Evaluating Workplace Wellbeing: A Framework Using Biometric Data and Neural Networks

Abstract (250 words)

The post-pandemic work landscape demands a paradigm shift in workplace design, prioritizing both employee wellbeing and rejuvenation. While a substantial number of remote working Americans are refusing to return to the office, studies show that hybrid or remote workers experience a higher rate of mental health issues due to constant interruptions at work. Therefore, the role of a dedicated workplace becomes more crucial, as it provides a designated space to focus and supports a work-life boundary essential for both physical and mental health. This paper posits that effective workplace design, particularly through visual elements such as spatial forms, materials, and coloros, can improve physiological and psychological responses. These design elements influence the long term wellbeing through sensory stimulations and by shaping perceptions of space. Although standards such as the WELL standard and LBC 4.0 provide design guidelines that promote health, they fall short in quantitatively demonstrating their impact on occupants' wellbeing, as well as being diluted by broader issues just as environmental sustainability and building performance. Addressing this gap, our research aims to develop a framework using convolutional neural networks (CNNs) trained on images of interior environments, as well as human valence and arousal responses. This model will test against the evaluation based on electroencephalogram (EEG), electrodermal activities (EDA), and resting heart rate (RHR) to quantitatively evaluate how visual design elements influence employee wellbeing. This approach promises to inform more effective workplace design that not only meets operational and performance needs but also promotes employee wellbeing and facilitates rejuvenation.

Keyword: *workplace wellbeing; Convolutional Neural Networks; biometric data; visual stimuli, human responses*

1. Introduction

In the aftermath of the COVID-19 pandemic, the traditional workplace has undergone a paradigm shift in which remote or hybrid work schedule has become a norm. According to WFHResearch, 12.7% and 28.2% of employees work remotely and in hybrid respectively (Aksoy et al. 2023). Reli Exchange's survey highlights that nearly 26% of remote workers resist returning to the office, despite potential job loss risks . This shift has stirred debates around productivity and mental health impacts. For instance, while the Harvard Business Review suggests remote workers enjoy heightened productivity and satisfaction, Stanford SIEPR notes a 10% productivity drop in fully remote settings due to communication and motivation challenges (Barrero et al., 2023). Nonetheless, the reduced costs from space savings and global hiring ostensibly compensate for these deficits, suggesting an overall productivity gain.

However, the mental health implications of prolonged remote work are concerning. Studies by Costin et al. reveal an increase in burnout, stress, and emotional exhaustion among remote workers, exacerbated by inadequate training and technology adaptation challenges (Costin et al., 2023). The lack of workplace and social connections, alongside increased job demands and insufficient organizational support, further deteriorates mental health and job satisfaction. These findings underscore the importance of the physical workplace not just for operational efficiency but also as a vital component of employee wellbeing. Amidst this backdrop, the design of the workspace has gained unprecedented importance. Current building standards, primarily focused on sustainability and environmental impacts, often overlook human health and wellness aspects. The WELL Building Standard emerges as a notable exception. Yet, its guidance on promoting workplace wellbeing through design remains vague in associating the effects of design aspects. For example, the materials chapter includes safety guidelines to reduce human exposure to toxic chemicals such as asbestos, mercury, and lead. However, promoting safety of a building is not the same as promoting human health, as it does not address how materials impact occupants physiological and psychological wellbeing as discovered in studies as Medhat Assem et al. (2023), Yin et al. (2019, 2020), and Elbaiuomy et al. (2019).

This research aims to bridge the gap in understanding how architectural design influences employee wellbeing by utilizing artificial intelligence (AI), particularly machine learning (ML) to analyze biometric data related to visual stimuli. This innovative approach seeks to establish a clear correlation between design elements and wellbeing indicators, moving beyond traditional evaluation methods to offer a quantifiable analysis of design impacts. The study's objectives are twofold: to harness AI in identifying specific design elements that not only meet operational needs but also support the mental and physical health of occupants, a crucial consideration in the post-pandemic era. Through a comprehensive analysis of existing research and digital tools, this study aims to influence future design practices significantly, advocating for workspaces that nurture occupant wellbeing in every aspect.

2. State of the Art

2.1. Workplace Wellbeing Definition

The concept of "workplace wellbeing" related to the built environment first emerged in 1911 and has grown significantly over the century. Aryanti et al. (2020) explore workplace wellbeing as the sense of prosperity employees derive from work, promoting sustainable retention, productivity, and psychological health. Litchfield (2021) describes it as the holistic health status of employees within a work environment, encompassing physical, mental, and social aspects. Myerson et al. (2017) define it as the balance between employees' psychological, physical, and social resources and external challenges, ensuring productivity and mental health.

WELL 2.0 and LBC 4.0 IEQ are pioneering standards that incorporate wellbeing into their design guidelines. WELL 2.0 emphasizes spatial geometries for movement, biophilic design principles, and materials to enhance wellbeing. LBC 4.0 focuses on spatial design flexibility, aesthetically pleasing colors reflecting the local environment, and the environmental impact of materials. However, these guidelines are limited in evaluating the effectiveness of proposed strategies.

Synthesizing these findings, workplace wellbeing integrates physical, mental, and social aspects influenced by the built environment. In our study, wellbeing is significantly affected by design elements such as form, material, and color. These controllable aspects play a crucial role in shaping environments that promote positive valence (emotional positivity) and optimal arousal (alertness and engagement). Understanding these design aspects aims to create workplaces that enhance productivity, satisfaction, and overall mental and physical health.

2.2. Metrics for Quantifying Wellbeing

To fill the missing gap, biometric measurements have been utilized to provide a comprehensive insight into how architectural designs influence human emotional and physiological states through biometric feedback and neural responses. Kim J. and Kim N. (2022) establish how design elements affect emotional responses using biometric indicators. A. Mostafavi (2022) extends this to virtual environments, enhancing control over experiment variables and replicability of findings as shown in Fig.2.1. Sandra G.L. Persiani et al. (2021), contribute by detailing real-time physiological impacts of design features, enhancing the temporal resolution of data on occupant wellbeing. Sameh Azzay et al. (2021) add depth by focusing on the neural impacts of these environments, emphasizing the need for neuro-architectural research to fill knowledge gaps. These studies demonstrate design features elicit quantifiable changes in brain and physiological functions and underscores the critical role of architectural design in enhancing human wellbeing. The common metrics are electroencephalogram (EEG), heart rate variability (HRV), and electrodermal activity (EDA).

