

Research Article

Recognition of Design Fixation via Body Language Using Computer Vision

Zhongliang Yang,¹ Yumiao Chen ,² and Song Zhang³

¹College of Mechanical Engineering, Donghua University, Shanghai 201620, China

²School of Art, Design and Media, East China University of Science and Technology, Shanghai 200237, China

³Department of Materials, The University of Manchester, Manchester M13 9PL, UK

Correspondence should be addressed to Yumiao Chen; 181603042@qq.com

Received 21 December 2020; Revised 19 June 2021; Accepted 12 August 2021; Published 27 August 2021

Academic Editor: Kai Zhang

Copyright © 2021 Zhongliang Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The main objective of this study is to recognize design fixation accurately and effectively. First, we conducted an experiment to record the videos of design process and design sketches from 12 designers for 15 minutes. Then, we executed a video analysis of body language in designers, correlating body language to the presence of design fixation, as judged by a panel of six experts. We found that three body language types were significantly correlated to fixation. A two-step hybrid recognition model of design fixation based on body language was proposed. The first-step recognition model of body language using transfer learning based on a pretrained VGG-16 convolutional neural network was constructed. The average recognition rate achieved by the VGG-16 model was 92.03%. Then, the frames of recognized body language were used as input vectors to the second-step fixation classification model based on support vector machine (SVM). The average recognition rate for the fixation state achieved by the SVM model was 79.11%. The impact of the work could be that the fixation can be detected not only by the sketch outcomes but also by monitoring the movements, expressions, and gestures of designers, as it is happening by monitoring the movements, expressions, and gestures of designers.

1. Introduction

Administrator, professionals, teaching staff, and statesmen have indicated that making innovations is the key to our future [1]. Innovative products can make headlines and generate significant economic returns [2]. The focus of innovation is always on generating new ideas of products and/or services by designers [3]. More and more scholars have adopted an innovative method called “design thinking” [4]. However, design fixation is an inevitable phenomenon, which will have a negative impact on the design results, especially in the concept stage of design process [5].

The appearance of fixed design will make the designers tend to focus on the features of existing designs and then produce new designs similar to the previous example, which will lead to the reduction of novelty and diversity of ideas

[6, 7]. In recent years, research on design fixation has attracted more and more attention, especially in the field of design. Researchers around the world are studying the causes, influencing factors, effects, and indicators of design fixation [5, 8]. Some methods have been developed to combat design fixation [5]. Some studies suggest that the representation of triggers can effectively reduce design fixation and improve creativity [5, 7–10]. Tseng et al. [11] found that the effectiveness of trigger in design depends on the timing of when the inspiring information is given. Design fixation is hard to combat because it occurs unconsciously [7]. Designers, even those who study and teach design regularly, do not know when they are being influenced or fixated by inadequate or misleading information [8]. Existing literature provides very limited insights regarding how to recognize fixation effectively. Based on these concerns, this study will focus on the recognition of design

fixation based on the body language during the ideation stage of a design process.

This paper proposes a body-language-based method for the recognition of design fixation from body movements and presents a controlled experiment with designers, recording the videos of the design process and design sketches separately for 15 min. We selected eight common types of body language during the design processes according to the recorded videos. Subsequently, three types of body language with significant differences between fixation and no fixation were found. We converted videos to images and used them as input vectors to a body language recognition model based on the VGG-16 convolutional neural network. Besides, the frames of three types of body language during each sketch were extracted and used as input vectors to a design fixation classification model based on a support vector machine (SVM).

2. Background/Previous Work

2.1. Design Fixation. The term “fixated” first appeared in the experimental psychology literature, referring to a person’s subconscious mind that focuses on a particular aspect of an object or task and ignores other aspects [12]. Definitions of fixation differ with the context of human activity, field of knowledge, or design objectives [5]. In 1991, Jansson and Smith first proposed the concept of “design fixation” [6, 13], which expanded the fixation research from psychology to design. The design fixation was described as “the blind, sometimes counterproductive, adherence to a limited set of ideas in the design process.” The novelty and variety of ideas will reduce the adherence to a limited set of solutions. Previous research has shown that introducing graphical examples of existing solutions limits a designer’s ability to generate new solutions [14]. Jansson and Smith [13] showed designers an example solution to a problem and found that this reduced the designers’ propensity to move effectively between the conceptual space (of abstract ideas) and the configuration space (of potential solutions).

The interdisciplinary interest in design fixation from designers, cognitive scientists, engineers, computational modelers, architects, educators, and others around the emerging field of design science has resulted in increasingly broad definitions of design fixation [15, 16]. In a narrow sense, design fixation can be considered as designers’ overreliance on the functions and features in some cases [16], or their tendency to change the new design scheme to cohere with the familiar design paradigm [17]. In a broad sense, design fixation can refer to any cognitive intervention that influences design results [18] or any method that can influence design activities [19]. The broader definition holds that design fixation is a manifestation of the low level of creativity [20]. However, what is entirely consistent is that fixation is framed as an unfavorable phenomenon, with most of the studies presenting ways to avoid, mitigate, or overcome it [12].

Some methods have been developed to combat design fixation. Youmans and Arciszewski [10, 16] describe that the increasingly broad definitions of the phenomenon

might be undermining empirical research efforts, educational efforts to minimize fixation, and the acquisition and dissemination of transdisciplinary knowledge about fixation effects. To address these issues, they categorized the fixation phenomena into one of six classifications. Next, they proposed a system of orders of design fixation and recommended targeted methods for reducing fixation in inventive design. Moreno et al. [5] presented a review of defixation approaches and the metrics employed to understand and account for design fixation. The study then explored the relevant ideation approach of design-by-analogy (DbA) to overcome the design’s fixation. Sio et al. [21] conducted a meta-analytical review of design studies examining whether and under what conditions the presence of examples will induce fixation or inspiration. Crilly [3] pointed out that recognizing fixation episodes and reflecting on them was described as the means by which designers could guard against such episodes in the future and thus be more creative. In 2017, Crilly and Cardoso [22] held an international workshop, and nine questions were outlined to stimulate renewed thinking about fixation and promote debate about where research should head next. In 2019, Crilly [23] promoted methodological diversity and theoretical integration for design fixation research.

2.2. Body Language in Design. Behaviorism is primarily concerned with observable behavior, as opposed to internal events like thinking and emotion. Design fixation is also a mental state that is very difficult to identify by any means. We think that the internal design fixation state can be mapped through external body language.

Body language, as a form of nonverbal cues used in human–human interactions almost entirely subconsciously, accounts for more than half of human communication [24–26]. There are a vast number of different information channels for body language, for example, facial expressions, gestures, body movements, eye gaze, head movement, posture, distance, and so on [27–29].

The probably two most relevant aspects of body language for design research are the ability to enact physical concepts and ideas and communicate emotion [29, 30]. Some scholars have paid attention to design thinking research based on body language. Bezawada et al. [31] proposed a machine learning model of automatic facial feature extraction for predicting designers’ comfort with engineering equipment during prototype creation. Through a brief overview of existing work on the role of body language in engineering design and common tools for body language analysis not limited to engineering design, Wulvik et al. [29] called for further study of body language in an engineering design context using automatic data gathering tools. Cash and Anja [30] explored the many roles of gesture in the communication of design concepts through observing and video-coding four teams of engineering graduates during an ideation session. Sun et al. [32] reported an experiment that recorded the participants’ eye movements to analyze their perception and examine whether the designers’ perception

during sketching is in accord with the creative segment theory.

Although previous research has utilized expensive and intrusive devices, such as physiological sensors, the automatic image recognition of gesture has become the most common because of the cost efficiency and noninvasiveness of cameras. Zhao [24] combined image information from the body language of the head to recognize the emotional and cognitive states using soft computing techniques. Behoora et al. [33] proposed a machine learning recognition method of emotional states of individual design team members based on capturing multiple skeletal joint images of body language. The methodology uses the link between body language and emotions to detect emotional states with accuracies above 98%. Recently, the development of deep learning has greatly improved the accuracy of image-based gesture recognition. Yun et al. [34] proposed data-driven convolutional neural networks (CNNs) based on the engagement recognition method that uses only facial images from input videos.

Crilly and Cardoso [22] called for more objective ways to capture design activities during fixation experiments, leading us to the hypothesis that design fixation might be correlated with gesture. This manuscript focused on detecting design fixation states from the body language using computer vision techniques and transfer learning techniques. Python and Keras are used to implement the proposed methods and evaluate the results.

2.3. Summary and Breaking New Ground. CNN was first proposed by Lecun et al. [35] and was applied to handwriting recognition. It is a variant of multilayer perception (MLP), which has been widely used in image recognition [36, 37]. CNN can extract features through image input directly and avoid information loss caused by traditional manual feature extraction methods [38].

It is clear from this literature review that experiments have been deployed for studying fixation effects, underlying causes, and mitigating strategies for alleviating fixation within designers. We build on these studies and the associated results in this paper. Although much progress has been made, the much fertile ground has yet to be explored, especially in the domain of automatic recognition of design fixation. Our body language-based method addresses these limitations, and we propose a two-step hybrid model for design fixation classification on the foundation of the previous research and advanced convolutional neural network (CNN).

The main objective of this study is to use multiple subtle behavioral cues as an indicator of design fixation state to construct the relationship between the typical body language during design process and the fixation states from sketch results. We conducted an experiment where we recorded a video of participants generating ideas for a very simple design problem. Using expert judgment, we identified the ideas that were resulting from design fixation. We found that when the participants fixate, they use certain body gestures; these body gestures show a significant relationship with design fixation. Further, we try to develop a CNN-based computational model to recognize design fixation in future design tasks.

