Interlacing Latent Features: Synthesis of Past and Present in Architectural Design through Artificial Intelligence in a Case Study of Japanese Houses

Interlacing Latent Features: Synthesis of Past and Present in Architectural Design through Artificial Intelligence in a Case Study of Japanese Houses

A Thesis Submitted to the Department of Architecture
Harvard University Graduate School of Design,

by

Rio Kobayashi

In Partial Fulfillment of the Requirements for the Degree of
Master of Architecture

December 2023

(signature of author)

Rio Kobayashi

Rio Kobayashi

(signature of advisor)

Andrew Witt

Interlacing Latent Features: Synthesis of Past and Present in Architectural Design through Artificial Intelligence in a Case Study of Japanese Houses

Harvard University
Graduate School of Design
Master of Architecture I

Rio Kobayashi

Advisor: Andrew Witt

Tapestry of Architectural Data

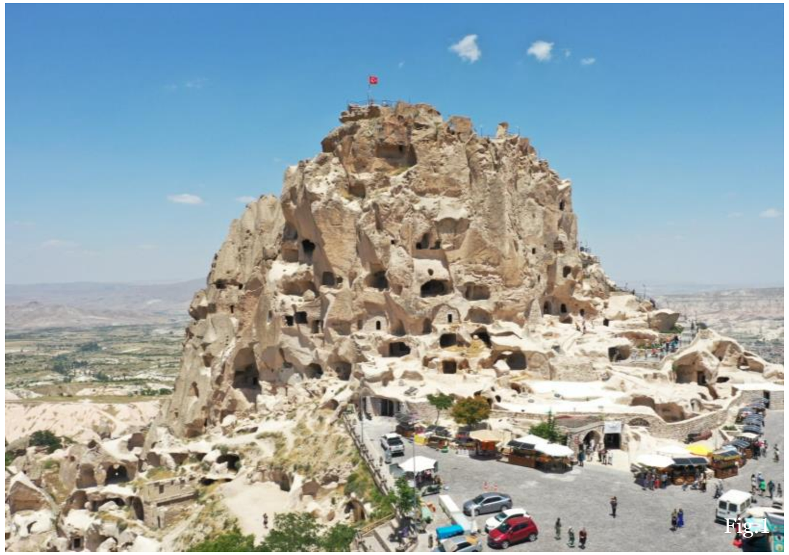


Fig.1


Fig.2


Fig.3


Fig.4

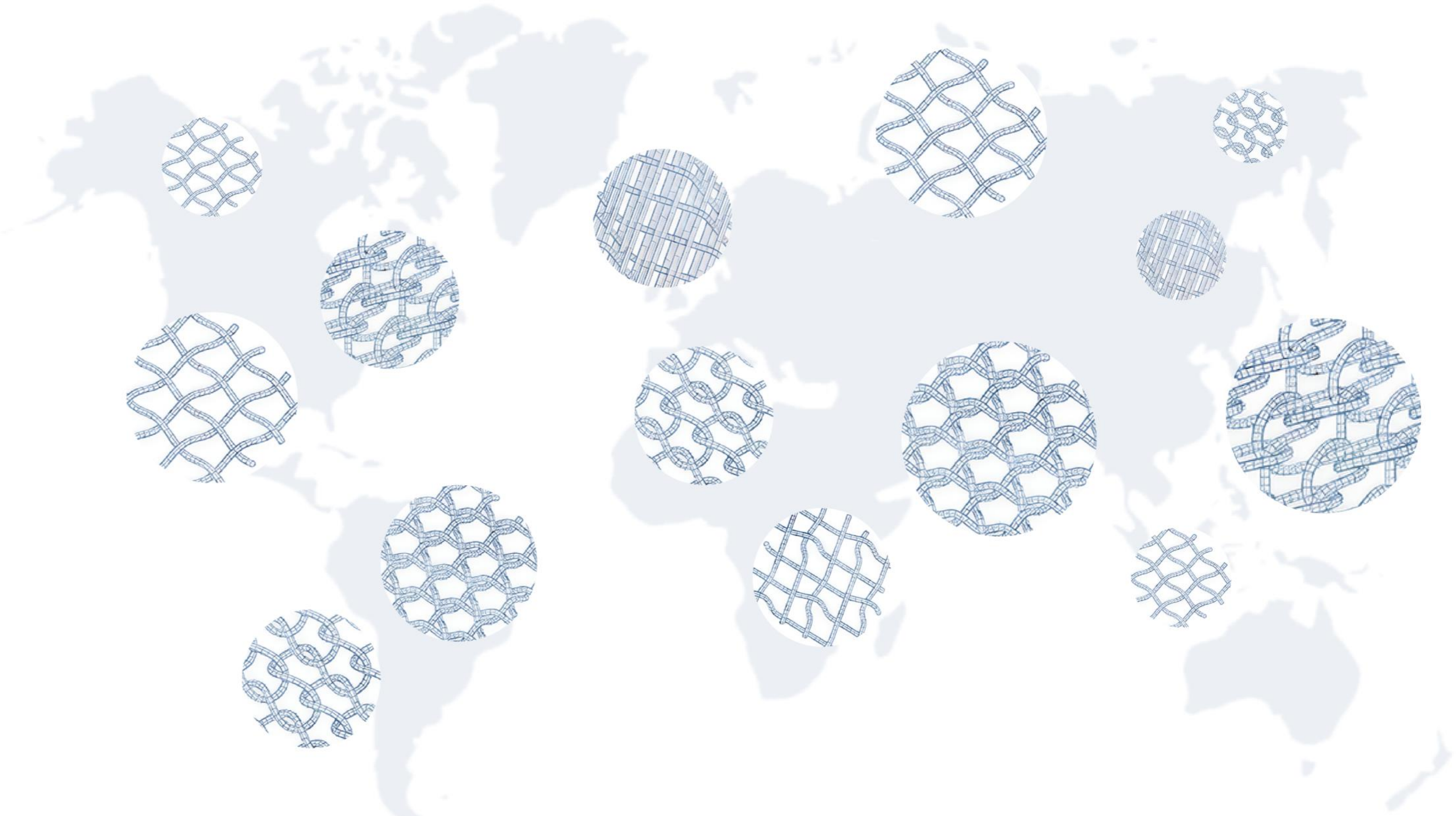

Fig.5


Fig.6

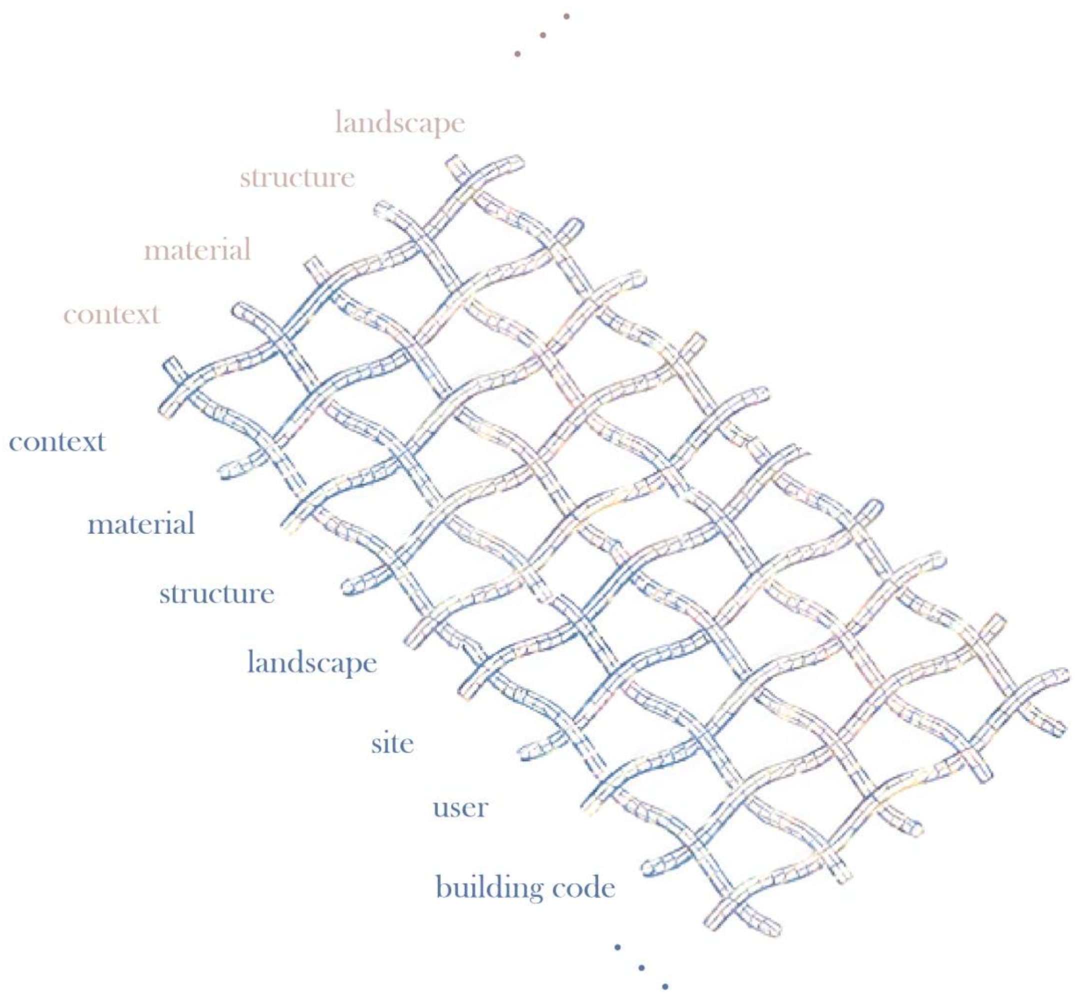
Throughout history, humans have created and accumulated rich stores of architectural knowledge and data.
This vast repository (of drawings, images, writings, and other media representations) includes architecture with and without architects.

While some forms of knowledge and data fade into oblivion, others persist and continue to evolve within contemporary design landscapes.

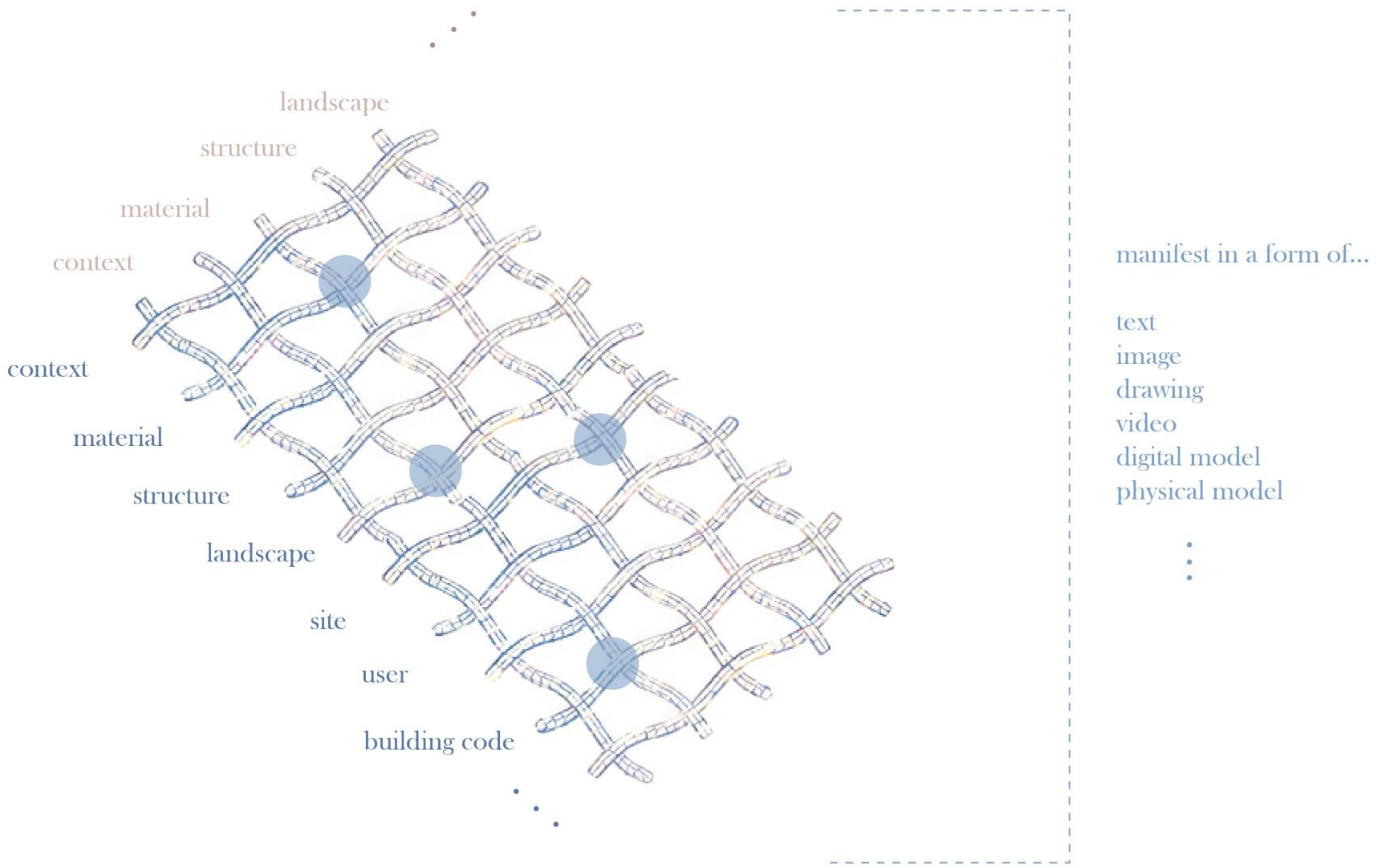
Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt



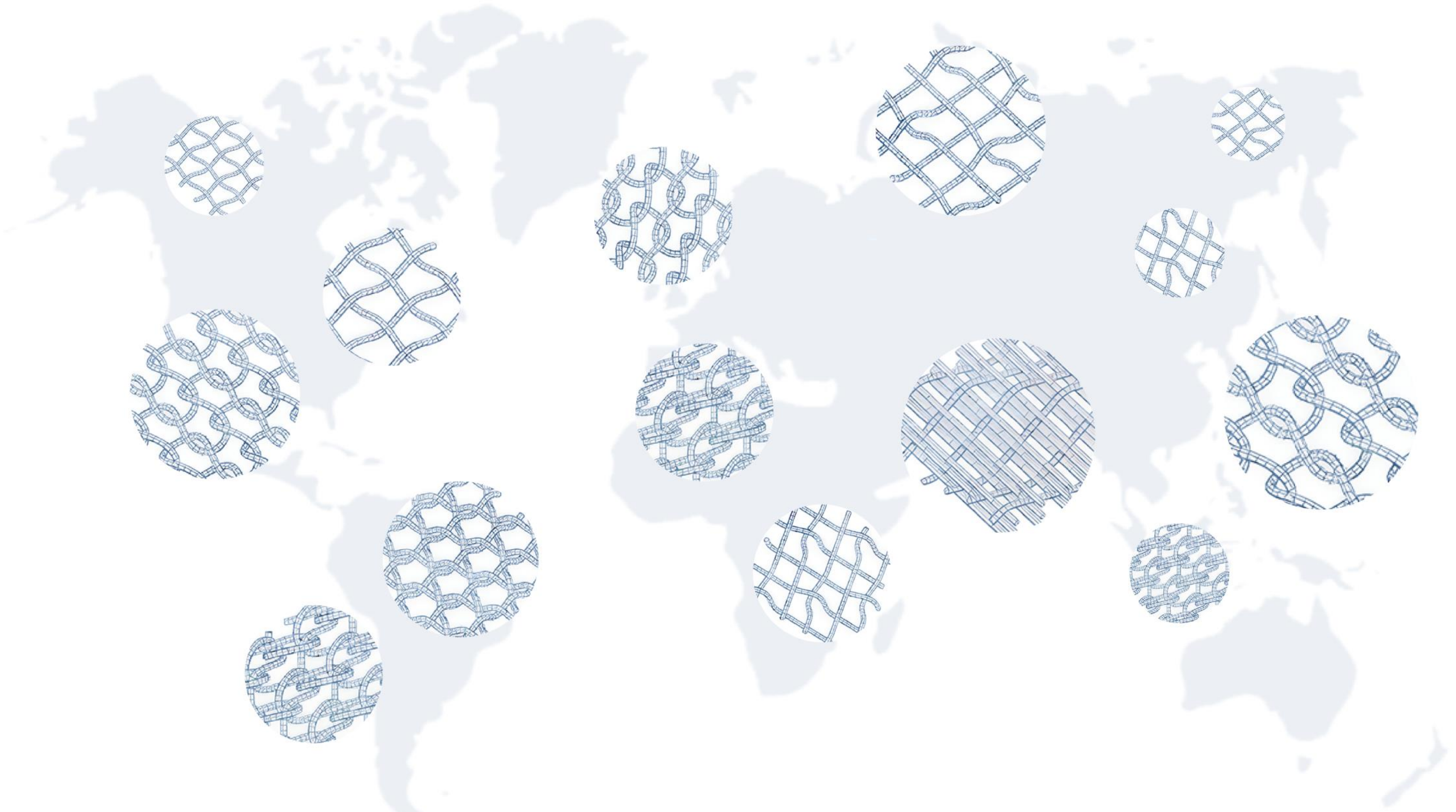
Knowledge and data inherited from the past can be deeply embedded in our conscious and subconscious thinking.
It stretches across the temporal and spatial expanse of human civilization like a grand tapestry.



My project likens the web of knowledge to such a tapestry consisting of diverse threads each representing a distinct architectural environment:
a line for form, a cross-stitch for function, delicate silk filament for aesthetics, sturdy twine for structure.



Threads of information manifest in various forms, from text to 3D model,
while the manner of weaving can be likened to how every design practitioner revisits and reinterprets each thread and tool they come across..



Threads are interlaced to create patterns and motifs. Those that are bold and pronounced remind me of grand edifices and monuments, while others that are more subtle or imperceptible might represent the nuances of vernacular dwellings and modest charm of utilitarian spaces. Various movements have taken place in time, and the weaving manner has changed many times in the past.

Contextual
Architecture

Contextual architecture, also known as Contextualism is a philosophical approach in architectural theory that refers to **the designing of a structure in response to the literal and abstract characteristics of the environment in which it is built.** Contextual architecture contrasts modernist architecture, which value the imposition of their own characteristics and values upon the built environment. **Contextual architecture is usually divided into three categories: vernacular architecture, regional architecture, and critical regionalism**, all of which also inform the complementary architecture movement. [1]

Vernacular
Architecture

Vernacular architecture is building done **outside any academic tradition, and without professional guidance.** This category encompasses a wide range and variety of building types, **with differing methods of construction, from around the world, both historical and extant, representing the majority of buildings and settlements created in pre-industrial societies.** It constitutes 95% of the world's built environment, as estimated in 1995 by Amos Rapoport, as measured against the small percentage of new buildings every year designed by architects and built by engineers. **It usually serves immediate, local needs; is constrained by the materials available in its particular region; and reflects local traditions and cultural practices.** Traditionally, the study of vernacular architecture did not examine formally schooled architects, but instead that of the design skills and tradition of local builders, who were rarely given any attribution for the work. [2]

Indigenous
Architecture

The field of Indigenous architecture refers to the study and practice of architecture **of, for and by Indigenous people**. It is a field of study and practice in the United States, Australia, Aotearoa/New Zealand, Canada, Arctic area of Sápmi and many other countries where **Indigenous people have a built tradition or aspire translate or to have their cultures translated in the built environment.** This has been extended to landscape architecture, urban design, planning, public art, placemaking and other ways of contributing to the design of built environments. [3]

Traditional
Architecture

Architecture **based on a way of thinking, behaving, or doing something that has been used by the people in a particular group, family, society, etc., for a long time** : following the tradition of a certain group or culture. [4]

Regional
Architecture

The main idea inherent in the concept of regional architecture/regionalism is **context- specific architecture.** This, in turn, is **based on knowledge of the history of a place, climatic conditions, concerns, materiality, topology, ecology, environmental conditions, culture and traditions, skills, tools, and technology available in a particular area.** The driving idea behind Critical Regionalism is resistance to the standardization of Architecture. The increasing standardization is a modern phenomenon caused by globalization. [5]

Knowledge/data can often be biased and misrepresentative, a reminder that architectural identity is closely intertwined with context, time, and place. Both architectural form and knowledge can be influenced through the invention and popularization of new terms and connotations associated with them: contextual vernacular, regional, indigenous, traditional

Fig.7

Tokyo, Japan

Fig.8

Belair, Luxembourg

Fig.9

London, United Kingdom

Fig.10

Colombes, France

Today, one can argue that much of our knowledge/data and access to it has been largely overwritten by capitalist motives and urbanist biases thus resulting in the homogenization of design and deemphasizing sustainability and diversity in design, a homogeneity that benefits from anonymity.