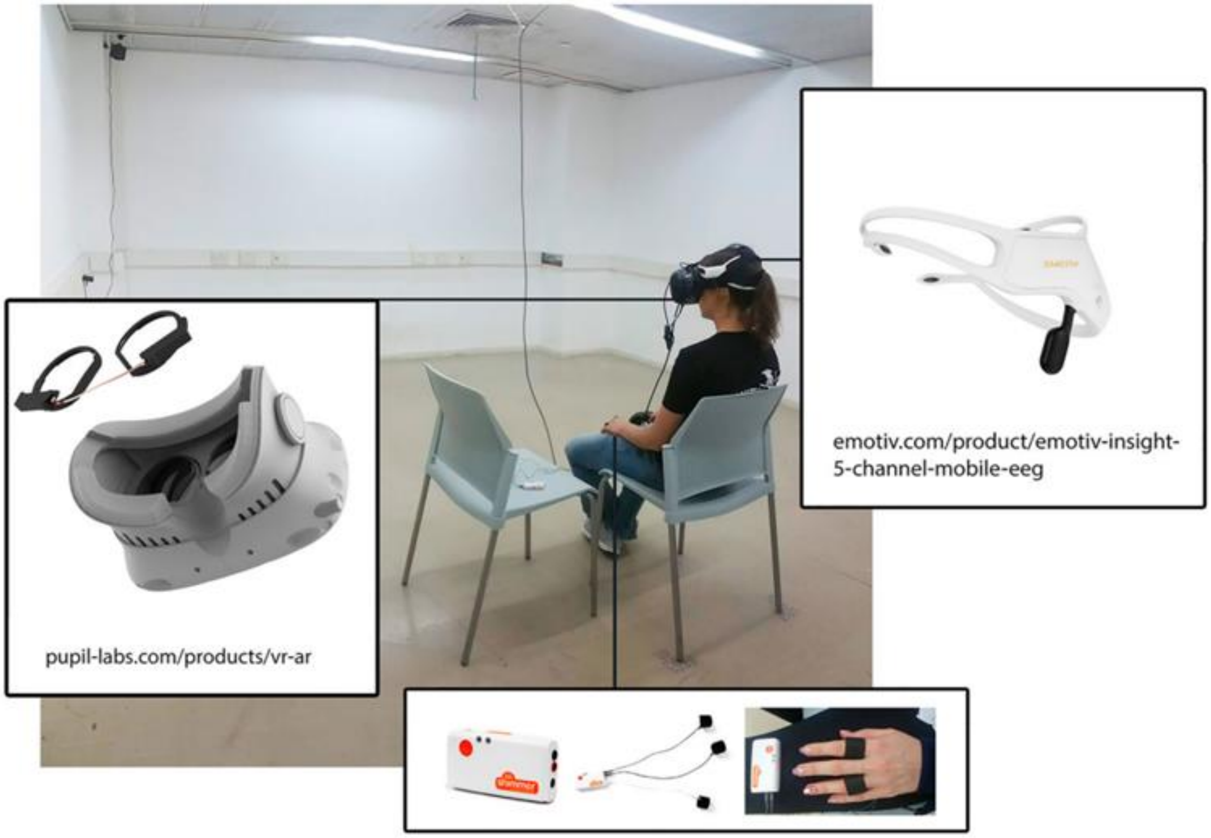


Fig.2.1: Data triangulation between architecture and biometric data (Mostafavi 2022)

2.2.1 Methods for Quantifying Wellbeing (Measuring EEG, HRV, EDA)

Electroencephalogram (EEG) is an instrumental tool in neurophysiological diagnostics, reflecting the electrical activity generated by neural function within the brain (). There are five major frequency bandwidths: Delta (δ) at 0.5 – 4 Hz, associated with sleeping and dreaming state; Theta (θ) at 4 – 8 Hz, associated with deeply relaxed and meditation state; Alpha (α) at 8 – 12 Hz, associated reflective and restful state; Beta (β) at 12 – 35 Hz, associated with busy and active mind; and Gamma (γ) at 35+ Hz, associated with problem solving and concentration mind (Abhang et al. 2016). These bandwidths can be further translated into various psychological states based on valence and arousal scale (Shemesh et al. 2022). Among many other EEG headsets, this study selects Emotiv Insight 2.0 that offers a wireless, multi-channel device that captures real-time brain activity. This device is equipped with five sensors that measure EEG data across different channels, facilitating the analysis of emotional and cognitive responses through non-invasive methods. Its portability, ease of use, and cost-effectiveness make it an excellent tool for both scientific research and consumer applications. It has been validated in various studies such as Zabcikova (2019), Partama et al. (2020), and Shemesh et al. (2017, 2021) that highlight its reliability and comparative effectiveness to other EEG devices.

Heart Rate Variability (HRV) is a sophisticated metric used to gauge the variation in time intervals between consecutive heartbeats, known as inter-beat intervals (IBIs). It considers the exact moment-to-moment changes in heartbeats, providing deeper insights into autonomic nervous systems (ANS) functions, as well as the ability to respond to stress and environmental changes (Arza et al. 2015; Ge et al. 2020). Resting Heart Rate (RHR) is another valuable physiological marker for assessing stress level, as stress typically activates the sympathetic nervous system, leading to an increased heart rate. Studies like Chalmers et al. (2021) and Altini and Plews (2021) have shown individuals under stress exhibit significantly higher RHR compared to their baseline level.

Electrodermal activity (EDA), also known as galvanic skin response (GSR), is a method at measuring the electrical conductance of the skin, which varies with its moisture level. This physiological phenomenon is primarily used to gauge ANS activity, particularly the sympathetic branch that largely controls sweat gland activity. It is extensively used in psychology to study emotional responses, particularly because it is directly affected by arousal states which cannot be easily masked or controlled voluntarily (Affanni 2020).

2.2.4. Valence and Arousal Scale

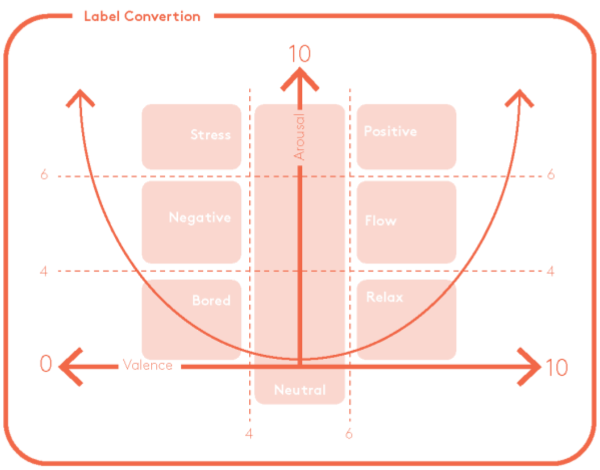


Fig.2.2: Mental state labels on Valence and Arousal Scale (VAS)

Valence and arousal are two fundamental dimensions used to describe emotional experiences and are integral to the psychological assessment of affective states. Valence refers to the pleasantness or unpleasantness of an experience, while arousal indicates the degree of activation or excitement that the experience generates. This two-dimensional approach as shown in Fig.2.2 is critical in fields like psychology, neuroscience, and human-computer interaction to quantify and evaluate emotional responses. EEG studies often focus on specific brain regions and frequencies that correlate with arousal and valence. For instance, frontal asymmetry in the alpha band is commonly linked to valence, whereas changes in the beta and theta bands can indicate variations in arousal. This capability of EEG to detect subtle changes in brain activity related to emotional states makes it invaluable for research in neuromarketing, psychiatric assessments, and interface design, where understanding emotional responses is crucial. For example, in neuromarketing, companies analyze consumer reactions to products or advertisements at the neurological level, using EEG to gauge emotional responses based on arousal and valence scales. This approach helps in refining marketing strategies to better align with consumer emotions, enhancing engagement and effectiveness. One pivotal study by Schmidt et al. (2001) outlines methods for quantifying emotional valence using EEG by examining frontal brain asymmetry, providing a foundational approach for subsequent research in this area. Another significant contribution by Posner et al. (2005) details the use of EEG to assess both arousal and valence, further cementing the importance of these emotional dimensions in psychological and neuroscientific research.

2.2.4. Visual Perception of Spatial Forms, Materials, and Colors

This study focuses on "environmental conditions", specifically the impacts on occupants' wellbeing through visual stimuli in relation to spatial geometries, colors, and materials.This is because they are few of the most fundamental decisions in design that have proven to be effective in shaping occupants' physiological responses. For instance, curved and symmetrical spaces tend to be associated with higher levels of positive emotions and lower arousal, making them feel more pleasant and safe. Conversely, sharp-edged and asymmetrical spaces often elicit higher arousal and less positive valence, potentially due to their association with discomfort or threat (Shemesh et al. 2017, 2021). Natural materials like wood and stone are often associated with warmth and comfort, leading to positive valence and a calming effect, reducing arousal (Medhat Assem et al. 2023; Yin et al. 2019, 2020). In contrast, materials such as glass and steel can evoke feelings of modernity and coldness, potentially increasing arousal due to their association with high-tech environments but possibly decreasing positive valence(Medhat Assem et al. 2023; Elbaiuomy et al 2019). In relation to colors, studies have shown that bright and saturated colors, such as red and yellow, tend to increase arousal and can evoke feelings of excitement or urgency. Conversely, cool colors like blue and green are associated with calming effects, reducing arousal and promoting relaxation. Warm colors (reds, oranges, and yellows) often elicit positive valence in the form of feelings of warmth and comfort, while cool colors can either be calming (positive valence) or perceived as cold and distant (negative valence) depending on the context. Additionally, colors such as gray or beige are often perceived as neutral or dull, leading to lower arousal and neutral or slightly negative valence (Kuller et al. 2009; Yoon and Wise. 2014; Cha et al. 2020).

