Automatic Question Generation from Handwritten Lecture Notes on KeyBERT-indexed T5-TrOCR Pipeline with Gemini Context Correction*

Rommel John Ronduen[1], Jan Adrian Manzanero[2], Analyn Yumang[3]
School of Electrical, Electronics, and Computer Engineering
Mapua University
Manila, Philippines
[1]rjhronduen@mymail.mapua.edu.ph
[2]jacmanzanero@mymail.mapua.edu.ph
[3]anyumang@mapua.edu.ph

***Abstract*—In this research, a system capable of performing optical character recognition (OCR) on handwritten lecture notes and consequently using the extracted text for automatic question genertion (AQG) was conceived. The system utilized the Text-to-Text Transformer (T5) for AQG and the Transformer-based Optical Character Recognition model (TrOCR) for OCR. The base version T5 was fine-tuned using the SQuADv1.1 dataset while the pre-trained handwritten base version for TrOCR was used. The system was evaluated using the word error rate (WER) for OCR evaluation, while the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) and Bilingual Language Understanding Evaluation (BLEU) were used for AQG evaluation. The system achieved a WER of 0.40 and a question validity rate of 70%.**

***Index Terms*—Large Language Model, Optical Character Recognition, Automatic Question Generation, Handwritten Lecture Notes, Raspberry Pi**

I. Introduction

Handwritten lecture notes are still widely used in educational institutions. Since the advent of artifical intelligence (AI), education has become the primary focus of AI research. AI has been used to automate several processes, and one which is the generation of learning materials such as questions in the form of quizzes. Literature defines automatic question generation (AQG) as the process of generating questions from a given text.

Despite existing implementations of AQG, no attempt has been given for utilizing handwritten lecture notes as a source for AQG. There have been several approaches to AQG, such as rule-based, template-based, and neural-based, but none of these approaches have been applied to handwritten lecture notes as a direct source for AQG through image capturing and OCR. It is noted by (Arbaaeen) AQG is defined by the methods used to generate questions. The AQG system is composed of three main components: the context extraction and parsing, the question generation, and the question validation. AQG systems differ in the method of extracting information and also the type of information that is being extracted. In fact, it was noted by them that it is suggested that AQG implementations utilize other forms of media sources for context bases. For instance, (Gaur) conceived of an AQG system that utilizes programming source code as the context for question generation based on selected keywords from the code snippets. However, these code snippets were assumed to be transcribed as an input to the system and not from the handwritten medium. Another use case through (Ou) involved the use of recorded videos as a source for AQG that led to question-answer pairs through the Bidirection and Auto-Regressive Transformer (BART) model for the search of sentences in the text. They utilized an algorithm of indexing or searching through the extracted context for specific question generations, which will be explained further in the succeeding text. Despite this, the researchers utilized a recorded video as a source. On the other hand, (Moron) allowed for an implementation of AQG through the use of named entity recognition (NER) and semantic rules for question generation in aiding the learning of English that led to the generation of questions that primarily answered the questions of ”what” and ”who”. However, the AQG implementation relied on the manual input of the user in text boxes for the context. In total, AQG systems are capable of generating questions from a given context, but no system has been developed to generate questions from handwritten lecture notes. There exists a variety in the algorithms used for AQG, such as the use of the BART model for question generation through the use of a search algorithm and the use of NER and semantic rules for question generation. An example includes the use of treebanks for providing for inferences (Padilla). However, one such study by (Tsai) utilized the use of the Text-to-Text Transformer (T5) model for AQG and the evaluation utilizing the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) and the Bilingual Evaluation Understudy (BLEU) metrics. The Stanford Question Answering Dataset (SQuAD) was used for the fine-tuning of the T5 model that led to satisfactory similarity of model-generated questions to the ground truth questions with a ROUGE-L score of 0.613. However, they noted a limitation as the model had a BLEU score of 0.567 due to the processes of the T5 model mixing

the syntax and form of the context used for AQG. Due to the capability of the T5 model to generate questions from a given context while being fine-tuned to a specific dataset, it was chosen as the model for AQG in this research. In bridging the gap for AQG from handwritten lecture notes, the Transformer-based Optical Character Recognition (TrOCR) model was used for the optical character recognition (OCR) of the handwritten lecture notes. Starting first with OCR, it is often associated with convolutional neural networks (CNN) and recurrent neural networks (RNN). As noted by (Manlises) the use of CNNs for OCR is often used for the detection of text in images and even objects such as that of the implementation for the detection of the different types of mushrooms (Caya). CNNs have also been used for recognizing text through the transformation of shorthand terminologies to English text (Vitug). In fact, it was shown by (Ligsay) that it was possible to recognize text in Baybayin (a Filipino lettering system) using CNNs. Another involved the identification of expiry dates on canned goods (Manlises). In terms of evaluation, these CNNs are often evaluated using confusion matrices. These confusion matrices are used to evaluate the performance of the OCR model in terms of the true positive, true negative, false positive, and false negative values (Villaverde). However, a transformer-based approach has been used for OCR, such as the TrOCR model (Li) which has been used in recognition of text from scanned receipts (Zhang) and even in the recognition of text from images of Arabic text (Mortadi). The model has been exemplary over the use of CNNs by its encoder-decoder framework with pre-trained weights for the recognition of text.