Fig.11


Fig.12


Fig.13

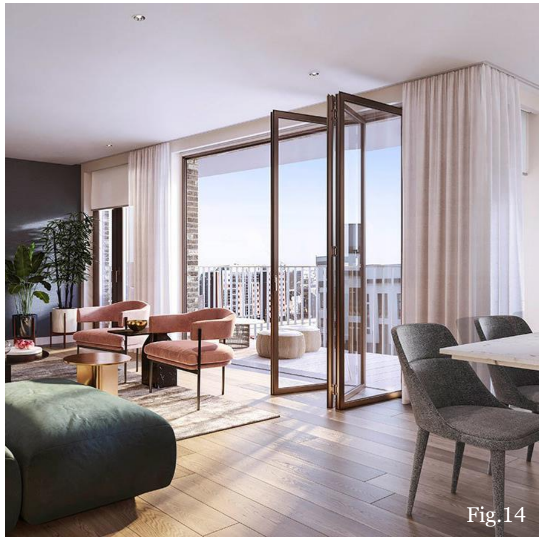


Fig.14

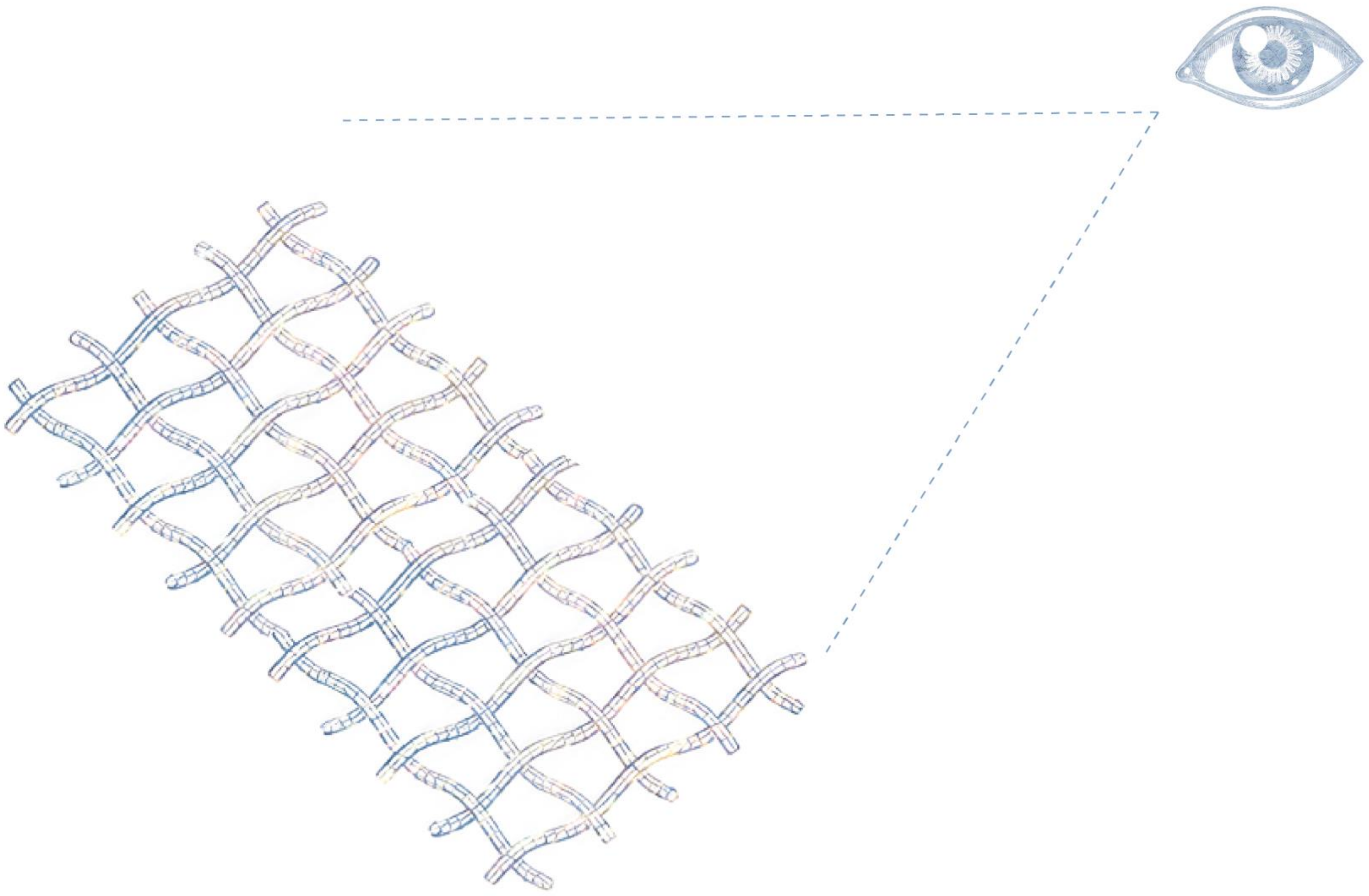
Sustainable, Diverse

Cost-effective, Uniform

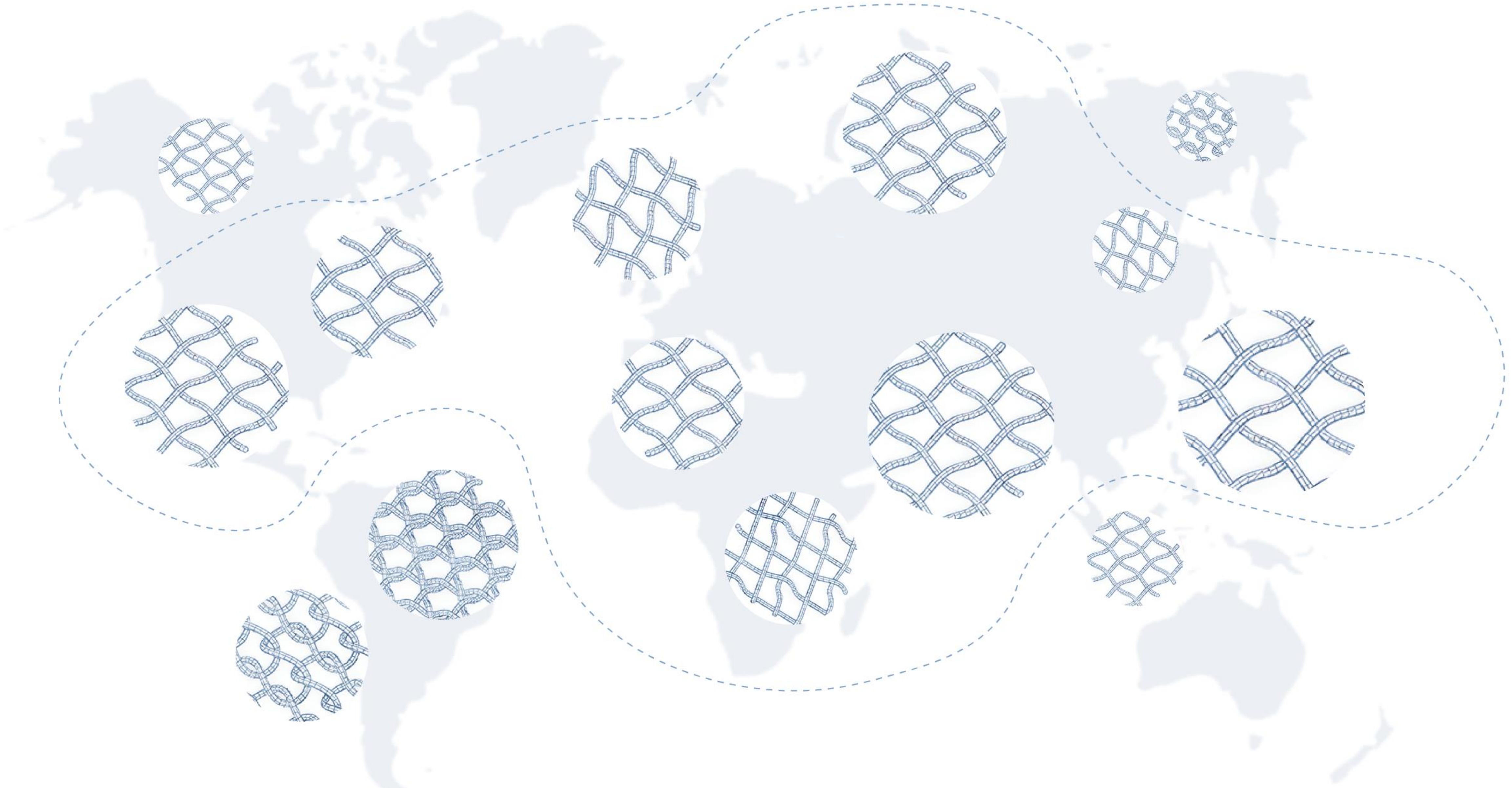
?

?

We question; will our society ever see a residential building like Cappadocia again?



As designers, we learn and digest information both individually and collectively. It is intimate and cognitively demanding work,
and there is a limit to our computing capacity. Hence the creation from our own learning is inevitably exposed as biased and limited as well.



This period of exhaustion is likely where the neglect and erasure of data/novelty happens.
Can there be a way to balance our current architectural tapestry from the encroachments of globalization and homogenization?

Machine Intelligence and Perception

Fig.15

Machines have served humans in various ways for a long time and as predicted by experts decades ago,
it has achieved an intelligence via neural networks that now closely mirrors the cumulative nature of collective human intelligence.

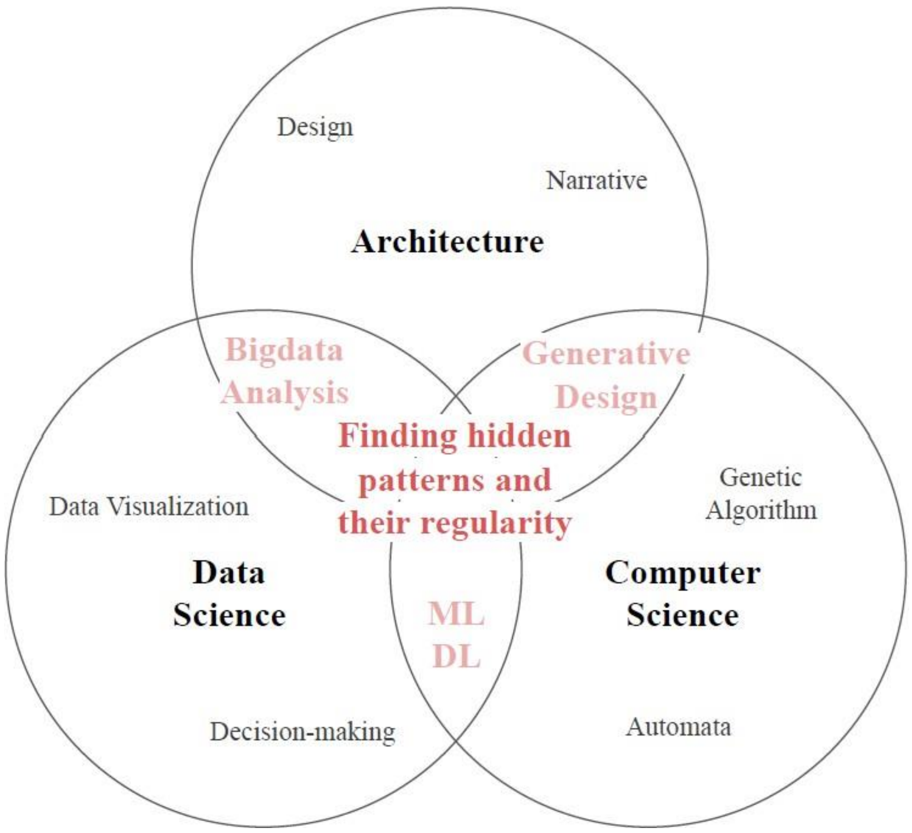


Fig.16

Architecture is hardly exempt, as we have come to appreciate and acknowledge
the added value these machine learning technologies and digital innovations bring to design as a discipline and practice.

"With the advent of computer vision, for the first time, art and design can be quantified. Never before have we had the power to attribute artwork with the support of AI to confirm artist technique with data. We are in an exciting period in technological design with wide implications of this innovation in a variety of fields." [6]

The advent of machine brain and vision now allows art and design to be quantified.

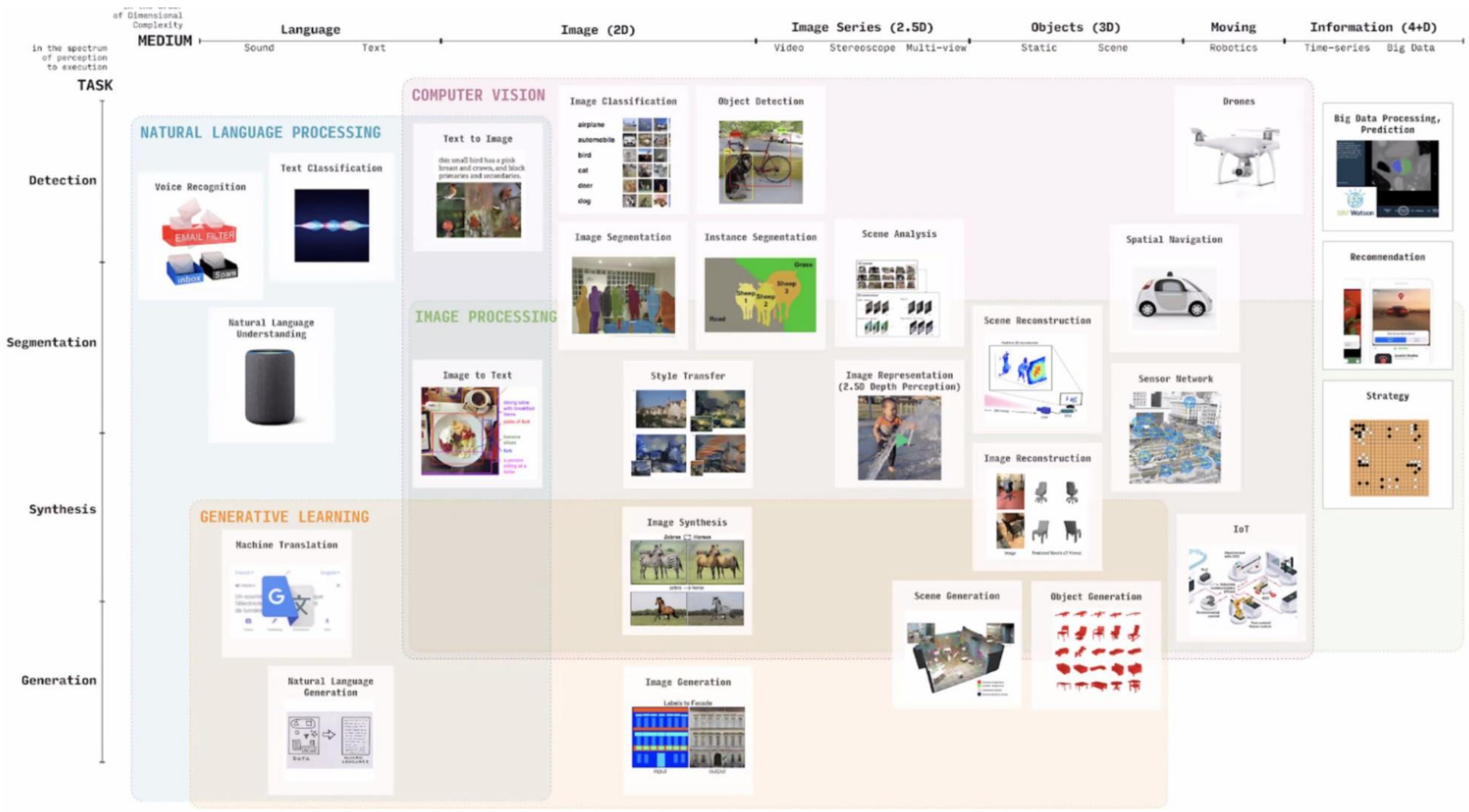
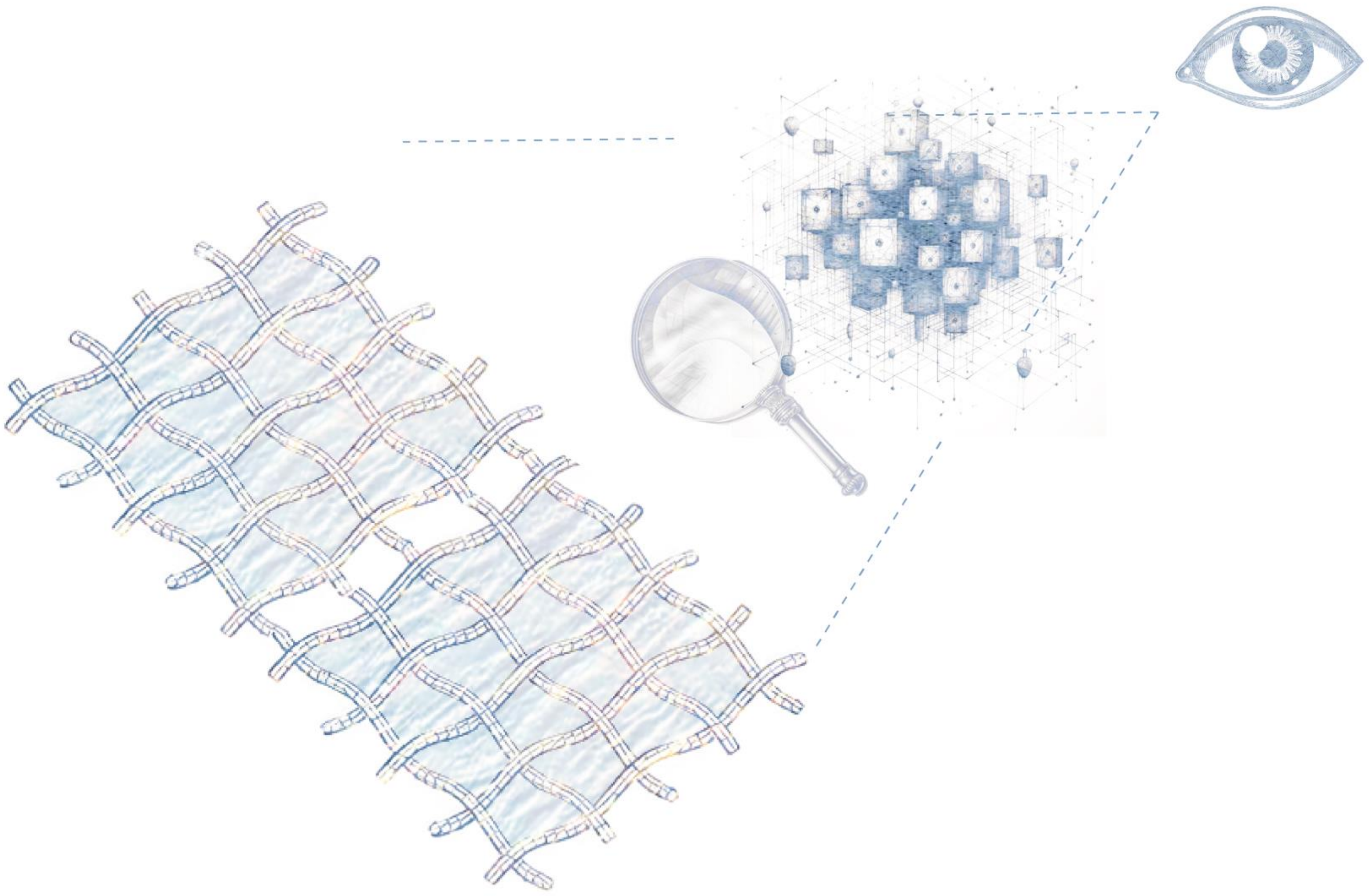


Fig.17

Machine intelligence can take varied forms: Classification, evaluation, optimization, generation.

 Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt



One of the more interesting capacities of machine intelligence is
its ability to identify hidden features and fragments that are typically imperceptible to humans, and then transmit them into their creations.

Open Access Article

Deep Learning Model for Form Recognition and Structural Member Classification of East Asian Traditional Buildings

by Seung-Yeul Ji and Han-Jong Jun *

School of Architecture, Hanyang University, Seoul 04763, Korea

* Author to whom correspondence should be addressed.

Sustainability **2020**, *12*(13), 5292; **https://doi.org/10.3390/su12135292**

Submission received: 17 May 2020 / Revised: 17 June 2020 / Accepted: 22 June 2020 / Published: 30 June 2020

(This article belongs to the Special Issue The Exploration of Sustainability in Traditional Rural Buildings)

Download Browse Figures Versions Notes

Abstract

The unique characteristics of traditional buildings can provide fresh insights for sustainable building development. In this study, a deep learning model and methodology were developed for classifying traditional buildings by using artificial intelligence (AI)-based image analysis technology. The model was constructed based on expert knowledge of East Asian buildings. Videos and images from Korea, Japan, and China were used to determine building types and classify and locate structural members. Two deep learning algorithms were applied to object recognition: a region-based convolutional neural network (R-CNN) to distinguish traditional buildings by country and you only look once (YOLO) to recognise structural members. A cloud environment was used to develop a practical model that can handle various environments in real time.

Keywords: East Asia; traditional buildings; deep learning; artificial intelligence; region-based convolutional neural network (R-CNN); you only look once (YOLO); cloud computing

1. Introduction

Artificial intelligence (AI) is considered one of the greatest revolutions in human history [1]. To some degree, AI has transcended human judgement at classifying and making decisions [2]. In this study, AI deep learning technology was applied to traditional buildings, which has lagged behind other field [3] in terms of applications of computer technology.

Although East Asian countries can trace their cultures to Chinese civilisation, they have evolved with their own unique characteristics. For example, the traditional architectural style of each country varies according to purpose. In China, the country's vast landmass means that the style changes regionally according to the climatic conditions. In northern China, which has little rainfall and people tend to be frugal, roofs have a slightly emphasised curvature. South of the Yangtze River, which receives heavy rainfall and has a mild climate, the curves are more elaborate and rise up around the eaves. In Japan, wooden architecture techniques were altered to help buildings withstand earthquakes. Korea placed importance on heating and insulation because of its four distinct seasons and emphasised simplicity owing to Confucian philosophy [4].

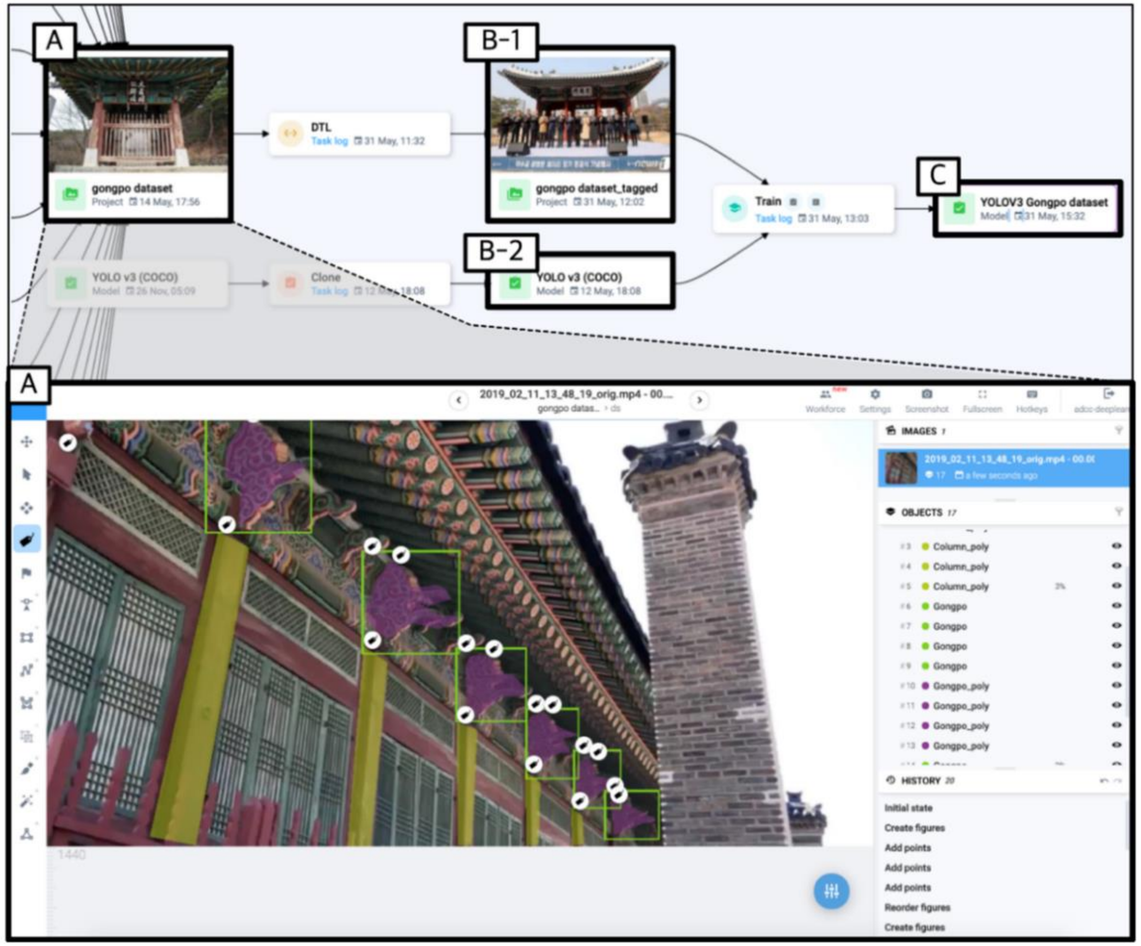
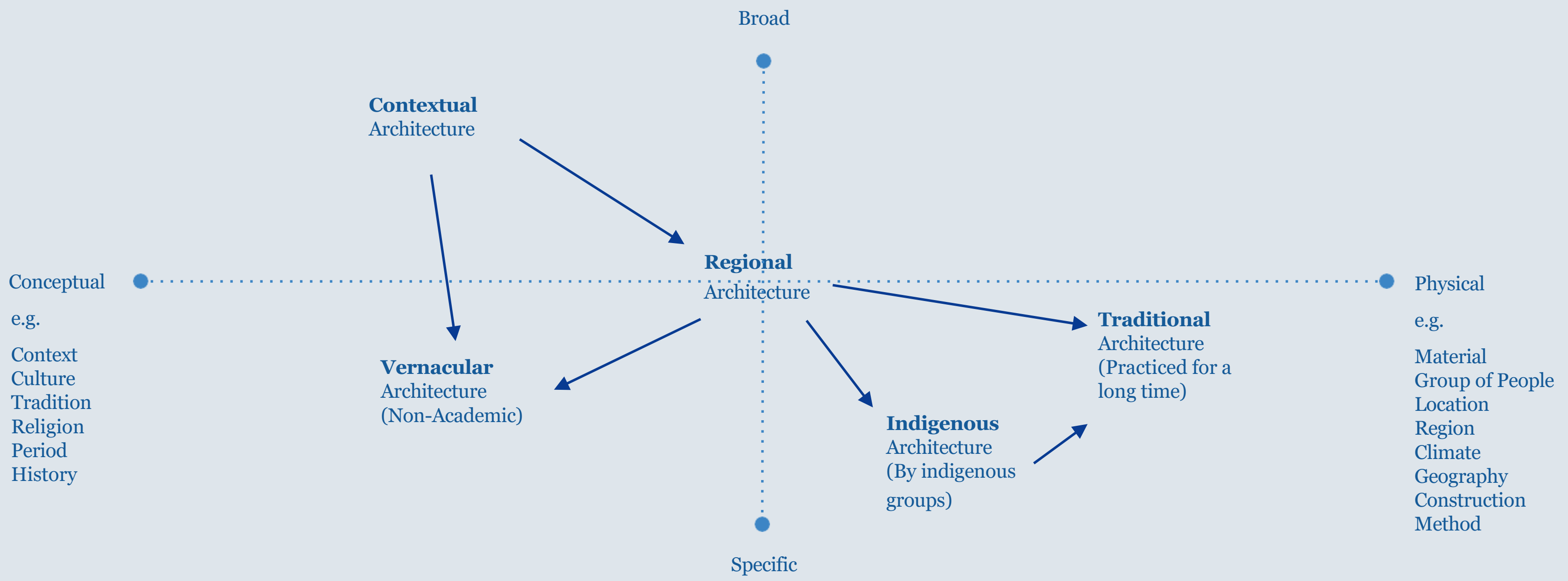
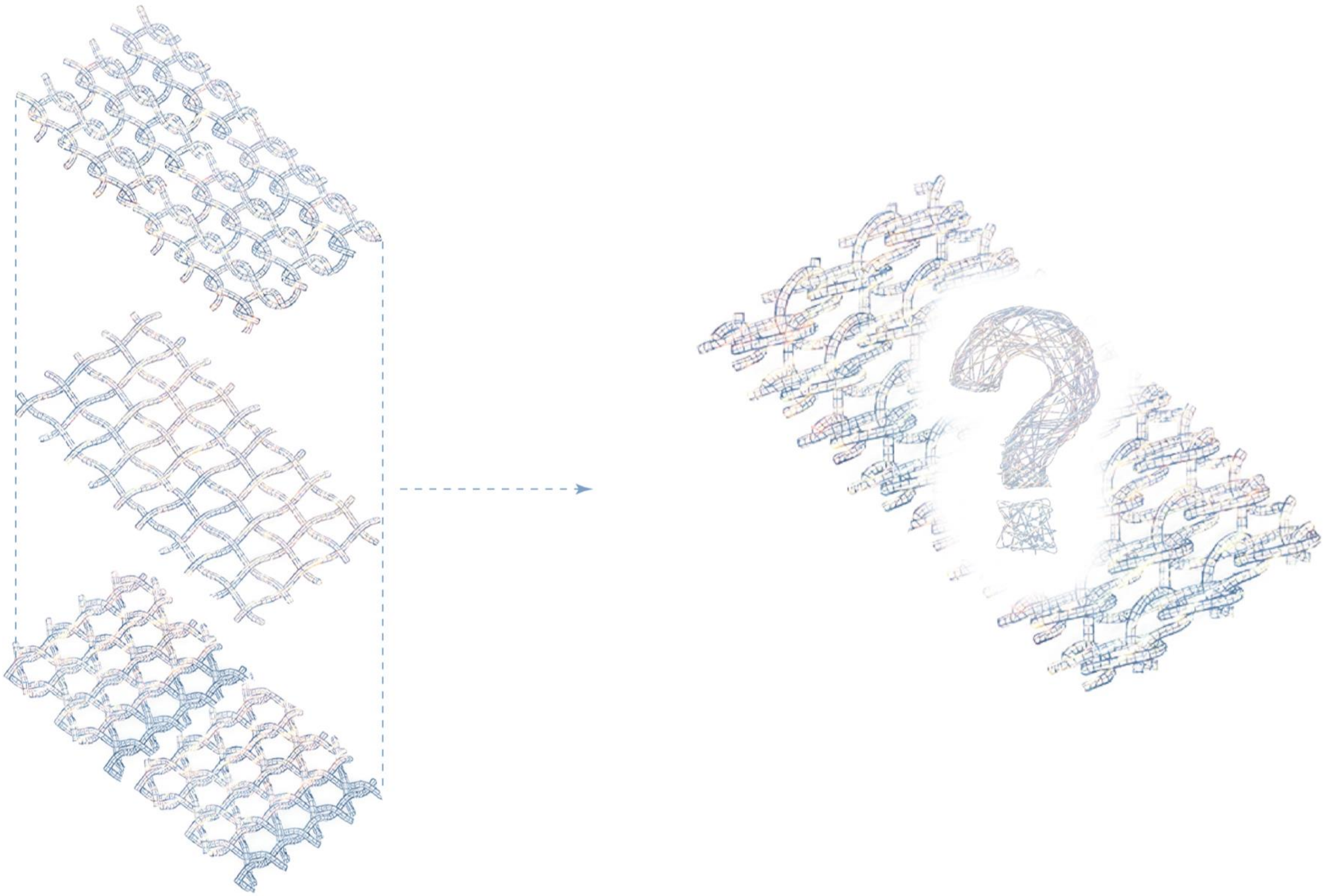