3. Materials and Methods

In this study, the two-step hybrid model was used to recognize design fixation. Before the recognition model, we need to find out the types of body language with significant differences between fixation and nonfixation. The model was applied to the case of shape design of a mug. The conceptual framework of the study is shown in Figure 1.

3.1. Participants. Twelve senior undergraduate students (mean = 21; SD = 1.3; range: 19–24 years) from the College of Mechanical Engineering at Donghua University volunteered to participate in the experiment. There were four male participants and eight female participants. All the undergraduate students were recruited from the senior design classes. All subjects are majoring in industrial design and have basic design skills and experience in product design. The participants were offered payment as compensation at the end of the experiment. They were instructed not to discuss any aspects of the experiment with their classmates to avoid bias.

3.2. Equipment and Materials. This experiment was carried out in the industrial design laboratory, and the experimental environment is shown in Figure 2. During the experiment, to make subjects concentrate on the design task, it is required to keep the room quiet and keep no idle people in or out of the room. In Figure 2, “a” represents the camera for recording the video of sketches, “b” represents the camera for recording the video of body language during each design process, and “c” represents the green background.

3.2.1. Design Task. The purpose of this experiment is to extract the body language of the subjects during the design process, rather than to test design ability, so the difficulty of the design task should not be too high. To ensure the emergence of design fixation, we chose the common products as the case. To ensure the variety of shape, we selected the simple products as the case. Thus, we chose “mug” as the case of the design task to ensure the emergence of design fixation and the variety of shape design.

All subjects were asked to deal with the same design tasks individually. The design task given in this study is displayed in Figure 3.

3.2.2. High-Definition Camera (HD Camera). Two cameras (Logitech C270, Suzhou Logitech electronics Ltd., China) were used to collect the videos of the design process and sketches separately. Camera a (Figure 2) for recording sketch video was fixed directly above the desktop, and the coverage can be shown between the green dotted lines in Figure 2. Camera b (Figure 2) for recording body language during the design process was placed in front of the subjects, and the coverage can be shown between the purple dotted lines.

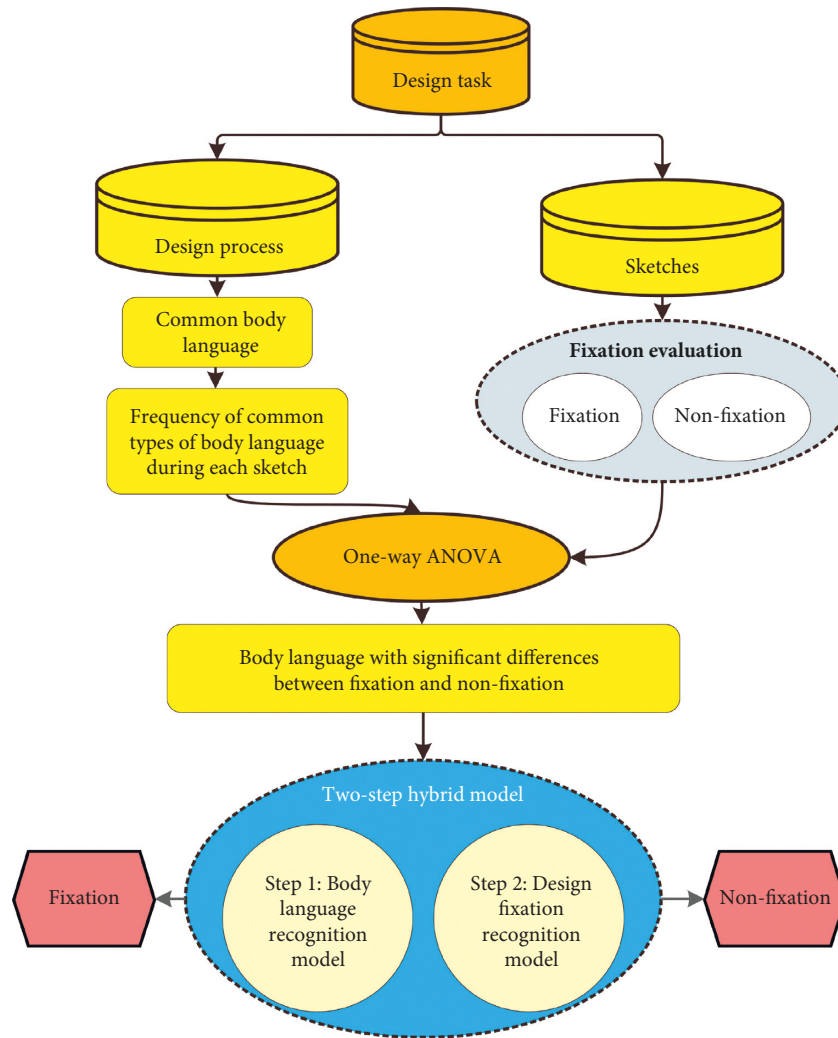


FIGURE 1: The conceptual framework of the study.

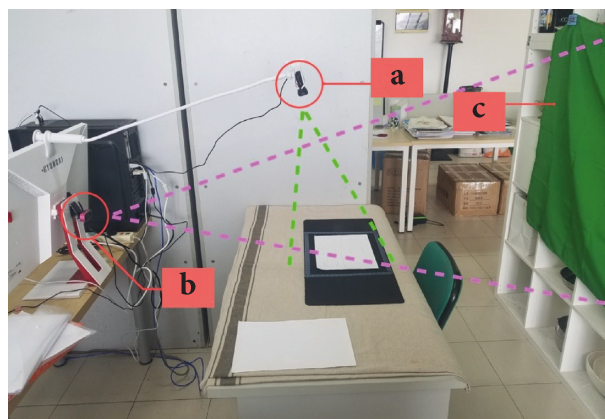


FIGURE 2: Experimental environment (a, camera a; b, camera b; c, green background).

3.2.3. Green Background. To allow easy recognition of the body language, the video background was set as green (c in Figure 2). The green screen background was made from a green fabric (100% polyester). During video acquisition, the subject is always within the range of the green screen.


3.2.4. Open Broadcaster Software (OBS). The videos were recorded by Open Broadcaster Software (OBS; version: 21.0.1). The software is shown in Figure 4. Figure 4(a) shows the picture taken by camera a. Figure 4(b) shows the picture taken by camera b. We can manually adjust focusing

Design task-Mug

You are hired to design the novelty shapes of a mug, which are not fixated upon the most common mug (right figure). This task stressed innovation and variety.

For the task, 2 min were allotted to the task understanding. 15 min were dedicated to the execution of the conceptual design in the form of sketches.

Due to the limited time, you do not need to consider material or details on the sketch, just enough to express the design concept.



The most common mug

FIGURE 3: Design task for the experiment.

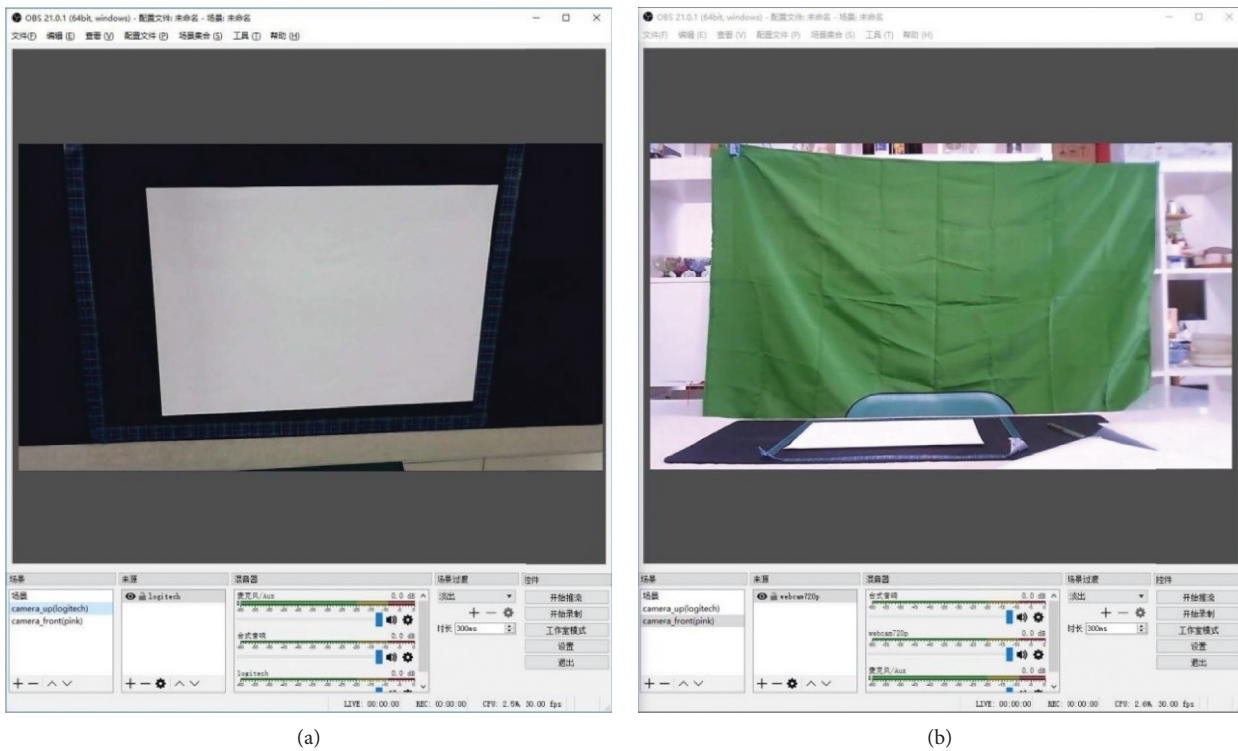


FIGURE 4: The interface of OBS.

distance, exposure, white balance, color, and other configuration of the camera to make sure video works well. Moreover, H.264 high-definition video encoding is used for transcoding storage to ensure the clarity of the video.

