International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 334-347

https://doi.org/10.70517/ijhsa46324

Safety Assessment and Design Improvement Strategy of Children's Products Based on Image Recognition Algorithm in the Digital Era

Jun Wang[1] and Xin Guo[1,*]
[1] School of Industrial Design, Hubei University of Technology, Wuhan, Hubei, 430068, China
Corresponding authors: (e-mail: 13173353393@163.com).

Abstract The safety of children's products is closely related to children's health and life safety, and is an important part of the design process that cannot be ignored. This paper takes children's products as the research object, and explores the method of children's product safety assessment from the perspective of image feature recognition and classification. The images of children's products are preprocessed by geometric transformation, grayscaling and image enhancement to extract their color, texture and shape features. Relying on the image features of children's products, a VGG-based image feature recognition model of children's products is constructed, and the model performance is improved by the increase of residual module and the equivalent conversion of multi-branching model into a single-path model, so as to realize the safety assessment of children's products. In the safety assessment experiments of children's products, the model in this paper achieves optimal results in four aspects, namely, accuracy, precision, recall and F1 value, which reach 98.34%, 98.34%, 98.33% and 98.34%, respectively. Compared with the original ResNet18 model, the model also has better recognition accuracy in the face of different categories of images of children's products, and can play an effective role in the work of children's product safety assessment.

Index Terms VGG network, image recognition, image feature extraction, children's products, safety assessment

I. Introduction

Children's products are products manufactured to meet the needs of children's life, learning, and games, such as toys, school supplies, bedding, etc. [1], [2]. Since children's physical development is not yet mature and their ability to adapt to the external environment is weak, the safety requirements for children's products are higher, so it is evident that the safety assessment of children's products is an important link that cannot be ignored for children's healthy growth [3]-[5].

Children's product safety standards need to provide for the safety of children's toys, children's beds, children's strollers, infant safety seats, children's bicycles and other children's products [6], [7]. Among them, it is required that the product design, material selection, process manufacturing and other aspects of strict compliance with national standards, the product must be through the relevant testing and certification, qualified before sales [8], [9]. In the use of products, but also requires products equipped with instructions to remind consumers to comply with the product use specifications, to ensure that the product is safe and reliable [10]. At the same time, children's products in the sales process should also comply with the requirements of relevant laws and regulations, to prevent the emergence of false propaganda, unqualified products and other issues, to protect the health and safety of children consumers [11-13]. A series of safety performance of children's products are systematically evaluated and tested to determine whether the products can meet the safety standards and requirements [14], [15]. Through product safety assessment, safety risks in the product design and manufacturing process can be identified, and these risks can be prevented and controlled to improve the safety performance of the product [16], [17]. Meanwhile, product safety assessment can also provide improvement strategies for product safety use guidelines, product design, and product safety warnings to protect children's safety rights and interests [18], [19].

The safety of children's products is directly related to children's physical health and has an importance that cannot be ignored. In order to realize the safety assessment of children's products, this paper first preprocesses the images of children's products, carries out unified geometric transformation operations, improves the processing speed of images through image grayscaling, and utilizes the image enhancement method to make the blurred images clear and highlight the image features. After completing the image preprocessing work, the features of color, texture and shape of children's product images are extracted. On the basis of preprocessing and feature extraction of children's product images, an improved children's product image feature recognition model of VGG is constructed, which

realizes the assessment of children's product safety by recognizing and classifying the features of children's product images. It adds two residual modules to the basic structure of VGG to alleviate the problem of gradient vanishing and obtain a more robust feature representation. A 1×1 convolutional kernel is introduced to enhance the fitting ability of the model, and model reparameterization is used to convert the improved multi-branch model into a single path model to improve the inference speed. The performance test of the improved VGG model in this paper is carried out, and children's product safety assessment experiments are carried out to test its effectiveness in children's product safety identification and classification work. Finally, the safety of children's products is considered from the perspective of product design, and the safety design strategy of children's products is proposed to fully protect the safety of children.

II. Image pre-processing and feature extraction for children's products

With the continuous improvement of people's living standards, toys have been transformed from the former luxury goods to the necessities of most children today. However, because children do not have the ability to recognize, prevent and avoid danger, the quality and safety of some modern toy products are directly related to children's health and safety. In the development and design of toys, it is necessary to consider the nature and characteristics of children's toys to ensure that children are safe and happy in the process of playing.

In this paper, the safety of children's products will be evaluated based on image recognition algorithms. This chapter will start with preprocessing and feature extraction of children's product images.

II. A.Pre-processing of images of children's products

The image preprocessing methods used in this paper for children's products are: image several transforms, image gray scale change, and image enhancement.

II. A. 1) Geometric transformations of images

Image geometric changes are generally located in the first step of image processing, commonly used image geometric transformations are panning, scaling, mirroring, rotating, transposing, and so on.

Image panning. Image panning can be understood as deleting image pixels in a set direction and letting the pixels in the opposite direction pan to the vacant position, while the blank space left after panning is filled with black pixels or white pixels. For example, a pixel in the image has coordinates (x_0, y_0), then the coordinates after $(\Delta x, \Delta y)$ panning are (x_1, y_1), and the formula for the panning transformation is equation (1):

$$\begin{cases} x_1 = x_0 + \Delta x \\ y_1 = y_0 + \Delta y \end{cases} (x_2, y_2) \tag{1}$$

Image Scaling. Assuming that the scaling factor along the x-axis is k_x and the scaling factor along the y-axis is k_z, the relationship between the points in the image (x_1, y_1) and the original pixel point coordinates after doing the scaling transformation is shown in equation (2):

$$\begin{cases} x_1 = x_0 \times k_x \\ y_1 = y_0 \times k_y \end{cases} \tag{2}$$

Image Mirroring. Image mirroring has horizontal mirroring and vertical mirroring, if the lower left corner of the image as the origin (0, 0), h for the height of the picture, w for the length of the picture, then the original picture to $y = h/2$ this straight line symmetry over the picture is the original picture of the vertical mirror, the original picture to $x = w/2$ as the axis of symmetry over the picture is the original picture of the horizontal mirror, the original image to $x = w/2$ as the axis of symmetry.