Fig.18

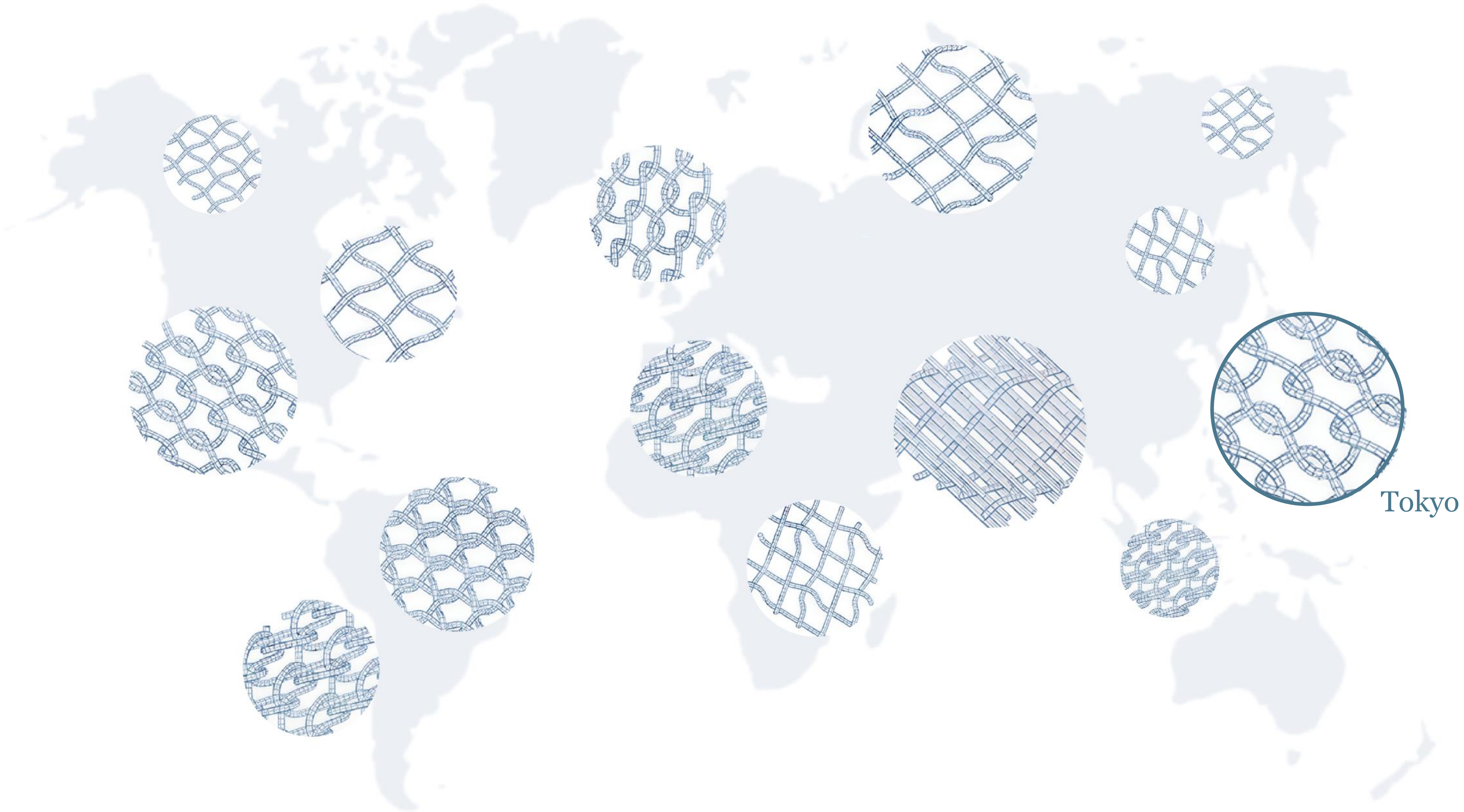
Today researchers leverage machine intelligence in order to augment design perceptions and thinking.



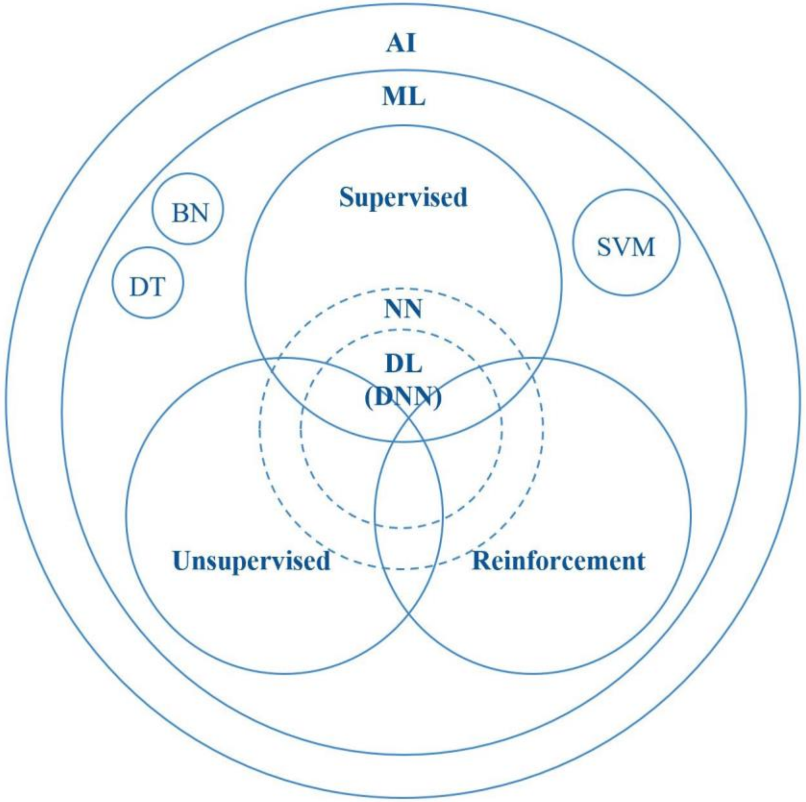
Language and text is a way of storing, retrieving, and generating meaning in architecture and is an essential input into machine learning as well.

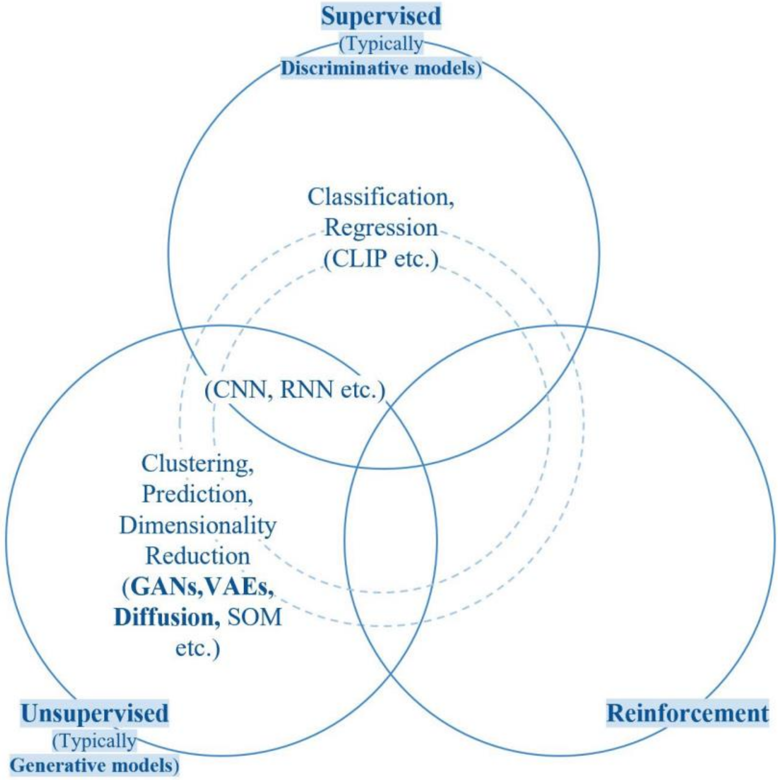


Using machine learning as tool and inspiration, my project seeks to track and archive AI's ability to read hidden features in order to synthesize the past and the present fabric of architecture across various places and times, and in doing so, produce intriguing and evocative forms.



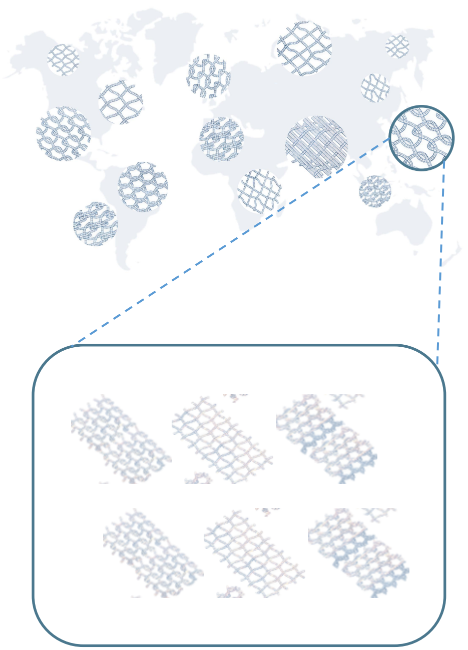
My project features three case studies in Tokyo which aim to reimagine the modern, westernized home...

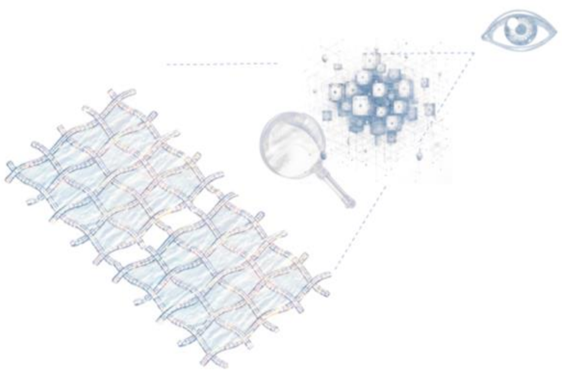


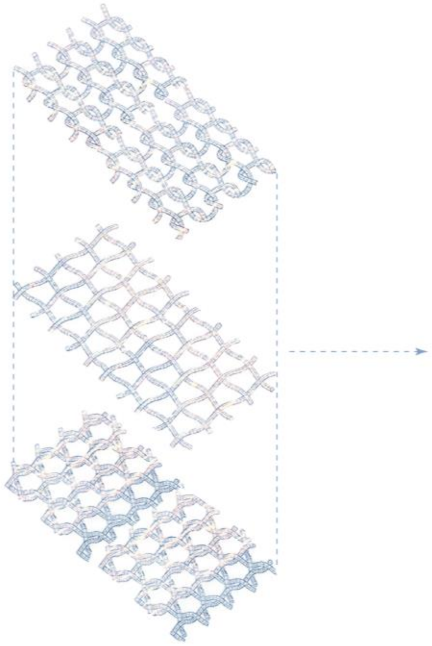


.... through the introduction of machine leaning models, classification, and image generation.

Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt







"Collect Fabrics"

Data
Collection, Curation, Annotation

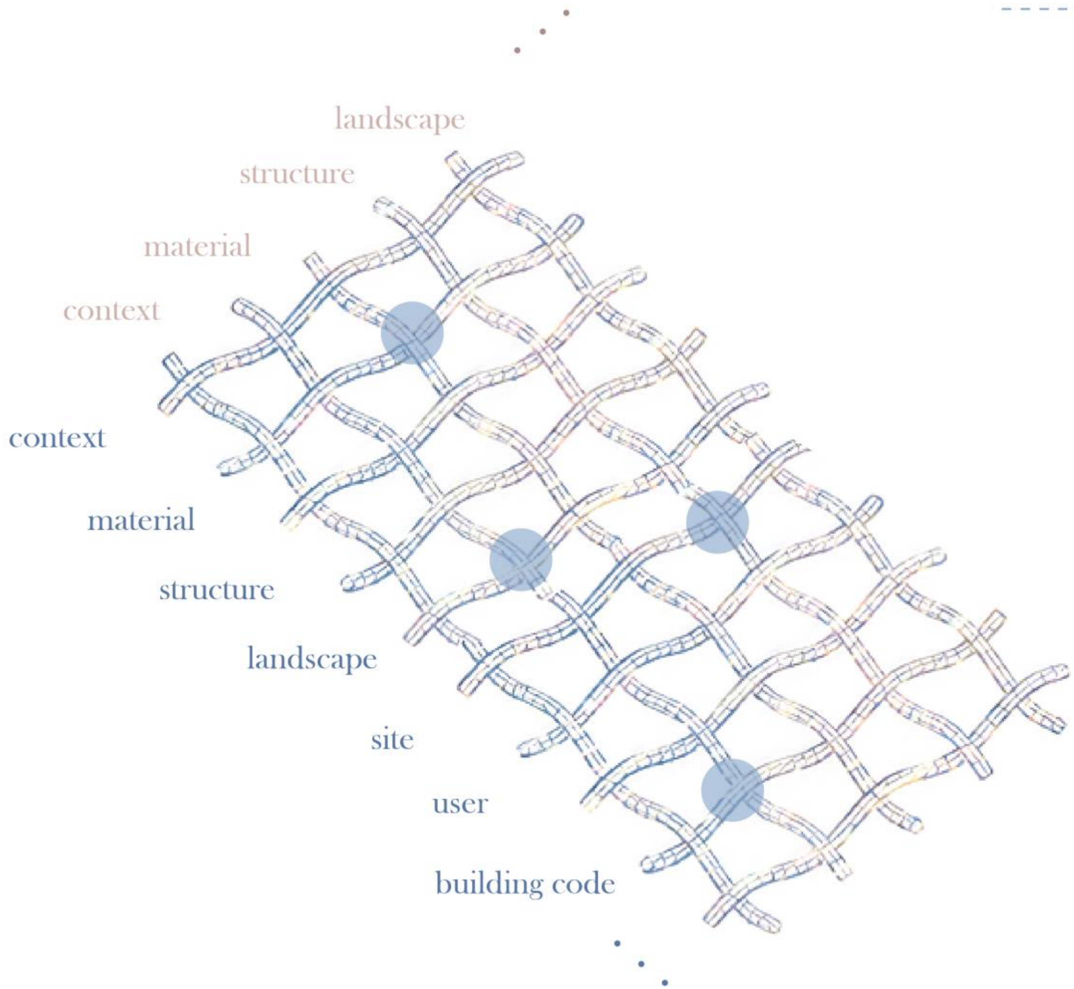
"Read between Threads"

Image Classification
(Discriminative AI)

"Synthesize Fabrics"

Image Generation
(Generative AI)

My design process began with collecting data fabrics and defining synthetic fabric qualities.
I followed this up with the classification and generation phases – reading between threads and synthesizing the fabrics.



manifest in a form of...

text
image
drawing
video
digital model
physical model

My initial data collection focused on plans and facades in order to highlight spatial features that face the local context.

Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt

Japanese Houses and Metadata


Fig.19


Fig.20

Japan's postwar economic growth is emblematic in several of that period's architypes.
Housing, offices, and commercial spaces were deeply influenced by mass production and standardization.

Fig.21

Spaces began emphasizing contemporary values of living such as speed and privacy. Interior spaces became reclusive and more enclosed, with smaller windows.
This rapid transition can be owed to the predominance of wooden construction and market preferences for new construction over second-hand properties, as the latter's value drops significantly as soon as it goes on sale.

Era will be Heisei, named for universal peace

CHIEF CABINET SECRETARY Keizo Obuchi announces the name of the new Imperial era Saturday.

"Heisei," which can be translated as "achievement of universal peace," will be the new Imperial era name, succeeding "Showa," or "enlightened peace," the government announced Saturday. The new era begins today.

Explaining the new era name, Chief Cabinet Secretary Keizo Obuchi told reporters that it was based on the hope that peace will be achieved both in Japan and around the world.

The word was taken from two Chinese historical classics, the "Shu Jing" ("Book of History") and the "Shi Ji" ("Historical Memoirs"). The name was chosen under a government guideline specifying that it must be simple and composed of two Chinese characters.

The new designation was made following discussions Saturday by an eight-member special forum comprised of representatives from the mass media and academia.

It was approved at an emergency Cabinet meeting that afternoon.

The forum members were Yoshizo Ikeda, president of NHK (Japan Broadcasting Corp.); Yosoji Kobayashi, president of the Japan Newspaper Publishers and Editors Association; Osamu Nakagawa, president of the National Association of Commercial Broadcasters in Japan; Haruo Nishihara, chairman of the Federation of Associations of Private Universities; Yoko Nuita, a scholar; Wataru Mori, chairman of the Association of National Universities; and Ryogo Kubo and Hajime Nakamura, recipients of the Order of Culture.

The government declined to disclose who had actually selected the new era name, although a special task force headed by Deputy Chief Cabinet Secretary Nobuo Ishihara oversaw the designation process.

Government sources said the task force had been asked in 1979 to begin the process of name selection and to keep the matter confidential.

The era name system has been in use since around the 5th century, with some modifications. The final year of the late Emperor's reign was referred to as the 64th year of Showa.

Both the Western calendar and the era name, or gengo, calendar are widely used in Japan, but the era name is required on all government documents, under the controversial 1979 Gengo Law.

Fig.22

Japan's year calendar system aligns with its imperial lineage, beginning a new era with each emperor's ascent.
Each era thus represents one human's life span, and consequently corresponds to the architectural style of that time.

Japan

National Isolation
1641 - 1853

Camera introduced to Japan
1841

Baby Boom
1947

"Small House" Origin of prefabricated housing
1959

First McDonald's (Beginning of Westernization)
1971

大同
Daido
806

(Oldest Existing House)

江戸
Edo
1603 - 1867

大正・明治
Taisho, Meiji
1868 - 1926

昭和
Showa
1926 - 1989

平成
Heisei
1989 - 2019

令和
Reiwa
2019 ~

Elevations and plans of Japanese houses can be traced back to the Edo period which started 420 years ago, covering 5 periods until now.


Fig.23

江戸
Edo
1603 - 1867

A thick straw roof and shoji doors are typical in the Edo period.


Fig.24

平成
Heisei
1989 - 2019

Whereas white and gray exterior colors with small windows begin to proliferate in the modern period.

