Readability of Non-Text Images on The World Wide Web (WWW)

Jaisurya S
Dept of MCA (Master of Computer Applications),
Jayam college of engineering and technology, Anna University, Dharmapuri.

Abstract- *The World Wide Web associated the world in a manner that was unrealistic previously and made it a lot more straightforward for users to get data, share and impart. But, irrelevant non-text images on the web pages equally specify poor readability, disrupting the people from the emphasis of the reading. The main goal of this paper is to evaluate the impact of irrelevant or low-quality non-text images on the readability of the webpage. An automatic methodology has been proposed to compute the relevancy of the non-text images. This methodology merges different approaches to extract information from non-text images and read text from websites in order to find relevancy between them. This technique was used to analyze fifty different educational websites to automatically find the relevancy of their non-text images. A user study has been carried out to evaluate the proposed methodology with different types of questions. The results agree with the fact that the relevant non-text images enhance the readability of the web page. This research work will help web designers to improve readability by considering only the relevant content of a web page, without relying on expert judgment.*

Keywords- Web page, non-text images, readability, relevancy, evaluation.

I. INTRODUCTION

With the increase in the growth of the internet since 1990, the World Wide Web has gained worldwide popularity. The web has become an ultimate source of information worldwide [1], [2], [3], [4]. Readability is the straightforwardness with which a user can perceive passages, sentences, and words [5], [6], [7], [8]. In this paper, we will highlight the non-text images of websites and how they play an important role in the readability of web pages. Generally, non-text images are more successful than simply the composed text in web readability because our brain can decipher graphical contents much faster than text, which is why images can impart an item, administration, or brand in a flash [9], [10], [11], [12]. Furthermore, non-text images give profundity and setting to a depiction or story and a significantly more vivid experience than text alone. It's the reason why websites need good and relevant graphical content [13], [14], [15].

The associate editor coordinating the review of this manuscript and approving it for publication was Claudio Zunino.

Not all non-text images are appropriate for enhancing the attractiveness and readability of the text they accompany, such as decorative non-text images or graphical contents not relevant for the webpage content itself. Furthermore, different factors could affect the readability of the non-text images on web pages. For example, poor resolution of graphical content, wrong aspect ratio, the improper color combination of graph- ical content itself, etc., and the World Wide Web Consor- tium (W3C) suggested dissimilar recommendations for these problems [16], [17], [18]. These recommendations suggest low contrast, alternate text, proper color combination, and enhanced resolution. However, one of the most basic issues is that the irrelevancy of non-text images with the text of the website could badly affect web readability. The research workers pondered only on the textual contents of the websites while evaluating the readability of web pages and suggested different assessment tools for this. They, however, did not work on the non-text image relevancy evaluation of the web pages [19], [20]. In this paper, a new methodology has been proposed that computes the relevance of non-text images on websites and hypothesized that relevant non-text images could increase web readability. A user survey was performed to validate the hypothesis.

The paper is structured as follows: Section 2 defines a literature review recognizing a research gap. Section 3 dis- cusses the hypothesis; Section 4 explains the proposed research methodology to compute non-text image relevancy. Section 5 describes the evaluation process. Section 6 dis- cusses evaluation results and Section 7 finishes the paper with the important findings and future work.

II. LITERATURE REVIEW

Most researchers have focused on the readability of website text and have conducted dissimilar user evaluations to

esti- mate web readability. Researchers presented research on the automatic and manual use of thirty-nine readability guidelines on the webpage. This research analyzed the ground-truth readability for a set of fifty web pages by eye-tracking with dyslexic and average readers. The outcomes validated twenty-two guidelines as being connected to readability. The assessment among programmed and human-based results likewise uncovered a complex structure: calculations were better than or as great as human pros at assessing site pages on explicit guidelines - especially those about low-level highlights of website page legibility and text organizing. However, a few guidelines actually require human judgment related to deciphering and understanding site page content. These results add to a guideline characterization laying the ground for future design assessment strategies [21].