3.3. Experimental Protocol. The experiments were carried out for 12 days in February in 2018. They were performed from 10:00 to 11:00 every day for each subject randomly. The experiment was divided into four stages: welcome stage, preparation stage, task stage, and final stage.

During the welcome stage, the design task, procedures, and equipment used in the experiment were introduced to the participants. The design task instructions were read to

the participants, and they were given a brief tutorial on how to complete the task. All participants signed a consent form and completed a questionnaire asking about age, height, and weight.

During the preparation stage, the subject was asked to take a seat as shown in Figure 2. A rest of 3 min was used to make him/her relaxed and familiar with the design task and environment. The position of desk, chair, pen, and paper was adjusted to suit the subject. For the task stage, we prepared enough paper and pen for each subject. Then, the position and angle of the two cameras were adjusted properly to ensure the coverage of the subject matter (Figure 5).

During the task stage, after the subject fully understood the experimental task, he/she was required to complete the



FIGURE 5: Sample images from the cameras. (a) One sample image from the camera a; (b) one sample image from camera b.

design task for 15 min. Meanwhile, the experimenter started the OBS to record the videos from two cameras synchronously. When the remaining time was 5 min, the subject was given a timely reminder.

During the final stage, after 15 min, the experimenter saved the collected video data. Then, the sketches drawn by the subject were numbered and archived in order.

3.4. Manual Video Coding. Studies of the role of body language in engineering design heavily rely on manual video coding. Most of them are focused on which role gesture plays in design activities [29, 30, 39]. In this study, after video data collection, we sorted out the body language in the design activities through manual video coding. Body language mainly can be divided into four categories: gesture, facial expression, eye contact, and posture [33, 40, 41]. Thus, during the behavioral analysis of the design process, we paid attention to the movements and gestures of hands, eyebrows, mouth, and head.

During the manual coding stage, to avoid observer bias, the reliability analysis shall be preceded through comparing observations by different coders [42]. The video recordings were segmented according to the participant's behaviors. A behavior starts when the coder detects its appearance, and it ends when a new behavior is detected. All videos were coded by one independent coder who was a postgraduate student majoring in industrial design. A second independent coder analyzed 25% of the data, for coding validation. Each recording was time consuming and persisted for approximately 1–2 h for each coder to code. When the results of the two independent coders show reliability more than 85%, credibility across the data set can be ensured [42].

Note that when two behaviors overlap, we will end the recording of the previous behavior and start to record the later one. If there are two consecutive drawing behaviors, we will segment them by the simultaneous sketch video recorded by camera a. A drawing behavior starts when the participant starts drawing a sketch, and it ends after the sketch is completed.

3.5. Selection of the Body Language Related to Design Fixation

3.5.1. Fixation Evaluation of Design Sketches

(1) Evaluation Methods. Subjective evaluation is a common evaluation method in psychological experiments, which has the advantages of low cost, simple procedure, and being noninvasive [43]. In this study, six professional staff members from industrial design companies were selected to compose an expert panel. They subjectively evaluated whether the sketches were fixated or not through professional judgments.

Before the experiments, the experts were trained to look for the attributes or metrics to identify fixation. The sketches with lower variety, novelty, and originality indicate fixated designs [5, 44, 45]. We used a picture of the most common mug as a stimulus (Figure 3) to fixate upon. The experts were trained to find out the sketches fixated on the stimulus.

(2) Evaluation Process. The evaluation experiments were carried out for 12 days in March in 2018. They were performed from 09:00 to 12:00 every day for each subject's sketches randomly. Every day, six experts take turns to vote for the fixated sketches of one subject produced during the design fixation to ensure that the evaluations are conducted independently and not influenced by others. The evaluation experiment can be divided into four stages: welcome stage, preparation stage, task stage, and final stage.

During the welcome stage, the experimenter provided an expert with enough sticky notes. An expert was invited to determine which design sketches were produced in the case of design fixation and put a note on the fixated sketches.

During the preparation stage, the experimenter attached one subject's sketches to the wall in chronological order. In one day, the sketches of one randomly selected subject were posted.

During the task stage, after the expert fully understood the evaluation process, he/she was required to complete the evaluation in 30 min.

During the final stage, the experimenter recorded and managed the voting results for each sketch in chronological order. After the 12-day experiment, the experts were thanked and offered payment as compensation.

3.5.2. Selecting the Body Language Significantly Affected by Fixation. After the fixation evaluation of design sketches, the sketches voted by more than 50% of experts (4–6 votes) were classified into the fixation set, and the remaining sketches were classified into the nonfixation set. Then, the videos of the design process were checked again to record the frequency of common types of body language during each sketch. To evaluate the significance of differences of two fixation states (fixation and nonfixation) in terms of frequency of common types of body language, one-way ANOVA was conducted in SPSS 20, and a statistical significance was accepted when $P < 0.05$.

3.5.3. Inter-Rater Agreement Statistic. Multiple experts rated the sketches of the participants for design fixation. An inter-rater agreement statistic was conducted to judge the extent of agreement between the experts who rated the data.

We calculated the percentage agreement for design fixation. This measure indicates how often raters who rated the fixation item on the same sketches choose the same response category. We considered the highest number of similar ratings per sketch as agreement, and the other ratings as nonagreement. The percentage agreement was calculated by dividing the number of ratings with agreement on all sketches by the total number of ratings on all sketches for which that measurement property (fixation state) was assessed. A percentage agreement $>80\%$ was considered appropriate (arbitrarily chosen).

In addition, we used kappa coefficients to calculate the reliability of the items. This is a measure that indicates how well sketches can be distinguished from each other based on the given item score. Dichotomous items were analyzed using intraclass kappa coefficients [46]; fixation was given a score of one and nonfixation, zero.

3.6. Two-Step Hybrid Model

3.6.1. Schematic Procedure of the Hybrid Model. The schematic procedure of the two-step hybrid model is illustrated in Figure 6. The deep convolution neural network (CNN) was used for the construction of body language recognition model based on the images converted from videos of camera b. The support vector machine (SVM) was used for the construction of design fixation classification model based on the behavior statistics that is calculating the frequency of recognized body language during each sketch.

The tuning parameters included learning rate, activation function, normalization, and data segmentation ratio. A single variable was firstly changed to obtain a range of its best influences. Then, a balanced optimal solution was obtained according to the coordination relationship between the variables above.

3.6.2. Step 1: Body Language Recognition Model

(1) Video Preprocessing. The data for the recognition model construction were design behavior videos collected from camera b. To improve the robustness of the recognition model, the original data need to be preprocessed.

At first, data cleansing is needed, because some videos are not qualified as training data due to the poor shooting angle and the subjects' failure to conduct experiments as required. To improve the recognition accuracy, these videos were excluded.

Then, three body language types that were significantly correlated to fixation were set as the target vectors. Besides, we set an extra category "others," to contain other remaining body languages.

After data cleansing, Adobe Premiere CC 2017 (Figure 7) was used to cut out the video segments containing the selected types of body language with significant differences between fixation and nonfixation and rename them manually for labeling. The remaining video segments were marked as "others." Finally, these marked video segments were sorted and classified according to the labeling.

The original video was collected at a rate of 30 frames per second. However, the design behaviors are always low-speed, with large repetition of adjacent frames. The video contains a large number of redundant frames. To reduce redundant data and accelerate the training process, the video needs to be resampled with a low frame rate. Python image processing libraries PIL and Imageio were used to resample the video segments at a rate of 8 frames/second, and then, the data were saved as JPG images. Each 1-minute video was converted to 480 images in terms of time series, and these images were saved.

(2) Software and Hardware. The construction of this model was based on Python language and Keras. Keras [47] is a high-level neural network application programming interface (API) developed by Google artificial intelligence expert, Francois Chollet. Keras support convolutional neural network (CNN), and GPU can be used to accelerate model training. Keras is a good choice, especially for beginners since it can use either TensorFlow or Theano as a backend, and it provides a simpler model for development [48].

Our body language recognition model was based on CNN. The training process requires a computer with high performance.

(3) Transfer Learning Based on VGG-16. At first, according to holdout cross-validation, the dataset was divided into three parts, which are, respectively, used for training, validation, and testing. Data augmentation was used to expand the data volume of the training set. Then, based on the pretrained VGG-16 model, a deep neural network was constructed for transfer learning. The model was trained with the data of the training set, and the training effect was tested with the data of the validation set. Finally, the data of the test set was used to simulate the real environment for testing the performance of the model.

During the holdout cross-validation [49], after resampling, the image data were disordered and randomly divided into three subsets in 60%, 20%, and 20% proportions, which were, respectively, used for training, validation, and testing to ensure the effective performance evaluation of the recognition model.