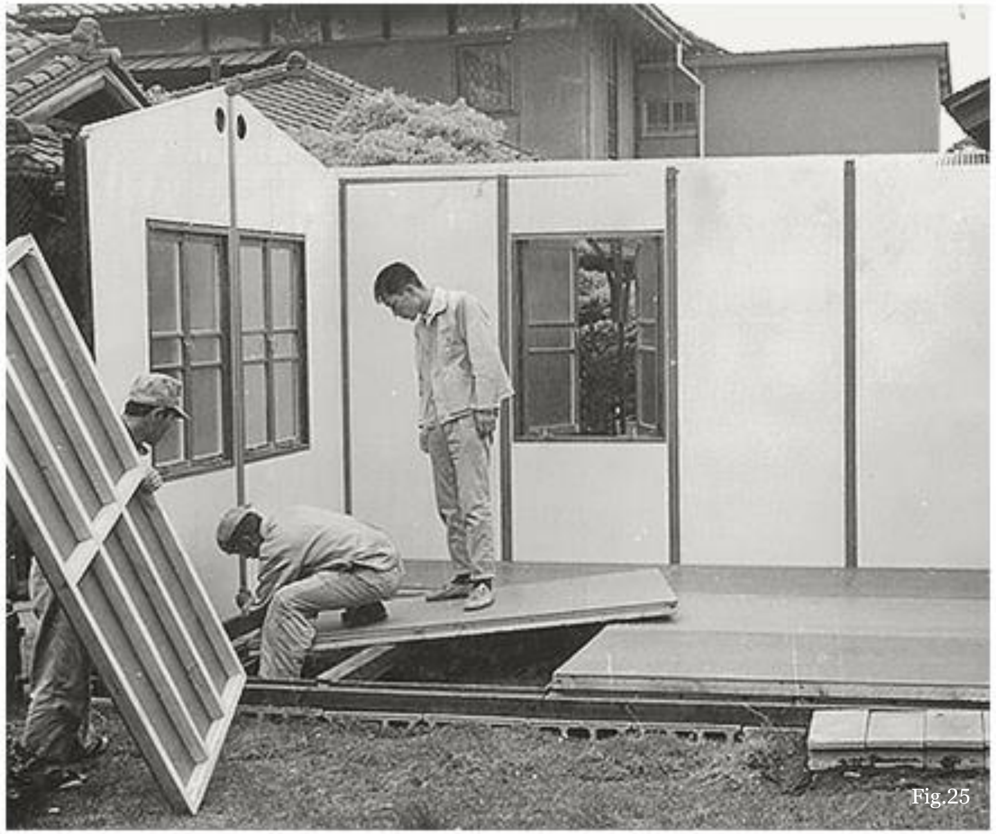


Fig.25

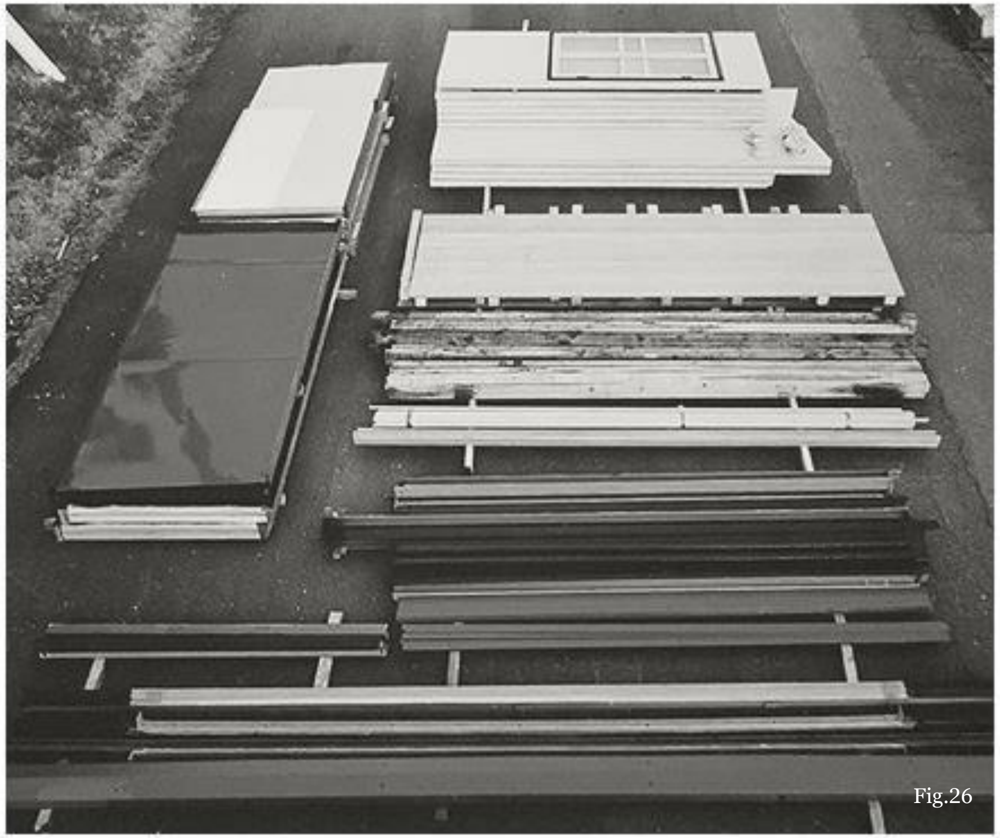


Fig.26

A key indicator of the shift from traditional to modern living can be seen with the introduction of the "small house," the first prefabricated building in Japan which was marketed as a solution to the shortage of schools and housing due to the baby boom era of the fifties and sixties. This pre-fabricated system would go on to significantly influence Japan's architectural landscape thereafter.

World

- **Camera invented** 1826
- **Crystal Palace (First Steel and Glass)** 1856
- **Oriel chambers by Peter Ellis (First Curtain Wall)** 1864
- **Unity Temple by Frank Lloyd Wright (First RC)** 1908
- **"International Style"** 1920s, 1930s
- **"Architecture without Architect" by Rudofsky** 1964
- **"Critical Regionalism" by Frampton** 1981

JAPAN

- **National Isolation** 1641 - 1853
- **Camera introduced to Japan** 1841
- **Baby Boom** 1947
- **"Small House" Origin of prefabricated housing** 1959
- **First McDonald's (Beginning of Westernization)** 1971





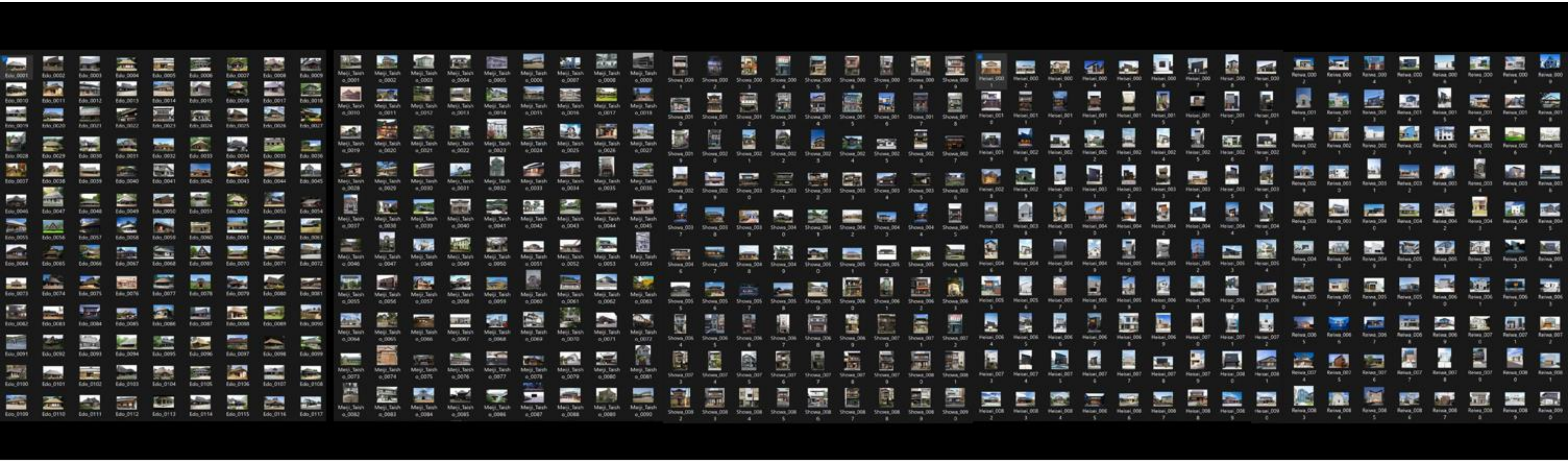

大同 Daido **806**	江戸 Edo **1603 - 1867**	大正・明治 Taisho, Meiji **1868 - 1926**	昭和 Showa **1926 - 1989**	平成 Heisei **1989 - 2019**	令和 Reiwa **2019 ~**

(Oldest Existing House)

Regional Architecture: Minka

Non-Regional Architecture: Jutaku

This project uses distinct historic eras as labels and prompts that teach AI how to distinguish the features from one period to the next.



大同 Daido **806** (Oldest Existing House)	江戸 Edo **1603 - 1867**	大正・明治 Taisho, Meiji **1868 - 1926**	昭和 Showa **1926 - 1989**	平成 Heisei **1989 - 2019**	令和 Reiwa **2019 ~**	
	117	245	177	153	171	= 863

Plans and façade images are likewise collected from the internet and input into the machine.

"Architecture"
Random region and type

647

"OtherRegions"
House in other regions

354

Some random architecture images were also added to the dataset to augment machine understanding of houses in the Japanese context in general.

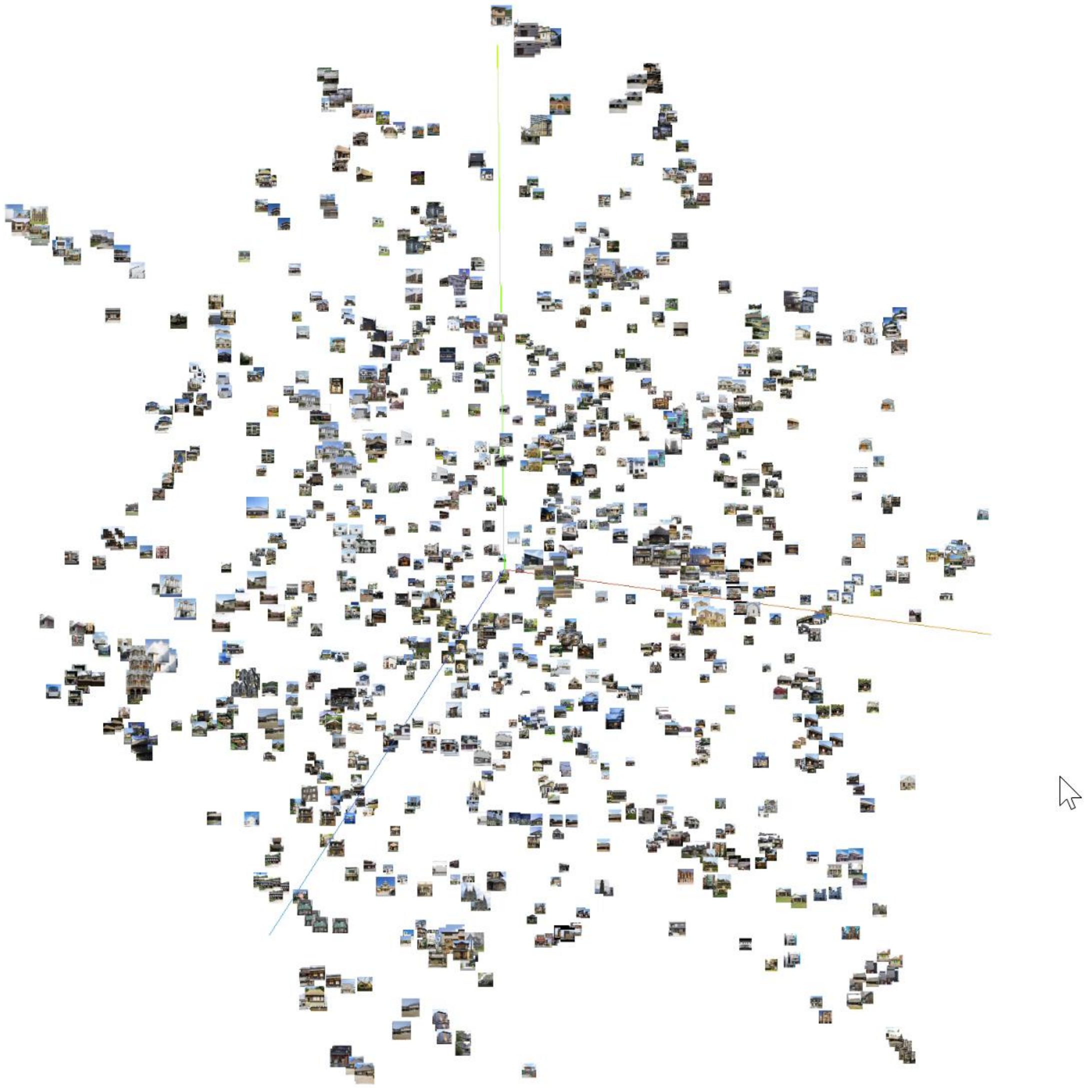
Rio Kobayashi (MArchI 2024’), Advisor: Andrew Witt

Photo

Monochrome

Outlines

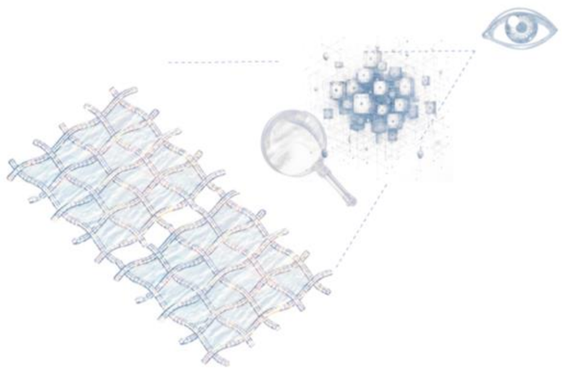
During the data curation phase, image styles were optimized to highlight different expressions of architectural elements.



*GIF

This dataset and period labels allowed machine learning to nuance the features of Japanese architecture.

Model Finding Journey



"Read between Threads"

Image Classification
(Discriminative AI)

Classifier based image generation model →

Based on classification model provided in SCI 6487: Machine Aesthetics: The Binary and the Spectrum
By Panagiotis Michalatos
at Harvard GSD Spring 2023

Classification assessment model

Based on classification model provided in SCI 6485: Introduction to Generative Artificial Intelligence
By Sabrina Osmany
at Harvard GSD Fall 2023

Here it begins to read what is between the threads – revealing hidden features

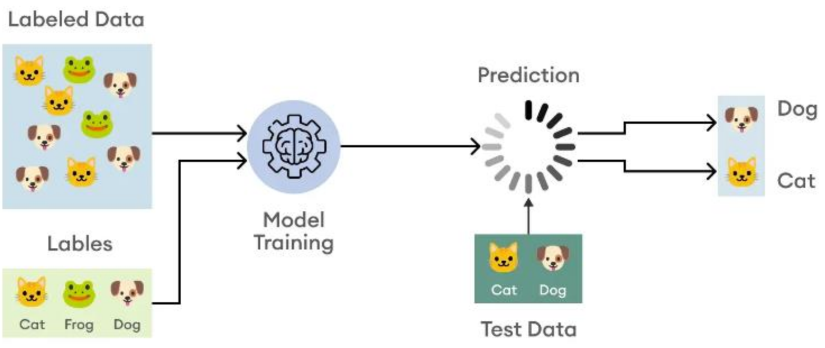


Fig.27 Concept Image of Classification Model

```
import torch.nn as nn
import torch.nn.functional as F

# This line defines a new Python class named Net, which is a neural network model
# It inherits from the nn.Module class, which is a base class for all PyTorch neural network modules
class CNNnet(nn.Module):

    # the constructor method for the Net class
    def __init__(self):

        # super().__init__() calls the constructor of the parent class (nn.Module) to ensure that necessary initialization is performed
        super().__init__()

        # the lines below define the layers in the NN. you could customize the name of the layer by changing the string after the 'self.'
        # This line creates the first convolutional layer (conv1) with 3 input channels (RGB images), 6 output channels and 5 is the convolutional filters size
        self.conv1 = nn.Conv2d(in_channels=3, out_channels=10, kernel_size=7)

        # This line specifies that max-pooling will be applied with a 2x2 window size and a stride of 2.
        self.pool = nn.MaxPool2d(kernel_size=2, stride=2)

        # same as the last convolutional layer
        self.conv2 = nn.Conv2d(in_channels=10, out_channels=16, kernel_size=5)

        # This line creates the first fully connected (linear) layer (fc1).
        # It specifies that it takes an input of size 16 * 5 * 5 (output from the previous convolutional layers) and outputs a tensor of size 120.
        # 9 = ((50-(7-1))/2-(5-1))/2. For more information, you can refer to https://pytorch.org/docs/stable/generated/torch.nn.Conv2d.html
        self.fc1 = nn.Linear(16 * 9 * 9, 120)

        # same as the last linear layer but with an input size of 120 and an output size of 84.
        self.fc2 = nn.Linear(120, 84)

        # same as the last linear layer but with an input size of 84 and an output size of 10 (which is the total number of classes).
        self.fc3 = nn.Linear(84, 10)


    # The forward method is where the actual computation of the neural network occurs.
    def forward(self, x):
        # F is typically an alias for the PyTorch module torch.nn.
        # you could also add relu layer in the model architecture configuration part
        x = F.relu(self.conv1(x))

        #  It performs max-pooling on the ReLU-activated feature map using the pool layer defined earlier.
        x = self.pool(x)

        # same as above
        x = F.relu(self.conv2(x))
        x = self.pool(x)

        # flatten all dimensions except batch. It converts the 2D feature maps from the convolutional layers into a 1D tensor suitable
        x = torch.flatten(x, 1)

        # We apply the ReLU activation function to the result of the first fully connected layer
        x = F.relu(self.fc1(x))

        # We apply the ReLU activation function to the result of the second fully connected layer
        x = F.relu(self.fc2(x))

        # We apply the third fully connected layer (fc3) to the output of the previous layer
        x = self.fc3(x)

        # return the tensor x, which contains the model's output after passing through all the layers
        return x

        model = torch.hub.load('pytorch/vision:v0.10.0', 'alexnet', pretrained=True)
        model.eval()
        print(model)
```

Classification assessment model

Based on classification model
SCI 6485: Introduction to Generative
Artificial Intelligence
By Sabrina Osmany

To identify the features the "fabrics" and "threads" contains in the scraped dataset, a supervised machine learning called classification is used here.

Total epoch = 20
32 data points
= images for one batch
3 channels = RGB
50 x 50 pixel

Total epoch = 20
32 data points
= images for one batch
3 channels = RGB
256 x 256 pixel

Total epoch = 50
32 data points
= images for one batch
3 channels = RGB
256 x 256 pixel

Classification assessment model

Based on classification model
SCI 6485: Introduction to Generative
Artificial Intelligence
By Sabrina Osmany

The model tries to predict the correct label of a given input data, in this case periodic information.
By optimizing some parameters in the code that controls the machine's learning capabilities and the influence weight of the dataset, the classification model provides more accurate predictions.

Edo: 1.94%
Meiji_Taisho: 4.21%
Showa: 86.69%
Heisei: 0.58%
Reiwa: 0.17%
Others: 6.36%

Edo: 2.13%
Meiji_Taisho: 33.35%
Showa: 0.08%
Heisei: 48.93%
Reiwa: 8.74%
Others: 6.76%

Edo: 0.40%
Meiji_Taisho: 14.20%
Showa: 13.19%
Heisei: 3.09%
Reiwa: 1.20%
Others: 67.90%

Edo: 5.48%
Meiji_Taisho: 81.13% Showa: 2.37% Heisei: 1.44% Reiwa: 1.84% Others: 7.64%

Edo: 3.65%
Showa: 0.19%
Meiji_Taisho: 79.27% Heisei: 6.22%
Reiwa: 2.42%
Others: 8.24%

Edo: 0.01%
Meiji_Taisho: 0.16%
Showa: 0.02%
Heisei: 81.36%
Reiwa: 18.39%
Others: 0.05%

Edo: 49.56%
Meiji_Taisho: 27.88%
Showa: 4.08%
Heisei: 11.35%
Reiwa: 0.22%
Others: 6.86%

Edo: 0.09%
Meiji_Taisho: 5.32%
Showa: 88.19%
Heisei: 0.93%
Reiwa: 0.50%
Others: 4.96%

Edo: 3.56%
Meiji_Taisho: 17.37%
Showa: 1.56%
Heisei: 1.35%
Reiwa: 0.89%
Others: 75.16%

Edo: 1.22%
Meiji_Taisho: 5.04%
Showa: 55.90%
Heisei: 0.28%
Reiwa: 0.06%
Others: 37.48%

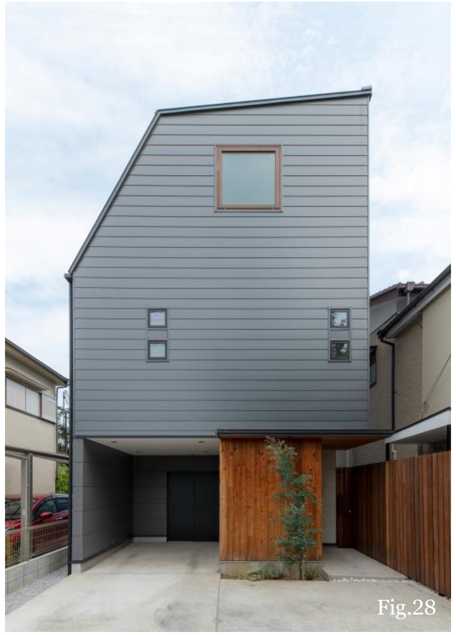


Fig.28

Edo: 4.05%
Meiji_Taisho: 9.42%
Showa: 0.06%
Heisei: 54.51%
Reiwa: 13.28%
Others: 18.63%

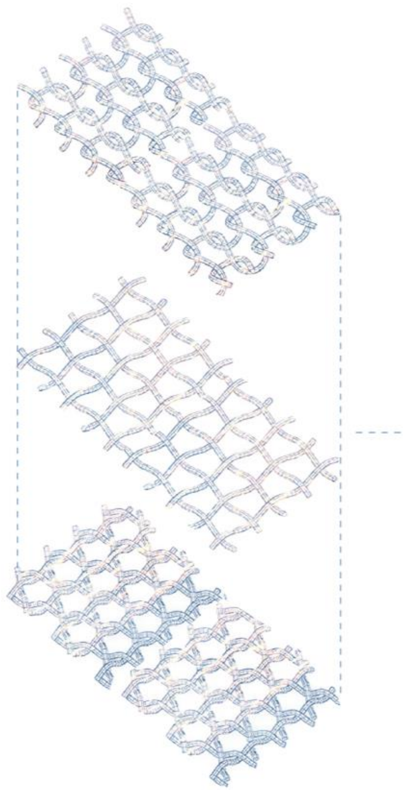

Fig.29

Edo: 18.47%
Meiji_Taisho: 3.40%
Showa: 0.10%
Heisei: 52.19%
Reiwa: 23.90%
Others: 1.88%


Fig.30

Edo: 0.01%
Meiji_Taisho: 0.89%
Showa: 0.05%
Heisei: 13.69%
Reiwa: 77.27%
Others: 8.08%

The percentages in the images provide a more detailed quantitative prediction, indicating the likelihood of the houses belonging to each historical period.



"Synthesize Fabrics"

Image Generation
(Generative AI)

Image generative AI tools
Available online →

- Runway ML
- DALL-E 2,3
- Mid journey
- ChatGPT4

Classifier based image generation model →

Based on classification model provided in SCI 6487: Machine Aesthetics: The Binary and the Spectrum
By Panagiotis Michalatos
at Harvard GSD Spring 2023

CycleGAN model →

Built based on:
"Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks"
Jun-Yan Zhu
Taesung Park
Phillip Isola
Alexei A. Efros
UC Berkeley
In ICCV 2017

Stable Diffusion (SD) model (text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

My project took experimental steps for generation process to find the most effective methods
including popular generative AI tools, classification-based modeling, CycleGAN, and stable diffusion models.









