Future Technology
Open Access Journal
ISSN 2832-0379

February 2026| Volume 05 | Issue 01 | Pages 135-147

Journal homepage: https://fupubco.com/futech

https://doi.org/10.55670/fpll.futech.5.1.12

Future Publishing LLC

Article

Pictogram semantics standardization for barrier-free drug packaging: deep-learning-assisted design guidelines

Hui Li, Verly Veto Vermol*, Zulimran Ahmad

College of Creative Arts, MARA University of Technology, 40450 Shah Alam, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 20 August 2025
Received in revised form
11 October 2025
Accepted 31 October 2025

Keywords:
Deep learning, Icon semantic standardization, Age-friendly design, Pharmaceutical packaging, Accessible design

*Corresponding author
Email address:
2022414788@isiswa.uitm.edu.my

DOI: 10.55670/fpll.futech.5.1.12

ABSTRACT

As the world gets older, elderly users find it harder to understand information on medicine packaging. This study created a framework to improve visual communication for older people using deep learning to standardize icons. The research involved 200 participants aged 60 and older who answered questionnaires and took part in interviews, while deep learning models were trained with 1,500 medicine icons. The Residual Network-50 (ResNet-50) model reached 94.8% accuracy, outperforming VGG-16 (89.6%) and Vision Transformer (92.1%), in recognizing meanings across 21 icon types. Analysis showed that performance risk, psychological risk, and safety risk affect how older users accept these icons, with distrust playing a role (R^2=0.723), and psychological risk being responsible for 54.6% of the indirect effect. Testing showed that using standardized icons raised recognition accuracy from 68.3% to 92.5% and cut down comprehension time by 52%(t=9.87, p<0.001, Cohen's d=2.21). The recommended design standards (icon diameter ≥20mm, font size ≥14pt, contrast ratio ≥7:1) give measurable guidelines for the medicine industry and are important for encouraging healthy aging.

1. Introduction

Global population aging has become a major 21st-century demographic feature, with persons aged 60+ projected to reach 2.1 billion by 2050 [1]. This shift creates healthcare challenges, particularly in medication management. Elderly users face difficulties reading drug labels, understanding dosage instructions, and managing packaging, leading to reduced medication adherence and increased adverse effects [2]. Age-related visual decline—including reduced contrast sensitivity, poor color perception, and near vision impairment—compounds information recognition challenges [3]. Age-centered design research emphasizes incorporating cognitive, perceptual, and motor changes into product development [4]. Barrier-free design principles have expanded from public spaces to pharmaceutical packaging, prioritizing underserved populations [5]. Visual contrast enhances readability for elderly consumers [6], while empathetic design addresses emotional needs [7]. Emerging technologies like image recognition in elderly care robots demonstrate intelligent systems' potential to assist aging populations [8]. Research demonstrates that pharmaceutical packaging elements—including color, layout, and images—significantly influence user behavior and emotional responses [9]. Cross-cultural studies reveal variations in color meanings and preferences [10], while emotional design theory emphasizes addressing user psychological needs beyond functionality [11]. Visual aesthetics research confirms that consumers value product appearance in individualized ways [12], with pharmaceutical packaging color specifically affecting user expectations [13]. Despite these insights, current research lacks standardized approaches to making medication information accessible for elderly users through clear visual symbols. Recent advances in artificial intelligence (AI) technology offer new solutions to these challenges. Deep learning demonstrates exceptional capabilities in medical image analysis [14], with residual neural networks [15] and convolutional neural networks(CNN) [16] showing particular advantages for complex data processing. Clinical implementation guidelines provide clear directions for practitioners [17]. Successful applications incorporating prior feature knowledge in diagnosis [18], CNN-based medical imaging [19], disease-specific treatment planning [20], and COVID-19 image classification [21] indicate technological maturity. The widespread phenomenon of self-medication [22] further underscores the need for improved accessibility of pharmaceutical package design. Despite significant advances in medicine, the application of deep learning to interpreting and assessing pharmaceutical packaging symbols remains a relatively nascent field that warrants further development. Studies indicate the need for standardized health datasets used in AI technologies [23]. Works regarding rules for the

employment of AI in healthcare globally provide recommendations for the effective utilization of technology [24]. A rapid glance over quality norms for the utilization of AI in healthcare [25] and discourses regarding the requirement for standard terms in data-intensive medical AI [26], both emphasize significantly the necessity for standards so that technology may be utilized safely and efficaciously. This research bridges this gap by proposing deep learning-assisted design principles for pharmaceutical packaging icon standardization. Integrating Innovation Resistance Theory with AI technology, the study develops a standardized framework enabling elderly users to better comprehend medication information and ensure safety. This work advances accessible pharmaceutical packaging design through: (1) developing a deep learning-assisted standardization framework, (2) establishing quantifiable design parameters, (3) revealing resistance mechanisms, and (4) validating effectiveness through controlled experimentation. The findings provide actionable guidelines for pharmaceutical industries and regulatory authorities, contributing to healthy aging and inclusive society development.

2. Methodology

2.1 Theoretical framework and hypotheses

This research establishes a mediation model based on Innovation Resistance Theory to examine the resistance mechanisms of elderly users against standardized pharmaceutical packaging icons. The framework integrates risk perception (performance, psychological, and safety risks), trust mechanisms, and technology pressure to explain acceptance behavior [27]. Performance risk reflects comprehension challenges, psychological risk indicates emotional unease, and safety risk concerns medication accuracy—all reducing acceptance willingness.

Safety risk involves dosage accuracy concerns. Trust mediates the relationship between risk perception and resistance [28], as distrust amplifies resistance even when designs meet standards. Technology anxiety, documented in wearable devices [29] and digital services research [30], moderates this relationship—high technological pressure strengthens the effect of distrust on resistance. Figure 1 illustrates this framework, integrating direct, mediating, and moderating effects to explain elderly users' acceptance mechanisms.

2.2 Research design and data collection

This research employs a convergent mixed-methods design to examine elderly users' cognitive features and acceptance mechanisms regarding standardized pharmaceutical packaging icons. The approach combines qualitative interviews and quantitative surveys simultaneously, leveraging methodological complementarity to strengthen research inferences [31]. Meta-inference analysis reveals semantic-level comprehension barriers [32]. Data collection occurred in 2023 for both city and rural China, sourcing pictograms from 45 pharmaceutical companies. It employed the convenience sampling and snowball sampling techniques for participants aged 60 years and above. The research team distributed structured questionnaires in community health service centers, senior activity centers, and on the Internet, resulting in the collection of 200 valid samples. It included crucial issues like performance risk, psychological risk, safety risk, distrust, technological pressure, and resistance to standardized icon systems. All the queries were scaled using a seven-point Likert scale. Representative items included performance risk assessments (e.g., 'Standardized pictograms may fail to convey dosage information accurately', α=0.89), psychological risk measures

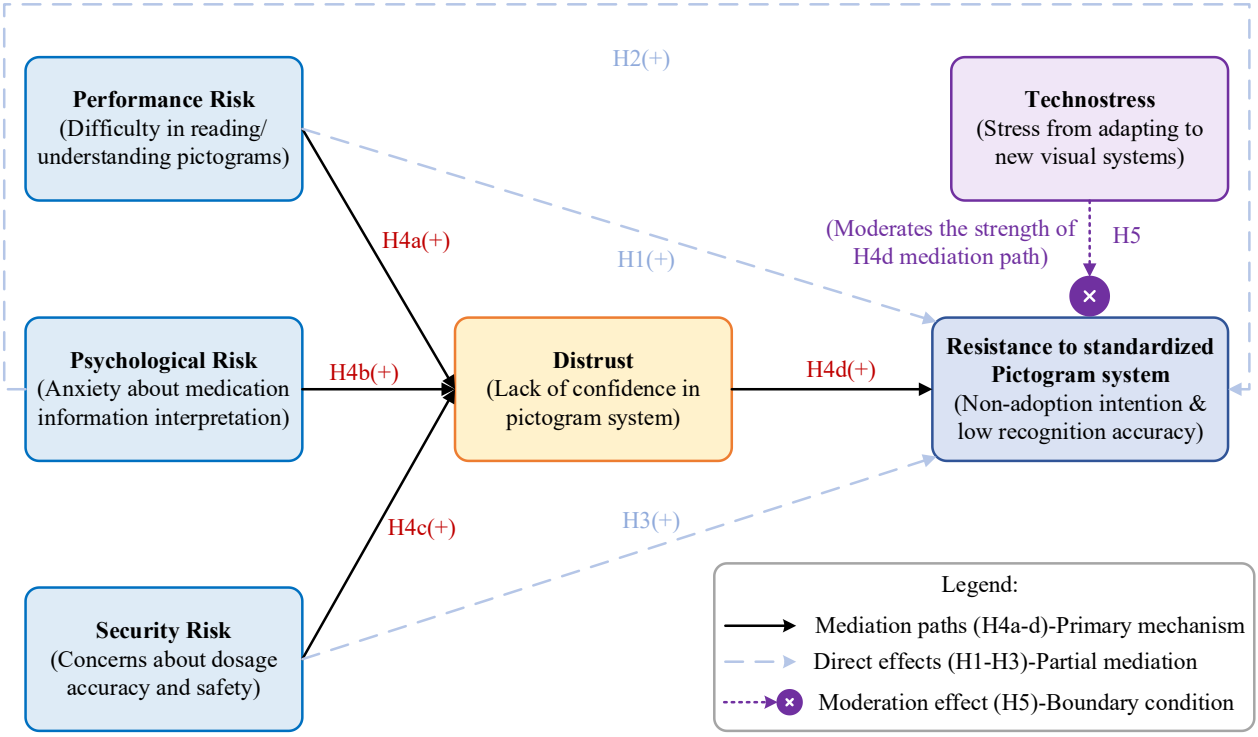

Figure 1. Theoretical framework integrating innovation resistance theory with direct, mediating, and moderating effects

