CRFNN: CONTEXTUAL REASONING-BASED FUZZY NEURAL NETWORKS FOR THE UNCERTAINTY OUTPUT IN OBJECT RECOGNITION

ANDY L YANG[†] AND FENG GU[‡]

Abstract. Object recognition is an important task in image processing, and fuzzy neural networks (FNNs) are a commonly used tool employed for this purpose. However, when applied to datasets containing complex or ambiguous features, the accuracy of FNNs decreases and their uncertainty of output increases. To address this issue, in this paper, I propose a novel contextual reasoning-based fuzzy neural network method called CRFNN. The idea is to utilize the FNN model as an initial classifier to identify objects and acquire the predicted values of these objects, which then serve as the basis for subsequent contextual comparisons. The network output is integrated with a local searching strategy that involves feature ranking and the evaluation of the distance between contextual knowledge and the objects, ultimately leading to the determination of object categories. The developed algorithm's performance was assessed using three datasets, each characterized by complex and fuzzy image features. The experimental results demonstrate that the proposed CRFNN method achieves significantly higher accuracies, outperforming the traditional FNNs by 19.8%, 9.8%, and 11.2% on the three respective datasets. An additional analysis is further conducted using a comprehensive evalution metrics including accuracy (ACC), Hamming loss (HL), average precision (AP), weighted accuracy (WA), Recall, and the area under the ROC Curve (AUC), all of which consistently highlight the superior performance of the proposed CRFNN method over traditional FNNs as well as other existing baseline algorithms.

1. Introduction

Image processing involves the conversion of an image into a digital format, facilitating the utilization of image features such as color space regularization, image fusion, image filtering, and more, to achieve the desired outcome [14]. Recognizing objects in images is a fundamental task in the field of image processing [6], yet the inherent fuzziness of images poses significant challenges for accurate object recognition, prompting many studies to investigate the integration of fuzzy attributes to enhance recognition accuracy. Some well-known examples include the implementation of fuzzy logic to solve the 3D object recognition problem, where the fuzzy model is constructed utilizing

Key words and phrases. Image processing; Object recognition; Fuzzy neural network; Contextual knowledge; Contextual reasoning.

[†]Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA; Email: alyang@mit.edu.

[‡]Department of Computer Science, College of Staten Island, 2800 Victory Boulevard, Staten Island, NY 10304, USA; Email: feng.gu@csi.cuny.edu.

the invariant moments (Affine/Zernike/Hu) [30], the use of a neuro-fuzzy framework for the recognition of building objects and materials through a semi-automatic image-driven system [22], and the application of adaptive fuzzy-based network topology coupled with deep convolutional neural networks to automatically recognize predefined objects [4].

In recent years, fuzzy neural networks (FNNs), as a combination of fuzzy theory and neural networks, have received extensive attention in the field of computer vision. Their effectiveness has been demonstrated in a broad spectrum of engineering and scientific applications [18]. For instance, researchers have utilized FNNs to perform numerical diagnosis and identification of crankshaft bearings in different states, as detailed in [16], as well as in pattern classification, including the application of supervised fuzzy clustering and pruning algorithms to determine the exact number of clusters with appropriate centroids and widths [20]. In the context of object recognition, the handling and analysis of vague information have been at the forefront of FNNs research since the fuzzy set was first introduced by Zadeh [41], in which the fuzzy phenomenon was described with accurate numerical values.

In the field of image processing using neural network models, uncertainty arises from multiple sources, including the complex nature of object features, disparities between training and testing data, data incompleteness, and noise [2]. Moreover, additional uncertainties are introduced by expert experience and boundary conditions, see [35]. In light of these challenges, extensive efforts have been dedicated to object recognition using FNN models to mitigate the impact of uncertainty in image data. Here, we briefly summarize the pros and cons of these established methods with diverse application scenarios in image processing.

