International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

December 2024 | Issue 61 | Volume 16 | Number 4 | Pages 322-335

AI-DRIVEN AUTOMATION IN SEO: ENHANCING IMAGE DESCRIPTIONS WITH BLIP-PROCESSOR MODELS

Z. Ayaz

Department of Management Information Systems, Gazi University, Ankara, Turkey, zafer@gazi.edu.tr

Abstract- Automating Search Engine Optimization (SEO) for HTML images represents a significant innovation by streamlining the generation of alt text descriptions for image elements, a critical component of improving search engine visibility and website traffic. This study leverages the technical capabilities of BLIP-Processor models, specifically the "base" and "large" versions of the "Salesforce/blip-image-captioning" dataset, within the Python and Hugging Face platforms. By automating the "Image-to-Text" process, the system efficiently generates SEO-compatible annotations and file names for HTML images, significantly reducing the time and effort required for manual description creation in image-intensive domains such as e-commerce and news platforms. The findings highlight the technical trade-offs between the models: the "base" model provides rapid and sufficiently descriptive outputs. In contrast, albeit slower, the "large" model generates more nuanced and detailed annotations. Beyond enhancing SEO practices, the automation of image descriptions has practical implications for digital marketing efficiency and web accessibility. By enabling visually impaired users to understand image content better and simplifying the content management process for businesses, this approach enhances inclusivity while driving efficiency in digital content management.

Keywords: Search Engine Optimization, SEO, Artificial Intelligence, BLIP-Processor, Image-to-Text, Text Extraction from Image.

1. INTRODUCTION

Today, the increase in internet use and the number of users connected to the global vast area network is directly related to the rapid rise of the World Wide Web [15]. However, users may need help finding the page they are looking for because they come across many web pages when they search, and the loading time of these pages may need to be more predictable [21]. The proliferation of search engines has made them indispensable to daily life [2]. Search engines that use complex algorithms provide search services by crawling many websites, but sometimes they can show irrelevant results. To solve this problem, researchers have developed search engine optimization techniques to increase website visibility and make it more relevant to search results [21].

SEO consists of strategies that increase a website's inbound traffic and help it rank high in organic search results [7]. Therefore, growing and changing a website's visibility levels in search engine results page rankings and creating online content targeting have become essential aspects of marketing today [2]. The SEO-friendliness of a website ensures that the site is listed highly by search engines while also playing a critical role in increasing the company's business [7]. SEO, which helps ensure that it appears among the leading results in search engines such as Google or Bing, has become a critical business strategy for audiences with increasing digital use, especially for digital journalism [12]. Many company websites need SEO strategies, limiting their search engine visibility and significantly reducing traffic if no alternative digital marketing methods are employed [7].

While SEO metrics provide valuable insights into a site's search engine performance, the impact of these metrics may only sometimes be constant due to the continuous change of search engine algorithms. Today, SEO has become an integral part of the online content development process. It has increased the visibility of digital content and enabled it to reach large audiences [2]. Analyzing the features of top-ranked websites and evaluating their impact on rankings are two primary research methods accepted in SEO research [20].

SEO is divided into on-page and off-page optimization. On-page optimization entails configuring site elements, including dividing site content into separate themes and categories (SILOS), site performance, meta tags, and sitemaps, to improve search engine compatibility. Off-page optimization focuses on increasing domain authority through content creation and backlinks, utilizing methods like social media, press releases, and directory submissions. Tools such as Google Analytics and Google PageSpeed Tool support SEO improvement efforts [21].

This study examined the success status of naming and automatic generation of sub-title HTML "*image*" elements, one of the SEO on-page techniques, using Python programming language and Hugging Face platform libraries and pre-trained "*base and large*" data sets. The Hugging Face platform defines itself as "*the A.I. community that builds the future*" and is one of the

Registered: 240515 | Article Info: 1928 | Accepted: 241122
Reserved: 240928 | Available: 241231

platforms that offers many analysis libraries, especially natural language processing (NLP), which is reported to aim to democratize the machine learning process [10]. This study aims to contribute innovatively to the SEO field by supporting the definitions needed by SEO-oriented HTML "*image*" elements with an artificial intelligence-supported automatic description model. The contribution of image optimization to SEO is widely discussed in the literature. However, studies on developing tools that automate this contribution and speed up the coding process are minimal. This study offers a new approach in this field by integrating object recognition algorithms with an SEO-friendly description process.

2. CONCEPTUAL FRAMEWORK

This section discusses the general principles of SEO, the effect of using images by the rule on search rankings, and the benefits of supporting images with alt texts.

2.1. Search Engine Optimization (SEO)

Initially, users accessed websites directly due to their limited number. Rapid technological advances have allowed fast access to information through numerous search engines, digital libraries, and other databases as the Internet and its applications have grown [30]. With billions of websites today, search engines serve as intermediaries, directing users to relevant content through refined algorithms. As these ranking criteria remain undisclosed, SEO professionals rely on observation and practice to optimize website visibility [17].

SEO techniques and practices aim to rank a website first on search results pages (SERPs). Websites can increase their visibility in search engines by combining techniques such as keyword density, on-page optimizations, and metadata optimization [25].

Search engine optimization is a tactic commonly used in Internet marketing and aims to rank high in search engine results using specific keywords. SEO requires optimizing web pages by search engine algorithms, focusing on organic results. Web admins carry out this process with technical knowledge and an in-depth understanding of algorithms. While there is no official certification for SEO, advertising certifications such as Google Ads are available. Search engines consider many criteria in their ranking algorithms, and while most of these criteria are known, their impact on ranking needs to be clearly defined. These factors fall into two main categories: on-page and off-page elements that affect ranking [20].

On-page factors are actions taken on the content of a web page that a webmaster owns and has complete control over. A web page's content includes text, URL addresses, file names, external links, photos, site navigation structure, and HTML codes. Off-page elements depend on various external factors and have little to do with the content of a web page but strongly affect modern search engine algorithms or features.

These are issuing a web page creator has no control [20]. Some known best practices for search engine compliance are given in Table 1. Since the main focus of the research is on-page factors, off-page factors are not discussed in detail in this study.

Table 1. Example of on-page and off-page SEO factors [20]

On-Page SEO Factors	Off-Page SEO Factors
Text integrity (including quality of information)	Inbound and outbound connections and the quality of those connections
Easy navigation	Social media recommendations
The titles of the pages ("title" tags in the HTML source)	Domain age and other outsourced features
The "meta description" tag in the HTML source	etc.
Using the H tag to indicate headings (H1, H2, H3, etc.)	
ALT title attributes of images	
Anchor texts of links	
URL addresses of the pages	
Loading speed of the website	
Code ratio of HTML content	

2.2. The Role of SEO Techniques in Website Visibility

In the information-rich digital era, search engines are vital for users, with surveys indicating that 61% of global users search for products online, and 44% utilize search engines for this purpose. Users predominantly focus on top-ranking results on the first pages. While paid advertisements often dominate these rankings, SEO techniques aim to enhance a website's position organically, driving cost-free traffic through improved search visibility [17]. Applying SEO techniques can improve a website's visibility and make it more likely to appear in top search results. According to a study, websites on the first page of Google receive 71.33% of all organic search traffic [14].

Integrating SEO and front-end (UX) designs combines user experience and higher search engine visibility and increases website usability. SEO focuses on attracting organic traffic through keyword and metadata optimization, boosting SERP rankings. Meanwhile, UX design emphasizes intuitive interfaces, seamless navigation, and visual appeal to increase user engagement. This integration ensures that websites rank well while delivering valuable and satisfying user experiences [25].