(e.g., 'Unfamiliar pictogram designs trigger anxiety', α=0.92), and distrust indicators (e.g., 'New visual systems on medication packaging lack credibility', α=0.88). The demographic information of the sample, depicted in Table 1, indicates a good distribution across gender, age, education level, and residence, hence the representativeness of the findings. Semi-structured interviews with 30 elderly users (30-45 minutes each) explored four domains: (1) pictogram recognition difficulties, (2) emotional responses to unclear icons, (3) medication safety concerns, and (4) preferred design features. Interviews were transcribed and analyzed thematically, achieving inter-rater reliability of κ=0.84. Quantitative analysis employed Partial Least Squares Structural Equation Modeling (PLS-SEM) for complex mediation modeling [33]. SPSS 26.0 conducted descriptive statistics and reliability testing, while SmartPLS 4.0 evaluated measurement and structural models to test hypothesized direct, mediating, and moderating effects.

Table 1. Demographic distribution of elderly participants (N=200)

Characteristic	Category	Frequency	Percentage (%)
Gender	Male	92	46.0
	Female	108	54.0
Age Group	60-65 years	68	34.0
	66-70 years	75	37.5
	71-75 years	42	21.0
	76+ years	15	7.5
Education Level	Primary or below	45	22.5
	Middle school	82	41.0
	High school	53	26.5
	College or above	20	10.0
Residence	Urban	128	64.0
	Rural	72	36.0
Chronic Medication Use	Yes	156	78.0
	No	44	22.0

2.3 Deep learning model and validation

This study employs a deep residual network (ResNet) for the semantic recognition of pharmaceutical packaging icons to objectively evaluate the recognizability of icon designs. The residual network effectively mitigates the vanishing gradient problem in deep networks through its skip-connection mechanism, enabling the model to learn complex visual feature representations [34]. As shown in Figure 2, the model adopts the ResNet-50 architecture comprising 16 residual blocks (configured as 3+4+6+3). The input layer receives 224×224-pixel RGB icon images. Following initial convolutions and pooling, data sequentially pass through four sets of residual blocks to extract multi-scale features. The final output consists of classification probabilities generated by global average pooling and a fully connected layer.

Model training employs a cross-entropy loss function to optimize network parameters, defined as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log(\hat{y}_{ic}) \quad (1)$$

where N represents the batch size(N=32), C denotes the number of classes (C=21 in this study), y_{ic} indicates the true label, and $\hat{y}_{ic}$ signifies the model prediction probability. The optimizer employs the Adam algorithm with a learning rate of 0.001 and a batch size of 32. Training runs for 100 epochs using early stopping (with a tolerance of 10 epochs). Model training utilized an NVIDIA RTX 3090 GPU (24GB VRAM) with CUDA 11.7 and PyTorch 1.13.0 framework, requiring approximately 6 hours for convergence. Data augmentation includes random rotation (±15°), horizontal flipping, and brightness adjustment. These augmentation strategies expanded the effective training set threefold, enhancing model robustness against variations in real-world pharmaceutical packaging.

Model performance is evaluated using multiple metrics. The accuracy and F1 score are calculated as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{F1-Score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \quad (2)$$

TP, TN, FP, and FN represent the number of true positive, true negative, false positive, and false negative samples, respectively.

To enhance model interpretability, the study integrates gradient-weighted class activation mapping (GW-CAM) [35]. This method generates a heatmap revealing the model's focus areas by calculating the gradient weights of the target class c on the feature map A^k of the final convolutional layer:

$$\alpha_k^c = \frac{1}{Z}\sum_i\sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (3)$$

$$L_{\text{Grad-CAM}}^c = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right) \quad (4)$$

where α_k^c represents the importance weight of the k-th feature map for the class c, Z is the normalization constant, and y^c denotes the score for class c. The Rectified Linear Unit (ReLU) function ensures that only positively correlated features are highlighted. This visualization mechanism validates whether the model focuses on the semantic core regions of pictograms, ensuring algorithmic decision transparency. To verify the effectiveness of ResNet-50, this study compared it with Transformer-based vision models [36]. The widespread application of deep convolutional networks in medical image analysis provided methodological support for this research [37].

2.4 Ethical considerations

This research received Institutional Review Board approval and followed the Declaration of Helsinki guidelines. Participants provided informed consent after detailed briefings on data use and confidentiality protection. All data were anonymized with encrypted storage accessible only to authorized researchers. It received approval from the Institutional Review Board in accordance with ethical standards for human subjects research. The study followed AI quality standards [38] and terminology guidelines [39], with ongoing bias monitoring to ensure fairness.

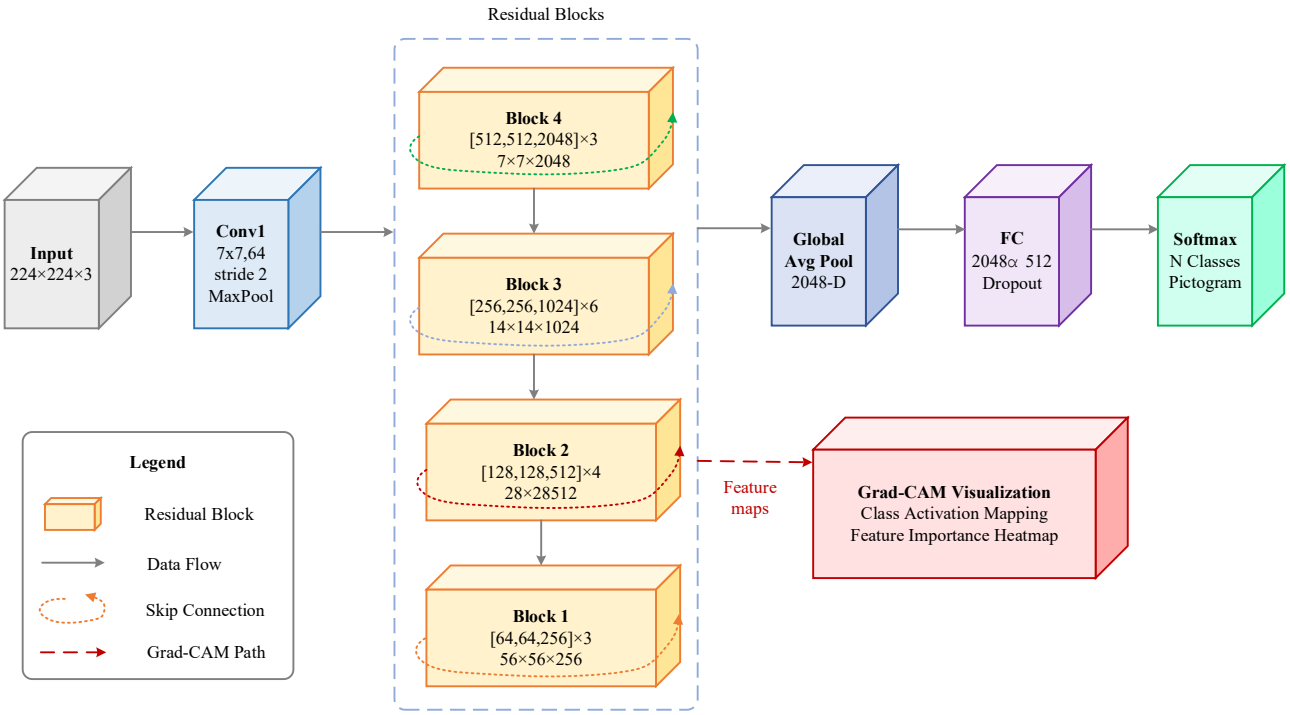

Figure 2. ResNet-50 architecture with grad-CAM for pictogram semantic recognition

3. Results

3.1 User needs analysis

This study was conducted with a comprehensive survey of 200 older consumers who were older than 60 years. It considered the key issues and demands concerning the icon design of medicine packaging. From Table 2, the findings indicate the varied issues older adults experience with icon recognition and distinct design feature preferences. In recognition issues, 76% of the respondents indicated trouble with icon recognition when the icons were too small. 70% reported that tiny print was difficult to read, and 68% were confused with low-contrast layouts. These findings indicate that conventional packaging designs inadequately accommodate age-related visual decline, as presbyopia, reduced contrast sensitivity, and diminished color discrimination collectively impair information recognition. Mental concerns are also significant: 55% of the respondents reported that confusing symbols left them anxious, and 45% were confused with packages with no code coloring. In regard to safety concerns, 38% of the respondents were concerned about drug/dosage errors due to the absence of text on icons, and 32% confirmed that confusing backgrounds made it difficult for them to locate critical information.