Assuming that the coordinates of a pixel point in the image are (x_0, y_0), and the corresponding coordinates are (x_1, y_1) after the mirroring transformations, the vertical mirroring transformations and the horizontal mirroring transformations are shown in equations (3) and (4):

$$\begin{cases} x_1 = x_0 \\ y_1 = h - y_0 \end{cases} \tag{3}$$

$$\begin{cases} x_1 = w - x_0 \\ y_1 = y_0 \end{cases} \tag{4}$$

Image Rotation. Image rotation takes a point in the image as the center of rotation and rotates it by a certain angle clockwise or counterclockwise by the rotation formula. Now suppose that the point (x_0, y_0) in the image is

the center of rotation, and the angle of rotation is $\theta(\theta \in [-\pi, \pi])$, then the coordinates of the pixel point after the rotation are (x_1, y_1). The correspondence of the coordinates before and after rotation is shown in equation (5):

$$\begin{cases} x_1 = x_0 \cos\theta - y_0 \sin\theta \\ y_1 = x_0 \sin\theta + y_0 \cos\theta \end{cases} \tag{5}$$

Image Transpose. Image transposition can be imagined as the image after the picture has been symmetrically transformed along the $y = x$ line and its horizontal and vertical coordinates have been exchanged. Now suppose that the pixel point (x_0, y_0) in the image is (x_1, y_1) after the transposition, then the image transposition is described in mathematical language as in Equation (6):

$$\begin{cases} x_1 = y_0 \\ y_1 = x_0 \end{cases} \tag{6}$$

II. A. 2) Image Grayscaling

In order to increase the processing speed of the image and without reducing the sample size, grayscaling the image is very effective method. The weighted average expression (7) for image grayscaling is given below [20]:

$$f(i, j) = 0.30R(i, j) + 0.59G(i, j) + 0.11B(i, j) \tag{7}$$

In Eq. (7), i, j denotes the coordinates of the pixel point in the image, and $R(i, j), G(i, j), B(i, j)$ denote the values of the three primary colors of RGB and are weighted with different weights to obtain (i, j), which is the pixel value of the grayed image at the point (i, j).

II. A. 3) Image enhancement

In this paper, histogram equalization will be used for image enhancement respectively [21]. Gray scale histogram represents the probabilistic statistics of various gray levels in an image and uses statistical ideas to represent the features of an image. Assuming that r is any gray level in the gray range of the image, $P(r)$ represents the probability that a pixel with a gray level of r occurs in the image. The relationship between r and $P(r)$ can be expressed by Equation (8), and the function image of expression (8) is the gray level histogram of the current image. Then:

$$P(r) = \frac{N_r}{N} \tag{8}$$

In Eq. (8), N_r denotes the number of pixels with gray level value r in the target image, and N represents the total number of pixels.

According to expression (8), the relationship between the normalized image gray level r, the probability density function $P_r(r)$ and the gray level s of the target image after equalization is shown in expression (9), and $T(r)$ in the expression is the gray level transformation function:

$$s = T(r) = \int_0^r P_r(r) dr \tag{9}$$

The steps for histogram equalization with the given target image data are as follows:

1) Statistical target image histogram is obtained:

$$P_r(r_k) = \frac{N_k}{N} \tag{10}$$

2) Histogram cumulative distribution curve:

$$s_k = T(r_k) = \sum_{i=0}^{k} P_r(r_i) = \sum_{i=0}^{k} \frac{N_k}{N} \quad k = 0, 1, 2, \ldots, L-1 \tag{11}$$

II. B.Feature extraction of children's product images

Feature extraction is needed before image recognition because image recognition is performed through the visual features of the image, which mainly include shape features, color features, spatial features, texture features and so on. Currently the main feature extraction methods are color features, texture features and shape features.

II. B. 1) Color characteristics

Color features are visual features that can be easily captured by an observer and can reveal the nature of an object's surface.

Color histogram is a method of representing the color features of an image. Now assume that an image has a distribution of pixel values $s(x, y)$, then its color histogram formula can be represented by equation (12):

$$h_c[m]=\sum_{x=0}^{W-1}\sum_{y=0}^{H-1}\begin{cases}1, & Q_c(T_c s(x,y))=m\\ 0, & \text{otherwise}\end{cases} \tag{12}$$

Where W denotes the width of the image, H denotes the height of the image, Q_c denotes the quantization method, and T_c denotes the transform form.

Non-uniform quantization is used for the RGB color space, i.e., the three components R, G and B are quantized to 8 levels, and the quantized three components are merged into one bit feature vector L=R+G+B, where each feature L is a bin in the histogram, and the quantized image can be obtained as a color histogram of 72 bins.

Equation (13) is the level formula for the three components R, G and B:

$$R=\begin{cases}0 & \text{if} \quad r\in(200,30]\\ 1 & \text{if} \quad r\in(30,50]\\ 2 & \text{if} \quad r\in(50,75]\\ 3 & \text{if} \quad r\in(75,95]\\ 4 & \text{if} \quad r\in(95,115]\\ 5 & \text{if} \quad r\in(115,135]\\ 6 & \text{if} \quad r\in(135,155]\\ 7 & \text{if} \quad r\in(155,180]\\ 8 & \text{if} \quad r\in(180,200]\end{cases} G=\begin{cases}0 & \text{if} \quad g\in(210,28]\\ 1 & \text{if} \quad g\in(28,55]\\ 2 & \text{if} \quad g\in(55,80]\\ 3 & \text{if} \quad g\in(80,95]\\ 4 & \text{if} \quad g\in(95,120]\\ 5 & \text{if} \quad g\in(120,135\\ 6 & \text{if} \quad g\in(135,155\\ 7 & \text{if} \quad g\in(155,180]\\ 8 & \text{if} \quad g\in(180,210]\end{cases} B=\begin{cases}0 & \text{if} \quad r\in(200,20]\\ 1 & \text{if} \quad r\in(20,50]\\ 2 & \text{if} \quad r\in(50,70]\\ 3 & \text{if} \quad r\in(70,95]\\ 4 & \text{if} \quad r\in(95,115]\\ 5 & \text{if} \quad r\in(115,135]\\ 6 & \text{if} \quad r\in(135,155]\\ 7 & \text{if} \quad r\in(155,180]\\ 8 & \text{if} \quad r\in(180,200]\end{cases} \tag{13}$$

Color histogram to extract color features are mainly the following:

1) Mean value: it can reflect the color tone of the image as a whole. The expression is:

$$\mu_L=\sum_{k=0}^{L}kp(k) \tag{14}$$

2) Variance: can reflect the color richness in the image. The expression is:

$$\sigma_L^2=\sum_{k=0}^{L}(k-u_L)^2p(k) \tag{15}$$

3) Energy: the second-order moments of the individual gray values with respect to the origin. The expression is:

$$L_g=\sum_k[p(k)]^2 \tag{16}$$

4) Peak: reflect the color histogram distribution distance from the mean degree of proximity. The expression is:

$$K=\frac{1}{\sigma_L^4}\sum_{k=0}^{L-1}(k-\mu)^4p(k) \tag{17}$$

5) Entropy: the amount of information contained in the image. Expressed as:

$$E=-\sum_b p(b)\log p(b) \tag{18}$$

II. B. 2) Texture Characterization

The texture features of an image can demonstrate some kind of homogenized image visual features in the image.