In lieu of the gap on AQG from utilizing handwritten lecture notes, the general objective of this research to develop a system of allowing for automatic question generation from handwritten lecture notes on KeyBERT-indexed T5-TrOCR pipeline with Gemini context correction. The specific objectives of this research to utilize a fine-tuned T5 model for AQG on SQuADv1.1 while using KeyBERT for indexing the context of the handwritten lecture notes; to utilize the base handwritten TrOCR model for the OCR of the handwritten lecture; to evaluate the system using the word error rate (WER) for OCR evaluation and the ROUGE and BLEU metrics for AQG evaluation; and to utilize a Raspberry Pi 5 within a constructed enclosure with proper illumination for the facilitation and procurement of the system processes, a web camera for image capture, and a touchscreen monitor for the display of the generated questions.

This research is mainly limited by the consideration of only single-column, diagram and equation free, and English handwritten lecture notes of no erasures on strictly letter-sized plain sheet paper. Moreover, the system is confined to the use of the T5-TrOCR pipeline utilizing the base versions while also utilizing the SQuADv1.1 dataset for the fine-tuning. It is also important to note that the system is limited to the facilitation of the processes in the hardware of the 8-gigabyte version of the Raspberry Pi 5. The use of a higher parameter count for the T5 and TrOCR models and the use of a different dataset for fine-tuning may allow for differing results. Moreover, the utilization of a different hardware setup may also lead to alternative, if not, better results in terms of processing time.

II. Materials and Methods

A. Hardware Development

1) System Block Diagram:

2) Experimental Setup:

B. Software Development

1) System Flowchart:

2) Model Fine-tuning:

C. Data Gathering

D. Testing and Evaluation

III. Results and Discussion

IV. Conclusion and Recommendations

A. Maintaining the Integrity of the Specifications

The IEEEtran class file is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

V. Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections V-A–V-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads—LaTeX will do that for you.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.

- Do not mix complete spellings and abbreviations of units: "Wb/m^2" or "webers per square meter", not "webers/m^2". Spell out units when they appear in text: ". . . a few henries", not ". . . a few H".
- Use a zero before decimal points: "0.25", not ".25". Use "cm^3", not "cc".)

C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \tag{1}$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not "Eq. (1)" or "equation (1)", except at the beginning of a sentence: "Equation (1) is . . ."

D. LaTeX-Specific Advice

Please use "soft" (e.g., `\eqref{Eq}`) cross references instead of "hard" references (e.g., `(1)`). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don't use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in LaTeX will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you've discovered a new method of counting.

BibTeX does not work by magic. It doesn't get the bibliographic data from thin air but from .bib files. If you use BibTeX to produce a bibliography you must send the .bib files.

LaTeX can't read your mind. If you assign the same label to a subsubsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

LaTeX does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it's supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won't be any anyway) and it might stop a wanted equation number in the surrounding equation.

E. Some Common Mistakes

- The word "data" is plural, not singular.
- The subscript for the permeability of vacuum μ_0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter "o".
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an "inset", not an "insert". The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).
- Do not use the word "essentially" to mean "approximately" or "effectively".
- In your paper title, if the words "that uses" can accurately replace the word "using", capitalize the "u"; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones "affect" and "effect", "complement" and "compliment", "discreet" and "discrete", "principal" and "principle".
- Do not confuse "imply" and "infer".
- The prefix "non" is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the "et" in the Latin abbreviation "et al.".
- The abbreviation "i.e." means "that is", and the abbreviation "e.g." means "for example".

An excellent style manual for science writers is [7].

F. Authors and Affiliations

The class file is designed for, but not limited to, six authors. A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

G. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is "Heading 5". Use "figure caption" for your Figure captions, and "table head" for your table title. Run-in heads, such as "Abstract", will require you

to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced.

H. Figures and Tables

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation "Fig. 1", even at the beginning of a sentence.

TABLE I
TABLE TYPE STYLES

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy[a]		

[a]Sample of a Table footnote.

Fig. 1. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

Please number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.
[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
[4] K. Elissa, "Title of paper if known," unpublished.
[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.
[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.