2.3. Optimizing Images for Search Engines

Web pages often feature images alongside text, and properly formatted images with suitable file sizes improve user experience and search engine rankings. File names and alt tags are essential for SEO, offering descriptive, deterministic image information. Many users search for images to inform their decisions, so these elements must accurately reflect the image content. Title tags also play a vital role in describing internal or external links, helping users and search engines navigate pages efficiently [17]. Below is how to create an SEO-friendly HTML image element:

```
<! -- Not SEO Friendly Image Tag -->
<img src= "img/image3.jpg">
<! -- SEO Friendly Image Tag -->
<img src= "img/macbook-pro-mjlq2.jpg"
      alt= "Macbook Pro MJLQ2">
```

Google indexes by extracting information about the image's subject from the page's content, including the title and caption. The file names of the image elements in the page content can give Google a clue about the image's subject. For this reason, instead of naming the file with a short and meaningless name like IMG00023.JPG, it would be a good idea to name the file "my-new-black-kitten.jpg". Consider automating the image naming process if you have many images on your site. A demonstration of the excellent use case reported by Google is shown below [9]:

Bad (missing alt text):

```
<img src="puppy.jpg"/>
```

Best:

```
<img src="puppy.jpg" alt="Dalmatian puppy playing fetch"/>
```

2.4. Artificial Intelligence Tools for Image-to-Text Captioning

Artificial intelligence (AI) is significantly transforming the future of SEO by enhancing search algorithms and content strategies. AI-powered algorithms, such as Google's RankBrain, improve search result accuracy and personalization, making SEO more complex yet rewarding. Generative AI has revolutionized content creation and optimization, saved time and boosted content quality. Balancing AI's capabilities with human creativity is essential for effective SEO strategies. Adapting to AI-driven tools and focusing on user intent and personalization is crucial for maintaining relevance and improving search rankings [32].

Generative AI, powered by deep learning, revolutionizes image processing by enabling semantic extraction and interpretation of visual data. It is crucial in automating descriptive metadata generation and enhancing image SEO efficiency. By analyzing extensive datasets, Generative AI uncovers complex interactions between images and keywords, enabling the creation of SEO-compliant alt texts that accurately reflect visual content [1].

Lopezosa, et al. present a method to quantify uncertainties in generative AI (G.A.I.), addressing its occasional production of unexpected results in text, image, and video rendering. The study aims to enhance confidence in G.A.I. performance by identifying areas of ambiguity. Text inputs, machine translation metrics, and word embedding similarities of text-image models are compared with the captions generated by four image-text models using BLIP, BLIP2, GIT, and InstructBLIP. Image-to-text generation, a key area in computer vision, creates descriptive captions for images, utilizing datasets like Microsoft COCO and Stanford Image-Paragraph. Techniques such as encoder-decoder architectures, Variational Auto Encoders (VAEs), Generative Adversarial Networks (GANs), and reinforcement learning underpin model development. Companies like Microsoft and Salesforce have contributed significantly with models like GIT and BLIP, advancing image captioning and visual question-answering capabilities. This research field continues to evolve rapidly [12].

In the paper by Li, et al., Image-to-text caption generation is a multifaceted process requiring the integration of several components, including object recognition, attribute identification, image comprehension, feature extraction, natural language processing, and caption evaluation. This process is vital for facilitating effective human-computer interaction. Traditional machine learning (ML) approaches, which rely on manual feature extraction, have shown limitations in efficiency, mainly when applied to large and diverse datasets. In contrast, deep learning (DL) frameworks have advanced significantly over the past decade and offer more robust solutions by leveraging automated feature extraction and pattern recognition. A notable advancement within DL is the development of Vision-Language Pre-Training (VLP) models, which have demonstrated the best performance in generating descriptive and contextually accurate image captions. These models integrate visual and linguistic data to produce coherent textual descriptions. The evaluation of image-to-text algorithms usually includes quantitative metrics to measure the quality of the produced subtitles; the Bilingual Evaluation Assistant (BLEU) scale is widely used to evaluate the success and agreement between the produced and reference subtitles. This paradigm shifts from traditional ML to advanced DL frameworks, particularly VLP models, underscores the evolving sophistication of image-to-text caption generation techniques [4].

BLIP is a versatile Vision-Language Pre-training (VLP) framework that understands and generates tasks by leveraging synthetic captions and noise filtering, achieving state-of-the-art performance in vision-language and zero-shot video-language tasks. The study by Li, et al. demonstrates that BLIP achieves state-of-the-art performance across diverse vision-language tasks, including caption generation, visual question answering, reasoning, and dialogue, by effectively integrating caption generation and filtering mechanisms to enhance downstream task performance. Vision-language pre-training (VLP) enhances vision and language tasks by using large-scale image-text pairs, often sourced from the web, due to the high cost of human annotation. Li, et al.'s paper highlights the suboptimal nature of noisy parts. The working web texts for VLP introduce CapFilt, a method to utilize these datasets more effectively [31].

In addition, the model proposed in Li, et al.'s BLIP paper is a unified vision-language model that works in three modes: (1) a unimodal encoder aligns vision and language representations using image-text comparison loss, (2) an image-based text encoder models vision-language interactions using image-text matching loss, and (3) an image-based text decoder generates captions using language modeling loss. The model uses shared cross-attention layers and feedforward networks for efficient processing. The schematic drawing of this model is shown in Figure 1 [31].

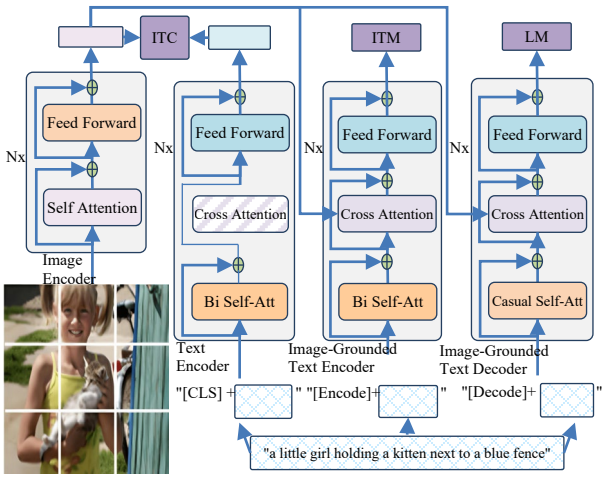

Figure 1. Pre-training model architecture and goals of BLIP proposed new BLIP model by Li, et al. [31]

2.5. Related Studies

To find national and international studies related to the research, 14 records were listed on the Scopus database on 12.11.2024, and the words "image-to-text AND search" were listed. Likewise, to find national studies conducted in Turkey, the words "search AND engine AND compliance" were searched on the Dergipark database on 12.11.2024, and it was seen that none of the three records listed were related to our research. In another search, when the words "search AND engine AND optimization" were searched, 42 records were listed. When these studies were examined, it was seen that the studies in Table 2 were relatively close to the subject of our research, but study that solved problem within the scope of our research could not be reached.