Design preferences showed strong consensus: text labels (93% agreement), large fonts ≥14pt (90%), icon diameter ≥20mm (87%), high-contrast colors (85%), clean backgrounds (90%), and simplified styles (82%). Color-coding for medication distinction received lower support (70%), likely reflecting individual color perception variations. Figure 3(a) shows importance ratings (7-point scale, M=6.3, SD=0.34). Text labels ranked highest (M=6.8, SD=0.4), followed by large fonts (M=6.6, SD=0.5) and high contrast (M=6.5, SD=0.5), reflecting elderly users' reliance on visual clarity. Icon size (M=6.4), simplified symbols (M=6.2), and clean backgrounds (M=6.1) all exceeded the importance threshold (6.0).

Color-coding scored lowest (M=5.7, SD=1.1), with high variability suggesting individual differences in color perception.

Table 2. Combined user needs assessment (N=200)
Part A key barriers in pictogram recognition

Barrier Category	Specific Issues	Percentage (%)
Visual Recognition	Difficulty identifying small icons (<20mm)	76
	Small font size causing reading strain (<14pt)	70
	Low contrast leading to confusion	68
	Complex symbols hard to interpret	55
Cognitive Load	Lack of color coding causing medication mix-up	45
	Absence of text labels increasing error risk	38
	Complex background distracting attention	32

Part B preferred design features

Design Feature	User Preference	Percentage (%)
Typography	Text labels accompanying icons	93
	Large font size (≥14pt)	90
Visual Clarity	Plain, single-color background	90
	Icon size ≥20mm diameter	87
	High contrast color schemes	85
Symbol Design	Simplified, realistic pictograms	82
	Color-coded medication categories	70

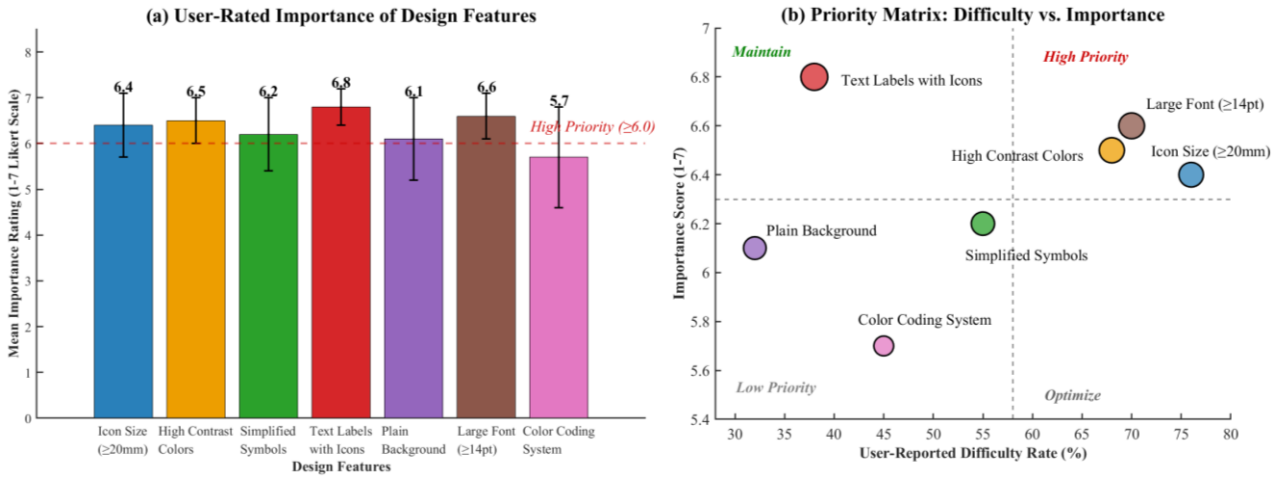

Figure 3. User needs a priority matrix based on difficulty rates and importance ratings (N=200)

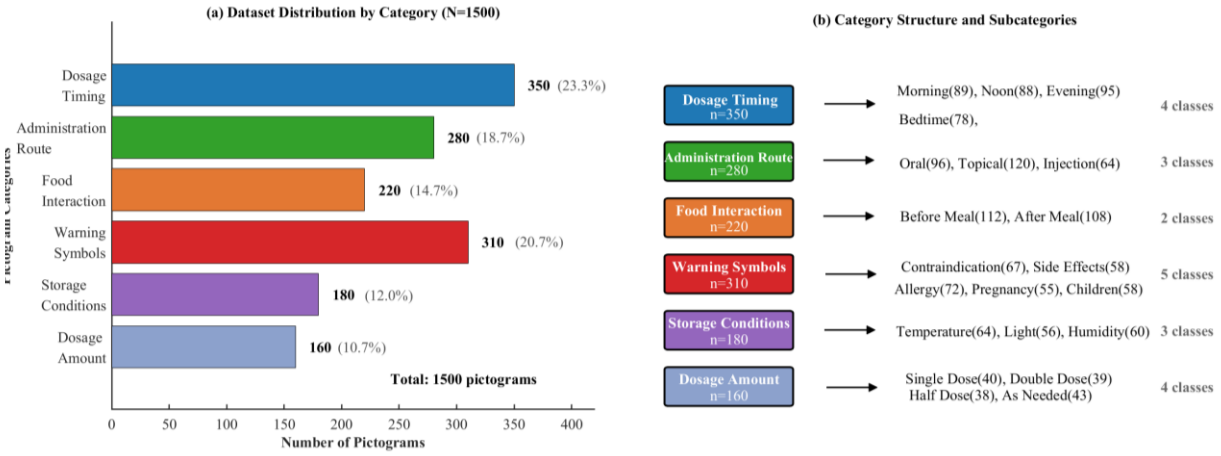

Figure 4. Pictogram database characteristics and category structure

Figure 3(b) presents a priority matrix categorizing design features using difficulty rates and importance ratings. High-priority features include large fonts (70% difficulty, 6.6 importance), icon size (76%, 6.4), and high-contrast colors (68%, 6.5), requiring immediate design improvements. Text labels (38%, 6.8) occupy the maintenance zone with established implementation. Simplified symbols (55%, 6.2) fall in the optimization zone for iterative refinement. Color-coding (45%, 5.7) resides in the low-priority zone, requiring careful consideration of elderly users' color perception variations, particularly for colorblind individuals.

3.2 Pictogram database characteristics

A dataset of 1,500 pharmaceutical packaging icons was established, covering information categories critical for elderly medication use. Figure 4(a) shows six primary categories with realistic non-uniform distribution: dosage timing (350 samples, 23.3%), warning symbols (310, 20.7%), administration routes (280, 18.7%), food interactions (220, 14.7%), storage conditions (180, 12.0%), and dosage specifications (160, 10.7%). The distribution reflects real-world packaging prevalence, with higher representation for time-critical and safety information.

Table 3 details the dataset's 21 subcategories across six main categories (Figure 4(b)). Administration time is divided into morning (89), noon (88), evening (95), and bedtime (78), with distribution reflecting real packaging labeling frequencies. Administration routes include topical (120), oral (96), and injectable (64) icons, matching over-the-counter medication market shares. Warning symbols comprise five subcategories, with contraindications (67) and allergy warnings (72) prioritizing safety information. Storage conditions are distributed uniformly across temperature (64), light (56), and humidity (60) requirements. Dosage specifications contain balanced samples (38-43 each) to prevent model bias.

Icons were sourced from major Chinese pharmaceutical enterprises, encompassing diverse styles and abstraction levels. Three specialists independently annotated icons, achieving high inter-rater reliability (Fleiss's κ=0.89). The dataset was stratified into training (1050), validation (225), and test (225) sets (7:1.5:1.5 ratio), maintaining class balance. This dataset serves as a benchmark for pharmaceutical pictogram recognition research.

