



# Convolutional neural network-based automatic classification for incomplete antibody reaction intensity in solid phase anti-human globulin test image

KeQing Wu<sup>1,2</sup> · ShengBao Duan<sup>3</sup> · YuJue Wang<sup>3</sup> · HongMei Wang<sup>3</sup> · Xin Gao<sup>2,4</sup> 

Received: 26 October 2021 / Accepted: 30 January 2022 / Published online: 7 March 2022  
© International Federation for Medical and Biological Engineering 2022

## Abstract

The precise classification of incomplete antibody reaction intensity (IARI) in hydrogel chromatography medium high density medium solid-phase Coombs test is essential for haemolytic disease screening. However, an automatic and contactless method is required for accurate classification of IARI. Here, we present a deep ensemble learning model that integrates five different convolutional neural networks into a single model for IARI classification. A dataset, including 1628 IARI images and corresponding labels of IARI categories ((-), (1+), (2+), (3+), and (4+)), was used. We trained our model using 1302 IARIs and validated its performance using 326 IARIs. The proposed model achieved 100%, 99.4%, 99.4%, 100%, and 100% accuracies in the (-), (1+), (2+), (3+), and (4+) categories, respectively. The results were compared with those of manual classification by immunologists (average accuracy: 99.8% vs. 88.3%,  $p < 0.01$ ). Following model assistance, all three immunologists achieved increased accuracy (average accuracy: +6.1%), with the average accuracy of junior immunologists maximum increasing by 11.3%. The time required for model classification was 0.094 s·image<sup>-1</sup>, whereas that required manually was 5.528 s·image<sup>-1</sup>. The proposed model can thus substantially improve the accuracy and efficiency of IARI classification and facilitate the automation of haemolytic disease screening equipment.

**Keywords** Antibody · Coombs test · Neural network · Blood transfusion

## 1 Introduction

Acute blood loss, anaemia, and coagulopathy are treated using blood transfusion [1, 2]. Improper blood transfusion or incompatible transfusion increases the risk of haemolytic diseases (e.g., haemolytic disease in new-borns, autoimmune

haemolytic disease, drug immune haemolytic diseases), renal failure, and even death [3–5]. The incomplete antibody reaction intensity (IARI) is the main factor responsible for incompatible transfusion. IARI is divided into five categories ((-), (1+), (2+), (3+), (4+)), with the higher intensity category causing more serious incompatible transfusion [6]. Therefore, IARI multi-classification tests are essential before blood transfusion [7–9].

Currently, the haemolytic IARI multi-classification test mainly uses the micro-column gel immune-assay Coombs test (MGIA-Coombs test) as it has high sensitivity and strong interpretability [10]. However, fibrin in the plasma can erroneously trap red blood cells (RBCs) at the top of the gel column, leading to high false-positives with the MGIA-Coombs test [11–14]. The solid-phase red cell adherence Coombs test (SPRCA Coombs test) was then proposed, involving pre-coating of anti-human globulin (AHG) on U-bottom microwells to prevent RBCs from binding to fibrin, which reduces false-positive results [15]. However, the SPRCA Coombs test requires a tedious and error-prone washing process for RBC suspensions to separate sensitised

✉ HongMei Wang  
wanghm@sibet.ac.cn

✉ Xin Gao  
xingaosam@163.com

<sup>1</sup> School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China

<sup>2</sup> Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou 215163, China

<sup>3</sup> CAS Key Lab of Bio-Medical Diagnostics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou 215163, China

<sup>4</sup> Jinan Guoke Medical Engineering and Technology Development Co., Ltd., Pharmaceutical Valley New Drug Creation Platform, Jinan 250109, Shandong, China

RBCs from free fibrin, in turn causing false-negative results [16–20]. Recently, the hydrogel chromatography medium high density medium solid-phase Coombs test (HCM-HDMS Coombs test) was proposed. It involves hydrogel chromatography medium (HCM) as the separation solution in the reaction-and-separation chamber for separating sensitised RBCs from free fibrin, thus eliminating the washing process and effectively reducing the false-negative results [21]. However, this chamber obstructs the view during the process of observing incompatible IARI in the HCM-HDMS Coombs test, thus affecting accurate visual classification, whereas moving the chamber away results in reagent contamination and leakage. Further, the observation process by immunologists is subjective and diverse [22]. Therefore, contactless, automatic, and intelligent multi-classification methods are needed to enhance the practical value of the HCM-HDMS Coombs test.

Deep learning has led to the achievement of remarkable success in medical image classification [23, 24]. In particular, with the advent of convolutional neural networks (CNNs), high-level semantic features of images can be automatically and effectively extracted to reduce the necessity of handcrafted feature processes. Recently, Liu, et al. applied a CNN model to the multi-classification of COVID-19 pneumonia, other common pneumonia, and normal controls using CT images and achieved an accuracy of 92.49% [25]. In another study, CNN models were utilized in the automated multi-classification of cells in the epithelial tissue of oral squamous cell carcinoma, with an accuracy of 97.5% [26]. Tessema et al. demonstrated the potential of integrating the deep learning-based automatic model into the quantitative multi-classification of blood cells with an average accuracy of 80.6% [27]. Thus, we hypothesized that CNN methods have the potential to achieve automatic and intelligent IARI multi-classification, and this will assist immunologists and clinicians to improve clinical efficiency and accuracy.

However, among the five IARI categories, the number of poor positive samples in (1 +), (2 +), and (3 +) is much smaller than that in (-) and (4 +), demonstrating the sample imbalance distribution of IARI categories, which will result in a shift in the decision boundary of the CNN networks in the training process. Further, the distinguishing characteristics between adjacent categories such as (-) and (1 +); (1 +), (2 +), and (3 +); and (3 +) and (4 +) are not particularly obvious, thus seriously affecting the ability of CNN networks to automatically learn and identify. Moreover, there are several bubbles, particulates, and other artefacts in reaction mixture of HCM-HDMS Coombs test that hinder the classification task. Therefore, the above existing CNN models cannot be directly used for effectively solving the IARI multi-classification.

In this study, we aimed at establishing a novel deep learning model for the automatic classification of IARI, which is unaffected by the influence of the sample imbalance distribution of IARI categories and the interference of artefacts in the HCM-HDMS Coombs test. An ensemble learning framework is used to reduce the influence of sample imbalance distribution and obtain accurate classification results. A convolutional block attention module (CBAM) is used to avoid the interference of artefacts by combining pixel-level channel interaction relationships and spatial location information.