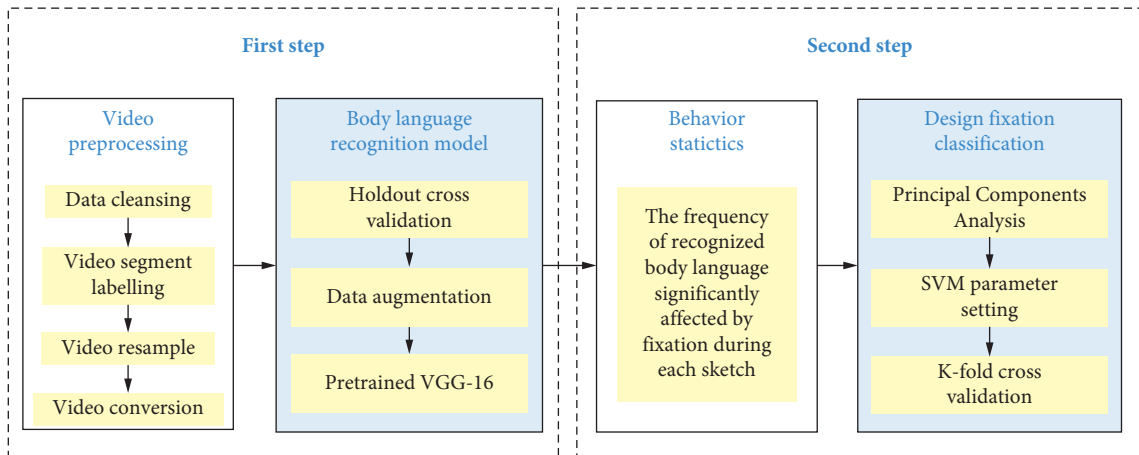


FIGURE 6: The schematic procedure of the two-step hybrid model.

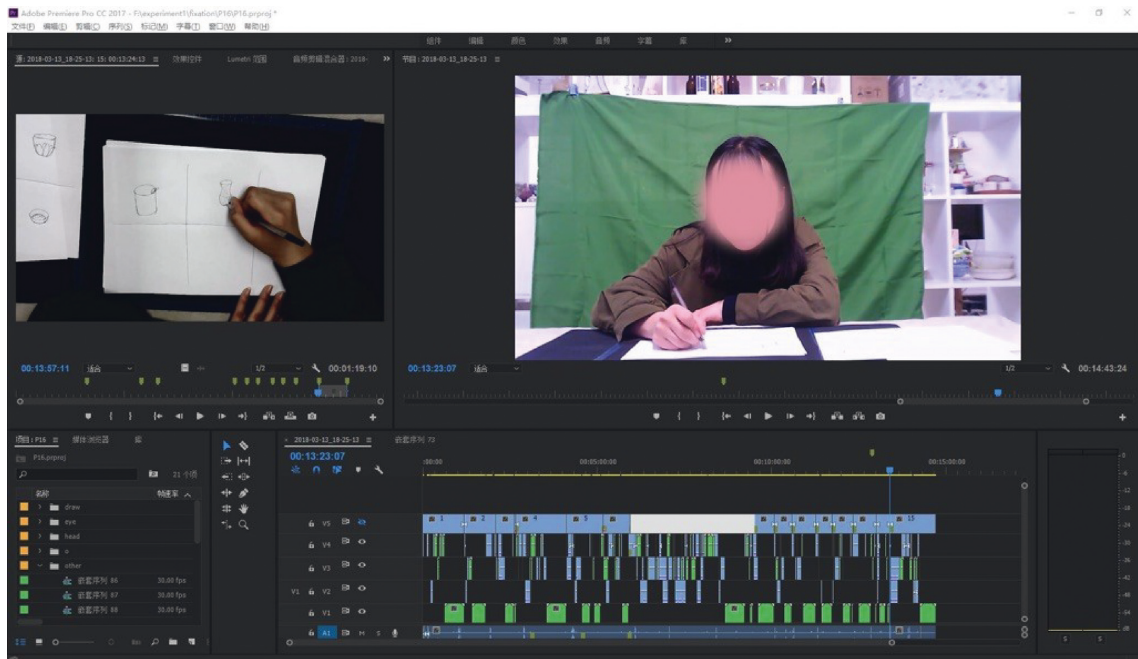


FIGURE 7: Premiere interface.

Data augmentation [50] is a conventional data processing technique used by machine learning in the field of image recognition. After random scaling, rotation, inversion, and cutting of the original image, new images were generated to expand the data volume of the training set. The ImageDataGenerator module in Keras was used to enhance the data volume in the training set. The parameter settings of data augmentation are shown in Table 1.

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task [51]. It can be used to speed up training and improve the performance of deep learning model. In this study, we will use the transfer learning using VGG-16 convolutional neural network model. The VGG-16 model was proposed by Simonyan et al. [52]. It is a CNN with 13 convolutional layers and three fully connected

TABLE 1: The parameter settings of data augmentation.

Image processing operation	Parameter setting
Rotation	40°
Lateral translation	20%
Longitudinal translation	20%
Beveling	20%
Scaling	20%
Horizontal flip	Yes

layers. The model is pretrained on a subset of the ImageNet database, which is used in the ImageNet large-scale visual recognition challenge (ILSVRC) [53].

This manuscript will use the transfer learning of VGG-16 on ImageNet dataset, the 13 convolution layers at the bottom and their weights are retained, and the three fully connected layers at the top are removed. The output features of VGG-

16 are vectored through the flatten layer and then as the input vector to the hidden layer 1, which contains 256 nerve units. The rectified linear unit (ReLU) function is used as the activation function [54]. The output result of hidden layer 1 is then input into hidden layer 2, which contains 128 nerve cells and also uses ReLU function as activation function. Finally, the output results of hidden layer 2 are mapped to the output layer by softmax function. The final outputs are probability values with the sum of 1. Figure 8 shows the overall structure of the neural network used in this model.

Adam [55] was selected as the optimization tool for model training, and the recognition accuracy was set as the performance evaluation index, with the number of epochs set to 300. The learning rate of the optimiser was set at 0.001. Default settings for hyperparameters ($\alpha = 0.0001$, $\beta_1 = 0.9$, and $\beta_2 = 0.999$) were used in the experiments. After each iteration, the recognition accuracy of the training set and validation set were recorded, respectively. After the training stage, the model was used to predict the data of the test set, and the recognition accuracy was set as the performance evaluation metric.

To investigate to which degree the results depended on the specific image-based CNN models used, we constructed the other recognition model using VGG-19 [52] and DenseNet-169 [56]. The depth of these CNN models is different. The depth of VGG-16 is 23, and that of DenseNet-169 is 169.

3.6.3. Step 2: Design Fixation Classification

(1) *Software and Hardware.* The construction and training of this model were based on Python and Scikit-learn. Scikit-learn is an open source machine learning library that was born at the Google summer camp. It contains almost all common algorithms for classification, regression, and clustering. The data used in the training model come from the solidification degree evaluation and behavior statistics obtained in the third chapter.

(2) *Behavior Statistics.* During behavior statistics, the frequency of selected types of body language significantly affected by fixation during each sketch was calculated. The remaining body language was set as “others,” and the frequency during each sketch was also calculated.

(3) *The Support Vector Machine Classifier.* At first, principal components analysis (PCA) [57] was carried out for the feature variables (the occurrence frequency of each selected type of body language), and the main components were extracted as the new feature variables, to reduce the data dimension and improve the accuracy of small sample training. The support vector machine (SVM) is then used as the classifier for design fixation classification. Finally, k-fold cross-validation was used to verify the recognition accuracy of the model.

The input vectors are the frequency of body language related to design fixation during each sketch. The remaining body languages were combined into one file as “others,” and

the frequency of “others” also represents one input vector. The output vector contains fixation and nonfixation, labeled 1 and 0, respectively.

In this paper, the recognition of design fixation is transformed into a binary classification problem, so support vector classification (SVC) or LinearSVC were selected as the classifier, and the final parameter setting can be determined through training and validation of different parameters. First, four different models were set up for training by four different parameter settings. Then, the one with the best validation results was selected as the final model. The specific parameters are shown in Table 2.

We used the k-fold cross-validation for the SVM classifier. The data were randomly divided into ten subsets, one of which was used as the validation set and the other nine as the training set. Each model is trained ten times, using a different subset as the validation set for each time. Finally, the mean value of 10 validation results was used to evaluate the recognition performance of the model.

4. Results

4.1. *Results of Data Acquisition.* In this experiment, we collected videos with a total time of 6 hours and 48 minutes, including more than 320,000 valid body language images, and a body language image database during the design process was established. Meanwhile, 156 sketches were collected for fixation evaluation. In conclusion, the total number of concepts was 156, and 13 concepts were generated per participant.

4.2. *Results of Video Analysis.* After preliminary observation, eight common types of body language were selected. They are coded in Table 3. The design behaviors can be shown in Figure 9.

The definitions of the eight behaviors are as follows: (a) drawing: drawing a sketch on a piece of paper with a pen, also including writing words to describe a scheme; (b) small movements: hand movements with no practical significance, such as turning a pen and tapping the table with a finger; (c) touching head: scratching the head with the hand, touching the hair, and touching the face with the hand; (d) eye moving: rapid eye rotation due to large eye focus movements, often accompanied by frequent blinking; (e) eyebrow moving: changes in the shape of the eyebrows due to changes in facial expressions, usually caused by a frown; (f) mouth moving: the corners of the mouth sink, and the lips stretch or move back and forth; (g) turning head: the head moves and is accompanied by a shift of vision; and (h) holding up head with hand: supporting the head with the palm, arms, and other parts of the hand and holding it still for a long time.

The eight behaviors showed a total of 1,158 times in the experimental process of all subjects, ranging from 42 times to 333 times ($M = 144.75$; $SD = 103.61$). The total frequency of each behavior is shown in Figure 10, and the descriptive statistics are shown in Table 4. The eye movement occupied a high proportion. It might be because the control of eyes is more casual, conveying sensitive changes in mentality.

