Analysis of Text-to-Image Artificial Intelligence Systems in Terms of Contribution to Interior Coloring

Araştırma Makalesi/Research Article

Muhterem HOŞER*, Erdem KÖYMEN

Mühendislik ve Doğa Bilimleri Fakültesi, S. Zaim Üniversitesi, İstanbul, Türkiye

muhteremhoser@gmail.com ,erdem.koymen@izu.edu.tr

(Geliş/Received:19.02.2023; Kabul/Accepted:03.09.2023)

DOI: 10.17671/gazibtd.1252993

Abstract— In this article, based on its potential contribution to architectural design processes, research has been made on the "text-to-image" systems of artificial intelligence. In the research, the four most common systems Craiyon, Dall-E, Midjourney, and Stable Diffusion were selected, and these systems were tested for coloring a pre-school education space. First of all, the "kindergarten" text was presented to the systems and according to this text, four alternative images were produced from each system. Afterward, the dominant color coding of the images was analyzed in the computer environment. The 3D model of preschool space was colored with the obtained color codes. The 16 images that emerged because of coloring were presented to 62 expert participants, consisting of preschool teaching and architecture/interior architecture department members, accompanied by a survey. In the survey, the experts were asked to evaluate the colored images in "entertainment" and "academic" contexts. As a result of the statistical analysis of the survey data showed that the Craiyon system used colors more successfully than other systems in terms of coloring a preschool education space. This study measured the ability of artificial intelligence systems from text-to-image to interpret the text in terms of the production of color codes suitable for the type of space. However, it has been tried to articulate such systems to architectural design areas and to open the door from a unique perspective.

Keywords— artificial Intelligence, text-to-image, coloring, interior space

İç Mekân Renklendirmesine Katkı Açısından Metinden Görüntüye Yapay Zekâ Sistemlerinin İncelenmesi

Özet— Bu makalede mimari tasarım süreçlerine potansiyel katkısından hareketle yapay zekânın "metinden görüntüye" sistemleri üzerine bir araştırma yapılmıştır. Araştırmada en yaygın dört sistem olan Craiyon, Dall-E, Midjourney ve Stable Diffusion seçilmiş ve bu sistemler bir okul öncesi eğitim mekânının renklendirmesi için denenmiştir. Öncelikle sistemlere "kindergarten" (anaokulu) metni sunulmuş ve bu metne göre her bir sistemden dörder alternatif görsel üretilmiştir. Sonrasında görsellerin baskın renk kodlamaları bilgisayar ortamında çözümlenmiştir. Elde edilen renk kodları ile bir okul öncesi mekânın 3B modeli renklendirilmiştir. Renklendirme sonucu ortaya çıkan 16 görsel, okul öncesi öğretmenliği ve mimarlık/iç mimarlık bölüm mensuplarından oluşan 62 uzman katılımcıya bir anket eşliğinde sunulmuştur. Ankette uzmanlardan renklendirilmiş görselleri "eğlence" ve "akademik" bağlamlarda değerlendirmeleri istenmiştir. Anket verilerinin istatistiksel analizleri sonucu Craiyon sisteminin bir okul öncesi eğitim mekânının renklendirmesi açısından renkleri diğer sistemlere göre daha başarılı kullandığı izlenmiştir. Bu çalışma ile metinden görüntüye yapay zekâ sistemlerinin metni, mekânın türüne uygun renk kodlarının üretimi açısından yorumlama becerisi ölçülmüştür. Bununla birlikte bu gibi sistemlerin mimari tasarım alanlarına eklemlenmesine, özgün bir açıdan kapı açılmaya çalışılmıştır.

Anahtar Kelimeler— yapay zekâ, metinden görüntüye, mekân tasarımı, mimari tasarım

1. INTRODUCTION

Artificial intelligence, an interdisciplinary concept, involves modeling living and human behavior systems [1]. It can also be defined as a method that mimics human thinking ability and the functioning of the brain through computer simulation, generating solutions by utilizing the necessary information when confronted with a specific situation or problem [30]. For centuries after the Industrial Revolution, humans focused on machine production. With increasing mechanization, artificial intelligence has become one of the leading mottos of the digitized world [2]. Artificial intelligence is a cognitive science that enables the exploration of many "intelligent" ways to model human perception and thought processes. Data sets containing various types of text, image, and audio data are analyzed using artificial intelligence, which achieves its effectiveness through learning and training with algorithms. As the data is analyzed, the algorithms are also trained to produce more accurate content [3].

Artificial intelligence has entered almost all areas of life, such as automotive, healthcare, gaming, entertainment, aviation, e-commerce, and computer-aided translation. Moreover, the use of artificial intelligence in architecture, interior design, and various branches of art is increasing day by day. Today, 2D and 3D digital design tools are complemented by systems such as energy and acoustics calculation programs with artificial intelligence, cost estimation, and environmental simulation [4]. The opening of these AI-powered services to shared virtual worlds, nourished by a multitude of technologies through metadata, is defined as a more advanced objective by experts [31].