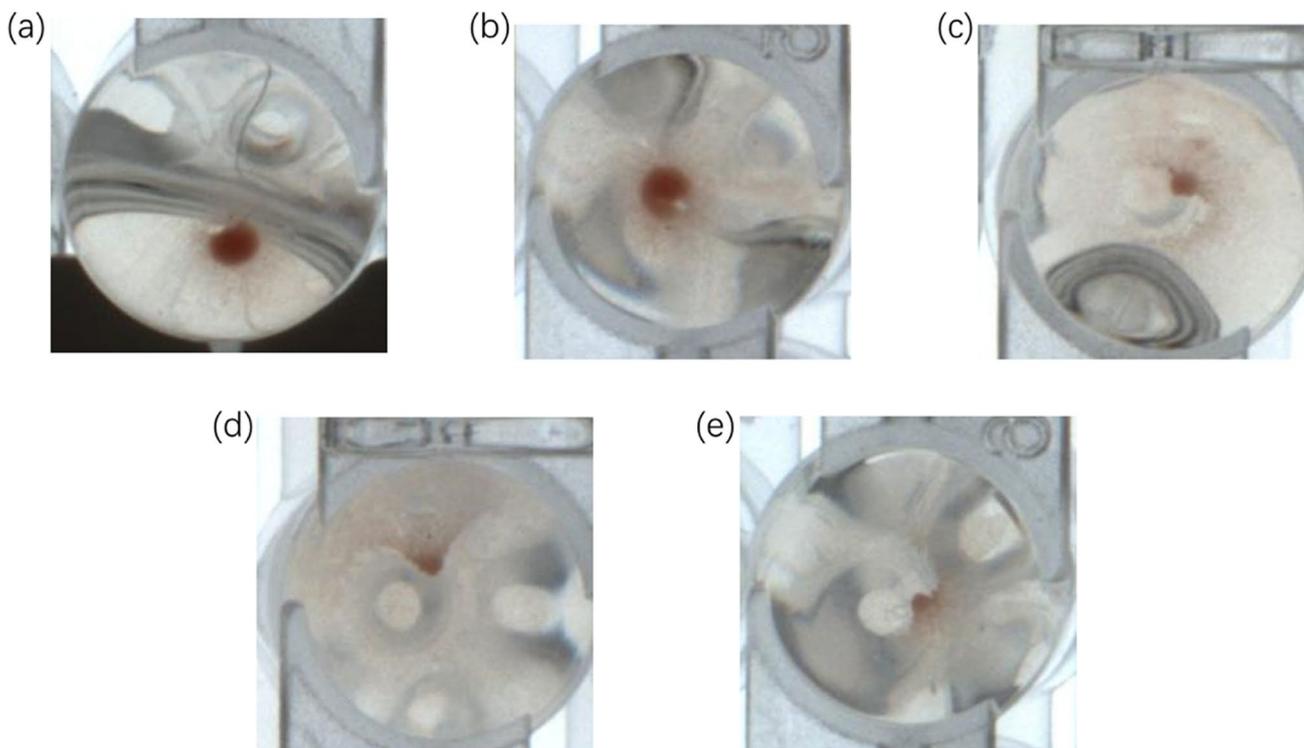
## 2 Methods

### 2.1 Dataset

In total, 1725 blood samples were collected from the Suzhou Blood Centre and the First Affiliated Hospital of Soochow University, China; these were kept at 4 °C and used within 1 week. The corresponding IARI of the blood samples were obtained using the HCM-HDMS Coombs test, and 97 samples (5.62%) whose IARI category could not be obtained accurately were excluded. A total of 1628 IARI samples (94.38%) were selected, and the number of each IARI category was as follows: 650 (-), 230 (1 +), 68 (2 +), 130 (3 +), and 550 (4 +). IARI images were captured from U-microplate bottoms in a closed image acquisition space with a stable light field, using a digital camera. Each image had a size of 229 × 230 pixels. The images for the five IARI category samples are shown in Fig. 1.

To ensure that all IARI samples were correctly classified, the labels were determined by three professional immunologists. If there was a difference among the labels, the suspected samples were re-tested using the MGIA-Coombs test to obtain the correct category. In total, 1302 IARIs (80%) were used as a training dataset to develop the deep learning model, and another 326 IARIs (20%) were used as a testing dataset for model evaluation [28]. Table 1 lists the number of labelled samples for each category in the datasets.

As we used a deep neural network-based model for classification, the training dataset was not sufficient to achieve invariances and robustness for the network model. Considering that data augmentation is a common procedure for generating sufficient training data for CNN-based models, we utilised the data augmentation package from Torchvision-Transform (<https://pytorch-cn.readthedocs.io/zh/latest/torchvision/torchvision-transform/>) and augmented the training dataset by adopting image cropping, flipping horizontally and vertically, rotating at four fixed angles of 0°, 90°, 180°, and 270° and zooming.



**Fig. 1** Samples of the five IARI intensity categories (a) Morphology of the (-) category; (b) Morphology of the (1+) category; (c) Morphology of the (2+) category; (d) Morphology of the (3+) category; (e) Morphology of the (4+) category

**Table 1** Details of the IARI image dataset used for the experiments

Category name	(-)	(1+)	(2+)	(3+)	(4+)
Dataset	650 (40%)	230 (14%)	68 (4%)	130 (8%)	550 (34%)
Training set (80%)	520	184	54	104	440
Testing set (20%)	130	46	14	26	110

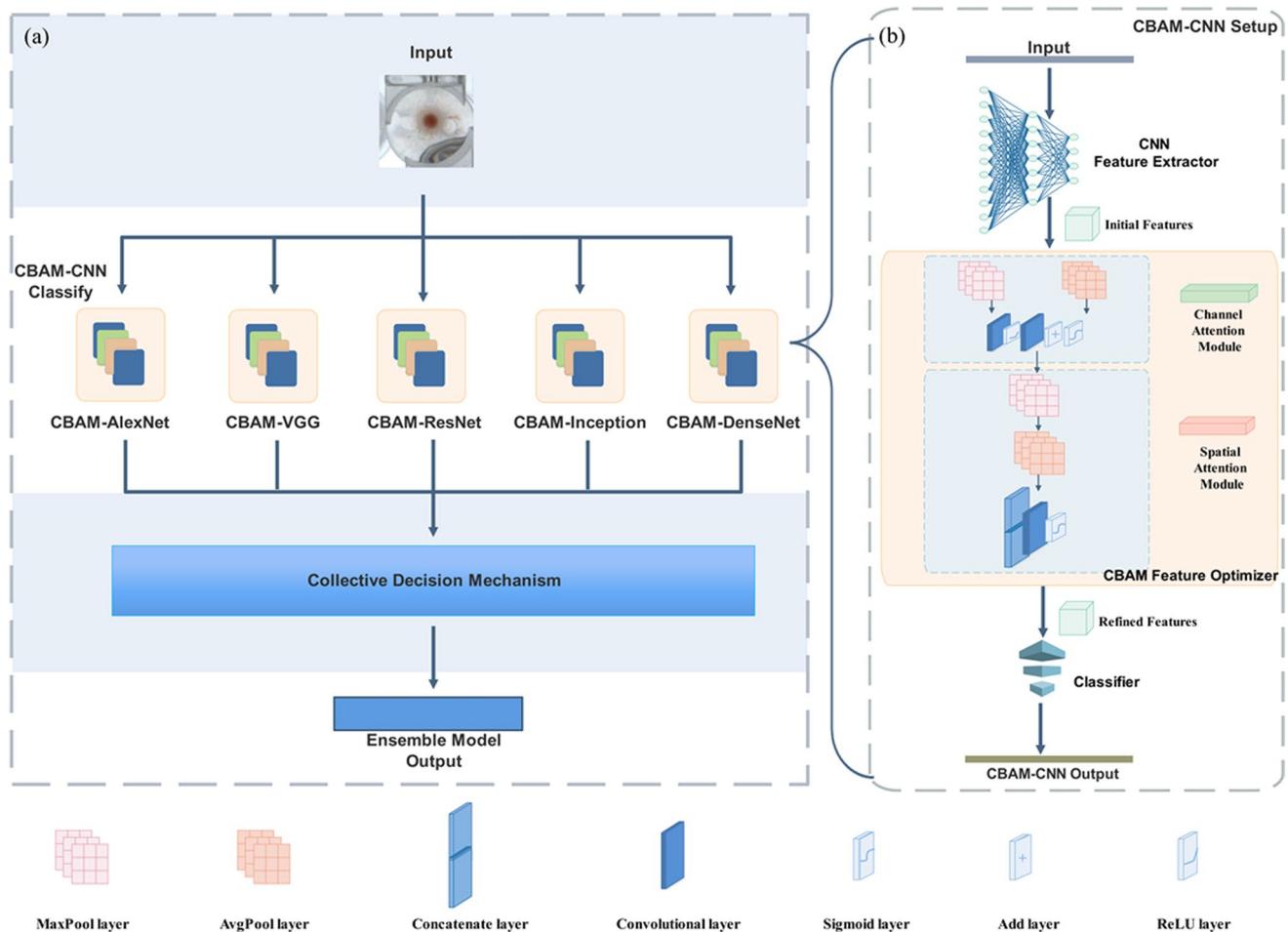
### 2.2 Classification model

In this study, we proposed an end-to-end deep learning model based on CNNs to classify the IARI, as shown in Fig. 2a, via two main stages. First, five sub-models were built by associating the IARI status. Second, the sub-models were combined into an ensemble model in parallel to obtain the final category using a collective decision mechanism. The details are as presented below.