Researchers compared the effectiveness and efficiency of heuristic evaluation and user testing in assessing four dis- similar commercial web pages. The outcomes exhibited that both heuristic evaluation and user testing addressed dissim- ilar usability issues. For instance, analysis by the severity of issues found and diminishing return analysis model on the association among the number of new issues exposed with users and assessors used exhibited. These momentous changes found among these two approaches recommended that the two approaches are harmonizing and should not be competing [22].

Another research was conducted to measure the readability and quality of websites provided that information regarding orthodontic clear aligners to prospective patients [23]. Thirty stroke training sites were investigated utilizing readability, responsibility, and unwavering quality measures. Eleven health experts and fifteen clients analyzed six sites for content, design and usability. The site pages habitually met responsibility models; however, their unwavering qual- ity scores were low and their comprehensibility was high. Consumers' perspectives were reliably higher than health experts, yet scores showed their inclinations for explicit pages, especially as far as design. The meaning of thinking about consumers' inclinations while designing and suggest- ing site pages are featured [24].

Also, another research was led to assess the impacts of serif and san serif textual styles in the classification of screen textual styles and print textual styles, as far as Malay text readability on sites. For this reason, four textual styles were chosen, specifically Georgia (serif) and Verdana (san serif) for the primary respondents, and Times New Roman (serif) and Arial (san serif) for the subsequent respondent. Georgia and Verdana were intended for PC screens in the meantime, Times New Roman and Arial were initially intended for print media. A comprehensibility test on a PC screen was con- ducted on 48 students. Generally, the outcomes showed that there was no huge contrast between the clarity of serif and san serif textual style of both screens show class and print display classification. Appropriately, the exploration discoveries and the writing outline, propose Verdana and followed by Georgia as the better decision in showing long text on sites. Similarly, as expected, Times New Roman and Arial text styles are favorable to provide great readability for print media, which builds up their status as the printing text style class [25].

People worked to assess the usability of advanced educa- tion sites in Asia. Initially, a web-based Google application overview structure was planned to utilize Google Forms and utilized for the assessment of web ease of use and under-study reaction. After an intensive examination, a compact model was intended to assess the ease of use of instructive sites called the ''Web Usability Evaluation Model'' (WUEM). In this examination, the ten highest level designing colleges in Asia against the elements recorded in the WUEM. The assessment investigation shows that the scholarly sites are halfway usable in their instructive design, route, and furthermore feeble unavailability. The assessment shows a point- by-point primary depiction of what should be worked on in these sites to improve their ease of use. The proposed WUEM helps in a compelling and simple assessment of sites by web designers. The examination will help scholastic web designers to upgrade the ease of use of their sites by considering such straightforward elements recorded in WUEM [26].

Another research focused on how to make web pages more usable for dissimilar age groups in terms of readability. This research focused on eight timeless readability factors for example shading contrast, blank area, line dispersing, text style, text dimension, text width, headings, designs, and liveliness. These eight variables are looked at how changed age gatherings act with the web applications by shifting these eight elements [27].

Different researchers have worked on the evaluation of the text on the web from different perspectives, e.g., to assess the readability according to the guidelines [21], to assess the quality and readability of websites [23], to measure the readability of stroke education websites [24], check the read- ability according to the factors content, style, design, and structure [27]. But up to our knowledge, none of the works found in the previous research are focused on computing the relevancy of non-text images on the website by using Google API Services from a readability perspective. This paper is focused on proposing to compute the relevancy of non-text

images on web pages and a user study carried out with different types of questions for evaluation.

III. HYPOTHESIS

The fundamental hypothesis of this paper is that non-text images can enhance web readability when non-text images are related to the textual part of the web page. Non-text images being used on the websites should indicate transferring the context of the page in a more viable manner. Relevant and quality non-text images can play this role. Irrelevant non-text images seriously affect the readability of websites. Further, the hypothesis can be divided into two sub-hypothesis:

- The utilization of non-text images relevant to the tex- tual part of the website can enhance the readability of the website.
- The extensive usage of non-text images irrelevant to the textual content of the website indicates low web readability.