This article highlights text-to-image data processing systems, which are a current application area of artificial intelligence, and questions the success of using color in the production of imagery in providing output for architectural design. "The inability of architectural and interior design sciences to keep up with the rapid advancements in various domains of artificial intelligence and assistive systems" has been considered as a problem in this study. Fueled by the concerns stemming from this problem, the hypothesis that "artificial intelligence-supported data processing systems, transitioning from text to image, can contribute to interior spatial coloring in a scientifically accurate manner, thereby enriching the fields of architecture and interior design" has been formulated. Underlying this main hypothesis, subsidiary hypotheses have been articulated as follows: "A method for evaluating the efficiency of prominent systems can be developed, and based on these evaluations, designers can be directed towards suitable systems." At the outset, the term "kindergarten" was entered into four distinct artificial intelligence models, facilitating the generation of four corresponding images to elucidate its interpretation by each system. Following this, RGB color codes were extracted from the cumulative 16 images obtained, and these 16 distinct images were employed to colorize the interior of a 3D model representing a preschool space. Consequently, the same set of 16 images, pertaining to the identical spatial context, was presented to a group of 62 experts consisting of specialized architects/interior architects, and preschool educators. Through a survey employing Likert-type questions, these experts were prompted to assess the volumes in terms of "academic instruction" and "entertainment". The resulting data were then subjected to statistical methods for subsequent analysis and interpretation.

1.1. Data Visualization with Text-to-Image AI Systems

Text to Image can be defined as a "machine learning model" [5]. It refers to computer approaches that can convert human-written text descriptions, such as keywords or sentences, into visually depicted concepts that have the same semantic meaning as the text. [6]. The development of such models began in the mid-2010s as a result of scientific studies in the field of deep neural networks [7]. These models also offer variations of the concept of text-to-image conversion by drawing on various artificial intelligence techniques. These technologies, instead of requiring programmers to make every decision, create their own models by utilizing training data. This allows programmers to set goals and provide data to artificial intelligence for problem-solving, often establishing connections in ways that humans may not envision [33]. In general, in text-to-image models, artificial intelligence interprets the text provided to the system, producing images as close to the description as possible. Yıldırım states that text-image conversion systems have the revolutionary potential for interior design. Thanks to this system capability, faster architectural visualizations, and more effective client communication can be achieved. Moreover, with the accessible and intuitive methods provided by the system, the design process can become more accessible to people who do not have technical skills or expertise [8]. These tools have quickly given rise to various content creation communities, thereby connecting amateur and professional users across different platforms and contributing to the advancement of systems [32].

Today, there are many examples of these systems, such as Craiyon, DALL-E, MidJourney, Stable Diffusion, Motionleap, and DifusionBee [9]. In the context of this article, four of the most well-known text-to-image artificial intelligence systems are examined below.

Dall-E

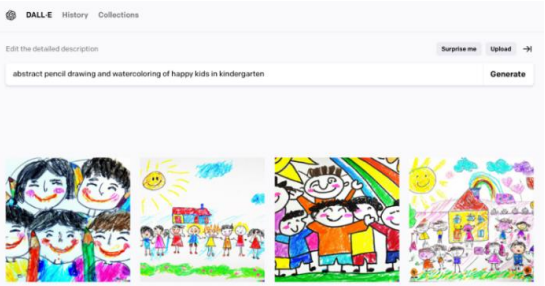

Figure 1. A view from Dall-E results screen [10].

The name of the software is a combination of the names of Wall-E, a Pixar character, and Salvador Dalí, a surrealist artist [11]. Dall-E, announced by OpenAI in January 2021, uses an updated version of GPT-3 for rendering [12]. In April 2022, OpenAI announced Dall-E 2, which can combine concepts, attributes, and styles to produce more realistic images with higher resolution (Figure 1). The company claimed to have trained a neural network that generates images from text titles for a variety of concepts that can be expressed in natural language [13]. The system has also attracted attention from the publishing world, and the cover of Cosmopolitan magazine was designed with an image generated by Dall-E [14]. Dall-E's system uses a process called "diffusion", which starts with a pattern of dots and gradually changes that pattern toward an image as it recognizes certain aspects of the image [15].

Stable Diffusion

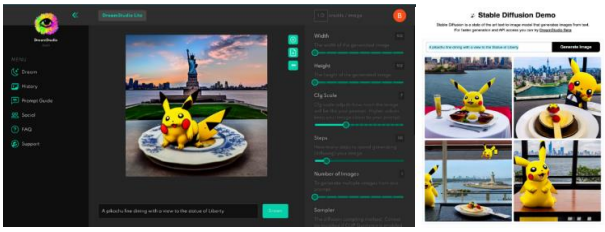

Figure 2. Result views from two beta versions of Stable Diffusion [16].