In the first stage, to address the problem of a single model not being able to fully capture the detailed features distinguishing between adjacent IARI categories, five different CNN-based frameworks, including Alex Deep Convolutional Neural Network (AlexNet), Visual Geometry Group (VGG) Network, Residual Network (ResNet), Inception Network, and Dense Convolutional Network (DenseNet),

were adapted to classify the IARIs with improved classification performance. AlexNet transforms the linear mapping between features into a nonlinear relationship to simulate any polynomial [29]. The VGG Network reduces the computation of each convolution layer and captures more abundant features using the stacked convolution core [30]. ResNet adds residual blocks and eliminates overfitting [31]. The Inception Network balances the network depth and width and reasonably reduces the dimensions [32]. DenseNet utilises feature information more efficiently through dense connections and reduces gradient vanishing [33].

Further, a large number of bubbles, particulates, or other artefacts in IARIs contribute to the useless features extracted and hinder classification. For increased focus on effective areas and to suppress useless features, a CBAM was added to each CNN framework, as a hybrid attention mechanism capable of combining channel dimensions and spatial dimensions [34]. In the channel dimension, average pooling was used to aggregate channel interaction information, and maximum pooling was used to infer the finer channel information to further improve the representation power of the network. In the spatial dimension, average pooling and maximum pooling were concatenated to generate an efficient feature descriptor for extracting valid feature location information. As shown in Fig. 2b, compared with the original CNN, CBAM was only inserted between the feature



**Fig. 2** The proposed deep learning model design (a) A pipeline of the model; (b) The CBAM-CNN setup: CBAM is a module used before the classifier as part of improved CNN frameworks (AlexNet, VGG, ResNet, Inception, and DenseNet)

extractor and classifier of each CNN, instead of in the feature extraction process, which emphasises the crucial feature information and ensures effective feature extraction of five CBAM-CNNs (CBAM-AlexNet, CBAM-VGG, CBAM-ResNet, CBAM-Inception, and CBAM-DenseNet).

In the second stage, five CBAM-CNNs were used to form an ensemble model with a parallel combination [35]. A collective decision mechanism, referred to as relative plurality voting (RPV) module, was also constructed and added to the ensemble model. Based on the RPV module, the intensity category with the most votes among all sub-models was identified as the final classified category, providing a more reasonable decision boundary for the model. The ensemble model with the RPV module showed the advantage of making full use of multiple networks to offset the limitations of a single network and reduce the overall classification error rate [36].

For IARI images, a corresponding ensemble classification model was developed. To achieve quick convergence of the proposed model, a training dataset comprising five

IARI categories was used to train the CBAM-CNN model. The loss function Loss, for training was cross-entropy, which can be represented as follows:

$$\text{Loss} = \frac{1}{C} \sum_{j=1}^C l_c \quad (1)$$

$$l_c = -\frac{1}{N} \sum_{i=1}^N [y_i \log p(\phi(x_i); \omega) + (1 - y_i) \log (1 - p(\phi(x_i); \omega))] \quad (2)$$

where  $C$  is the number of CBAM-CNN models,  $l_c$  is the loss of one CBAM-CNN model,  $N$  is the number of IARI images,  $x_i$  is one IARI image,  $\phi(x_i)$  is the nonlinear transformation of  $x_i$ ,  $y_i$  is the IARI category corresponding to  $x_i$ ,  $\omega$  is the parameter set of the model, and  $p$  is the probability output from the model.

Relying on PyTorch open-source libraries as a back end, the ensemble model was implemented on an Ubuntu 16.04

computer with one Intel Xeon CPU, using an NVIDIA RTX 2080 Ti GPU, with 32 GB available RAM.

### 2.3 Performance metrics

Herein, four metrics, *Accuracy*, *Precision*, *Recall*, and *F-score*, were used to quantitatively evaluate the performance of the model for each IARI category classification, and these are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

$$F - \text{score} = \frac{(1 + \beta^2)\text{Precision} \times \text{Recall}}{\beta^2\text{Precision} + \text{Recall}} \tag{6}$$

where true positive (*TP*) represents the number of positives correctly predicted by the classification discriminant model, and true negative (*TN*) represents the number of negatives predicted correctly; false positives (*FP*) and false negatives (*FN*) denote the number of positive and negative misjudgements by the classification model, respectively.  $\beta$  is the weight in the *F-score* calculation to balance the proportion of *Precision* and *Recall*, and is assigned as equal to 1.

In the imbalanced datasets, macro-averaged metrics were computed the average overall categories and gave equal weights to each category, which were represented fairly for each category and regardless of its frequency [37, 38]. The macroperformance of the overall ensemble model for all categories ( $n=5$ ) was evaluated using macro-average  $X_{avg}$  [39–41], as follows:

$$X_{avg} = \frac{\sum_{i=1}^n X_i}{n}, X \in R \tag{7}$$

where  $R = \{Accuracy, Precision, Recall, F1-score\}$ .

To quantify the comparison of classification performance between the model and the immunologists, the *Kappa* coefficient was used to measure the consistency between the predictive values and true values as follows [42]:

$$\text{Kappa} = \frac{p_0 - p_e}{1 - p_e} \tag{8}$$

$$p_e = \frac{\sum_{i=1}^n (TP_i + FP_i) \times (TP_i + FN_i)}{(TP + TN + FP + FN)^2} \tag{9}$$

where  $p_0$  represents *Accuracy*.

### 2.4 Statistical analysis

Statistical analysis was conducted using R software (version 3.5.1, <https://www.r-project.org/>). *Accuracy*, *Precision*, *Recall*, *F-score*, and *Kappa* coefficient were used to evaluate the performance of CNN models and immunologists. *Kappa* coefficient utilised the `cohen.kappa()` from the `concord` package in the consistency analysis. Pearson’s chi-square test was applied to assess the differences in performance between the manual classification and the proposed ensemble learning model utilised the `chisq.test`. Statistical significance was set at  $p < 0.01$ .

## 3 Results

The implementation details of the CBAM-CNN model training were as follows: batch size, 32; epoch number, 50; and adaptive moment estimation (Adam) optimiser [43] was used to tune the parameter set  $\omega$ ; the initial learning rate was set to  $5e - 4$ .