Gray scale covariance matrix (GLCM) is a second order statistical method used to analyze the texture features of an image [22]. Now define a grayscale image $f(x,y)$ of size M×N and the level of grayscale can be expressed as Ng, then the grayscale covariance matrix can be defined as Equation (19):

$$p(i,j)=\#\{(x_1,y_1),(x_2,y_2)\in M\times N\mid f(x_1,y_1)=i,f(x_2,y_2)=j \tag{19}$$

In the equation #{} denotes the number of elements in the set {}, and P is a matrix of gray levels Ng×Ng. Let the distance between the points (x_1,y_1) and (x_2,y_2) be d, and the angle with the horizontal axis of the coordinates be θ, then according to these variables, we can get the grayscale covariance matrices with different spacing as well as the angle $P(i,j,d,\theta)$. In general $d=1$, θ can be taken as 0°, 45°, 90°, 135°, then the definition expression of the grayscale covariance matrix is as follows:

$$
\begin{aligned}
&p(i,j,d,0°) = \#\{[(k,l),(m,n)] \in (M \times N) \times (M \times N) \mid k-m=0 \\
&\mid l-n \mid = d, f(k,l) = i, f(m,n) = j\} \\
&p(i,j,d,45°) = \#\{[(k,l),(m,n)] \in (M \times N) \times (M \times N) \mid k-m=d \\
&l-n=-d, f(k,l) = i, f(m,n) = j\} \\
&p(i,j,d,90°) = \#\{[(k,l),(m,n)] \in (M \times N) \times (M \times N) \mid\mid k-m \mid = d \\
&l-n=0, f(k,l) = i, f(m,n) = j\} \\
&p(i,j,d,135°) = \#\{[(k,l),(m,n)] \in (M \times N) \times (M \times N) \mid k-m=d \\
&l-n=d, f(k,l) = i, f(m,n) = j\} \\
&k,m = 0,1,2,\ldots,L_y; l,n = 0,1,2,\ldots,L_x; i,j = 0,1,2,\ldots,N_g - 1
\end{aligned}
\tag{20}
$$

In Eq. (20), #{} denotes the number of elements in the set, (m,n) and $(k,1)$ denote the coordinates of the pixel points, and $f(m,n)$ and $f(k,1)$ denote the gray values of the pixel points.

Definitional formulas for angular second-order moments, moments of inertia, entropy, and correlation are given below:

1) Angular second-order moments: used to measure the degree of texture coarseness and uniformity of the gray scale distribution of the image. The definition formula is as follows:

$$Asm = \sum_i \sum_j p(i,j,d,\theta)^2 \tag{21}$$

2) Moment of inertia: Used to measure the degree of image clarity and image texture depth. The definition formula is as follows:

$$Con = \sum_i \sum_j (i-j)^2 p(i,j,d,\theta) \tag{22}$$

3) Entropy: Used to measure the amount of information contained in an image. The definition formula is as follows:

$$Ent = -\sum_i \sum_j p(i,j,d,\theta) \log p(i,j,d,\theta) \tag{23}$$

4) Correlation: used to measure the gray-scale correlation between local positions in an image. The definition formula is as follows:

$$Corr = \left[\sum_i \sum_j ((ij) p(i,j)) - \mu_x \mu_y \right] / {\sigma_x}^2 {\sigma_y}^2 \tag{24}$$

II. B. 3) Shape characteristics

The external representation pattern of an object is a key feature of the object.

1) Rectangularity: indicates the ratio of the area of the pest spot image to the area of its smallest outer rectangle, which can reflect the ratio of the pest spot image to its smallest outer matrix. The defined formula is as follows:

$$R_j = A / S_{MER} \tag{25}$$

2) Narrowness: indicates the degree of narrowness of the target image. The defined formula is as follows:

$$L = \frac{M_{\theta \max}}{M_{\theta \min}} \tag{26}$$

where M_θ is given by:

$$M_\theta = \sum_{i=1}^{M} \sum_{j=1}^{N} (j \cos\theta - i \sin\theta)^2 f(i,j) = \mu_{02} \cos^2\theta + \mu_{20} \sin^2\theta - \mu_{11} \sin 2\theta \tag{27}$$

M_θ denotes the moment of inertia f of the image around $y = x \tan\theta$, and $M_{\theta \max}$ and $M_{\theta \min}$ are the maximum and minimum values of the moment of inertia, respectively.

3) Circularity: it is used to measure the similarity between the target image and a circle. The defined formula is as follows:

$$R = 4\pi A / PL \tag{28}$$

Where A denotes the area of the infested spot, P denotes the perimeter of the infested spot, and L is the narrowness.

4) Eccentricity: used to measure the compactness of the target image. The defined formula is as follows:

$$E = p / q \tag{29}$$

where p denotes the short axis length of the rotational inertia of the target region and q denotes the long axis length of the rotational inertia of the target region.

$$p=\sqrt{\frac{2}{(\mu_1+\mu_2)+\sqrt{(\mu_1-\mu_2)^2+4\mu_3^2}}} \quad q=\sqrt{\frac{2}{(\mu_1+\mu_2)-\sqrt{(\mu_1-\mu_2)^2+4\mu_3^2}}} \tag{30}$$

5) Shape complexity discrete index: describes the complexity of the shape of the target area. The defined formula is as follows:

$$e=P^2/A \tag{31}$$

6) Hu invariant moments

Hu invariant moments, also called geometric moments, can be maintained under the original image rotation, translation, scaling and other transformations, assuming that an image function is represented as $f(x,y)$, the size of M × N, then $p+q$ order moments are:

$$m_{pq}=\iint x^p y^q f(x,y)dxdy \quad p,q=0,1,2,... \tag{32}$$

The corresponding central moments are:

$$\mu_{pq}=\iint (x-x_0)^p (y-y_0)^q f(x,y)dxdy \tag{33}$$

where x_0 and y_0 denote the horizontal and vertical coordinate points of the center of gravity of the image, respectively.