2.3 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed for processing structured grid data, such as images (LeCun et al. 2015; Alzubaidi et al. 2021). CNNs have revolutionized computer vision tasks by effectively capturing spatial hierarchies in data through convolutional layers (LeCun et al. 2015). These networks are particularly powerful for tasks that involve images as they can automatically and adaptively learn spatial hierarchies from raw pixel data (Krizhevsky et al. 2017). Key components of CNNs include convolutional layers, pooling layers, fully connected layers, and activation functions. Convolution layers apply convolution operations to the input data using filters (kernels) to extract local features, such as edges, textures, and patterns (Krizhevsky et al. 2017). Pooling layers reduce the spatial dimensions of data which helps make the model invariant to small translations and reduce computational complexity (Simonyan et al. 2015). Fully connected layers perform the final classification or regression tasks (LeCun et al. 2015). Non-linear activation functions like ReLU (Rectified Linear Unit) introduce non-linearity into the model, enabling it to learn complex patterns (Krizhevsky et al. 2017). CNNs have been widely used in various domains, demonstrating their versatility and effectiveness. For instance, they are used in image classification to identify objects within images and assign them to predefined categories. For image segmentation, CNNs divide an image into segments or regions, which is particularly useful in image analysis.

In the context of classifying images of interior environments based on their impacts on occupants' wellbeing, CNNs offer significant potential. By analyzing the visual characteristics of interior space, CNNs can predicate valence and arousal responses, which are crucial indicators of emotional wellbeing. Therefore, our research focuses on utilizing CNNs to address the gap between environment designs and their wellbeing impact, leveraging their capabilities to effectively analyze complex data of the interior images.

3. Method

3.1 Framework

The study employs a mixed-methods research design, integrating both quantitative and qualitative approaches to capture a holistic view of workplace wellbeing. It first collects 900 images of interior environments to create a pilot survey, to collect qualitative insights from participants feedback for a dataset to train a CNNs model. Besides, the study photographs actual environments and collects participants' physiological responses in terms of EEG, EDA, and RHR. These photographs and biometric data are later used in testing the performance of the CNNs model. Moveover, this study performs a series of regression learning algorithms to discover additional insights, such as the relationships between design elements and their wellbeing impact to occupants, measured in terms of valence and arousal. This framework, shown in Fig.3.1, facilitates a detailed analysis of biometric data while contextualizing findings within the broader workplace environment.

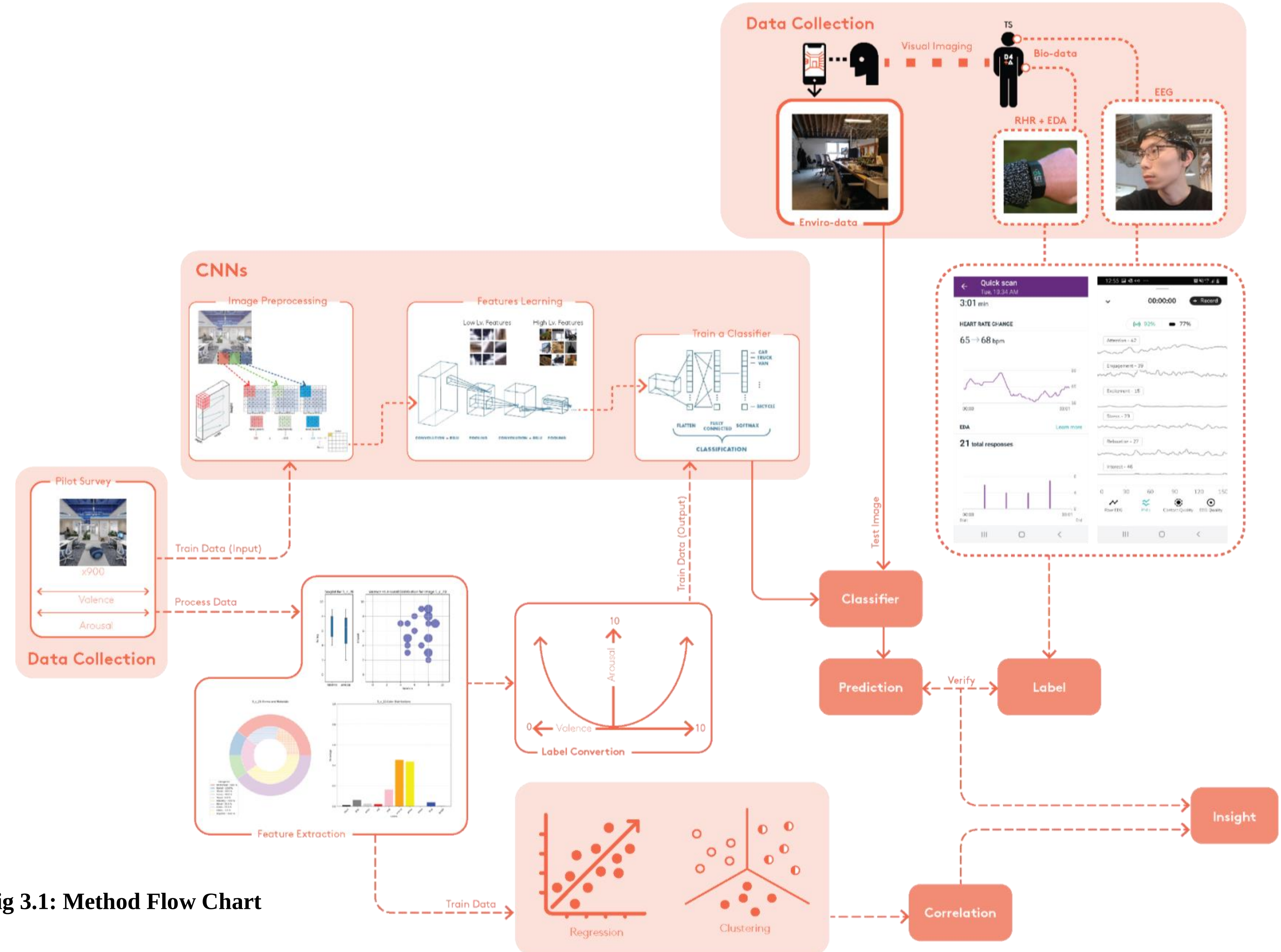


Fig 3.1: Method Flow Chart

3.2 Data Collection and Processing

3.2.1 Pilot Survey and Data Processing

In the pilot survey, this study has collected 900 images of different interior environments to collect human responses. These 900 images are distributed equally under 3 major categories: “Form”, “Material”, and “Color”. Each category contains 300 images as shown in Fig.3.2. “Form” category contains black and white images of interior spaces to minimize the appearance of material texture and color that may distract participants' rating. “Material” category contains images of interior spaces that emphasize on material textures. The “Color” category contains images that feature the use of 10 colors: “Black”, “Gray” , “White”, “Red”, “Pink”, “Orange”, “Yellow”, “Green”, “Blue”, “Purple”. Respondents are asked to imagine those images as their work environments, and rate the images accordingly in terms of valence and arousal in a 0 to 10 scale.

To train the CNNs model, this study collects survey responses and maps the valence and arousal distribution on a 2D Valence-Arousal Scale (VAS) as shown in Fig.3.3. For labels classification, the responses at the 25th percentile will be used as the benchmark compared with the labeling graph shown in Fig.3.3. For example, the interior environment in the image 1_c_12 has both valence and arousal reading at 5 as the benchmark, which is classified as “Neutral”. After that, these 900 images will be used as input dataset while their corresponding labels will be used as the output dataset.