### 3.1 Classification performance of the ensemble model

To demonstrate the effectiveness of our proposed model, six metrics (*Accuracy<sub>avg</sub>*, *Precision<sub>avg</sub>*, *Recall<sub>avg</sub>*, *F1-score<sub>avg</sub>*, *Kappa*, and *Time*) were used to evaluate the classification performance of five independent CBAM-CNN models and the ensemble model. The corresponding performance of all the models is listed in Table 2 [44]. The classification performance of independent original models was treated as the baseline for comparison with the CBAM-CNN models. The details in each category are described in the Supplementary Material Table S1. In the independent model, CBAM-CNN achieved a better performance than the original CNN, and the highest accuracy was achieved by the CBAM-Inception model (*Accuracy<sub>avg</sub>*=94.6%; *F1-score<sub>avg</sub>*=0.951; *Kappa*=0.925). The ensemble model with CBAM as the ideal optimisation method yielded the highest *Accuracy<sub>avg</sub>* of 99.8% and *F1-score<sub>avg</sub>* of 0.983 and *Kappa* of 0.991, compared to all the models.

### 3.2 Clinical utility of ensemble model classification assistance

To verify the clinical utility of our proposed model, we conducted a mind-machine comparison experiment [45]. This experiment contrasted the classification performance differences among the proposed model, immunologists, and immunologists and re-classified using model assistance

**Table 2** Performance of CNN models used for IARI intensity classification

Method	$Accuracy_{avg}$ (%)		$Precision_{avg}$		$Recall_{avg}$		$F1-score_{avg}$		$Kappa$		$Time(s)$	
	w/o	w	w/o	w	w/o	w	w/o	w	w/o	w	w/o	w
AlexNet	92.6	92.8 ↑	0.905	0.897	0.955	0.965 ↑	0.929	0.930 ↑	0.919	0.921 ↑	0.019	<b>0.006</b>
VGG	92.4	90.5	0.903	0.872	0.942	0.984 ↑	0.922	0.925 ↑	0.917	0.896	0.070	0.091
ResNet	91.3	93.0 ↑	0.878	0.907 ↑	0.954	0.960 ↑	0.914	0.933 ↑	0.905	0.924 ↑	0.029	0.032
Inception	94.6	94.6	0.922	0.929 ↑	0.974	0.975 ↑	0.948	0.951 ↑	0.941	0.941	0.023	0.077
DenseNet	92.9	93.1 ↑	0.908	0.919 ↑	0.953	0.927	0.930	0.923	0.923	0.925 ↑	0.043	0.049
Ensemble Model	99.6	<b>99.8</b> ↑	0.972	<b>0.975</b> ↑	0.987	<b>0.991</b> ↑	0.979	<b>0.983</b> ↑	0.987	<b>0.991</b> ↑	0.078	0.094

Note: w/o represents CNN without CBAM; w represents CBAM-CNN; Ensemble Model denotes the results of the proposed model; “↑” indicates that the result of CBAM-CNN is better than that of the original CNN

(human-machine integration experiment), and especially, the classification performance of each category. The classification performances for each category and the average of the proposed model, three immunologists (Immunologist-1 and Immunologist-2 with about 2 years of experience each, and Immunologist-3 with 5 years of experience), and that of immunologists with model assistance are shown in Table 3. The proposed model had a much higher classification performance than the three immunologists ( $Accuracy_{avg}$ : 99.8% vs. 83.8%, 85.5%, 95.6%;  $F1-score_{avg}$ : 0.983 vs. 0.478, 0.456, 0.845;  $Kappa$ : 0.991 vs. 0.469, 0.500, 0.844;  $p < 0.01$ ). Notably, in the classification of the (3+) category, our proposed model showed a more remarkable classification performance than the immunologist-avg (99.4% vs. 86%;  $F1-score_{avg}$ : 0.933 vs. 0.487;  $p < 0.01$ ). With model-assisted prediction, the performance was highly improved for all three immunologists ( $Accuracy_{avg}$ : 88.3% vs. 94.4%;  $Precision_{avg}$ : 0.588 vs. 0.749;  $Recall_{avg}$ : 0.599 vs. 0.776;  $F1-score_{avg}$ : 0.593 vs. 0.757;  $Kappa$ : 0.604 vs. 0.805;  $p < 0.01$ ). For the (4+) category classification, the  $F1-score$  values achieved by the immunologists all exceeded 0.9 ( $F1-score$ : 0.954, 0.925, and 0.922;  $p < 0.01$ ). In particular, the time taken for model classification was  $0.094s \cdot image^{-1}$ , which was approximately 60 times faster than that taken by the immunologists ( $Immunologist_{avg} = 5.528s \cdot image^{-1}$ ).

Further, to reflect the results intuitively, confusion matrices [46] were generated for the classification results of the ensemble model, immunologists, and immunologists with model assistance as shown in Fig. 3. Confusion matrices are cross-compared, predicted, and used for actual decision classes as a widespread approach in deep learning. In confusion matrices, the abscissa represents the true label, and the ordinate represents the predicted label. In the red area, the numbers indicate the amount of data predicted for each category. The deeper red area indicates the larger amount of data predicted in the category, the deeper red area on the main diagonal indicates the higher accuracy of prediction, and the other lighter areas indicate a lower error rate. In the green area, the percentages in the right-most column

represent the  $Precision$  of each category, the percentages in the bottom row represent the  $Recall$  of each category, and the percentage in the lower right corner represents total  $Accuracy$ . The deeper green area indicates better prediction performance. As shown in Fig. 3, the total accuracy of the immunologists was also improved ( $Accuracy$ : Imm-1, 59.5% vs. 73.3%, Imm-2, 89.0% vs. 92.3% and Imm-3, 89.0% vs. 92.3%). In the model-assisted experiment, Immunologist-3 achieved the highest accuracy ( $Accuracy_{total} = 92.3\%$ ), which was close that of the model ( $Accuracy_{total} = 99.4\%$ ).

## 4 Discussion

This study presents a fast, fully automatic deep learning model based on CNNs for IARI classification. It is an end-to-end hybrid processing method combining the ensemble model with the CBAM and RPV modules that can accurately divide the intensity of IARI into five categories. This decision-level fusion design dramatically improved the classification efficiency and precisely fitted the IARI dataset compared with the independent model. More importantly, the proposed model achieved better classification performance than the immunologists and effectively improved the classification accuracy of the immunologists.

Table 2 shows the results of the macro-averaged metrics of the single and ensemble models for IARI multi-classification. In the IARI classification, all the models achieve an  $Accuracy_{avg}$  of more than 90% in the overall categories because the deep learning models can automatically mine the subtle and deep features related to the IARI, which cannot be perceived manually. However, there are differences in the performance of the single models in different category classifications, as shown in the Supplementary Material Table S1, the accuracy of the ResNet model in the (-) and (3+) categories were 77.6% and 84.7%, respectively, and the accuracy of the DenseNet model in the (-) category was only 79.2%. Additionally, compared with the single model, the ensemble model substantially improves the accuracy of

