

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**AN AI-DRIVEN COMPUTATIONAL MODEL FOR EVALUATING OBJECT
MEMORABILITY**



M.Sc. THESIS

Fatma SEVEN

Department of Informatics

Architectural Design Computing Programme

JUNE 2025

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**AN AI-DRIVEN COMPUTATIONAL MODEL FOR EVALUATING OBJECT
MEMORABILITY**



M.Sc. THESIS

**Fatma SEVEN
(523221005)**

Department of Informatics

Architectural Design Computing Programme

Thesis Advisor: Assoc. Prof. Dr. Sema ALAÇAM

JUNE 2025

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**NESNE HATIRLANABİLİRLİĞİNİ DEĞERLENDİRMEK İÇİN YAPAY ZEKA
TABANLI HESAPLAMALI BİR MODEL**

YÜKSEK LİSANS TEZİ

**Fatma SEVEN
(523221005)**

Bilişim Anabilim Dalı

Mimari Tasarımda Bilişim Programı

Tez Danışmanı: Doç. Dr. Sema ALAÇAM

"

HAZİRAN 2025

Fatma SEVEN, a M.Sc. student of ITU Graduate School student ID 523221005, successfully defended the thesis/dissertation entitled “AN AI-DRIVEN COMPUTATIONAL MODEL FOR EVALUATING OBJECT MEMORABILITY”, which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc.Prof. Dr. Sema ALAÇAM**
İstanbul Technical University

Jury Members : **Assoc.Prof. Dr. Orkan Zeynel GÜZELCİ**
İstanbul Technical University

Assoc.Prof. Dr. Özlem KANDEMİR
Eskisehir Technical University

Date of Submission : 30 May 2025
Date of Defense : 26 June 2025





To the people who make my life pleasant,



FOREWORD

In the early phases of this study, the project titled "Modelling Visual Perception of Timevarying Scenes Using Likelihood" conducted by Assoc. Prof. Dr. Sema Alaçam at Istanbul Technical University Scientific Research Projects Department made a significant contribution.

I would like to thank my thesis advisor Assoc. Prof. Dr. Sema Alaçam, who guided me with her positive approach, knowledge and patience throughout my master's degree and helped me get through this process in the best way possible. I would also like to thank the jury members, Assoc. Prof. Dr. Orkan Zeynel Güzelci and Assoc. Prof. Dr. Özlem Kandemir, for their valuable feedback. I thank all my lecturers who supported me throughout my education. I am also grateful to my mother Sadiye Seven, my brother Muhammet Kazım Seven and my entire family who have always been there for me with their patience, love and support throughout this process. I would like to thank my friends and Gülsüm Altunbaş who motivated me during the research process.

June 2025

Fatma SEVEN
(Architect)



TABLE OF CONTENTS

	<u>Page</u>
FOREWORD	ix
TABLE OF CONTENTS.....	xi
ABBREVIATIONS	xiii
SYMBOLS	xv
LIST OF TABLES	xvii
LIST OF FIGURES	xix
SUMMARY	xxi
ÖZET	xxv
1. INTRODUCTION.....	1
1.1 Purpose and Research Questions.....	2
1.2 Methodology of Thesis.....	3
2. THEORETICAL BACKGROUND	5
2.1 Visual Perception	5
2.2 Factors Affecting Visual Perception	9
2.2.1 Size.....	9
2.2.2 Texture	10
2.2.3 Form and shape.....	11
2.2.4 Light and shadow	13
2.2.5 Colour.....	13
2.2.6 Semantic and spatial relations.....	15
2.2.7 Individual differences.....	15
3. EXAMINING VISUAL PERCEPTION THROUGH COMPUTATIONAL MODELS.....	17
3.1 Probabilistic and Statistical Applications.....	19
3.2 VR Analysis	20
3.3 Eye-Tracking and EEG Studies.....	22
3.4 Machine Learning Applications	24
4. METHODOLOGY.....	27
4.1 Generation of the Experimental Environment.....	28
4.2 Development of the Computational Model.....	28
4.3 Evaluation and Implementation of the Model.....	30
5. CASE STUDY.....	33
5.1 Creation of the Experimental Environment.....	33
5.2 Development of the Computational Model and Computation of Scores	35
5.3 Data Collection and Analysis	40
5.4 Findings.....	42
6. CONCLUSIONS AND RECOMMENDATIONS	47
REFERENCES	51
APPENDICES	57
APPENDIX A	58
CURRICULUM VITAE	61



ABBREVIATIONS

3D	: Three-Dimensional
2D	: Two-Dimensional
CNN	: Convolutional Neural Network
DCNN	: Deep Convolutional Neural Network
EEG	: Electroencephalography
FPS	: Frames Per Second
VR	: Virtual Reality
YOLOv8	: You Only Look Once Version 8



SYMBOLS

n	: Number of image frames
A_i	: Bounding box area in the i -th frame
A_t	: Total image area
Sm	: Memorability score based on size
D_i	: Euclidean distance of the detected object
(x_0, y_0)	: Coordinates of the object's center
(x_c, y_c)	: Coordinates of the image's center
c	: Constant



LIST OF TABLES

	<u>Page</u>
Table 3.1 : Related studies and their approaches	17
Table 5.1 : Size ratios for all frames	37
Table 5.2 : Memorability scores based on size	38
Table 5.3 : Centrality values for all frames	39
Table 5.4 : Centrality scores for objects	40
Table 5.5 : Average memorability scores for objects	42
Table 5.6 : Linear regression results	45



LIST OF FIGURES

	<u>Page</u>
Figure 2.1 : Objects with different sizes	10
Figure 2.2 : Large two forms sculpture by Henry Moorean example of abstract forms	11
Figure 2.3 : Elements that form shape	12
Figure 2.4 : Staircase attracts attention with it's colour	14
Figure 4.1 : Methodology of proposed model.....	27
Figure 5.1 : Image taken from the experimental environment	34
Figure 5.2 : Placement of objects and camera movement.....	35
Figure 5.3 : Object labeling process	36
Figure 5.4 : YOLOv8 results after training.....	36
Figure 5.5 : Determining the size ratios of objects	37
Figure 5.6 : Determining centrality of objects	39
Figure 5.7 : Experiment and data collection process	41
Figure 5.8 : Memorability question for suitcase	41
Figure 5.9 : Participants' recall graph.....	42
Figure 5.10 : Memorability scores based on the model	43
Figure 5.11 : Average memorability scores obtained from participants.....	43
Figure 5.12 : Size-based memorability score and survey averages	44
Figure 5.13 : Memorability scores based on centrality	46



AN AI-DRIVEN COMPUTATIONAL MODEL FOR EVALUATING OBJECT MEMORABILITY

SUMMARY

People make their lives meaningful and sustain them with the information they gather from their environment. Perception is a complex process that takes place in milliseconds. In this process, stimuli from the external world are interpreted by the brain and gain meaning. For this reason, perception is not just a process created by the senses. Visual perception, which is the highest level of perception used by humans, is an important factor in architectural design processes that determines the effect of object placement and spatial arrangements on users. It is a subject that should be given priority in fields where shaping user perception is critical, such as the design of museums, exhibition spaces, and educational buildings. There are many factors that affect visual perception. These include size, texture, form and shape, light and shadow, colour, and semantics. In addition to the characteristics of the stimulus, the personal characteristics, cultural background, and experiences of the perceiver also play a role in this process.

With advances in technology, intuitive approaches used to evaluate visual perception have been replaced by analyzable and measurable approaches. In this context, computational visual perception studies have been conducted. These studies have used probabilistic and statistical approaches, VR (Virtual Reality) analyses, eye-tracking and EEG (electroencephalogram), machine learning, and deep learning. In these studies, visual perception evaluate through factors such as memorability, visual attention, recognition, and emotional state.

Remembering is a situation related to awareness and perception. In an environment where there are many objects, it may not be possible to remember all of them even though the light from each object reaches the eye. Studies have shown that there are certain objects that people tend to forget and remember. For this reason, studies are being conducted to examine the features that make objects memorable or forgettable. Previous studies have examined the memorability of objects based on object categories, colour, visual salience, location, and size properties. Studies in the literature have examined object memorability using single images. In this study, computations were carried out taking into account the changing perspective of the perceiver throughout an experience.

The aim of this thesis is to examine the relationship between memorability and object size and location and to quantify visual perception evaluation through a computational model. This thesis aims to contribute to user-centered design processes. The developed model is a tool that can provide feedback to architects during the design process. A comprehensive study on visual perception and computational visual perception was conducted within the scope of this thesis. Based on the information obtained from this section, a computational visual perception model was developed. A case study was conducted with the permission of the Istanbul Technical University Social and Human Sciences Scientific Research and Publication Ethics Committee. The data collected in

the case study and the results produced by the model were compared, and correlation analyses of the model were conducted.

A virtual experimental environment has been prepared for the case study. Nine objects of different sizes have been selected to be placed in the experimental environment. Objects that people see in their daily lives, such as an oven, stool, drawer, suitcase, briefcase, wardrobe, radio, vending machine, and monitor, have been used for the experiment. Texture, color and material factors affecting visual perception are kept equal for all objects.

Screen recording was taken in the designed experimental environment to be shown to participants. These screen recording was also used to develop the model. The screen recording was separated into frames using VLC Media Player. The size and location of the objects in these frames were examined and scored. This process was carried out for all frames. The YOLOv8 object detection model was used in this process. For training the YOLOv8 model, a part of the image frames taken from the experimental environment was used to label objects through Roboflow and create a dataset. After this step, the objects in the image frames were detected. During object detection, some objects in some images were detected incorrectly or were missing. In this case, the computations were carried out manually.

In the first step, it was aimed to examine the relationship between size and memorability. The bounding box area of each detected object was calculated. The size ratios were obtained by dividing this area by the total image area. These values were noted separately for each frame and each object, then transferred to an excel file. By summing these values, the memorability score of the objects depending on their size was calculated.

In the second step, a computation method was presented to examine the relationship between memorability and object location. The center point of the bounding box of each object detected in the image frames and the center point of the whole image were calculated. Then, the distance between these two points was calculated. The calculated value was subtracted from a constant number. Thus, objects located near the center of the image received higher scores, while objects located away from the center received lower scores. Objects not included in the image frame were assigned a score of 0. These values were also calculated and noted for each image frame on an object-by-object basis. The location-based memorability scores were obtained by summing the noted values.

The case study was conducted with the participation of 43 individuals aged between 20 and 30. Before starting, the study was explained to the participants, and their consent was obtained. The screen recording taken from the experiment environment was shown to the participants once. After waiting for 20 minutes, a questionnaire was directed to the participants to measure the memorability of each object. The images of the objects were placed in this questionnaire and participants were asked to rate the memorability of each object on a scale of 1 to 7.

To evaluate the performance of the model, the relationship between the data obtained from the participants and the memorability scores obtained from the model was analyzed using statistical methods. As a result of these analyses, a strong positive relationship was found between object size and participants' data, while no significant relationship was observed between object location and memorability.

This thesis explains the role of visual perception in architectural design processes and how it can be included into this process. The computational model developed in this study makes it possible to predict the effect of object placement in interior spaces on perception. The findings presented regarding object size can contribute to fields such as architectural design, advertising, and education. The study also has some limitations. These include the small number of participants in the case study, the use of abstract representations of objects, and the number of objects used in the experimental environment. Furthermore, the study is limited to the factors of size and location. However, features such as colour, texture, material, and semantics also shape visual perception. This study was made possible through the analysis of image data obtained from a video recording. In future studies, a model capable of real-time computation can be developed. To minimize the effect of individual differences, the study can be conducted with more participants from the same profession.





NESNE HATIRLANABİLİRLİĞİNİ DEĞERLENDİRMEK İÇİN YAPAY ZEKA TABANLI HESAPLAMALI BİR MODEL

ÖZET

İnsanlar çevreden topladıkları bilgilerle yaşamlarını anlamlı kılar ve sürdürürler. Algı milisaniyeler içerisinde gerçekleşen karmaşık bir süreçtir. Bu süreçte dış dünyadan gelen uyaran zihinde yorumlanır ve anlam kazanır. Bu sebeple algı sadece duyuların oluşturduğu bir süreç değildir. İnsanların en üst düzeyde kullandığı görsel algı, mimari tasarım süreçlerinde nesne yerleşimlerinin ve mekânsal düzenlemelerin kullanıcı üzerindeki etkisini belirleyen önemli bir unsurdur. Özellikle müze, sergi alanları, eğitim yapılarının tasarımı gibi kullanıcı algısını şekillendirmenin kritik olduğu alanlarda öncelik verilmesi gereken bir konudur. Görsel algıyı etkileyen birçok unsur bulunmaktadır. Bunlar, boyut, doku, form ve biçim, ışık- gölge, renk ve semantiktir. Uyarının özelliklerine ek olarak, algılayanın kişilik özellikleri, kültürel geçmişi ve deneyimleri de bu süreçte rol oynar.

Teknolojideki ilerlemeler ile birlikte görsel algı değerlendirmesi için kullanılan sezgisel yaklaşımlar yerini analiz edilebilir ve sayısal olarak ölçülebilir yaklaşımlara bırakmıştır. Bu bağlamda hesaplamalı görsel algı çalışmaları yapılmıştır. Çalışmalarda olasılıksal ve istatistiksel yaklaşımlar, sanal gerçeklik ortamının deneyimlenmesine ilişkin analizler, göz izleme ve EEG (elektroensefalografi) verilerinin analizi, makine öğrenmesi ve derin öğrenmeden yararlanılmıştır. Bu çalışmalarda görsel algı değerlendirmesi hatırlanabilirlik, görsel dikkat, tanıma, duygu durumu faktörleri üzerinden değerlendirilmiştir.

Hatırlama nesnenin farkında olmakla, algıyla ilgili bir durumdur. Çok sayıda nesnenin yer aldığı bir ortamda her nesneden gelen ışık göze ulaşmasına rağmen hepsini hatırlamak mümkün olmayabilir. Yapılan çalışmalarda insanların unuttukları ve hatırladıkları nesnelerin ortak olduğu gözlemlenmiştir. Bu gözlem nesneleri unutulmaya veya hatırlanmaya eğilimli kılan özellikleri inceleyen çalışmaların önünü açmıştır. Önceki çalışmalarda nesnenin hatırlanabilirliği nesne kategorileri, renk, görsel belirginlik, konum, boyut özellikleri üzerinden incelenmiştir. Literatürdeki çalışmalar nesne hatırlanabilirliğini durağan görseller üzerinde yapmışlardır. Bu çalışmada bir deneyim süreci boyunca algılayanın değişen bakış açısı dikkate alınarak hesaplama yapılmıştır.

Bu tezin amacı, algılayanın değişen bakış açısını dikkate alarak hatırlanabilirlik ile nesne boyutu ve konumu arasındaki ilişkiyi incelemek ve görsel algı değerlendirmesini hesaplamalı bir model yoluyla gerçekleştirmektir. Bu tez özellikle kullanıcı odaklı tasarım süreçlerine katkı sağlamayı hedeflemektedir. Geliştirilen model, tasarım sürecinde mimarlara geri bildirim sağlayabilecek bir araçtır. Tez kapsamında görsel algı ve hesaplamalı görsel algı üzerine kapsamlı bir araştırma yapılmıştır. Bu bölümden edinilen bilgiler ışığında hesaplamalı görsel algı modeli geliştirilmiştir. Geliştirilen modelin uygulanması için İstanbul Teknik Üniversitesi

Sosyal ve Beşeri Bilimler İnsan Araştırmaları Etik Kurulu izniyle bir vaka çalışması gerçekleştirilmiştir. Vaka çalışmasında toplanan veriler ve modelin ürettiği çıktılar analiz edilip aralarındaki ilişki incelenmiştir.

Vaka çalışması için sanal bir deney ortamı kurulmuştur. Deney ortamına yerleştirilmek üzere farklı boyutlarda 9 nesne seçilmiştir. İnsanların günlük yaşamlarında gördükleri fırın, tabure, şifonyer, bavul, evrak çantası, gardırop, radyo, otomat ve monitör nesneleri deney için kullanılmıştır. Görsel algı üzerinde etkili olabilecek doku, renk ve malzeme faktörleri tüm nesneler için eşdeğer tutulmuştur.