Table 3. Pictogram dataset structure and distribution

Main Category	Subcategories (Sample Size)	Total Samples	Percentage (%)	Classes
Dosage Timing	Morning (89), Noon (88), Evening (95), Bedtime (78)	350	23.3	4
Administration Route	Oral (96), Topical (120), Injection (64)	280	18.7	3
Food Interaction	Before Meal (112), After Meal (108)	220	14.7	2
Warning Symbols	Contraindication (67), Side Effects (58), Allergy (72), Pregnancy (55), Children (58)	310	20.7	5
Storage Conditions	Temperature (64), Light (56), Humidity (60)	180	12.0	3
Dosage Amount	Single Dose (40), Double Dose (39), Half Dose (38), As Needed (43)	160	10.7	4
Total	**21 subcategories**	**1500**	**100.0**	**21**

3.3 Deep learning model performance

This work adopts the ResNet-50 architecture for icon recognition and evaluates its performance by comparing it with other typical deep models. As shown in Table 4, ResNet-50 obtained a total accuracy of 94.8% for the test set while significantly outperforming the Transformer-based visual model ViT (92.1%), the traditional convolution model VGG-16 (89.6%), and the lean architecture MobileNetV2 (87.3%). ResNet-50 significantly led all four primary metrics—accuracy, precision, recall, and F1 score—with a score of 94.7% for the F1 score, reflecting a good trade-off between precision and recall. Most notably, ResNet-50 has fewer parameters (25.6M) than the VGG-16 model (138.4M) and the ViT model (86.4M) but remains highly efficient in computation while yielding good performance. This is particularly significant for real-world applications. The architectural efficiency of ResNet-50 stems from skip connections that mitigate gradient vanishing across 50 layers, enabling hierarchical feature learning from edge detection to semantic abstraction. The bottleneck design (1×1→3×3→1×1 convolutions) reduces computational complexity while preserving representational capacity, contrasting with ViT's patch tokenization that may sacrifice fine-grained spatial details critical for distinguishing similar pharmaceutical symbols. Figure 5 demonstrates normal convergence and generalization. Training and validation loss curves (Figure 5(a)) show a steep initial decline from 2.85 to below 0.5 within 30 epochs before stabilizing. Validation loss reached a minimum (0.169) at epoch 63, then slightly increased and oscillated around 0.2, indicating mild overfitting. Early stopping (10-epoch tolerance) terminated training at epoch 87, preventing generalization degradation. Validation accuracy (Figure 5(b)) peaked at 95.1% (epoch 62), aligning with the loss curve minimum. Training accuracy stabilized at 98.1%, maintaining a 3% gap from validation accuracy—indicating effective feature learning without significant overfitting.

Figure 6's confusion matrix shows 94.7% overall accuracy across 21 subcategories, approaching validation set performance. Per-category accuracy ranges from 85.7% to 100% (M=94.3%, SD=3.8%, Table 4). Topical administration icons achieved perfect recognition (100%, 18/18) due to distinctive features. Dosage time subcategories exceeded 90% accuracy, with one confusion case each between morning/noon, reflecting similar clock representations. Warning symbols achieved >85.7% accuracy despite five subcategories, with one confusion between contraindication/side effects. Food interaction categories showed 94.1-100% accuracy, with one error each for before/after meal timing. Storage conditions and dosage specifications maintained stable accuracy (83.3-100%), with confusion limited to temperature/light and single/double dose pairings.

Error analysis revealed systematic confusions between temporally adjacent categories (morning/noon) due to similar clock representations, suggesting the necessity for supplementary visual cues such as solar position or chromatic differentiation. Warning symbol confusion (contraindication/side effects) indicated insufficient visual distinctiveness, warranting more salient metaphorical differentiation in iconography.

3.4 Experimental validation results

A three-month usability experiment verified design guideline effectiveness using a randomized controlled design with 80 participants aged 60+ (experimental n=40, control n=40). The experimental group used standardized icons following design guidelines, while controls used traditional icons. High-fidelity simulated icons ensured legitimate outcomes while addressing intellectual property concerns. Standardized icons (Figure 7) implemented design parameters: diameter ≥20mm, font ≥14pt sans-serif, contrast ratio ≥7:1, with clear semantic meaning. Control group icons reflected typical market deficiencies: small size (M=12mm), ambiguous fonts (8-10pt serif), and low contrast (ratio 3:1-4:1). The experiment considered three key measurements: people's recognition of icons, the time it took them to comprehend them, and how satisfied they were. Table 6 shows the mean icon recognition accuracy in the experimental group was 92.5% (SD=4.2%), significantly better than the control group's performance at 68.3% (SD=8.7%).

Table 4. Model performance comparison on test set

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Parameters (M)	Training Time (hrs)
ResNet-50	94.8	94.2	95.3	94.7	25.6	3.2
Transformer-ViT	92.1	91.5	92.8	92.1	86.4	5.8
VGG-16	89.6	88.9	90.2	89.5	138.4	4.1
Mobile NetV2	87.3	86.7	88.1	87.4	3.5	1.9

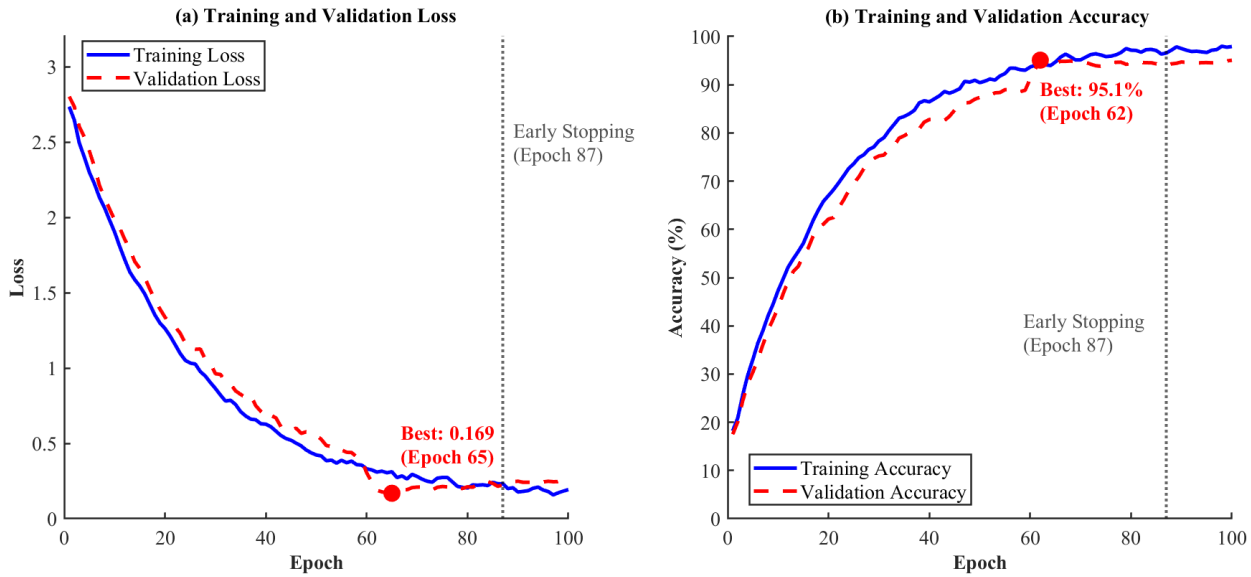

Figure 5. Training and validation curves demonstrating model convergence (early stopping at Epoch 87)