The central moment of the image function $f(x,y)$ is normalized:

$$\eta_{pq}=\frac{\mu_{pq}}{\mu_{00}^r}, r=(p+q)/2+1 \tag{34}$$

Then seven invariant moments are constructed using the second- and third-order normalized central moments:

$$\begin{aligned}
\phi_1&=\eta_{20}+\eta_{02}\\
\phi_2&=(\eta_{20}+\eta_{02})^2+4\eta_{11}^2\\
\phi_3&=(\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\eta_{03})^2\\
\phi_4&=(\eta_{30}+\eta_{12})^2+(\eta_{21}+\eta_{03})^2\\
\phi_5&=(\eta_{30}-3\eta_{12})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]\\
&+(3\eta_{21}-\eta_{03})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]\\
\phi_8&=(\eta_{20}-\eta_{02})\left[(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]+4\eta_{11}(\eta_{30}+\eta_{12})(\eta_{21}+\eta_{03})\\
\phi_7&=(3\eta_{21}-\eta_{03})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]\\
&+(3\eta_{21}-\eta_{30})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]
\end{aligned} \tag{35}$$

III. VGG-based image feature recognition model for children's products

On the basis of children's product image preprocessing and feature extraction, this chapter will use the VGG network to recognize children's product image features, adding two residual modules to the basic structure of VGG to realize the assessment of children's product safety based on image feature recognition and classification [23].

III. A. VGG Convolutional Neural Networks

The VGG network consists of five convolutional layers (each using a 3×3 convolutional kernel), three fully connected layers, and a Softmax output layer. The layers in the network are divided among themselves using Max-pooling, and the activation units of all the hidden layers use the ReLU function.VGG reduces the parameters by using multiple convolutional layers with small convolutional kernels (3×3) instead of large convolutional layers with convolutional kernels, which corresponds to more nonlinear mapping and improves the fitting/expression ability of the network. Since the convolutional kernel concentrates on the expansion of the number of channels and the pooling concentrates on the reduction of the width and height, it makes the model architecture wider in terms of depth and width, and at the same time suppresses the size of the computational growth.

III. B. Multi-branch structure

The first appearance of multibranch structure was in Inception, where high performance gains were obtained, plus different sensory fields could be obtained after applying different convolution kernels to different branches. The next network, ResNet, also has a multibranch structure for its residuals. However, it is worth noting that the multi-branch structure needs to save the intermediate results, which will significantly increase the memory usage.

The more branches the network model has, the slower the processing speed of the model, and the less branches the network model has, the faster the processing speed. It can be concluded that although the appearance of multiple branches will bring the network model high performance gains, the memory occupation will increase significantly, and the inference speed will be reduced to a certain extent, which is not the current industrial scenarios desired. Especially when dealing with images of children's products, the number of product samples that need to be processed is huge, which requires fast processing to improve efficiency. It is concluded from the study that this problem can be solved by using multiplex fusion method.

III. C. High-performance components

As deep learning continues to be explored and researched, many network components with excellent performance have emerged. For example, deep separable convolution, grouped convolution, and so on. All of these components can significantly increase the performance of the network large, in terms of grouped convolution, when a network model has been grouped convolution, the more groups the network performance will be better.

Suppose the number of groups is 2, the size of the feature map is $W \times H \times C$, the size of the filter is $k \times k \times C$, the number is K. The feature map is divided into G groups, so the number of feature maps per group is equal to C/K, and the filters are grouped together and the number of each group is equal to C/K. In the process of convolution, each group of filters processes the corresponding feature map.

The number of parameters that its filters have is:

$$p = k \times k \times \left(\frac{K}{G}\right) \times \left(\frac{C}{G}\right) \times G \tag{36}$$

If no grouping is performed, the number of parameters of the filter is:

$$p_1 = k \times k \times K \times C \tag{37}$$

Calculated from Eqs. (36) and (37), the final number of parameters for the filter of grouped convolution is 1/G of the number of parameters required for normal convolution, which greatly reduces the number of parameters required.

Although deep separable convolution can significantly reduce the number of floating point operations per second (FLOPs), it has the same increase in memory access cost, which ultimately leads to a slower model.

III. D. Improved VGG network structure

This chapter proposes an improved network model based on VGG from the perspective of feature extraction relevance, inspired by ResNet, the model adds two residual modules to the basic structure of VGG, thanks to the multi-branching structure the network mitigates the gradient vanishing problem of the deep network during training and obtains a more robust feature representation. Residual module a contains only conv3×3 and conv1×1 branches; residual module b contains conv3×3, conv1×1 and Identity branches. The main ideas are as follows.

(1) The information of children's product images is relatively trivial, and it is difficult to extract the main feature information. The introduction of the residual module is equivalent to applying the features in ResNet network to VGG network, which reduces the loss rate in the process of image information extraction and alleviates the problem of gradient vanishing. In this paper, we use VGG network as a benchmark network to recognize children's product image data.

2) The introduction of 1×1 convolutional kernel, without affecting the output dimension, the use of small convolutional kernel in the same sense of the field of the case of the computational volume becomes smaller, you can introduce more nonlinear transformations, so as to make the model fit better, increase the expressive ability of the network, reduce the computational amount of the network.

3) In the training and inference phases of the network this paper made adjustments and used different network architectures. In the training phase we focus more on the accuracy of the recognition, while in the inference phase we focus more on the speed.

Compared to ResNet made different additions by adding each branch on a 3×3 conv. Divided into two Block blocks, residual module a contains only conv3×3 and conv1×1 branches; residual module b contains conv3×3, conv1×1 and Identity branches.

The computations performed within each Block are:

$$Out = F(X) + G(X) + X \tag{38}$$

In Eq. (38) F(X) denotes 3 × 3 convolution and G(X) denotes 1 × 1 convolution.

III. E. Conversion of multiplexed models to single-pass models

The following details how to equivalently convert the trained multiplexed model to a single-pass model, with a BN layer added in each branch before conversion.

The first step first fuses the 3×3 convolution and BN into a 3×3 convolution, then fuses the 1×1 convolution and BN into a 3×3 convolution, and then converts the BN into a 3×3 convolution. The second step fuses three 3×3 convolutions into one 3×3 convolution.