Deney ortamı Spatial.io platformu üzerinden hazırlanmıştır. Bu platform çeşitli mekan şablonları ve avatarlar barındırır. Deney için tasarlanan ortamda seçilen avatar aracılığıyla sanal bir tur düzenlenmiştir. Tur sırasında daha sonra katılımcılara izletilmek üzere ekran kaydı alınmıştır. Aynı zamanda bu ekran kaydı modelin geliştirilmesi için de kullanılmıştır. Ekran kaydı VLC Media Player programı aracılığıyla görüntü karelerine bölünmüştür. Google Colab ortamında Python dili kullanılarak karelerdeki nesnelerin boyut ve konum bilgisi incelenip puanlandırılmıştır. Bu işlem tüm kareler için yapılmıştır. Bu süreçte YOLOv8 nesne tespit modelinden yararlanılmıştır. YOLOv8 modelinin eğitimi için deney ortamından alınan görüntü karelerinin bir bölümü kullanılarak Roboflow aracılığıyla nesneler etiketlenmiş ve bir veri seti oluşturulmuştur. Bu aşamadan sonra görüntü karelerinde yer alan nesneler tespit edilmiştir. Nesne algılama sırasında bazı görüntülerdeki nesnelerin eksik veya yanlış olarak algılandığı gözlemlenmiştir. Bu durumda hesaplama işlemleri manuel yapılmıştır.

İlk aşamada boyut ve hatırlanabilirlik ilişkisini incelemek istenmiştir. Görüntü kareleri üzerinde nesne tespiti yapılmıştır. Tespit edilen her bir nesnenin sınırlayıcı kutu alanı hesaplanmıştır. Bu alanın bütün görüntü alanına bölünmesiyle boyut oranları elde edilmiştir. Bu değerler her kare ve her nesne için ayrı ayrı not alınarak bir excel tablosuna aktarılmıştır. Bu değerlerin toplanması ile nesnelerin boyuta bağlı hatırlanabilirlik puanları hesaplanmıştır.

İkinci aşamada nesne konumu ve hatırlanabilirlik ilişkisini incelemek üzere bir hesaplama yöntemi sunulmuştur. Görüntü kareleri üzerinde nesne tespiti gerçekleştirilmiştir. Görüntü karelerinde tespit edilen her bir nesnenin sınırlayıcı kutusunun merkez noktası ve bütün görüntünün merkez noktası hesaplanmıştır. Daha sonra bu iki nokta arasındaki uzaklık hesaplanmıştır. Hesaplanan değer sabit bir sayıdan çıkarılmıştır. Böylelikle görüntü merkezinin yakınında yer alan nesneler daha yüksek puan alırken, uzakta yer alan nesneler daha düşük bir puan almıştır. Görüntü karesinde yer almayan nesnelerin puanları 0 olarak belirlenmiştir. Bu değerler de her görüntü karesi için nesne bazlı olarak hesaplanıp not edilmiştir. Not edilen değerlerin toplanmasıyla konuma bağlı hatırlanabilirlik puanları elde edilmiştir.

Vaka çalışması 20-30 yaş aralığında 43 kişinin katılımıyla gerçekleştirilmiştir. Çalışmaya başlamadan önce katılımcılara çalışma hakkında bilgi verilip, onayları alınmıştır. Deney ortamından alınan ekran kaydı katılımcılara bir kez izletilmiştir. 20 dakika beklendikten sonra her nesnenin hatırlanma durumunu ölçmek için katılımcılara 9 soruluk bir anket yönlendirilmiştir. Bu ankete nesnelerin görselleri yerleştirilmiş ve her nesne için hatırlanma durumunun 1-7 arasında ölçeklendirilmesi istenmiştir. 1 puan kesinlikle hatırlamama durumunu ifade ederken 7 puan kesinlikle hatırlama durumunu ifade eder.

Modelin başarısını değerlendirmek amacıyla katılımcılardan anket yoluyla elde edilen veriler ve modelden elde edilen hatırlanabilirlik puanları arasındaki ilişki istatistiksel

yöntemler kullanılarak analiz edilmiştir. Bu süreçte katılımcı anketlerinden her nesne için ortalama hatırlanma durumu elde edilmiştir. Anketlerden elde edilen ortalama puanlar ve modelden elde edilen boyut ve konuma bağlı hatırlanabilirlik puanları arasında Pearson ve Spearman korelasyon analizleri ve basit doğrusal regresyon analizi yapılmıştır. Bu analizler sonucunda nesne boyutu ve katılımcı verileri arasında güçlü pozitif ve anlamlı bir ilişki bulunurken, nesne konumu ve hatırlanabilirlik arasında anlamlı bir ilişki gözlemlenmemiştir.

Bu tezde mimari tasarım süreçlerinde görsel algının yeri ve bu sürece görsel algının nasıl dahil edilebileceği açıklanmıştır. Çalışma kapsamında geliştirilen hesaplamalı model yoluyla iç mekandaki nesne yerleşimlerinin algıdaki etkisi öngörülebilir. Nesne boyutu ve konumuyla ilgili sunulan bulgular mimari tasarım, reklamcılık, eğitim gibi alanlara katkı sağlar. Çalışmanın bazı kısıtları vardır. Bunlar; vaka çalışmasındaki katılımcı sayısının az olması, nesnelerin soyut temsillerle kullanılması ve deney ortamında kullanılan nesne sayısıdır. Ayrıca çalışma boyut ve konum faktörleri ile sınırlandırılmıştır. Oysaki renk, doku, malzeme gibi temel özellikler ve semantik gibi yüksek seviyeli özellikler de görsel algıyı şekillendirmektedir. Nesne tespiti bazı görüntü karelerindeki nesneleri tespit edememiştir veya yanlış tespit yapmıştır. Bu çalışma bir video kaydından alınan görüntü verilerinin analizi ile mümkün olmuştur. Gelecek çalışmalarda gerçek zamanlı hesaplama yapan bir model geliştirilebilir. Bireysel farklılıkların etkisini azaltmak için çalışma aynı meslek gruplarına sahip daha fazla katılımcıyla gerçekleştirilebilir. Ayrıca YOLOv8 nesne tespit modelinin eğitimi için daha geniş bir veri seti hazırlanabilir, böylelikle modelin performansı iyileştirilebilir.



1. INTRODUCTION

People perceive stimuli from the external world through perception and sustain their lives through it. It is known that perception involves the processes of stimuli from the environment reaching the sensory organs and being interpreted and understood in subjective consciousness. In this process, our sensory organs, nervous system, and experiences play a role. For this reason, perception is an active evaluation rather than passive observation (Goldstein & Brockmole, 2002). It has been stated that the information obtained through the eyes, ears, or other sensory organs is interpreted in the brain based on the individual's culture, personality characteristics, and experiences, and for this reason, perception is a multidimensional formation influenced by many factors (Castelhano & Krzyś, 2020; Yılmaz, 2004). Two people in the same environment might act differently and make different inferences based on their interests, the aspects they have developed, or their purpose for being in that environment.

It is common knowledge that humans obtain a significant amount of information about their environment through their eyes. In addition, visual perception is the most effective form of perception, accounting for approximately 90% of an individual's perception of their environment (Ardıçoğlu, 2024; Özçevik & Yüksel Can, 2013). It has been stated that the area of the brain linked to visual perception is larger than the total area responsible for other types of perception (Yılmaz, 2004).

In order to shape an individual's experience within a space, it is necessary to have a deep understanding of the concept of perception and to interpret this concept in the design process. It is known that architectural spaces are perceived mostly through visual perception. Therefore, visual perception and the factors that shape it have a leading role in creating spatial perception (Ardıçoğlu, 2024). The way in which elements that form the space come together, their characteristics such as colour, size, texture, material, shape, and semantic properties, have an effect on visual perception (Małgorzata & Michał, 2018). Predicting how visual perception changes according to

these characteristics enables the user's perception to be shaped. In order to transform this prediction from an intuitive dimension into a measurable form, computational visual perception models are required. In line with this need, studies are being conducted using probabilistic and statistical methods, VR (Virtual Reality) analyses, EEG (Electroencephalography) and eye-tracking, and machine learning. In this thesis, visual perception is studied using a computational model. The reason for this is to handle the visual perception process, which involves complexity and subjectivity, in a quantitative way and to make the effect of the object on visual perception measurable.

Visual perception, memorability, and attention are interconnected concepts. The memorability of an object is related to its perception. Studies have also observed that memorable objects are attention-grabbing (Çelikkale et al., 2015; Çiftcioglu & Bittermann, 2013). For this reason, visual perception evaluation was conducted in relation to the memorability of objects within the scope of this thesis.

1.1 Purpose and Research Questions

After visiting an exhibition, people don't remember all the artworks. After walking down a street, the only thing people can remember is a large poster on a building. After looking at a shop window, people can realize that they only remember a few objects. An environment contains many objects. Even if the light reflected from all of these objects reaches the eye, it may not be possible to remember all of them in every case. Because just seeing is not enough to remember; remembering is related to perception (Çelikkale et al., 2015; Çiftcioglu & Bittermann, 2013). Perception is the interpretation of sensory data in the mind, and not all objects are equally prominent in this process. Which objects in a visual scene are more easily forgotten or remembered depends on the characteristics of the objects and the perceiver. Although this process is subjective, studies show that the objects people remember and forget in a scene are mostly the similar. This finding has led to studies into the features that make objects memorable or forgettable.

The first study on object memorability in images was conducted by Dubey et al. (2019). In this study, the researchers created a dataset to examine object memorability and used this dataset to analyze the effects of visual factors such as object category,

shape, color, and salience on memorability. The study found that the effects of these factors on memorability, from weakest to strongest, were: shape and color, salience, and object category. Basavaraju et al. (2019) later analyzed this dataset and predicted that object size and location could also be important factors affecting memorability. Their analyses revealed that the size and location of an object in the visual field significantly increased memorability estimation performance. In addition, Turgay Ziraman and İmamoğlu (2020) revealed in their study on exhibition objects that the size and location of an object have a decisive effect on visual attention. In light of these findings and due to the limited research on this subject in the literature, this study examined the effect of object size and location on memorability.

In previous studies on memorability, scene analyses were conducted on single images and the memorability of objects in the scenes was examined. This study aims to evaluate visual perception in the process of experiencing a space. In line with a similar approach by Alaçam (2019), the participant's changing field of vision is taken into account. The purpose of this thesis is to reveal the effect of object size and object location on memorability in the process of experiencing a space. In line with this aim, a computational model that evaluates visual perception through memorability has been designed. The study's research questions are as follows:

- During the experience of a space, are larger objects in an individual's field of vision remembered more?
- During the experience of a space, are objects located at the center of an individual's field of vision remembered more?

1.2 Methodology of Thesis

The methodology of the thesis consists of the following steps: literature review, developing the computational model, implementing the model through a case study, collecting data from participants, analyzing the relationship between participants' data and model's scores, and presenting the findings.

This thesis consists of six chapters. The first chapter describes the subject, purpose, hypotheses, and methodology of the thesis.

The second chapter of the thesis focuses on the concept of visual perception. The role of this concept in architecture is discussed. Additionally, the factors influencing visual perception are examined within the scope of the characteristics of the perceived and the characteristics of the perceiver. In the context of the perceived's features, size, texture, form and shape, light and shadow, colour, and semantics are analyzed.

In the third chapter of the thesis, studies using computational methods related to the evaluation of visual perception are examined. The contributions of these studies to the model developed within the scope of the thesis and the similarities and differences of the model are explained.

In the fourth chapter, the method of the computational model developed in line with the thesis objective is explained. In this chapter, information about the tools and environments used is also presented. The study was conducted by setting up a virtual experimental environment. An object detection model was used in the study, and computational methods for object size and location were developed.

The developed model was implemented through a case study. The fifth chapter describes the details of this case study. It covers the setting up of the experimental environment, the steps for generating the model, the computation of scores, the process of obtaining data from participants, and the data analysis process. Finally, the findings of the study are presented.

In the last section, the study is summarized. The limitations of the study and future studies are discussed.

2. THEORETICAL BACKGROUND

2.1 Visual Perception

People make sense of and continue their lives through the information they obtain from their environment. The interpretation and understanding of information obtained through the senses in the mind is called perception (Goldstein & Brockmole, 2002). Seeing a cat a short distance away and deciding to close or move away from it happens in a very short period of time. Because perception occurs in milliseconds, many people do not consider this process to be complex. However, in the 1950s, computer scientists understood that there was a misunderstanding on this subject. They realized that it would take a long time to design machines that could easily perceive the environment (Goldstein & Brockmole, 2002).

The relationship between the actual environment and the perceived environment has been discussed since perception results from the subjective interaction between an individual and their surroundings (Özer, 2004). As a consequence, different definitions of perception have emerged in the fields of philosophy, psychology, and art. Perception has been defined in philosophy as the main way to reach the world outside. This definition raises questions about how subjective or objective our perceptions are and how our senses represent reality (Morales & Firestone, 2023; Tural, 2017). In psychology, perception is defined as a dynamic process in which the brain organizes and interprets the information it receives from stimuli, making it meaningful. This process is not limited to sensory data; emotions, memories, and other cognitive factors also play an important role in this process. It has also been defined as an essential function that enables an individual to understand their environment and interact with it (Goldstein & Brockmole, 2002). In art, perception has been studied in close relation to subjective experiences. Perception is defined as an active process of interpretation and emotional response to sensory data. In this definition, the perception process is at the center of the art experience, and the perceiver's mind plays a role in the construction of meaning (Beyoğlu, 2015).

Studies generally consider the five senses: sight, hearing, touch, smell, and taste (Erişti et al., 2013; Reghukumar, 2019; Özer, 2004; Wu & Zhao, 2023; Yu et al., 2018). Different types of perception are affected by the stimulation of sensory organs. The process that enables the information obtained through the sense of sight to be interpreted is defined as visual perception (Erişti et al., 2013). Visual perception is the most dominant type of perception during the perception of the environment. The fact that the total area of the brain dedicated to senses other than visual perception is smaller than the area dedicated to visual perception is also related to this situation (Özer, 2004). Light is essential for visual perception to occur (Yılmaz & Kaya, 2023). Visual perception encompasses physical, physiological, and psychological processes. The physical process involves light from the environment reaching the eye, the physiological process involves the refraction of light and the image reaching the fovea, and the psychological process involves the perception of the image in the brain (Akkurt, 2019). In his book, Gordon (2004) mentions six important theories that aim to explain this multi-layered process. These are Gestalt theory, Brunswick's Probabilistic Functionalism, Direct Perception and Gibson's Ecological Optic, Empiricism, Neuropsychological Approach, and Marr's Computational Approach.

The Gestalt theory, which began with Max Wertheimer in the 1900s, was developed with contributions from Kurt Koffka and Wolfgang Köhler. It is based on the whole and is known as the pioneer of contemporary perception studies (Özer, 2004). This theory is most simply expressed by the sentence the whole is beyond the sum of its parts. Perception is not considered as a part affected by a single factor, but rather as a whole affected by many factors (Çağlayan et al., 2014; Erişti et al., 2013). Perception requires more than single information from the senses, according to this theory. The perceived object does not create perception on its own. In order for perception to occur, the brain interprets the information obtained (Eryayar, 2011).

Brunswick proposed the Probabilistic Functionalism theory regarding visual perception. The theory focuses on how individuals obtain and process a large amount of environmental information cognitively and perceptually (Yılmaz & Kaya, 2023). In this theory, the characteristics of the perceiver, such as gender, age, and cognitive state, are important. According to this theory, the expectations of the perceiver should shape the design. The theory also emphasizes that the object characteristics offer clues about how they are perceived (Erişti et al., 2013; Gordon, 2004; Brunswick, 1955). The

concept of uncertainty is discussed in this theory. The perception of an object depends on its characteristics, the environment in which it is located, the mediator between the perceiver and the object, and the perceiver. Individuals perceive information from the environment by selecting clues that they find useful from their own perspective. Therefore, perception is shaped by the individuals' personal perspectives and the conditions they are in (Erişti et al., 2010).

The Direct Perception and Ecological Optics Theory proposed by James J. Gibson and Eleanor Gibson in 1950 suggests that individual perception consists of two parts. These are schematic perception and literal perception. Literal perception is the direct perception of an object or environment as it is. Schematic perception, on the other hand, is concerned with the perception of objects that have the potential to be useful or meaningful to the perceiver (Özer, 2004; Gibson, 1954). Perception is studied from an environmental point of view in the ecological approach. Gibsons argues that perception is affected by relationships and the organization of the environment in addition to sensory data. This approach states that people get an understanding of their surroundings by directly experiencing it and that environmental context has a significant influence on perceptual processes. The Gibsons included a hypothesis for the objects of perception in addition to a theory about how perception occurs (Withagen & van der Kamp, 2010; Erişti et al., 2010). They integrated colour, texture, shape, and surface differences in the perception process and emphasized the factor of movement. This theory attracted the attention of architects for including the physical characteristics of objects and making perception more understandable (Özer, 2004).

One of the theories presented in Gordon (2004) book is the Empiricism. This theory is based on criticisms of the passive role of the individual in Gestalt Theory. In this theory, the meaning given to an image is explained by examining the eye movements of the individual during the perception process. It is stated that the perception process is related to the individual's cognitive processes (Gordon, 2004). This process is active, and individual differences such as knowledge, motivation, mental and cognitive skills play a role (Yılmaz & Kaya, 2023). According to this theory, the perceptual process is shaped by the construction of the stimulus based on prior knowledge and experience. Although it is possible to have many stimuli in an environment, the stimuli that the perceiver associates with his/her past and experiences are perceived (Erişti et al., 2016; Shankar, 2007). From a design perspective, this view can be interpreted as considering

the interaction between the perceiver and the perceived and creating interactive designs for the user.