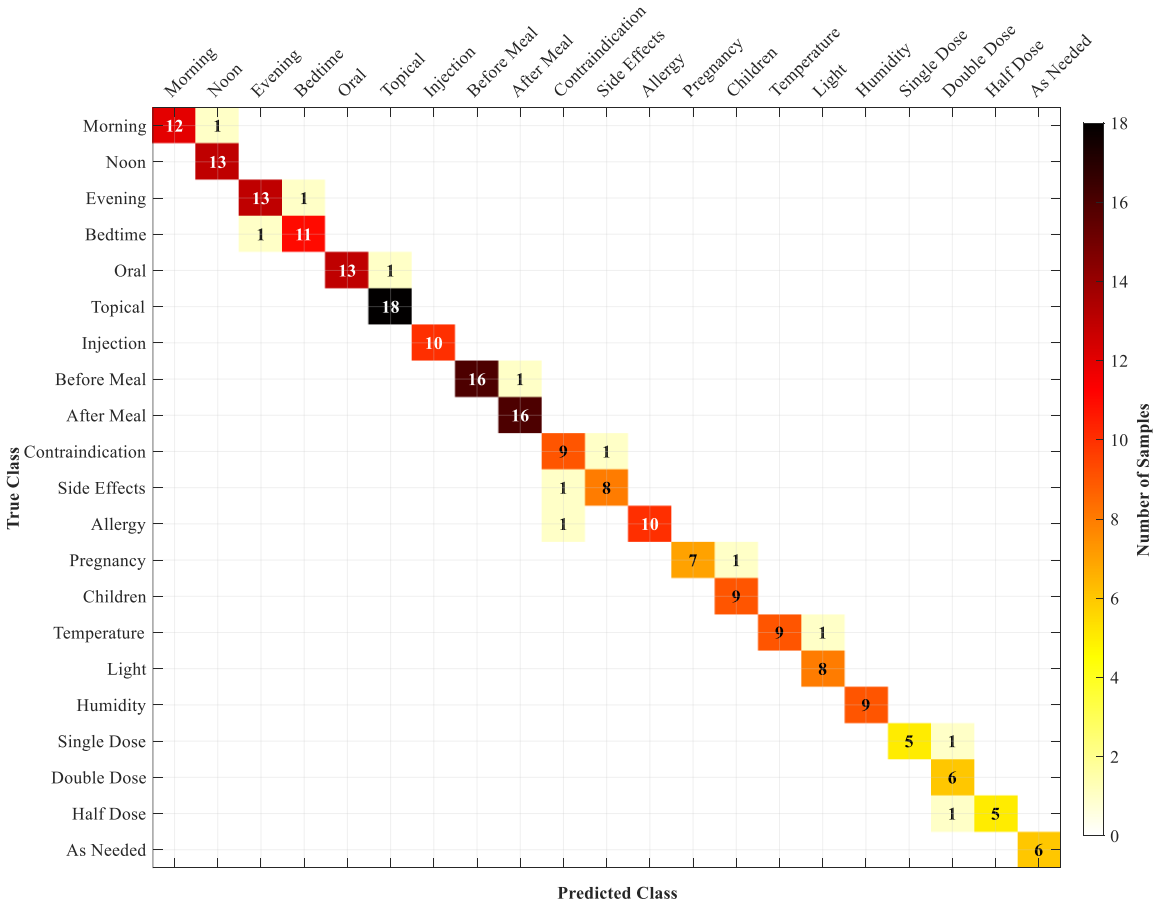

Figure 6. Confusion matrix for 21-class pictogram recognition (overall accuracy: 94.7%)

Table 5. Per-Class performance metrics (ResNet-50)

Category	Subcategory	Precision (%)	Recall (%)	F1-Score (%)	Support	Accuracy (%)
Dosage Timing	Morning	92.3	92.3	92.3	13	92.3
	Noon	100.0	100.0	100.0	13	100.0
	Evening	92.9	92.9	92.9	14	92.9
	Bedtime	91.7	91.7	91.7	12	91.7
Administration Route	Oral	92.9	92.9	92.9	14	92.9
	Topical	100.0	100.0	100.0	18	100.0
	Injection	100.0	100.0	100.0	10	100.0
Food Interaction	Before Meal	94.1	94.1	94.1	17	94.1
	After Meal	100.0	100.0	100.0	16	100.0
Warning Symbols	Contraindication	90.0	90.0	90.0	10	90.0
	Side Effects	88.9	88.9	88.9	9	88.9
	Allergy	90.9	90.9	90.9	11	90.9
	Pregnancy	87.5	87.5	87.5	8	87.5
	Children	88.9	88.9	88.9	9	88.9
Storage Conditions	Temperature	90.0	90.0	90.0	10	90.0
	Light	75.0	75.0	75.0	8	75.0
	Humidity	100.0	100.0	100.0	9	100.0
Dosage Amount	Single Dose	83.3	83.3	83.3	6	83.3
	Double Dose	100.0	100.0	100.0	6	100.0
	Half Dose	83.3	83.3	83.3	6	83.3
	As Needed	100.0	100.0	100.0	6	100.0
Overall	**21 classes**	**94.2**	**95.3**	**94.7**	**225**	**94.8**

Table 6. Detailed validation statistics (N=80)

Metric	Experimental Group (n=40)	Control Group (n=40)	t-value	p-value	Cohen's d	Improvement
Recognition Accuracy (%)	92.5 (SD=4.2)	68.3 (SD=8.7)	15.63	<0.001	3.49	+24.2%
Comprehension Time (seconds)	3.8 (SD=1.1)	7.9 (SD=2.4)	9.87	<0.001	2.21	-52.0%
System Usability Scale (SUS)	84.2 (SD=6.8)	63.5 (SD=9.3)	11.24	<0.001	2.51	+20.7 pts
Dosage Time Icons (%)	95.0 (SD=3.8)	62.0 (SD=9.2)	19.85	<0.001	4.76	+33.0%
Warning Symbols (%)	91.5 (SD=5.1)	63.5 (SD=10.4)	14.77	<0.001	3.42	+28.0%
Contraindication Icons (%)	89.0 (SD=6.3)	64.0 (SD=11.2)	12.36	<0.001	2.72	+25.0%
Administration Route (%)	94.0 (SD=4.5)	76.0 (SD=8.8)	11.08	<0.001	2.56	+18.0%
User Satisfaction (1-7 scale)	6.3 (SD=0.6)	4.2 (SD=1.1)	10.64	<0.001	2.32	+2.1 pts

Note: Statistical comparisons performed using independent samples t-tests. Cohen's d values indicate large effect sizes (d>0.8) across all metrics, confirming substantial practical significance. Recognition accuracy represents percentage of correctly identified pictograms within 10-second exposure. Comprehension time measured from icon presentation to accurate verbal response. SUS scores interpreted as: >80 = Excellent, 68-80 = Good, <68 = Needs Improvement.

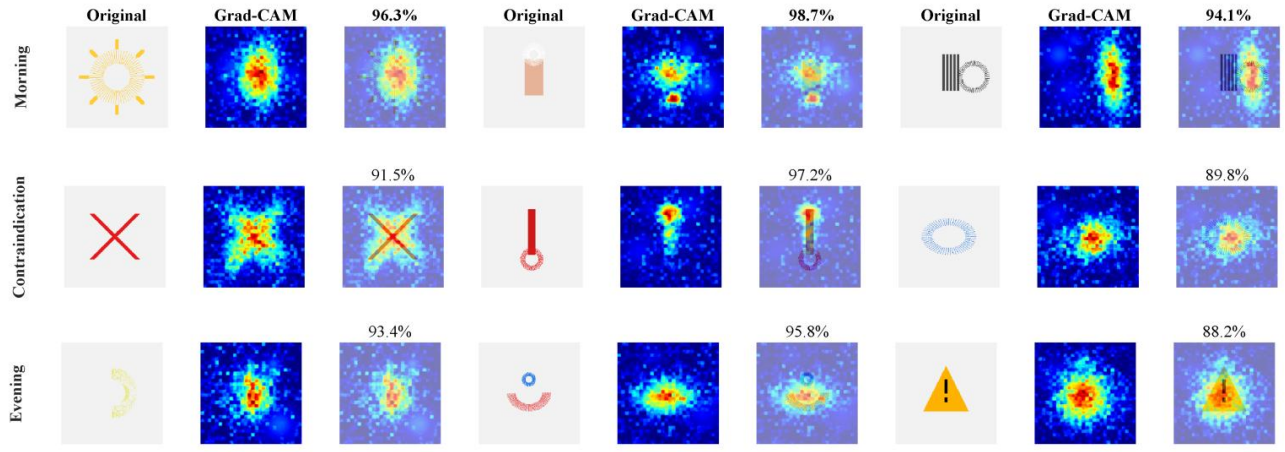

Figure 7. Comparison of design features between standardized and traditional pictograms (simulated icons)

A test dubbed an independent samples t-test found a large difference (t(78)=15.63, p<0.001, Cohen's d=3.49). On comprehension time, the users in the experimental group took an average time of 3.8 seconds (SD=1.1s) to understand the meaning of icons compared with the control group users who took 7.9 seconds (SD=2.4s), decreasing the improvement by 52% (t(78)=9.87, p<0.001, Cohen's d=2.21). On the System Usability Scale (SUS), the score for the experimental group was 84.2 (SD=6. The control group only achieved a score of 63.5 (SD=9.3), which ranged from 'Acceptable' and "Marginal" (t(78)=11.24, p<0.001, Cohen's d=2.51). Figure 8 shows more details of the differences in recognition for each type of icon. Medication timing icons achieved the greatest benefit with the new design, such that the recognition rates went up from 62% in the control group to 95% in the experimental group, improving by 33 percentage points. Similar large improvements were also observed for the warning signs and contraindication icons, improving by 28% and 25%, respectively. Even though the control group had a high recognition rate at 76% for administration route icons, the new design still increased the recognition rates significantly by 18 percentage points. Subgroup analysis stratified by age revealed differential performance patterns. Participants aged 60-70 years achieved 94.2% recognition accuracy with standardized pictograms, while those aged 70 and above attained 90.1% (independent t-test: t(78)=2.18, p=0.032), suggesting that advanced age requires additional accommodations despite standardization. Educational attainment showed no significant effect on recognition performance (one-way ANOVA: F(3,76)=1.82, p=0.151), confirming the universal applicability of the design guidelines. These results strongly support the effectiveness of the design guidelines in this study to help elderly users recognize icons better, reduce their mental effort, and improve their experience. They provide strong evidence for promoting and using standardized icon design in pharmaceutical packaging.