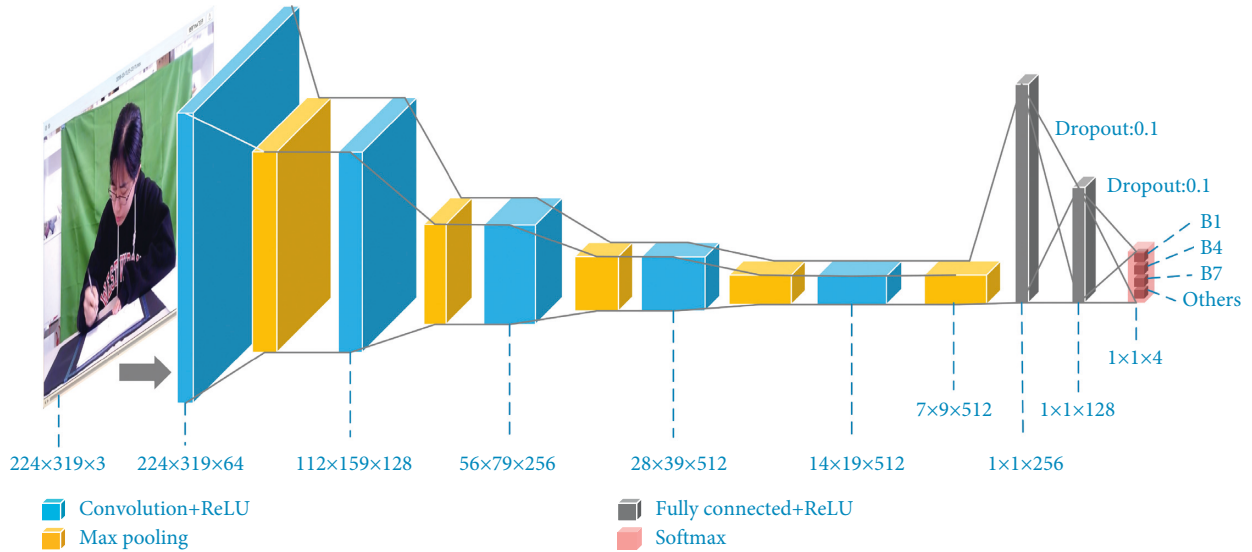


FIGURE 8: The overall structure of the neural network.

TABLE 2: SVM parameter settings.

Model number	Classifier type	Kernel function
Model A	SVC	Linear
Model B	SVC	RBF
Model C	SVC	Poly
Model D	Linear SVC	None

TABLE 3: The eight common types of body language during a design process.

Code number	Body language	Region
B1	Drawing	Hand
B2	Small movements	Hand
B3	Touching head	Hand
B4	Eye moving	Eyes
B5	Eyebrow moving	Eyebrows
B6	Mouth moving	Mouth
B7	Turning head	Head
B8	Hold up head with hand	Head

Holding the head happened less due to the large range of the action. The participants might be more limited to showing their solidification in a test.

4.3. Results of Fixation Evaluation. Figure 11 shows some examples of what was rated as a fixated design and what was not. The fixated forms have more repeated features and lower variety, novelty, and originality than the nonfixated ones.

Six experts evaluated the design fixation state of 156 design sketches of 12 designers. The voting results are shown in Table 5. One-way ANOVA was performed on the behavior frequency and fixation state of 156 sketches, and the results are shown in Table 6. The results of ANOVA show significant effects of fixation states on B1 ($F = 5.406$; $P = 0.005$), B4 ($F = 3.884$; $P = 0.023$), and B7 ($F = 6.624$; $P = 0.002$). No significant differences were found in other

behaviors. Thus, B1, B4, and B7 were selected as the body languages with significant differences between fixation and nonfixation.

The inter-rater agreement and reliability of the questions regarded whether the fixation occurred in a concept generation. The measurement property (fixation state) had a high percentage agreement (86%), and it had an excellent kappa coefficient (0.74).

The frequency data of B1, B4, and B7 for fixation and nonfixation separately is shown in Table 7. Figure 12 provides a comprehensive list of ratio statistics for behavior frequency/fixation state frequency (B/F). B4 represented the highest proportion of fixation and nonfixation, which formed a huge gap with B1 and B7. B4 highlighted and magnified the performance of the design fixation.

The ratios for behavior frequency/fixation frequency are always higher than those for behavior frequency/nonfixation frequency, although B1, B4, and B7 behaviors occur more frequently during nonfixation than those during fixation, as shown in, as shown in Tables 5 and 7 and Figure 12. In other words, B1, B4, and B7 behaviors for each fixated sketch appear more frequently than those for each nonfixated sketch. In addition, the average duration (76.86 s) of each fixated sketch is higher than that (62.41 s) of each nonfixated sketch.

4.4. Results of Body Language Recognition. As the number of iterations increases, the training accuracy and validation accuracy of the model gradually increase and tend to converge. The changes in these two indicators after each iteration are shown in Figure 13. It can be observed that the training accuracy rate and verification accuracy rate are not significantly different, nor are they at a low level at the same time. Therefore, it can be concluded that there is no occurrence of underfitting or overfitting in the model, and the training results are robust. The recognition accuracy rate of the model for the test set is 92.03%, and it can recognize the



FIGURE 9: Eight behaviors in conceptual design process: (a) drawing; (b) small movements; (c) touching head; (d) eye moving; (e) eyebrow moving; (f) mouth moving; (g) turning head; (h) holding up head with hand.

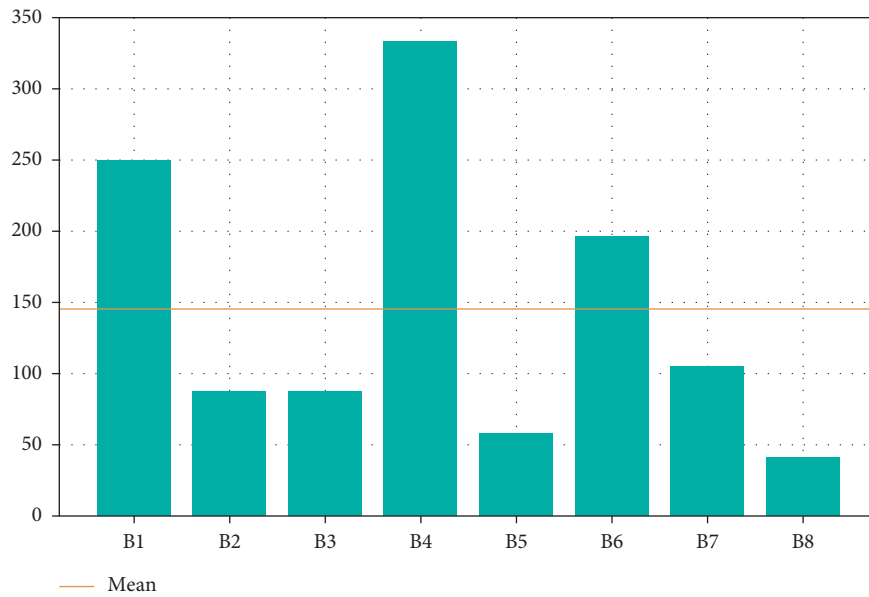


FIGURE 10: The total frequency of each behavior.

TABLE 4: The descriptive statistics of eight body language behaviors for 12 subjects.

Body language	Minimum	Maximum	Sum	Mean	SD
B1 Drawing	13	28	249	20.75	5.15
B2 Small movements	1	19	87	7.25	6.50
B3 Touching head	0	17	88	7.33	5.40
B4 Eye moving	8	53	333	27.75	13.96
B5 Eyebrow moving	0	27	58	4.83	7.93
B6 Mouth moving	3	35	196	16.33	8.94
B7 Turning head	2	19	105	8.75	4.88
B8 Holding up head with hand	0	10	42	3.50	3.29



FIGURE 11: Some examples of what was considered as a fixated/nonfixated design: (a) fixated sketches; (b) nonfixated sketches.

TABLE 5: The voting results of fixated/nonfixated sketches.

Fixation state	Numbers	Percentage	Mean value of votes	Variation
Nonfixation	112	71.8	1.8	0.82
Fixation	44	28.2	4.6	0.61

TABLE 6: The results of one-way ANOVA.

	Behavior	F value	P value
B1	Drawing	5.406	0.005*
B2	Small movements	1.618	0.481
B3	Touching head	1.618	0.202
B4	Eye moving	3.884	0.023*
B5	Eyebrow moving	0.192	0.825
B6	Mouth moving	2.635	0.075
B7	Turning head	6.624	0.002*
B8	Holding up head with hand	0.408	0.666

* $P < 0.05$.

TABLE 7: The frequency of B1, B4, and B7 for fixation states.

	B1 (times)	B4 (times)	B7 (times)
Nonfixation	164	217	59
Fixation	85	116	46

four types of body language (B1, B4, B7, and others) accurately.

For comparison purposes, we used the same datasets for the other image-based CNN recognition model. The results of training and validation are shown in Figure 13. For the test set, the accuracy of VGG-19 is 91.09%, and that of DenseNet-169 is 73.06%. Hence, the accuracies of the training, validation, and test sets of VGG-19 model are slightly lower than those obtained with the VGG-16 model, respectively. The DenseNet-169 model performs the poorest and has the problem of overfitting. It could be seen that the performance of VGG-16 was more stable, and the overall accuracy was the highest among the three.

It could be seen from Table 8 that the training time of DenseNet-169 was the shortest in the three aspects under the same computing ability, while VGG-19 was the longest. The time costs of three algorithms represented by these three criteria were consistent. Although the duration of VGG-16 was longer than that of DenseNet-169, its performance was improved. The recognition of VGG-16 was not only more

accurate, but also of a less computational complexity compared with VGG-19.

The comparison between models A, B, and C is shown in Figure 14. It could be seen that the average value of model B was the highest and the fluctuation was the lowest. There was a significant difference compared with model C. Therefore, model B was selected for the second step.