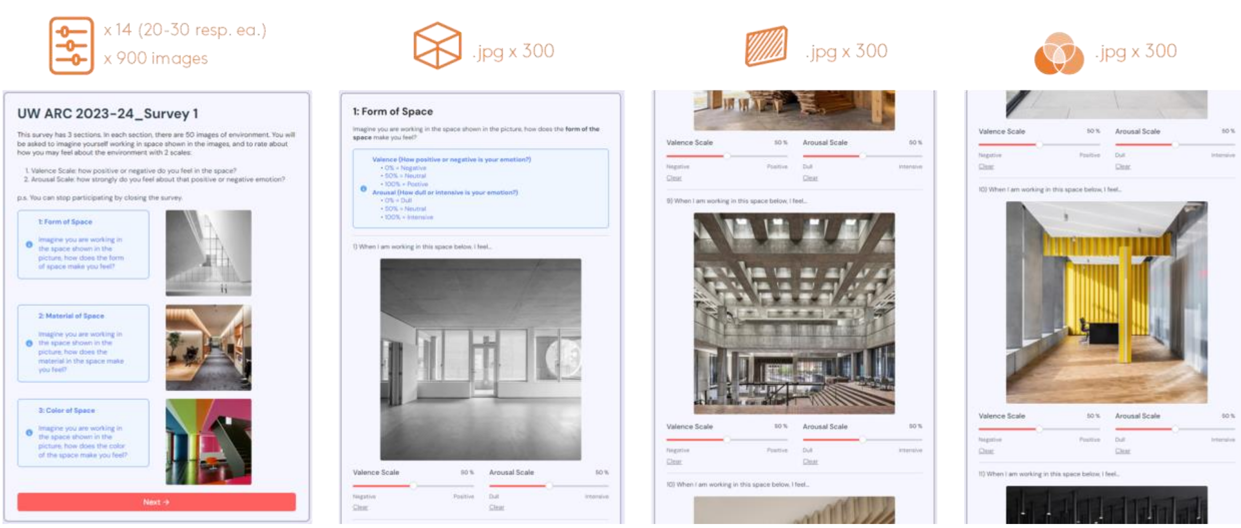


Fig.3.2: Survey Format

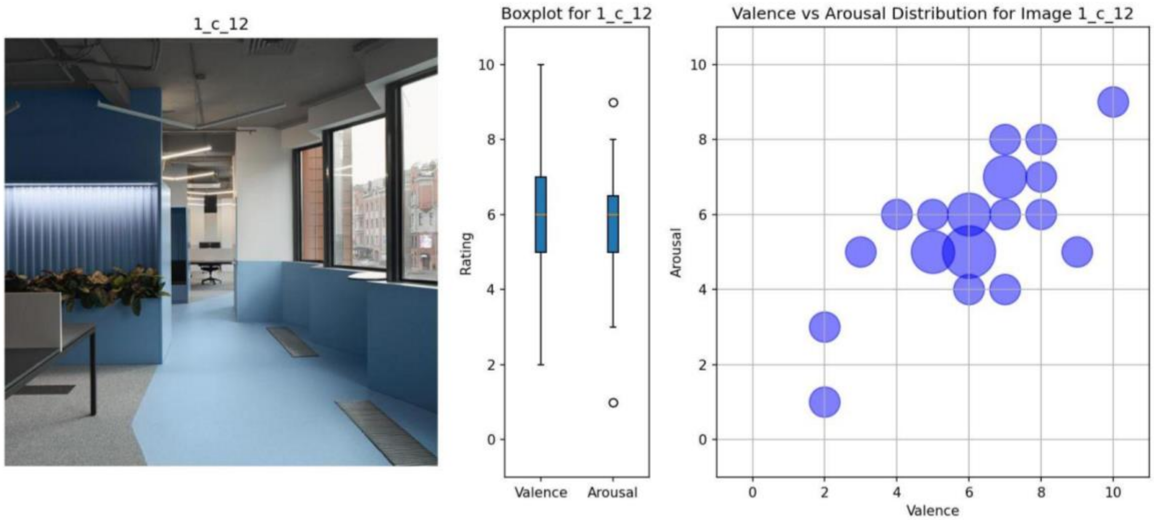


Fig.3.3: Example: Survey response data visualization

For further insight extractions, this study first analyzes each image and computes the proportion of over 20 features in 3 major categories. “Form” category contains features of “Rectilinear, Sharp, Round, Curvy” based on the spatial classification in Shemesh et al. studies (2021). This study first utilizes the image segmentation techniques to outline the spatial geometries of the captured environments as shown in Fig.3.4. Then, these images are sent to ChatGPT-4 developed by OpenAI, an AI model with advanced image analysis features, to analyze the spatial geometric proportions between “Form” features. “Material” category contains features of “Wood, Masonry, Metal, Glass, Fabric, Biophilic”. Images are also sent to ChatGPT-4 to analyze the material proportions of each environment. “Color” category contains features of “Black, Gray, White, Red, Pink, Orange, Yellow, Green, Blue, Purple”. This study performs color distribution analysis based on HSV color segmentation via “OpenCV”, a library of programming functions for computer vision, as shown in Fig.3.5.

Fig.3.4: Geometric Segmentation

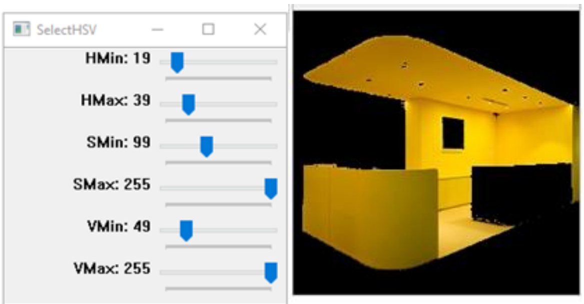


Fig.3.5: Color Segmentation

These features computations are later compiled into one dataset as shown in Fig.3.6, and further used in a variety of supervised regression analysis such as linear regression, decision tree, and random forest, etc. Additionally, these images are also used in unsupervised learning analysis using K-means clustering to help further extract correlations between design elements and human valence and arousal responses.