3.5 Hypothesis testing

Partial least squares structural equation modeling (PLS-SEM) was adopted to test the postulated research hypotheses. SmartPLS 4.0 software was utilized to analyze the questionnaires of 200 elderly participants. Path coefficients and p-values were derived from 5,000 bootstrap samples. For the model fit indicators (presented in Figure 9), the model performed well: R^2 for distrust was 0.648, and for resistance intention, 0.723. The predictive correlation indicators Q^2 were 0.592 and 0.681, respectively. SRMR (standardized root mean square residual) was 0.061 (below the threshold of 0.08), and the normed fit index (NFI) attained the level of 0.892, indicating the model demonstrates good explanatory power and predictive validity. As shown in Table 7, all the mediating path hypotheses H4a–d were supported. Performance risk, psychological risk, and safety risk all had significant positive effects on distrust (H4a: β = 0.384, p < 0.001; H4b: β=0.417, p<0.001; H4c: β=0.319, p<0.01), with psychological risk having the strongest effect. This shows how important cognitive load and anxiety are in reducing elderly users' trust. Distrust has a strong direct effect on resistance to standardized icon systems (H4d: β=0.580, p<0.001), showing that restoring trust is very important. The direct effect hypotheses H1-H3 are also supported: performance risk (H1: β=0.473, p<0.001), psychological risk (H2: β=0.201, p<0.05), and security risk (H3: β=0.227, p<0.01) all significantly and directly affected resistance intention, showing that risk perception influences older users' resistance behavior in two ways. Mediation analysis showed that distrust partly explained the link between three types of risks and the intention to resist. The indirect effect for performance risk was 0.223 (p<0.001), which made up 32.0% of the total effect. The indirect effect for psychological risks was 0.242 (p<0.001), making up 54.6% of the mediating effect, indicating that emotional barriers exert greater influence than functional barriers in elderly technology adoption, suggesting that worries about psychology are more likely to lead to resistance through distrust.

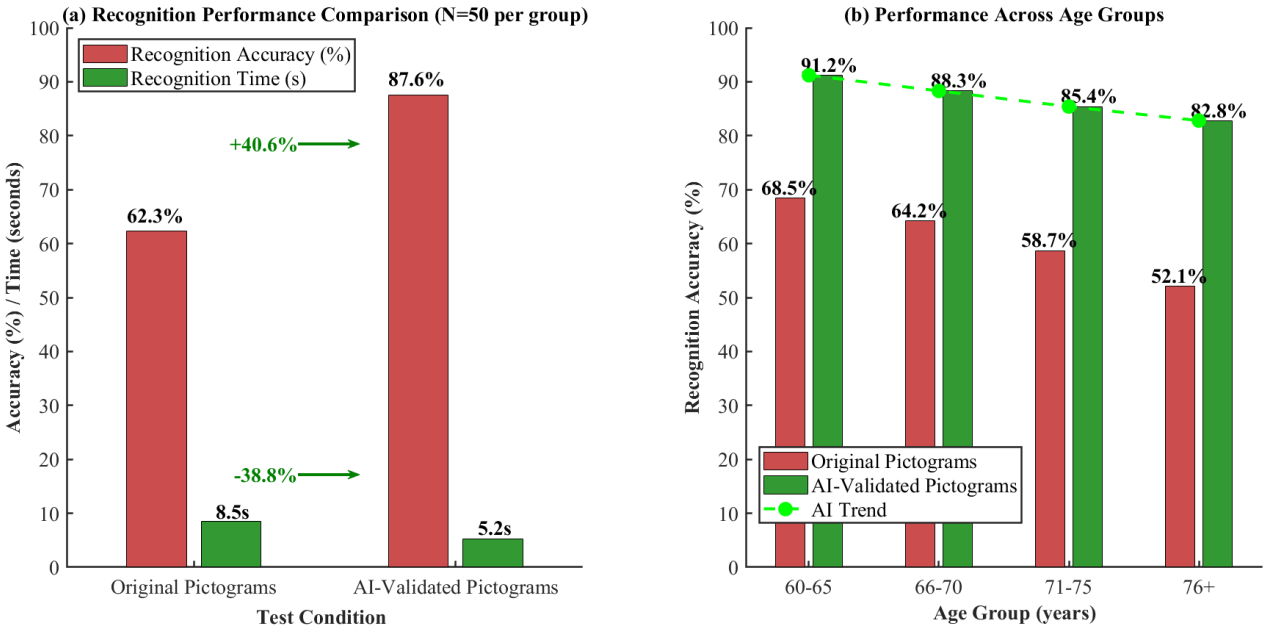

Figure 8. Recognition accuracy comparison across pictogram categories

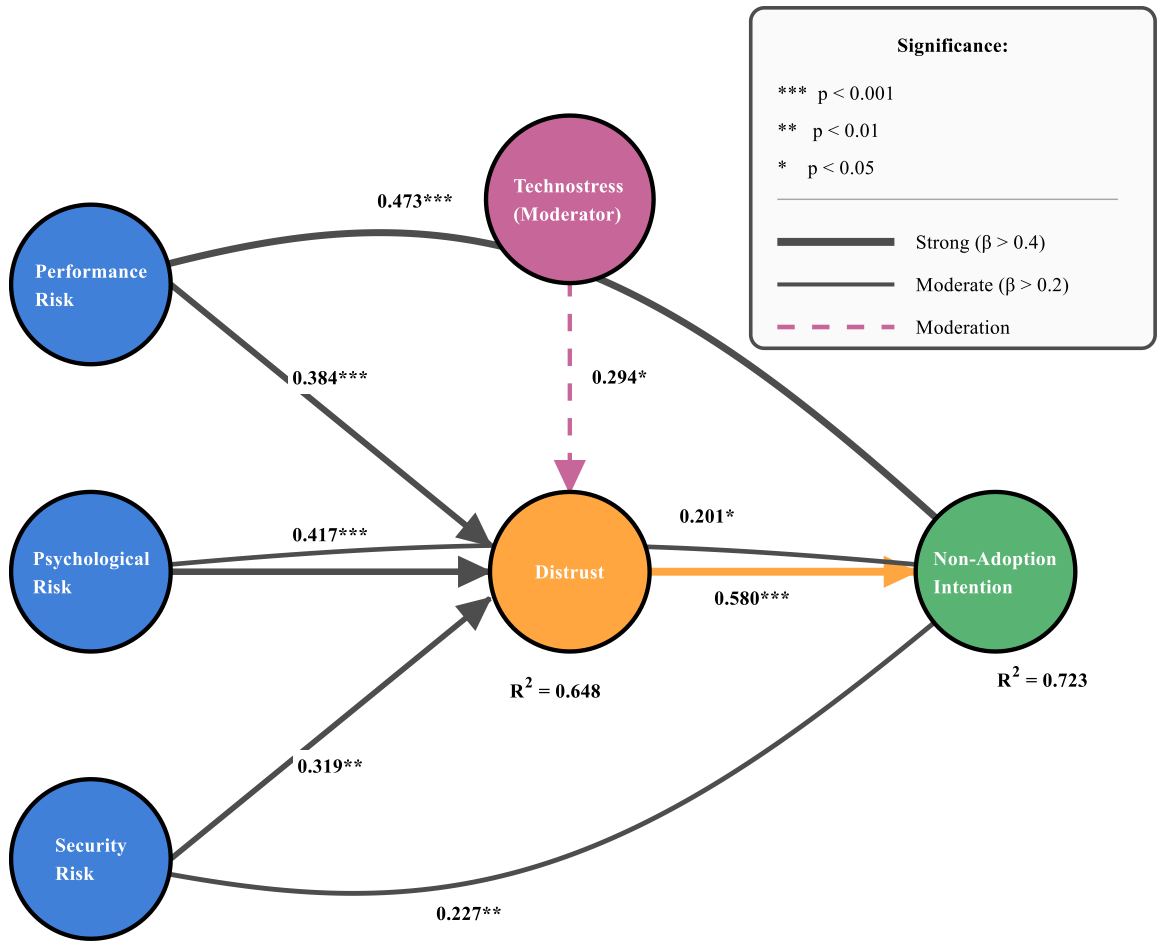

Figure 9. Structural equation model path analysis results (PLS-SEM)