Table 2. Existing Studies Similar to the Research Topic Theme

Theme	Author(s), Study and Details
Visual-Based Identifications on Social Media Platforms	The automatic captioning system has been developed for profile pictures of online social network users. The system works with pre-training on large data sets [4] Research has focused on combining artificial intelligence, computer vision, and natural language processing to generate natural language descriptions from images [24] Image-to-text multimodal conversion is studied to create smooth captioning appropriate to a known image [27]
Multimodal Analysis in Fashion Products	By addressing the multimodal recall problem of fashion products, methods of assigning similarity scores to state-of-the-art text-image pairs were examined [13]
Social Media Data and Language Models	A new approach for translating text from stage pictures has been developed. The system includes text detection, recognition, translation, and reconstruction steps [5]. Hugging Face BERT models and machine learning techniques were compared using social media data and language models to predict symptoms of mental disorders [16]
SEO (General Topics)	Search engine optimization analysis of web pages in Turkey and abroad was conducted; Meta, Title, and Description commands were examined [19] The data of the Dokuz Eylül University website before and after SEO improvement were analyzed [22]. The harms of incorrect and excessive use of SEO techniques on websites have been examined [26]
SEO and Competitive Advantage	An evaluation was made on the contribution of SEO to brand image and competitive advantage in marketing [28]
Digital Marketing Strategies	SEO and SEM applications of e-commerce sites such as Trendyol and Morhipo were analyzed [6] The role of SEO in increasing brand awareness and sales is discussed; advantages, challenges, and strategies are discussed [8] The effect of Google algorithms on consumer purchasing behavior and the role of SEO in e-commerce strategies were examined [3]
SEO & Google Search Ranking	The effect of SEO methods on Google search result rankings was analyzed using a one-hundred-day experiment [23]
Functionality of Search Engines	The functionality and accessibility of various search engines in terms of artificial intelligence have been analyzed [18]

3. METHOD

In this study, a "design-based research model" was used. Design-based research examines the process of developing a new product or process in an academic framework to solve a real problem. In this study, to reduce the manual workload and increase accuracy in SEO processes, a description tool that analyzes images and can turn them into text has been developed by adopting natural language processing methods using two data sets, "base and large" belonging to Python and Hugging Face library. The schematic representation of the developed system is shown in Figure 2.

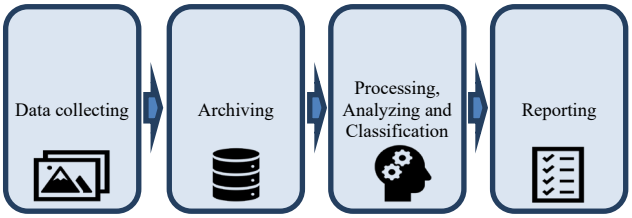

Figure 2. Schematic representation of the processes of the research

The originality of this study is the development of an SEO-oriented automatic description process and its integration with HTML coding. In the literature, while the alignment of images with SEO is usually done manually, a more efficient solution is proposed here: providing automation. This solution is one of the rare studies that deals with the contribution of object recognition-based descriptive algorithms to SEO and offers an innovative approach.

3.1. Purpose of the Research

The primary purpose of the research is to investigate whether the description of HTML image elements, which is one of the factors in making a website SEO compatible, can be automated with the help of artificial intelligence-supported tools instead of a time-consuming workforce. In this context, our research questions are as follows:

RQ1- Can the thing or objects depicted in a picture be explained with the help of artificial intelligence?

RQ2- Is there an efficiency aspect to the image annotation function made with the help of artificial intelligence?

3.2. Data Universe and Data Collecting

The data universe of the research is all image files on the internet. However, to limit the research universe, a dataset was created by compiling fifty images with different content from the "Flickr Image Dataset" collection hosted on the Kaggle portal and available for research under the "Creative Common: Public Domain - CCO" license [11]. The images were selected randomly, and attention was paid to the condition that they would cover different topics in terms of content. The fifty images obtained were copied to the ". /images" folder in the location where the Python code would be run.

3.3. The Python Tool Developed for Application

When the Python code developed within the scope of the research was run, a ". /Images-output" folder was created in the folder where the code was located. Two more folders, "English" and "Turkish", were made under this folder, and the processed images were copied to these folders. The workflow diagram of the prepared Python code is shown in Figure 3.

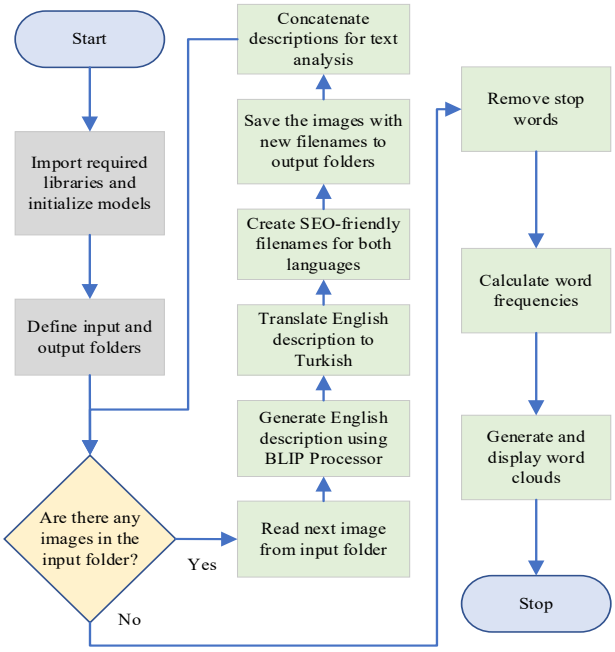

Figure 3. Schematic representation of the processes of the research

The Python codes used in the research and developed by the author are shown in Appendix 3.

3.4. Processing, Analyzing and Reporting

The analysis processes were carried out using the Python programming language and the "base and large" data sets of the HuggingFace "Salesforce/blip-image-captioning" model. The model can make verbal descriptions using a pre-trained data set of pictures. Each image in the "images" folder was subjected to the analysis process, and the files' names were changed according to SEO techniques depending on the texts suggested by the model and automatically copied to the "images-output" folder.

The developed script also measures the time spent on each image to measure the system's performance, combines all the texts, and provides visualization in Word cloud format with the help of the Python Word2Cloud library. The descriptions in English produced by HuggingFace models were converted into Turkish text using the "GoogleTrans" library, which is also one of the libraries in the Python programming language. Since using Turkish characters in file names will create problems in terms of SEO and visibility, Turkish dotted letters have been converted to undotted letters. After all the characters in the file name are converted to lowercase, words are combined with "-" sign to be SEO-compatible.

4. APPLICATION AND FINDINGS

Within the scope of the research, the application was made on a desktop computer with a capacity of 16Gb RAM using an Intel i7-9700T 2.00Ghz CPU and Microsoft 10 (64-bit) operating system. Application details are explained in the following articles.

4.1. Application

An example image from the original version of the fifty image files used as the research dataset is shown in Figure 4. As can be seen, the file name of the image is "COCO_test2014_000000005082.jpg", which contains meaningless letters and some numerical codes but does not contain any detailed information about the content of the photo. If these images are taken and used directly with these names, it will create a problem with SEO compatibility.

Figure 4. An example image from the Flickr test database is used in the process [11]

After the process is completed, the name of this file is determined as "several-young-children-are-playing-soccer-on-a-field-with-other-kids.jpg", which is a very satisfactory result regarding SEO-friendliness.

4.2. Findings

The total analysis time of fifty images was 54.34 seconds when the model was used, and the average processing time per image was 1.09 seconds. When the large model was used, the total analysis time was 132.16 seconds, and the average processing time per image was 2.64 seconds. Figure 5 is presented as a graph comparing the running times of the base and large models.