In the words of Stable Diffusion's developers, it is a text-to-image AI model that will "get billions of people to create stunning works of art in seconds" (Figure 2). Model: AI developer Katherine Crowson builds on the CompVis and Runway teams' work on the widely used latent diffusion model with insights from the conditional diffusion models of Dall-E 2 from Open AI and Imagen from Google Brain. Stable Diffusion is currently being tested at scale with more than 10,000 beta testers generating 1.7 million daily views [17].

Two beta versions are available on the Stable Diffusion website. One can be accessed without membership, while the other is accessible through a detailed interface that is active after membership. Since this is a beta version, only the words that the system predicts should be selected. In this context, it is necessary to use terms such as object, place, space, and color, which are accepted by all but do not encompass everything that is desired.

Craiyon (Dall-E Mini)

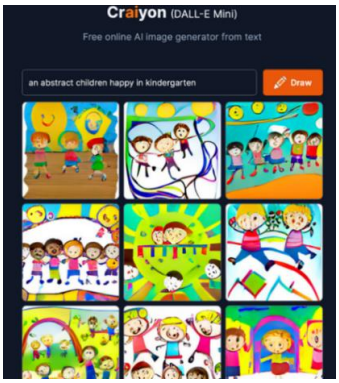

Figure 3. A view from the results screen of Craiyon. [18]

Craiyon, formerly "Dall-E Mini", is an artificial intelligence model that can draw images from any text input [18]. According to OpenAI's explanation of Dall-E Mini, this system was developed to replicate Dall-E's results with a smaller architecture. With this in mind, Craiyon's architecture and memory requirements were simplified and reduced to 27 times the size of Dall-E. The system demonstrates that impressive results can be obtained, albeit of lower quality, when constrained by much smaller hardware resources. Craiyon, which uses the same algorithm as Dall-E, differs only in its interface (Figure 3). After the text is input, it can be printed in about 1 minute [19].

Midjourney

Figure 4. Views from Midjourney's Discord [20].

In the words of Midjourney's developers, it's an independent research lab that "explores new worlds of thought and expands the productive powers of the human species." It was founded in 2021 as a small, self-funded team focused on design, human infrastructure, and artificial intelligence [21]. Midjourney is distinctive in that its artificial intelligence bot is accessible over the Internet through Discord, a social platform for voice protocols and instant messaging, rather than through its web application. The same types of industry terms are used, such as "common models" and "colloquial pre-education (CLIP)" [22].

The system displays four alternative images on the screen after the text is entered through a channel opened in Discord. After this stage, the alternatives are detailed using additional commands (Figure 4).

2. FURTHER WORKS

In his study, Ali Borji quantitatively compares Stable Diffusion, Midjourney and Dall-E 2 in terms of their ability to create photorealistic faces and finds Stable Diffusion more successful than others [23].

In their article, Ploennigs and Berger examined the applicability of diffusion-based models for the early stages of architectural design, which includes the sketching and modeling phase. Comparing Midjourney, Dall-E 2 and Stable Diffusion platforms in this respect, the team analyzed a dataset consisting of 40 million Midjourney queries with NLP methods to extract common usage patterns and revealed how they are currently used [24].

In Dudani's study, metaphors were used in the discovery of artificial intelligence, which are complex systems, and in terms of providing a fun experiment, creating a specific descriptive language in the use of artificial intelligence programs such as Dall-E and Midjourney, exploring the relationship between the participants and how alternative metaphorical visualizations between words and visual representations, multiple and has analyzed how they can affect design designs as prompts to imagine versatile [25].

In his research, Vermillion tested various diffusion model platforms, including Dall-E, MidJourney, and Stable Diffusion, which leverage artificial intelligence to generate architectural concepts, ideas, and images, to iterate conceptual design ideas and question their generative capacity [26].

Abduljawad and Alsalmani discussed the Dall-E 2, Stable Diffusion and Midjourney artificial intelligence models in their article, and they found that the images produced by these models differ due to the difference in software architectures and the data they train. In their study, they found the overall performance of their model positive, but they stated that Dall-E 2 offered the best performance in the tests, followed by Stable Diffusion and finally Midjourney. [27].

3. INVESTIGATION OF TEXT-TO-IMAGE ARTIFICIAL INTELLIGENCE SYSTEMS IN TERMS OF INTERIOR COLORING

In the study, Dall-E, Craiyon, Midjourney and Stable Diffusion, selected from text-to-image artificial intelligence systems, were compared in terms of their success in interpreting colors according to text. Pre-school education space was preferred as the plot area in the research. The colors used by the systems in terms of compliance with the "entertainment" and "academic education" actions taking place in this space were questioned with a survey directed to the users. The methods and steps used in the research are listed below.