IV. METHODOLOGY

A new technique to measure the relevancy of non-text images on web pages has been proposed to evaluate above mentioned hypothesis. Once relevancy is measured, a user survey has been conducted for hypothesis validation. In order to compute non-text images relevancy, these are the fundamental steps followed:

A. CORPUS GENERATION

Corpus generation is the first step in which text and images are extracted from the fifty different arbitrarily chosen edu- cational websites in Pakistan. The lists of webpages used for the corpus are listed in Table 1. On the one hand, the images have been extracted by utilizing an image web scraping technique [28]. Almost 500 pictures have been gathered, out of which 180 were non-text images. Different types of images have been extracted by using this and only considered non-text images in this research. After the extraction of non-text images, each non-text image is passed to Google Vision AI services [29]. Google Vision AI assigns labels to graphical contents and rapidly categorizes them into millions of predefined classes. This service identifies objects and faces, reads printed and handwritten text, and constructs valuable metadata into your picture catalog. The confidence score is provided by the service, which represents the accuracy of the results. For instance, in Fig. 1, extracted information from the non-text image is ‘‘Smile, Trousers, Plant, Grass, Leisure, Recreation, Fun, Competition Event, Event, Lawn, Crowd, Team, Happy, Public Event, Academic Institution, Player, T-shirt, Sitting, Campus, University, Tourism’’ with different confidence values. On the other hand, at the same time, the text from the webpage was extracted.

B. DATA PRE-PROCESSING

Once the text is extracted from the web page. A pre-processing has been performed because it is essential to clean the data and get it into a structure that is unsurprising and

FIGURE 1. Non-text image with extracted information.

analyzable for relevance evaluation. The following phases have been executed during the pre-processing:

1) Tokenization - Texts extracted from the webpage have been tokenized for terms identification and it’s the initial phase of the processing of text. Tokens, words acquired subsequent to dividing crude text, help with grasping the specific situation or fostering the model for natural language processing.
2) Eliminate Stop Words - It conveys very slight or no data and typically are eliminated so a calculation could think about just meaningful words. For this reason, we just fabricated a set of stop words like ’is’, ’the’, ’and’, ’are’, ’an’, ’a’, and so on and involved it in our tech- nique. The similarity has been implemented iteratively for all token words and any token word found in this list has been eliminated.
3) Lemmatization and Stemming - To decrease the inflection of words inside the extracted text from a webpage with their root structure we have utilized lemmatization and stemming. A typical word has one root-base structure however could have various varieties. For example, ‘‘help’’ is a root-base word, and helping, helped, and helps are the various types of a single word. Lemmati- zation and Stemming assist us with accomplishing the root structures.
4) Uniform Case - Taking into account the way that the handling of information is case-sensitive on a machine, extracted information must be transformed into a uniform case. Similar words with various meanings, for instance, Apple and apple are controlled in different

ways by machines. In this way, we really want to create the text in a similar case ideally in lowercase.

5) Punctuation Letters Removal - Letters are $, ?, ‘‘, !, etc. The C# language function makes available the list of punctuation letters. Punctuation letters have been elim- inated because they did not provide any information related to semantic similarity.
6) Non-ASCII Letters Removal - Like punctuation, non-ASCII letters are not valuable to capture semantic similarity.

C. FEATURES EXTRACTION

It’s the representation of a sequence of sentences or words into a numeric vector. Term Frequency and Word2Vec have been utilized.