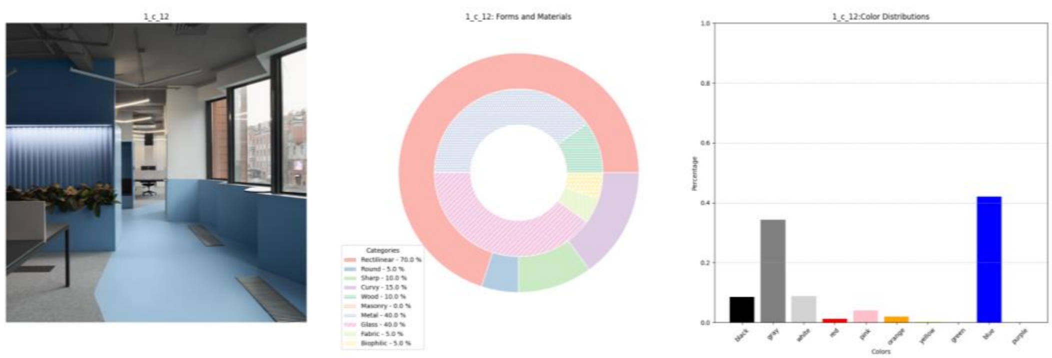


Fig.3.6: Example: Feature extraction data visualization

3.2.2 Biometric Data collection and Data Processing

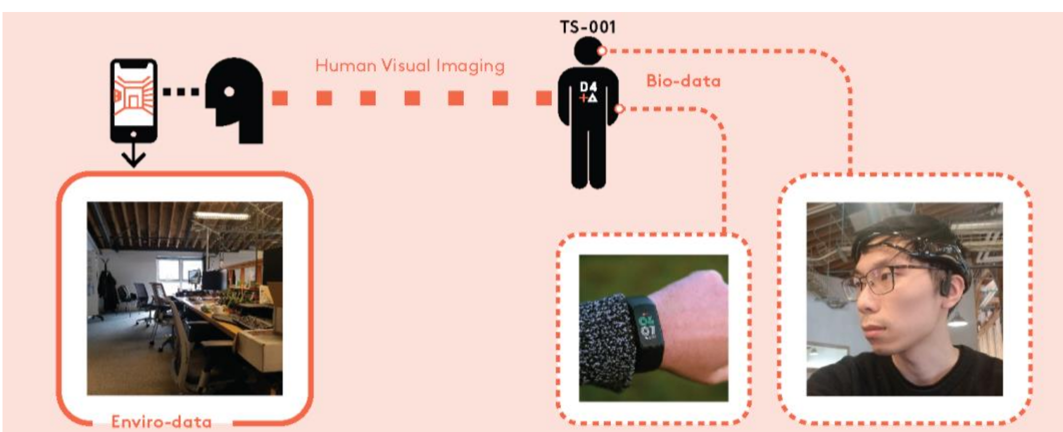


Fig.3.7: Environmental data and Biometric data collection

To collect the biometric data, this study uses a noninvasive wireless EEG headset called Insight 2.0 developed by Emotiv to collect EEG data; and uses a smartwatch called “Fitbit Charge 6” developed by Google to collect EDA and RHR data. as shown in Fig.3.8. Following the IRB protocol for human subject tests and the reliability of collected data, this study only invites people who are at least 21 years old without any health conditions (such as neurological disorders or use of psychoactive medication…etc.) that may affect physiological or cognitive functioning to be Test Subjects(TSs). After the devices’ calibration, TSs will keep wearing the devices until the whole process is over. First, they will select a location to perform daily work tasks for at least 10-15 minutes. Then, take 3-5 photos to cover the environment in their visual range. Next, they can start by recording their EDA and RHR data as instructed by the smartwatch. After 10-15 minutes they will record EDA and RHR data again before moving to a new location to continue the next data collection or ending the data collection process.

To process the EEG data, this study utilizes Emotiv analysis software “EmotivPro”to translate the raw EEG data into mental state via performance metrics as shown in Fig.3.8. This study primarily focuses on “Engagement(En)”, “Relax(Re)”, and “Stress(St)” indexes as they are associated with the label graph as shown in Fig.2.1. RHR and EDA data will be measured by a smartwatch called “Fitbit Charge 6” developed by Google. Inc as shown in Fig.3.7. After that, EEG, RHR, and EDA will be used to comprehensively compute the labels of the corresponding environments captured in the photos, and later be used to test the CNNs performance.

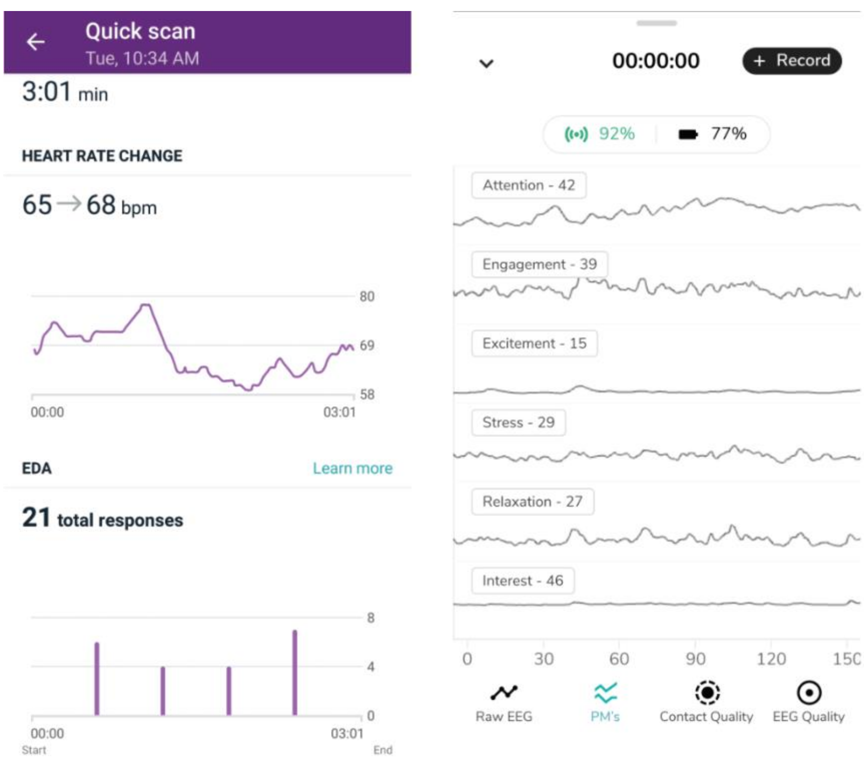


Fig.3.8: Example of EDA, RHR, EEG recording

3.3 Building CNNs model

This study uses the MobileNetV2 (Sandler et al. 2018) CNN model development begins with importing libraries including TensorFlow, numpy, pandas, and scikit-learn, along with specific modules for image processing and model building. The data is loaded from a CSV file containing image IDs and associated features, which is read using pandas. Images are preprocessed to a target size of 224x224 pixels to match the input requirements of the MobileNetV2 model. Helper functions are used to load and preprocess these images, ensuring they are compatible with the model. To address class imbalance, data augmentation techniques, including rotation, zoom, width and height shifts, shear, and horizontal flips, are applied using the "ImageDataGenerator". The MobileNetV2 model, pre-trained on the ImageNet dataset, is employed for feature extraction. Custom layers are added on top of this base, combining image features with additional CSV features. The architecture is defined by flattening the base model's output, concatenating it with CSV inputs, and passing it through dense layers with ReLU activation functions, culminating in a softmax output layer for classification. The model is compiled using the Adam optimizer and categorical cross-entropy loss function and is then trained using the augmentation dataset. The training process is validated through a separate validation set, and the model's performance is evaluated, yielding metrics such as accuracy to gauge its effectiveness in classifying images based on their impact on occupants' wellbeing.

4. Result and Discussion

4.1 CNNs Modeling Result

After processing the label classification based on the collected survey responses, the environments were categorized as follows: 501 as “Neutral,” 114 as “Relax,” 107 as “Negative,” 95 as “Bored,” 68 as “Flow,” 11 as “Positive,” and 4 as “Stress” (Fig.4.1). To address class imbalance, 150 additional images were augmented, focusing on underrepresented classes. The dataset was split into 870 training samples and approximately 180 validation samples, following an 80/20 train-test split. The model was trained over 10 epochs with a batch size of 32. Evaluation on the validation set resulted in an accuracy of 53.33%, while the training accuracy reached 99.43% (Fig.4.2), indicating overfitting, where the model fails to generalize unseen data. This underperforming issue can be caused by factors such as limited and homogeneous dataset; high model complexity; and insufficient data augmentation, given by the sample size of this research. The dataset, with only 870 training samples and approximately 180 validation samples, might not fully capture the diversity and complexity of interior environments. This limited data may lead to overfitting and hinder the model's ability to generalize to new environments. Additionally, data augmentation techniques, such as rotation, zoom, and flips, may introduce unrealistic variations that do not reflect natural changes in physical spaces, potentially affecting the model's learning process. There are several potential steps to improve the CNN models. Applying regularization techniques such as introducing dropout layers during the training process, or L2 regularization to encourage the model to maintain simpler weights, can reduce the risk of overfitting. Also, other model architectures or simplification methods can be explored to prevent it from becoming too complex and thus prone to overfitting. Besides, more data to enhance the diversity and comprehensiveness of the training set is much needed to improve its generalization capability.