Preparation of 3D Model

Figure 5. Pre-school education volume modeled in Max.

The 3D model used within the scope of the research was modeled in the 3dsMax to appeal to the 3-6 age group. In the literature, it is recommended that pre-school education spaces should be at least 2.7 meters high and a minimum of 1.5 square meters should be allocated for each child [28]. Adhering to these limitations in the modeling, the floor height of the space is 3.6 m and the dimensions 7.8x6 m. In the design, there are reading, working and painting spaces, libraries and children's playgrounds used in the academic education of children. In order to expand the space and provide a vertical perspective, a ceiling in the light sky model was designed and suspended objects in the form of cloud/planet were added to this ceiling. The environment is enriched with various pillows, toys and paintings (Figure 5).

Identifying the Texts

Text-to-image artificial intelligence systems, as explained above, work on the basis of sentences and words. Therefore, in order for the systems to present visuals related to the pre-school space, a search for compatible concepts has been made. As a result of the trials, the word "kindergarten" was entered into the systems for getting the ideal visuals.

Resolution of Color Codes

According to the word "kindergarten", 4 alternative images were obtained in each system. Then, another study was started to obtain the RGB codes for the dominant colors preferred by the artificial intelligence in these images. Each image was analyzed using CorelDraw software, and the RGB color codes were listed according to the dominance ratio of the colors. This list is limited to the 5 most dominant colors, matching the number of model types to be overlaid in the 3D model. Below are the images produced with 4 different systems and the dominant RGB color codes (Figure 6).

Figure 6. Images and dominant RGB color codes produced by four different artificial intelligence systems.

3D Scene Coloring

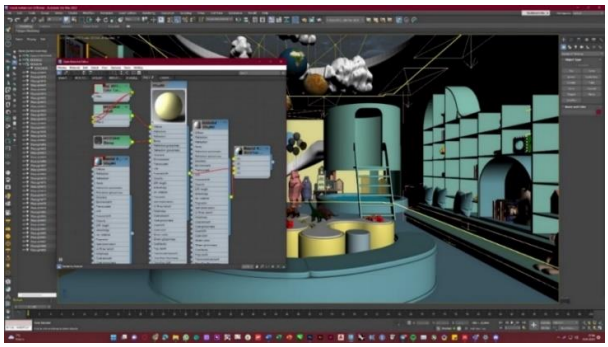

Figure 7. Coloring stage of 3D model in 3ds Max environment

Afterward, the colors obtained from the visuals were transferred to the 3D model as a material color (Figure 7). Thus, a total of 16 views of the same space were obtained from 4 different systems. These views are listed below (Figure 8).

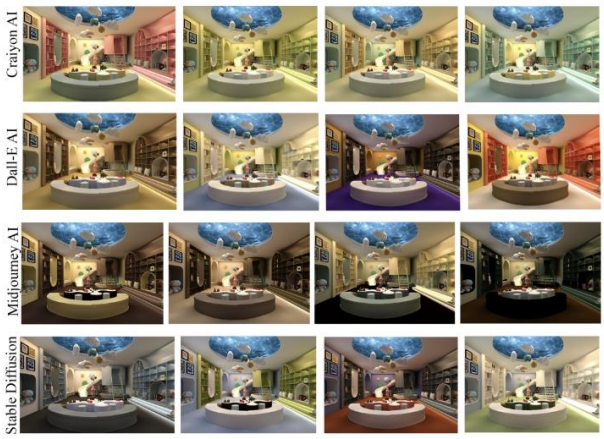

Figure 8. Images and dominant RGB color codes produced by four different artificial intelligence systems.

Survey Study

A survey was organized to evaluate the images obtained after the above preparation steps. The next section will explain the scope, method, evaluation, and interpretation phases of the study.

3.1. Statistical Analysis Study

Scope and Method

The survey study includes a mixed group of graduates and academics in the fields of pre-school teaching and architecture/interior architecture. The appropriateness of the color schemes in the images in the contexts of "education" and "entertainment" for a pre-school educational space was measured using the "relational screening model". For this purpose, an attitude scale was applied to the participants of the study group. 50.8% of the study group are pre-school teachers, and 49.2% are members of the architectural and interior design department. To measure the attitudes of the participants, a total of 32 questions were asked, which consisted of 2 items in 10 Likert-type items associated with each of the 16 images. The result of Cronbach's Alpha reliability analysis of the survey was found to be 0.918. The data obtained in the research were analyzed with the "t-test for independent groups" using SPSS 26 software. Depending on the variables, "arithmetic mean", "standard deviation", and "kurtosis-skewness" data were used in the study in addition to the "t-test for independent samples". Whether the data were suitable for normal distribution was tested with the normality test on the $p<0.05$ plane and based on the hypothesis that the group variances were "homogeneous" with 95% confidence. Thus, parametric tests were performed.