In the Neurophysiological Theory of Perception, it is stated that memory, perception and thinking are carried out through the central nervous system. This theory examines learning, memory, and how learning occurs through neurons and the connections between neurons. It is argued that learning is possible through interactions between neurons and that knowledge becomes permanent and automatic through the repetition of connections formed as a result of these interactions. (Yılmaz & Kaya, 2023; Gordon, 2004). According to this theory, the brain's response to a stimulus is related to the information stored in the individual's memory. In this perception theory, the perceiver is prioritized over the perceived. Perception is constructed not based on the reality of what is perceived, but on the realities defined by the perceiver (Erişti, 2016; Gordon, 2004). Based on the neurophysiological approach it can be said that design should be composed in a way that activates the information in memory and attracts the individual's attention (Yılmaz & Kaya, 2023).

The computational perception approach proposed by David Marr focuses on the activities of neural structures during perceptual processes and the concept of integrity. The colour, shape, movement, and similar physical properties of objects, instead of being evaluated separately, come together to form a whole according to individuals' cognitive properties (Erişti et al., 2013; Gordon, 2004). According to this theory, an individual's characteristics, particularly memory, integrate with the design elements of the environment to form a structure. The meaning is derived from this structure (Yılmaz & Kaya, 2023; Gordon, 2004). This approach argues that perception and vision are information processing processes. This is reflected in the design field by taking into account the user's past experiences. In addition, this theory states that visual information is first processed at a low level, then recreated as a representation of the object and scene, and integrated in a way that is appropriate for perception. From this, an inference can be made about simplicity and clarity (Erişti et al., 2010). Finally, the understanding that objects need to be surrounded by distinct boundaries in order to be easily distinguished and identified emphasizes the use of delimiting elements such as borders and lines in design (Erişti et al., 2010).

2.2 Factors Affecting Visual Perception

The environment of an individual consists of many objects. However, not all objects receive the same amount of attention. Because some objects are more prominent in visual perception. Visual perception of objects is affected by the properties of the perceived object, the environment, and individual differences. While colour, texture, size, form, light-shadow, and semantics are included in the group of factors related to the object and environment, individual differences include many factors such as experience, age, emotional and psychological state.

2.2.1 Size

In the architectural design, object size plays a formative role in shaping spatial organization and is one of the factors that influence visual perception. Various studies have been conducted in the fields of architecture, psychology, and computer science to examine the effect of size. The common findings of these three disciplines indicate that an increase in object size enhances visual perception. The larger objects in an environment are more memorable and attention-grabbing than smaller objects (Basavaraju et al., 2019; Castelhana et al., 2020; Dubey et al., 2015; Turgay Zıraman & Imamoglu, 2020). In psychology, visual perception and size have been studied to understand cognitive processes such as spatial awareness, memory, and attention (Castelhana et al., 2020). In computer science, the size factor has been investigated for image processing, user interface designs, and the creation of human-like perception systems (Dubey et al., 2015).

In architecture, it is important to investigate the size factor for the design of commercial, educational, exhibition, retail spaces and facade design. Figure 2.1 shows objects of different sizes in an exhibition. Larger objects in the environment attract more attention (Lisińska-Kuśnierz et al., 2018). In interior design, the size of objects has a great impact on spatial hierarchy. The concept of size brings with it the concept of perceived size. Objects can be perceived as smaller or larger than they are depending on the context. Perspective, light intensity relationship and nearby object sizes create changes in perceived size (Turgay Zıraman et al., 2020). The human brain evaluates objects comparatively and makes sense of them. Reference scales such as doors, windows or other objects in the environment affect perceived size. In addition, the distance of the object to the observer also affects perception. Close objects are

perceived as larger, while distant objects are perceived as smaller (Bittermann & Ciftcioglu, 2008). As a result, there is a multi-layered and complex relationship between the size factor and visual perception. It is important for the designer to consider the size factor with this awareness and use this factor effectively in the design process. This allows for design decisions that direct and enrich the user experience and the development of aesthetically and functionally powerful spatial solutions.



Figure 2.1: Objects with different sizes.

2.2.2 Texture

The materials used in design affects the perception of the experiencer. There are two types of texture: Natural and artificial. Tree bark, leaf veins, animal skins are examples of natural textures (Aslan et al., 2015). Plastic reliefs and concrete mold marks produced by humans are examples of artificial textures. In terms of surface properties, two types of textures are mentioned: soft and hard. Surface properties also have a psychological effect on people. Soft surfaces create a more inviting and positive effect, while hard surfaces create a negative and more distant effect (Ör & Özen, 2024). In addition, texture affects the perceived colour. Tone differences are seen due to texture differences. Texture differences also affect the distance-closeness relationship. Since soft textures are perceived as farther away, spaces where these textures are used feel larger. On the contrary, hard textures are felt closer and spaces where these textures are used are perceived as smaller (Özer, 2004).

2.2.3 Form and shape

The concepts of form and shape are also among the factors that affect visual perception (Basu & Ghosh ,2017). Although the concepts of form and shape generally seem synonymous, form differs in that it is more comprehensive in space design. Form is defined as the mass expression of an object (Ör & Özen, 2024). In other words, form is the general appearance of each entity reflected to the outside, depending on the internal and external boundaries (Demirarslan, 2006). The concept of form is examined in three different groups: symmetrical and asymmetrical, organic and inorganic, abstract and concrete (Ör & Özen, 2024).

Forms perceived by the senses are called concrete forms, while forms perceived through thought are called abstract forms. Therefore, abstract forms are open to individual interpretation, while concrete forms convey more general meanings (Yüce, 2018). Figure 2.2 shows an example of an abstract form. Forms are also classified as organic and inorganic. Organic forms are explained as a movement-based reflection of wholeness and unity. Organic forms are related to interior organization characterized by curved and circular shapes (Ör & Özen, 2024). Inorganic forms, on the other hand, are described as rigid, angular forms. In the final group, forms were analyzed based on their symmetrical structure. Symmetrical forms are defined as more monotonous, balanced, and stable structures. In contrast, asymmetrical forms are defined as more unbalanced and dynamic structures with a dynamic energy point (Ör & Özen, 2024).

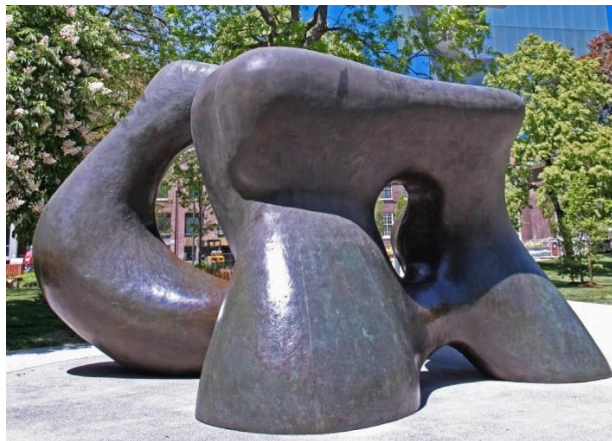


Figure 2.2: Large two forms sculpture by Henry Moorean example of abstract form.

(Url-1)

The shape is defined as the sum of all the physical properties of an object perceived in a characteristic order (Yüce, 2018). Shape is related to the change in form over time. For example, if the static image of a moving object at its initial position is referred to as form, then all the positions it passes through while moving are referred to as shape (Yüce, 2018).

Point, line, surface, and volume are the elements that form shapes (Figure 2.3). The point is the smallest element that forms a shape. Other elements that form shapes are created by combining points. Points have a dynamic structure. For instance, points can create different effects when used in light or dark tones, big or small, random or organized (Yüce, 2018). Lines are formed when points put together side by side. It is defined in geometry as a one-dimensional shape with no depth or width. Lines come together to create edges and surfaces. They also define the boundaries of objects. Depending on the characteristics of the lines that define the boundaries of an object, shapes with soft, rounded, sharp, and pointed characteristics are formed. Different uses of lines create different psychological effects on people. This difference is most evident between straight and curved lines. Curved lines give a more natural impression, while straight lines feel artificial and man-made. Vertical and horizontal lines evoke a sense of calm, while broken lines create a more energetic and dynamic atmosphere; slanted lines can increase the effect of movement depending on their use, creating a sense of liveliness and naturalness. The surface is formed by the combination of lines. Surfaces are classified into two categories: curved surfaces and flat surfaces. Volume, the final element that forms shape, is the space occupied by an object in relation to its height, width, and length. In architecture, it is defined as the body that emerges by covering the space (Şensoy, 2024).

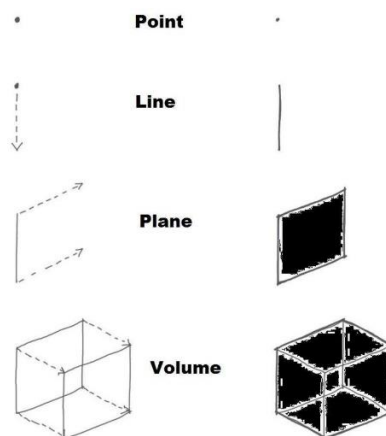


Figure 2.3: Elements that form shape. (Url-2)

Different shapes invoke different meanings in art and nature. For example, narrow-angled shapes create a sense of movement, while rectangular shapes convey a clear and determined expression. Geometric curves evoke a sense of continuity. Free forms emphasize volumetric integrity and continuity (Yüce, 2018; Özer, 2004).

2.2.4 Light and shadow

The first condition for the emergence of visual perception is the presence of light. Light makes it possible to perceive the colour, texture, and shape of objects. Light makes details visible or hides them. The atmosphere of a space, its three dimensions, and depth can be perceived through light (Jaglarz, 2011).

The concept of light brings with it the concept of shadow. A shadow is defined as the dark area formed behind an object by blocking the passage of light. Shadows are also an important factor in the perception of objects. Shadows create depth and dimension in shapes (Akkurt, 2019; Özer, 2004). Light sources can be classified into two categories: natural or artificial. Examples of natural light include sunlight, while artificial light sources include objects that emit light produced by humans (Aslan et al., 2015). Different parts of an object may appear brighter or darker depending on their distance to the light source. The part of the object where light hits directly and most intensely is called the light tone, the part where light hits least is called the dark tone, and the area between these two regions is called the midtone (Yüce, 2018).

The amount of light (low-high) and its temperature (warm-cold) as well as the way it comes in (direct-indirect) have different effects on visual perception. For example, using bright light directly creates high contrast. This makes the distinction between black and white more obvious and increases the perceived volume (Aslan et al., 2015). Light that filters through and does not strike the surface directly creates a sense of calm, introspection, and isolation (Yüce, 2018).

2.2.5 Colour

Colours are sensations that affect the eye depending on the structure of light rays and their reflection from objects (Castelhano & Henderson, 2008). Colours are represented in physics by different wavelengths. Every object reflects or absorbs the colours coming from the sun in its own unique way (Özer, 2004). The colours of spaces or

objects have a psychological effect on people. For example, pale and pastel colours can make people feel calmer, while bright and vibrant colours can make them feel more excited and cheerful (Ör & Özen, 2024). When people see colours, they first associate them with coldness and warmth. Cold colours lower blood pressure, while warm colours raise it, thereby affecting the human body physiologically.

Colours evoke different meanings for people. Blue is a cool colour and brings to mind concepts such as purity and calmness. It also increases creativity while calming the nerves and providing peace of mind. Orange and yellow have a warm and energizing effect. Purple evokes mystery. As the colour of the unknown, it can create a tension effect when used over a wide area. White is interpreted as representing abstraction, purity, and intelligence (Özsavaş, 2016). Colours can be used to define spaces, emphasize objects in a space, or reduce their impact. These processes are achieved through colour differences. In the example in Figure 2.4 the staircase stands out with its colour and form.



Figure 2.4: Staircase attracts attention with it's colour. (Url-3)

2.2.6 Semantic and spatial relations

Semantic information enables the brain to interpret images. Beyond only visually perceiving the objects, or events in a scene, it enables us to understand their functions, meanings, contextual appropriateness, and relationships (Shomstein et al., 2019). For example, recognizing that a spherical object is a ball and associating it with a playing concept in the brain is related to semantics.

Spatial relationships and semantic effects in visual perception have been studied, mainly in scene analysis studies. These studies conducted over the past 60 years have greatly contributed to the understanding of scene perception (Castelhano et al., 2020). The effects of scene context and object's relevance on visual attention, visual search, and memory have been investigated in these studies. These studies have observed that knowing the function of objects improves visual search performance. At the same time, objects that are appropriate to the scene context are found more quickly in the visual search process (Castelhano et al., 2020). For example, a bowl in a kitchen image is easier to find than a bowl in a store image. In addition, it has been observed that objects with spatial relationships between them are more easily remembered. These findings indicate that examining physical elements alone is not sufficient to analyze the visual perception of objects and environments.

2.2.7 Individual Differences

The process of visual perception depends on the individual. Gender, age, cultural background, education, experience, and emotional state all influence how objects, places, and environments are perceived (Wade & Swanson, 2013). In studies on visual attention, the effects of individual differences on age (Yoshida et al., 2017) and gender (Kim et al., 2021; Yoshida et al., 2017) were examined and it was revealed that they play a role in the attention process. People's experiences give them visual selectivity. In addition, people tend to acquire information that is relevant to their goals (Özer, 2004).

One person may find it enjoyable to look at a busy city view from a high-rise building, while another may prefer a quieter house with a natural view. The interest in an object seen every day is different from the interest in an object experienced for the first time (Özer, 2004). A person visiting Istanbul's historic peninsula to explore would prioritize

historical buildings, while someone visiting the same place to eat would have a focus on restaurants. This situation is also valid for the perception of objects. Within the same store, someone interested in visual arts would be drawn to brushes, while someone interested in music would be attracted to musical instruments.



3. EXAMINING VISUAL PERCEPTION THROUGH COMPUTATIONAL MODELS

The visual perception process is a complex process that involves subjectivity. Different computational methods have been used in the fields of design, computer science, and psychology to make this process measurable and analyzable (Seven & Alaçam, 2025). These studies have examined visual perception in virtual environments, physical environments, or through images. In the studies, visual perception has been evaluated by analyzing emotional state, visual search, memorability, and visual attention. These studies contribute to the development of the field through the methods they use and the findings they reveal. Within the scope of this thesis, computational visual perception models in the literature were examined and divided into four groups according to their approaches. These groups are as follows: Probabilistic and statistical approaches, virtual reality analyses, eye tracking and EEG studies, and machine learning applications. The following sections of this thesis describe the studies related to these four topics. 3.1 shows the reviewed studies and their approaches.

Table 3.1: Related studies and their approaches.

Study Name	Authors-Year	Approach
Fusion of perceptions in architectural design	Bitterman et al., 2013	Probabilistic Approach
Visitor attention in exhibitions: the impact of exhibit objects' ordinal position, relative size, and proximity to larger objects	Turgay Zıraman et al., 2020	Statistical Application
Revealing Visual Perception Threshold Regarding Object Size and Visibility in Spatial.io Environment	Seven, 2024	Statistical Application
The Impact of Color and Object Size on Spatial Cognition and Object Recognition in Virtual Reality	Diaz et al, 2024	VR Analysis
Memorability of Spatial Features In Virtual Reality Süleymaniye Experience	Şen Bayram et al., 2023	VR Analysis

Table 3.1 (continued): Related studies and their approaches.

Study Name	Authors-Year	Approach
The method of loci in virtual reality: explicit binding of objects to spatial contexts enhances subsequent memory recall	Reggente et al., 2020	VR Analysis
Determination of the effects of visual and auditory stimuli on spatial perception in enclosed public spaces using fuzzy logic	Cantemir, 2025	EEG & Fuzzy Logic
Towards emotionally intelligent buildings: A convolutional neural network based approach to classify human emotional experience in virtual built environments	Zou et al., 2023	CNN & EEG
Neural representation of different 3D architectural images: An EEG study	Grima et al., 2019	EEG Analysis
Assessing consumer attention and arousal using eye-tracking technology in virtual retail environment	Kim et al., 2021	Eye Tracking Analysis
Visual attention in retail environments: design analysis using HMD based VR system integrated eye-tracking	Kim et al., 2020	Eye Tracking Analysis
Eye tracking in research on perception of objects and spaces	Lisińska-Kuśnierz et al., 2018	Eye Tracking
The features underlying the memorability of objects	Kramer et al., 2023	Deep Learning
Object memorability prediction using deep learning: Location and size bias	Basavaraju et al., 2019	Deep Learning
Predicting memorability of images using attention-driven spatial pooling and image semantics	Celikkale et al., 2015	Deep Learning
What makes an object memorable?	Dubey et al., 2015	Deep Learning
What makes an image memorable?	Isola et al., 2011	Machine Learning

3.1 Probabilistic and Statistical Applications

The difficulty of modeling human visual perception is that the brain plays a role in the perception process, not only the eye. For this reason, even though light reflected from every object on a scene hits the eye, it may not be possible to perceive or remember every object. This becomes particularly evident when there are more than seven objects (Çiftcioglu & Bittermann, 2013). However, the unremembering of an object does not mean that it was not seen. The relationship between seeing and remembering is complex and indistinguishable. The object that is paid attention to and remembered is also related. It is difficult to distinguish between attention, perception, and remembrance. For this reason, an isovist or visibility graph is not sufficient for a perception model. Some studies in the literature have used a probabilistic method for modeling perception, considering it appropriate for the complexity of the human visual perception process (Bittermann & Ciftcioglu, 2008).