**Table 3** Comparison between the ensemble models and three immunologists in each sub-category

Method	Category	Accuracy (%)		Precision		Recall		F1-score		Kappa		Time(s)
		w/o	w	w/o	w	w/o	w	w/o	w	w/o	w	
Imm-1	(-)	84.4	88.7 ↑	0.965	0.989 ↑	0.631	0.723 ↑	0.763	0.835 ↑	-	-	-
	(1+)	75.5	80.7 ↑	0.146	0.327 ↑	0.152	0.348 ↑	0.149	0.337 ↑	-	-	-
	(2+)	86.5	93.3 ↑	0.059	0.000	0.143	0.000	0.084	-	-	-	-
	(3+)	81.3	87.1 ↑	0.254	0.382 ↑	0.692	1.000 ↑	0.372	0.553 ↑	-	-	-
	(4+)	91.4	96.9 ↑	0.966	0.972 ↑	0.773	0.936 ↑	0.859	0.954 ↑	-	-	-
	avg		83.8	89.3 ↑	0.478	0.534 ↑	0.478	0.601 ↑	0.478	0.566 ↑	0.469	0.637 ↑
Imm-2	(-)	83.1	99.1 ↑	0.838	0.985 ↑	0.715	0.992 ↑	0.772	0.988 ↑	-	-	-
	(1+)	81.3	98.2 ↑	0.174	0.935 ↑	0.087	0.935 ↑	0.116	0.935 ↑	-	-	-
	(2+)	89.6	97.2 ↑	0.045	0.778 ↑	0.071	0.500 ↑	0.055	0.609 ↑	-	-	-
	(3+)	83.4	94.5 ↑	0.274	0.611 ↑	0.654	0.846 ↑	0.386	0.710 ↑	-	-	-
	(4+)	90.2	95.1 ↑	0.861	0.952 ↑	0.845	0.900 ↑	0.853	0.925 ↑	-	-	-
	avg		85.5	96.8 ↑	0.438	0.852 ↑	0.474	0.835 ↑	0.456	0.843 ↑	0.500	0.886 ↑
Imm-3	(-)	96.6	100 ↑	0.922	1.000 ↑	1.000	1.000	0.959	1.000 ↑	-	-	-
	(1+)	95.1	98.5 ↑	1.000	1.000	0.652	0.891 ↑	0.789	0.942 ↑	-	-	-
	(2+)	97.9	97.9	0.769	0.769	0.714	0.714	0.740	0.740	-	-	-
	(3+)	93.3	93.3	0.542	0.542	1.000	1.000	0.703	0.703	-	-	-
	(4+)	95.1	95.1	1.000	1.000	0.855	0.855	0.922	0.922	-	-	-
	avg		95.6	97.0 ↑	0.847	0.862 ↑	0.844	0.892 ↑	0.845	0.861 ↑	0.844	0.892 ↑
Imm-avg	(-)	88	95.9 ↑	0.908	0.991 ↑	0.782	0.905 ↑	0.831	0.941 ↑	-	-	-
	(1+)	84	92.5 ↑	0.440	0.754 ↑	0.297	0.725 ↑	0.351	0.738 ↑	-	-	-
	(2+)	91.3	96.1 ↑	0.291	0.516 ↑	0.309	0.405 ↑	0.293	0.675 ↑	-	-	-
	(3+)	86	91.6 ↑	0.357	0.512 ↑	0.782	0.949 ↑	0.487	0.655 ↑	-	-	-
	(4+)	92.2	95.7 ↑	0.942	0.959 ↑	0.824	0.912 ↑	0.878	0.935 ↑	-	-	-
	avg		88.3	94.4 ↑	0.588	0.749 ↑	0.599	0.776 ↑	0.593	0.757 ↑	0.604	0.805 ↑
Model	(-)	100	100	1	1	1	1	1	1	-	-	-
	(1+)	99.1	99.4 ↑	1	1	0.935	0.957 ↑	0.966	0.978 ↑	-	-	-
	(2+)	99.7	99.4	0.933	0.875	1	1	0.965	0.933	-	-	-
	(3+)	99.3	100 ↑	0.929	1 ↑	1	1	0.963	1 ↑	-	-	-
	(4+)	100	100	1	1	1	1	1	1	-	-	-
	avg		99.6	<b>99.8</b> ↑	0.972	<b>0.975</b> ↑	0.987	<b>0.991</b> ↑	0.979	<b>0.983</b> ↑	0.987	<b>0.991</b> ↑

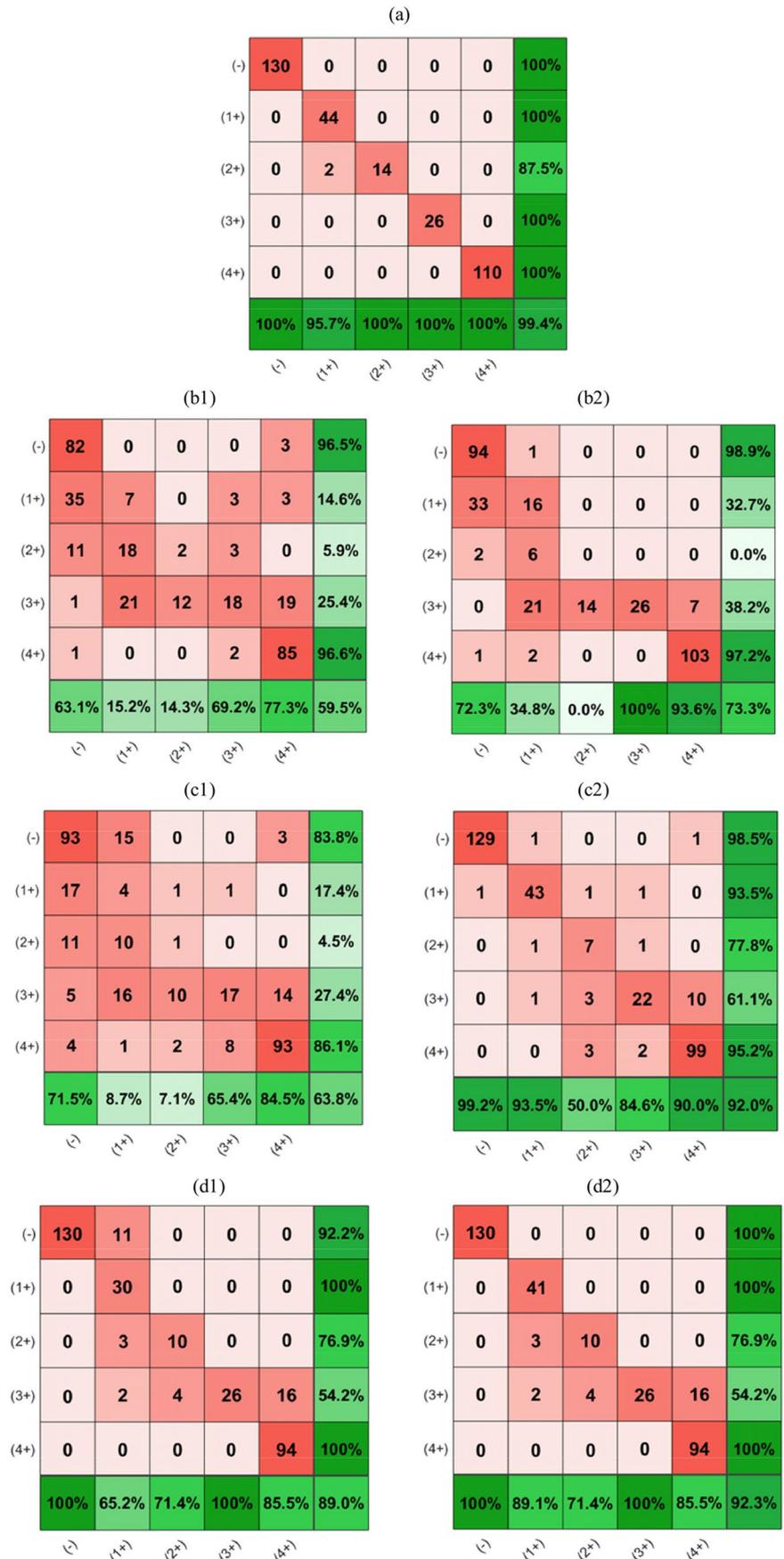
Notes: Imm-n denotes Immunologist-1, Immunologist-2, Immunologist-3, and Immunologist-avg; w/o represents the immunologist without model assistance; w represents the immunologist with model assistance; “↑” indicates that the result of the immunologist with model assistance is better than that of the immunologists