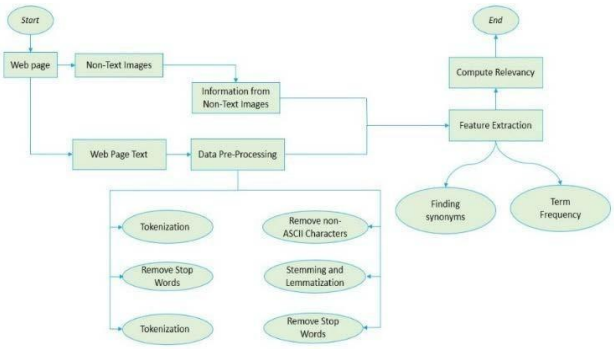

FIGURE 2. Workflow of relevancy computation.

- Finding Synonyms – Synonyms words have been found against every term and word2vec has been utilized for this in our methodology. A set of words got subsequent to stemming was passed as a contribution to word2vec and a set of their synonyms was acquired.
- Term Frequency – It is defined as the proportion of a word’s presence in the text to the all-out number of words in the text. Information extracted from non-text images is connected with a number that characterizes how related every word is to the text of the web- site. Non-text images and website text with matching, related words will have similar vectors, which is what we are seeing for a cosine similarity approach.

D. RELEVANCY COMPUTATION

The main goal of this research is to decide how much the extracted information from non-text images is relevant to the text of the website. The cosine similarity strategy is utilized that checks the relevancy between two vectors. In this way, extracted information from non-text images and website text is described by what is known as the vectors of term frequency. The relevancy of the non-text image in Fig. 1 with its web page text is 0.68. After computing the relevancy of every non-text image, the overall relevancy of the non-text images of a website is computed by using the average of their non-text image’s relevancies. The workflow of non-text image relevancy computation is shown in Fig. 2.

V. EVALUATION

Once the relevancy of non-text images of the webpages is measured, we have evaluated the hypothesis of whether the relevant non-text images on the web could increase readabil- ity, and an online user survey consisting of user testing and heuristics evaluation (experts) has been conducted. For this research work, the two web pages with a better relevancy score and the two web pages with a worse relevancy score according to the methodology proposed were selected. A total of 712 users were enlisted for final user testing and 32 read- ability experts for heuristic investigation for every one of the sites. Clients for user testing were enlisted in light of a profile that was laid out by reviewing an agent test of the client

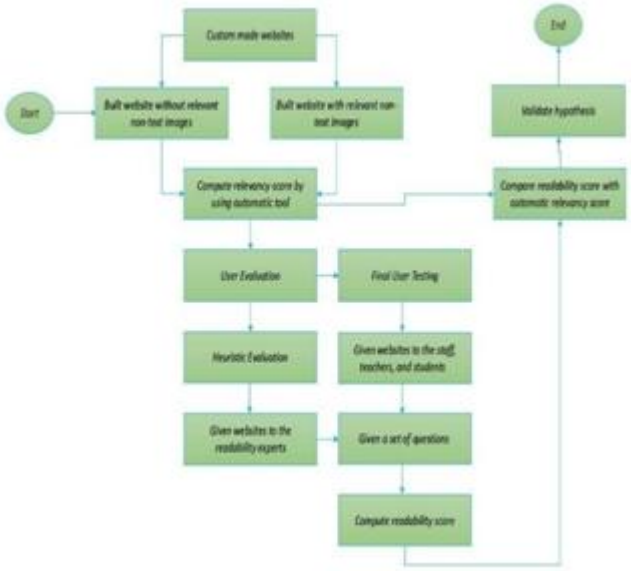

FIGURE 3. Workflow of relevancy evaluation.

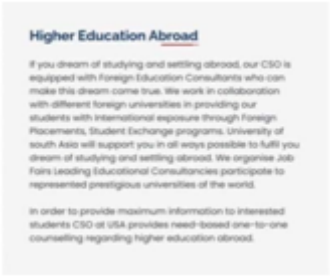

FIGURE 4. Webpage without non-text images.

populace. These clients were non-readability specialists and non-power clients, and that implies they have not had any web assessment encounters yet with some involvement with riding the web. Readability specialists were enlisted to per- form the heuristic examination. For this review, a specialist was characterized as one, who had graduate-level coursework in human-PC collaboration, and brutal variables of website architecture, and who had previously been taught and taken an interest in somewhere around one heuristic web assessment project. This is predictable with the thought that master eval- uators ought to be utilized for heuristic assessment, as they give better outcomes [30]. Our evaluation design consists of the following steps:

A. QUESTIONNAIRE

For validation of the hypothesis, different types of questions which consist of control questions, questions related to the user's understanding, and finally, questions relative to the user's feelings have been asked in the user survey.