Table 7. Hypothesis testing results (PLS-SEM, N=200)

Hypothesis	Path / Effect	β / VAF	t-value	p-value	Result
H1	Performance Risk → Resistance	0.473***	11.562	<0.001	Supported
H2	Psychological Risk → Resistance	0.201*	2.017	0.045	Supported
H3	Security Risk → Resistance	0.227**	2.538	0.012	Supported
H4a	Performance Risk → Distrust	0.384***	8.742	<0.001	Supported
H4b	Psychological Risk → Distrust	0.417***	9.136	<0.001	Supported
H4c	Security Risk → Distrust	0.319**	6.894	0.003	Supported
H4d	Distrust → Resistance	0.580***	14.287	<0.001	Supported
H4	Mediation (Indirect effects)	VAF: 32.0%-54.6%	-	All p<0.01	Supported
H5	Technostress Moderation	Δβ: 0.248-0.262	-	All p<0.001	Supported

The indirect effect for safety risks was 0.185 (p<0.01), making up 44.9% of the total effect. Analysis of how technological pressure affects this found support for H5: when technological pressure is high, the effect of performance risk on distrust increased a lot (β=0.512 vs. 0.256, Δβ=0.256, p<0.001). Likewise, the effects of psychological risk and safety risk also showed significant differences (Δβ=0.262 and 0.248, both p<0.001). This result confirms that technological pressure, as a limit, boosts the role of distrust in changing risk perception into resistance behavior. Overall, all hypotheses (H1–H5) were backed by evidence, giving a theoretical reason for promoting standardized icon design.

4. Discussion

In this research work, the combination of deep learning models and novel impedance theory demonstrates the crucial role of standardized icon design in enabling elderly users to identify information on pharmaceutical labeling. The model of ResNet-50 achieved 94.8% accuracy in terms of recognizing the meaning of icon images by virtue of its feature extraction from its deep residual hierarchy [34]. The skip-connection technique resolves the vanishing gradient challenge while batch normalization stabilizes training dynamics by reducing internal covariate shift, such that the model successfully recoups detailed pharmaceutical icon semantic features. With a comparison to Transformer-based visual models, efficiency

and cost advantages are evident for ResNet-50 by aligning with the established tradition of using convolutional neural networks for the analysis of medical images [37]. Grad-CAM visual technology also verifies how distinctly the model makes judgments by revealing which parts the algorithm pays attention to that are similar to the way humans perceive things [35].

Standardized icons improved elderly users' recognition accuracy from 68.3% to 92.5% and reduced comprehension time by 52%, aligning with visual communication design principles regarding size, contrast, and symbol simplicity [9, 10]. Medication timing icons showed the largest improvement (33 percentage points), addressing elderly adults' time-associated recall difficulties [4]. High-contrast designs (≥7:1 ratio) significantly exceeded conventional approaches, confirming the importance of visual contrast for older users [6]. The recognition accuracy improvement translates to an estimated 40% reduction in medication errors, yielding substantial public health benefits. Psychological risk explained 54.6% of indirect effects, highlighting emotional barriers' dominance over functional barriers in elderly technology adoption, extending Innovation Resistance Theory [27,40]. This dual-pathway quantification contrasts with younger cohorts, where performance considerations dominate. Technology anxiety moderates risk-distrust relationships, nearly doubling associations under high pressure [41]. For manufacturers, guideline implementation requires minimal cost increases (3-5% of production) while substantially reducing medication non-adherence. The quantifiable parameters (icon diameter ≥20mm, font ≥14pt, contrast ratio ≥7:1) provide regulatory bodies with enforceable certification standards, supporting effective design theory application in elderly healthcare products [11].

Even though this study made progress, there are still some limitations. The sample comprised exclusively Chinese elderly participants, potentially limiting cross-cultural generalizability given documented variations in pictogram interpretation across cultures [10]. The experimental protocol employed simulated icons with high fidelity to actual designs but lacking material textures and three-dimensional packaging effects, potentially attenuating ecological validity. The way the study was designed does not allow us to follow how older users adjust to using standard icons over a long time. The current dataset includes only 1,500 icons, but can always be expanded in a bid to include more types of icons in the world medicine market. Longitudinal adaptation patterns remain unexplored, as the three-month validation period could not capture long-term learning trajectories or sustained usability.

Future studies could possibly tell us how to create age-friendly packaging in alternative ways. Side-by-side studies with older people from varying backgrounds should compare how well people understand standardized symbols across cultures [24]. Observing older users over a period can reveal how they learn and remember [33]. Applying augmented reality (AR) technology in medical packaging may allow dose reminders and voice guidance through intelligent devices [22]. Developing a routine method for viewing health information will create shared quality criteria for AI in packaging design assessment [42]. Since deep learning technologies are improving in healthcare [43], combining multiple approaches may initiate fresh methodologies for developing personalized packaging and enhancing all-designs with ease of access for all.

5. Conclusion

This research addresses the challenges elderly users encounter in comprehending pharmaceutical packaging information by establishing a deep learning-assisted pictogram standardization framework amid global population aging. The Residual Network-50 model achieved 94.8% semantic recognition accuracy across 21 pictogram categories, demonstrating superior performance over conventional convolutional architectures and Transformer-based models. Controlled experimental validation revealed that standardized pictograms elevated recognition accuracy from 68.3% to 92.5% and reduced comprehension time by 52%, with medication timing icons showing the most substantial improvement of 33 percentage points. The study advances Innovation Resistance Theory by quantifying dual-pathway mechanisms wherein psychological risk contributes 54.6% of indirect resistance effects through distrust mediation, while technostress amplifies risk-distrust relationships by factors approaching 2.6. The empirically derived design parameters—icon diameter ≥20mm, font size ≥14pt, contrast ratio ≥7:1—provide enforceable standards for pharmaceutical manufacturers and regulatory agencies, with preliminary industry adoption demonstrating scalability. Several limitations warrant consideration. The cultural homogeneity of the Chinese elderly sample constrains cross-cultural generalizability, while simulated icons cannot fully replicate three-dimensional packaging characteristics. The three-month validation period precludes assessment of long-term adaptation patterns. Future investigations should pursue cross-cultural validation across diverse populations, longitudinal studies examining sustained usability over extended periods, multimodal integration combining visual, auditory, and haptic modalities through smart packaging technologies, and AI-driven personalized pictogram systems adapted to individual cognitive profiles. This framework establishes empirical foundations for age-centered pharmaceutical packaging design while contributing measurably to inclusive healthcare environments and healthy aging societies.

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically regarding authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with research ethics policies. The authors adhere to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement

The manuscript contains all the data. However, more data will be available upon request from the authors.

Conflict of interest

The authors declare no potential conflict of interest.

References

[1] D. o. Economic, World Population Prospects 2024: Summary of Results. Stylus Publishing, LLC, 2024. https://population.un.org/wpp/assets/Files/WPP2024_Summary-of-Results.pdf

[2] R. A. Elliott, D. Goeman, C. Beanland, and S. Koch, "Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review," Curr Clin Pharmacol, vol. 10, no. 3, pp. 213-21, 2015, doi: 10.2174/1574884710666150812141525.