classification both in single categories and overall. As shown in Table 3, the accuracies of all categories were above 99%, and the maximum improvement in the overall accuracy was up to 8.3% ( $Accuracy_{avg}$ : ResNet 91.3% vs. ensemble model 99.6%). The ensemble model is efficient for the improvement of the model fit; however, it is not sensitive to outliers for reducing the decision boundary shift [47–51]. In addition, the ensemble model, by collective decision mechanism, focuses on synthesizing information from several sub-models with different structures and has been shown to reduce average error and combine the strengths of models in the exploration of diverse data patterns [52–54]. However, the addition of a poorly performing model will not reduce the overall model classification skill, because the ensemble

model has a net gain compared to the single models [55, 56]. Given the above, the ensemble model can reduce the risk of relying on a single prediction distribution and extract richer semantic feature information than the single CNN models (such as each sub-model in the training process has a different probability for boundary regions in pixel-level), which are beneficial in classification tasks to or the achievement of better performance to improve classification accuracy [57–62].

As shown in Table 2, CBAM has a limited effect on overall model performance improvement in that it slightly increases the accuracy of the models except that of the VGG and Inception models. But CBAM reduces the cross-adjacent category errors, especially those of the CBAM-CNN

**Fig. 3** Confusion matrices for IARI intensity classification; **(a)** The proposed ensemble model; **(b)** Immunologist-1 manually and with model assistance; **(c)** Immunologist-2 manually and with model assistance; **(d)** Immunologist-3 manually and with model assistance



models, thus improving the accuracy of the (4+) category shown in the Supplementary Material Table S1 and reducing the error rate of blood artefacts being mistaken as (4+) category. CBAM flexibly introduced into various models, partially reserves the channel interaction information and spatial location information while gathering clues about actual class object features and giving a meaningful focus for the input images by element-wise operations [63–74]. Thus, the CBAM-CNN models bring more robust and plausible classification decision-making. Table 3 shows that the ensemble model with CBAM, which we proposed, gives the best performance among all models both in the single categories and in the overall.

We further compared the classification performance of the immunologists and our deep learning model. As illustrated in Table 3, the performance of the model was higher than the average performance of three junior immunologists with varying experience. For the immunologists, (-) and (4+) categories were relatively easier to classify, whereas (1+), (2+), and (3+) classification were more prone to errors. The results show that the more experienced immunologist had a classification of higher accuracy. The immunologists also conducted reclassification with the assistance of our proposed model to verify the clinical utility of the model. In the (1+), (2+), and (3+) categories, the performance of immunologists was greatly improved; especially for immunologists with relatively low experience, the auxiliary effect is more obvious. Further, analysis of the time required for classification using the model and by the immunologists showed that the calculation time of the proposed model is at the millisecond level and is 60 times faster than the time needed manually. Thus, the model holds great potential for real-time assistance, especially for junior immunologists.

The confusion matrices of the ensemble model for classification tasks clearly and intuitively showed the classification performance of the models and immunologists across each category. Higher accuracy rates generally indicate better results, but *FP* and *FN* are also important and should not be ignored in clinical medicine. Reducing the ratio of *FP* to *FN* can significantly reduce the possibility of medical errors. There were no serious errors in the IARI classification using the ensemble models: no strongly positive IARI ((4+) category) was identified as poorly positive ((1+), (2+), (3+) categories) or negative category ((-) category), and no negative IARI was identified as a positive category. The model also did not show across-adjacent category errors within the poorly positive categories. However, immunologists usually misclassified these samples as (-) and (1+), (3+), and (4+), and the internal categories of poor positive samples, resulting in unsatisfactory classification results. Our analyses revealed that automatic classification is feasible and reliable and can significantly outperform the immunologists. We observed

that the performance of the immunologists was highly improved with assistance of the proposed model: the main diagonal became deeper, indicating that the accuracy of each category increased; the other areas became lighter indicating that the errors in the across-adjacent categories were reduced; the green area became deeper, indicating an improvement in *Accuracy<sub>total</sub>*, *Precision*, and *Recall*. These results demonstrate that our proposed model may be used as a reference for assisting immunologists.

In addition, in this experiment, batch size was set to 32 using the data-parallelization strategy to adapt to the IARI dataset, which can train the CNN models in the correct direction of change of gradients to be able to accurately classify IARI [75–78]. The learning rate of  $5e-4$  with the fixed batch size can keep the generalization performance from being degraded and makes CNN models achieve the best performance because of small batch sizes requiring small learning rate [79–81]. The epoch was set to 50 to terminate the training because the models all achieved stable convergence. The models using the well-designed parameters are robust and achieve good results.

However, our research had some limitations. First, our datasets were derived from the same source, and the AHG test was performed in the same laboratory, which contained limited variances; thus, the generalisation of the model needs to be externally verified at multiple centres. Second, the model could only differentiate the negative reaction, poor positive reaction, and strongly positive reaction. In clinical reality, there is a suspected category, ( $\pm$ ), which is indistinguishable from negative and poor positive reactions by immunologists that have insufficient images to build the CNN model. Thus, in future, extending the dataset to multivendor and multi-centre platforms may further improve the performance of the model. Simultaneously, we will qualitatively distinguish the boundary between ( $\pm$ ) and other classes described in this study.

## 5 Conclusion

In this study, we presented a deep ensemble learning model based on CNN models that can accurately classify IARI into multiple categories. This model can aid immunologists in differentiating distinct clinical patients by providing an objective and accurate evaluation of IARI categories, which could reduce the risk of haemolytic diseases. The model holds great potential in the field of fully automatic machinery and holds promise for promoting intelligent AHG test classification.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11517-022-02523-1>.

**Acknowledgements** We thank the Wiley Editing Services for editorial assistance.

**Author contribution** Xin Gao designed the study; HongMei Wang performed data collection; HongMei Wang ShengBao Duan, and YuJue Wang labelled data; and KeQing Wu performed the research and wrote the paper.

**Funding** This work was supported by the Development Program of Guangdong Province [grant number: 2019B010152001]; the Key Research and Development Program of Jiangsu [grant number: BE2021663]; the Research Fund of Jihua Laboratory [grant number: X190171TD190]; the Special fund for high-tech industrialisation of science and technology cooperation between Jilin Province and Chinese Academy of Sciences [grant number: 2020SYHZ0025]; the National Natural Science Foundation of China [grant number: 81871439]; and the Shandong Province Department of Science and Technology [grant number: ZR2020QF019].