- **Image fusion.** To address image uncertainties when dealing with image fusions, in [28], the authors introduced a probabilistic fuzzy logic-based image fusion approach for merging multiple radar images of the same scene to produce a more informative synthetic image, however, the proposed method tends to retain most of the clutter and background noise in the fusion image. In [17], a reformulated fuzzy local-information C-means clustering algorithm was designed for the analysis of uncertain regions in fused images, however, determining the optimal parameters of the fuzzy membership function becomes very time-consuming as the distributions may differ among images.
- **Color space regularization.** Fuzzy color space normalization methods had been devised to address the challenges related to color space regularization problems. For staining color normalization, in [24] and [8], the authors proposed a rough-fuzzy circular clustering algorithm and a fuzzy transformation method to deal with issues such as uncertainty, fuzzy

nature, and incompleteness of color spaces, respectively. However, it is worth noting that these techniques, while effective, exhibit computational inefficiency and often fall short of generating images that maintain their natural appearance.

- **Image filtering for noise reduction.** Image noise exhibits significant fuzzy characteristics. In [27], the authors introduced an iterative fuzzy filtering technique framework to reduce the impulse noise in color images. In [40], the authors proposed a fuzzy logic-based image filter for detail-preserving restoration of images corrupted by impulse noise. Although these filtering techniques exhibit commendable filtering performance, their effectiveness typically comes at the expense of a noteworthy escalation in computational complexity.
- **Image retrieval.** The focus of image retrieval lies on how to speed up the retrieval process and improve accuracy. In the image retrieval process, the inherent fuzziness and uncertainty of image texture, color, layout, and other factors make it necessary to integrate fuzzy theory into retrieval algorithms. In [39], the authors introduced a retrieval scheme based on fuzzy rules, while in [21], the authors established a method to calculate the distance between fuzzy sets, aiding in patterns detection and images matching to enhance performance and shorten retrieval time. These methods may, however, impose additional burdens on users, particularly when more extensive image information is required.
- **Image segmentation.** Image segmentation is a technique that divides an image into multiple specific regions with unique properties to locate objects of interest [32]. To improve the accuracy of image segmentation, fuzzy theory has been introduced in this application. In [25], the authors proposed a variational model based on fuzzy clustering and total variation regularization to address the limitations of classical hard-labeled methods. In [5], the authors proposed an intuitive center-free fuzzy C-means clustering method for segmenting magnetic resonance brain images, in which the membership function is adopted to define the pixel-to-cluster similarities. In [9], the authors proposed a continuous fuzzy logic approach based a Markov random field model for carbon-13 image segmentation. In [10], the authors presented fuzzy C-means for spatially coherent image segmentation. However, it is worth noting that these methods still exhibit limitations in achieving high accuracy, especially when applied to noisy images, which results in restricted performance when handling uncertain outputs.
- **Object recognition.** To solve the problem of automatic detection and recognition of predefined objects, in [4] the authors proposed an adaptive FNN topology integrated in parallel with a deep convolutional neural network. This method improves detection and recognition

rates while enabling both fast and qualitative decision-making. In [37], the authors provided a methodology employing fuzzy clustering radial basis function neural networks to detect distorted and incomplete images. To determine the boundaries of image objects, in [36], a multi-scale contrast fuzzy discriminative segmentation (MCFDS) method that leverages spatial information to measure image objects and accurately delineate their boundaries was proposed. FNNs have also been widely used in medical, military and industrial fields. To illustrate, in the medical sector, the authors in [33] proposed two neuro-fuzzy classifier networks to improve the accuracy of electrocardiogram (ECG) signal classification. In the military domain, the authors in [3] proposed a two-coordinate radar-based FNN to idenity air objects. In the industrial context, the authors in [34] conducted an adaptive moving target tracking control scheme based on dynamic Petri recurrent FNN, aiming to solve the chattering phenomenon in mobile robot vision.