In this context, the attitudes of the participants regarding the suitability of the spaces in the visuals were determined according to their departments. In addition, as sub-dimensions of the scale, the sub-contexts "entertainment" and "academic" were examined to determine which artificial intelligence system the participants considered

more successful in creating the color code. It was also calculated for which artificial intelligence system the departments adopted parallel settings.

3.2. Findings

In Table 1, sample size (N), lowest-highest attitude score (Min-Max), arithmetic mean (Mean), median mean (Median), standard deviation (SD) and skewness-kurtosis values are given.

Table 1. General attitude levels of participants towards visuals.

	N	*Min*	*Max*	*Mean*	*Median*	*ss*	*Skewness*	*Kurtosis*
Participants' attitude scale towards visuals	62	3.38	10	6.7298	6.6250	1.51126	**0.119**	**-0.614**

In Table 1, the "normality" distributions of the data obtained from the normality test were examined, and the values for skewness and kurtosis were presented. According to the results, the skewness value of the data ranged from 0.119 to 0.304, and the kurtosis value ranged from -0.614 to 0.599. According to Tabachnick and Fidell, the data can be considered normally distributed when these values range from -1.5 to +1.5 [29]. In addition, the lowest score that the participants achieved on the attitude scale was 3.38, and the highest score was 10. The mean score of the participants was 6.7298, and the median mean score was 6.6250. According to this value, it can be said that the participants generally have a "close to positive" attitude about the visuals. Afterwards, the "independent sample t-test" was conducted to determine whether there was a significant difference regarding the "department" variable in the attitudes of the participants, whose data were observed to be normally distributed, and the analysis results were shared in Table 2.

Table 2. Results of participants' assessment of attitudes toward visual media in the context of "entertainment" and "academic", determined by independent samples t-test by department.

Variables	*Departments*	*N*	*X*	*ss*	*T Test*		
					t	*sd*	*p*
Craiyon (Entertainment)	Arch./int. Arch	30	6.7333	1.36920	-1.923	62	.049
	Pre-school Teaching	32	7.4609	1.59271			
Dall-E (Entertainment)	Arch./int. Arch	30	5.3250	1.64939	-2.194	62	.032
	Pre-school Teaching	32	6.2656	1.72147			
Midjourney (Entertainment)	Arch./int. Arch	30	3.1500	1.61031	-1.206	62	**.232**
	Pre-school Teaching	32	3.7578	2.27671			
Stable D. (Entertainment)	Arch./int. Arch	30	5.7333	1.48547	-3.019	62	.004
	Pre-school Teaching	32	7.0391	1.88169			
Craiyon (Academic edu.)	Arch./int. Arch	30	5.4583	1.81003	-3.646	62	.001
	Pre-school Teaching	32	7.1875	1.91766			
DallE (Academic edu.)	Arch./int. Arch	30	4.6917	1.34818	-3.308	62	.002
	Pre-school Teaching	32	6.1484	2.02789			
Midjourney (Academic edu.)	Arch./int. Arch	30	3.9333	1.69041	-.171	62	**.865**
	Pre-school Teaching	32	4.0234	2.37031			
Stable D. (Academic edu.)	Arch./int. Arch	30	5.4583	1.74825	-3.387	62	.001
	Pre-school Teaching	32	7.0859	2.01567			

When Table 2 is examined, it is seen that 30 of the 62 participants who participated in the study were from architecture/interior architecture and 32 of them were individuals from the department of pre-school teaching. According to the table, **Craiyon-Entertainment** (t[62]=-1.923; p<.05), **DallE-Entertainment** (t[62]= -2.194; p<.05), **Stable Diffusion-Entertainment** (t[62]= -3.019; p<.05), **Craiyon-Academic education** (t[62]= -3.646; p<.05), **Dall-E-Academic education** (t[62]= -3.646; p<.05) and **Stable Diffusion-Academic education** (t[62]= -3.387; p<.05) significant differences were observed in their attitudes since the p value was less than 0.05. Attitude differences can be followed by looking at the mean (X). In addition, the p values of the participants' attitudes towards **Midjourney-Entertainment** (t[62]= -1.206; p>.05) and **Midjourney-Academic education** (t[62]= -.171; p>.05) are greater than 0.05. No significant differences were observed in the images under these contexts. From here, it was seen that the participants from both disciplines showed a common attitude toward the results of the **Midjourney** system.