For example, consider the best webpage as shown in Fig. 4 without relevant non-text images, please answers the following questions.

i. The webpage explains higher educational institutes.

FIGURE 5. Webpage with non-text images.

ii. Do you think the educational institute has a clean envi- ronment?
iii. Does it consist of male and female students?
iv.Different trees are surrounding the buildings.
v. Do you think the institute has huge buildings?
vi It has good sports grounds.
vii Do you think it has a friendly environment?

On the other hand, the same webpage, as shown in Fig. 5 with relevant non-text images, please answer the fol- lowing questions:

i. The webpage explains higher educational institutes.
ii. Do you think the educational institute has a clean envi- ronment?
iii. Does it consist of male and female students?
iv. Different trees are surrounding the buildings.
v. Do you think the institute has huge buildings?
vi. It has good sports grounds.
vii. Do you think it has a friendly environment?
viii. The new image added to the webpage helps me to understand better the web content.
ix. I prefer a webpage with relevant non-text images (web- page shown in Fig. 5).

B. USER EVALUATION

User evaluation consists of the following steps:

1) OBJECTIVE

The main goal of this evaluation is to evaluate the hypothesis that relevant graphical content could increase web readability for users.

2) ENVIRONMENT

An online survey through Google Forms has been conducted. Experts and users have the option to evaluate the web page at any place.

3) DEPENDENT AND INDEPENDENT VARIABLES

In our case, dependent variables are comprehension with possible levels of bad, fair, good, and excellent. This com- prehension depends on the following independent variables:

- The type of non-text image - chart, diagram, flow dia- gram, photo
- The visual quality of non-text images
- The relationship between non-text images and paragraph

4) PARTICIPANTS

In the heuristic evaluation, users from industry backgrounds were invited to be registered in the evaluation. Developers from different software houses in Pakistan were particularly requested to participate. In total, 32 participants (Male=16 and Fe-male=16) volunteered for our study. While in the final user evaluation, users from academic backgrounds were invited to be registered in the evaluation. Teachers, staff, and students from different educational institutions in Pakistan were specially requested to participate. In total, 712 partic- ipants (Male =356 and Female =356)

volunteered for our study, having been aged between 20 to 35 and at least hav-
ing graduated. This survey was shared and advertised using different social media platforms, and also emailed the links to academic users, and to industry people.

5) PROCEDURE

In this evaluation procedure, a set of questions have been asked from experts and users. Questions asked from the users are related to the images on the websites and it is more spe- cific to the hypothesis that relevant graphical content could increase web readability. This experiment has been conducted in two different groups. Firstly, we gave two websites (with non-text images and without non-text images) to half of the experts and users. Another set of two websites (with non-text images and without non-text images) was given to the other half of the experts and users. During the evaluation procedure, experts and users had the opportunity to clarify any doubts or problems. Experts and users checked the relevance of images with the webpage and answers to questions. User feedback has been recorded and this was used to check the relevancy of graphical content with the text of the web page and its readability.