4.5. Results of Design Fixation Classification

4.5.1. Results of PCA. The feature variables of the original data were analyzed to obtain four components. The variance percentage and the cumulative contribution rate of each component are shown in Table 9.

Figure 15 shows the scree plot of the analysis results of the PCA. The slopes of components 1 and 2 are the highest. According to Table 6 and Figure 15, components 1 and 2 were selected as the main components, and their cumulative contribution rate reached 90.92%. Mapping the feature variables of the original data to these two components for dimensionality reduction can not only reduce the dimension of features, but also retain most of the information in the original data. The composition of feature variables after dimension reduction is shown in Table 10. The features presented by this matrix showed that B4 had the highest component 1, which was similar to the features presented in Figure 13. The “others,” including more body language, came in a close second place to B7.

4.6. Results of Cross-Validation. The cross-validation results of the four models are shown in Table 11, including the accuracies of 10 cross-validation results and the mean accuracy (MA) for the last row. Model B has the highest average accuracy, so that we can select model B as the final one.

The ten validation results of model B are shown in Figure 16. The difference between each time is small, and the standard deviation is 3.77%, indicating that the model has a high robustness. The mean value of 10 validation results was 72.9%, indicating that the model shows promising recognition performance.

4.7. Results of Testing Set. The recognition accuracy rate of the model for the test set is 79.11%. Figure 17 shows the

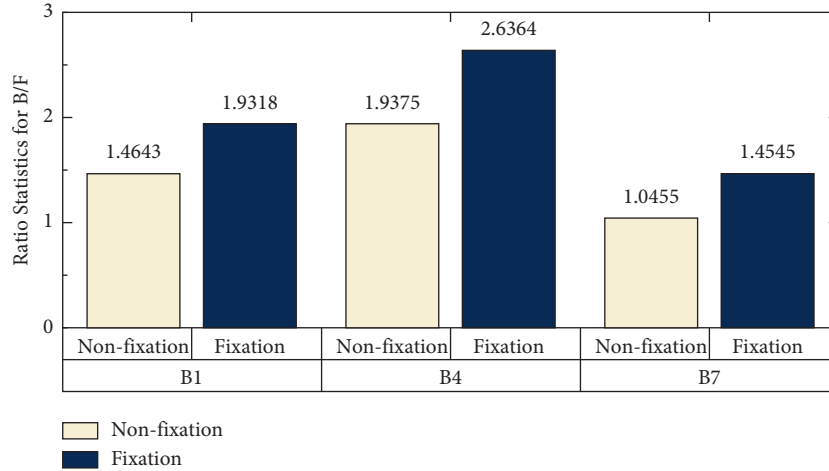


FIGURE 12: Ratio statistics for behaviors/fixation states (B/F).

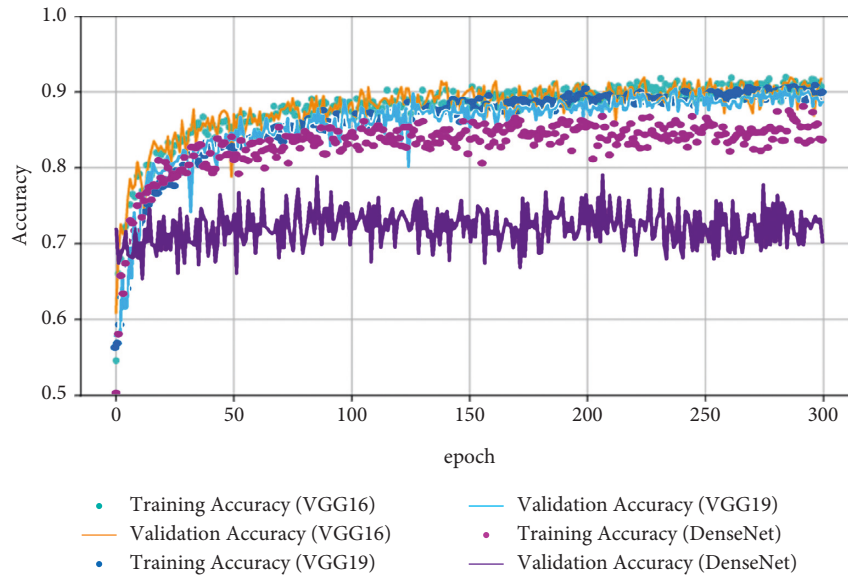


FIGURE 13: The results of the recognition model based on VGG-16, VGG-19, and DenseNet-169.

TABLE 8: The training time of the recognition model based on VGG-16, VGG-19, and DenseNet-169.

Training time criteria	VGG-16	VGG-19	DenseNet-169
Average time cost per epoch (s)	54.5	62	38.5
Longest time per epoch (s)	65	73	55
Median of time costs (s)	54	62	38

visualized results of the recognition model. In the figure, the horizontal and vertical axes represent components 1 (x1) and 2 (x2), respectively. The yellow point represents the nonfixation sample, and the blue point represents the fixation sample. The boundary between the light blue and green areas is the decision boundary of the model, and the samples within the boundary (the green area) are recognized as fixation by the model.

5. Discussion

Within the literature, when and how to combat design fixation have been the focus of researchers and designers. Until now, a number of approaches to overcome design fixation have been identified, such as defixation instructions [14] and analogical inspiration [58, 59]. However, to our knowledge, when to present the defixation methods remains unknown, because the perception of being fixated is unconscious and the designers are always unaware that they are being influenced by example solutions or previously generated solutions [60, 61]. The participants also cannot acknowledge, in retrospect, that they were fixated via our postinterview observations.

This study is the first to use computer vision and deep learning to recognize design fixation from body language during the design process. The results suggest that our two-step hybrid model performs well in fixation classification.

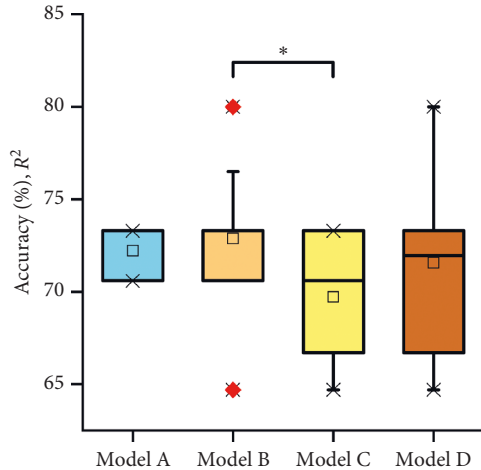


FIGURE 14: The comparison between models A, B, and C. The statistical differences among the accuracy were measured by the paired sample Wilcoxon signed-rank test, where “*” represents the significant difference in accuracy between two objects at the 0.05 level. Error bars denote variance of the means.

TABLE 9: The results of PCA.

Components	Percentage of variance	Cumulative contribution rate (%)
1	69.68	69.68
2	21.24	90.92
3	4.96	95.88
4	4.12	100

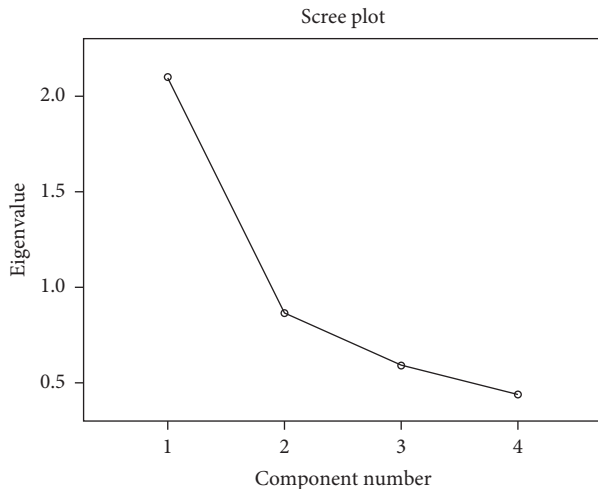


FIGURE 15: Scree plot of PCA.

TABLE 10: Component matrix.

Original feature variable	Component 1	Component 2
B1 frequency	0.5177	0.830
B4 frequency	0.807	-0.300
B7 frequency	0.763	-0.293
“Others” frequency	0.773	-0.047

The impact of the work could be that fixation could be detected not only by the design sketch outcomes, but as it is happening by monitoring the movements, expressions, and gestures of designers in the design sketch process before the sketch outcomes.

The characteristic of the design fixation recognition task is that the recognition of a single action is essential, while the fixation represented by the action does not correspond to the visual features. The direct recognition method leads to confusion in visual feature extraction. Therefore, the task was divided into two steps. First, the visual features were extracted, and second, the actions represented by the features were classified into two categories.

In the first step, the extractor with good performance and stability was used as the backbone network of the pretraining models, which increased the reliability of action recognition. As long as the results of the first step were accurate, the second step can meet the requirements of accurate final results without large fluctuations. Therefore, a more reliable SVM was chosen.

In this study, we found out that the body languages B1, B4, and B7 were significantly related to the design fixation. Then, we proposed a two-step hybrid model for design fixation classification. VGG-16 convolution neural network was used for body language recognition, and SVM was used for design fixation classification. The accuracy rate for body language recognition is 92.03%, and for design fixation classification, it is 79.11%. The recognition performance attained with our approach is promising. In most cases, four types of body language (drawing, eye movement, turning the head, and others) can be recognized accurately. Meanwhile, our SVM-based method can accurately recognize whether the designer is in a fixation state. Compared with the neural network, SVM has the advantage of high accuracy for small sample training [62, 63].