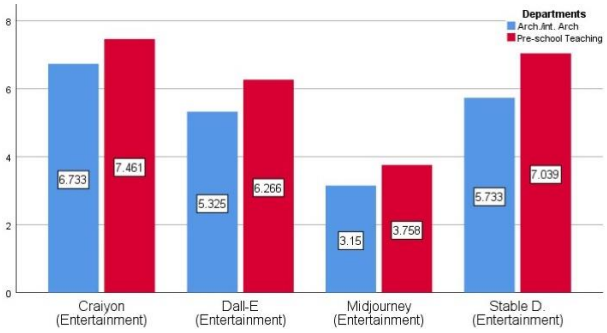

Figure 9. Comparison of results by departments in the context of "entertainment".

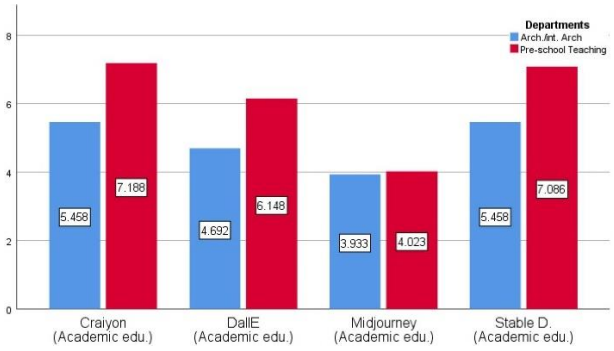

Figure 10. Comparison of results by departments in the context of "academic education".

Table 2 shows the average and overall average data scores resulting from the ratings of 16 different visual displays in the "entertainment" and "science" sections by members of each department. From the values, it can be seen that the members of the Architecture/Interior Design department found **Craiyon**'s color interpretation in the "Entertainment" context more successful than the others, with an average value of 6.7333 (Figure 9). Members of the pre-school education department answered the same context question by finding **Craiyon** successful with an average value of 7.4609.

In the "academic" context query, the members of the architecture/interior architecture department gave their preferences equally weighted to **Craiyon** and **Stable** systems, with an average of 5.4583. In the same context, the attitude of the pre-school teaching department was again the **Craiyon** system with an average of 7.1875 (Figure 10).

Table 3. Average totals of all participants in relation to attitudes toward visual aids by department.

	Craiyon (entertainment)	*DallE (enterta inment)*	*Midjourney (entertainme nt)*	*Stable D. (entertain ment)*	*Craiyon (academic edu.)*	*DallE (academi c edu.)*	*Midjourney (academic edu.)*	*Stable D. (academic edu.)*
Average totals of the departments	**7.1089**	5.8105	3.4637	6.4073	**6.3508**	5.4435	3.9798	6.2984

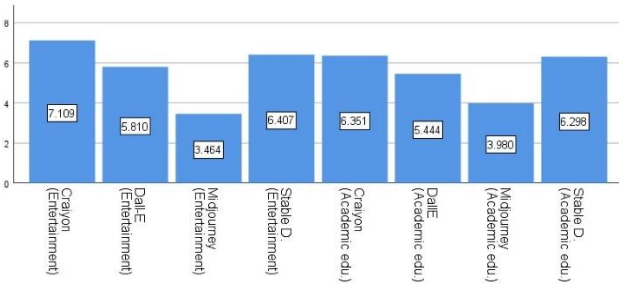

Figure 11. The average of the responses of all participants.

Table 3 presents the response averages of all participants for the four AI systems, regardless of the departments. According to the values, in both contexts, **Craiyon** system seems to be ahead of the others with 7.1089 and 6.3508 averages (Figure 11).

4. CONCLUSION

Artificial intelligence (AI), which has been on the agenda since the 1950s, makes its presence felt even more today. Due to the advancement of machine learning and the availability of data, the interest in artificial intelligence has increased considerably today. Today, artificial intelligence has been used in a wide variety of fields and applications, and has begun to have a significant impact on different industries and branches of science. Simple image variations could be produced with machines that "understand" the language developed by researchers in the 50s and early 60s. Thus, with the contributions of many researchers in the following years, detailed and realistic images began to be obtained with various synthesis methods by using today's deep learning techniques. The developments in artificial intelligence techniques have also attracted the attention of the field of architecture and led researchers to use artificial intelligence in architecture. Thus, artificial intelligence-oriented ways to support the design process by creating design alternatives, analysis/optimization or visualization have been opened.

In the study, four selected models of text-to-image artificial intelligence systems were compared in terms of their success in interior coloring. In the research, the pre-school education space was preferred as the plot area. First, 3D modeling of an experimental training space was performed. Afterwards, 4 images were produced from artificial intelligence systems according to the texts related to pre-school education. After that, the dominant color codes used in the formation of these images were obtained.