Website 1 and Website 2

- Group 1 –> with non-text images / questions / without non-text images / questions about preferences
- Group 2 –> without non-text images / questions / with non-text images / questions about preferences

Website 3 and Website 4

- Group 1 –> with non-text images / questions / without non-text images / questions about preferences
- Group 2 –> without non-text images / questions / with non-text images / questions about preferences

VI. RESULTS

Relevancy scores of fifty educational websites have been computed by using the automatic tool, and the outcomes were categorized into three different ranges. Extracted information from non-text images of 16 out of fifty websites was 50-60% matched with the text of the websites. This relevancy score was 61-70% for 20 websites whereas 14 residual websites were found to own non-text images 71-80% relevant to web- sites as shown in Fig. 6. For the user's studies, four of these

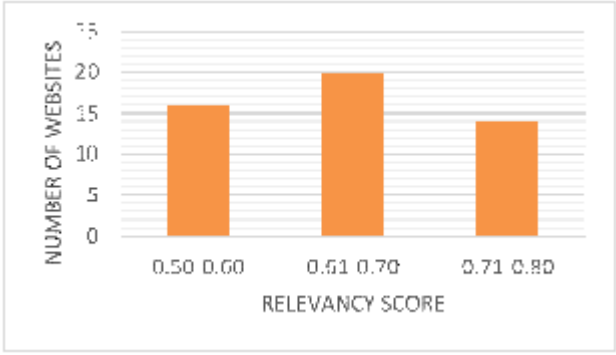

FIGURE 6. Automatic relevancy distribution.

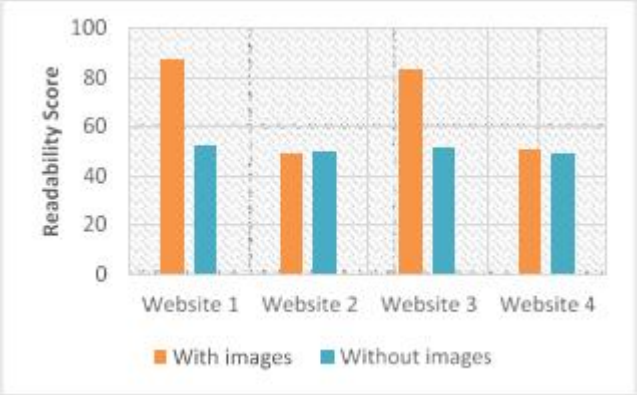

FIGURE 7. User's based readability scores with and without images.

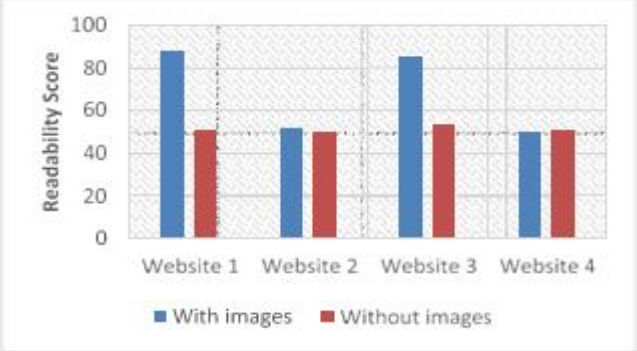

FIGURE 8. Experts-based readability scores with and without images.

FIGURE 9. Web Readability Time

websites were selected: the two ones with the best relevancy score and the two ones with the worst relevancy score. Out- comes were examined and assembled to present in statistical form. According to the user online results, on webpage 1, which has relevant graphical content mostly, the readability score of a webpage without non-text images is 52.01% while with non-text images it is 87.33%. The result was suggestive that being relevant, graphical contents served best for bet- ter understanding of the webpage however, without non-text images users found it difficult to perceive the concept of the same webpage.

Similar was the case with webpage 3 as well. Webpage 3 without non-text images has a readability score of 51.77% while the same webpage with non-text images has 83.45%.

On the other hand, when websites, which have irrelevant non- text images, were served to users without non-text images were found relativity easier to understand as compared to when those were served with graphical content. From the online results, webpage 2 without non-text images has a readability score of 49.11% while with non-text images this score is 50.01%. Similar was the case with webpage 4. Web- page 4 without non-text images has a readability score of 49.11% while the same webpage with non-text images has 50.67%. From the results, it's evident that irrelevant non-text images' negativity affects the readability shown in Fig. 7. Users perceived more accurately and quickly when irrelevant non-text images were removed.