Bittermann & Ciftcioglu (2008) applied this approach in their study. The relationship between distance and perception was illustrated with two architectural applications. The effect of geometry in the early vision stage was examined, and a new model for this stage was presented. Probability theory was used to measure the observer's awareness of the objects around them.

A computational model that combines perceptions of two different objects that could be perceived from many different perspectives, such as an exhibited object and a structure in an urban context, is presented. The approach of this study is based on Bayes' theorem. The starting point of the study is the stimulus on the stage, and the study covers the processing and interpretation stages that occur in the brain. In the study, three different viewpoints were determined for the objects, and the fusion of probabilities was obtained for each viewpoint. Using ray tracing, rays were sent into the environment, and the perception was obtained by dividing the number of rays hitting the object by the total number of rays sent (Ciftcioglu & Bittermann, 2013).

Statistical methods were also used to evaluate visual perception. In these studies, researchers collected data from people to analyze it, and evaluated visual perception. Turgay Zıraman & Imamoglu (2020) aimed to reveal how the relative sizes of exhibition objects, the presence of large objects nearby, and their positions during

viewing affect interest in the exhibition and visitor attention. For the study, seven exhibition objects were designed, one of which was larger than the others. In the study, the position of the large object was changed to create four different exhibition layouts. With the participants' permission, a video recording was made during their observations in the hall and analyzed in detail. AHT (average holding time) and AP (attention power) were determined for each object. The data obtained were analyzed using z-test, ANOVA, Scheffe test, and independent-samples t-test.

In another example of a statistical study, visual perception was evaluated based on object recall (Seven, 2024). The purpose of the study was to reveal the visual perception threshold depending on object size. The study was conducted by setting up a virtual experiment environment. During the case study, participants navigated the virtual experimental environment. Afterward, data on the recall of each object was collected through questionnaires. The collected data was compared with the surface area and volume information of the objects.

The examples mentioned above have evaluated visual perception through the use of probabilistic and statistical approaches. This study is similar to Turgay Zıraman & İmamoğlu (2020) and Seven (2024) in that it creates a case study using objects of different sizes and makes an evaluation of visual perception based on the case study data. In this study, averaging, correlation, and regression analyses were applied to evaluate visual perception. In addition, as in the study by Bittermann & Ciftcioglu (2008) objects were used without colour and texture features.

3.2 VR Analysis

VR (Virtual Reality) enables the simulation of environments with high realism and provides control over environmental variables. Traditional methods often have difficulties in being applied in real-world conditions. In contrast, VR offers a dynamic and interactive platform that enables architectural elements such as facade materials, landmarks, or spatial layouts to be systematically modified and tested. For these reasons, it is preferred in computational perception research (Kim & Lee, 2021). There are various studies in the literature on examining visual perception using VR. In this field, Diaz et al. (2024) conducted a study in a VR environment to examine the relationship between colour compatibility and object size with recall speed, object

placement accuracy, and recognition. In the study, 150 objects were divided into 6 groups and presented to participants. In the VR environment, participants were shown objects in a room. They were then asked to select which objects were not in the room. The study analyzed how colour differences between the first and second rooms affected visual perception. Participants were also asked to place the objects they saw. Thus, the accuracy of the positions where participants placed the objects and their placement speed were also examined. These parameters were analyzed in relation to object size. Additionally, researchers conducted a survey to examine participants' emotional states. The collected data revealed relationships between spatial cognition, colour compatibility, recognition and recall speed, object placement accuracy, and object size.

Another study in this field was conducted by Şen Bayram et al. (2023). The study aimed to evaluate and reveal the perception of spatial features during the user's cultural heritage experience in a virtual environment. The study was based on the three-fold imagery concept presented by Kaufmann. The concept states that visual imagery, linguistic representations, and exploratory activity interact strongly with each other during creative thinking and problem solving (Şen Bayram et al., 2023; Kaufmann, 1980). Süleymaniye Mosque was selected as cultural heritage in the study and was spatial memory examined through visual, linguistic, and experiential factors. After experiencing the VR environment, data was collected from participants through sketches, route extraction, and semi-structured interviews, and then analyzed.

In a study conducted by Reggente et al. (2020), the Loci Method, a technique used for memory, was researched in a VR environment. The study aims to present this method in an easy-to-use environment and to explore its effectiveness. Five different environments were designed for the experiment, and specific objects were placed in these environments. The 60 participants in the experiment were asked to collect these objects. Later, photographs of these objects were shown to the participants, and they were asked to label the objects' locations. The recall tests were conducted using Matlab and the Psychophysics Toolbox. The mentioned studies provided insight into the process of designing the experimental environment. For this study, a virtual environment was created using 3D models to evaluate visual perception. In this study, the images of the objects were shown and their memorability was measured, similar to

Diaz et al. (2024) however, participants rated their memorability. Although VR technology is an effective tool in perception research, it was not used in this study. The reasons for this include the technical complexity of VR systems, installation costs, and the possibility of causing effects such as nausea and dizziness in some participants. Additionally, previous studies have observed that errors in the perception of size and distance can occur in VR environments (Wilson & Soranzo, 2015).

3.3 Eye-Tracking and EEG Studies

Analyses using eye tracking and EEG (electroencephalography) are widely used for visual perception evaluation. Eye tracking records people's eye movements, gaze directions, and fixation points. This technology is used in fields such as cognitive psychology, product development, and design to examine decision-making processes and visual attention and perception (Kim & Lee, 2021). EEG is a high-resolution measurement technique that records electrical activity in the brain using electrodes placed on the head. It is frequently preferred in studies investigating cognitive processes related to perception, such as memory, attention, and emotions (Zou & Ergun, 2023). In this section, current studies analyzing and evaluating eye-tracking and EEG data and their contributions to computational visual perception models are discussed.

Lisińska-Kuśnierz, & Krupa (2018) conducted a study on the visual perception of urban spaces and architectural objects. They used eye tracking to conduct a case study. Four city photographs were used for the case study. In these photographs, historical and modern buildings and objects not directly related to architecture were shown to participants for 15 seconds. The research results were analyzed by creating heat maps. The data obtained from these maps helped identify areas of interest. The analyses were conducted by examining the number of fixations, their order, the average gaze duration per individual participant, and the total gaze duration.

There are also studies in the literature that use VR and eye-tracking together. One such study conducted a quantitative assessment of visual perception in VR for retail environments (Kim & Lee, 2020). Researchers designed an experiment using eye tracking and VR. A 360-degree simulation of a car showroom was prepared for the

experiment. Data was collected from participants via a questionnaire at the end of the experiment to evaluate retail cognition. Visual attention analysis was performed based on the participants' gaze fixation order, number, and duration. A similar study was conducted by the same researchers a year later (Kim & Lee, 2021). In this study, participants' visual attention and emotional arousal were examined in a test environment. In difference from the previous study, the effect of gender differences on the results was also examined.

EEG is an important tool in urban design evaluation, analyzing the impact of design elements on human perception, and improving design decisions (Zou & Ergan, 2023). Grima Murcia et al. (2019) investigated how images with different levels of detail are processed in the brain using EEG. The study aimed to compare computer-generated 3D representations with photographs. In their study, participants were shown a photograph of a bedroom and different levels of detail from a 3D model of the same room (these images varied in colour, pattern, and number of objects). Participants were then asked to provide feedback on their valence and arousal levels. Additionally, EEG recordings were used to examine brain regions during the viewing of these images. Differences in brain regions were observed, particularly in the early stages, indicating that people perceive realistic images as more enjoyable and that aesthetic judgment may be influenced by this. This suggests that human acceptance can be enhanced by producing familiar, realistic visuals.

The approach of Zou & Ergan (2023) is CNN-based. Two-dimensional EEG images were used as input. The model's performance was tested using recall, precision, F-1 score, and area under the curve (AUC). They examined three different sets of emotions that architecture evokes in people: stress and anxiety, restorativeness, and aesthetic pleasure. Different virtual environments were designed in the study, and participants' experiences were monitored using EEG. Bipolar environments were designed for the three emotion groups, with positive and negative sides. After exploring both virtual environments, participants evaluated their emotional experiences through a survey. Once the data was collected, the raw EEG data was cleaned of noise and converted

into 2D images, which were then used as input for the classification model to predict human emotions (Zou & Ergan, 2023).

In another study, EEG and fuzzy logic were used to determine auditory and visual perception in a closed public space (Cantemir, 2025). In this study, EEG data and fuzzy logic were used to analyze the effect of different music genres on spatial perception based on colour and material characteristics. Eight virtual spaces were designed with material and colour changes and presented to participants accompanied by three different types of music. EEG device was used to collect participants' stress, attention, excitement, attachment, interest, and relaxation data at specific time intervals. These data and music types were used as inputs in fuzzy logic to obtain spatial outputs (Cantemir, 2025).

Eye-tracking technologies are commonly employed to monitor participants' visual attentional focus and duration. However, the majority of eye-tracking technologies create a static experimental setting by requiring participants to keep their eyes fixated on a screen (Kim & Lee, 2021). Even though eye-tracking is highly effective for capturing momentary attention and gaze patterns, it primarily measures immediate visual focus and does not directly evaluate longer-term memory (Kiefer et al., 2017). EEG is effective in measuring rapid brain activity in visual cognition research; however, it is costly. Moreover, it requires an experimental setup that limits participant movement and demands expertise for data interpretation (Wilson et al., 2024). For these reasons, EEG and eye tracking were not preferred in this study.

3.4 Machine Learning Applications

With the development of computational methods and artificial intelligence, models using machine learning have begun to emerge for visual data analysis. Advancements in computer vision have contributed to the quantitative assessment of users' perception of scenes, spaces and objects. In this context, attention maps, CNNs, and object recognition models are effectively used to detect visual perception. This section of the

thesis explains computational visual perception models based on machine learning and deep learning.

Studies using machine learning have primarily evaluated visual perception through object memorability. These studies show that the objects people remember and forget in an image are similar (Dubey et al., 2015). Studies in the literature reveal that there are no categorical similarities (such as humans, animals, plants in object recognition studies) between commonly remembered images. So what is the common feature that makes objects in images memorable for most people? Celikkale et al. (2015) mention in their study that people remember only the gist of a scene, not its entire representation. Studies on memorability also generally focus on investigating this gist (Basavaraju et al., 2019; Celikkale et al., 2015; Dubey et al., 2015).

Isola et al. (2011) were the first researchers to examine this situation in a computational manner. In their study, they designed a visual memory experiment and examined memorability using 2222 images. The images were shown to 665 participants for 1 second. The images were shown again shortly after, and feedback was obtained from the participants. Based on the feedback, a memorability score was calculated for each image. This study showed that machines can perform well in predicting the intrinsic memorability of images. (Celikkale et al., 2015; Isola et al., 2011)

Dubey et al. (2015) analyzed image memorability at the object level in their study. The aim was to examine the impact of factors affecting object memorability, such as colour, visual salience, and object categories, on image memorability. Ground-truth scores were collected for a large dataset. These scores were used for memorability analysis and for a deep learning prediction model. In their work, they were the first to demonstrate that object and visual memorability of image are related.

A similar study was conducted by Celikkale et al. (2015). In this study, researchers aimed to reveal the effect of image semantics and visual attention on memorability. The authors identify salient regions of images and use attention-driven pooling instead of spatial pooling. Features extracted from salient regions in images led to better results than previous studies. Furthermore, meta-level information was analyzed in terms of

semantic features such as scene features, evoked emotions, and object categories, and a combined model was proposed.

Another study examined the memorability of objects after they were seen once (Basavaraju et al., 2019). Different from previous studies, this study developed a baseline model to examine the effect of spatial size and location on object memorability. In addition, a model was developed using deep learning for automatic learning of these features. The dataset used by Dubey et al. (2015) was utilized in this study. The results demonstrate that spatial location and size improve recall prediction performance and that the developed model performs better than those in the literature.

A more recent study was conducted by Kramer et al. (2023). A larger data set was used than in previous studies. The object feature model created predicts the memorability of images and examines its relationship with object typicality. In this study, the relationship between memorability and typicality was examined using three measurement methods. These are multidimensional object features, human judgments, and DCNN predictions. Using multiple regression, memorability was predicted from semantic and visual features, and it was found that visual features were less effective. Subsequently, memorability scores and object typicality-related features were analyzed. The relationship between typicality and memorability was revealed through this approach.

Machine learning and deep learning are prominent for memorability prediction. The limitations of these methods are that they require large-scale datasets, high-performance computing resources, and their outputs are difficult to interpret (Tsirtsakis et al., 2025; Wang, 2016). The similarity between this thesis and the studies in this section is as follows: A CNN-based deep learning model YOLOv8 was used. In addition, memorability was examined on image frames. However, the difference in this study is that image frames obtained from a virtual experiment environment were used instead of photos taken from daily life. Furthermore, the analyzed images contain abstract objects. In addition, the data set and the number of participants are limited.

4. METHODOLOGY

The methodology of the model created in this thesis is abstracted into three main stages. In the first stage, the experimental environment was generated, in the second stage, the computational model was developed. Finally, a case study was conducted to evaluate the implementation of the model. The data obtained from the case study and the computational model were analyzed. Each of these stages is described in detail in this section. The methodology of the proposed model and tools used are shown in Figure 4.1.

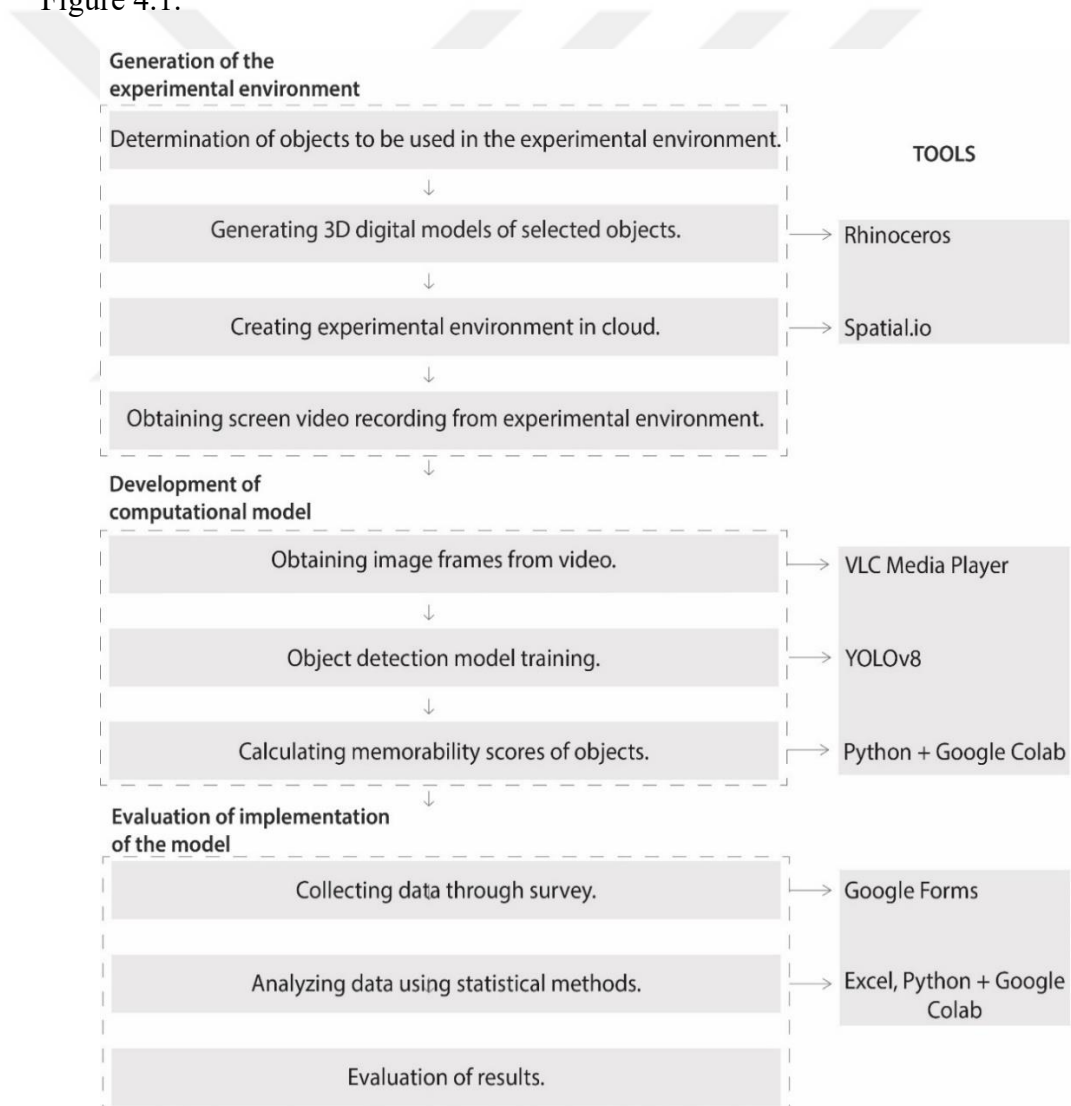


Figure 4.1: Methodology of proposed model.