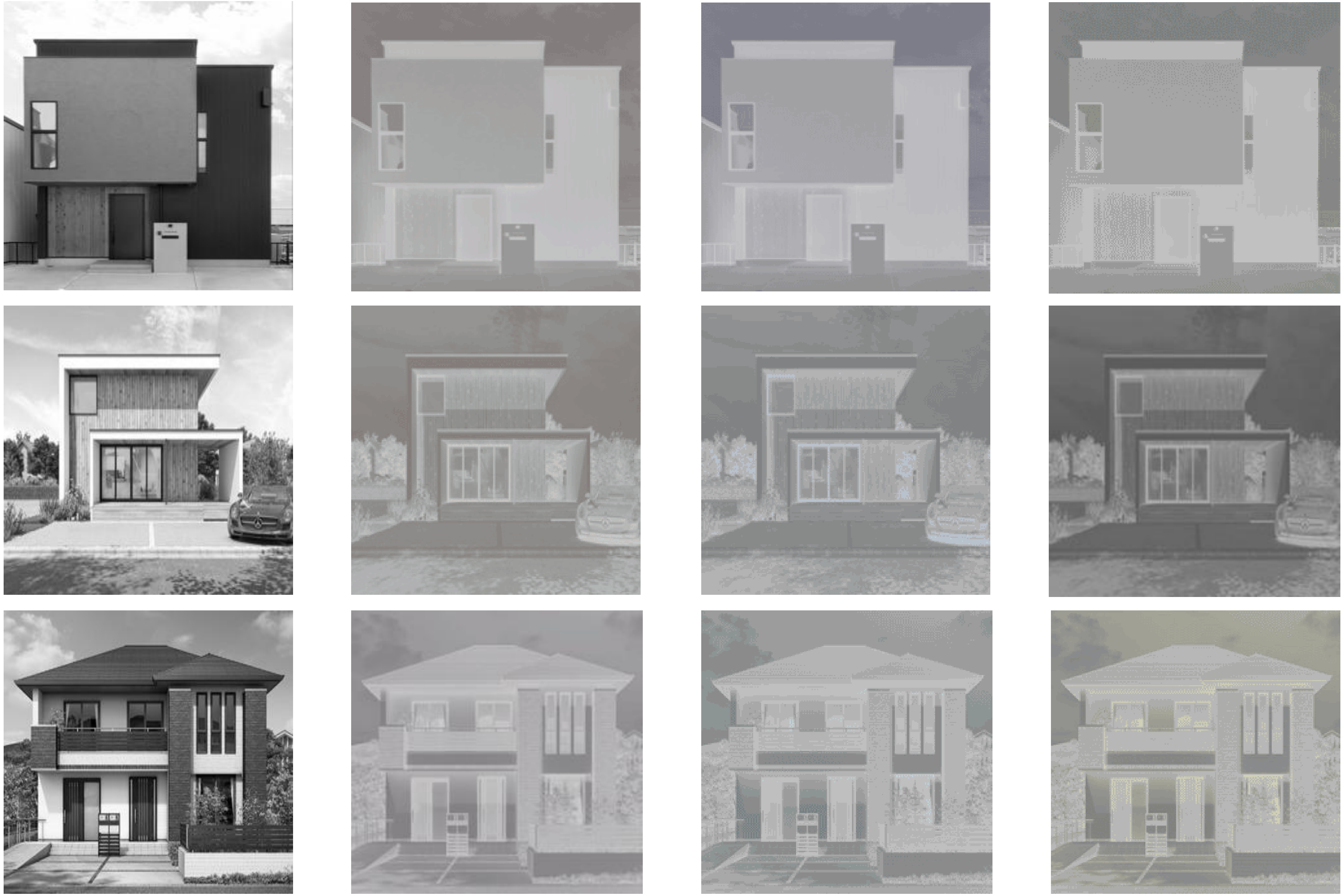








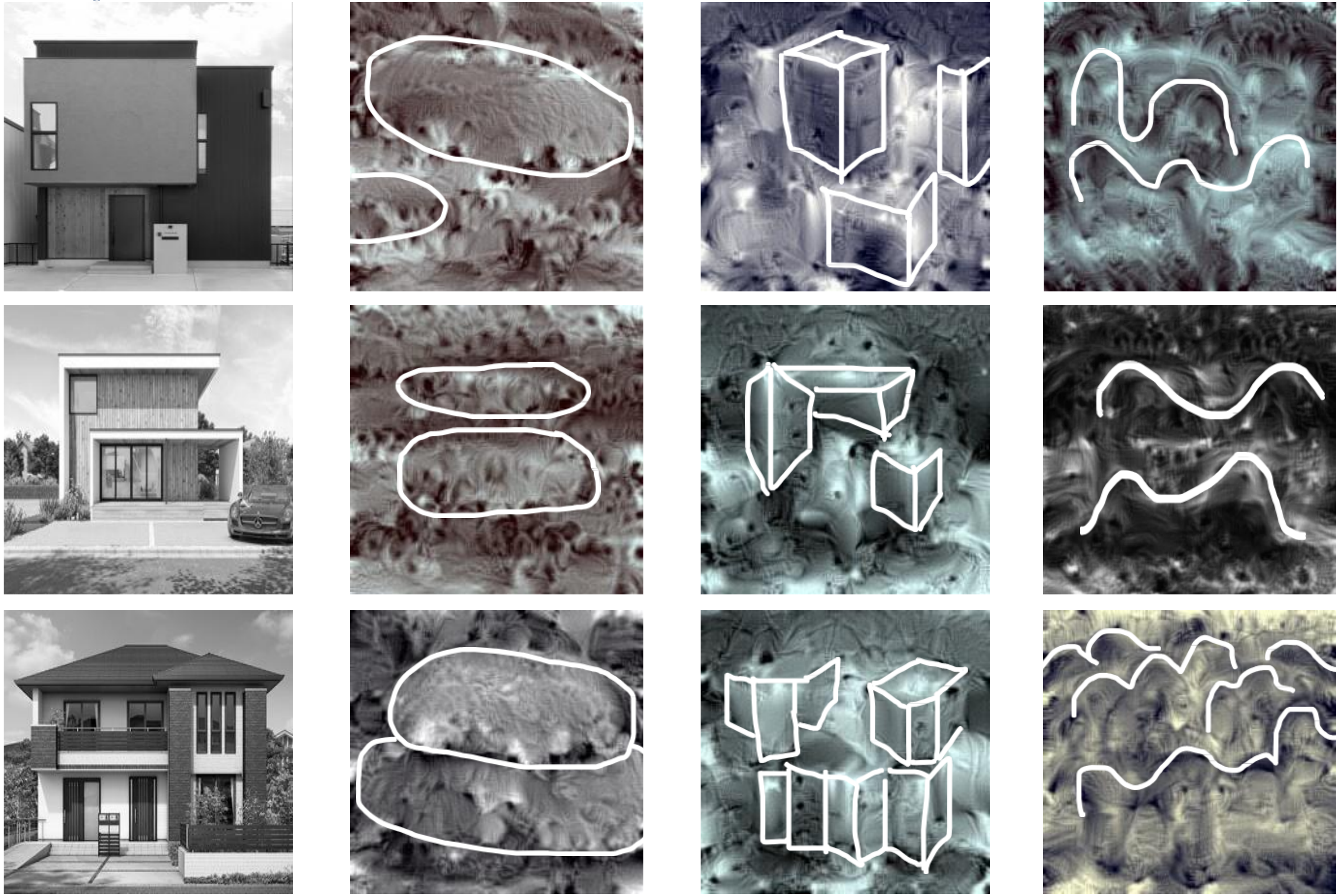

In the experimental iterations that used AI tools like DALL-E2 and Midjourney,
the only way to synthesize data is mainly through texts in natural language, which forces one to give up most of design control.



*GIF

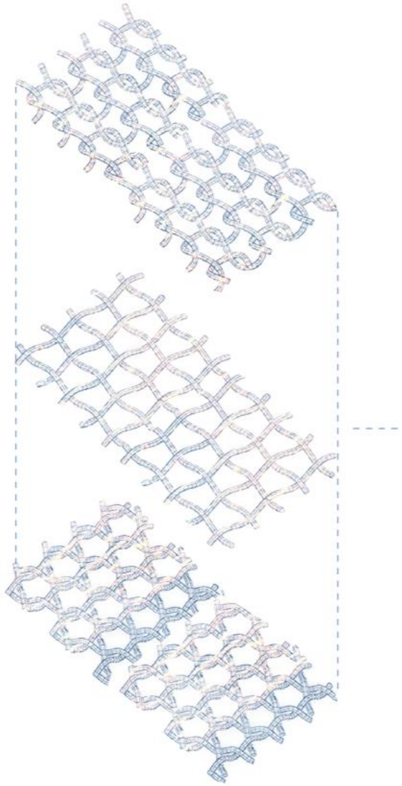
Edo Reiwa Other

The next model I explored is classification based generative model, which transforms features from one label period to another.



Edo Reiwa Other

The results highlighted the predominant features of each period, thereby validating the quality of this dataset and machine's understanding of architectural features.



Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

This project landed on stable diffusion modeling, which has been the most popular image generative model over the last few years.

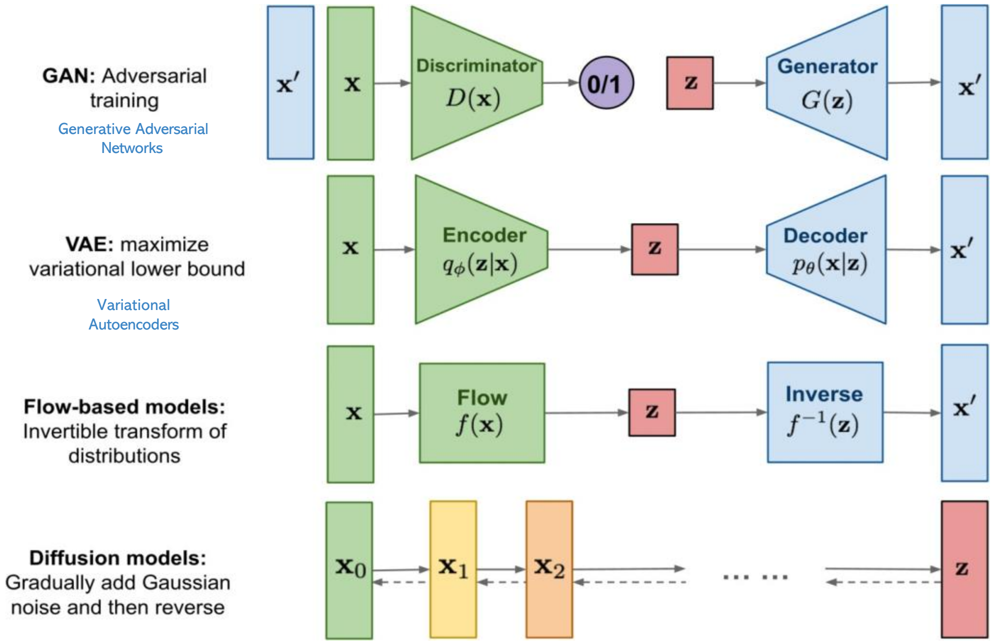
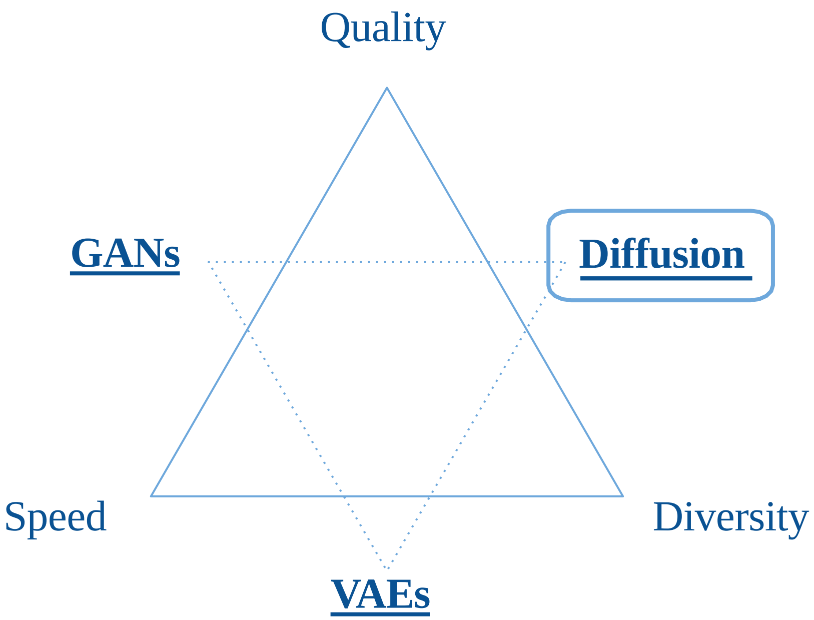


Fig.31 Overview of different types of generative models



This model is known for its high-quality and diverse output compared to GAN and autoencoder.

```
[ ] import os
    from diffusers import StableDiffusionPipeline
    from PIL import Image
```

```
[ ] repo_id = "CompVis/stable-diffusion-v1-4" # Use V100 GPU in Notebook Setting
    #repo_id = "stabilityai/stable-diffusion-2-1" # Use A100 GPU in Notebook Setting
    pipe = StableDiffusionPipeline.from_pretrained(pretrained_model_name_or_path=repo_id)
    _ = pipe.to("cuda")
```

```
[ ] def image_grid(imgs, rows, cols):
        assert len(imgs) == rows*cols
        w, h = imgs[0].size
        grid = Image.new('RGB', size=(cols*w, rows*h))
        grid_w, grid_h = grid.size
        for i, img in enumerate(imgs):
            grid.paste(img, box=(i%cols*w, i//cols*h))
        return grid
```

```
[ ] num_images = 3
    prompt = ["contemporary japanese house"] * num_images
    images = pipe(prompt).images
    grid = image_grid(images, rows=1, cols=3)
```

```
[ ] grid
```

e.g. prompt = ["contemporary Japanese house"]

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

Starting from text to image generative model (.. which is basically an armature version of the AI tools mentioned earlier)

Stable Diffusion text-to-image fine-tuning

The `train_text_to_image.py` script shows how to fine-tune the stable diffusion model on your own dataset.

> The text-to-image fine-tuning script is experimental. It's easy to overfit and run into issues like catastrophic forgetting. We recommend to explore different hyperparameters to get the best results on your dataset.

Running locally

Installing the dependencies

Before running the scripts, make sure to install the library's training dependencies:

```
pip install git+https://github.com/huggingface/diffusers.git
pip install -U -r requirements.txt
```

And initialize an 🤗 Accelerate environment with:

```
accelerate config
```

Fig.32

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

And to make machine understand more specific subject and style, fine-tuned the stable diffusion model, first with Pokemon image dataset as a test

```
[ ] import os
    from diffusers import StableDiffusionPipeline
    from PIL import Image
    import torch
    from safetensors.torch import load_file

[ ] repo_id = "CompVis/stable-diffusion-v1-4" # Use V100 GPU in Notebook Setting
    #repo_id = "stabilityai/stable-diffusion-2-1" # Use A100 GPU in Notebook Setting
    pipe = StableDiffusionPipeline.from_pretrained(pretrained_model_name_or_path=repo_id)
    _ = pipe.to("cuda")

[ ] unet_weights = load_file("sd-pokemon-model/checkpoint-15000/unet/diffusion_pytorch_model.safetensors")
    pipe.unet.load_state_dict(unet_weights)

[ ] def image_grid(imgs, rows, cols):
        assert len(imgs) == rows*cols

        w, h = imgs[0].size
        grid = Image.new('RGB', size=(cols*w, rows*h))
        grid_w, grid_h = grid.size

        for i, img in enumerate(imgs):
            grid.paste(img, box=(i%cols*w, i//cols*h))
        return grid

[ ] num_images = 3
    prompt = ["a photo of two cats"] * num_images
    images = pipe(prompt).images
    grid = image_grid(images, rows=1, cols=3)

[ ] grid
```


Fig.33


Fig.33

Successfully created new Pokemon through text as an experiment

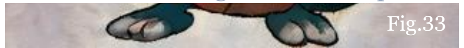


Fig.33

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

Stable Diffusion text-to-image fine-tuning

```
import os
from diffusers import StableDiffusionPipeline
from PIL import Image
import torch
from safetensors.torch import load_file
import warnings
import random
warnings.filterwarnings("ignore")
```

```
[ ] repo_id = "CompVis/stable-diffusion-v1-4" # Use V100 GPU in Notebook Setting
    #repo_id = "stabilityai/stable-diffusion-2-1" # Use A100 GPU in Notebook Setting
    pipe = StableDiffusionPipeline.from_pretrained(pretrained_model_name_or_path=repo_id)
    _ = pipe.to("cuda")
```

```
[ ] unet_weights = load_file(file_path)
    pipe.unet.load_state_dict(unet_weights)
```

```
[ ] def image_grid(imgs, rows, cols):
        assert len(imgs) == rows*cols

        w, h = imgs[0].size
        grid = Image.new('RGB', size=(cols*w, rows*h))
        grid_w, grid_h = grid.size

        for i, img in enumerate(imgs):
            grid.paste(img, box=(i%cols*w, i//cols*h))
        return grid
```

```
[ ] eras = ['Edo', 'Heisei', 'Meiji_Taisho','Reiwa','Showa']
```

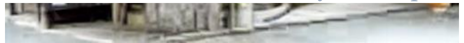
```
[ ] num_images = 3
    era = "Reiwa"
    prompt = [f"a photo of a house in the {era} period"] * num_images
    images = pipe(prompt).images
    grid = image_grid(images, rows=1, cols=3)
```

```
[ ] grid
```

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

Referencing the methods, built a text to image stable diffusion model extra-trained with my own Japanese housing dataset with periodic labels



```
import os
from diffusers import StableDiffusionPipeline
from PIL import Image
import torch
from safetensors.torch import load_file
import warnings
import random
from PIL import Image
from IPython.display import Image as IImage

warnings.filterwarnings("ignore")
```

```
repo_id = "CompVis/stable-diffusion-v1-4" # Use V100 GPU in Notebook Setting
#repo_id = "stabilityai/stable-diffusion-2-1" # Use A100 GPU in Notebook Setting
pipe = StableDiffusionPipeline.from_pretrained(pretrained_model_name_or_path=repo_id)
_ = pipe.to("cuda")
```

```
unet_weights = load_file(file_path)
pipe.unet.load_state_dict(unet_weights)
```

```
def image_grid(imgs, rows, cols):
    assert len(imgs) == rows*cols

    w, h = imgs[0].size
    grid = Image.new('RGB', size=(cols*w, rows*h))
    grid_w, grid_h = grid.size

    for i, img in enumerate(imgs):
        grid.paste(img, box=(i%cols*w, i//cols*h))
    return grid
```

```
eras = ['Edo', 'Heisei', 'Meiji_Taisho','Reiwa','Showa']
```

```
num_images = 3
era_1 = "Reiwa"
era_2 = "Showa"
prompt_1 = [f"a photo of a house in the {era_1} period"] * num_images
prompt_2 = [f"a photo of a house in the {era_2} period"] * num_images

images = pipe(prompt).images
```

```
grid
```

e.g.
era_1 = “Edo"
era_2 = “Heisei"
prompt_1 = [f"a photo of a house in the {era_1} period"]
prompt_2 = [f"a photo of a house in the {era_2} period"]

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
**Transformation SD model
between prompts - GIF**
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

Simultaneously created a text-to-gif model

*GIF

To see a walk the model is taking, transformation process that reflects machine's understanding and suggestions.
Machine walks around between the information threads and intersections pondering where to go and taking a detour – called latent walk.

Image-to-image

The Stable Diffusion model can also be applied to image-to-image generation by passing a text prompt and an initial image to condition the generation of new images.

The StableDiffusionImg2ImgPipeline uses the diffusion-denoising mechanism proposed in SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations by Chenlin Meng, Yutong He, Yang Song, Jiaming Song, Jiajun Wu, Jun-Yan Zhu, Stefano Ermon.

The abstract from the paper is:

Guided image synthesis enables everyday users to create and edit photo-realistic images with minimum effort. The key challenge is balancing faithfulness to the user input (e.g., hand-drawn colored strokes) and realism of the synthesized image. Existing GAN-based methods attempt to achieve such balance using either conditional GANs or GAN inversions, which are challenging and often require additional training data or loss functions for individual applications. To address these issues, we introduce a new image synthesis and editing method, Stochastic Differential Editing (SDEdit), based on a diffusion model generative prior, which synthesizes realistic images by iteratively denoising through a stochastic differential equation (SDE). Given an input image with user guide of any type, SDEdit first adds noise to the input, then subsequently denoises the resulting image through the SDE prior to increase its realism. SDEdit does not require task-specific training or inversions and can naturally achieve the balance between realism and faithfulness. SDEdit significantly outperforms state-of-the-art GAN-based methods by up to 98.09% on realism and 91.72% on overall satisfaction scores, according to a human perception study, on multiple tasks, including stroke-based image synthesis and editing as well as image compositing.

> Make sure to check out the Stable Diffusion Tips section to learn how to explore the tradeoff between scheduler speed and quality, and how to reuse pipeline components efficiently!

StableDiffusionImg2ImgPipeline

Fig.34

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

I further finetuned the model to make it image-to-image so that it can be effectively applied to the case study.

```
#launch jupyter online
#jupyter lab

#Change Env! Somehow mamba does not work...
#conda activate text2image

!pip install git+https://github.com/huggingface/diffusers.git
#!pip install -U -r requirements.txt
!pip install transformers
```

```
from google.colab import drive
# Mount Google Drive
drive.mount('/content/drive')
```

```
# Access a file in Google Drive
file_path = '/content/drive/My Drive/text2image/diffusion_pytorch_model.safetensors'
file_path_image = '/content/drive/My Drive/text2image/contemporary_house/house_001.jpg'
#file_path = '/content/drive/My Drive/text2image/houses-model/checkpoint-15000/unet/diffusion_pytorch_model.safetensors'
```

```
import requests
from PIL import Image
from io import BytesIO

from diffusers import StableDiffusionPipeline
from diffusers import StableDiffusionImg2ImgPipeline

import os

import torch
from torchvision import transforms
from safetensors.torch import load_file

import random
from PIL import Image
from IPython.display import Image as IImage

import warnings
warnings.filterwarnings("ignore")
```

```
device = "cuda"
repo_id = "CompVis/stable-diffusion-v1-4" # Use V100 GPU in Notebook Setting
#repo_id = "stabilityai/stable-diffusion-2-1" # Use A100 GPU in Notebook Setting
pipe = StableDiffusionImg2ImgPipeline.from_pretrained(repo_id, torch_dtype=torch.float16)
pipe = pipe.to(device)
```

```
#Finetuned Unet model
#unet_weights = load_file(file_path)
#pipe.unet.load_state_dict(unet_weights)
```

```
#ls data/contemporary_house/
```

```
init_image = Image.open(file_path_image)
init_image = init_image.resize((768, 512))
```

```
init_image
```

```
prompt = "a photo of a house in the traditional japanese architecture Edo period"
generator = torch.Generator(device=device).manual_seed(1024)
image = pipe(prompt=prompt, image=init_image, strength=0.7, guidance_scale=7.5).images[0]
```

```
image
```

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
**Fine-tuned image2image SD model
with my dataset**

The stable diffusion model transforms images by referencing the dataset of Japanese houses and capturing the nature of the hidden qualities.

Case Study:
Transforming Japanese Moden Houses

Fig.35

Fig.36

I return to my case study of three houses which are assumed to be located at the intersection of new central areas and an old residential town in northern Tokyo.
Responding to Tokyo’s high population density and the limited space, residential areas near central Tokyo often consist of compact apartments and homes jostling for space.

Fig.28

Fig.29

Fig.30

Each house has two or three stories designed for a single family - a couple and children.
This minimalistic façade indicates that they were newly built in the Reiwa period, during the last several years.


Fig.37

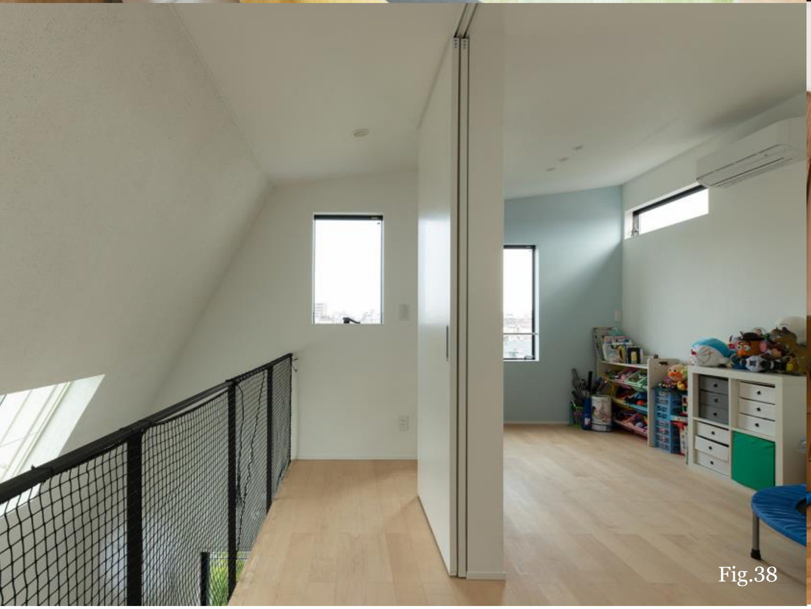


Fig.38


Fig.39


Fig.40


Fig.41


Fig.42

Their interiors use generic materials, adhering to modern housing design conventions.

Ground Floor Second Floor Third Floor

House 1

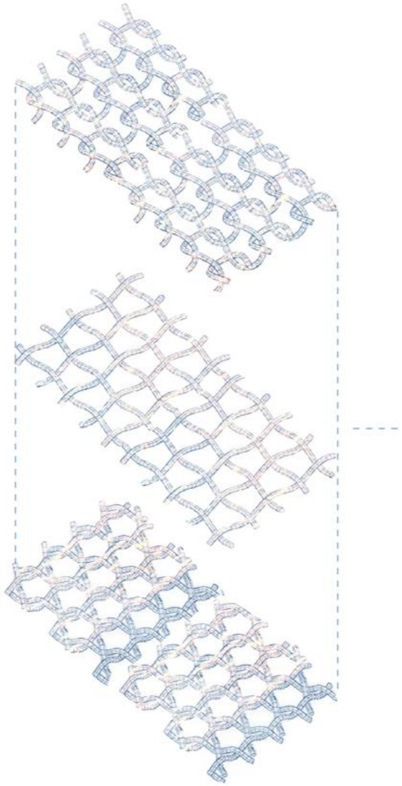
Ground Floor Second Floor

House 2

Ground Floor Second Floor

House 3

Their assumed plans and façades exhibit typical patterns seen in modern Japanese housing.