Evaluation by the experts was not much different. Web- page 1 without non-text images has a readability score of 51.17% while the same webpage with non-text images has a readability score of 88.23%. Webpage 3 without non-text images has a readability score of 53.13% while the same web- page with non-text images has 85.11%. A webpage 2 without non-text images has a readability score of 50.13% while the same webpage with non-text images has 51.67%. Webpage 4 without irrelevant non-text images has a readability score of 51.15% while the same webpage with non-text images has 49.63% as shown in Fig. 8

We have observed that the results of the final user evalu- ation are close to the heuristic evaluation, and also observed that the websites have high relevancy scores and high read- ability scores in the user evaluation. On the other hand, Users understand web pages quickly in the case of relevant non-text images as compared to irrelevant non-text images on the web as shown in Fig.9. So, the results validate the hypothesis that relevant non-text images could enhance web readability.

VII. CONCLUSION

The non-text image being used on a webpage that is relevant to the webpage supports the greater readability of the web- page. Dissimilar assessment techniques are presented that evaluate the textual content of the web. However, to the best of our knowledge, no work has been done on evaluating the readability of non-text images on the web. This paper proposes a new methodology to measure the relevancy of non-text images on websites. In this approach, Google ser- vices are used to extract information from non-text images, and the cosine similarity approach is used to compute the relevancy of the extracted information with the webpage text. Fifty websites were evaluated using this technique, and the outcomes specify that non-text images that are irrelevant to the context of the page cause worse relevancy scores, whereasrelevant non-text images result in greater relevancy scores. After measuring relevancy, we evaluated the hypothesis thatrelevant non-text images could increase web readability usinga user survey by considering four websites out of fifty. The survey consists of different types of questions. The results show that the more the graphical content on a webpage is rele- vant to the webpage text, the better the readability score of the webpage in the user evaluation that verifies our hypothesis. This research has focused on educational websites and non-text images. Currently, we are working on studying the otherapplication domains and countries.

APPENDIX A

TABLE 1. Educational websites.

Sr #	Website	Sr #	Website	Sr #	Website
1	http://nu.edu.pk/	18	https://superior.edu.pk/	35	http://www.uop.edu.pk/
2	https://nust.edu.pk/	19	https://www.ucp.edu.pk/	36	https://gcu.edu.pk/
3	https://lums.edu.pk/	20	https://iiu.edu.pk/	37	https://usindh.edu.pk/
4	https://www.giki.edu.pk/	21	http://www.uaf.edu.pk/	38	https://www.hamdard.edu.pk/
		22	https://www.uok.edu.pk/	39	https://www.uhs.edu.pk/
6	http://www.pu.edu.pk/	23	https://gcuf.edu.pk/	40	https://uog.edu.pk/main.php
7	https://www.aku.edu/	24	https://www.bzu.edu.pk/		https://www.hup.edu.pk/
8	https://www.umt.edu.pk/	25	https://su.edu.pk/		
9	https://www.comsats.edu.pk/	26	https://www.neduet.edu.pk/	43	https://fjwu.edu.pk/
	https://uol.edu.pk/	27	https://www.muet.edu.pk/	44	https://iqra.edu.pk/

5	https://itu.edu.pk/	28	https://www.fccollege.edu.pk/	45	https://www.aup.edu.pk/
12	https://www.iba.edu.pk/	29	https://www.kmu.edu.pk/	46	https://web.uettaxila.edu.pk/
	https://www.riphah.edu.pk/		https://szabist.edu.pk	47	https://ssuet.edu.pk/
			https://ndu.edu.pk/		
15	https://uet.edu.pk/	32	https://www.uaar.edu.pk/		
16	https://www.iub.edu.pk/	33	https://www.pide.org.pk/	50	https://www.ntu.edu.pk/
17	https://www.bahria.edu.pk/ https://qau.edu.pk/	34	https://www.numl.edu.pk/		

ACKNOWLEDGMENT

The author would like to thank jayam college of engineering and technology, Dharmapuri.