The present study also found that the recognition results of the VGG-16-based and VGG-19-based models were better than those of the DenseNet-169-based model. The reason may be that the model complexity of DenseNet-169 is mismatching the data complexity of our experiments. This finding suggests that the VGG models are more appropriate than the DenseNet-169 for design fixation recognition via body language in this study. The recognition results of the VGG-16-based model were slightly better than those of the VGG-19-based model. However, the performance differences between the two classifiers were so small that they, for practical purposes, can be regarded as equivalent.

The method proposed in this study could be used in the reminder applications to correct the design fixation of designers and improve the outputs and efficiency. In addition, the two-step algorithm can provide an idea for other studies to solve the problem of recognition failure caused by the hybrid visual features in large categories.

Nevertheless, our findings and general approach have several limitations:

- (1) During the experiment, two cameras were employed. One was installed in front of the subjects to record the body language during the design process, and the

TABLE 11: The cross-validation results of each model.

Model	Accuracies of 10 cross-validation results (%)										MA (%)
	1	2	3	4	5	6	7	8	9	10	
A	70.6	70.6	70.6	70.6	73.3	73.3	73.3	73.3	73.3	73.3	72.24
B	76.5	64.7	70.6	70.6	73.3	80	73.3	73.3	73.3	73.3	72.9
C	70.6	64.7	70.6	64.7	73.3	73.3	66.7	73.3	66.7	73.3	69.73
D	76.5	70.6	70.6	64.7	73.3	73.3	66.7	73.3	66.7	80	71.57

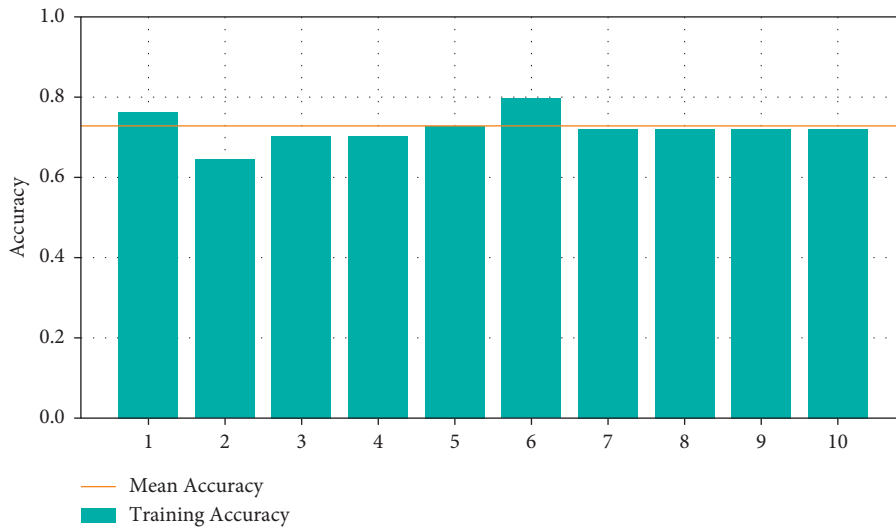


FIGURE 16: The 10 cross-validation results of model B.

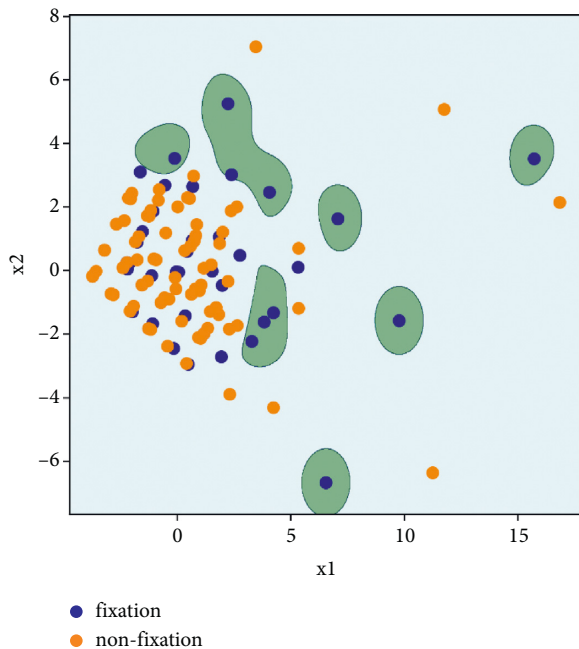


FIGURE 17: The visualized results of the recognition model.

other one was installed above the head to record the sketches. The upper one can clearly record the sketches. However, the video quality of the front one was vulnerable to sitting problems of some subjects. As a result, facial movements and expressions were

often not recorded. In our future work, a third camera to capture the facial movements will be added.

- (2) The behavior image database is big enough to support the training of body language recognition model based on computer vision and VGG-16. However, the sketch database still needs to be expanded, although this problem has been alleviated through dimension reduction and cross-validation. In order to improve the accuracy rate of design fixation classification model, we plan to test more subjects to fill in the data pond.
- (3) Through manual coding, we identified eight common behaviors during the design process. However, manual video coding is time and resource consuming [29], and it is inevitable that there exist some errors and uncertainties in the manual observation statistics. In future work, we plan to make certain parts of data collection and analysis automated.
- (4) The body language database can be obtained by objective evaluation with small error. However, the fixation evaluation database comes from subjective evaluation, which may be influenced by error and bias in human judgments [43]. Morgan et al. [64] pointed out that the quantitative objective evaluation should be prior to the subjective opinions, comments, and ratings. In our future work, we can conduct an objective evaluation of the fixation state.

- (5) In our study, we only classified two fixation states (fixation and nonfixation). In future work, we plan to test whether our method is also able to detect fixation states of different intensities, such as low-intensity, medium-intensity, and high-intensity fixation states, and distinguish, for example, low-intensity fixation states from medium-intensity fixation states.
- (6) For the design task, the participants were only designing the form of a coffee mug. Much of the design fixation research has focused on the functionality of products [65]. The focus on form design only may impact the results here. In future work, we will explore the functional fixedness in problem-solving situations.
- (7) The design problem employed here is simple. The relatively low frequency and short sketch duration may affect experimental results. In future work, we will consider the problem is significantly more complex and has multiple functions associated with it.
- (8) There was a limitation in the proposed method. For example, the selection of the basic algorithm has a delaying attribute. All the VGG, DenseNet, and SVM were proposed for several years. We should continue to try the latest networks, like YOLO (You Only Look Once) on target recognition, or Capsule Networks with vectorized properties.

6. Conclusion

We proposed a two-step hybrid recognition model of design fixation based on body language. While the results are encouraging, additional research is needed to further develop the method. Our future work will concentrate on the development of an artificial-intelligence-aided design system that can be used to recognize the fixation state of the designer automatically via objective body language during sketch process ahead of sketch outcomes and give the appropriate inspiration in time to help combat fixation easily.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This study was partly supported by the National Natural Science Foundation of China (no. 51905175), Shanghai Pujiang Talent Program (no. 2019PJC021), the Cooperative Education Program of Ministry of Education (no. 202002SZ05), and the Shanghai Soft Science Key Project (no. 21692196800).

References

- [1] A. Cabello, *Human-centered Innovation Processes: The Case of Design Thinking in Nascent and Large Firms*, Ph.D. thesis, EPFL College, Lausanne, Switzerland, 2015.
- [2] G. Scott, L. E. Leritz, and M. D. Mumford, "The effectiveness of creativity training: a quantitative review," *Creativity Research Journal*, vol. 16, no. 4, pp. 361–388, 2004.
- [3] N. Crilly, "Fixation and creativity in concept development: the attitudes and practices of expert designers," *Design Studies*, vol. 38, pp. 54–91, 2015.
- [4] S. C. Stewart, "Design thinking: understanding how designers think and work," *Design Studies*, vol. 32, no. 6, pp. 608–609, 2011.
- [5] D. P. Moreno, M. C. Yang, A. A. Hernández, J. S. Linsey, and K. L. Wood, *A Step Beyond to Overcome Design Fixation: A Design-by-Analogy Approach*, Springer International Publishing, Cham, Switzerland, 2015.
- [6] A. T. Purcell and J. S. Gero, "Design and other types of fixation," *Design Studies*, vol. 17, no. 4, pp. 363–383, 1996.
- [7] R. J. Youmans, "The effects of physical prototyping and group work on the reduction of design fixation," *Design Studies*, vol. 32, no. 2, pp. 115–138, 2011.
- [8] J. S. Linsey, I. Tseng, K. Fu, J. Cagan, K. L. Wood, and C. Schunn, "A study of design fixation, its mitigation and perception in engineering design faculty," *Journal of Mechanical Design*, vol. 132, 2010.
- [9] S. Yilmaz, C. M. Seifert, and R. Gonzalez, *Cognitive Heuristics in Design: Instructional Strategies to Increase Creativity in Idea Generation*, Cambridge University Press, Cambridge, UK, 2010.
- [10] R. J. Youmans and T. Arciszewski, *Design Fixation: A Cloak of Many Colors*, Springer Netherlands, Dordrecht, Netherlands, 2014.
- [11] I. Tseng, J. Moss, J. Cagan, and K. Kotovsky, "The role of timing and analogical similarity in the stimulation of idea generation in design," *Design Studies*, vol. 29, no. 3, pp. 203–221, 2008.
- [12] L. A. Vasconcelos and N. Crilly, "Inspiration and fixation: questions, methods, findings, and challenges," *Design Studies*, vol. 42, pp. 1–32, 2016.
- [13] D. G. Jansson and S. M. Smith, "Design fixation," *Design Studies*, vol. 12, no. 1, pp. 3–11, 1991.
- [14] E. G. Chrysikou and R. W. Weisberg, "Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 31, no. 5, pp. 1134–1148, 2005.
- [15] J. S. Gero, "Research methods for design science research: computational and cognitive approaches," in *Proceedings of the ANZASCA Conference*, Sydney, Australia, November 2000.
- [16] R. J. Youmans and T. Arciszewski, "Design fixation: classifications and modern methods of prevention," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 28, no. 2, pp. 129–137, 2014.
- [17] R. A. Finke, "Imagery, creativity, and emergent structure," *Consciousness and Cognition*, vol. 5, no. 3, pp. 381–393, 1996.
- [18] M. Perttula and P. Sipilä, "The idea exposure paradigm in design idea generation," *Journal of Engineering Design*, vol. 18, no. 1, pp. 93–102, 2007.
- [19] D. W. Dahl and P. Moreau, "The influence and value of analogical thinking during new product ideation," *Journal of Marketing Research*, vol. 39, no. 1, pp. 47–60, 2002.
- [20] D. Zahner, J. V. Nickerson, B. Tversky, J. E. Corter, and J. Ma, "A fix for fixation? rerepresenting and abstracting as creative processes in the design of information systems," *Artificial*

- Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 24, no. 2, pp. 231–244, 2010.
- [21] U. N. Sio, K. Kotovsky, and J. Cagan, “Fixation or inspiration? a meta-analytic review of the role of examples on design processes,” *Design Studies*, vol. 39, pp. 70–99, 2015.
- [22] N. Crilly and C. Cardoso, “Where next for research on fixation, inspiration and creativity in design?” *Design Studies*, vol. 50, pp. 1–38, 2017.
- [23] N. Crilly, “Methodological diversity and theoretical integration: research in design fixation as an example of fixation in research design?” *Design Studies*, vol. 65, pp. 78–106, 2019.
- [24] Y. Zhao, *Human Emotion Recognition from Body Language of the Head Using Soft Computing Techniques*, Ph.D. thesis, Université d’Ottawa/University of Ottawa, Ottawa, Canada, 2012.
- [25] A. Mehrabian, *Nonverbal Communication*, Routledge, London, UK, 2017.
- [26] R. L. Birdwhistell, *Kinesics and Context: Essays on Body Motion Communication*, University of Pennsylvania press, Philadelphia, PA, USA, 2010.
- [27] Z. Yang and Y. Chen, “An semg approach to recognize the body language of the head based on the gga-elman network,” *CAA Transactions on Intelligent Systems*, vol. 9, no. 4, pp. 385–391, 2014.
- [28] D. E. Matsumoto, H. C. Hwang, and M. G. Frank, *APA Handbook of Nonverbal Communication*, American Psychological Association, Washington, DC, USA, 2016.
- [29] A. Wulvik, J. Erichsen, M. Steinert et al., “Capturing body language in engineering design—tools and technologies,” in *Proceedings of the NordDesign 2016*, vol. 1, pp. 165–174, Trondheim, Norway, August 2016.
- [30] P. Cash and A. Maier, “Prototyping with your hands: the many roles of gesture in the communication of design concepts,” *Journal of Engineering Design*, vol. 27, no. 1–3, pp. 118–145, 2016.
- [31] S. Bezawada, Q. Hu, A. Gray, T. Brick, and C. Tucker, “Automatic facial feature extraction for predicting designers’ comfort with engineering equipment during prototype creation,” *Journal of Mechanical Design*, vol. 139, Article ID 021102, 2017.
- [32] L. Sun, W. Xiang, C. Chai, Z. Yang, and K. Zhang, “Designers’ perception during sketching: an examination of Creative Segment theory using eye movements,” *Design Studies*, vol. 35, no. 6, pp. 593–613, 2014.
- [33] I. Behoora and C. S. Tucker, “Machine learning classification of design team members’ body language patterns for real time emotional state detection,” *Design Studies*, vol. 39, pp. 100–127, 2015.
- [34] W.-H. Yun, D. Lee, C. Park, J. Kim, and J. Kim, “Automatic recognition of children engagement from facial video using convolutional neural networks,” *IEEE Transactions on Affective Computing*, vol. 11, no. 4, pp. 696–707, 2020.
- [35] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [36] S. Lawrence, C. L. Giles, A. C. Ah Chung Tsoi, and A. D. Back, “Face recognition: a convolutional neural-network approach,” *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, 1997.
- [37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, Lake Tahoe, NV, USA, January 2012.
- [38] D. Cireşan, U. Meier, and J. Schmidhuber, “Multi-column deep neural networks for image classification,” in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 13–18, Providence, RI, USA, June 2012.
- [39] O. Eris, N. Martelaro, and P. Badke-Schaub, “A comparative analysis of multimodal communication during design sketching in co-located and distributed environments,” *Design Studies*, vol. 35, no. 6, pp. 559–592, 2014.
- [40] J. Borg, *Body Language: 7 Easy Lessons to Master the Silent Language*, FT press, Upper Saddle River, NJ, USA, 2009.
- [41] Z. Yang, A. Metallinou, and S. Narayanan, “Analysis and predictive modeling of body language behavior in dyadic interactions from multimodal interlocutor cues,” *IEEE Transactions on Multimedia*, vol. 16, pp. 1766–1778, 2014.
- [42] C. Chai, F. Cen, W. Ruan, C. Yang, and H. Li, “Behavioral analysis of analogical reasoning in design: differences among designers with different expertise levels,” *Design Studies*, vol. 36, pp. 3–30, 2015.
- [43] J. Annett, “Subjective rating scales: science or art?” *Ergonomics*, vol. 45, no. 14, pp. 966–987, 2002.
- [44] K. Grantham, G. Okudan, T. W. Simpson, and O. Ashour, “A study on situated cognition: product dissection’s effect on redesign activities,” in *Proceedings of the ASME 2010 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. 617–626, American Society of Mechanical Engineers (ASME), Quebec, Canada, August 2010.
- [45] M. Agogué, A. Kazakçi, B. Weil, and M. Cassotti, “The impact of examples on creative design: explaining fixation and stimulation effects,” in *DS 68-2: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 2: Design Theory and Research Methodology, Lyngby/Copenhagen*, Copenhagen, Denmark, August 2011.
- [46] L. B. Mokkink, C. B. Terwee, E. Gibbons et al., “Inter-rater agreement and reliability of the cosmin (consensus-based standards for the selection of health status measurement instruments) checklist,” *BMC Medical Research Methodology*, vol. 10, no. 1, p. 82, 2010.
- [47] A. Gulli and S. Pal, *Deep Learning with Keras*, Packt Publishing Ltd, Birmingham, UK, 2017.
- [48] J. Lemley, S. Bazrafkan, and P. Corcoran, “Deep learning for consumer devices and services: pushing the limits for machine learning, artificial intelligence, and computer vision,” *IEEE Consumer Electronics Magazine*, vol. 6, no. 2, pp. 48–56, 2017.
- [49] A. Koul, C. Becchio, and A. Cavallo, “Cross-validation approaches for replicability in psychology,” *Frontiers in Psychology*, vol. 9, p. 1117, 2018.
- [50] J. Salamon and J. P. Bello, “Deep convolutional neural networks and data augmentation for environmental sound classification,” *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 279–283, 2017.
- [51] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [52] K. Simonyan and A. Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*, <http://arxiv.org/abs/1409.1556>, 2014.
- [53] O. Russakovsky, J. Deng, H. Su et al., “Imagenet large scale visual recognition challenge,” *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

- [54] C. Xing, L. Ma, and X. Yang, "Stacked denoise autoencoder based feature extraction and classification for hyperspectral images," *Journal of Sensors*, vol. 2016, 2016.
- [55] D. P. Kingma, J. Ba, and B. Adam, "A method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, December 2015.
- [56] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, Honolulu, HI, USA, July 2017.
- [57] Ł. Sadowski, M. Nikoo, and M. Nikoo, "Principal component analysis combined with a self organization feature map to determine the pull-off adhesion between concrete layers," *Construction and Building Materials*, vol. 78, pp. 386–396, 2015.
- [58] H. Casakin and G. Goldschmidt, "Expertise and the use of visual analogy: implications for design education," *Design Studies*, vol. 20, no. 2, pp. 153–175, 1999.
- [59] C. BT and S. CD, "The relationship of analogical distance to analogical function and preinventive structure: the case of engineering design," *Memory & Cognition*, vol. 35, p. 29, 2007.
- [60] R. L. Marsh, T. B. Ward, and J. D. Landau, "The inadvertent use of prior knowledge in a generative cognitive task," *Memory & Cognition*, vol. 27, no. 1, pp. 94–105, 1999.
- [61] S. M. Smith, T. B. Ward, and J. S. Schumacher, "Constraining effects of examples in a creative generation task," *Memory & Cognition*, vol. 21, no. 6, pp. 837–845, 1993.
- [62] N. Loukeris and I. Eleftheriadis, "Support vector machines neural networks to a hybrid neuro-genetic svm form in corporate financial analysis," in *Proceedings of the 15th CSCC International Conference on Systems*, pp. 410–416, Corfu, Greece, August 2011.
- [63] B. Xing, K. Zhang, S. Sun et al., "Emotion-driven Chinese folk music-image retrieval based on de-svm," *Neurocomputing*, vol. 148, pp. 619–627, 2015.
- [64] C. T. Morgan, J. S. Cook, A. Chapanis, and M. W. Lund, *Human Engineering Guide to Equipment Design*, FT press, Upper Saddle River, NJ, USA, 1963.
- [65] A. T. Purcell and J. S. Gero, "Effects of examples on the results of a design activity," *Knowledge-Based Systems*, vol. 5, no. 1, pp. 82–91, 1992.