N	12
Chi-Square	2.800
df	2
Asymp. Sig.	.247
Exact Sig.	.272
Point Probability	.037

a. Friedman Test

Test Statistics[a]

	Keywords - Title+Thumbnail	Summary - Keywords	Title+Thumbnail - Summary
Z	-1.732[b]	-1.508[c]	-.426[c]
Asymp. Sig. (2-tailed)	.083	.132	.670
Exact Sig. (2-tailed)	.125	.234	.805
Exact Sig. (1-tailed)	.063	.117	.402
Point Probability	.039	.086	.117

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.
c. Based on positive ranks.

Figure B.5: Preference for Video Surrogates: b) This design has all the functions and capabilities I expect it to have

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Title+Thumbnail	12	4.3333	.88763	2.00	5.00	4.0000	4.5000	5.0000
Keywords	12	4.2500	.96531	2.00	5.00	4.0000	4.5000	5.0000
Summary	12	3.2500	1.21543	1.00	5.00	2.2500	3.0000	4.0000

Ranks

		N	Mean Rank	Sum of Ranks
Keywords - Title+Thumbnail	Negative Ranks	3[a]	5.00	15.00
	Positive Ranks	4[b]	3.25	13.00
	Ties	5[c]		
	Total	12		
Summary - Keywords	Negative Ranks	9[d]	5.00	45.00
	Positive Ranks	0[e]	.00	.00
	Ties	3[f]		
	Total	12		
Title+Thumbnail - Summary	Negative Ranks	2[g]	3.00	6.00
	Positive Ranks	7[h]	5.57	39.00
	Ties	3[i]		
	Total	12		

a. Keywords < Title+Thumbnail
b. Keywords > Title+Thumbnail
c. Keywords = Title+Thumbnail
d. Summary < Keywords
e. Summary > Keywords
f. Summary = Keywords
g. Title+Thumbnail < Summary
h. Title+Thumbnail > Summary
i. Title+Thumbnail = Summary

Test Statistics[a]

N	12
Chi-Square	8.914
df	2
Asymp. Sig.	.012
Exact Sig.	.008
Point Probability	.001

a. Friedman Test

Test Statistics[a]

	Keywords - Title+Thumbnail	Summary - Keywords	Title+Thumbnail - Summary
Z	-.172[b]	-2.762[b]	-1.992[c]
Asymp. Sig. (2-tailed)	.863	.006	.046
Exact Sig. (2-tailed)	.938	.004	.066
Exact Sig. (1-tailed)	.469	.002	.033
Point Probability	.063	.002	.021

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Figure B.6: Preference for Video Surrogates: c) The organization of information on the layout screen is clear

B.3 Interview Transcripts

The interview transcripts of the user study with 12 participants can be accessed here[1].

B.4 Survey results

This section shows the survey data for the preference of temporal level and video surrogates in the retrieved video results as can be seen in the tables B.2, B.3. Later, the usability data in terms of Effectiveness, Efficiency and Satisfaction is also covered in the section in the tables. The results are derived from the 12-student experimental study discussed in chapter 4. The likert-scale data in all the results range from 1-Strongly disagree to 5-Strongly agree.

[1] https://ubcca-my.sharepoint.com/:f:/g/personal/ranjs92_student_ubc_ca/EtvgSRjJKktPqKHcXUxUauIBynZU9R26ePuSuwC-5YvZpA?e=Zl0y1u

Table B.2: Likert-scale data for temporal preference in Point, Interval and whole video

		Point			Interval			Video		
Participant	**Name**	**a**	**b**	**c**	**a**	**b**	**c**	**a**	**b**	**c**
P1	Fan	5	5	5	4	4	5	4	4	3
P2	Qian	3	4	4	3	3	3	3	3	3
P3	Rui	5	5	5	5	5	5	3	4	4
P4	Mazoud	3	4	3	2	2	2	4	4	4
P5	Praneeth	5	5	5	5	5	5	5	5	4
P6	Taslim	4	5	4	4	4	4	3	4	3
P7	Pramit	5	5	3	4	5	4	3	3	2
P8	Deb	5	3	2	5	5	5	5	4	4
P9	Achinth	5	5	3	5	5	4	4	5	2
P10	Shaunak	5	4	5	4	4	3	4	4	2
P11	Munira	5	5	5	3	3	4	1	2	2
P12	Lakshmi	4	5	3	5	5	5	4	4	3

- a - It is easy to find the information I need from this point
- b - The information is effective in helping me complete the tasks and scenarios for learning
- c - I'm satisfied with the retrieved video time point

Table B.3: Likert-scale data for video surrogate preference

		Title + Thumbnail			Title + Thumbnail + Keywords			Title + Thumbnail + Keywords + Summary		
Participant	**Name**	**a**	**b**	**c**	**a**	**b**	**c**	**a**	**b**	**c**
P1	Fan	4	5	5	4	4	3	3	3	2
P2	Qian	4	4	4	4	4	4	3	3	4
P3	Rui	4	4	5	5	5	5	4	4	3
P4	Mazoud	4	4	4	5	5	5	3	3	3
P5	Praneeth	5	5	4	5	5	5	5	5	4
P6	Taslim	4	3	4	3	3	4	4	3	3
P7	Pramit	4	3	4	4	5	5	5	5	5
P8	Deb	3	4	5	5	4	5	5	5	5
P9	Achinth	2	3	5	4	4	4	5	5	2
P10	Shaunak	4	5	5	4	4	2	3	3	1
P11	Munira	2	3	5	5	5	5	5	5	4
P12	Lakshmi	1	2	2	4	4	4	3	3	3

- a - It is easy to find the information I need from this point
- b - This design has all the functions and capabilities I expect it to have.
- c - The organization of information on the layout screen is clear

Table B.4: Likert-scale survey results for Effectiveness in Usability Evaluation

Effectiveness		
Easy to use	**Functions well integrated**	**Need technical support**
5	4	2
5	5	1
4	4	1
4	4	3
4	4	1
4	5	1
4	3	4
5	4	2
4	4	1
5	4	1
4	3	2
5	4	1

Table B.5: Likert-scale survey results for Efficiency in Usability Evaluation

Efficiency		
Quick producitivity	**Quick Learning**	**Lot to learn**
5	4	3
5	5	1
3	5	2
4	4	1
5	3	4
5	5	1
4	4	4
4	4	2
4	4	1
5	5	2
4	4	2
4	5	1

Table B.6: Likert-scale survey results for Satisfaction in Usability Evaluation

Satisfaction		
Frequent use	**Overall satisfaction**	**Complexity**
5	4	2
5	5	5
4	4	2
3	4	2
4	5	2
5	5	1
4	4	2
5	4	2
4	4	2
4	5	2
4	4	2
5	4	2

APPENDIX C

Implementation Details

The implementation details of the note recognition of each type that is text, formula, and figure and the mobile application interface design can be accessed here [1].

[1] https://learning.github.ubc.ca/ranjs92/NoteLink.git