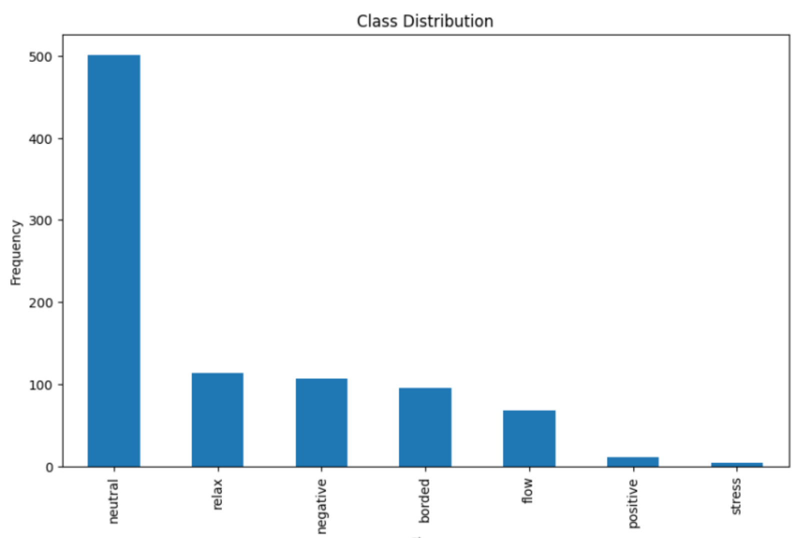


Fig.4.1: Image class distribution

```
Epoch 1/10
28/28 [==============================] - 52s 2s/step - loss: 5.9901 - accuracy: 0.3701 - val_loss: 3.0579 - val_accuracy: 0.4667
Epoch 2/10
28/28 [==============================] - 52s 2s/step - loss: 1.0233 - accuracy: 0.7678 - val_loss: 3.3154 - val_accuracy: 0.5556
Epoch 3/10
28/28 [==============================] - 69s 2s/step - loss: 0.2148 - accuracy: 0.9437 - val_loss: 3.0327 - val_accuracy: 0.5278
Epoch 4/10
28/28 [==============================] - 52s 2s/step - loss: 0.0761 - accuracy: 0.9816 - val_loss: 2.7377 - val_accuracy: 0.5222
Epoch 5/10
28/28 [==============================] - 60s 2s/step - loss: 0.0505 - accuracy: 0.9908 - val_loss: 3.2489 - val_accuracy: 0.5333
Epoch 6/10
28/28 [==============================] - 82s 3s/step - loss: 0.0398 - accuracy: 0.9885 - val_loss: 3.0866 - val_accuracy: 0.5333
Epoch 7/10
28/28 [==============================] - 65s 2s/step - loss: 0.0419 - accuracy: 0.9920 - val_loss: 3.3204 - val_accuracy: 0.5444
Epoch 8/10
28/28 [==============================] - 48s 2s/step - loss: 0.0270 - accuracy: 0.9920 - val_loss: 3.5504 - val_accuracy: 0.5444
Epoch 9/10
28/28 [==============================] - 73s 3s/step - loss: 0.0597 - accuracy: 0.9908 - val_loss: 3.0221 - val_accuracy: 0.4944
Epoch 10/10
28/28 [==============================] - 55s 2s/step - loss: 0.0175 - accuracy: 0.9943 - val_loss: 3.3228 - val_accuracy: 0.5333
```

Fig.4.2: CNN model training and evaluation

4.2 Insights from Regression Models

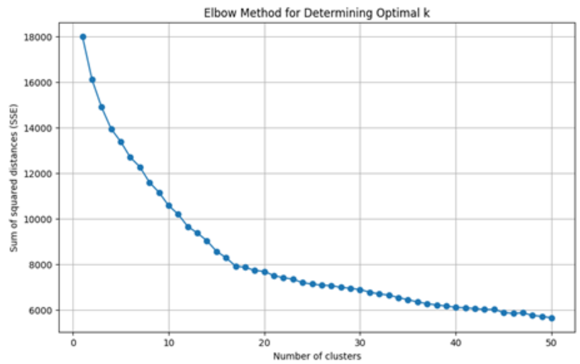
The regression model showed suboptimal performance, with almost all models (Fig.4.3) having negative R^2 scores, indicating that they perform worse than a simple mean prediction of the target variable. The Ridge Regression model is the only model with a slightly positive R^2 score but is close to zero (0.011312), indicating minimal explanatory power. The high Mean Squared Error (MSE), and Mean Absolute Error (MAE) across all models further highlights the significant prediction errors, and low R^2 suggest that the model does not capture much of the data variability and may not fully understand the underlying patterns. Other models such as Lasso Regression and ElasticNet showed similar performance to Ridge Regression but with slightly lower R^2 scores. Ensemble methods like Random Forest Regressor and Gradient Boosting Regressor provided moderate performance improvements, but their R^2 scores were still negative. CatBoost and LightGBM showed slightly better performance among the boosting algorithms, though their R^2 scores remained negative. These results highlight that significant improvements are needed. Several potential steps include data preprocessing such as noise reduction by removing irrelevant details from the image of the interior environments that might introduce noise; better feature engineering such as contextual features introducing object detection for materials and calculating their pixel proportion to the entire image; and deeper hyperparameter tuning. However, as the original data are essentially images, which the underlying structure and non-linearity issues could be inherently too difficult for regression models to learn the association between design elements and their wellbeing impacts.

	mean_squared_error	mean_absolute_error	R2_score
Ridge Regression	0.754045	0.679883	0.011312
Lasso Regression	0.768611	0.688511	-0.006191
ElasticNet	0.768611	0.688511	-0.006191
Linear Regression	0.782465	0.692989	-0.029794
Random Forest Regressor	0.792805	0.707229	-0.042428
Gradient Boosting Regressor	0.806049	0.711030	-0.060012
AdaBoost Regressor	0.819180	0.725575	-0.083287
CatBoost	0.847537	0.730277	-0.108338
LightGBM	0.867224	0.746579	-0.137587
Hist Gradient Boosting Regressor	0.897959	0.754003	-0.175506
XGBoost	0.950195	0.767330	-0.243682
Extra Tree Regressor	1.553682	0.969850	-1.045935
Decision Tree Regressor	1.680412	1.001631	-1.211216

Fig.4.3: Means squared errors of various regression models

4.3 Clustering and Principal Component Analysis

To uncover underlying structure and address non-linearity issues, K-means clustering and principal component analysis (PCA) were performed. Seventeen clusters were suggested based on the elbow method (Fig.4.4), and "Cluster 14" was selected as an example for PCA shown in the correlation matrix (Fig.4.5 and Fig.4.6). The feature extraction methods might not fully capture all relevant aspects of interior design elements. While effective for general image features, they may miss subtle, domain-specific nuances crucial for understanding the impact of interior environments on wellbeing. Advanced feature extraction techniques tailored to interior design could provide more accurate and insightful features. The analysis revealed several interesting relationships between design elements and their impacts on occupants' valence and arousal responses. For valence, rectilinear design elements showed a positive correlation (0.25), suggesting that straight lines and orderly designs are associated with positive valence. Glass elements also exhibited positive emotions. In contrast, sharp design elements had a negative correlation with valence (-0.39), implying that they are perceived negatively. Surprisingly, biophilic elements, typically associated with positive wellbeing outcomes, showed a negative correlation (-0.46) with valence in this dataset. For arousal, wood (0.18) and glass (0.22) elements were positively correlated, suggesting these materials are associated with higher arousal levels, enhancing alertness and engagement. Conversely, biophilic elements displayed a negative correlation (-0.32) with arousal, indicating a calming effect, aligning with the general understanding of biophilic design promoting relaxation.