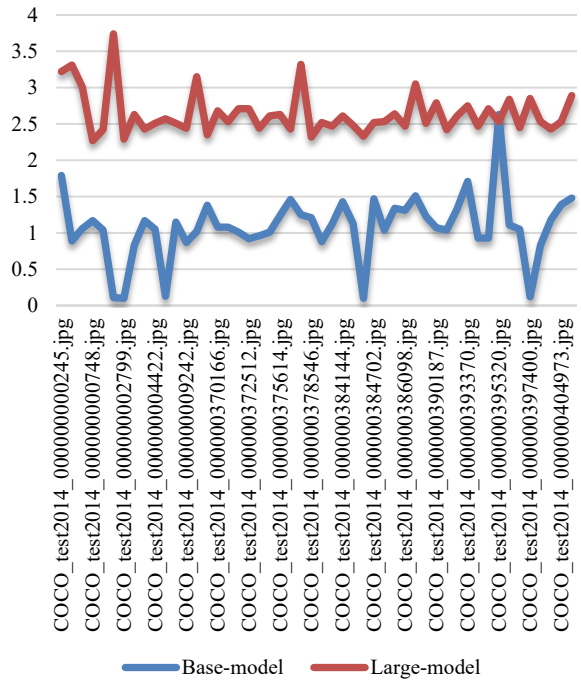

Figure 5. Processing times of image files for base and large model

Using the information obtained after running the developed Python tool, word frequencies are received from the descriptions in the data set, and word cloud visualization can be made accordingly. Word frequencies obtained using the "base model" are given in Table 3.

Table 3. Word frequencies table (Base-model top 10 items)

Words	Frequency
hot	15
dog	15
man	8
field	6
group	5
plate	4
standing	4
red	4
street	3
sitting	3

In Figure 6, word cloud visualizations depend on these word frequencies. The word cloud display makes it easy to understand what images are mainly used as data sets.

Figure 6. Base-model word cloud visual (created by author)

Word frequencies obtained using the "large model" are given in Table 4. In Figure 7, word cloud visualizations depend on the word frequencies obtained from the "large" model.

Table 4. Word frequencies table (Large-model top 10 items)

Words	Frequency
a	85
there	31
is	29
on	26
in	25
with	22
the	22
of	15
are	13
and	11

Figure 7. Large-model word cloud visual (created by author)

4.3. Validity and Reliability of Findings

After the process, a "human evaluation" was performed to examine the quality level of the explanations automatically generated from the images by the model. In this context, the pictures of the data set were reviewed for any missing or incorrect parts in the explanations generated by the "base" and "large" models. In the evaluation, a precision score was calculated for each model. In addition, after the calculation, the average precision score was calculated for each model, which was considered the translation quality score. The ratio of correct words in the translation using BLIP to the number of words in the reference target sentence was calculated using Equations (1) and (2). After calculating the translation quality score for each sentence, the translation quality calculation based on the model was made using Equation (3). Finally, Equation (4) was used to calculate the average translation quality score on a model basis.

$$w_s = \left[w_{s1}, w_{s2}, ..., w_{sn} \right] \tag{1}$$

$$w_t = \left[w_{t1}, w_{t2}, ..., w_{tn} \right] \tag{2}$$

$$r_i = \frac{w_{si}}{w_{ti}} \tag{3}$$

$$Avg_{qs} = \frac{1}{n} \sum_{i=1}^{n} r_i \tag{4}$$

A comparative representation of the scores calculated for both models within the scope of human evaluation is shown in Figure 8. A detailed explanation of the scores used in the graphical drawing for each image is presented in the tables Appendices 1 and 2.

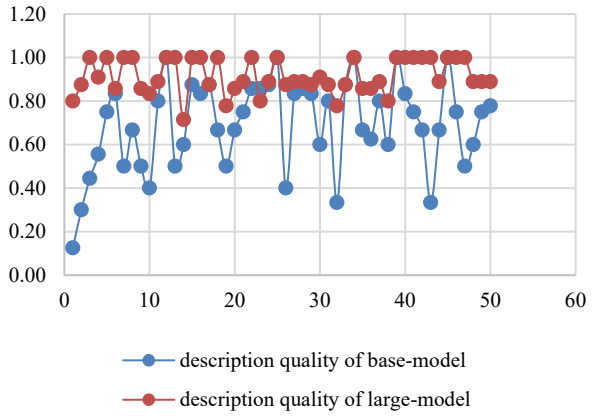

Figure 8. Description quality scores comprising both models

The average of the evaluation scores was calculated separately for both models, and the average translation 1-gram word quality score was found to be 0.69 for the "base" model and 0.91 for the "large" model. The graphical representation of the calculation is shown in Figure 9.

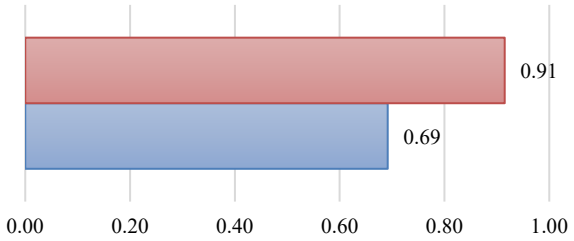

Figure 9. Image captioning quality comparison of both model

Below is a sample HTML coding generated based on information obtained from images using the "base model":

Below is a sample HTML coding generated based on information obtained from images using the "*large model*":

5. CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

Within the scope of the research, a series of analyses were carried out for Image-to-Text translation using the "Blip-Processor" model from the Python programming language and HuggingFace platform libraries and "Salesforce/blip-image-captioning" "base and large" data sets and the images were copied into two different folders in English and Turkish. In this process, automatic descriptions were made using the data set of the relevant model trained for this process without any help other than the visual of the images, and according to the description information, texts were produced for the naming of the files and the explanation of the HTML IMG element (ALT title feature).

Under normal conditions, using human effort, fifty images will take much time to make this classification. Even on a desktop computer that did not have the type of PC and very high-speed features used in the study, it took 54.34 seconds to analyze and name the images with the "base model". The average speed per picture was 1.09 seconds. When using the "large model", the process took 132.16 seconds, and the average processing time per image was 2.64 seconds. Examining the pictures and describing them in such a short time would take much work.

During the analysis process, for the image named COCO_test2014_000000395320.jpg, which has many hot-dog sausages on a grill, the "base" model definition was "hot dog … hot dog" with many repetitions of words. However, the "large" model output accurately translated the same image: "there are many hot dogs on a grill with tongs". The "large" model seems more accurate based on other interpretations.

The average evaluation scores were calculated separately for both models, and the average translation 1 gram word quality score was found to be 0.69 for the "basic" model and 0.91 for the "large" model. The "base" model generally made mistakes in complex colored and multi-object images. Still, the success rate was high in images with more contrasting color tones and more distinct outlines of shapes. On the other hand, the "large" model mostly succeeded in this complexity and correctly identified most of the objects that were close to each other in color. For example, In the image named COCO_test2014_000000003780.jpg, there was a flower in a red vase in front of a glass window. The "base" model defined this image as "a vase with flowers," while the "large" model defined it as "there is a red vase with a flower in it sitting on a window sill" and achieved complete success.

In a study conducted by Reiter (2018), when comparing BLEU with human evaluation, the level of correlation that would be considered sufficient against human evaluation was determined as 0.85 and above as high, 0.70 to 0.85 as medium, 0 to 0.70 as low, and below 0 as unfavorable [33]. Inspired by this, it was observed that the "base" model was close to the starting level of the medium level, and the "large" model was included in the high-level group and produced more consistent results.