4.1 Generation of the Experimental Environment

The experiment was designed in a virtual environment. The reason for this was to enable the objects used in the experiment to be easily modeled and placed in the environment and to conduct the experiment with more participants due to easier use. The Spatial.io platform was chosen to create the experimental environment. Spatial.io is a web and mobile-compatible platform that works on augmented and virtual reality. The created spaces are exploreable through selected avatars. Screen recording can be taken from the space. The platform includes ready-made space templates. Users can also import their own designed 3D models (Martí-Testón et al., 2023). The objects to be used in the study were identified, and their 3D models were created. They were placed in the selected space template. A screen recording was taken during the tour of the space.

4.2 Development Of The Computational Model

The study aims to evaluate the relationship between visual perception and object size using a computational model based on memorability. The model was developed in three steps. In the first step, image frames were obtained from the video recording. The VLC Media Player program was used for this purpose. A specified number of image frames were obtained for each second.

In the second step, object detection was carried out on the image frames obtained. YOLOv8, introduced by Ultralytics in 2023, was selected due to its compact design and high accuracy. The model has six working modes: train, val (validation), predict, export, track, and benchmark (Sohan et al., 2024). The model performs tasks such as object detection, segmentation, and pose estimation. This model can be easily adapted to low or high hardware. It is also possible to train the model in the cloud (Sohan et al., 2024). In this study, YOLOv8 was used for the object detection task.

In the final step, computational processes were conducted. Scores were determined based on the sizes and centrality of objects. Size ratios were obtained by dividing the area covered by each object detected in the image frame by the image area. Equation 4.1 shows the formula of memorability score based on size. This process was repeated

for each frame. Memorability scores based on size were created by summing the ratios obtained for each frame.

According to the formula of memorability score based on size:

- n is total number of image frames.
- A_i is detected area of the object in the i -th frame (bounding box area).
- At is total image area (constant across all frames).
- Sm is memorability score based on size.

$$Sm = \sum_{i=1}^n (A_i / At) \quad (4.1)$$

Basavaraju et al. (2019) mention in their studies that the location of objects in an image affects memorability. They have found that objects in the center of the image are more memorable. In this study, the distance of objects from the center of the image was examined under the value of centrality. A similar approach to size computation was used in the centrality computation. The euclidean distance between the center of each object detected in the image frame and the center of the image was measured. This step is explained in the equation 4.2.

According to the formula of object's center distance:

- D_i is euclidean distance between the center of the detected object and the center of the image in the i -th frame.
- (x_0, y_0) are coordinates of the object's center.
- (x_c, y_c) are coordinates of the image frame's center.

$$D_i = \sqrt{(x_0 - x_c)^2 + (y_0 - y_c)^2} \quad (4.2)$$

This process was repeated for each frame. The values obtained were subtracted from a constant number to create centrality values on a per-frame basis. Memorability scores were created by summing the centrality values obtained from each frame. Equation 4.3 shows the formula of memorability score based on centrality.

In the formula of memorability scores based on centrality:

- c is constant defined according to the frame size.
- n is total number of image frames.

$$\sum_{i=1}^n (c - D_i) \quad (4.3)$$

4.3 Evaluation of Implementation of the Model

A data was collected from people to test the performance of the developed computational model. For this purpose, a case study was conducted. A recall experiment was conducted within the scope of the case study. A quantitative data collection method was used. A video recorded in the experimental environment, was shown to the participants. Seven (2024) showed 26 objects and asked the participants to take a recall survey 15 minutes after showing the objects. Since there were fewer objects in this experiment, this period was extended and the questionnaire was conducted 20 minutes after the video was shown. The questionnaire included a recall score from 1 to 7 for each object. The options were 1 = “I definitely do not remember,” and 7 = “I definitely remember.” The questionnaire was carried out using Google Forms.

The relationship between the data obtained from the developed computational model and the data obtained from the questionnaires was examined. First, the average recall scores collected from the participants’ surveys were calculated for each object. The relationship between these average scores and the scores obtained from the computational model (memorability scores based on centrality and size) was examined. For this purpose, correlation analyses were performed. Correlation analysis

was used to understand the relationship between variables. This statistical method was used to examine the direction and strength of the relationship between two variables. Pearson correlation, and Spearman correlation widely used method in this field, were applied. The r value was used to examine the strength and direction of the relationship, while the p value was used to determine whether the relationship was statistically significant or coincidental. Additionally, simple linear regression analysis was carried out to determine the effect of independent variables on the dependent variable.





5. CASE STUDY

The case study was conducted with the participation of 43 people from different professional groups between the ages of 20-30. It was conducted with the approval of the ITU Social and Human Sciences Scientific Research and Publication Ethics Committee (Appendix A). All participants were informed about the study process and their consent was obtained. The experimental environment was first set up, then a computational model was developed and memorability scores were calculated. Finally, data was collected from participants via a survey and analyzed. In this section, these stages and findings are explained.

5.1 Creation of the Experimental Environment

The Spatial.io platform was used in the study due to its ease of use, variety of space templates and avatars, accessibility, and ability to import models from external sources. One of the templates provided by Spatial.io, Aeries Gallery, was selected for this study. The selected environment is appropriate for conducting the experiment. The reason for this is there are no obvious differences such as light, texture, colour, or material that would affect the perception of objects.

Nine objects of different sizes have been placed in the space template for the experiment. Çiftcioglu & Bittermann (2013) stated in their study that in an environment with more than seven objects, remembering all the objects becomes even more difficult. Based on this, the number of objects was kept above seven; however, to avoid increasing the complexity of the experiment and the cognitive load on participants, the number of objects was set to nine. Objects that people are likely to see in their daily lives were selected for this study. These objects are a briefcase, monitor, suitcase, stool, dresser, radio, oven, vending machine, and wardrobe. The reason for selecting these objects is that they are familiar to the participants and do not lead to subjective interpretations like art objects. This allowed for a reduction in individual interpretations and a more controlled experiment. Except for the stool, the object models were obtained from 3D Warehouse website (Url-4). Adjustments have

been made to the models in Rhinoceros. Since differences in colour, texture, and material of the objects could affect visual perception, these factors were kept equivalent for all objects. The objects were placed in the experimental environment in a heterogeneous spatial arrangement without size classification. Figure 5.1 shows an example screenshot taken from the experimental environment.

A virtual tour was conducted in the designed experimental environment. This tour lasts 20.5 seconds and allows users to see all objects. Screen recording was taken during the tour, resulting in a 30 FPS video. This video was later used to calculate memorability scores and to be shown to participants. Figure 5.2 shows the placement of objects and camera movement.

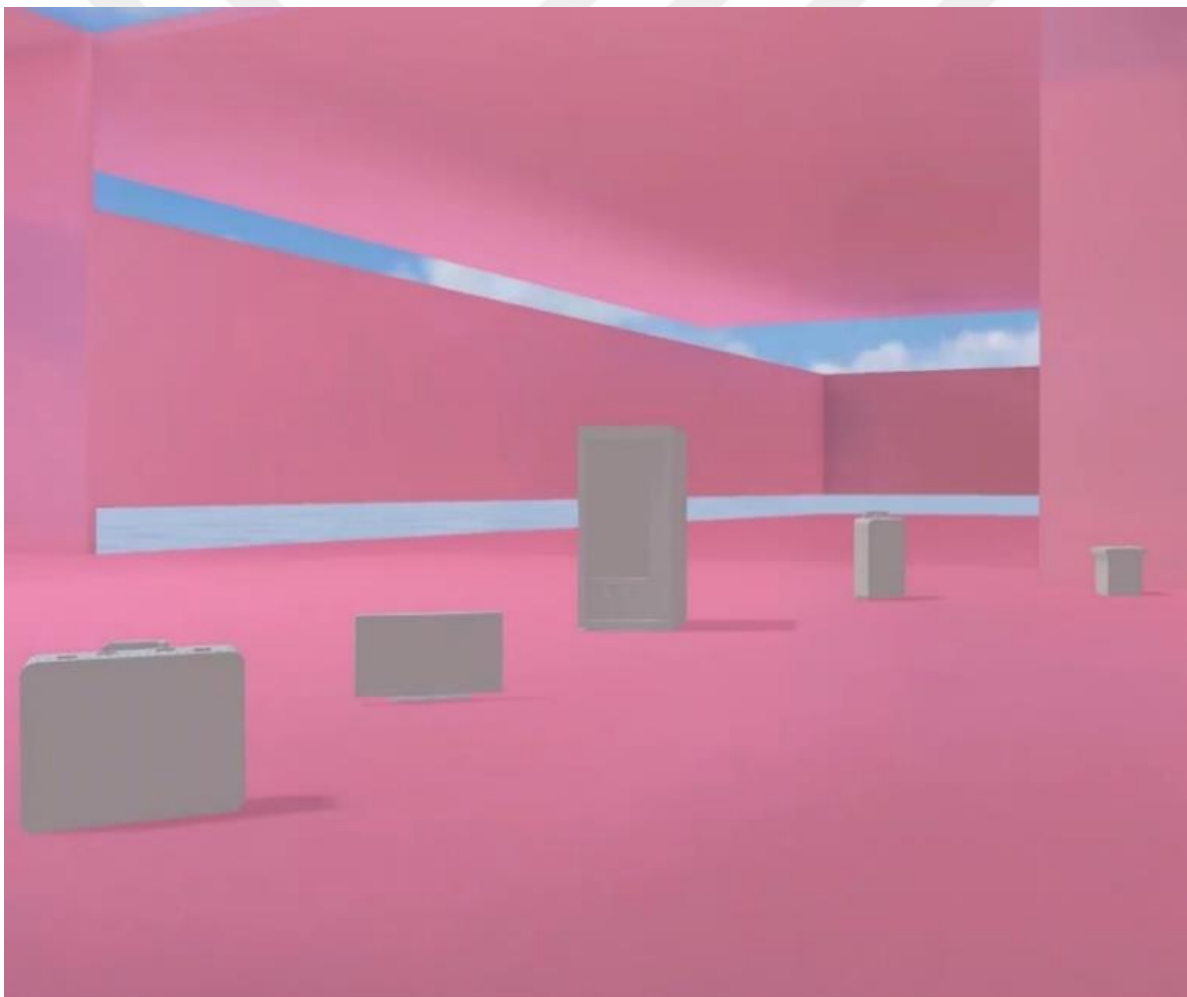


Figure 5.1 : Image taken from the experimental environment.

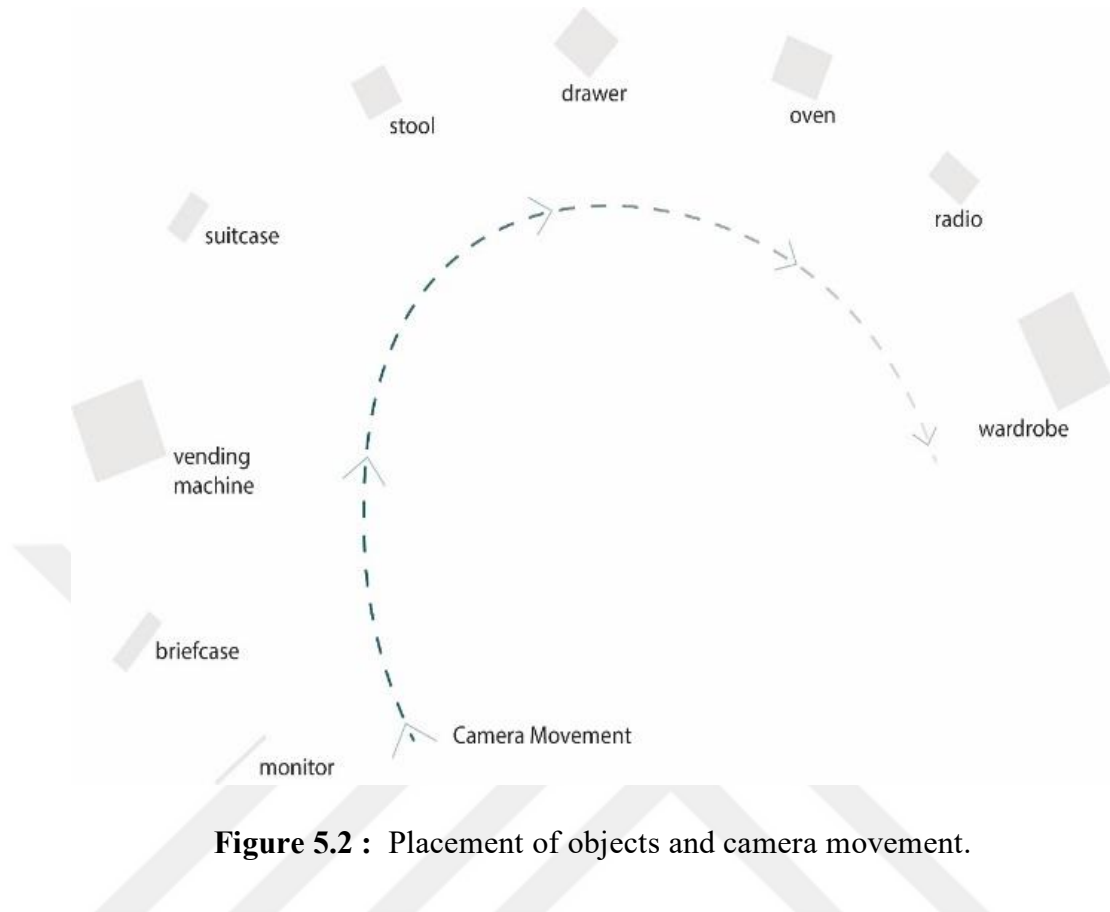


Figure 5.2 : Placement of objects and camera movement.

5.2 Development of the Computational Model and Computation of Scores

The obtained video was split into image frames at a rate of 2 frames per second. A total of 41 image frames with a resolution of 1920x880 pixels were obtained. The obtained image frames were transferred to Google Colab for object detection. First, the YOLOv8 model trained with the COCO dataset was used for the study. However, it was observed that this model was not sufficiently successful in detecting objects in the image frames. For this reason, a dataset was created for training by using the image frames through Roboflow. Nine classes were determined for the selected 21 frames according to object categories. Object labeling was made. Figure 5.3 shows the object labeling process. At the end of the labeling process, the number of images in the dataset was increased to 51 with the addition of rotations of frames. The dataset was separated into three parts: training, validation, and testing.

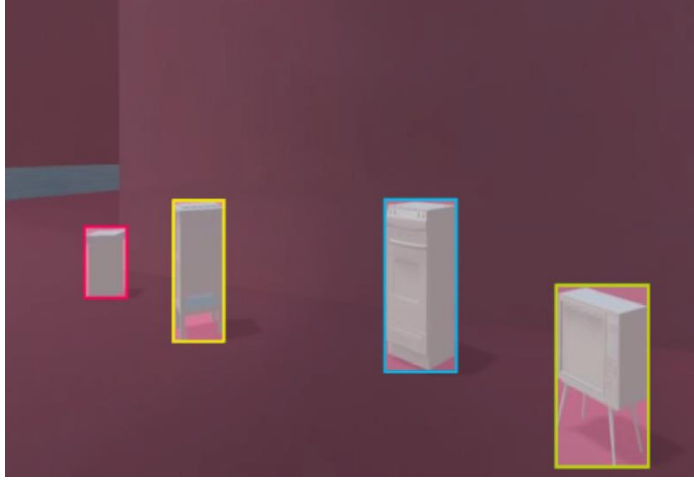


Figure 5.3: Object labeling process

After this step, object detection was performed for each frame. First, the Ultralytics library was downloaded and the YOLO model was imported from this library. The dataset prepared in Roboflow was uploaded to the Colab environment. Then, the training process began. The task was set to detect, the mode was set to train, and the number of epochs was set to 25. The four example images obtained from the training results are shown in Figure 5.4.