## Declarations

**Competing interests** The authors declare no competing interests.

## References

- Vincent JLBJ, Reinhart K et al (2002) Anemia and blood transfusion in critically ill patients. *JAMA* 288(12):1499–1507. <https://doi.org/10.1001/jama.288.12.1499>
- Drews RE (2003) Critical issues in hematology: anemia, thrombocytopenia, coagulopathy, and blood product transfusions in critically ill patients. *Clin Chest Med* 24:607–622. [https://doi.org/10.1016/s0272-5231\(03\)00100-x](https://doi.org/10.1016/s0272-5231(03)00100-x)
- White J (2009) Pre-transfusion testing. *ISBT Sci Ser* 4(1):37–44. <https://doi.org/10.1111/j.1751-2824.2009.01211.x>
- Brand A (2002) Immunological aspects of blood transfusions. *Transpl Immunol* 10(2–3):183–190
- Wilhelm DKH, Klouche M et al (1995) Impact of allergy screening for blood donors: relationship to nonhemolytic transfusion reactions. *Vox Sang* 69(3):217–221
- Wang HM, Chen YZ, Ding SH et al (2015) A new approach to detection of incomplete antibodies using hydrogel chromatography medium. *Transfus Apheres Sci* 53:337–341. <https://doi.org/10.1016/j.transci.2015.06.003>
- Coombs RRA, Mourant AE, Race RR (1945) A new test for the detection of weak and incomplete Rh agglutinins. *Br J Exp Pathol* 26(4):255
- Race RR (1944) An “incomplete” antibody in human serum. *Nature* 153(3895):771–2
- Nedelcu E (2013) Pre-analytical issues and interferences in transfusion medicine tests. In: *Accurate Results in the Clinical Laboratory: A Guide to Error Detection and Correction*. Elsevier Inc., pp 273–294. <https://doi.org/10.1016/B978-0-12-415783-5.00017-7>
- Lapierre Y, Rigal D, Adam J, Josef D, Meyer F, Greber S, Drot C (1990) The gel test: a new way to detect red cell antigen-antibody reactions. *Transfusion* 30:109–113. <https://doi.org/10.1046/j.1537-2995.1990.30290162894.x>
- Shastri S, Murugesan M, Bhat S (2014) Mixed field agglutination: unusual causes and serological approach. *Asian J Transfus Sci* 8:4. <https://doi.org/10.4103/0973-6247.126680>
- Alwar V, Devi AMS, Sitalakshmi S et al (2012) Evaluation of the use of gel card system for assessment of direct coombs test: weighing the pros and cons. *Indian J Hematol Blood Transfus* 28(1):15–18. <https://doi.org/10.1007/s12288-011-0098-7>
- Liu Q, Cheng XN, Fei HX (2011) Effects of micro-magnetic field at the surface of 316L and NiTi alloy on blood compatibility. *Med Biol Eng Comput* 49:359–364. <https://doi.org/10.1007/s11517-010-0685-z> 2010/10/12
- Lemm WUV, Bücherl ES (1980) Blood compatibility of polymers: in vitro and in vivo tests. *Med Biol Eng Comput* 18(4):521–526
- Plapp FV, Sinor LT, Rachel JM et al (1984) A solid phase antibody screen. *Am J Clin Pathol* 82(6):719–721. <https://doi.org/10.1093/ajcp/82.6.719>
- LA Sandler SG, Rumsey DH, Novak SC (2000) A solid phase and microtiter plate hemagglutination method for pretransfusion compatibility testing. *Haematologia* 30(3):149–157. <https://doi.org/10.1163/156855900300109143>
- Guignier F, Domy M, Angue M, Richaud P, Chatelain P (1988) Comparison between a solid-phase low-ionic-strength solution antiglobulin test and conventional low-ionic-strength antiglobulin test: assessment for the screening of antierythrocyte antibodies. *Vox Sang* 55(1):30–34
- Sigdel A, Chalise G, Bolideei M et al (2021) Comparison between the manual method of indirect coombs via gel technology and solid phase red cell adherence. *Maedica (Bucur)* 16:200–206. <https://doi.org/10.26574/maedica.2021.16.2.200>
- Cate JC IV, Reilly N (1999) Evaluation and implementation of the gel test for indirect antiglobulin testing in a community hospital laboratory. *Arch Pathol Lab Med*. 123(8):693–697
- Aristov A, Nosova E (2017) Method of evaluation of process of red blood cell sedimentation based on photometry of droplet samples. *Med Biol Eng Comput* 55:621–630. <https://doi.org/10.1007/s11517-016-1542-5> 2016/07/13
- Chen Y, Wang M, Ding S et al (2019) A new reliable test for crossmatching: microplate hydrogel immunoassay technology. *Transfus Med* 29(5):344–350. <https://doi.org/10.1111/tme.12631>
- Roman L, Armstrong B, Smart E (2020) Principles of laboratory techniques. *ISBT Sci Ser* 15:81–111. <https://doi.org/10.1111/voxs.12591>
- Er Saw P, Jiang S (2020) The significance of interdisciplinary integration in academic research and application. *BIO Integration* 1:2–5. <https://doi.org/10.15212/bioi-2020-0005>
- Liu C, Jia D, Liu Z (2020) Artificial intelligence (AI)-aided disease prediction. *BIO Integration* 1:130–136
- Zhang K, Liu X, Shen J et al (2020) Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* 181:1423–1433.e1411. <https://doi.org/10.1016/j.cell.2020.04.045>
- Das N, Hussain E, Mahanta LB (2020) Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network. *Neural Netw* 128:47–60. <https://doi.org/10.1016/j.neu-net.2020.05.003>
- Tessema AW, Mohammed MA, Simegn GL et al (2021) Quantitative analysis of blood cells from microscopic images using convolutional neural network. *Med Biol Eng Comput* 59:143–152. <https://doi.org/10.1007/s11517-020-02291-w>
- Mohanty SP, Hughes DP, Salathe M (2016) Using deep learning for image-based plant disease detection. *Front Plant Sci* 7:1419. <https://doi.org/10.3389/fpls.2016.01419>
- Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. *Commun ACM* 60(6):84–90. <https://doi.org/10.1145/3065386>
- Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings

31. He K, Zhang X, Ren S et al (2016) Deep residual learning for image recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition
32. Szegedy C, Vanhoucke V, Ioffe S et al (2016) Rethinking the inception architecture for computer vision. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition
33. Huang G, Liu Z, Van Der Maaten L et al (2017) Densely connected convolutional networks. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017
34. Woo S, Park J, Lee JY et al (2018) CBAM: Convolutional block attention module. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)
35. Demrekler M, Altncay H (2002) Plurality voting-based multiple classifier systems: statistically independent with respect to dependent classifier sets. *Pattern Recogn* 35:2365–2379. [https://doi.org/10.1016/S0031-3203\(01\)00227-8](https://doi.org/10.1016/S0031-3203(01)00227-8)
36. Mu X, Watta P, Hassoun MH (2009) Analysis of a plurality voting-based combination of classifiers. *Neural Process Lett* 29(2):89–107. <https://doi.org/10.1007/s11063-009-9097-1>
37. Özgür A ÖL, Güngör T (2005) Text categorization with class-based and corpus-based keyword selection. *International Symposium on Computer and Information Sciences Springer, Berlin, Heidelberg*; 606–615
38. Yang Y LX (1999) A re-examination of text categorization methods. Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval; 42–49
39. Lewis DD, Schapire RE, Callan JP, Papka R (1996) Training algorithms for linear text classifiers. Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval; 298–306
40. Harman DK (1995) Overview of the third text retrieval conference (TREC-3). DIANE Publishing
41. Phoungphol PZY, Zhao Y (2012) Robust multiclass classification for learning from imbalanced biomedical data. *Tsinghua Sci Technol* 17(6):619–628
42. Hardin PJ, Shumway JM (1997) Statistical significance and normalized confusion matrices. *Photogramm Eng Remote Sensing* 63(6):735–9
43. Kingma DP, Ba JL (2015) Adam: A method for stochastic optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings
44. Grandini M BE, Visani G (2020) Metrics for multi-class classification: an overview. arXiv preprint arXiv; 2008:05756
45. Zhao X, Wang X, Xia W et al (2020) A cross-modal 3D deep learning for accurate lymph node metastasis prediction in clinical stage T1 lung adenocarcinoma. *Lung Cancer* 145:10–17. <https://doi.org/10.1016/j.lungcan.2020.04.014>
46. Düntsch I, Gediga G (2019) Confusion matrices and rough set data analysis. *J Phys: Conf Ser* 1229(1):012055. <https://doi.org/10.1088/1742-6596/1229/1/012055>
47. Hu R, Zhou S, Liu Y et al (2019) Margin-based pareto ensemble pruning: an ensemble pruning algorithm that learns to search optimized ensembles. *Comput Intell Neurosci* 2019:7560872. <https://doi.org/10.1155/2019/7560872>
48. Cruz RMO, Sabourin R, Cavalcanti GDC (2018) Dynamic classifier selection: recent advances and perspectives. *Information Fusion* 41:195–216. <https://doi.org/10.1016/j.inffus.2017.09.010>
49. Zhang C-X, Zhang J-S (2011) A survey of selective ensemble learning algorithms. *Chin J Comput* 34(8):1399–1410
50. Y L (2012) New discoveries in balanced ensemble learning. The 2012 International Joint Conference on Neural Networks (IJCNN); IEEE, 2012:1–8
51. Huang J, Gao J (2017) An ensemble simulation approach for artificial neural network: an example from chlorophyll a simulation in Lake Poyang, China. *Ecol Inform* 37:52–58. <https://doi.org/10.1016/j.ecoinf.2016.11.012>
52. Weigel AP, Liniger MA, Appenzeller C (2008) Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Q J R Meteorol Soc* 134:241–260. <https://doi.org/10.1002/qj.210>
53. Hacker JP, Lei L (2015) Nudging, ensemble, and nudging ensembles for data assimilation in the presence of model error. *Mon Weather Rev* 143:2600–2610. <https://doi.org/10.1175/mwr-d-14-00295.1>
54. Yoo JH, Kang I-S (2005) Theoretical examination of a multi-model composite for seasonal prediction. *Geophys Res Lett* 32:n/a-n/a. <https://doi.org/10.1029/2005gl023513>
55. Zador Paul LD (2013) Evaluating perturbation impact on key travel models, census statistical disclosure control research project 1. westat
56. Wu H, Levinson D (2021) The ensemble approach to forecasting: a review and synthesis. *Transport Res Part C: Emerg Technol* 132:103357. <https://doi.org/10.1016/j.trc.2021.103357>
57. Webb GIZZ (2004) Multistrategy ensemble learning: reducing error by combining ensemble learning techniques. *IEEE Trans Knowl Data Eng* 16(8):980–991
58. Dietterich TG (2000) An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. *Mach Learn* 40(2):139–157
59. Lee B (1998) Arcing classifiers. *Ann Stat* 26(3):801–849
60. Elder SLaJF (1997) Bundling heterogeneous classifiers with advisor perceptrons, tech report 97–1, Elder Research, Charlottesville
61. Margineantu DDTG (1997) Pruning adaptive boosting. *ICML* 97:211–218
62. Kamiya R SK, Hotta K (2019) Ensemble of training models for road and building segmentation. 2019 Digital image computing: techniques and applications (DICTA); IEEE, 2019:1–6
63. Lin M CQ, Yan S (2013) Network in network. arXiv preprint arXiv; 1312.4400
64. Hu J SL, Sun G (2018) Squeeze-and-excitation networks. Proceedings of the IEEE conference on computer vision and pattern recognition; 7132–7141
65. Zhang T, Zhang X (2022) A polarization fusion network with geometric feature embedding for SAR ship classification. *Pattern Recogn* 123:108365. <https://doi.org/10.1016/j.patcog.2021.108365>
66. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A (2016) Learning deep features for discriminative localization. Proceedings of the IEEE conference on computer vision and pattern recognition; 2921–2929
67. Jarrett K KK, Ranzato MA, LeCun Y (2009) What is the best multi-stage architecture for object recognition? IEEE 12th international conference on computer vision; 2146–2153
68. Yang J, Yu K, Gong Y, Huang T (2009) Linear spatial pyramid matching using sparse coding for image classification. IEEE Conference on computer vision and pattern recognition; 1794–1801
69. Serre T WL, Poggio T (2005) Object recognition with features inspired by visual cortex. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05); 2: 994–1000
70. Xue G, Liu S, Ren L et al (2022) Forecasting the subway passenger flow under event occurrences with multivariate disturbances. *Exp Syst Appl* 188:116057. <https://doi.org/10.1016/j.eswa.2021.116057>
71. Chen F, Tsou JY (2021) DRNet: Novel architecture for small patch and low-resolution remote sensing image scene classification. *Int J Appl Earth Observ Geoinform* 104:102577. <https://doi.org/10.1016/j.jag.2021.102577>
72. Boureau YL, Bach F, LeCun Y, Ponce J (2010) Learning mid-level features for recognition. IEEE computer society conference on computer vision and pattern recognition; 2559–2566

73. Yang L, Zhang F, Wang PS-P et al (2022) Multi-scale spatial-spectral fusion based on multi-input fusion calculation and coordinate attention for hyperspectral image classification. *Pattern Recogn* 122:108348. <https://doi.org/10.1016/j.patcog.2021.108348>
74. Yu W, Xu H (2022) Co-attentive multi-task convolutional neural network for facial expression recognition. *Pattern Recogn* 123:108401. <https://doi.org/10.1016/j.patcog.2021.108401>
75. Lee S, Kang Q, Madireddy S, Balaprakash P, Agrawal A, Choudhary A, Liao WK (2019) Improving scalability of parallel CNN training by adjusting mini-batch size at run-time. 2019 IEEE International Conference on Big Data (Big Data); 830–839
76. Kandel I, Castelli M (2020) The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express* 6:312–315. <https://doi.org/10.1016/j.icte.2020.04.010>
77. Bengio Y (2012) Practical recommendations for gradient-based training of deep architectures. *Neural networks: Tricks of the trade*; 437–478
78. Xu H vGJ, Xiong D, Liu, Q (2020) Dynamically adjusting transformer batch size by monitoring gradient direction change. *arXiv preprint arXiv*; 2005.02008
79. Sutskever I, Martens J, Dahl G, Hinton G (2013) On the importance of initialization and momentum in deep learning. *International conference on machine learning*; 1139–1147
80. Radiuk PM (2017) Impact of training set batch size on the performance of convolutional neural networks for diverse datasets. *Inf Technol Manag Sci* 20. <https://doi.org/10.1515/itms-2017-0003>
81. Masters D LC (2018) Revisiting small batch training for deep neural networks. *arXiv preprint arXiv*; *arXiv preprint arXiv*

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**KeQing Wu** is a master student in Shanghai University and Suzhou Institute of Biomedical Engineering and Technology, whose main research direction is precision medicine based on intelligent computing.



**ShengBao Duan** is a professor of CAS Key Lab of Bio-Medical Diagnostics, mainly researching the development of in vitro diagnostic reagents related to blood transfusion safety.



**YuJue Wang** is a member of CAS Key Lab of Bio-Medical Diagnostics, mainly researching the development of in vitro diagnostic reagents related to blood transfusion safety.



**HongMei Wang** is a professor at CAS Key Lab of Bio-Medical Diagnostics, mainly researching the development of in vitro diagnostic reagents related to blood transfusion safety.



**Xin Gao** is a professor at the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science.