1) Merging of 3×3 convolutional kernel and BN layer

The network model proposed in this paper improves the inference speed after merging both the convolutional and BN layers. The convolutional layer is calculated as in equation (39):

$$Conv(x) = W(x) + b \tag{39}$$

And the computation of BN layer in inference stage involves four sets of parameters, each of which has the same number of parameters and the same dimensions of features. The BN layer computation formula is shown in equation (40):

$$BN(x) = \gamma * \frac{(x - mean)}{\sqrt{var}} + \beta \tag{40}$$

γ represents the weights, $mean$ represents the mean, var represents the variance, and β represents the bias. The result obtained from Eq. (40) is then brought into the BN layer formula to obtain Eq. (41):

$$BN(Con(x)) = \gamma * \frac{W(x) + b - mean}{\sqrt{var}} + \beta \tag{41}$$

Simplification of the resulting equation yields:

$$BN(Con(x)) = \frac{\gamma * W(x)}{\sqrt{var}} + \left(\frac{\gamma * (b - mean)}{\sqrt{var}} + \beta \right) \tag{42}$$

We can think of Eq. (43) as a convolutional layer, except that the parameters of Eq. (39) are considered. For convenience the following order:

$$W_{fused} = \frac{\gamma * W(x)}{\sqrt{var}} \tag{43}$$

$$B_{fused} = \frac{\gamma * (b - mean)}{\sqrt{var}} + \beta \tag{44}$$

The final fusion result is equation (45):

$$BN(Conv(x)) = W_{fused} + B_{fused} \tag{45}$$

2) Merging of BN layer and 1×1 convolution kernel

The overall fusion idea is to first convert Conv1×1+BN to Conv3×3+BN and then fuse it to Conv3×3. In order to convert to a 3×3 convolution operation, zeros can be complemented around the 1×1 convolution kernel to make it a 3×3 convolution kernel, and then fusion calculations can be carried out using the merging method of the 3×3 convolution kernel and the BN layer mentioned above.

3) Conversion of BN layer to 3×3 convolution kernel

The overall fusion idea is to transform the BN to Conv1×1 then convert Conv1×1+BN to Conv3×3+BN then fuse to Conv3×3.

4) Multi-branch Conv3×3 fused into one Conv3×3

The convolution operation process is shown in equation (46):

$$Output = (I \otimes K_1 + B_1) + (I \otimes K_2 + B_2) + (I \otimes K_3 + B_3) \tag{46}$$

$Output$ is the fused output, $\otimes$ denotes the convolution operation, I represents the input features, K_i represents the weights of the corresponding 3×3 convolution, and B_i denotes the corresponding bias, and then $I \otimes K_i + B_i$ denotes the corresponding convolution kernels for the computation. This is simplified to Eq. (47):

$$Output = I \otimes (K_1 + K_2 + K_3) + (B_1 + B_2 + B_3) \tag{47}$$

From equation (47) it can be seen to consider the fused output as the sum of the convolution operation on the input feature map plus all the branch weights multiplied by the corresponding branch outputs plus all the branch biases.

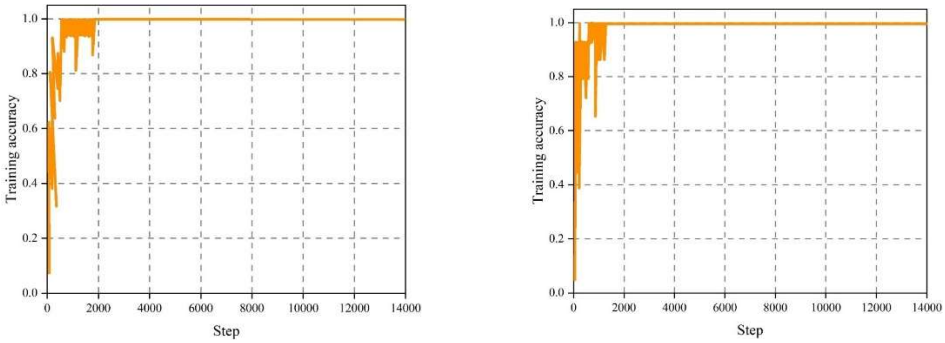
IV. Experiments on the performance of image feature recognition models for children's products

In this chapter, the performance of the proposed VGG-based image feature recognition model for children's products (hereinafter referred to as the “VGG improved model”) is tested. The experiments can be divided into three parts, respectively, to explore the prediction accuracy and loss function variation of the improved VGG model in this paper, and to compare it with other children's product image feature recognition and monthly classification techniques.

IV. A. Analysis of forecast accuracy

The main concerns in evaluating the merits of a model are the accuracy of training, the degree of decline of the loss function, the speed of convergence of the model, and the presence of oscillations.In this section, we will analyze the performance of this paper's VGG improved model in terms of the accuracy obtained from training. The accuracy reflects the ability of the network model to recognize the image features of children's products; the loss value reflects the error level of the network model in recognizing the image features of children's products; the degree of oscillation of the model reflects the stability of the model and whether there is the phenomenon of gradient explosion.

The training accuracy of the original VGG model and the training accuracy change of the VGG improved model in this paper are shown in Fig. 1, Figs. (a) and (b) correspond to the original VGG model and the VGG improved model in this paper, respectively. From the change of training accuracy shown in the figure, it can be seen that the convergence speed as well as the training accuracy of the VGG improved model of this paper are higher than that of the unimproved original VGG model during the 120 rounds of training. In terms of convergence speed, the original model starts to converge in the 15th round, while the improved model is close to convergence in the 10th round, which is significantly faster than the original model. Combining the training accuracy and convergence speed, the improved network model effectively improves the performance and can obtain a higher training accuracy in a shorter time.

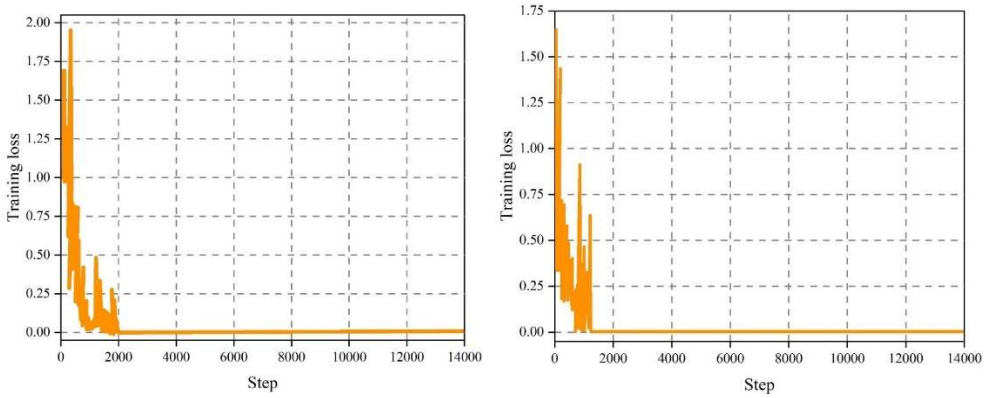


(a) The original VGG model (b) The improved VGG model in this paper

Figure 1: The change curve of model prediction accuracy

IV. B. Analysis of loss function changes

To evaluate the merits of a model algorithm, in addition to focusing on the change of its accuracy in network training, we also need to focus on the change of the loss function. The changes of loss values of the original VGG model and the VGG improved model of this paper in training are shown in Fig. 2, and Figs. (a) and (b) correspond to the original VGG model and the VGG improved model of this paper, respectively. As can be seen from the changes in the loss values of the models shown in Fig. 2, the loss function values of both models decrease rapidly in a relatively short period of time after training until they drop to a level close to 0. However, the loss function values of the improved neural network model decrease more rapidly. Although there are local oscillations in the loss value of the improved model in the pre-training period, and the stability of the model is a little bit poorer, the size of the loss value and the convergence speed of the improved network model play a positive role in the change of the training loss value.