Thanks to the developed methodology, website owners, news agencies, image galleries, museums, and many other similar institutions/organizations and businesses will be able to automatically adapt to search engine-compatible pages, especially those in the position of many image-based digital visual content providers. The texts to be used in both the file name and the ALT title [9] have been prepared by obtaining images as recommended by Google to give an example of good practice in the file names produced.

The image produced with the ALT title description "there is a cow that is eating grass in the field" is seen in Figure 10. When the produced descriptions were searched in the Google search engine, very close results were obtained, as seen in Figure 11.

Figure 10. An example original image from the dataset [11]

Figure 11. Screen image from Google search results page [29]

The model's advantages include that it is easy to use, simple, and understandable for an entry-level software developer. Since it is written in Python, it provides a wide range of platform support and performance levels.

Recently, AI has been applied in various fields, such as real-time text recognition for the visually impaired, tourist apps, and robot detection. With rapid advancements in deep learning, text translation technologies have significant potential for improvement [5].

The automatic image description produced by the model supports the creation of SEO-compatible pages. It provides a feature for the screen reader software used by visually impaired individuals to make audio descriptions. Thus, those who integrate this methodology into their web pages will be able to get more traffic through search engines by making the pages where a large number of images are presented together SEO compatible, and they will be able to gain accessibility by making disabled individuals visit the pages more efficiently. In addition, with the development of the model, it has been seen that uncategorized pictures can be easily categorized and classified using descriptive texts.

The most significant limitation of the study is that it did not extract written text from the images, and the performance comparison of all other image-to-text models needed to be made.

5.2. Recommendations

Only two models from the Hugging Face library were used in the study. The same library contains hundreds of natural language processing models and pre-trained datasets. Our proposed model is designed to describe images and is not currently coded to analyze text expressions outside images. It is seen that the use and success of the tool can be increased by using optical character recognition (OCR) models, among other models. On the other hand, thanks to the different models in the Hugging Face library, operations such as scanning an object in a picture, identifying what is being said in a video, and answering questions about the image video can be done quickly.

Image captions are vital in improving web accessibility for visually impaired individuals. They can be used to automatically create explanatory captions (alt text) that provide meaningful context for images. Recent developments in this field focus on improving web accessibility through crowdsourcing, machine-based methods, and hybrid technologies, allowing greater autonomy and participation in daily tasks and social activities. The procedures performed in this study contribute to making images more understandable, especially for visually impaired individuals [34]. However, it has been understood that in future research, the OCR factor should be added to the caption details to improve the performance of websites containing images for the visually impaired and to make images on e-commerce sites more accessible. The author recommends using details such as OCR and tables in the image captioning processes in future studies. It is recommended that researchers conduct future studies to improve these areas.

APPENDICES

Appendix 1. Base-Model Captioning Procession Statistics Table

Image File Name	Processing Time (seconds)	Created Captions from Images	Caption Creating Quality Scores
COCO_test2014_000000000245.jpg	1.79	a woman in a bikini	0.13
COCO_test2014_000000000439.jpg	0.89	a man playing tennis	0.30
COCO_test2014_000000000535.jpg	1.06	a truck with bicycles on the back	0.44
COCO_test2014_000000000748.jpg	1.17	a baby eating food in a high chair	0.56
COCO_test2014_000000000840.jpg	1.04	two zebras walking in the snow	0.75
COCO_test2014_000000002082.jpg	0.11	a group of cows grazing in a field	0.83
COCO_test2014_000000002799.jpg	0.10	a man in a red jacket	0.50
COCO_test2014_000000003780.jpg	0.83	a vase with flowers	0.67
COCO_test2014_000000004311.jpg	1.17	a plate of fruit and a fried chicken	0.50
COCO_test2014_000000004422.jpg	1.05	a book on a bed	0.40
COCO_test2014_000000005082.jpg	0.13	a group of young children playing soccer on a field	0.80
COCO_test2014_000000006249.jpg	1.15	a zebra standing in a field of grass	1.00
COCO_test2014_000000009242.jpg	0.87	a group of horses	0.50
COCO_test2014_000000369938.jpg	1.02	a cow grazing in a field	0.60
COCO_test2014_000000370057.jpg	1.38	a lighthouse on a rocky shore with a lighthouse in the background	0.88
COCO_test2014_000000370166.jpg	1.08	a gife standing in a field	0.83
COCO_test2014_000000370336.jpg	1.08	a little girl sitting on a sink	0.88
COCO_test2014_000000372370.jpg	1.01	a bus on the road	0.67
COCO_test2014_000000372512.jpg	0.92	man wearing a bow tie	0.50
COCO_test2014_000000372995.jpg	0.96	a bunch of apples	0.67
COCO_test2014_000000375578.jpg	1.01	a flock of birds	0.75
COCO_test2014_000000375614.jpg	1.24	a person riding a wave on a surfboard	0.86
COCO_test2014_000000375723.jpg	1.46	a plate of food with a sandwich and a drink	0.86
COCO_test2014_000000377739.jpg	1.25	a young man holding a baseball bat on a street	0.88
COCO_test2014_000000378546.jpg	1.21	a stop signs with a stop sign underneath it	1.00
COCO_test2014_000000381414.jpg	0.88	a body of water	0.40
COCO_test2014_000000382052.jpg	1.12	a boy brushing his teeth in a bathroom	0.83
COCO_test2014_000000384144.jpg	1.43	a red fire hydrant sitting in the middle of a field	0.86
COCO_test2014_000000384265.jpg	1.13	a knife cutting an orange into pieces	0.83
COCO_test2014_000000384481.jpg	0.10	a boat on the water	0.60
COCO_test2014_000000384702.jpg	1.47	a little girl eating a dough dough	0.80
COCO_test2014_000000384914.jpg	1.04	a plate of food	0.33
COCO_test2014_000000385788.jpg	1.34	a boat is in the water near a rocky island	0.88
COCO_test2014_000000386098.jpg	1.31	a swan and its baby swimming in a pond	1.00
COCO_test2014_000000387668.jpg	151	a plate of carrots and broccoles on a table	0.67
COCO_test2014_000000388979.jpg	1.23	a group of people standing around a street	0.63
COCO_test2014_000000390187.jpg	1.07	a man riding a red motorcycle	0.80
COCO_test2014_000000390242.jpg	1.04	a cat is watching a television	0.60
COCO_test2014_000000390755.jpg	133	a red jacket with a blue and white striped shirt	1.00
COCO_test2014_000000393370.jpg	1.71	a group of people standing around a cross	0.83
COCO_test2014_000000393581.jpg	0.93	a green and beige bus	0.75
COCO_test2014_000000394758.jpg	0.93	a bunch of lemons	0.67
COCO_test2014_000000395320.jpg	2.63	hot dog … hot dog	0.33
COCO_test2014_000000396131.jpg	1.11	a cat is sleeping on a bed	0.67
COCO_test2014_000000396849.jpg	105	a man eating a hot dog	1.00
COCO_test2014_000000397400.jpg	0.12	a woman is blowing candles on a cake	0.75
COCO_test2014_000000401039.jpg	0.83	a baby cow	0.50
COCO_test2014_000000404032.jpg	1.18	a woman and child on a motorcycle	0.60
COCO_test2014_000000404973.jpg	1.39	a man walking a herd of sheep down a street	0.75
COCO_test2014_000000405908.jpg	1.48	a man sitting on a blanket with a umbrella over his head	0.78
Total processing time (seconds)	54.34		
Average processing time (seconds)	1.09	Average captioning quality score	0.69