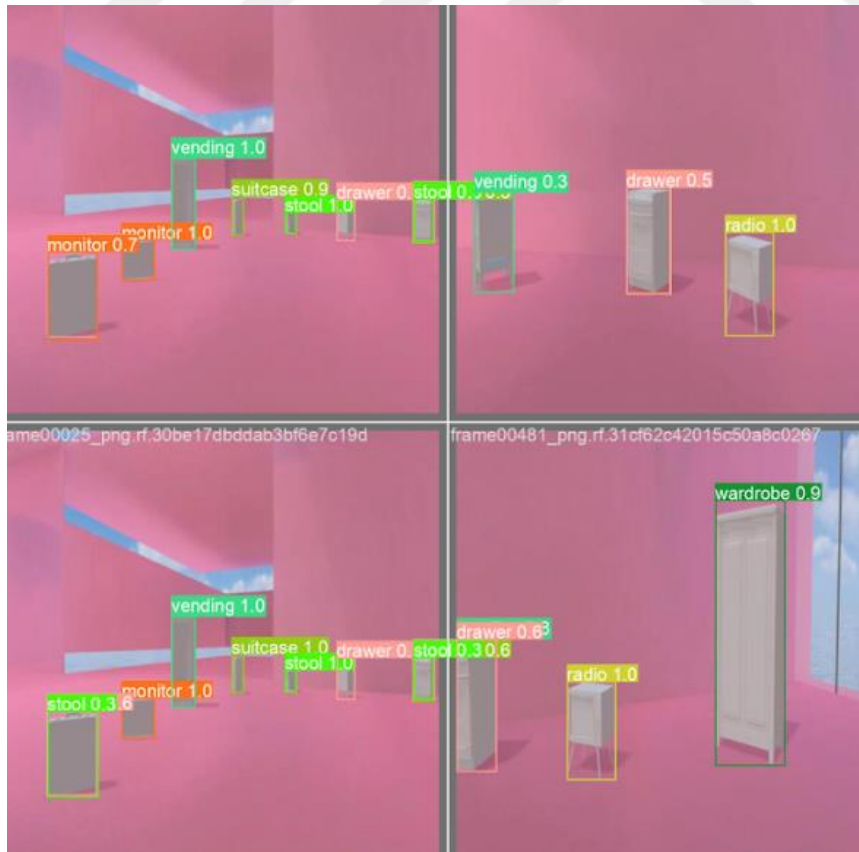


Figure 5.4: YOLOv8 results after training.

After the training phase, a Python code was written to calculate the ratio of all objects to the image. First, the image file to be processed was read and its dimensions were calculated. Then, the trained object detection model (YOLOv8) was run. The bounding box information for the objects detected by the model in the images was assigned to the variable named result.

Then, the width and height of the objects were calculated using the bounding box information. The area covered by the detected objects was determined by multiplying the obtained width and height values. The size ratios were calculated by dividing the bounding box areas of the objects by the entire image area. Figure 5.5 shows an example image frame with the calculated ratios. This process was performed for all frames and every object. It was observed that the model incorrectly detected or missed objects in some frames. In these cases, the ratios of the objects were calculated manually. The calculated ratios were transferred to a table (Table 5.1). The memorability scores based on size obtained for each object were summed up, considering all frames. Table 5.2 shows the memorability scores based on size of the objects.

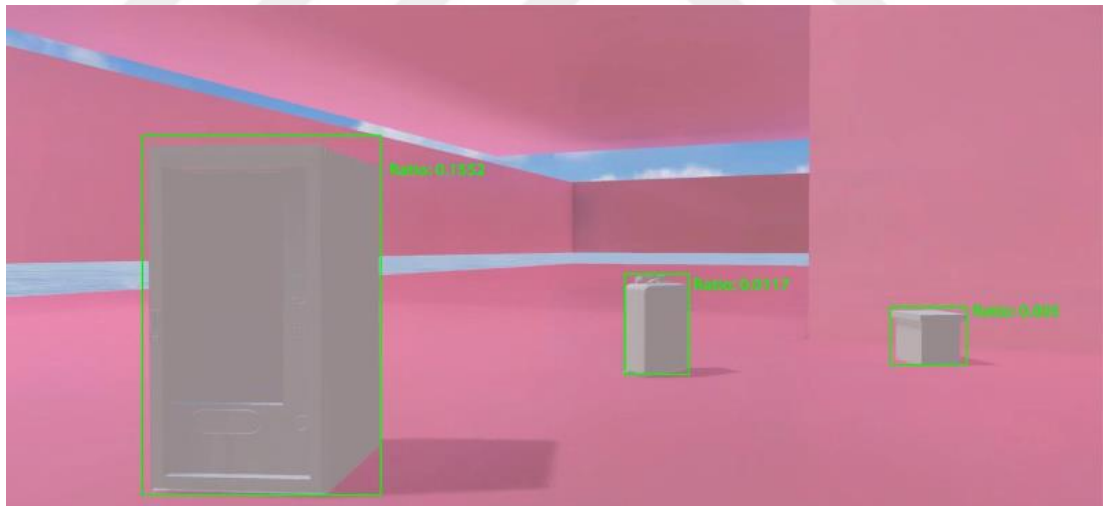


Figure 5.5: Determining the size ratios of objects

Table 5.1: Size ratios for all frames.

Frame Name	Briefcase	Monitor	Vending Machine	Suitcase	Stool	Drawer	Oven	Radio	Wardrobe
F1	0,024	0,007	0,012	0,002	0,001	0,004	0,004	0	0
F2	0,024	0,007	0,012	0,002	0,001	0,004	0,004	0	0
F3	0,03	0,008	0,013	0,002	0,002	0,004	0,004	0	0
F4	0,039	0,01	0,014	0,003	0,002	0,005	0,001	0	0
F5	0,025	0,013	0,016	0,003	0,002	0,005	0	0	0

Table 5.1 (continued): Size ratios for all frames.

Frame Name	Briefcase	Monitor	Vending Machine	Suitcase	Stool	Drawer	Oven	Radio	Wardrobe
F6	0,004	0,015	0,018	0,003	0,002	0,006	0	0	0
F7	0	0,025	0,025	0,004	0,002	0,007	0	0	0
F8	0	0,035	0,029	0,005	0,003	0,008	0	0	0
F9	0	0	0,048	0,006	0,004	0,008	0	0	0
F10	0	0	0,062	0,007	0,004	0,003	0	0	0
F11	0	0	0,056	0,007	0,006	0	0	0	0
F12	0	0	0,155	0,012	0,008	0	0	0	0
F13	0	0	0,190	0,014	0,009	0	0	0	0
F14	0	0	0,155	0,014	0,009	0	0	0	0
F15	0	0	0,099	0,014	0,009	0	0	0	0
F16	0	0	0	0,014	0,007	0,032	0	0	0
F17	0	0	0	0,014	0,007	0,030	0	0	0
F18	0	0	0	0,014	0,006	0,029	0	0	0
F19	0	0	0	0,014	0,006	0,029	0	0	0
F20	0	0	0	0,014	0,006	0,028	0	0	0
F21	0	0	0	0,017	0,006	0,027	0	0	0
F22	0	0	0	0,098	0,007	0,027	0,012	0	0
F23	0	0	0	0	0,007	0,026	0,045	0	0
F24	0	0	0	0	0,007	0,026	0,044	0	0
F25	0	0	0	0	0,008	0,025	0,044	0	0
F26	0	0	0	0	0,007	0,019	0,035	0	0
F27	0	0	0	0	0,006	0,017	0,029	0	0
F28	0	0	0	0	0,005	0,015	0,026	0,004	0
F29	0	0	0	0	0,005	0,016	0,026	0,026	0
F30	0	0	0	0	0,006	0,015	0,026	0,035	0
F31	0	0	0	0	0,006	0,015	0,025	0,033	0
F32	0	0	0	0	0,006	0,015	0,025	0,033	0
F33	0	0	0	0	0	0,010	0,026	0,023	0,042
F34	0	0	0	0	0	0	0,026	0,023	0,098
F35	0	0	0	0	0	0	0,032	0,023	0,145
F36	0	0	0	0	0	0	0,034	0,023	0,123
F37	0	0	0	0	0	0	0,034	0,023	0,12
F38	0	0	0	0	0	0	0,033	0,023	0,12
F39	0	0	0	0	0	0	0,034	0,023	0,116
F40	0	0	0	0	0	0	0,036	0,025	0,096
F41	0	0	0	0	0	0	0,026	0,025	0,095

Table 5.2: Memorability scores based on size.

	Briefcase	Monitor	Vending Machine	Suitcase	Stool	Drawer	Oven	Radio	Wardrobe
Score	0,144	0,120	0,904	0,284	0,173	0,453	0,630	0,34	0,955

The memorability scores based on centrality of the objects were also computed. For this computation, the center point of the image was first defined. Then, the center point of the bounding box of each object detected in the image frame was determined. The distance between these two points was calculated using Python code. This process was carried out for each image frame. Figure 5.6 shows an example image frame for which centrality is calculated. The obtained values were noted. The purpose is for objects closer to the center to receive a higher score. For this reason, the distance

measurements were subtracted from a fixed number (1100). This constant value was determined based on the maximum center distance value in the frames. The highest value was 1001.3 pixels for the briefcase in frame 6. This value was rounded to 1100. The centrality value of objects not appearing in the image frame was set to 0. The values obtained for each frame from this calculation are shown in Table 5.3. The memorability scores based on centrality of each object was determined based on the total value obtained from all frames. Table 5.4 shows the centrality calculation results.



Figure 5.6: Determining the centrality of objects.

Table 5.3: Centrality values for all frames

Frame Name	Briefcase	Monitor	Vending Machine	Suitcase	Stool	Drawer	Oven	Radio	Wardrobe
F1	423,5	719,4	940,1	1011,4	785,2	544,7	207,6	0	0
F2	426,2	719,4	940,1	1011,4	785,2	544,7	206,6	0	0
F3	359,1	696,6	933,9	1008,3	777,1	526,7	185,9	0	0
F4	229,1	648,1	922,6	1004,2	758,5	491,5	150,2	0	0
F5	168,2	606,9	910,5	999,3	745,4	463,2	0	0	0
F6	98,7	558,7	900	992,6	730,7	434,8	0	0	0
F7	0	415,5	873,6	980,8	694	361,3	0	0	0
F8	0	292,2	854,2	972,6	670,3	311	0	0	0
F9	0	0	788,3	947,8	601,3	185,5	0	0	0
F10	0	0	756,7	938,2	568,5	150,1	0	0	0
F11	0	0	700,2	920,8	520,5	0	0	0	0
F12	0	0	577,9	886,7	433,8	0	0	0	0
F13	0	0	515,5	873,4	393	0	0	0	0
F14	0	0	328,4	963,7	506,1	0	0	0	0
F15	0	0	258,8	968,1	578,3	0	0	0	0
F16	0	0	0	45,5	832,1	253	0	0	0
F17	0	0	0	633,7	913,4	397,8	0	0	0
F18	0	0	0	498,4	952,8	548,2	0	0	0
F19	0	0	0	421	933,5	637,6	0	0	0
F20	0	0	0	303	860,6	763	0	0	0
F21	0	0	0	202	785,5	858,2	0	0	0

Table 5.3 (continued): Centrality values for all frames

Frame Name	Briefcase	Monitor	Vending Machine	Suitcase	Stool	Drawer	Oven	Radio	Wardrobe
F22	0	0	0	177,1	738,2	898,3	70	0	0
F23	0	0	0	0	612,6	915,4	274,1	0	0
F24	0	0	0	0	597,3	906,2	309,8	0	0
F25	0	0	0	0	507,7	826,5	446,4	0	0
F26	0	0	0	0	544,6	841,3	543,2	0	0
F27	0	0	0	0	577,1	861,4	598,2	0	0
F28	0	0	0	0	590,7	871,9	614,2	129,7	0
F29	0	0	0	0	516,7	787,7	722	196,9	0
F30	0	0	0	0	418,4	672,4	845	347	0
F31	0	0	0	0	389,5	639,4	877,2	396,9	0
F32	0	0	0	0	173,2	443,3	989,9	614,5	0
F33	0	0	0	0	0	180,8	843,6	839,3	203,7
F34	0	0	0	0	0	0	743,2	914,1	276,7
F35	0	0	0	0	0	0	592,9	947,9	375,6
F36	0	0	0	0	0	0	485	908	496,1
F37	0	0	0	0	0	0	475,6	902,6	509
F38	0	0	0	0	0	0	475,4	902,7	510
F39	0	0	0	0	0	0	475,4	902,7	510
F40	0	0	0	0	0	0	231,1	710,7	758,8
F41	0	0	0	0	0	0	219,5	695	777,6

Table 5.4: Centrality scores for objects.

	Briefcase	Monitor	Vending Machine	Suitcase	Stool	Drawer	Oven	Radio	Wardrobe
Score	1,281	3,937	11,201	16,76	20,49 1	16,316	11,51 7	9,371	4,476

5.3 Data Collection and Analysis

To evaluate the performance of the model, the memorability scores of the model were compared with the memorability data obtained from humans. For this reason, 43 participants were informed about the experiment and their consent was obtained. Then, they were shown the video taken from the experimental environment. Participants were asked to complete a questionnaire 20 minutes after the experiment was completed. Figure 5.7 shows the experiment and data collection process. The questionnaire consisted of 13 questions. It includes images of 9 objects in the experimental environment and 4 objects not in the experimental environment. These objects were placed in the survey regardless of the order in which they appeared in the experimental setting. Participants were asked to rate how well they remembered these objects. Figure 5.8 shows an example question.



Figure 5.7: Experiment and data collection process.

Do you remember that the object in the image was there in the experimental environment? *

(The level of recall is rated between 1 and 7. 1: I definitely do not remember, 7: I definitely remember.)



1 2 3 4 5 6 7

I definitely do not remember ☐ ☐ ☐ ☐ ☐ ☐ ☐ I definitely remember

Figure 5.8: Memorability question for suitcase.

Participants rated their recall status for each object on a scale of 1 to 7. Figure 5.9 shows the distribution of participants' recall scores for the stool object. An average memorability score was obtained for each object from the data collected from the participants. Table 5.5 shows the average memorability score calculated for each object.

Do you remember that the object in the image was there in the experimental environment? (The level of recall is rated between 1 and 7. 1: I definitely do not remember, 7: I definitely remember.)
43 yanıt

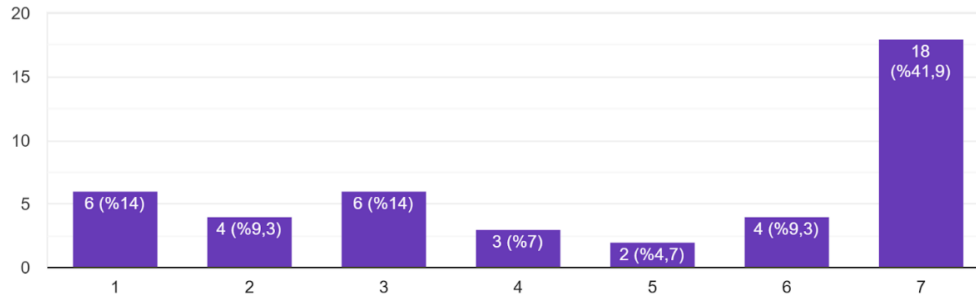


Figure 5.9: Participants' recall graph for stool.

Table 5.5: Average memorability scores for objects.

	Briefcase	Monitor	Vending Machine	Suitcase	Stool	Drawer	Oven	Radio	Wardrobe
Average Values	4,18	2,55	5	16,76	20,49	16,316	11,51	9,371	4,476

In this study, the results obtained by computational methods were analyzed with the data obtained from participants' questionnaires. The memorability scores calculated based on size and centrality were compared with the average memorability scores obtained from participant questionnaires. The data collected from Excel was transferred to Google Colab, where statistical analyses were made.

Pearson and Spearman correlation analyses were applied to examine the relationship between variables. These methods were used to determine the strength and direction of the relationship between survey scores and calculated scores. Then, simple linear regression analysis was performed. These analyses were carried out separately for both memorability scores based on size and centrality.

5.4 Findings

This section presents the findings obtained from the analyses conducted within the scope of the thesis. According to the developed computational model, it was observed that the objects with the highest memorability scores were, in order, wardrobe, vending machine, and oven, while the objects with the lowest memorability scores were, in order, monitor, briefcase, and stool. Figure 5.10 shows the memorability scores.