(a) The original VGG model (b) The improved VGG model in this paper

Figure 2: The change curve of model loss value

IV. C. Comparative analysis of model performance

In order to better compare and test the performance of each network model, the more famous convolutional neural network model is used to train the data set of this paper for migration learning. The commonly used neural network models include AlexNet model, unimproved VGG-16, VGG-19 model, and GoogLeNet model. The same dataset images are input into several different network models and trained, the prediction accuracy of each model can be obtained separately, as shown in Table 1. From the table, we can see that the AlexNet model, the original VGG-16 model and the original VGG-19 model have similar accuracies, which are 91.35%, 91.12% and 91.04%, respectively, while the prediction accuracy of the GoogLeNet model is slightly higher than that of them, reaching 93.82%. The model in this paper, on the other hand, has the highest prediction accuracy among all models, up to 98.47%. In addition, the loss value of this paper's VGG-improved model is also the lowest among all models, with a loss value of only 0.0824. From the results, this VGG-improved model has more excellent performance in recognizing the features of children's product images.

Table 1: Performance comparison of models

Model	Accuracy(%)	loss
AlexNet	91.35%	0.2368
GoogLeNet	93.82%	0.1865
VGG16 (SGD)	91.12%	0.2689
VGG19	91.04%	0.2716
The model in this paper	98.47%	0.0824

V. Experiments to assess the safety of children's products

In this chapter, experiments on children's product safety assessment will be carried out to verify the effectiveness of the VGG-based image feature recognition model for children's products proposed in this paper in the work of children's product safety identification and classification.

V. A. Experimental environment and parameter settings

1) Experimental environment

The operating system used for the experiments in this chapter is Windows 10, the programming language is Python, and the PyTorch framework is used to build the VGG convolutional neural network model, and the detailed configuration is shown in Table 2.

2) Parameter settings

Through several adjustments, combined with the experimental environment and experimental subjects in this chapter, the various training parameter settings used in the training of different network models are shown in Table 3.

Table 2: Experimental environment configuration

Name	Parameters / Version Number
CPU model	AMD Ryzen 7 4800H with Radeon Graphics 2.90 GHz
GPU model	NVIDIA GeForce RTX 2060
Memory	16GB
Python	3.8.5
PyTorch	1.8.1

Table 3: Experimental super parameter setting

Hyperparameters	Numerical value
Initial learning rate	0.001
Batch Size	16
Optimizer	Adam
The number of training rounds (epoch)	200

V. B. Data set selection

Google's public dataset Open Image Dataset partially contains children's product images, and the children's product images in its dataset are integrated to form children's product image datasetchildren's products Dataset, which covers 1,588 children's product images.

V. C. Analysis of experimental results

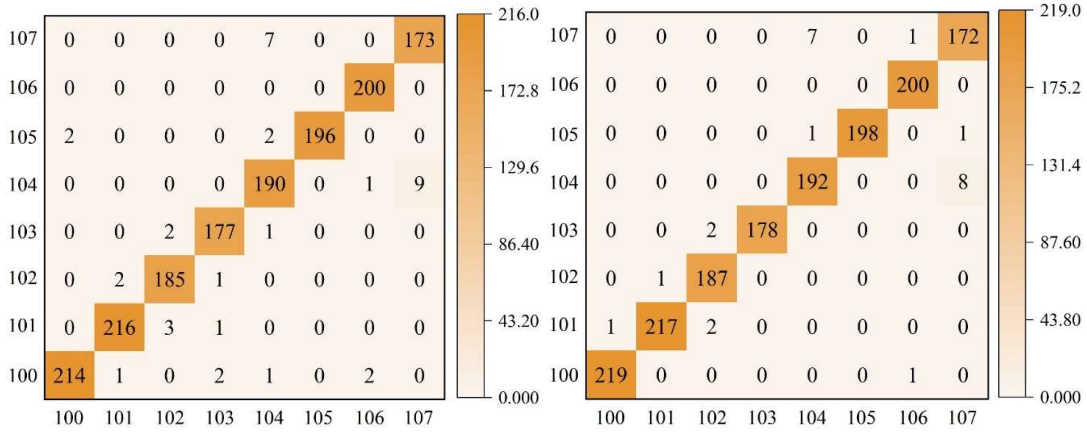
In order to prove the effectiveness of the proposed VGG-based image feature recognition model of children's products in the safety assessment of children's products, various network models are trained in the dataset, and the experimental results of the VGG improved network model in this paper are compared with those of several other classical convolutional neural network models (AlexNet, VGG16, GoogLeNet, DenseNet). Comparison is made to analyze the recognition performance of various network models from four aspects: accuracy, precision, recall and F1 value, and the experimental results are specifically shown in Table 4.

As can be seen from the table, for the recognition of children's product images, all six convolutional neural network models achieve better recognition results, but there are some differences in the recognition performance between the models. Among them, the recognition accuracy of AlexNet model reaches 94.6%, which has certain recognition ability for the safety of children's product images; GoogLeNet model achieves 95.45% recognition accuracy by virtue of its multi-scale feature extraction ability; VGG16 model simplifies the scale of the network model by stacking multiple 3×3 convolutional kernels, and the recognition accuracy reaches 96.68%; DenseNet model achieves 96.68% recognition accuracy by feature reweighting; and DenseNet model reaches 96.68% by feature reweighting; and DenseNet model reaches 96.68% by feature reweighting. ; the DenseNet model enhances the feature learning ability of the network model through feature reuse, and achieves a recognition accuracy of 97.45%; the ResNet18 model utilizes residual blocks to improve the recognition performance of the network model, and achieves a recognition accuracy of 97.72%; the improved model of VGG proposed in this paper adds two residual modules to the network structure of VGG, and achieves the highest recognition accuracy in children's products' safety recognition achieved the highest recognition accuracy of 98.34%. Compared with several other classical convolutional neural network models, the VGG improvement model in this paper achieves optimal results in the four aspects of accuracy, precision, recall and F1 value, which fully proves the effectiveness of the VGG improvement model in this paper for the assessment of the safety of children's products.