Appendix 2. Large-Model Captioning Procession Statistics Table

Image File Name	Processing Time (seconds)	Created Captions from Images	Caption Creating Quality Scores
COCO_test2014_000000000245.jpg	3.22	there are many people standing on the beach with umbrellas	0.80
COCO_test2014_000000000439.jpg	3.31	there is a man that is playing tennis on the court	0.88
COCO_test2014_000000000535.jpg	3.01	there are many bikes on the trailer on the street	1.00
COCO_test2014_000000000748.jpg	2.27	there is a baby sitting in a high chair eating food	0.91
COCO_test2014_000000000840.jpg	2.42	zebras walking in the snow near a wooden fence in the snow	1.00
COCO_test2014_000000002082.jpg	3.74	there are three cows standing in a field of grass with a stone wall in the background	0.86
COCO_test2014_000000002799.jpg	2.29	skier in orange jacket standing on snow covered slope with skis	1.00
COCO_test2014_000000003780.jpg	2.63	there is a red vase with a flower in it sitting on a window sill	1.00

COCO_test2014_000000004311.jpg	2.43	there is a plate of fruit and a hot dog on a plate	0.86
COCO_test2014_000000004422.jpg	2.51	arafy book with a red alarm clock on top of it	0.83
COCO_test2014_000000005082.jpg	2.57	several young children are playing soccer on a field with other kids	0.89
COCO_test2014_000000006249.jpg	2.51	zebra standing in a field of grass with trees in the background	1.00
COCO_test2014_000000009242.jpg	2.44	horses are walking on a dirt road in a wooded area	1.00
COCO_test2014_000000369938.jpg	3.15	there is a cow that is eating grass in the field	0.71
COCO_test2014_000000370057.jpg	2.35	there is a lighthouse on a rocky cliff next to the ocean	1.00
COCO_test2014_000000370166.jpg	2.68	there is a giraffe standing in a field of tall grass	1.00
COCO_test2014_000000370336.jpg	2.53	there is a little girl sitting on a sink brushing her teeth	0.88
COCO_test2014_000000372370.jpg	2.71	there is a bus and a car on the street next to each other	1.00
COCO_test2014_000000372512.jpg	271	araffe dressed in clown makeup and sunglasses holding a black umbrella	0.78
COCO_test2014_000000372995.jpg	2.44	there are many crates of apples in the store for sale	0.86
COCO_test2014_000000375578.jpg	2.61	flock of white birds sitting on a rocky cliff next to the ocean	0.89
COCO_test2014_000000375614.jpg	2.63	surfer riding a wave in the ocean with a mountain in the background	1.00
COCO_test2014_000000375723.jpg	2.43	there is a sandwich on a plate with a lemon wedge	0.80
COCO_test2014_000000377739.jpg	3.32	there is a man holding a baseball bat and a ball	0.89
COCO_test2014_000000378546.jpg	2.32	there is a stop sign with a traffic light attached to it	1.00
COCO_test2014_000000381414.jpg	2.52	boats are docked on the water near a house and a boat	0.88
COCO_test2014_000000382052.jpg	2.47	there is a young man brushing his teeth in the bathroom	0.89
COCO_test2014_000000384144.jpg	2.61	there is a red fire hydrant in the middle of a field	0.89
COCO_test2014_000000384265.jpg	2.48	there is a knife cutting an orange with a sharp blade	0.88
COCO_test2014_000000384481.jpg	2.33	boats are sailing in the water near a small island	0.91
COCO_test2014_000000384702.jpg	2.52	there is a little girl that is eating a donut	0.88
COCO_test2014_000000384914.jpg	2.53	there is a plate of food with a sandwich and a salad	0.78
COCO_test2014_000000385788.jpg	2.64	there is a boat that is floating in the water near a mountain	0.88
COCO_test2014_000000386098.jpg	2.47	there are two swans swimming in the water with their babies	1.00
COCO_test2014_000000387668.jpg	3.05	there are carrots and broccoli on a plate on a table	0.86
COCO_test2014_000000388979.jpg	2.51	people are standing around a table with paintings and umbrellas	0.86
COCO_test2014_000000390187.jpg	2.79	agraffe riding a red motorcycle in a parking lot with a helmet on	0.89
COCO_test2014_000000390242.jpg	2.42	there is a cat that is standing on a tv stand	0.80
COCO_test2014_000000390755.jpg	2.61	araphid red jacket with a blue shirt and a red tie	1.00
COCO_test2014_000000393370.jpg	2.75	skiers pose for a picture on top of a mountain with a cross	1.00
COCO_test2014_000000393581.jpg	2.47	there is a green and beige bus driving down the street	1.00
COCO_test2014_000000394758.jpg	2.71	araphid box of oranges with a blue pair of scissors	1.00
COCO_test2014_000000395320.jpg	2.53	there are many hot dogs on a grill with tongs	1.00
COCO_test2014_000000396131.jpg	2.84	there is a cat that is laying on a bed with a blue blanket	0.89
COCO_test2014_000000396849.jpg	2.45	there is a man eating a hot dog in a park	1.00
COCO_test2014_000000397400.jpg	2.85	araffe dog with a birthday hat on sitting in front of a cake	1.00
COCO_test2014_000000401039.jpg	2.53	there is a baby cow nursing from its mother in a field	1.00
COCO_test2014_000000404032.jpg	2.43	there are two women and a child riding on a motorcycle	0.89
COCO_test2014_000000404973.jpg	2.53	araphid man walking a herd of sheep down a street	0.89
COCO_test2014_000000405908.jpg	2.89	araphid man sitting on a table with a basket of fruit under an umbrella	0.89
Total processing time (seconds)	132.16		
Average processing time (seconds)	2.64	Average captioning quality score	0.91