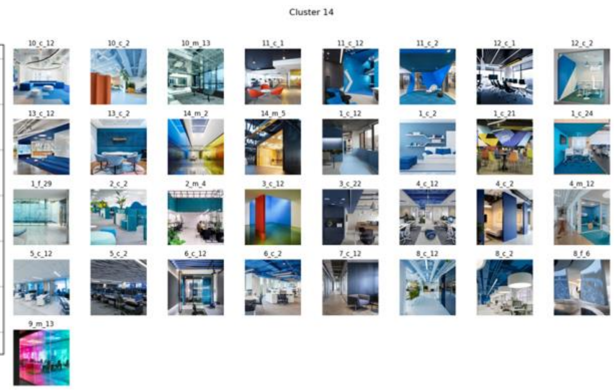


Fig.4.4: Ebow method

Fig.4.5: Cluster example

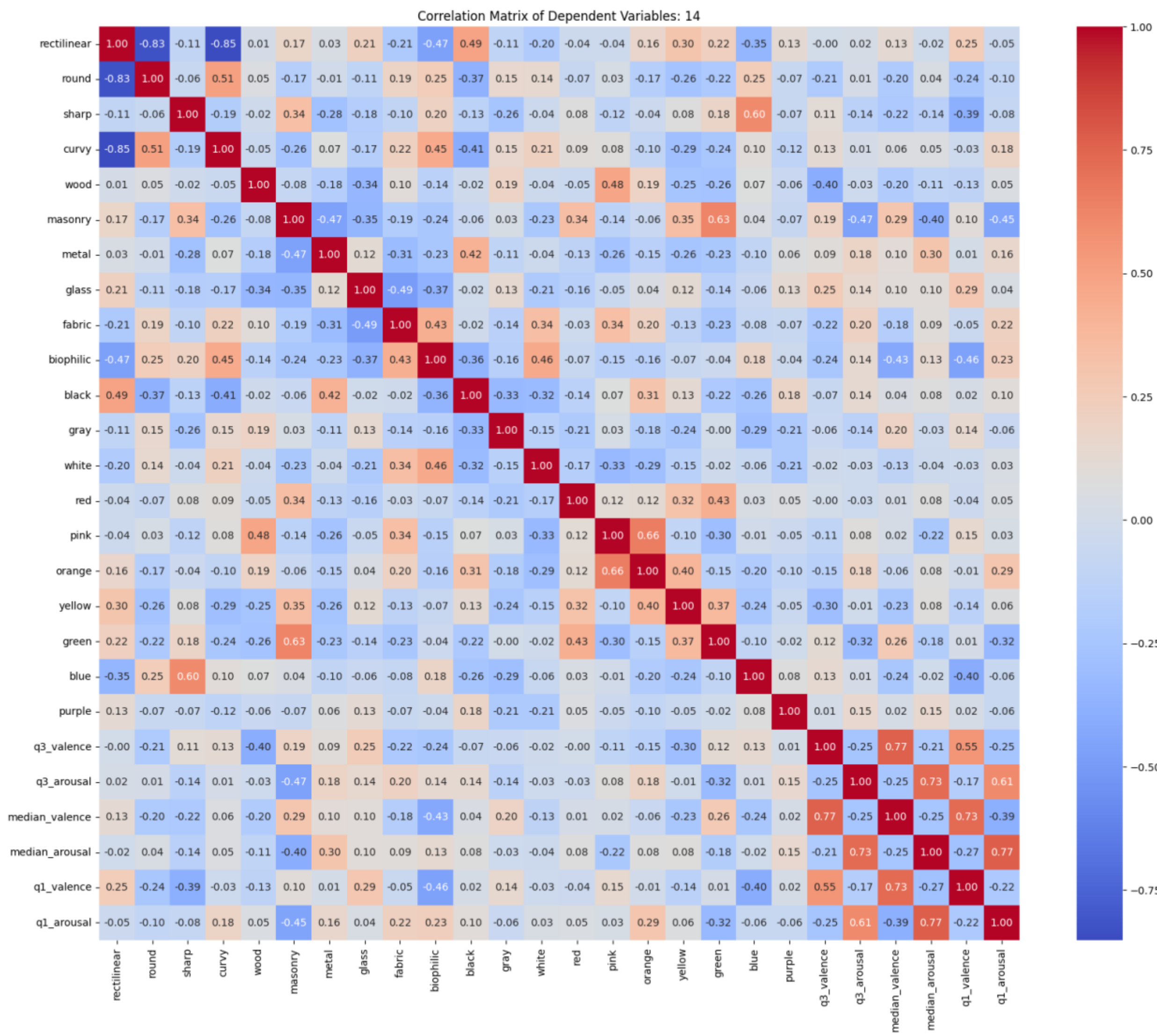


Fig.4.6: PCA analysis of cluster example

4.4 Key Findings and Their Implications

The findings indicate that certain design elements have a consistent impact on occupants' emotional responses. Glass elements are positively correlated with both valence and arousal, making them versatile for creating environments that are both engaging and emotionally positive. This suggests that incorporating glass in interior designs can enhance occupants' overall emotional wellbeing. Rectilinear elements also contribute positively to valence, suggesting that clean, straight lines are aesthetically pleasing and contribute to a positive emotional state. The negative correlation of sharp design elements with valence highlights their potential to evoke negative emotions, suggesting that softer, less aggressive design choices may be preferable for promoting positive emotions. The consistent negative correlation of biophilic elements with arousal suggests their effectiveness in creating calming, low-arousal environments. However, the unexpected negative correlation with valence warrants further investigation to understand the underlying factors that may contribute to this outcome in the specific context of this dataset. Despite promising findings, several limitations should be acknowledged to guide future research. The lack of control over survey participants' demographics and professional backgrounds may have introduced biases. Future studies should ensure a more representative and balanced participant pool.

5. Conclusion (600 words)

This study provides a comprehensive analysis of the influence of interior design elements on occupants' emotional responses through the integration of qualitative and quantitative methods. By categorizing environments based on survey responses into seven classes—“Neutral,” “Relax,” “Negative,” “Bored,” “Flow,” “Positive,” and “Stress”—it was found that a significant portion of the environments were perceived as “Neutral.” The CNN model, despite achieving a high training accuracy of 99.43%, exhibited a lower validation accuracy of 53.33%, suggesting potential overfitting and highlighting the need for a more diverse training dataset. Regression analysis revealed the non-linear nature imposes a significant challenge for regression models to effectively learn the relationship between design elements and emotional responses. K-means clustering and PCA provided further insights, such as the positive impact of rectilinear and glass elements on valence and the calming effect of biophilic elements on arousal. However, the unexpected negative correlation of biophilic elements with valence suggests a need for further investigation. These findings have practical implications for architectural design. Glass and rectilinear elements were found to enhance emotional positivity and engagement, making them valuable for creating appealing and emotionally supportive environments. Conversely, sharp design elements should be used cautiously due to their potential to evoke negative emotions, and the application of biophilic elements requires a nuanced approach to balance their dual effects on valence and arousal. Several limitations were identified, including the need for a more representative participant pool and a larger, more diverse dataset to capture the complexity of interior environments. Advanced feature extraction techniques and more sophisticated model tuning and optimization could also enhance predictive performance.

Overall, this study demonstrates the potential of mixed-methods research in understanding the impact of interior design on occupant wellbeing. The integration of CNN modeling, regression analysis, and clustering techniques provides a holistic view of how design features influence emotional responses. These insights can guide architects and designers in creating spaces that promote emotional wellbeing, ultimately contributing to healthier and more productive environments. The study's findings have significant potential to impact future design strategies, encouraging the development of environments that not only meet aesthetic and functional needs but also support the psychological and emotional health of their occupants.

6. Acknowledgement

This study was conducted under the auspices of the University of Washington's Applied Research Consortium (ARC) program, with invaluable support from our partnering firm, Mithun. We extend our heartfelt gratitude to Dr. Narjes Abbasabadi, Assistant Professor in the Department of Architecture at the University of Washington and head of the Sustainable Intelligence Lab, for her expert guidance and leadership of this research. Her extensive research in computation in the built environment has been instrumental in shaping this study. We also wish to acknowledge Jason Steiner and Katie Stege for their unwavering support and expertise, who lead this computation design research at Mithun. Their contributions have been critical in driving the ARC research and development efforts. Their collective insights and encouragement have been vital to the success of this research.