Table 4: Comparison of Experimental Results of Different Models

Model	Accuracy	Precision	Recall	F1-score
AlexNet	94.60%	94.64%	94.53%	94.49%
GoogLeNet	95.45%	95.39%	95.36%	95.42%
VGG16	96.68%	96.54%	96.58%	96.45%
DenseNet	97.45%	97.35%	97.48%	97.38%
ResNet18	97.72%	97.65%	97.60%	97.61%
Model of this article	98.34%	98.34%	98.33%	98.34%

Through the analysis, it can be seen that among all the compared network models, the performance of ResNet18 model is only second to the VGG improved model of this paper. In order to reflect more intuitively from the data that the recognition effect of this paper's VGG improved model on children's product images is better than the original ResNet18 model, the confusion matrix and experimental results of these two network models for recognizing different kinds of children's product images are shown separately to analyze the recognition results of the models and the specific misclassification in detail. The confusion matrix of the original ResNet18 model and the improved VGG model in this paper is shown in Figure 3. In the figure, 100~107 represent toys, baby products, children's clothing and apparel, learning and education, children's furniture, children's food, health and care, and sports and outdoor products in children's products, respectively. Observation of the confusion matrix shows that the proposed ECAResNet model correctly recognizes more images of children's products than the original ResNet18 model, which reduces the number of incorrect predictions. Taking the images of children's furniture products with category 105 as an example, the original ResNet18 model incorrectly misclassified 2 images to toy products with category 100 and 2 images to learning and education products with category 104, compared to the misclassification of the VGG improved model in this paper, which occurs less often.



(a) ResNet18 model (b) Model of this article

Figure 3: Confusion matrix

The experimental results of the original ResNet18 model and the VGG improved model of this paper for different categories of children's products recognition are shown in Table 5. Observing the table, it can be seen that the VGG improved model of this paper performs better and more reliably in recognizing different categories of children's products, and for children's product categories with more obvious and special product features, such as toys (100), baby products (101), children's clothing and apparel (102), children's food products (105), and health and care products (106), the model can achieve a recognition accuracy of 99%~100% recognition accuracy, while for children's products with similar product features and easy to be confused, such as learning and education products (103), children's furniture products (104), and sports and outdoor products (107), the model's recognition accuracy will be reduced, but all of them perform better than the original ResNet18 model.

Table 5: Comparison of recognition results of different types of children 's products

Categories of children 's products	ResNet18 model			Model of this article		
	Precision	Recall	F1-score	Precision	Recall	F1-score
100	99.08%	97.38%	98.28%	100.00%	98.78%	99.55%
101	98.52%	98.27%	98.52%	99.65%	98.75%	99.18%
102	97.26%	98.45%	97.77%	99.66%	99.44%	99.56%
103	97.68%	98.22%	98.15%	98.01%	99.58%	98.78%
104	94.64%	95.12%	94.86%	96.02%	96.42%	96.12%
105	100.00%	98.02%	98.88%	100.00%	99.20%	99.62%
106	98.63%	100.00%	99.15%	99.30%	100.00%	99.86%
107	95.14%	96.22%	96.58%	95.14%	95.69%	95.38%

VI. Strategies for improving the safety design of children's products

In the above paper, this paper constructs the image feature recognition model of children's products based on VGG, and realizes the evaluation of the safety of children's products by classifying and recognizing the image features of children's products. In this chapter, the safety of children's products will be considered and the corresponding design improvement direction will be proposed.

1) Focus on the safety of chemical composition of children's products

Children's products in the process may use a variety of dyes, additives, these chemicals may contain substances harmful to human health, under certain conditions may pose a risk to children's health. Therefore, in the choice of leather fabrics, designers should pay special attention to children's products of chemical composition safety standards.

2) Pay attention to the physical safety of children's products

Children's products in the use of physical safety is also very important. This includes whether the edges of the product are smooth to prevent scratches; whether there are sharp parts in the product, such as zippers and buttons.

3) Concern about the biological safety of children's products

It is very important to ensure that the biosafety of children's products is properly handled during the production and use of the products. Manufacturers need to comply with relevant safety standards and regulations, and use environmentally friendly and safe raw materials and processes to protect consumer health and reduce the impact on the environment. In addition, designers should consider the colors and textures of the products and avoid using overly bright colors, which may contain high levels of chemical additives. Natural colors or neutral tones are usually safer choices.

4) Design considerations for children's product sizes

Children's products should be sized according to the age and size of the child. Designers need to ensure that seams and edges are properly finished to accommodate smaller home product components, while avoiding sharp edges or corners to minimize the risk of injury to children.

5) Design considerations for children's product shapes

The shapes of children's products should be simple and rounded, avoiding sharp edges and corners, which helps to minimize injuries caused by collisions when children use the products.

6) Design consideration for the functionality of children's products

Children's products should have a variety of functions to accommodate the changing needs of children. Consider using durable and easy-to-clean materials to ease the wear and tear associated with frequent use of children's products.

7) Consider the friendliness of children's products.

When designing children's products, designers need to consider the physical and psychological development of children. When designing children's products, consider using materials with warm and soft characteristics to provide a comfortable environment that encourages children to explore, learn and grow, while ensuring their safety.

VII. Conclusion

In this paper, we use VGG neural network to recognize and classify the image features of children's products based on the deep learning method, and construct a VGG-based children's product image feature recognition model to realize the assessment of children's product safety. The model has a faster convergence speed on the training accuracy curve than the original VGG model, and the performance of the size of the loss value and the convergence speed is better. Compared with AlexNet model, unimproved VGG-16, VGG-19 model, and GoogLeNet model, the model in this paper has the highest prediction accuracy of 98.47% among all the models, and the lowest loss value of 0.0824. Obviously, the model in this paper performs well in the performance test, and the model performance is outstanding. Carrying out experiments on children's product safety assessment to verify the effectiveness of this model in children's product safety identification and classification work, it achieved optimal results in four aspects: accuracy, precision, recall and F1 value, which reached 98.34%, 98.34%, 98.33% and 98.34%, respectively. Comparing with the original ResNet18 model, the number of correctly recognized images of children's products in this paper's model is higher than that of the original ResNet18 model, the number of incorrect predictions is less, and misclassification occurs less often. For children's product categories with obvious or special product features, the recognition accuracy rate of this method can be as high as 99%~100%, while in the face of children and products with similar product features, the recognition accuracy rate will be reduced, but its performance is still better than the original ResNet18 model. After verifying the utility of the VGG-based image feature recognition model of children's products proposed in this paper in the assessment of children's product safety, based on the consideration of children's product safety, we propose the improvement strategy of children's product safety design from the aspects of chemical composition safety, physical safety, and biological safety.