Appendix 3. The Python Code Developed by Author

```
import pandas as pd
import time
from collections import Counter
from transformers import BlipProcessor,
BlipForConditionalGeneration
from PIL import Image
import os
import shutil
from wordcloud import WordCloud
import nltk
from nltk.corpus import stopwords
import matplotlib.pyplot as plt
from googletrans import Translator

# Function to replace Turkish characters with English
equivalents
def turkish_to_english(text):
    replacements = {
        'ç': 'c', 'ğ': 'g', 'ı': 'i', 'ö': 'o', 'ş': 's', 'ü': 'u',
        'Ç': 'C', 'Ğ': 'G', 'İ': 'I', 'Ö': 'O', 'Ş': 'S', 'Ü': 'U'
    }
    for turkish_char, english_char in replacements.items():
        text = text.replace(turkish_char, english_char)
    return text
```

```
# Function to create a filename based on text and original file extension
def create_image_filename(text, extension):
    text = text.lower()
    text = turkish_to_english(text)
    tokens = text.split()
    filename = "-". join(tokens) + extension # Add the original file extension
    return filename

# Load the BLIP processor and model
#processor = BlipProcessor.from_pretrained ("Salesforce/blip-image-captioning-base", clean_up_tokenization_spaces=True)
#model = BlipForConditionalGeneration.from_pretrained("Salesforce/blip-image-captioning-base")

processor = BlipProcessor.from_pretrained ("Salesforce/blip-image-captioning-large", clean_up_tokenization_spaces=True)
model = BlipForConditionalGeneration.from_pretrained("Salesforce/blip-image-captioning-large")

# Initialize the Google Translator
translator = Translator ()

# Define input and output folder paths
input_folder = ". /images"                  # Folder containing original images
output_folder_en = ". /Images-output/english" # Folder for saving English named images
output_folder_tr = ". /Images-output/turkish" # Folder for saving Turkish named images

# Create the output subfolders if they don't exist
os.makedirs(output_folder_en, exist_ok=True)
os.makedirs(output_folder_tr, exist_ok=True)

# Initialize lists for storing performance data, descriptions, and processing times
image_names = []
processing_times = []
english_descriptions = []
turkish_descriptions = []
image_htmls = []

# Process each image in the folder
for image_name in os.listdir(input_folder):
    if image_name.lower().endswith((('.png', '.jpg', '.jpeg')): # Check for image file extensions
        # Start timing
        start_time = time.time()

        # Open image
        image_path = os.path.join(input_folder, image_name)
        image = Image.open(image_path). convert("RGB")

        # Get the original file extension
        _, extension = os.path.splitext(image_name)

        # Process the image to generate a description
        inputs = processor (image, return_tensors="pt")
        max_length = 500
        out = model.generate(**inputs, max_length=max_length)
        description_en = processor.decode(out[0], skip_special_tokens=True)
        english_descriptions.append(description_en)

        # Translate the description to Turkish
        description_tr = translator.translate(description_en, src='en', dest='tr').text
        turkish_descriptions.append(description_tr)

        # Generate filenames based on descriptions and original file extension
        filename_en = create_image_filename(description_en, extension)
        filename_tr = create_image_filename(description_tr, extension)

        # Define full output paths for each language
        output_path_en = os.path.join(output_folder_en, filename_en)
        output_path_tr = os.path.join(output_folder_tr, filename_tr)

        # Copy the original image to the respective language folder with new filenames
        shutil.copy(image_path, output_path_en)
        shutil.copy(image_path, output_path_tr)

        # End timing
        end_time = time.time()
        processing_time = end_time - start_time
        processing_times.append(round(processing_time, 2))
        image name. append (image name)

        print (f"<img src=\"{filename_en} \" alt=\"{description_en} \">")
        #print (f"Image: {image_name} processed in {processing_time:.2f} seconds")

# Function to generate and display a word cloud
def generate_word_cloud(text, title):
    wordcloud = WordCloud(width=800, height=400, background_color="white").generate(text)
    plt.figure(figsize=(10, 6))
    plt.imshow(wordcloud, interpolation="bilinear")
    plt.axis("off")
    plt.title(title)
    plt.show()

# Remove stop words from English and Turkish descriptions
stop_words_en = set(stopwords.words('english'))
```

```
stop_words_tr = set(stopwords.words('turkish'))

# Concatenate all English and Turkish descriptions
all_text_en = " ". join(english_descriptions)
all_text_tr = " ". join(turkish_descriptions)

filtered_text_en = " ". join ([word for word in
all_text_en.lower().split() if word not in stop_words_en])
filtered_text_tr = " ". join ([word for word in
all_text_tr.lower().split() if word not in stop_words_tr])

# Calculate word frequencies for English and Turkish
descriptions
word_freq_en = Counter(filtered_text_en.split())
word_freq_tr = Counter(filtered_text_tr.split())

# Generate word clouds
generate_word_cloud (filtered_text_en, "Word Cloud for
English Descriptions")
generate_word_cloud(filtered_text_tr, "Word Cloud for
Turkish Descriptions")

print ("Word frequencies for English descriptions:",
word_freq_en)
print ("Word frequencies for Turkish descriptions:",
word_freq_tr)

# Print word frequencies
print ("English Word Frequencies:", filtered_text_en)
print ("Turkish Word Frequencies:", filtered_text_tr)

# Convert the word frequencies to Data Frames
df_en = datagram(word_freq_en.items(),
columns=["Word", "Frequency"]).sort values
(by="Frequency", ascending=False)
df_tr = datagram(word_freq_tr.items(),
columns=["Word", "Frequency"]).sort values
(by="Frequency", ascending=False)

# Combine processing times, English, and Turkish
descriptions into a DataFrame
df_processing_times = pd.DataFrame({
   "Image Name": image_names,
   "Processing Time (seconds)": processing_times,
   "English Description": english_descriptions,
   "Turkish Description": turkish_descriptions
})

# Save the tables to CSV files for further use
df_en.to_csv ("english_word_frequency.csv",
index=False)
df_tr.to_csv ("turkish_word_frequency.csv",
index=False)
df_processing_times.to_csv
("image_processing_times.csv", index=False)

# Print the tables for word frequencies and processing
times
print ("\nEnglish Word Frequency Table:")
print(df_en)

print ("\nTurkish Word Frequency Table:")
print(df_tr)

print ("\nImage Processing Times Table:")
print(df_processing_times)
```

NOMENCLATURES

Acronyms

AI.	Artificial Intelligence
G.A.I.	Generative Artificial Intelligence
HTML	Hyper Text Markup Language
SEO	Search Engine Optimization
URL	Unified Resource Locator

REFERENCES

[1] M. Amara, R. Zaghdoud, O.B. Rhaiem, E. Althubiti, "Enhancing Image SEO Using Deep Learning Algorithms: A Research Approach", Engineering, Technology and Applied Science Research, Issue 4, Vol. 14, pp. 15194-15200, August 2024.

[2] A.H. Ataei Monfared, A.H. Avizheh, "Re-Rank Search Engine Results Via Machine Learning", 2023 9th International Conference on Web Research, ICWR 2023, pp. 253-258, Tehran, Islamic Republic of Iran, May 2023.

[3] A. Ay, S. Bekar, B. Caglar Demir, "Search Engines and Google Algorithms as Digital Marketing and E-Commerce Tools", Hatay Mustafa Kemal University Journal of Social Sciences Institute, Issue 52, Vol. 20, pp. 295-317, December 2023.

[4] S. Bharne, P. Bhaladhare, "Enhancing User Profile Authenticity Through Automatic Image Caption Generation Using a Bootstrapping Language–Image Pre-Training Model", Engineering Proceedings, Issue 1, Vol. 59, p. 182, January 2024.

[5] Y. Cheema, M.N. Cheema, A. Nazir, F.A. Khokhar, P. Li, A. Ahmed, "A Novel Approach for Improving Open Scene Text Translation with Modified GAN", Visual Computer, April 2024.

[6] S. Dogan, "Use of Internet Search Engines in Digital Marketing: Comparative Analysis of Morhipo and Trendyol Sites", Inonu University Electronic Journal of the Faculty of Communication (INIF E-Journal), Issue 1, Vol. 6, pp. 123-139, May 2021.

[7] J.E. Garcia, R. Lima, M.J.S. Fonseca, "Search Engine Optimization (SEO) for a Company Website: A Case Study", In: A. Rocha, H. Adeli, G. Dzemyda, F. Moreira, (eds.,) Information Systems and Technologies, WorldCIST 2022, Lecture Notes in Networks and Systems, Vol. 470. Springer, Cham, Budva, Montenegro, May 2022.

[8] Y. Gedik, "Search Engine Optimization (SEO): A Conceptual Assessment of Its Advantages, Challenges, and Strategies", International Journal of Social Sciences and Education, Issue 6, Vol. 4, pp. 145-176, April 2022.