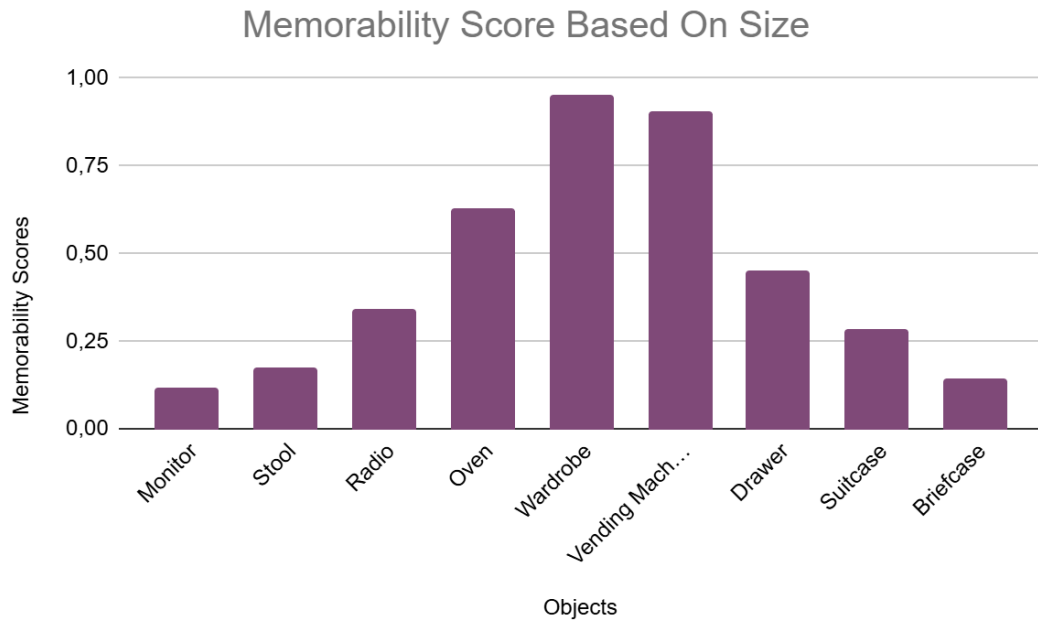


Figure 5.10: Memorability scores based on the model.

According to the average memorability scores obtained from the questionnaire, it was observed that the most memorable objects were, in order, the wardrobe, oven, and radio, while the least memorable objects were, in order, the monitor, suitcase, and drawer. Figure 5.11 shows the memorability levels of objects.

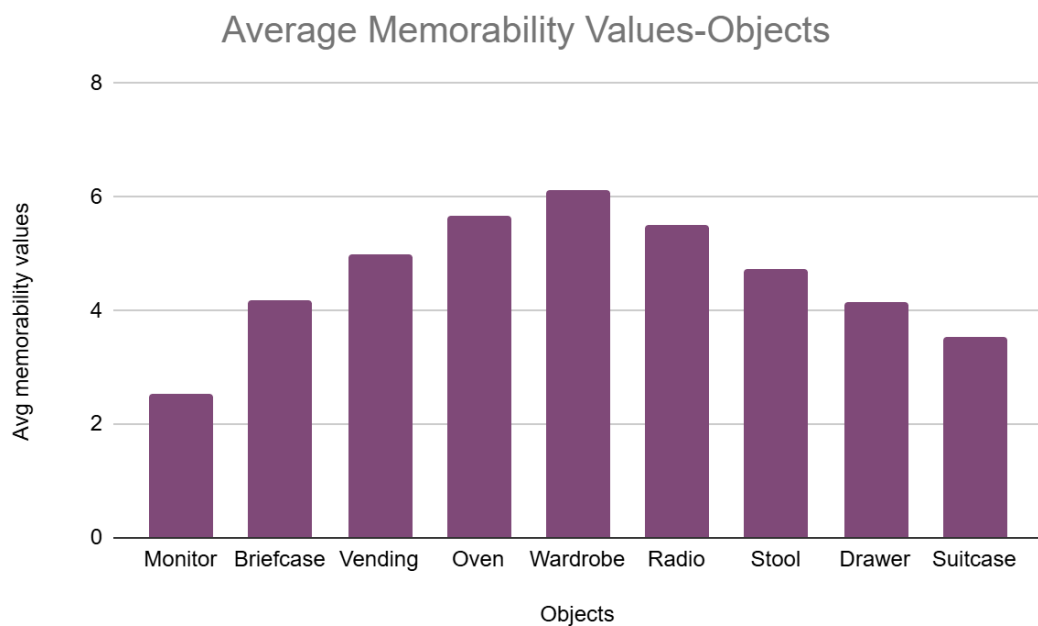


Figure 5.11: Average memorability scores obtained from participants.

According to both the computational model data and the questionnaire data, the most remembered object (wardrobe) and the least remembered object (monitor) are the same. The relationship between size-based memorability scores and average memorability values was examined using a line graph. Except for the part containing the stool and drawer objects, it was observed that both formed a similar pattern (Figure 5.12). It is observed that where size based memorability scores decrease, average memorability values also decrease, and where they increase, average memorability values also increase.

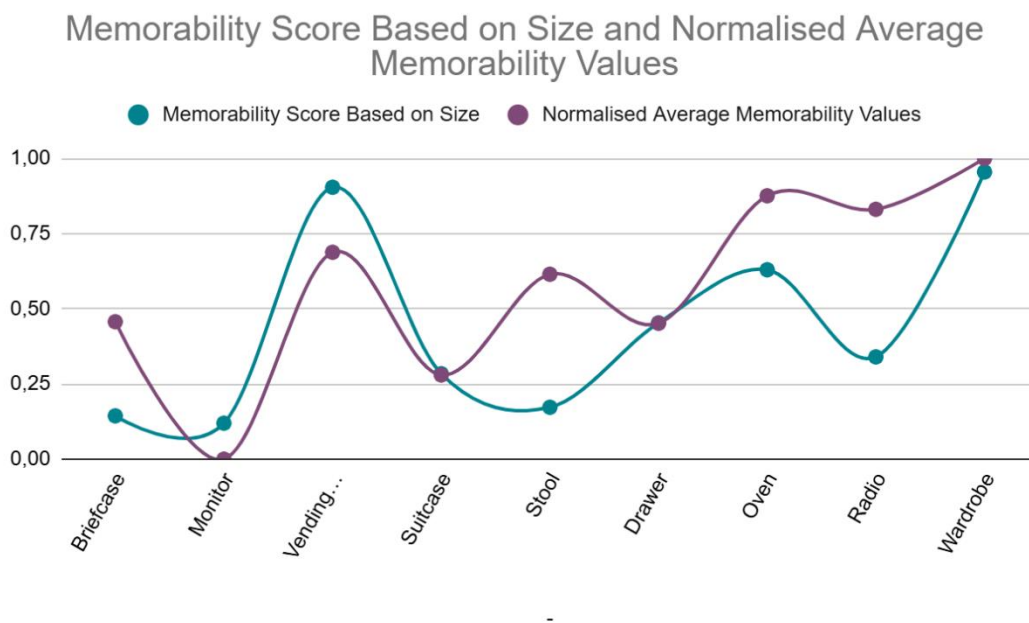


Figure 5.12: Size-based memorability score and survey averages.

Two different correlation analyses were performed to examine the relationship between participant questionnaire averages and the developed model in more detail. The distribution of the survey averages and the scores obtained from the model were determined first. For this purpose, the Shapiro-Wilk test was performed to examine the distribution of small sample groups. According to the test results, both the survey average data and the model data were found to be normally distributed. Therefore, Pearson correlation analysis and simple linear regression analysis were carried out. Pearson correlation is suitable for normal distributed data, but the size of the dataset in this study is very small. Pearson analysis is more sensitive to outliers. Additionally, the survey data are ordinal data. Considering this situation and the size of the data set, Spearman's correlation analysis was also conducted.

Firstly, memorability scores based on size and participant questionnaire data were compared. The results of the analysis showed that the Pearson correlation coefficient was $r=0.69$ and $p=0.038$. This result indicates that there is a strong positive relationship between the developed model and the data collected from participants. Additionally, this relationship is not coincidental.

Since there is a linear relationship between the variables and there is one independent variable (model scores based on size) and one dependent variable (averages from questionnaire), a simple linear regression analysis was performed. Regression analysis reveals the effect of the independent variable on the dependent variable. It is used to estimate the value that the dependent variable will take according to the value of the independent variable. Table 5.6 shows the regression analysis results. The determination coefficient (R^2) of the model is 0.545. In other words, the model scores can explain 54.5% of the questionnaire data. In addition, the slope coefficient is 3.014. This indicates that a one-unit change in the independent variable causes a 3.014-unit change in the dependent variable. The intercept value is 3.178. Therefore, when the independent variable is 0, the dependent variable is expected to be 3.178. Finally, the MSE value is 0.511. The squared difference between the predicted values and the actual values is 0.511.

Table 5.6: Linear regression results.

Intercept	Slope Coefficient	R^2	MSE
3.178	3.014	0.545	0.511

In the next step, memorability scores based on centrality and questionnaire data were compared. It was observed that the objects with the highest memorability scores were, in order, stool, suitcase, and drawer, while the objects with the lowest memorability scores were, in order, briefcase, monitor, and wardrobe. Figure 5.13 shows the memorability scores based on centrality. The Pearson correlation analysis was also carried to examine the relationship between these memorability scores and participants' data. The analysis revealed a correlation coefficient r of 0.14 and a p -value of 0.97. These values indicate that there is a very low, negligible relationship between memorability scores based on centrality and the data collected from participants. This relationship is also not statistically meaningful.

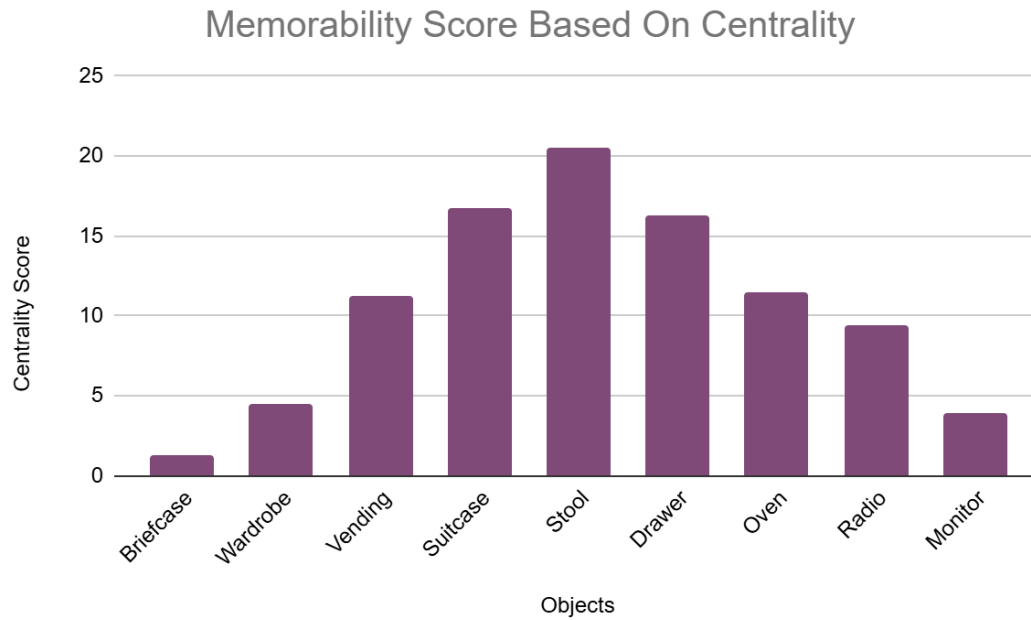


Figure 5.13: Memorability scores based on centrality.

The Spearman correlation analysis was also performed. The Spearman correlation coefficient between the model's memorability scores based on size and the questionnaire averages was found to be $r = 0.75$ with a p-value of 0.019. This indicates that there is a strong positive monotonic relationship between the scores obtained from the model and the questionnaire averages. Therefore, as the model score increases, the questionnaire averages are also expected to increase, but this increase is not linear. The Spearman correlation coefficient between the model's memorability scores based on centrality and questionnaire averages was found to be $r = -0.067$ with a p-value of 0.86. These values indicate that there is no significant relationship between these two variables.

The most important finding obtained from the analyses is that the size factor is an effective parameter for the memorability model. For this reason, it was decided that the model should only determine the memorability score based on the size layer.

6. CONCLUSIONS AND RECOMMENDATIONS

The role of visual perception in design processes and its impact on this process have been explained within a theoretical and practical framework within the scope of this thesis. This study aims to examine the change in visual perception depending on object size and location through a computational model. Computational methods previously used to measure visual perception have been examined in detail, and a new computational model has been developed based on existing approaches. A case study was carried out to evaluate the performance of the model. In the case study, objects of different sizes were placed in the experimental environment while keeping colour, texture, and material factors constant. A video recording was taken from the experimental environment and shown to the participants. Participants were then asked to complete a questionnaire indicating their recall of each object. The video used in the case study was split into frames. Object detection was performed on these frames using YOLOv8. Then, the ratios of the detected objects to the image size and their distances from the image center were calculated. Using these data, memorability scores were obtained for each object. The obtained scores and the memorability data obtained from the case study were analyzed using various statistical methods. Thus, the performance of the model was evaluated.

The Pearson and Spearman correlation coefficients obtained as a result of the correlation analyses show that there is a strong positive relationship between the size-based memorability scores obtained from the model and the average memorability scores obtained from the questionnaires. In addition, the coefficient of determination was examined to determine the level of explainability between the two variables. The coefficient of determination between these two variables was found to be 0.545. This means that 55% of the change observed in the survey results is explained by the change in the model results. Furthermore, it was observed that the most memorable and least memorable objects were the same according to both the model results and the survey results. This situation shows that the model has the potential to successfully determine the level of memorability on an object-by-object basis in addition to general patterns. One of the interesting findings of the study is that, unlike Basavaraju et al. (2019), no significant relationship was found in the correlation analyses between centrality scores and survey data. This may be due to the camera moving in accordance with the position

of objects in the video shown to participants. In the video, the camera is constantly moving, and each object comes into focus and then moves out of focus in turn. For this reason, the participants may have focused on objects closer to the camera rather than the object in the center of the visual field. This may have led to findings that differ from studies conducted with images.

In summary, the main findings of the study show that object size has a strong effect on memorability. However, the location of objects within an image was observed to have no statistically significant effect. A distinctive feature of the study is its evaluation of visual perception by taking into account the user's changing perspective throughout the experience of a space. The study's findings can be applied to the fields of advertising, education, and architectural design. For example, product or message sizing can be used as a strategic tool in advertising campaigns. In educational materials, visually highlighting critical information can support memory and enhance learning processes. The developed computational model enables the prediction of the effects of visual perception in architectural design and allows design decisions to be shaped accordingly. It defines the visual hierarchy within a space based on the size of the objects and can offer recommendations on which elements should be emphasized and which should remain in the background. In this way, designers can enhance the user's perception and experience of the space.

The study has a number of limitations. First, the scope of the study was limited to examining visual perception in relation to object size and location. The selected objects were used in white, abstract representations without using texture or material properties. However, it is known that these factors also affect visual perception. The study also did not examine the effect of semantic features. However, semantic features also have a significant impact on visual attention and memorability. Additionally, the number of objects was limited to nine. It was observed that the object detection model missed or incorrectly detected objects in some image frames. In such cases, manual calculations were carried out. This is also inefficient in terms of time usage and computation sensitivity. Furthermore, memorability computations were carried out on image frames taken from the video rather than directly from the video itself. Finally, the number of participants in the case study was limited. Consequently, the analyses

were conducted on a narrow data set. This situation is a factor that reduces the generalizability of the study findings.

Future studies could focus on a more developed model that performs real-time data analysis and processing. In addition, the study can be conducted with more participants from the same occupational groups to reduce the effect of individual differences. In this way, more distinct patterns may emerge from the data, and the reliability and generalizability of the model will increase. Furthermore, increasing the number of objects used in the experiment may make it possible to observe the size effect more clearly. Preparing a larger dataset for training the YOLO object detection model would be beneficial for improving object detection accuracy and accelerating the process. The colour, texture, semantic, and material properties of objects can also be included in the model and analyzed. These enhancements can enable the creation of a model that is closer to human visual perception.