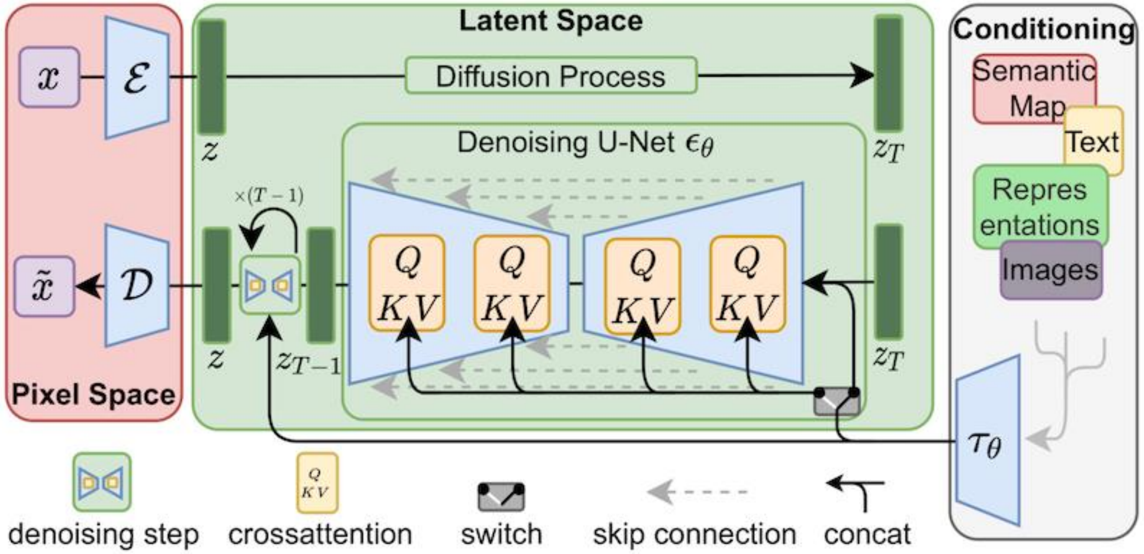


Fig.43

When taking architectural plans and façade images and feeding them into a fine-tuned stable diffusion model...

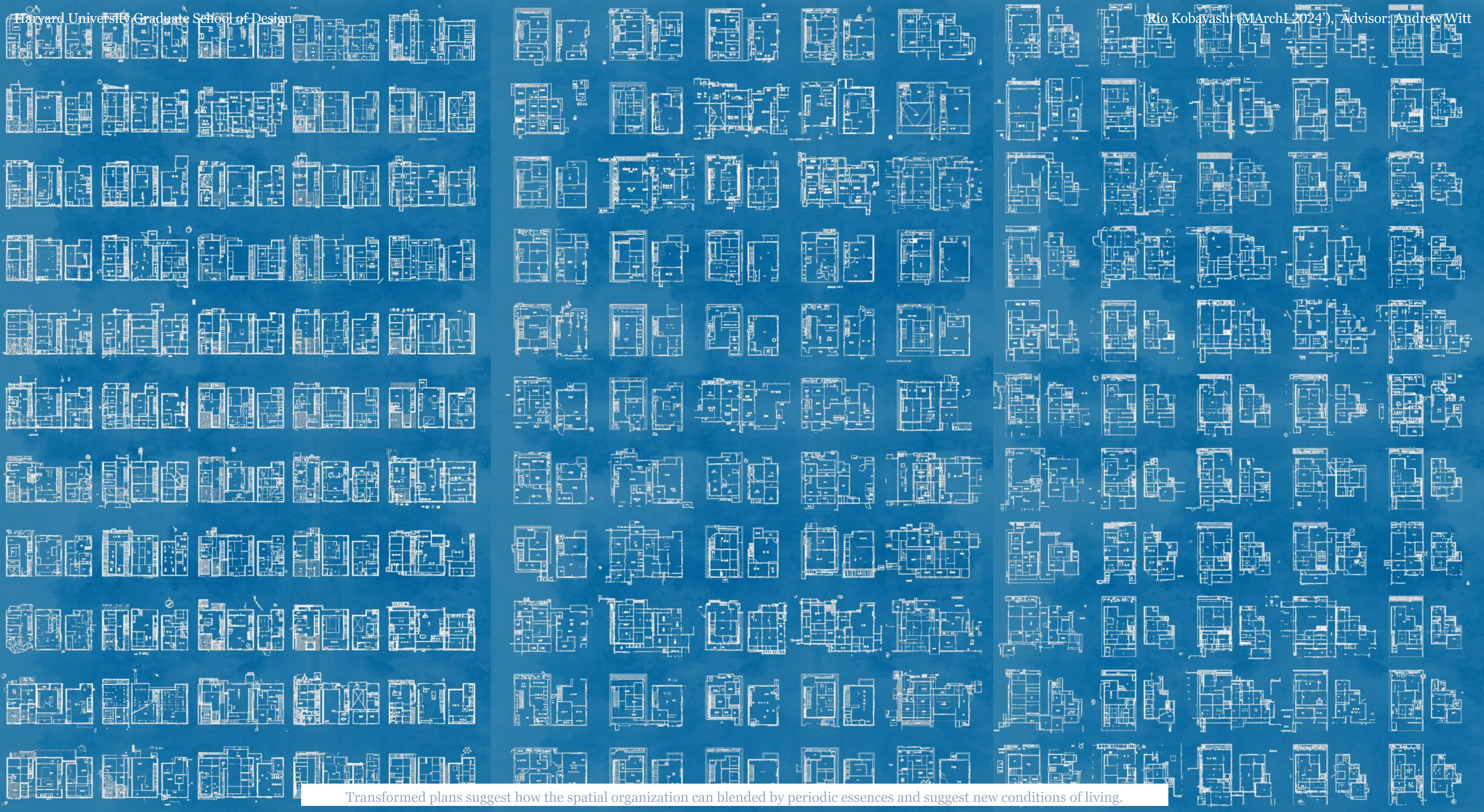
...the model manipulates period-specific design elements in the images
by blending, amplifying, or transforming them and then optimizing them for a coherent and relatively seamless outcome.

Reiwa-biased

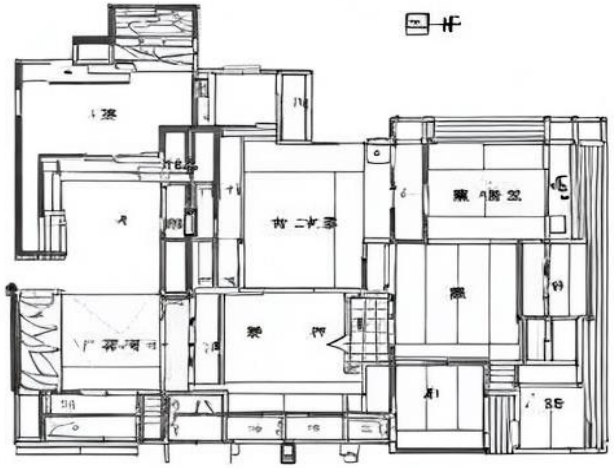
Taisho-biased

Edo-biased

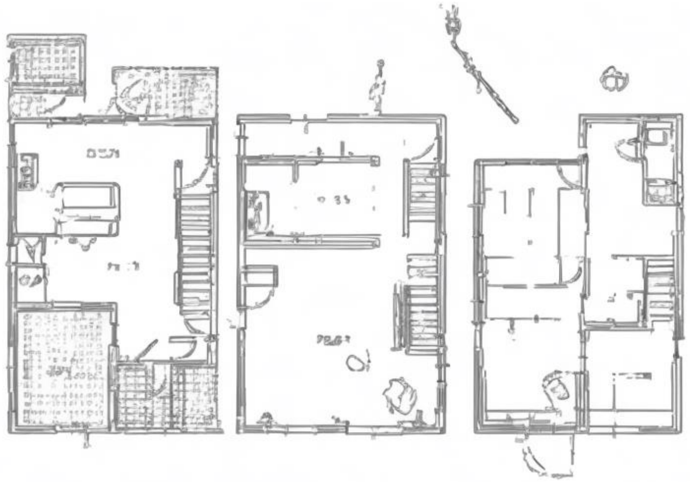
Each façade is generated by different period-based prompts which demonstrate
new hybridized possibilities of form, composition, and material use from eras like the Reiwa, Taisho, and Edo.



Transformed plans suggest how the spatial organization can blended by periodic essences and suggest new conditions of living.



Taisho and Edo biased



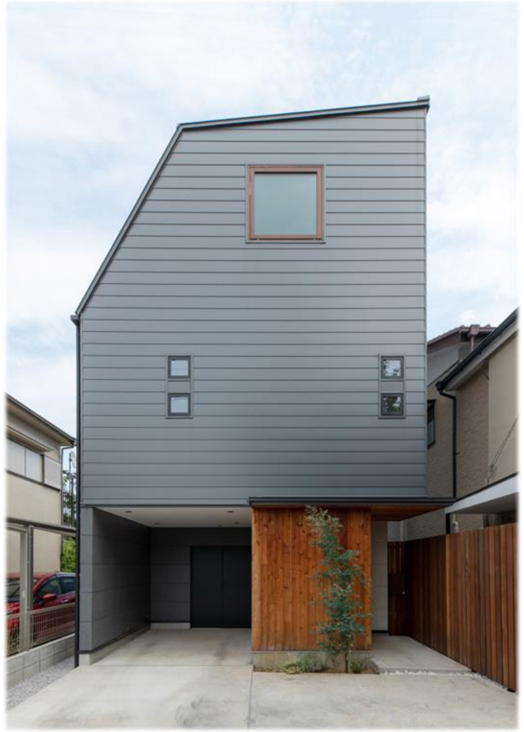
Showa and Reiwa based

In greater detail, some conventional plans absorbing hints of period based spatial logics.

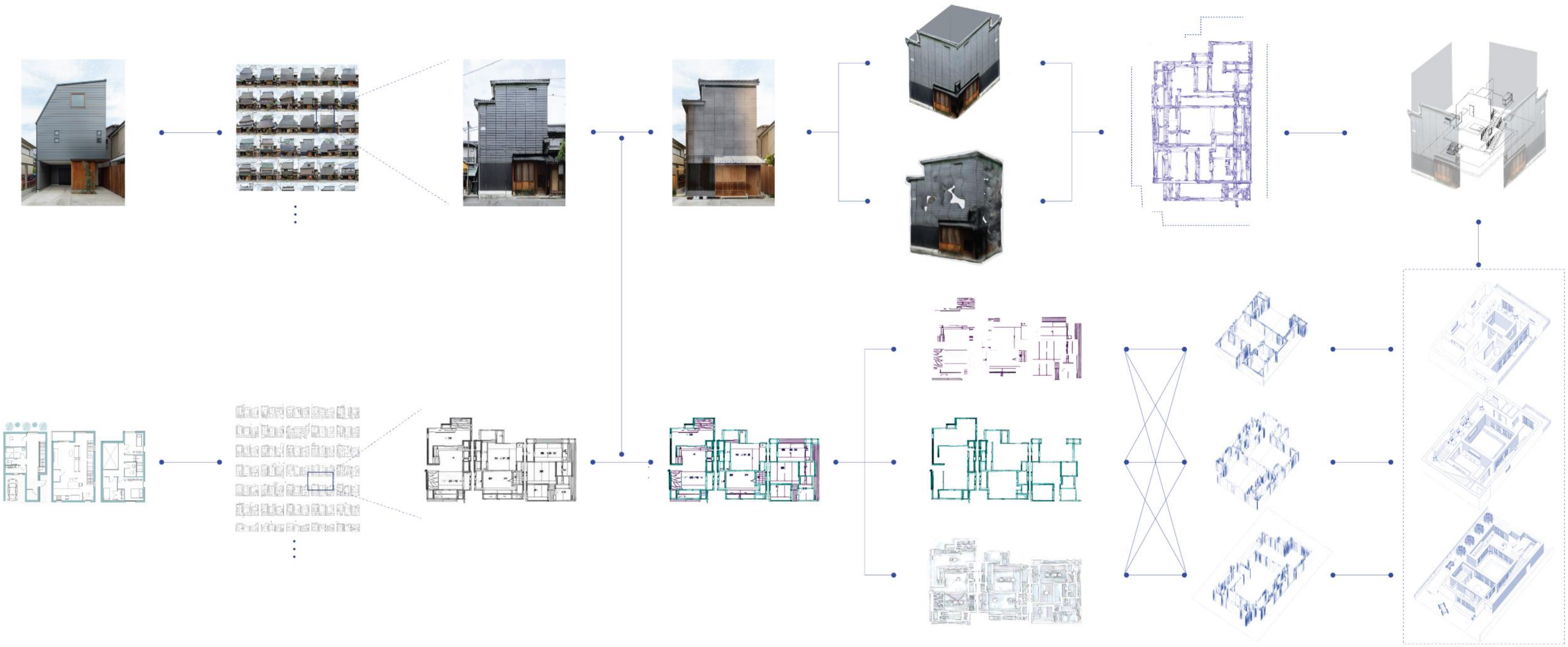
House 1
Synthesis of Edo, Meiji/Taisho, and Reiwa

Edo Meiji/Taisho Showa Heisei Reiwa

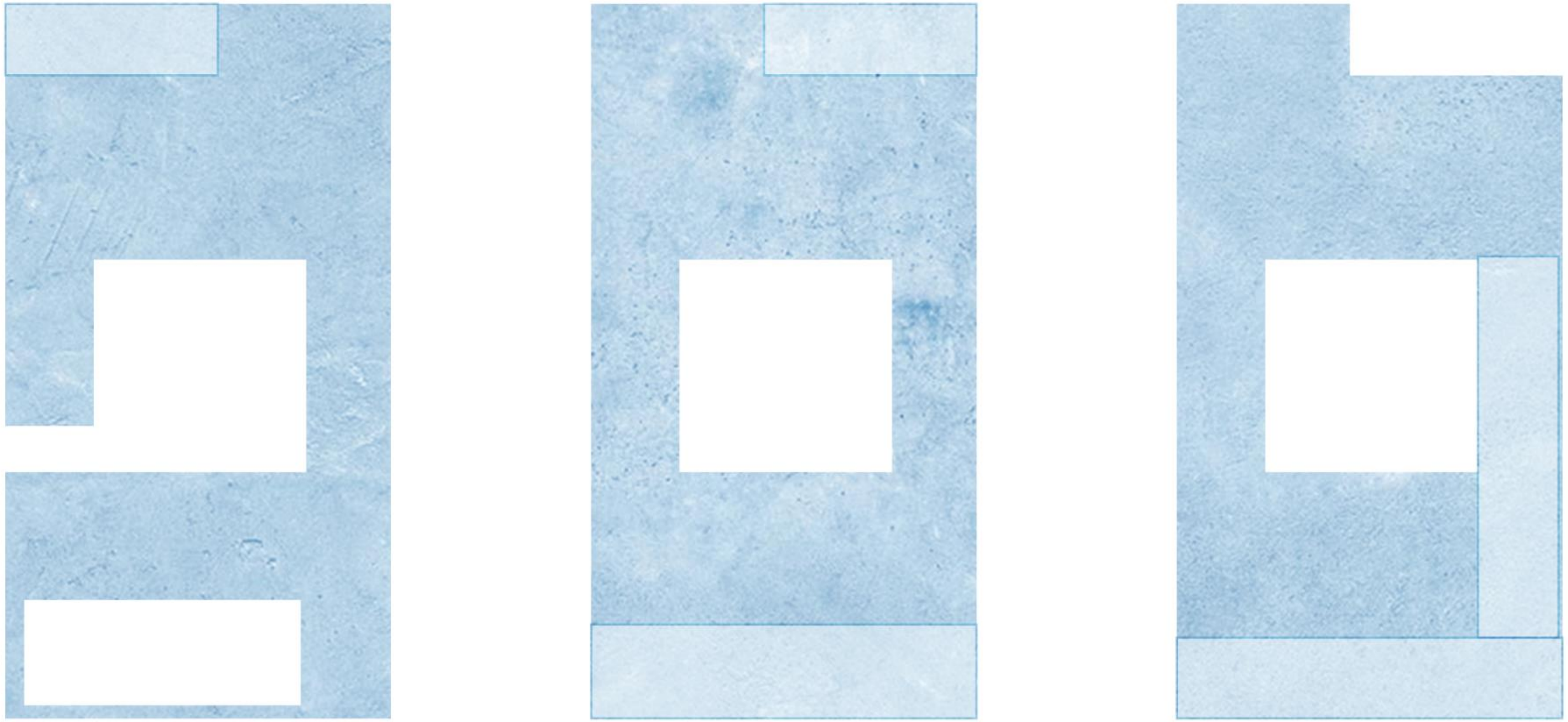
- 420 now



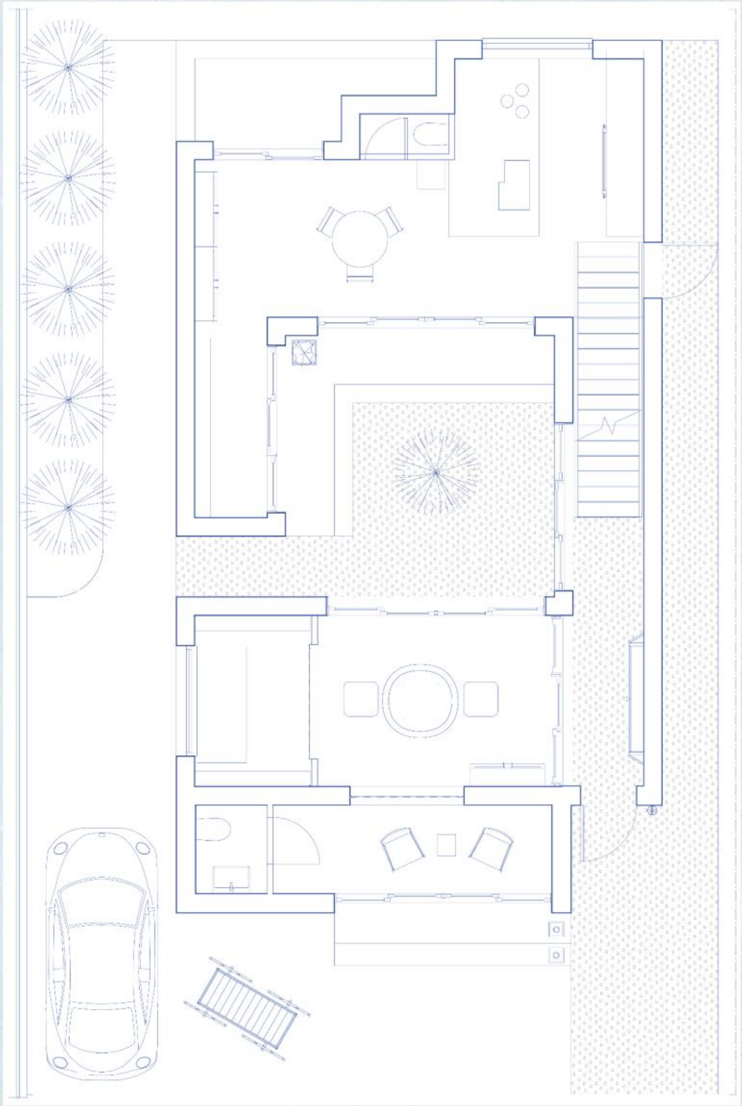
Original data of House 1

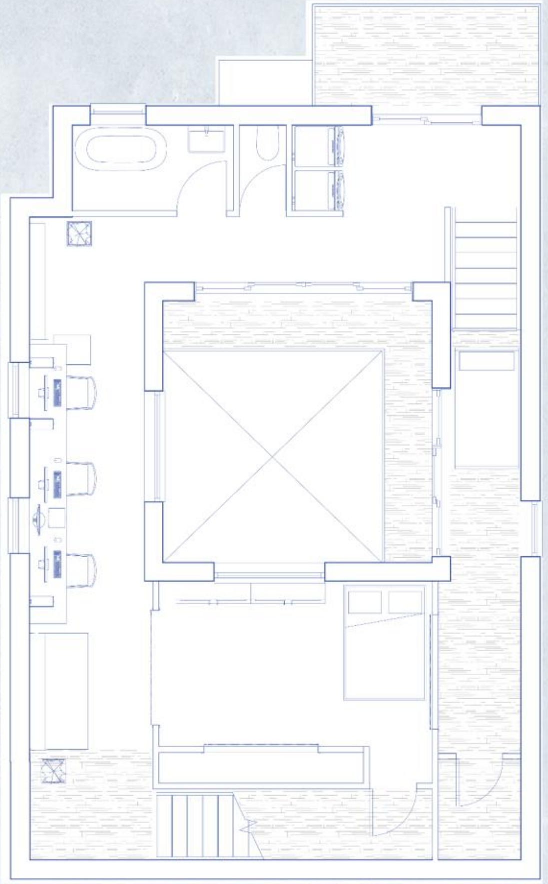


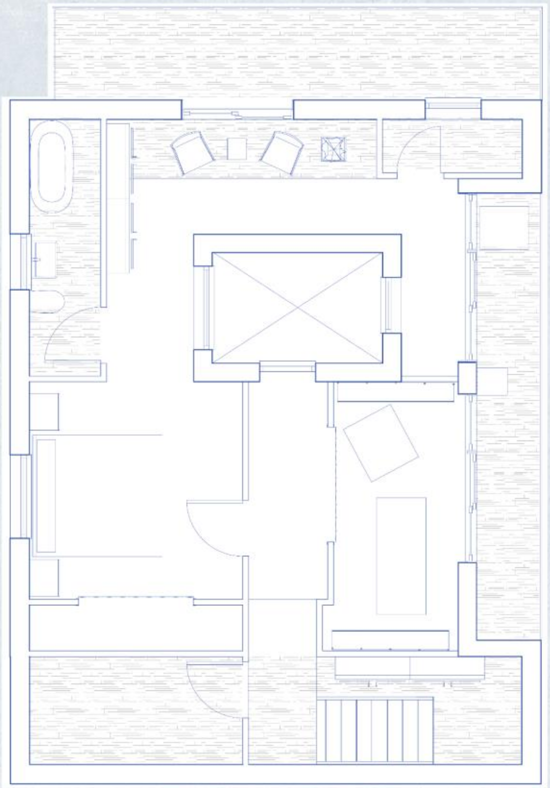
My process begins with an original building image that is fed into a stable diffusion model in order to generate suggestive images. I then curate and post-process those images to make them more legible while maintaining the suggestion of period influence. For the plans I sorted them into categories such as: wall structure, texture, spatial properties which were then extruded. Finally, I merged the hybridized façade onto the interior spaces.



In this case study, a distinct feature that is maintained throughout the diffusion transformation process is the courtyard space.