Funding

This work was supported by Open Fund Project of Hubei Cultural and Creative Industrialization Design and Research Center (HBCY1906).

References

[1] Page, T., & Thorsteinsson, G. (2017). Designing Toys to Support Children's Development. Journal on Educational Psychology, 11(2), 1-10.

[2] Todd, B. K., Fischer, R. A., Di Costa, S., Roestorf, A., Harbour, K., Hardiman, P., & Barry, J. A. (2018). Sex differences in children's toy preferences: A systematic review, meta-regression, and meta-analysis. Infant and Child Development, 27(2), e2064.

[3] John, D. R., & Chaplin, L. N. (2019). Children's understanding of the instrumental value of products and brands. Journal of consumer psychology, 29(2), 328-335.

[4] Schwebel, D. C., Wells, H., & Johnston, A. (2015). Children's recognition of dangerous household products: child development and poisoning risk. Journal of pediatric psychology, 40(2), 238-250.

[5] Lazzarini, R., Hafner, M. D. F. S., & Rangel, M. G. (2018). Evaluation of the presence of allergens in children's products available for sale in a big city. Anais brasileiros de dermatologia, 93(3), 457-459.

[6] Negev, M., Berman, T., Reicher, S., Balan, S., Soehl, A., Goulden, S., ... & Diamond, M. L. (2018). Regulation of chemicals in children's products: How US and EU regulation impacts small markets. Science of the Total Environment, 616, 462-471.

[7] Ismail, R., Haniff, W. A. A. W., Isa, S. M., Fadzil, R. M., AlSagoff, S. S., & Khalid, T. K. A. (2020). The approach to safety of children's toys in United States and European Union: A comparative study. Academic Journal of Interdisciplinary Studies, 9(1), 126.

[8] Niven, C., Mathews, B., & Vallmuur, K. (2022). Applying a public health approach to identify priorities for regulating child product safety. Australian and New Zealand journal of public health, 46(2), 142-148.

[9] Reynolds, K. M., Burnham, R. I., Delva-Clark, H., Green, J. L., & Dart, R. C. (2021). Impact of product safety changes on accidental exposures to liquid laundry packets in children. Clinical Toxicology, 59(5), 392-399.

[10] Constable, A., Mahadevan, B., Pressman, P., Garthoff, J. A., Meunier, L., Schrenk, D., ... & Hayes, A. W. (2017). An integrated approach to the safety assessment of food additives in early life. Toxicology Research and Application, 1, 2397847317707370.

[11] EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), Turck, D., Bresson, J. L., Burlingame, B., Dean, T., Fairweather-Tait, S., ... & van Loveren, H. (2017). Statement on the safety of synthetic l-ergothioneine as a novel food–supplementary dietary exposure and safety assessment for infants and young children, pregnant and breastfeeding women. EFSA Journal, 15(11), e05060.

[12] Alhusban, A. A., Ata, S. A., & Shraim, S. A. (2019). The safety assessment of toxic metals in commonly used pharmaceutical herbal products and traditional herbs for infants in Jordanian market. Biological trace element research, 187, 307-315.

[13] Vial, A., Assink, M., Stams, G. J. J., & van der Put, C. (2019). Safety and risk assessment in child welfare: A reliability study using multiple measures. Journal of Child and Family Studies, 28, 3533-3544.

[14] Sanjeevana, D., & Thenmozhi, R. (2022). Study on Product Safety of Children's Apparel. In Sustainable Approaches in Textiles and Fashion: Manufacturing Processes and Chemicals (pp. 157-169). Singapore: Springer Singapore.

[15] Dey, S., Purdon, M., Kirsch, T., Helbich, H., Kerr, K., Li, L., & Zhou, S. (2016). Exposure Factor considerations for safety evaluation of modern disposable diapers. Regulatory toxicology and pharmacology, 81, 183-193.

[16] Buckley, L. A., Salunke, S., Thompson, K., Baer, G., Fegley, D., & Turner, M. A. (2018). Challenges and strategies to facilitate formulation development of pediatric drug products: Safety qualification of excipients. International journal of pharmaceutics, 536(2), 563-569.

[17] Liu, X., Duan, S., Pei, F., Chen, Q., Liu, B., & Qiao, F. (2020, October). Safety risk assessment for children's products based on reinforcement learning. In Proceedings of the 4th International Conference on Advances in Artificial Intelligence (pp. 35-40).

[18] Żmudzińska, A., Puścion-Jakubik, A., Bielecka, J., Grabia, M., Soroczyńska, J., Mielcarek, K., & Socha, K. (2022). Health safety assessment of ready-to-eat products consumed by children aged 0.5–3 years on the Polish market. Nutrients, 14(11), 2325.

[19] Mak, S. L., & Lau, H. K. (2014, May). An implementation of toy safety assessment model. In 2014 IEEE Symposium on Product Compliance Engineering (ISPCE) (pp. 12-16). IEEE.

[20] Ivana Žeger,Ivan Šetka,Domagoj Marić & Sonja Grgic. (2024). Exploring Image Decolorization: Methods, Implementations, and Performance Assessment. Applied Sciences,14(23),11401-11401.

[21] D. Vijayalakshmi,Poonguzhali Elangovan,T. Sandhya Kumari & Malaya Kumar Nath. (2025). Optimized multi-scale framework for image enhancement using spatial information-based histogram equalization. The Imaging Science Journal,73(2),176-203.

[22] Kyuseok Kim & Youngjin Lee. (2025). Gray-Level Co-Occurrence Matrix Uniformity Correction Algorithm in Positron Emission Tomographic Image: A Phantom Study. Photonics,12(1),33-33.

[23] Xiaofang WANG,JiaLing WU,Xin CHEN,JunNiang HOU & Peichun CHEN. (2025). Optimized VGG Network with Dilated Residual Convolution and Path Enhancement for Crack Image Segmentation. INTERNATIONAL JOURNAL OF COMPUTERSCOMMUNICATIONS & CONTROL,20(2).