REFERENCES

- Akkurt, E.** (2019). Kullanıcının yapı cephelerindeki görsel algısında Gestalt Kuramının etkileri: Diyarbakır'da 3 farklı bulvar değerlendirmesi (Master's thesis, Institute of Science).
- Alaçam, S.** (2019). Modelling Visual Perception of Timevarying Scenes Using Likelihood. Istanbul Technical University Scientific Research Projects Department. Project no: mga-2019-41666
- Aslan, F., Aslan, E., & Atik, A.** (2015). İÇ MEKANDA ALGI. *İnönü Üniversitesi Sanat ve Tasarım Dergisi*, 5(11).
- Beyoğlu, A.** (2015). Sanat Eğitiminde Algı, Görsel Algı ve Yanılsama Victor Vasarely'nin Çalışmaları Üzerine Bir İnceleme. *Trakya Üniversitesi Sosyal Bilimler Dergisi*, 17(1), 333-348.
- Basavaraju, S., Gaj, S., & Sur, A.** (2019). Object memorability prediction using deep learning: Location and size bias. *Journal of Visual Communication and Image Representation*, 59, 117-127.
- Basu, T., & Ghosh, M.** (2017). Visual perception of space and parametric design. *GSTF Journal of Engineering*.
- Bittermann, M. S., & Ciftcioglu, O.** (2008). Visual perception model for architectural design. *Journal of Design Research*, 7(1), 35-60.
- Bittermann, M. S., Ciftcioglu, O., Bhatt, M., & Schultz, C.** (2013). Ambient environment analysis by means of perception. In 20th Int. Workshop Intelligent Computing in Engineering-EG-ICE.
- Brunswick, E.** (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62(3), 193-217
- Cantemir, E.** (2025). Determination of the effects of visual and auditory stimuli on spatial perception in enclosed public spaces using fuzzy logic. Doctoral thesis: Eskisehir Technical University.
- Castelhana, M. S., & Henderson, J. M.** (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology: Human Perception and Performance*, 34(3), 660.
- Castelhana, M. S., & Krzyś, K.** (2020). Rethinking space: A review of perception, attention, and memory in scene processing. *Annual Review of Vision Science*, 6(1), 563-586.
- Celikkale, B., Erdem, A., & Erdem, E.** (2015). Predicting memorability of images using attention-driven spatial pooling and image semantics. *Image and Vision Computing*, 42, 35-46.
- Ciftcioglu, O., & Bittermann, M. S.** (2013). Fusion of perceptions in architectural design. In eCAADe 2013: Computation and Performance –

Proceedings of the 31st International Conference on Education and research in Computer Aided Architectural Design in Europe.

- Çağlayan, S. & Korkmaz, M. ve Öktem, G.** (2014). Sanatta görsel algının literatür açısından değerlendirilmesi. *Eğitim ve Öğretim Araştırmaları Dergisi*, 3(1), 162-165.
- Demirarslan, D.** (2006). İç Mekan Tasarımına Giriş. Kocaeli Üni Yayın No: 228 / GSF Yayın no: 2; Kocaeli.
- Diaz, D., Duchowski, A. T., Volonte, M., Robb, A., Pagano, C. C., & Babu, S. V.** (2024). The Impact of Color and Object Size on Spatial Cognition and Object Recognition in Virtual Reality. In Proceedings of the 17th ACM SIGGRAPH Conference on Motion, Interaction, and Games (pp. 1-11).
- Dubey, R., Peterson, J., Khosla, A., Yang, M. H., & Ghanem, B.** (2015). What makes an object memorable? In Proceedings of the IEEE International Conference on Computer Vision (pp. 1089-1097).
- Erişti, S. D. B.** (2016). Görsel algı kuramlarına göre reklam içerikli tasarımların değerlendirilmesi. *Art-e Sanat Dergisi*, 9(17), 313-342.
- Eristi, S. D., Sahin-izmirli, O., Izmirli, S., Fırat, M., & Haseski, H. I.** (2010). An evaluation of educational website designs from the perspective of visual perception theories: A study on scale development. *Contemporary Educational Technology*, 1(4), 348-366.
- Erişti, S. D., Uluuysal, B., & Dindar, M.** (2013). Görsel Algı Kuramlarına Dayalı Etkileşimli Bir Öğretim Ortamı Tasarımı ve Ortama İlişkin Öğrenci Görüşleri. *Anadolu Journal of Educational Sciences International*, 3(1).
- Eryayar, E.** (2011). Endüstri ürünleri tasarımında Gestalt teorisi uygulaması. *Journal of World of Turks*, 3(2), 125-133.
- Gibson, J. J.** (1954). The visual perception of objective motion and subjective movement. *Psychological Review*, 61(5), 304.
- Goldstein, E. B., & Brockmole, J. R.** (2002). Sensation and perception (Vol. 90). Pacific Grove, CA, USA: Wadsworth-Thomson Learning. 9. baskı.
- Gordon, I. E.** (2004). Theories of visual perception. Psychology Press.
- Grima Murcia, M. D., Ortíz, M. J., López-Gordo, M. A., Ferrández, J. M., Sánchez Ferrer, F., & Fernández, E.** (2019). Neural representation of different 3D architectural images: An EEG study. *Integrated Computer-Aided Engineering*, 26(2), 197-205.
- Isola, P., Xiao, J., Torralba, A., & Oliva, A.** (2011). What makes an image memorable? In CVPR 2011 (pp. 145-152). IEEE.
- Jaglarz, A.** (2011). Perception and illusion in interior design. In Stephanidis, C. (Ed.), Universal Access in Human-Computer Interaction. Context Diversity: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011 (pp. 358-364). Springer Berlin Heidelberg.
- Kaufmann, G.** (1980). Imagery, language, and cognition: Toward a theory of symbolic activity in human problem-solving. Berlin: Universitetsforl.

- Kiefer, P., Giannopoulos, I., Raubal, M., & Duchowski, A. (2017).** Eye tracking for spatial research: Cognition, computation, challenges. *Spatial Cognition & Computation*, 17(1-2), 1-19.
- Kim, N., & Lee, H. (2020).** Visual attention in retail environments: design analysis using HMD based VR system integrated eye-tracking. In 25th International Conference on Computer-Aided Architectural Design Research in Asia, CAADRIA 2020 (pp. 631-640). CAADRIA.
- Kim, N., & Lee, H. (2021).** Assessing consumer attention and arousal using eye-tracking technology in virtual retail environment. *Frontiers in Psychology*, 12, 665658.
- Kramer, M. A., Hebart, M. N., Baker, C. I., & Bainbridge, W. A. (2023).** The features underlying the memorability of objects. *Science Advances*, 9(17), eadd2981.
- Lisińska-Kuśnierz, M., & Krupa, M. (2018).** Eye tracking in research on perception of objects and spaces. *Technical Transactions*, 2018(115), 5-22.
- Martí-Testón, A., Muñoz, A., Gracia, L., & Solanes, J. E. (2023).** Using WebXR Metaverse Platforms to Create Touristic Services and Cultural Promotion. *Applied Sciences*, 13(14), 8544. <https://doi.org/10.3390/app13148544>
- Morales, J., & Firestone, C. (2023).** Philosophy of Perception in the Psychologist's Laboratory. *Current Directions in Psychological Science*, 32(4), 307-317.
- Ör, R., & Özen, E. S. (2024).** İç Mekânda Görsel Algıyı Etkileyen Faktörler: Görsel Algı ve Aidiyet Kavramları Üzerinden Mekânı Okumak. *International Journal of Social and Humanities Sciences Research (JSHSR)*, 11(108), 1126-1145.
- Özsavaş, N. (2016).** İç Mekan Tasarımında Renk Algısı. *Art-e Sanat Dergisi*, 9(18), 449-460.
- Reggente, N., Essoe, J. K., Baek, H. Y., & Rissman, J. (2020).** The method of loci in virtual reality: Explicit binding of objects to spatial contexts enhances subsequent memory recall. *Journal of Cognitive Enhancement*, 4, 12–30.
- Reghukumar, A. (2019).** Sense and sensitivity in architecture-the use of five senses in space making. *International Research Journal of Architecture and Planning*, 4(3), 97-101.
- Rounds, J. D., Cruz-Garza, J. G., & Kalantari, S. (2020).** Using posterior EEG theta band to assess the effects of architectural designs on landmark recognition in an urban setting. *Frontiers in Human Neuroscience*, 14, 584385.
- Seven, F. (2024).** Spatial.io Ortamında Nesne Boyutu ve Görünürlüğe İlişkin Görsel Algı Eşiğinin Ortaya Çıkarılması. XVIII. Mimarlıkta Sayısal Tasarım Ulusal Sempozyumu, Balıkesir Üniversitesi Yayınları.

- Seven, F., & Alaçam, S. (2025).** 4th International Graduate Research Symposium IGRS'25 E-Abstract Book, May 12-14, 2025 ITU, Istanbul, Türkiye. (in proceeding)
- Shomstein, S., Malcolm, G. L., & Nah, J. C. (2019).** Intrusive effects of task-irrelevant information on visual selective attention: Semantics and size. *Current Opinion in Psychology*, 29, 153–159.
- Şen Bayram, A. K., Güzelci, O. Z., & Alaçam, S. (2023).** Memorability of Spatial Features In Virtual Reality Süleymaniye Experience. *Architecture and Planning Journal (APJ)*, 28(3), 36.
- Şensoy, E. S. (2024).** A Research on Factors Affecting Visual Perception in Architecture. *User Experience & Urban Creativity Journal (UXUC)*, 6(2).
- Sohan, M., Sai Ram, T., Reddy, R., & Venkata, C. (2024).** A review on YOLOv8 and its advancements. In *International Conference on Data Intelligence and Cognitive Informatics* (pp. 529–545). Springer, Singapore.
- Tsirtsakis, P., Zacharis, G., Maraslidis, G. S., & Fragulis, G. F. (2025).** Deep learning for object recognition: A comprehensive review of models and algorithms. *International Journal of Cognitive Computing in Engineering*.
- Tural, P. (2017).** George Berkeley'in Algı Teorisi (Master's thesis, Necmettin Erbakan University (Turkey)).
- Turgay Zıraman, A., & Imamoğlu, Ç. (2020).** Visitor attention in exhibitions: The impact of exhibit objects' ordinal position, relative size, and proximity to larger objects. *Environment and Behavior*, 52(4), 343–370.
- Wade, N., & Swanston, M. (2013).** Visual Perception: An Introduction. Psychology Press.
- Wang, X. (2016).** Deep learning in object recognition, detection, and segmentation. *Foundations and Trends® in Signal Processing*, 8(4), 217–382.
- Wilson, C. J., & Soranzo, A. (2015).** The use of virtual reality in psychology: A case study in visual perception. *Computational and mathematical methods in medicine*, 2015(1), 151702.
- Wilson, H., Chen, X., Golbabaee, M., Proulx, M. J., & O'Neill, E. (2024).** Feasibility of decoding visual information from EEG. *Brain-computer interfaces*, 11(1-2), 33-60.
- Withagen, R. & van der Kamp, J. (2010).** Towards a new ecological conception of perceptual information: Lessons from a developmental systems perspective. *Human Movement Science*, 29(1), 149-163.
- Wu, X., & Zhao, Y. (2023).** An integrated model of five sense perceptions and memorable tourism experiences: mediating effects of positive emotion. *Asia Pacific Journal of Tourism Research*, 28(8), 860-876.
- Yılmaz, A., & Kaya, A. (2023).** Görsel iletişim tasarımında görsel algı kuramlarının değerlendirilmesi. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi*, 13(3), 3009-3032.

- Yılmaz, Ö.** (2004). Mimari Mekanda Görsel Algı ve Manipölasyon İlişkilerinin İrdelenmesi. (154575) [Master's thesis, Yıldız Teknik Üniversitesi]. YÖK Tez. <https://tez.yok.gov.tr/UlusalTezMerkezi/>
- Yoshida, H., Furukawa, A., Ikegami, T., Fukuzumi, S. I., & Furuta, K.** (2017). Quantifying visual attention Modeling effects of object size. In *Advances in Ergonomics Modeling, Usability & Special Populations: Proceedings of the AHFE 2016 International Conference*, July 27–31, 2016, Walt Disney World®, Florida, USA (pp. 173–187). Springer.
- Yu, H., Bai, G., & Wu, L.** (2018). Application of perception theory in hotel interior design. *Open Journal of Applied Sciences*, 8(07), 285.
- Yüce, O.** (2018). Görsel algı ve tasarım kriterlerinin mekan olgusuna etkisi (Master's thesis, Mimar Sinan Güzel Sanatlar Üniversitesi).
- Zou, Z., & Ergan, S.** (2023). Towards emotionally intelligent buildings: A convolutional neural network based approach to classify human emotional experience in virtual built environments. *Advanced Engineering Informatics*, 55, 101868.
- Url-1** < <https://ago.ca/agoinsider/art-pick-week-large-two-forms>>, date retrieved 10.05.2025.
- Url-2** < <https://frontdesk.co.in/theory-of-design/elements-of-design/>>, date retrieved 10.05.2025.
- Url-3** < <https://maricamckee.com/modern-design-inspiration-pop-color/>>, date retrieved 10.05.2025.
- Url-4** < <https://3dwarehouse.sketchup.com/>>, date retrieved 10.05.2025.



APPENDICES

APPENDIX A: ITU Social and Human Sciences Scientific Research and Publication Ethics Committee Documents



APPENDIX A

İSTANBUL TEKNİK ÜNİVERSİTESİ
SOSYAL VE BEŞERİ BİLİMLER İNSAN ARAŞTIRMALARI ETİK KURULU
DEĞERLENDİRME SONUCU

PROJE NUMARASI: 025-03-660

TARİH:

1.	Projenin Adı ve Yürütücüsü: Hesaplamalı Görsel Algı Modeli - Doç.Dr. Sema Alaçam
2.	Projenin amacı, farklı boyutlarda nesnelerin hatırlanabilirliği üzerinde bir inceleme yaparak, nesne boyutunun görsel algı üzerindeki etkilerini detaylı bir şekilde analiz etmektir. Bu bağlamda, nesnelerin boyutlarının, kullanıcıların hatırlanabilirlik düzeylerini nasıl etkilediği üzerine bir çalışma yürütülecek ve görsel algının hesaplanması için gerekli veriler toplanacaktır. Bu çalışma, görsel algı ve bellek üzerine yapılan araştırmalara katkıda bulunmayı hedeflemektedir.
3.	Projenin Yöntemi (azami 100 kelime) : Farklı boyutta 9 nesneyi içeren sanal bir deney ortamı tasarlandı. Deney ortamında yapılan tur sırasında ekran kaydı alındı. Elde edilen video kareleri üzerinden nesne tespiti ve oran hesaplama işlemleri yoluyla nesnelerin hatırlanabilirlik dereceleri belirlendi. Etik kurul onayı verilmesi durumunda ekran kaydı katılımcılara gösterilecek ve bu aşamadan sonra katılımcılara nesnelerin hatırlanabilirliği üzerine hazırlanan anket yönlendirilecektir. Katılımcılardan toplanan veriler, hesaplama yoluyla elde edilen verilerle karşılaştırılacaktır.

Figure A.1: Approval of the ITU Social and Human Sciences Scientific Research and Publication Ethics Committee.

- ☒ Proje değerlendirilmiş ve etik açıdan uygun BULUNMUŞTUR.
☐ Proje değerlendirilmiş ve etik açıdan uygun BULUNMAMIŞTIR.
☐ Proje değerlendirilmiş ve etik kurul belgesine GEREK DUYULMAMIŞTIR.

Projenin Adı ve Yürütücüsü: Hesaplamalı Görsel Algı Modeli - Doç.Dr. Sema Alaçam

TARİH

Prof. Dr. Leman Figen GÜL
İstanbul Teknik Üniversitesi
Mimarlık Bölümü
(Başkan)

Prof. Dr. Gözde ÇOLAKOĞLU SARI
İstanbul Teknik Üniversitesi
Müzikoloji Bölümü

Prof. Dr. Mehmet ÖZGER
İstanbul Teknik Üniversitesi
İnşaat Mühendisliği Bölümü

Prof. Dr. Mehmet KUÇUKMEHMETOĞLU
İstanbul Teknik Üniversitesi
Şehir ve Bölge Planlama Bölümü

Doç. Dr. Eda YÜCESOY
İstanbul Teknik Üniversitesi
Şehir ve Bölge Planlama Bölümü

Prof. Dr. Burhaneddin İZGİ
İstanbul Teknik Üniversitesi
Matematik Mühendisliği Bölümü

Doç. Dr. Aysun HIZIROĞLU AYGÜN
İstanbul Teknik Üniversitesi
Ekonomi Bölümü

Dr. Öğr Üyesi Mehmet Ali DOĞAN
İstanbul Teknik Üniversitesi
Sosyoloji Bölümü

Av. Orhan Veli DOĞAN
İstanbul Teknik Üniversitesi
Hukuk Müşavirliği

Figure A.1 (continued): Approval of the ITU Social and Human Sciences Scientific Research and Publication Ethics Committee.



CURRICULUM VITAE

Name Surname : Fatma Seven

EDUCATION :

- **B.Sc.** : 2020-2024, Kadir Has University, Engineering and Natural Sciences, Computer Engineering
- **B.Sc.** : 2017-2022, Kadir Has University, Art and Design, Architecture
- **M.Sc.** : 2022-2025, Istanbul Technical University, Informatics, Architectural Design Computing

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

Seven, F., & Alaçam, S. (2025). 4th International Graduate Research Symposium, IGRS'25 E-Abstract Book, May 12-14, 2025 ITU, Istanbul, Türkiye. (in proceeding)

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

Seven, F. (2024). Spatial.io Ortamında Nesne Boyutu ve Görünürlüğe İlişkin Görsel Algı Eşiğinin Ortaya Çıkarılması. XVIII. Mimarlıkta Sayısal Tasarım Ulusal Sempozyumu, Balıkesir Üniversitesi Yayınları.