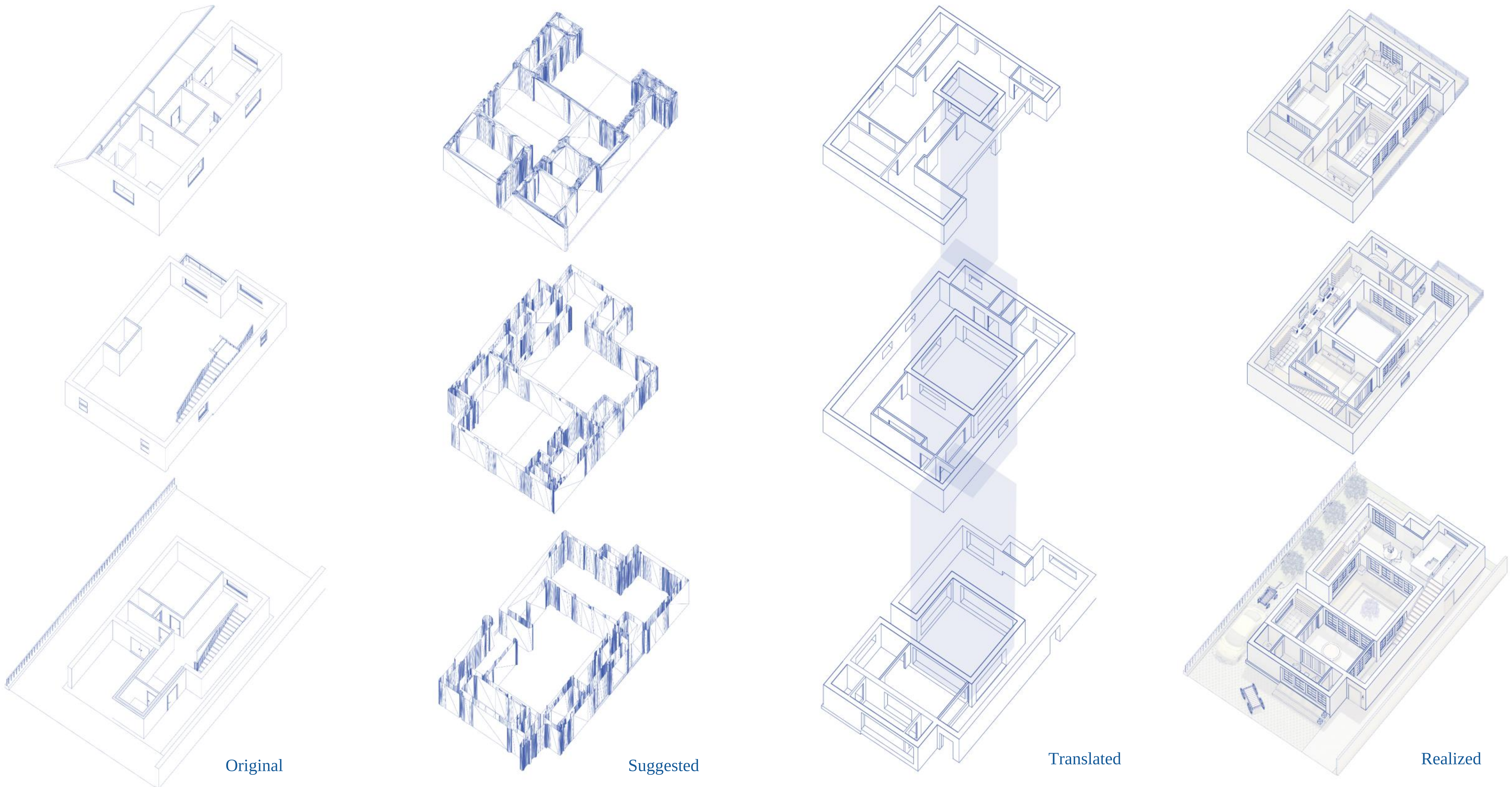




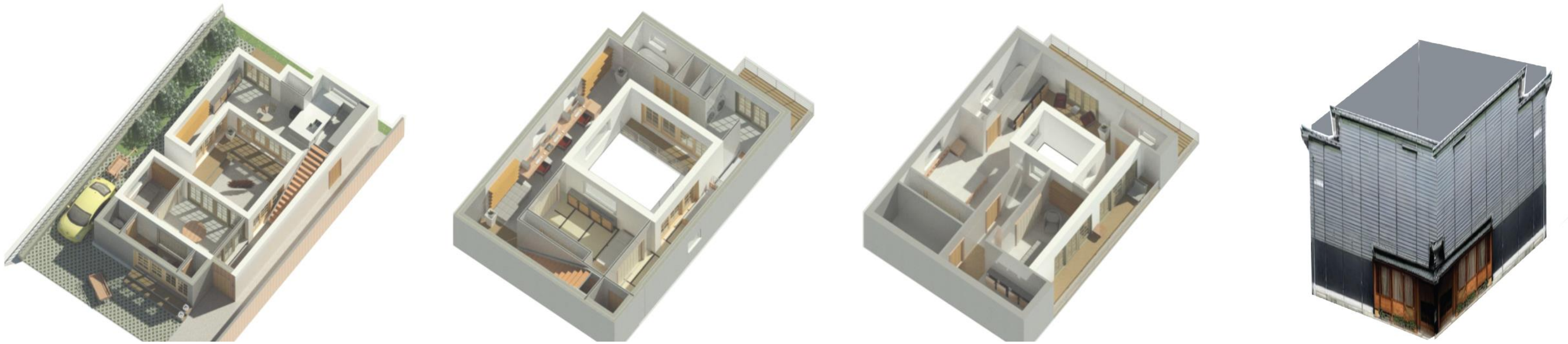


Interior spaces arranged around a central courtyard are typical in traditional Japanese homes and are known as 'tsubo-niwa'.

Transformed façade propose materiality palette mixed from the selected eras. Sugi-ban and wooden louver doors, ceramic tiles.



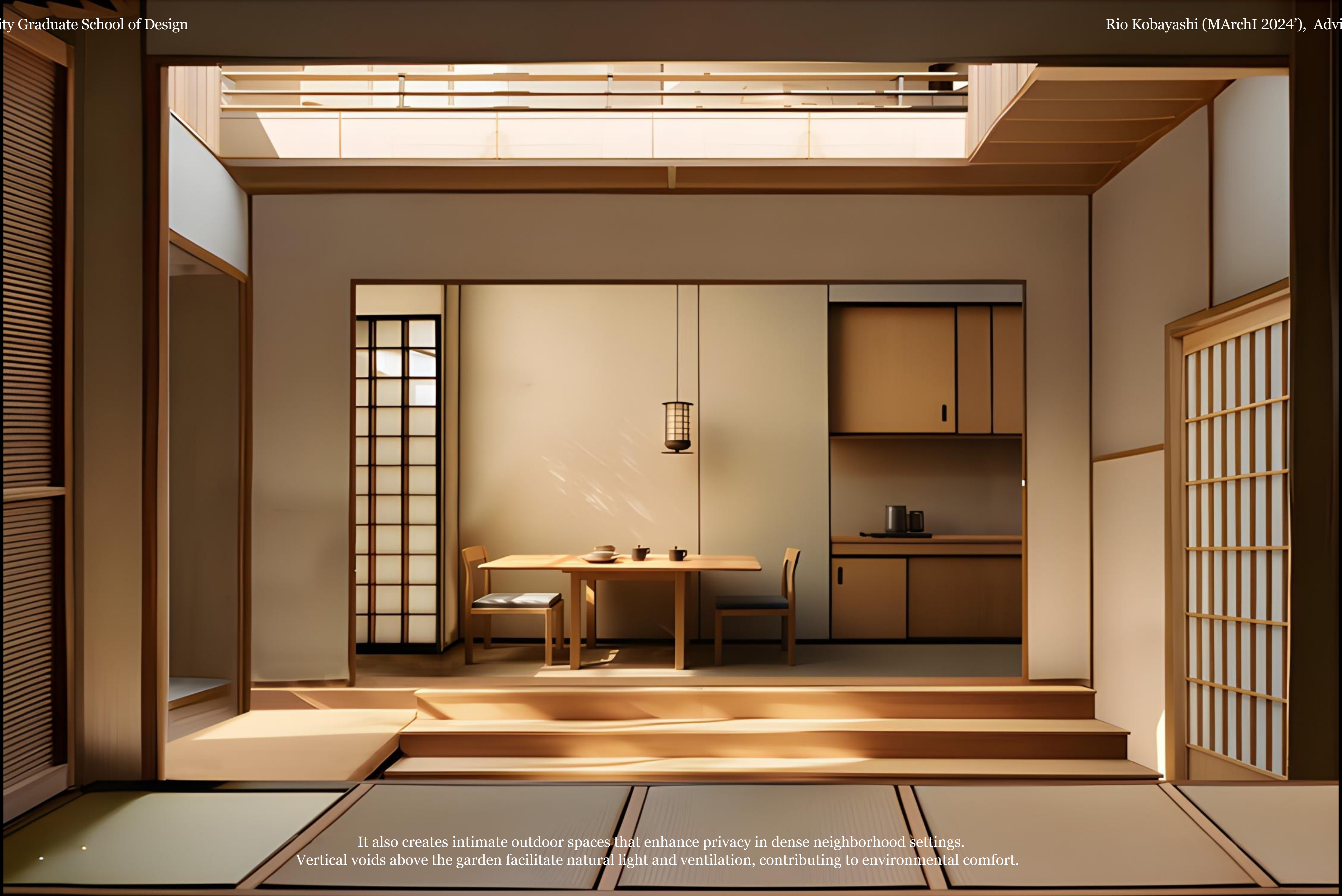
Further demonstrations of the transformation process: Here machine learning suggestions were translated and then realized in more specific and functional architectural language.



Here I further render machine-generated elements into detailed spaces and facades emphasizing an aesthetic of functionality and comfort.

Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt

The presence of an 'engawa', transitional space, further mediates between the personal and common areas, reinforcing the concept of semi-private zones that are open yet secluded. This approach allows the residents to experience the exterior environment without stepping outside, maintaining a delicate balance between exposure and privacy.



It also creates intimate outdoor spaces that enhance privacy in dense neighborhood settings.
Vertical voids above the garden facilitate natural light and ventilation, contributing to environmental comfort.

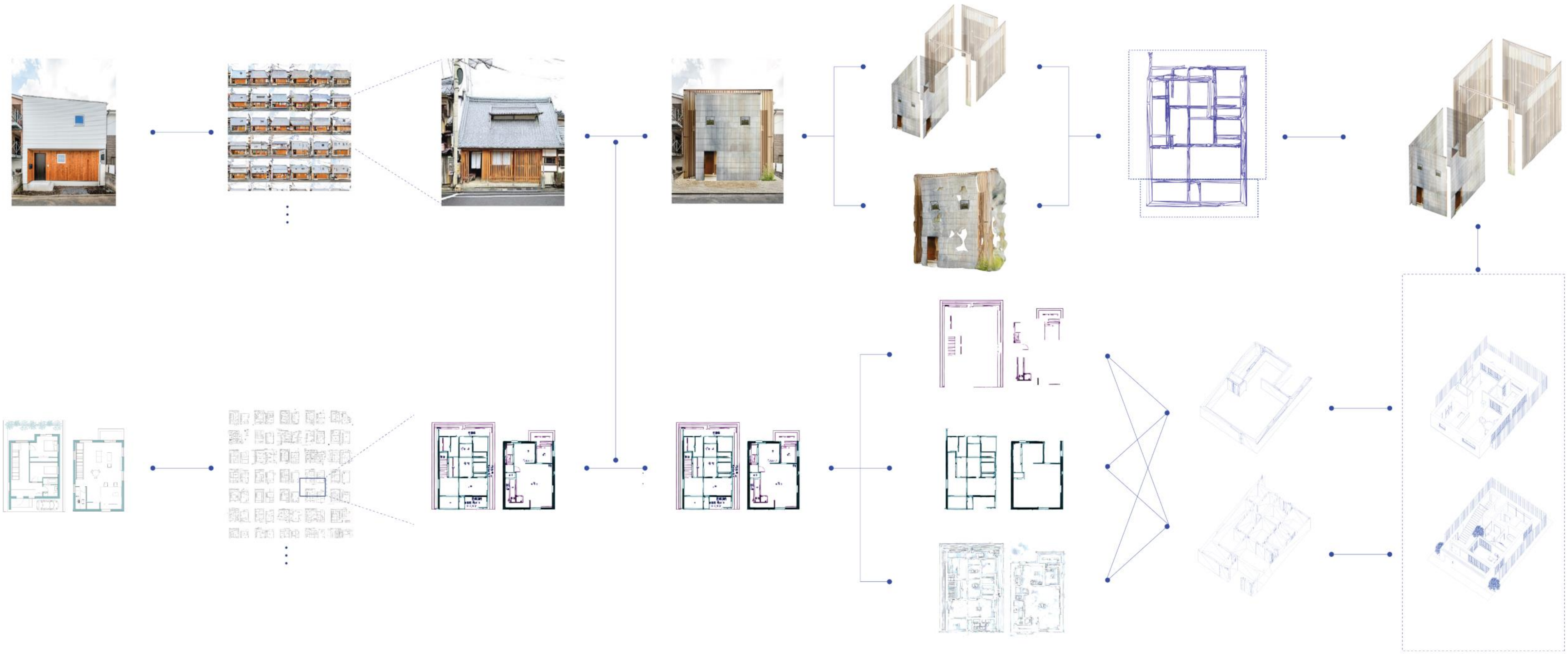
Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt

House 2
Synthesis of Showa, Heisei, and Reiwa

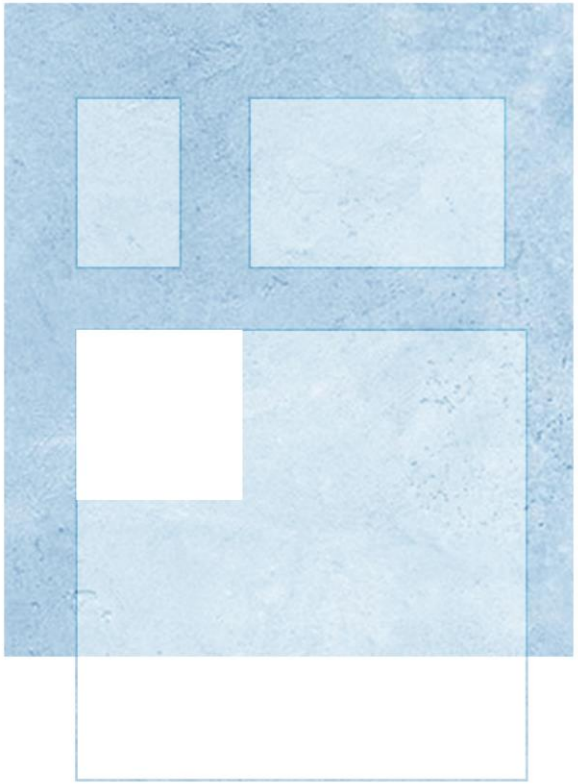
Edo Meiji/Taisho Showa Heisei Reiwa

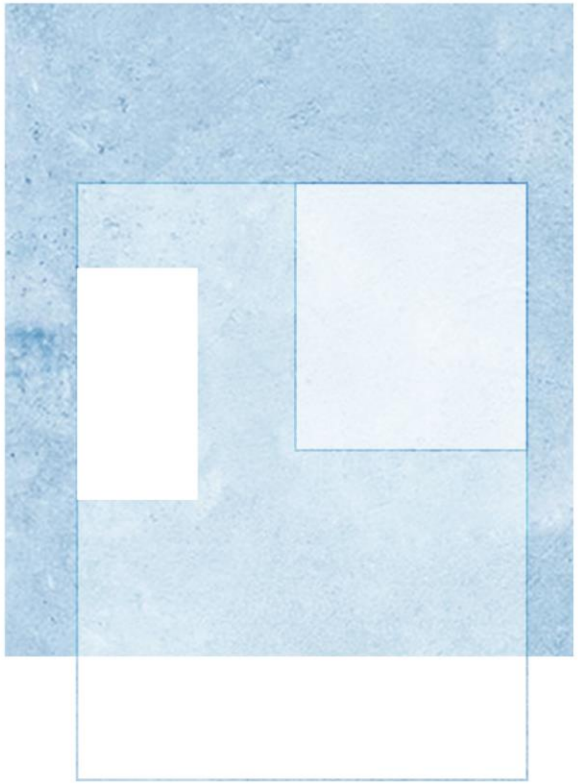
- 420 now

Original data of House 2



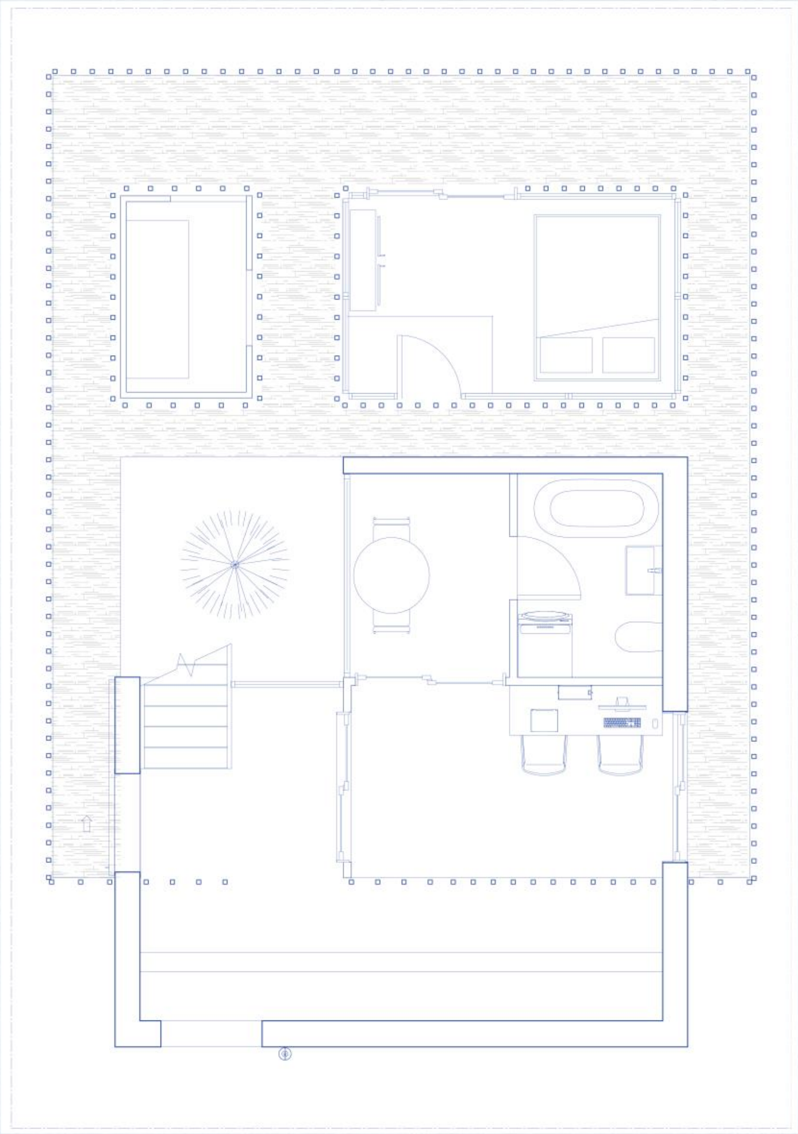
It follows the same process with the first case

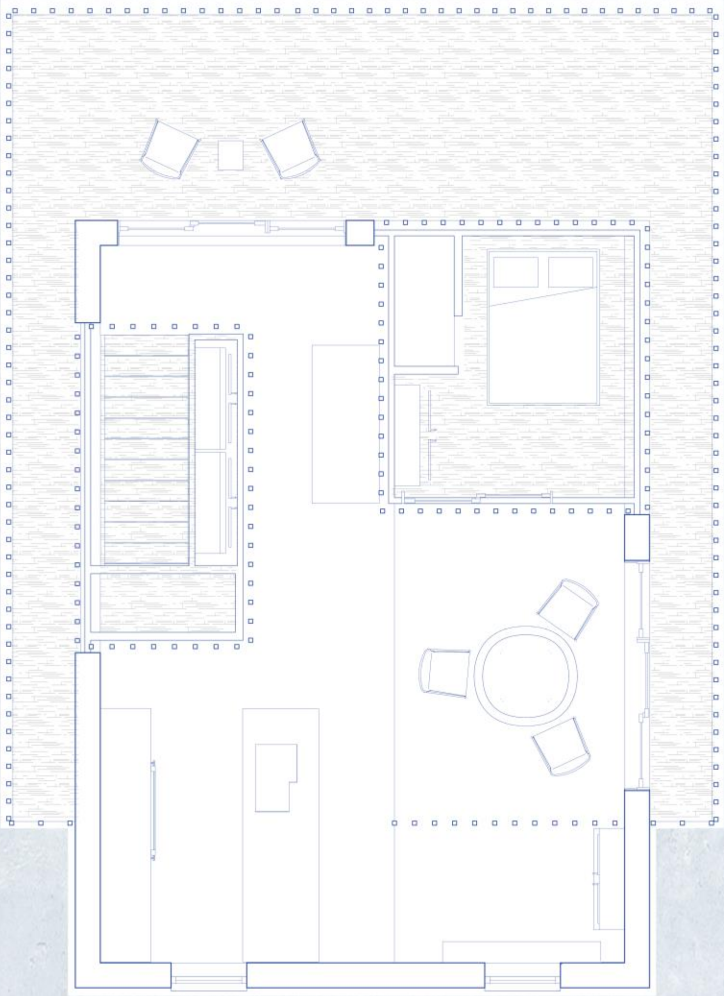




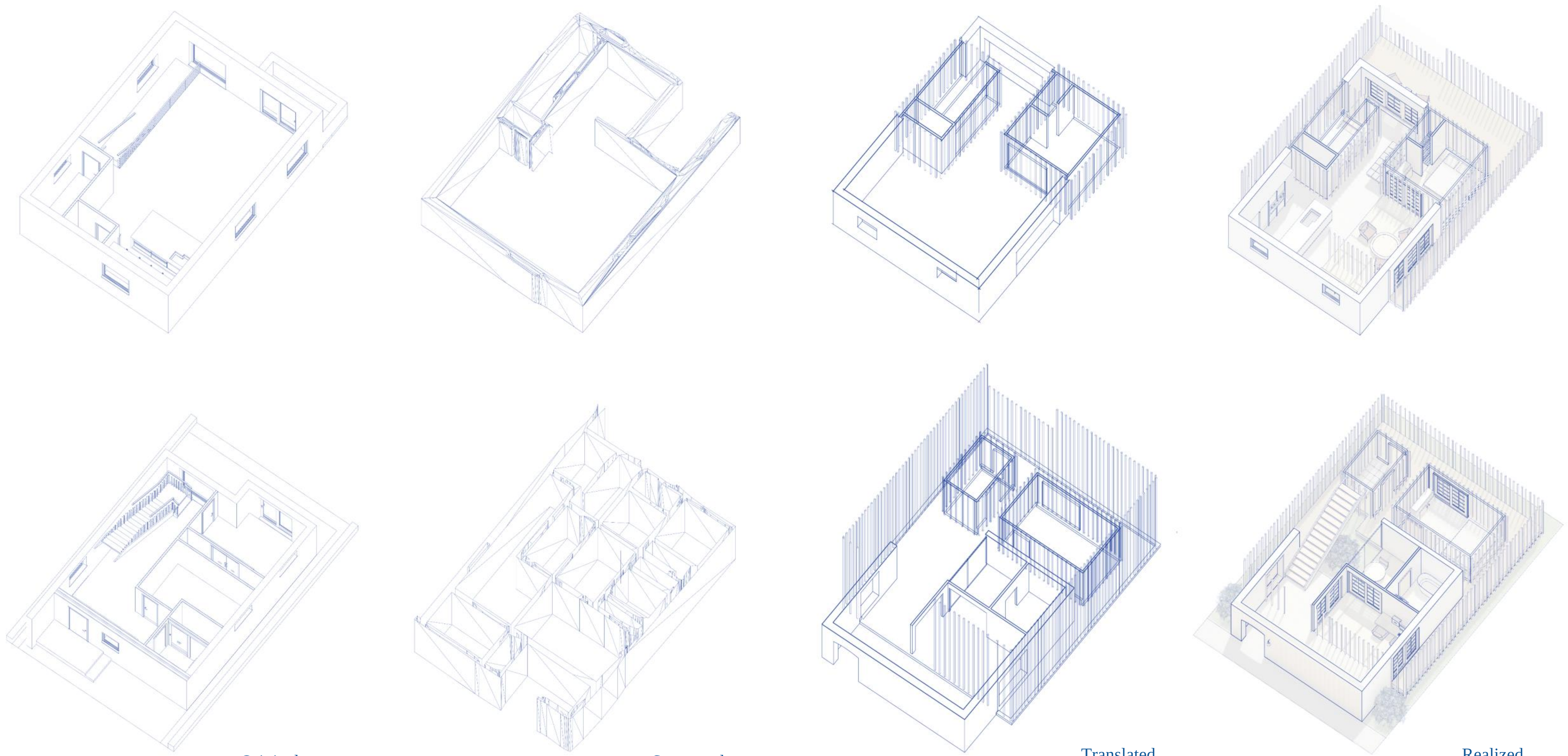
The organization of rooms in these images showcases a 'box-in-the-box' situation where spaces are defined yet retain a sense of openness.

Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt





The use of wooden louvers allows for a visual connection between spaces while still providing separation.