Then, the 3D model was colored with these color codes by relating the percentage of intensity of the colors in the images and the space occupied by the reinforcing elements and furnishings in the 3D space. With this approach, 4 views of a preschool education space were obtained with the color interpretation of 4 different systems. The images were presented to experts in the departments of pre-school and architecture/interior architecture, accompanied by a survay, and scored in terms of "entertainment" and "academic education". In both contexts, the images created with the Craiyon system color codes were rated by the participants as more successful than the other systems.

As it is known, "thinking with alternatives" is a very important approach in terms of design. This approach has a great impact on the emergence of richness and diversity in design. In this context, it is predicted that text-to-image artificial intelligence systems will make an important contribution to architectural practice in order to produce new design ideas or concept alternatives.

Today, these systems can be easily accessed from online platforms. Although complex mathematical calculations are performed in the background, the systems welcome users with very simple interfaces. With this access and ease of use, it is thought that every designer can easily use these systems and expand their imagination in this way.

Ploennigs and Berger state that in the future, the capabilities of these models and platforms will evolve, and as suitable interaction paradigms are established, the workflows between tools will converge [24]. In this regard, it is believed that the development of artificial intelligence models will lead to a better understanding of architectural structure and the relational network within contexts such as interior design, color, texture, and construction materials will be more effectively established in the future. It is also anticipated that these models will yield more refined outcomes in the field of architecture and interior design. The data and results obtained through the approach adopted within the scope of the article are thought to have the potential to influence the establishment of this network in a more sophisticated manner.

The systems update the data within the concept of "learning" with artificial intelligence bases. In this sense, text-to-image systems can offer the designer alternative ideas that are enriched day by day. In addition, by interpreting current scientific trends in a sociological framework, it can enrich design alternatives within this framework. In this sense, artistic approaches such as architecture or interior architecture, where intuitiveness comes to the fore, should be followed more closely.

The research conducted by Ali Borji has identified the need for further work in order for artificial intelligence-supported text-to-image translation models to be able to evaluate in detail according to various sub-categories [23]. Considering the diversity of spatial configurations in architecture and potential variations in spatial physiology based on psychic effects, it is envisaged that more sophisticated studies should be carried out for the advancement of AI-supported systems in the context of spatial and spatial psychological development. Furthermore, it is anticipated that the findings of this article will be re-evaluated from this perspective in the future.

Following a pursuit akin to that of this article, various research studies employing comparative methodologies have been examined in the literature, focusing on artificial intelligence systems for text-to-image conversion within the domain of architecture/interior architecture. In these investigations, it has been observed that certain models yield more successful outcomes compared to others [23], [26], [27]. The potential for such comparisons, along with similar ones, to contribute significantly to the advancement of models within their respective contexts is evident. Moreover, the collective evaluation of these studies on a unified platform by system developers is deemed essential. In alignment with this trend, it is considered important to both foster the development of these models and to monitor their implications for the field of architecture.

As mentioned above, this study focuses on the interpretation of colors by artificial intelligence systems from text-to-image. With similar scientific research approaches, it can also contribute to the more qualified development of systems on the axis of architecture.

REFERENCES

[1] Sucu, İ. & Ataman, E., "Dijital Evrenin Yeni Dünyası Olarak Yapay Zeka ve Her Filmi Üzerine Bir Çalışma", *Yeni Medya Elektronik Dergisi*, 4 (1), 40-52, 2020.

[2] Bayrak, E., **Yapay Zekâ ve Mekân Tasarımı Etkileşiminin Günümüz Tasarım Eğitiminde Değerlendirilmesi**. Yüksek Lisans Tezi, Hacettepe Üniversitesi, Güzel Sanatlar Enstitüsü, 2020.

[3] Deveci, M., "Yapay Zekâ Uygulamalarının Sanat ve Tasarım Alanlarına Yansıması", *Vankulu Sosyal Araştırmalar Dergisi*, 9, 119-140, 2022.

[4] Yıldırım, B., & Demirarslan, D., "İç Mimarlıkta Yapay Zekâ Uygulamalarının Tasarım Sürecine Faydalarının Değerlendirilmesi*", Humanities Sciences*, 15(2), 62-80, 2020.

[5] Reviriego, P., & Merino-Gómez, E., "Text to Image Generation: Leaving no Language Behind", *ArXiv Preprint ArXiv*, 2208.09333, 2022.

[6] He, X., & Deng, L., "Deep Learning for Image-to-Text Generation: A Technical Overview", *IEEE Signal Processing Magazine*, 34(6), 109–116, 2017.

[7] Internet: The Economist Newspaper. (n.d.), How a Computer Designed This Week's Cover, https://www.economist.com/news/2022/06/11/how-a-computer-designed-this-weeks-cover, 02.18.2023.