[9] Google, "Image SEO Best Practices", https://developers.google.com/search/docs/appearance/google-images?hl=tr, accessed 12 November 2024

[10] HuggingFace, "About Us", https://huggingface.co/huggingface, accessed 10 November 2024.
[11] Kaggle, "Flickr Image Dataset", https://www.kaggle.com/datasets/hsankesara/flickr-Image-Dataset, Accessed 10 November 2024.
[12] C. Lopezosa, D. Giomelakis, L. Pedrosa, L. Codina, "Google Discover: Uses, Applications and Challenges in the Digital Journalism of Spain, Brazil and Greece", Online Information Review: The International Journal of Digital Information, Research and Use, Issue 1, Vol. 48, pp. 123-143, January 2024.
[13] G. Moro, S. Salvatori, G. Frisoni, "Efficient Text-Image Semantic Search: A Multi-Modal Vision-Language Approach for Fashion Retrieval", Neurocomputing, Vol. 538, pp. 126-196, June 2023.
[14] J. Ologunebi, E.O. Taiwo, "The Importance of SEO and SEM in Improving Brand Visibility in E-Commerce Industry; A Study of Decathlon, Amazon and ASOS", https://ssrn.com/abstract=4638890, December 2023.
[15] V. Prifti, "Applications of Technology in Business in Digital Age", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 59, Vol. 16, No. 2, pp. 59-64, June 2024.
[16] A. Pourkeyvan, R. Safa, A. Sorourkhah, "Harnessing the Power of Hugging Face Transformers for Predicting Mental Health Disorders in Social Networks", IEEE Access, Vol. 12, pp. 28025-28035, February 2024.
[17] K.I. Roumeliotis, N.D. Tselikas, "An Effective SEO Techniques and Technologies Guide-Map", Journal of Web Engineering, Issue 5, Vol. 21, pp. 1603-1649, August 2022.
[18] D. Saritac, "Artificial Intelligence Impact and Functionality Analysis in Digital Search Engines", Journal of Communication Science Research, Issue 1, Vol. 4, pp. 17-31, December 2023.
[19] T. Sevindik, "Search Engines Optimization and General Usage", Education Sciences, Issue 3, Vol. 4, pp. 1136-1149, April 2009.
[20] S.S. Shaffi, I. Muthulakshmi, "Search Engine Optimization by Using Machine Learning for Web Page Classification", 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), pp. 342-349, Trichy, India, November 2022.
[21] M. Sithambranathan, N. Zainal, R.A. Sulaiman, "Comparative Study of Search Engine Optimization Algorithms on Retrieval Time and Precise Data Entry for Websites", Journal of Soft Computing and Data Mining, Issue 1, Vol. 4, pp. 53-62, May 2023.
[22] C. Tarhan, H. Koc, "Measuring the Changes Caused by Site Search Optimization in Website Traffic", Journal of Management Information Systems, Issue 2, Vol. 5, pp. 69-80, December 2019.
[23] E.G. Vuran, A. Alpkocak, "Analysis of Search Engine Optimization Methods", Dokuz Eylul University Faculty of Engineering Journal of Science and Engineering, Issue 65, Vol. 22, pp. 491-503, May 2020.
[24] X. He, L. Deng, "Deep Learning for Image-to-Text Generation: A Technical Overview", IEEE Signal Processing Magazine, Issue 6, Vol. 34, pp. 109-116, November 2017.
[25] W. Xinghai, "Reimagining Website Usability: A Conceptual Exploration of SEO and UX Design Integration", Journal of Digitainability, Realism and Mastery (DREAM), Issue 3, Vol. 2, pp. 60-66, March 2023.
[26] N. Yalcin, A. Kilic, "The Harms of Applying Excessive Seo Technique to Websites", AJIT-E: Academic Journal of Information Technology, Issue 24, Vol. 7, pp. 69-74, July 2016.
[27] Z. Yongqiang, J. Zhi, Z. Feng, Z. Haiyan, T. Zhengwei, D. Chengfeng, X. Xinhai, L. Donghong, "Deep-Learning-Based Image Captioning Analysis and Prospects", Journal of Image and Graphics, Issue 9, Vol. 28, pp. 2788-2816, September 2023.
[28] D. Yuksel, M. Tolon, "Search Engine Optimization as a Digital Marketing Strategy (SEO)", International Journal of 3D Printing Technologies and Digital Industry, Issue 3, Vol. 3, pp. 236-243, December 2019.
[29] Google, "Screen Image of Search Results Page Related to: There is a Cow That is Eating Grass in Field", https://www.google.com/search?q=there+is+a+cow+that+is+eating+grass+in+the+field&sca_esv=fb0328165e44674aandrlz=1C1CHBD_trTR917TR917&udm=2&biw=1257&bih=751&sxsrf=ADLYWII9SlddxYuEZ02aY63xWXIF2BK_MA%3A1732216912548&ei=UIg_Z5iIIfnlxc8PwcemyQY&ved=0ahUKEwiY8Pvtku6JAxX5cvEDHcGjKWkQ4dUDCBE&uact=5&oq=there+is+a+cow+that+is+eating+grass+in+the+field&gs_lp=EgNpbWciMHRoZXJlIGlzIGEgY293IHRoYXQgaXMgZWF0aW5nIGdyYXNzIGluIHRoZSBmaWVsZEiCflDZAliRfHABeACQAQCYAX6gAdUEqgEDMC41uAEDyAEA-AEB-AECmAIBoAIIqAIKwgIHECMYJxjqApgDCJIHATGgB-EB&sclient=img, 21 November 2024.
[30] M.N. Mansoor, M.S.H. Al Tamimi, "Plagiarism Detection System in Scientific Publication Using LSTM Networks", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 53, Vol. 14, No. 4, pp. 17-24, December 2022.
[31] J. Li, D. Li, C. Xiong, S. Hoi, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation", Computer Vision and Pattern Recognition, arXiv:2201.12086, February 2022.
[32] Forbes, "AI Is Driving the Future Of SEO: How to Adapt", https://www.forbes.com/councils/forbesagencycouncil/2024/02/02/ai-is-driving-the-future-of-seo-how-to-adapt/, accessed 23 November, 2024.
[33] E. Reiter, "A Structured Review of the Validity of BLEU", Computational Linguistics, Issue 3, Vol. 44, pp. 393-401, September 2018.
[34] T. Tiwary, R.P. Mahapatra, "An Accurate Generation of Image Captions for Blind People Using Extended Convolutional Atom Neural Network", Multimed Tools Appl, Issue 82, pp. 3801-3830, January 2023.

BIOGRAPHY

Name: **Zafer**
Surname: **Ayaz**
Birthday: 11.07.1974
Birthplace: Ankara, Turkey
Bachelor: Department of Business Administration, Faculty of Business Administration, Anadolu University, Eskisehir, Turkey, 2004
Master: Department of Management Information Systems, Graduate School of Informatics, Gazi University, Ankara, Turkey, 2008
Doctorate: Department of Management Information Systems, Graduate School of Informatics, Gazi University, Ankara, Turkey, 2020
The Last Scientific Position: Assist. Prof., Department of Management Information Systems, Faculty of Applied Sciences, Gazi University, Ankara, Turkey, June 2022
Research Interests: Social Media Marketing, Programming, Database Administrating, Artificial Intelligence, Web3 Technologies, e-Commerce, Information Systems, Datacenter Design and Management
Scientific Publications: 13 Papers, 1 Book Chapter, 2 Projects, 2 Theses
Scientific Memberships: Head of Gazi University Computer Sciences Research and Application Center, Gazi University Web3 Technologies Research Group