References

Taylor, Frederick W. The Principles of Scientific Management. New York: Harper & Brothers, 1911.

Priyanka A. Abhang, Bharti W. Gawali, and Suresh C. Mehrotra. "Technological Basics of EEG Recording and Operation of Apparatus." In *Introduction to EEG- and Speech-Based Emotion Recognition*, edited by Priyanka A. Abhang, Bharti W. Gawali, and Suresh C. Mehrotra, 19-50. Academic Press, 2016. https://doi.org/10.1016/B978-0-12-804490-2.00002-6.

Schmidt, Louis A. "Frontal brain electrical activity in shyness and sociability." *Psychological Science* 12, no. 4 (2001): 301-306. DOI:10.1111/1467-9280.00356.

Posner, Jonathan, James A. Russell, and Bradley S. Peterson. "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology." *Development and Psychopathology* 17, no. 3 (2005): 715-734. DOI:10.1017/S0954579405050340.

Barrero, Jose Maria, Steven J. Davis, and Nicholas Bloom. "The Evolution of Working From Home." *Journal of Economic Perspectives* 37, no. 4 (November 4, 2023): 1–6. https://doi.org/10.1257/jep.37.4.1.

Myerson, Jeremy, Philip Tidd, Catherine Greene, Imogen Privett, Gail Ramster, Andrew Thomson, Ankita Dwivedi, Namrata Krishna, and Zsuzsa Nagy. *Workplace & Wellbeing: Developing a Practical Framework for Workplace Design to Affect Employee Wellbeing*. Gensler, Helen Hamlyn Centre for Design at the Royal College of Art, 2017.

Aryanti, Ratih Devi, Erita Yuliasesti Diah Sari, and Herlina Siwi Widiana. "A Literature Review of Workplace Well-Being." In *Proceedings of the*

International Conference on Community Development (ICCD 2020), 477:605-609. Atlantis Press, 2020. https://doi.org/10.2991/aebmr.k.200312.085.

Byun, Kunjoon, Sara Aristizabal, Yihan Wu, Aidan F. Mullan, Jeremiah D. Carlin, Colin P. West, and Kevin A. Mazurek. "Investigating How Auditory and Visual Stimuli Promote Recovery After Stress With Potential Applications for Workplace Stress and Burnout: Protocol for a Randomized Trial." *Frontiers in Psychology* 13 (2022): 897241. https://doi.org/10.3389/fpsyg.2022.897241.

Kim, Jeongmin, and Nayeon Kim. "Quantifying Emotions in Architectural Environments Using Biometrics." *Applied Sciences* 12, no. 9998 (2022): 1-15.

Mostafavi, A. "Architecture, Biometrics, and Virtual Environments Triangulation: A Research Review." *Architectural Science Review* 65, no. 6 (2022): 504-521.

Persiani, Sandra G. L., Bilge Kobas, Sebastian Clark Koth, and Thomas Auer. "Biometric Data as Real-Time Measure of Physiological Reactions to Environmental Stimuli in the Built Environment." *Energies* 14, no. 232 (2021): 1-40.

Azzazy, Sameh, Amirhosein Ghaffarianhoseini, Ali GhaffarianHoseini, Nicola Naismith, and Zohreh Doborjeh. "A critical review on the impact of built environment on users' measured brain activity." *Architectural Science Review* (2021).

Costin, Alina, Alina Felicia Roman, and Raluca-Stefania Balica. "Remote Work Burnout, Professional Job Stress, and Employee Emotional Exhaustion during the COVID-19 Pandemic." *Frontiers in Psychology* 14 (June 1, 2023): 1193854. https://doi.org/10.3389/fpsyg.2023.1193854.

Arza, A., J. M. Garzon, A. Hemando, J. Aguilo, and R. Bailon. "Towards an Objective Measurement of Emotional Stress: Preliminary Analysis Based on Heart Rate Variability." In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 3331–34. Milan: IEEE, 2015. https://doi.org/10.1109/EMBC.2015.7319105.

Alzubaidi, Laith, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions." *Journal of Big Data* 8, no. 1 (March 31, 2021): 53. https://doi.org/10.1186/s40537-021-00444-8.

Ge, Fenfen, Minlan Yuan, Ying Li, and Wei Zhang. "Posttraumatic Stress Disorder and Alterations in Resting Heart Rate Variability: A Systematic Review and Meta-Analysis." *Psychiatry Investigation* 17, no. 1 (January 25, 2020): 9–20. https://doi.org/10.30773/pi.2019.0112.

Affanni, Antonio. "Wireless Sensors System for Stress Detection by Means of ECG and EDA Acquisition." *Sensors* 20, no. 7 (April 4, 2020): 2026. https://doi.org/10.3390/s20072026.

Medhat Assem, Hala, Laila Mohamed Khodeir, and Fatma Fathy. "Designing for Human Wellbeing: The Integration of Neuroarchitecture in Design – A Systematic Review." *Ain Shams Engineering Journal* 14, no. 6 (June 2023): 102102. https://doi.org/10.1016/j.asej.2022.102102.

Yin, Jie, Nastaran Arfaei, Piers MacNaughton, Paul J. Catalano, Joseph G. Allen, and John D. Spengler. "Effects of Biophilic Interventions in Office on Stress Reaction and Cognitive Function: A Randomized Crossover Study in Virtual Reality." *Indoor Air* 29, no. 6 (November 2019): 1028–39. https://doi.org/10.1111/ina.12593.

Yin, Jie, Jing Yuan, Nastaran Arfaei, Paul J. Catalano, Joseph G. Allen, and John D. Spengler. "Effects of Biophilic Indoor Environment on Stress and Anxiety Recovery: A between-Subjects Experiment in Virtual Reality." *Environment International* 136 (March 2020): 105427. https://doi.org/10.1016/j.envint.2019.105427.

Yoon, So-Yeon, and Kevin Wise. "Reading Emotion of Color Environments: Computer Simulations with Self-Reports and Physiological Signals." In *Industrial Applications of Affective Engineering*, edited by Junzo Watada, Hisao Shiizuka, Kun-Pyo Lee, Tsuyoshi Otani, and Chee-Peng Lim, 219–32. Cham: Springer International Publishing, 2014. https://doi.org/10.1007/978-3-319-04798-0_17.

Cha, Seung Hyun, Shaojie Zhang, and Tae Wan Kim. "Effects of Interior Color Schemes on Emotion, Task Performance, and Heart Rate in Immersive Virtual Environments." *Journal of Interior Design* 45, no. 4 (December 2020): 51–65. https://doi.org/10.1111/joid.12179.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep Learning." *Nature* 521, no. 7553 (May 28, 2015): 436–44. https://doi.org/10.1038/nature14539.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." *Communications of the ACM* 60, no. 6 (May 24, 2017): 84–90. https://doi.org/10.1145/3065386.

Simonyan, Karen, and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv, April 10, 2015. http://arxiv.org/abs/1409.1556.

Chalmers, Taryn, Blake Anthony Hickey, Phillip Newton, Chin-Teng Lin, David Sibbritt, Craig S. McLachlan, Roderick Clifton-Bligh, John Morley, and Sara Lal. "Stress Watch: The Use of Heart Rate and Heart Rate Variability to Detect Stress: A Pilot Study Using Smart Watch Wearables." *Sensors* 22, no. 1 (December 27, 2021): 151. https://doi.org/10.3390/s22010151.

Altini, Marco, and Daniel Plews. "What Is behind Changes in Resting Heart Rate and Heart Rate Variability? A Large-Scale Analysis of Longitudinal Measurements Acquired in Free-Living." *Sensors* 21, no. 23 (November 27, 2021): 7932. https://doi.org/10.3390/s21237932.

Sandler, Mark, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. "MobileNetV2: Inverted Residuals and Linear Bottlenecks." arXiv, March 21, 2019. http://arxiv.org/abs/1801.04381.