[8] Yıldırım, B. ve Emirarslan, S**., İç Mimarlıkta Yapay Zekâ: İnsana Öykünen Makineler Çağında Yapay Zekânın Mesleki Paydaşlığı, Yapay Zekâ ve Dijital Teknoloji,** İksad Publishing House, Ankara, s.101, 2021.

[9] Internet: Wikipedia, Text-to-image model, https://en.wikipedia.org/wiki/Text-to-image_model, 02.18.2023.

[10] Internet: Dall-E, Dall-E results screen, https://openai.com/dall-e-2/, 02.18.2023.

[11] Şen, E., "GPT3: DALL-E ve JL2P Ekseninde Veri Görselleştirme ve Hareketlendirme Üzerine Bir İnceleme", *USBAD Uluslararası Sosyal Bilimler Akademi Dergisi*, 3(5), 253-280, 2021.

[12] Internet: K. Johnson, OpenAI and Stanford researchers call for urgent action to address harms of large language models like GPT-3, https://venturebeat.com/ai/openai-and-stanford-researchers-call-for-urgent-action-to-address-harms-of-large-language-models-like-gpt-3/, 02.18.2023.

[13] Internet: OpenAI, Dall-E 2, https://openai.com/dall-e-2, 02.18.2023.

[14] Internet: G. Lui, Dall-E 2 made its first magazine cover. https://www.cosmopolitan.com/lifestyle/a40314356/dall-e-2-artificial-intelligence-cover, 02.18.2023.

[15] Internet: Y. Wu, How AI creates photorealistic images from text, https://blog.google/technology/research/how-ai-creates-photorealistic-images-from-text/, 02.18.2023.

[16] Internet: Stable Diffusion, Result views from two beta versions of Stable Diffusion, https://stablediffusionweb.com/, 02.18.2023.

[17] Internet: E. Mostaque, Stable diffusion launch announcement. Stability AI, https://stability.ai/blog/stable-diffusion-announcement, 02.18.2023.

[18] Internet: Craiyon, Results Screen of Craiyon, https://www.craiyon.com, 02.18.2023.

[19] Internet: B. Dayma et. al, Dalle-mini-hugging face, https://huggingface.co/dalle-mini/dalle-mini, 02.18.2023.

[20] Internet: Midjourney, Midjourney's Discord. https://discord.gg/midjourney, 02.18.2023.

[21] Internet: D. Holz, About. Midjourney, https://midjourney.com/home/?callbackUrl=%2Fapp%2F#about, 02.18.2023.

[22] Internet: S. Krishna, S.. Midjourney founder says 'the world needs more imagination', https://venturebeat.com/ai/midjourney-founder-says-the-world-needs-more-imagination/, 02.18.2023.

[23] Borji, A., "Generated faces in the wild: Quantitative Comparison Of Stable Diffusion, Midjourney And Dall-E 2*", ArXiv Preprint ArXiv*, 2210.00586, 2022.

[24] Ploennigs, J., & Berger, M., "AI Art in Architecture". *ArXiv Preprint ArXiv*, 2212.09399, 2022.

[25] Internet: P. Dudani, Wordings and Worldings: DALL-E wordplays to visualise alternative metaphors for complex systems, https://rsdsymposium.org/wordings-and-worldings/, 02.18.2023.

[26] Internet: Vermillion, J., Iterating the Design Process Using AI Diffusion Models, https://digitalscholarship.unlv.edu/cfa_collaborate/9/,02.18.2023.

[27] Abduljawad, M., & Alsalmani, A., "Towards Creating Exotic Remote Sensing Datasets using Image Generating AI", **In 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA)** (pp. 84-88). IEEE, 2022.

[28] Baran, M., Yılmaz, A., & Yıldırım, M, "Okul Öncesi Eğitimin Önemi ve Okul Öncesi Eğitim Yapılarındaki Kullanıcı Gereksinimleri Diyarbakır Huzurevleri Anaokulu Örneği", *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi,* (8), 27-44, 2007.

[29] Tabachnick, B. G. & Fidell, L. S., **Using Multivariate Statistics**. Pearson, 2013.

[30] Partigöç, N. S., "Afet Risk Yönetiminde Yapay Zekâ Kullanımının Rolü", *Bilişim Teknolojileri Dergisi* , 15 (4) , 401-411, 2022.

[31] Huynh-The, T., Pham, Q. V., Pham, X. Q., Nguyen, T. T., Han, Z., & Kim, D. S., "Artificial intelligence for the metaverse: A survey", *Engineering Applications of Artificial Intelligence*, 117, 105581, 2023.

[32] Vartiainen, H., & Tedre, M, "Using artificial intelligence in craft education: crafting with text-to-image generative models", *Digital Creativity*, 34(1), 1-21, 2023.

[33] Fernandez, P., "Technology Behind Text to Image Generators", *Library Hi Tech News*, 39(10), 1-4, 2022.