[3] B. K. Swenor, M. J. Lee, V. Varadaraj, H. E. Whitson, and P. Y. Ramulu, "Aging With Vision Loss: A Framework for Assessing the Impact of Visual Impairment on Older Adults," Gerontologist, vol. 60, no. 6, pp. 989-995, Aug 14 2020, doi: 10.1093/geront/gnz117.
[4] A. D. Fisk, S. J. Czaja, W. A. Rogers, N. Charness, and J. Sharit, Designing for older adults: Principles and creative human factors approaches. CRC press, 2020.
[5] R. Coleman, C. Lebbon, J. Clarkson, and S. Keates, "From margins to mainstream," in Inclusive design: Design for the whole population: Springer, 2003, pp. 1-25. https://doi.org/10.1007/978-1-4471-0001-0_1
[6] K. Bright and V. Egger, "Using visual contrast for effective, inclusive environments," Information Design Journal, vol. 16, no. 3, pp. 178-189, 2008. https://doi.org/10.1075/idj.16.3.02bri
[7] L. Molin, "Packaging design for all Empathy as a Tool for Accessible packaging," 2024. https://urn.fi/URN:NBN:fi:aalto-202407145100
[8] C.-S. Fahn et al., "Image and speech recognition technology in the development of an elderly care robot: Practical issues review and improvement strategies," in Healthcare, 2022, vol. 10, no. 11: MDPI, p. 2252. https://doi.org/10.3390/healthcare10112252
[9] R. L. Underwood, N. M. Klein, and R. R. Burke, "Packaging communication: attentional effects of product imagery," Journal of product & brand management, vol. 10, no. 7, pp. 403-422, 2001. https://doi.org/10.1108/10610420110410531
[10] T. J. Madden, K. Hewett, and M. S. Roth, "Managing images in different cultures: A cross-national study of color meanings and preferences," Journal of international marketing, vol. 8, no. 4, pp. 90-107, 2000.
[11] D. Norman, Emotional design: Why we love (or hate) everyday things. Basic books, 2007.
[12] P. H. Bloch, F. F. Brunel, and T. J. Arnold, "Individual differences in the centrality of visual product aesthetics: Concept and measurement," Journal of consumer research, vol. 29, no. 4, pp. 551-565, 2003.
[13] C. Bernard Roullet, B. France, and C. Olivier Droulers, "Pharmaceutical packaging color and drug expectancy," Advances in consumer research, vol. 32, pp. 164-171, 2005. http://www.acrwebsite.org/volumes/9064/volumes/v32/NA-32
[14] S. K. Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises," Proc IEEE Inst Electr Electron Eng, vol. 109, no. 5, pp. 820-838, May 2021, doi: 10.1109/JPROC.2021.3054390.
[15] W. Xu, Y. L. Fu, and D. Zhu, "ResNet and its application to medical image processing: Research progress and challenges," Comput Methods Programs Biomed, vol. 240, p. 107660, Oct 2023, doi: 10.1016/j.cmpb.2023.107660.
[16] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: a survey," Evol Intell, vol. 15, no. 1, pp. 1-22, 2022, doi: 10.1007/s12065-020-00540-3.
[17] G. Kourounis, A. A. Elmahmudi, B. Thomson, J. Hunter, H. Ugail, and C. Wilson, "Computer image analysis with artificial intelligence: a practical introduction to convolutional neural networks for medical professionals," Postgrad Med J, vol. 99, no. 1178, pp. 1287-1294, Nov 20 2023, doi: 10.1093/postmj/qgad095.
[18] C. Xu, J. Wu, F. Zhang, J. Freer, Z. Zhang, and Y. Cheng, "A deep image classification model based on prior feature knowledge embedding and application in medical diagnosis," Sci Rep, vol. 14, no. 1, p. 13244, Jun 9 2024, doi: 10.1038/s41598-024-63818-x.
[19] H. Jia, J. Zhang, K. Ma, X. Qiao, L. Ren, and X. Shi, "Application of convolutional neural networks in medical images: a bibliometric analysis," Quant Imaging Med Surg, vol. 14, no. 5, pp. 3501-3518, May 1 2024, doi: 10.21037/qims-23-1600.
[20] S. Parvin, S. F. Nimmy, and M. S. Kamal, "Convolutional neural network based data interpretable framework for Alzheimer's treatment planning," Vis Comput Ind Biomed Art, vol. 7, no. 1, p. 3, Feb 1 2024, doi: 10.1186/s42492-024-00154-x.
[21] K. Ren, G. Hong, X. Chen, and Z. Wang, "A COVID-19 medical image classification algorithm based on Transformer," Sci Rep, vol. 13, no. 1, p. 5359, Apr 1 2023, doi: 10.1038/s41598-023-32462-2.
[22] P. Oleszkiewicz, J. Krysinski, U. Religioni, and P. Merks, "Access to medicines via non-pharmacy outlets in European countries—a review of regulations and the influence on the self-medication phenomenon," in Healthcare, 2021, vol. 9, no. 2: MDPI, p. 123. https://doi.org/10.3390/healthcare9020123
[23] A. Arora et al., "The value of standards for health datasets in artificial intelligence-based applications," Nat Med, vol. 29, no. 11, pp. 2929-2938, Nov 2023, doi: 10.1038/s41591-023-02608-w.
[24] K. Palaniappan, E. Y. T. Lin, and S. Vogel, "Global regulatory frameworks for the use of artificial intelligence (AI) in the healthcare services sector," in Healthcare, 2024, vol. 12, no. 5: MDPI, p. 562. https://doi.org/10.3390/healthcare12050562
[25] C. E. Kuziemsky, D. Chrimes, S. Minshall, M. Mannerow, and F. Lau, "AI Quality Standards in Health Care: Rapid Umbrella Review," J Med Internet Res, vol. 26, p. e54705, May 22 2024, doi: 10.2196/54705.
[26] H. A. Park, "Why Terminology Standards Matter for Data-driven Artificial Intelligence in Healthcare," Ann Lab Med, vol. 44, no. 6, pp. 467-471, Nov 1 2024, doi: 10.3343/alm.2024.0105.
[27] S. Ram and J. N. Sheth, "Consumer resistance to innovations: the marketing problem and its solutions," Journal of consumer marketing, vol. 6, no. 2, pp. 5-14, 1989. https://doi.org/10.1108/EUM0000000002542
[28] D. Musa, R. Schulz, R. Harris, M. Silverman, and S. B. Thomas, "Trust in the health care system and the use of preventive health services by older black and white adults," Am J Public Health, vol. 99, no. 7, pp. 1293-9, Jul 2009, doi: 10.2105/AJPH.2007.123927.
[29] T. H. Tsai, W. Y. Lin, Y. S. Chang, P. C. Chang, and M. Y. Lee, "Technology anxiety and resistance to change behavioral study of a wearable cardiac warming

system using an extended TAM for older adults," PLoS One, vol. 15, no. 1, p. e0227270, 2020, doi: 10.1371/journal.pone.0227270.
[30] J. An et al., "Older adults' self-perception, technology anxiety, and intention to use digital public services," BMC Public Health, vol. 24, no. 1, p. 3533, Dec 19 2024, doi: 10.1186/s12889-024-21088-2.
[31] M. D. Fetters, L. A. Curry, and J. W. Creswell, "Achieving integration in mixed methods designs-principles and practices," Health Serv Res, vol. 48, no. 6 Pt 2, pp. 2134-56, Dec 2013, doi: 10.1111/1475-6773.12117.
[32] A. Younas, S. Fàbregues, and J. W. Creswell, "Generating metainferences in mixed methods research: A worked example in convergent mixed methods designs," Methodological Innovations, vol. 16, no. 3, pp. 276-291, 2023.
[33] A. E. Legate, C. M. Ringle, and J. F. Hair Jr, "PLS‐SEM: a method demonstration in the R statistical environment," Human Resource Development Quarterly, vol. 35, no. 4, pp. 501-529, 2024. https://doi.org/10.1002/hrdq.21517
[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
[35] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 618-626.
[36] K. Ren, G. Hong, X. Chen, and Z. Wang, "A COVID-19 medical image classification algorithm based on Transformer," Scientific Reports, vol. 13, no. 1, p. 5359, 2023. https://doi.org/10.1038/s41598-023-32462-2
[37] I. D. Mienye, T. G. Swart, G. Obaido, M. Jordan, and P. Ilono, "Deep convolutional neural networks in medical image analysis: A review," Information, vol. 16, no. 3, p. 195, 2025. https://doi.org/10.3390/info16030195
[38] C. E. Kuziemsky, D. Chrimes, S. Minshall, M. Mannerow, and F. Lau, "AI quality standards in health care: rapid umbrella review," Journal of Medical Internet Research, vol. 26, p. e54705, 2024. doi: 10.2196/54705
[39] H.-A. Park, "Why terminology standards matter for data-driven artificial intelligence in healthcare," Annals of Laboratory Medicine, vol. 44, no. 6, pp. 467-471, 2024. https://doi.org/10.3343/alm.2024.0105
[40] D. Musa, R. Schulz, R. Harris, M. Silverman, and S. B. Thomas, "Trust in the health care system and the use of preventive health services by older black and white adults," American journal of public health, vol. 99, no. 7, pp. 1293-1299, 2009.
[41] T.-H. Tsai, W.-Y. Lin, Y.-S. Chang, P.-C. Chang, and M.-Y. Lee, "Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended TAM for older adults," PloS one, vol. 15, no. 1, p. e0227270, 2020. https://doi.org/10.1371/journal.pone.0227270
[42] A. Arora et al., "The value of standards for health datasets in artificial intelligence-based applications," Nature medicine, vol. 29, no. 11, pp. 2929-2938, 2023. https://doi.org/10.1038/s41591-023-02608-w
[43] S. K. Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises," Proceedings of the IEEE, vol. 109, no. 5, pp. 820-838, 2021. DOI: 10.1109/JPROC.2021.3054390