Original

Suggested

Translated

Realized

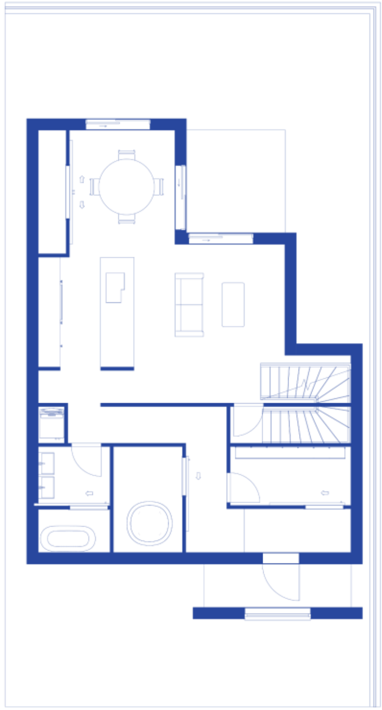
This arrangement is reflective of the Japanese principle of interconnectedness, where boundaries are suggested rather than explicitly stated,

allowing for flexibility and flow within the interior environment.

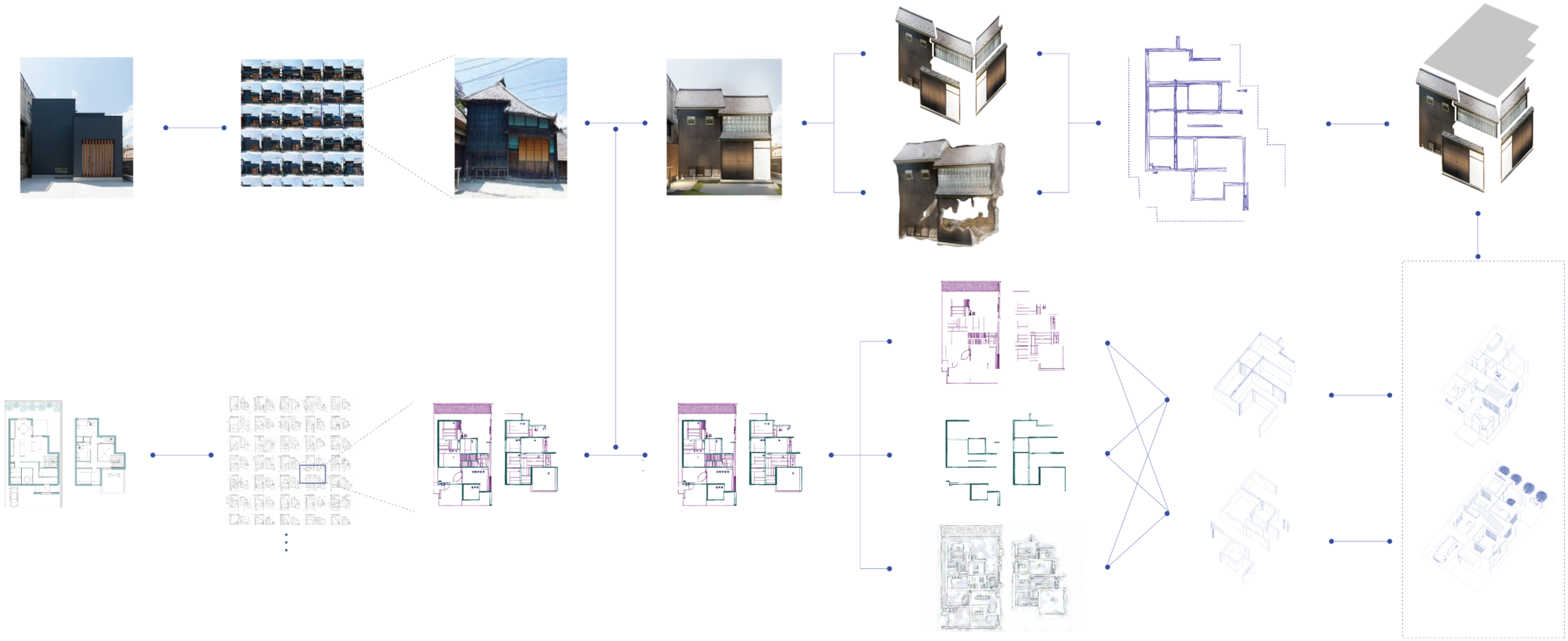
House 3
Synthesis of Meiji/Taisho, Showa and Heisei

Edo Meiji/Taisho Showa Heisei Reiwa

- 420 now

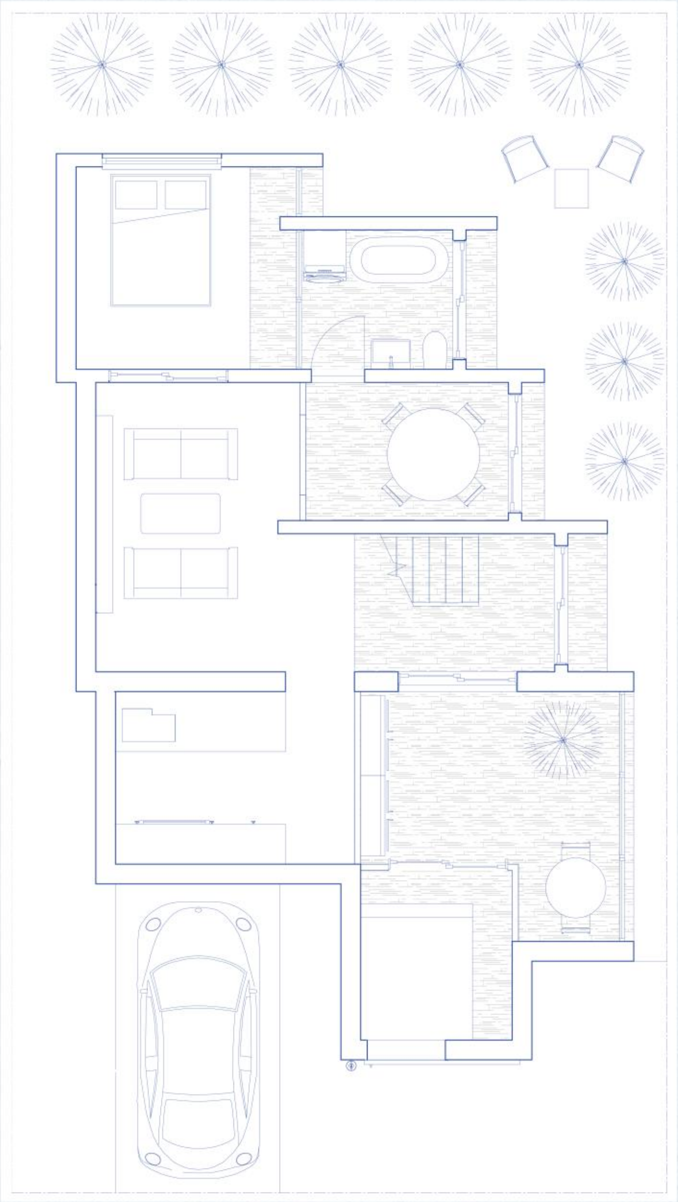


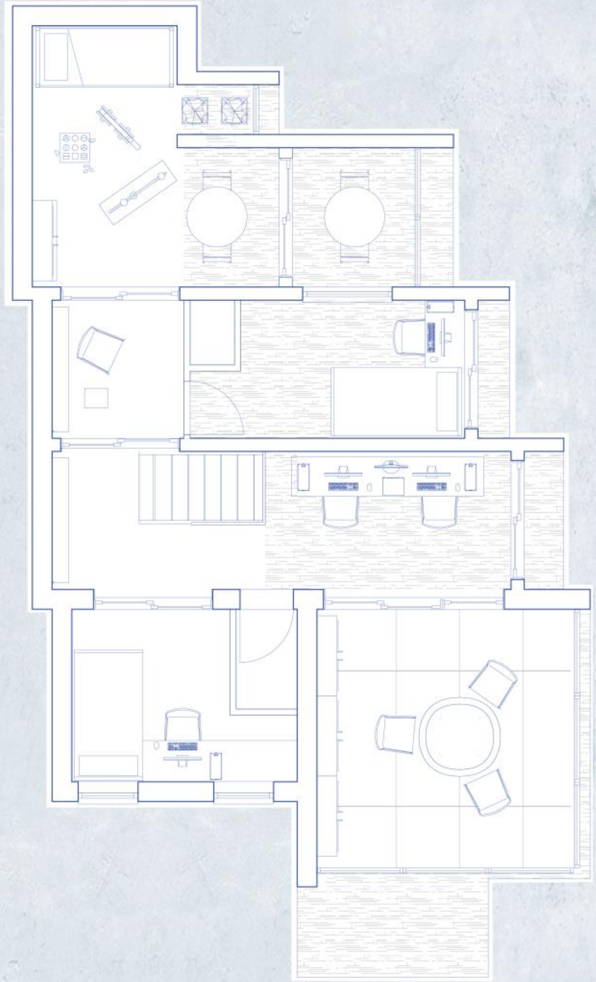
Original data of House 3



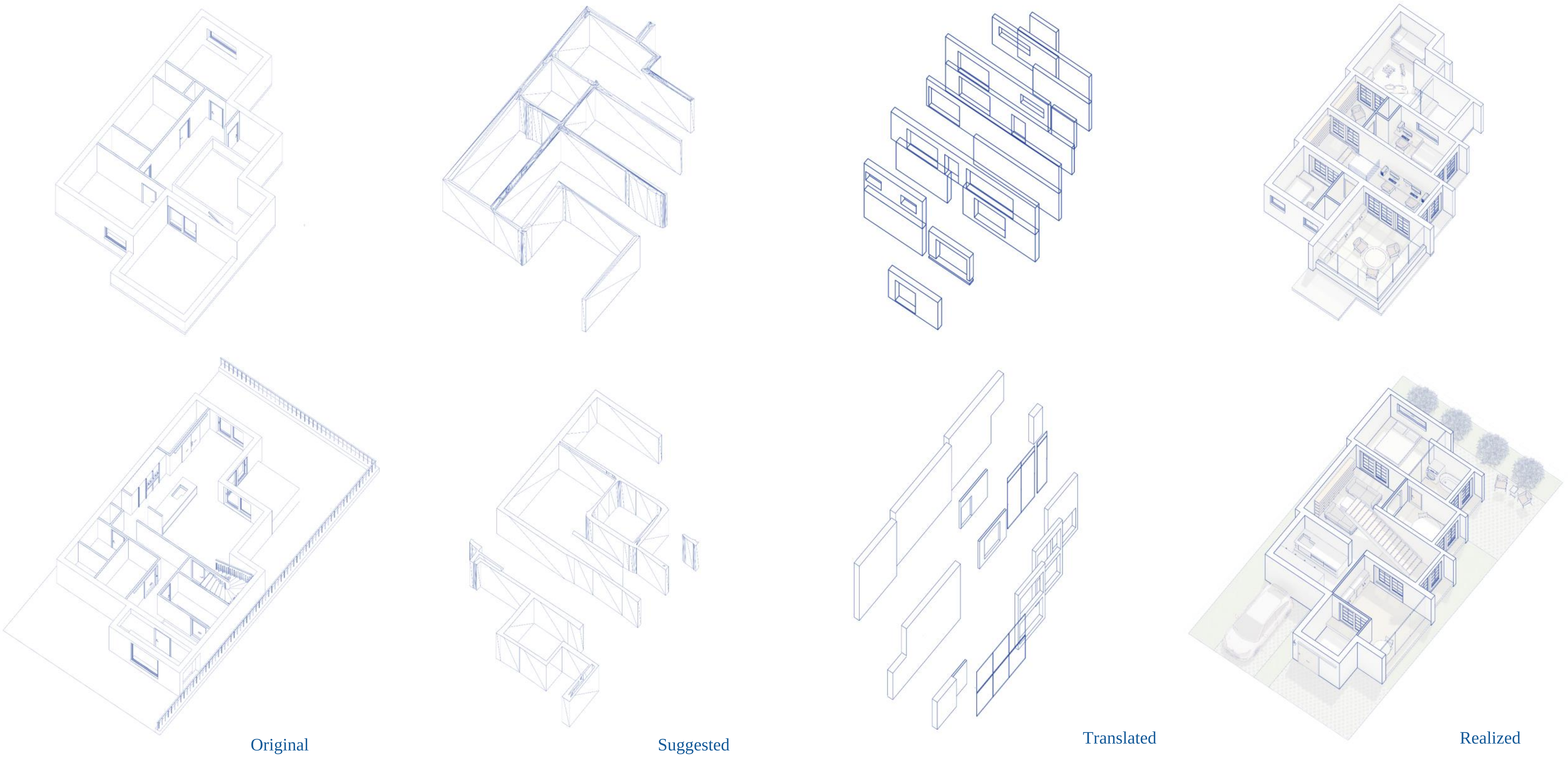
It follows the same process with the first case

In the process of translation of the suggested plan, this case exhibits the concept of layered walls,





The composition is linked to traditional Japanese architectural technique often seen in castles that creates depth and a dynamic effect through the use of Fusuma doors.

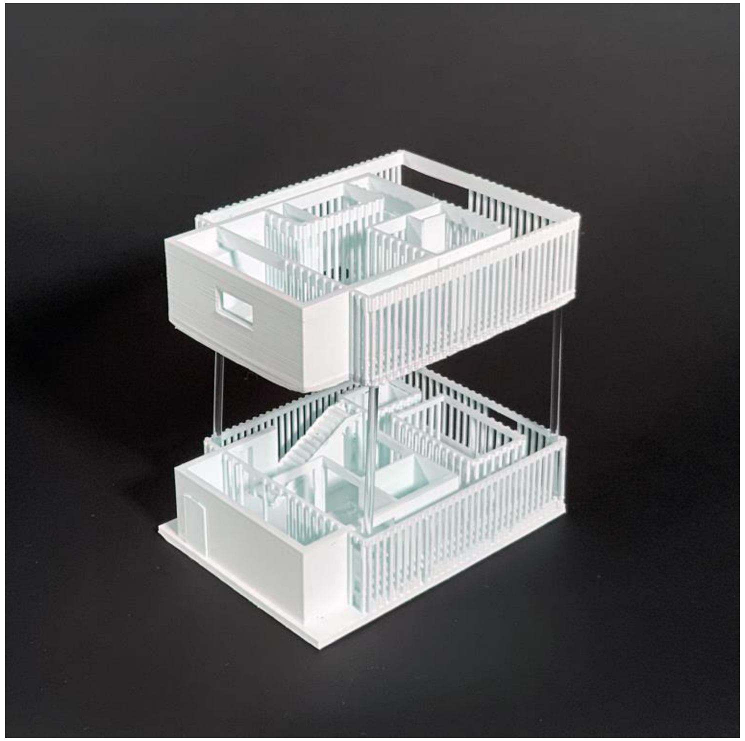

Original
Suggested
Translated
Realized

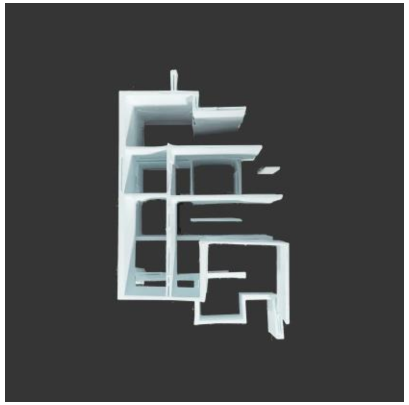
Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt

The sliding panels not only serve as functional room dividers but also contribute to the visual narrative of the space.
They create a sense of progression as one moves through the layers, offering varying degrees of privacy and interaction.

Rio Kobayashi (MArchI 2024'), Advisor: Andrew Witt

This arrangement can be a reflection of the Japanese spatial concept of 'ma', emphasizing
the importance of voids, or negative space, which are just as significant as the solid, or positive space.





Each image represents a nuanced interpretation of Japanese design principles through collaboration with machine intelligence and collaboration with the past.
The vast repository of architecture data embodies the diverse qualities of human existence that are closely intertwined with time and place,
allowing us to weave obsolete features and knowledge into contemporary design thinking and practice,
and in doing so, further enrich and support pluralistic projections far into the future.

This is just a humble introduction to prototypical workflow in the coming world of AI.

*GIF

e.g. 3D models generated by DreamGaussian [7]

The integration of AI in the design process has the theoretical potential to bypass the 2D phase completely, allowing for the direct generation of 3D data from 3D datasets.

However, these case studies indicate that 2D data-based generation process could be beneficial for desiners within this AI-enhanced design framework.

These cases illustrate the possibility of maintaining authorial control without compromising the dynamic interplay between design elements.
Machine intelligence is intended to enhance human creativity, not undermine it,
as long as we learn how to effectively merge it with our thinking mind.

AI in architecture functions much like an intricate loom networked to the vast archives of design knowledge.

It processes and proposes patterns that might elude the human eye, weaving new materials, technologies, and theories into the existing fabric.

This integration does not disrupt the continuity of the tapestry but rather enhances it,
introducing novel textures and colors that augment depth and nuance to the architectural landscape.

IMAGES

Fig.1 © IHA Photo, The Ughisar castle serves as the region's highest summit in Cappadocia, Tisrkiye, July 27, 2023. https://www.dailysabah.com/turkiye/castle-caves-in-turkiyes-cappadocia-offer-escape-to-tourists-from-heat/news.
Fig.2 Frearson, Amy. "Heydar Aliyev Center by Zaha Hadid Architects." Dezeen, July 11, 2013. https://www.dezeen.com/2014/07/01/designs-of-the-year-2014-zaha-hadid-saffet-kaya-bekiroglu-interview-heydar-aliyev/.
Fig.3 © Alexander Spatari via Getty Images
Fig.4 © Flickr user davidstanleytravel licensed under CC BY 2.00
Fig.5 © Dan Glasser. "Musée Yves Saint Laurent Marrakech / Studio KO" 25 Sep 2019. ArchDaily. https://www.archdaily.com/925363/yves-saint-laurent-museum-marrakech-studio-ko/. ISSN 0719-8884
Fig.6 © JINJA HONCHO. https://www.jinjahoncho.or.jp/en/ise/images/ind_im_001.jpg/.
Fig.7 © 2024 JAPAN PROPERTY CENTRAL K.K.. https://japanpropertycentral.com/2016/06/construction-starts-on-aoyamas-latest-luxury-condo/.
Fig.8 © Solum Real Estate https://solumre.com/en/projects/residencenook/.
Fig.9 IV Castle Lane Apartments / DROO" 09 May 2019. ArchDaily. https://www.archdaily.com/916606/iv-castle-lane-apartments-droo> ISSN 0719-8884/.
Fig.10 © Takuji Shimmura. Christophe ROUSSELLE architecte. https://archello.com/project/courbes /.
Fig.11 © Flickr user Curious Expeditions. https://www.flickr.com/photos/curiousexpeditions/3053929891/sizes/l/in/set-72157609976285956/.
Fig.12 Atilla's Cave House. Cappadocia, Turkey. https://www.airbnb.jp/hotels/6218207?guests=1&adults=1&s=67&unique_share_id=42e5a648-2b77-46b0-85f3-1be58eabf0e2/.
Fig.13 © DESIGNBLENDZ ARCHITECTURE. https://www.phillyvoice.com/kensington-apartments-lehigh-avenue-viaduct/.
Fig.14 © 2024 Taylor Wimpey Central London. https://www.taylorwimpeycentrallondon.com/postmark/plot/425/.
Fig.15 © Jason Allen. "Théâtre D'opéra Spatial," Jason Allen's A.I.-generated work
Fig.16 Kim, Dongyun. 2022. Latent Morphologies: Encoding Architectural Features and Decoding Their Structure through Artificial Intelligence. Master's thesis, Harvard Graduate School of Design.
Fig.17 Andrew Witt, 2022, Digital Media : Neural Bodie. Harvard Graduate School of Design.
Fig.18 Ji, Seung-Yeul, and Han-Jong Jun. 2020. "Deep Learning Model for Form Recognition and Structural Member Classification of East Asian Traditional Buildings" Sustainability 12, no. 13: 5292. https://doi.org/10.3390/su12135292 /.
Fig.19 archicraft. Yoshihusa Hukuda. https://www.archi-c.com/古民家-北本リフォーム/.
Fig.20 株式会社井上地所. https://www.inoue-chisho.co.jp/gallery/1377/.
Fig.21 LIVE JAPAN Perfect Guide. https://livejapan.com/en/article-a0000806/.
Fig.22 Miyagi, Taizo. "A Contextual Look at the Heisei Era." Japan Times. November 5, 2018. https://www.japantimes.co.jp/opinion/2018/11/05/commentary/japan-commentary/contextual-look-heisei-era/.
Fig.23 Hiyosi.net. "Construction from the Edo Period, the Oldest Existing Private House 'Sekike Residence' Open to 200 People on November 26 (Sat)." Yokohama Hiyoshi Newspaper, November 8, 2016. https://hiyosi.net/2016/11/08/cultural_property/.
Fig.24 nexthouse. https://nexthouse.jp/next/next-columns/simple-modern/.
Fig.25 -26 Daiwa House Industry Co., Ltd. "The Origin of Japan's Prefabricated House Industry: The Midget House." Daiwa House Group Technology Information. https://www.daiwahouse.com/English/tech/midgethouse.html.
Fig.27 SuperAnnotate. "What is Image Classification? Basics You Need to Know." SuperAnnotate Blog, May 30, 2023. https://www.superannotate.com/blog/image-classification-basics/.
Fig.28, 37, 38 Freedom Architects. "CASE746 Outlook." Design and Architecture: Freedom Architects. Construction Location: Kita, Tokyo. https://freedom.co.jp/architects/case746/.
Fig.29, 39, 40 Freedom Architects. "CASE690." Design and Architecture: Freedom Architects. Construction Location: Tokyo. https://freedom.co.jp/architects/case690/.
Fig.30, 41, 42 Freedom Architects. "CASE554 Blue Box House." Design and Architecture: Freedom Architects. Construction Location: Setagaya, Tokyo. https://freedom.co.jp/architects/case554/.
Fig.31 "What are Diffusion Models?" Lilian Weng , July 11, 2021 https://lilianweng.github.io/posts/2021-07-11-diffusion-models/.
Fig.32 "Stable Diffusion text-to-image fine-tuning." In Diffusers Documentation, v0.13.0. Hugging Face. https://huggingface.co/docs/diffusers/v0.13.0/en/training/text2image/.
Fig.33 Paul, Sayak, and Chansung Park. "Fine-tuning Stable Diffusion." Keras.io. Last modified January 13, 2023. https://keras.io/examples/generative/finetune_stable_diffusion/.
Fig.34 "Image-to-image - StableDiffusionImg2ImgPipeline." Diffusers Documentation. Hugging Face. https://huggingface.co/docs/diffusers/api/pipelines/stable_diffusion/img2img/.
Fig.35 © Alex Ziminski. https://tokyocheapo.com/entertainment/things-to-do-in-asakusa/
Fig.36 © Japan Wonder Travel. https://www.tripadvisor.com/AttractionProductReview-g1066441-d12648912-Yanaka_Shitamachi_Backstreets_Cultural_Walking_Tour-Arakawa_Tokyo_Tokyo_Prefectur.html
Fig.43 Rombach,Robin, et al. "High-Resolution Image Synthesis With Latent Diffusion Models." CVPR 2022.
https://openaccess.thecvf.com/content/CVPR2022/papers/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.pdf.

REFERENCES

[1] Wikipedia contributors. "Contextual Architecture." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Contextual_architecture.
[2] Wikipedia contributors. "Vernacular architecture." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Vernacular_architecture.
[3] Wikipedia contributors. "Indigenous architecture." Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Indigenous_architecture
[4] "Tradition." In Britannica Online Dictionary. https://www.britannica.com/dictionary/tradition.
[5] Rethinking The Future. 'What is the difference between regionalism and critical regionalism?'
https://www.re-thinkingthefuture.com/architectural-community/a8886-what-is-the-difference-between-regionalism-and-critical-regionalism/.
[6] Previously Unknown Major Painting by Raphael and Studio Discovered." ArtDaily, February 22, 2023. https://artdaily.com/news/154718/Previously-unknown-major-painting-by-Raphael-and-studio-discovered/.
[7] Tang, J., Ren, J., Zhou, H., Liu, Z., & Zeng, G. (2023). DreamGaussian: Generative Gaussian Splatting for Efficient 3D Content Creation. arXiv preprint arXiv:2309.16653.

Dedicated to

my parents, Mari and Naonori
my aunt and sister, Miki and Nanao
my grandparents, Yoshiaki, Atsumi, and Noriko
and all of my friends and advisors

Thank you;)