Despite substantial research and practical applications of fuzzy sets and FNNs in the field of object recognition, and the efforts of numerous models designed to address the uncertainty in FNN outputs, due to the absence of contextual elements (e.g., context reasoning [26, 42], context prediction [12], etc.), these FNN-based object recognition models, however, are still limited in their ability to deal with uncertainty outputs, notably affecting their accuracy in diverse scenarios.

Therefore, in this paper, we employ contextual reasoning to deal with the uncertainty of FNNs, proposing a novel, contextual reasoning-based fuzzy neural network model, called CRFNN. The idea is to initially use the FNN model as a classifier to identify objects, yielding predicted values for these objects, and using the predicted values as the basis for further context comparisons. Then, the network output is combined with a local search strategy, ultimately culminating in the definitive categorization of objects based on the ranking of features and the evaluation of the distance between contextual knowledge and the objects. To gauge the accuracy performance of the proposed CRFNN model, we conduct an assessment of the developed algorithm using three datasets, each characterized by complex and fuzzy image features. The obtained exprimental results indicate that the proposed CRFNN algorithm achieves significantly higher accuracies of 19.8%, 9.8%, and 11.2% when compared to the traditional FNN models when applied to the three respective datasets. Moreover, we perform additional analysis for the algorithm utilizing a variety of parametric metrics, such as accuracy (ACC), Hamming loss (HL), average precision (AP), weighted accuracy (WA), recall, and the Area under the ROC Curve (AUC), which consistently show that the proposed CRFNN model significantly outperforms both traditional FNNs and other established baseline algorithms.

The remainder of this article is structured as follows. Section 2 provides an introduction to fundamental concepts related to fuzzy sets and FNNs. Section 3 is dedicated to the discussion of output uncertainty, where key terms such as context, contextual knowledge, and reasoning are defined. Within this section, the design of the CRFNN model is presented, along with the detailed pseudo-code and a flowchart. In Section 4, we perform an empirical assessment of the proposed algorithm's accuracy using three distinct datasets consisting of images with inherent fuzziness. This evaluation involves a comparative analysis with FNN and other baseline algorithms, aiming to ascertain the CRFNN algorithm's superiority in relation to other well-established methods. Some concluding remarks are given in Section 5.

2. Preliminaries on fuzzy sets and FNNs

In this section, we provide a brief overview of the concepts and definitions pertaining to fuzzy sets and FNN. This preliminary knowledge will lay the groundwork for a more comprehensive understanding of the proposed CRFNN model in subsequent sections.

Definition 2.1. (Membership function and degree [41]) *Given a finite set U and a function μ maps the members in U to $[0,1]$, $\mu_A : A \to [0,1]$ is called the membership function of the subset $A \subseteq U$, and the value of μ_A is called the membership degree, which specifies the degree to which a given input belongs to the set A.*

Definition 2.2. (Fuzzy set [41]) *A fuzzy set is a pair (U, μ), where U is a set and $\mu : U \to [0,1]$ a membership function. For a finite set $A = \{x_1, x_2, \cdots, x_n\} \subseteq U$, the fuzzy set (A, μ_A) is often denoted by $\left\{\frac{\mu_A(x_1)}{x_1}, \frac{\mu_A(x_2)}{x_2}, \cdots, \frac{\mu_A(x_n)}{x_n}\right\}$.*

A fundamental network structure of FNNs consists of five key layers: a feature input layer, a fuzzy layer, a rule layer, a hidden layer, and an output layer [7, 15, 19]. Fig. 2.1 provides an illustrative example of a binary classification task, in which these five layers are visually represented. In the context of image processing, the input to the first layer is the image feature, denoted as θ^τ, and the overall input is expressed as $\theta = \{\theta^1, \theta^2, \cdots, \theta^m\}$. The output of the τ-th neuron is defined as:

$$x_\tau^{(1)} = \theta^\tau, \tau = 1, 2, \cdots, m. \tag{2.1}$$

The fuzzy layer uses membership functions corresponding to categories based on image feature. This layer has km neurons with l categories for the classification task. The input of the p-th neuron is $x_\tau^{(1)}$, and the output is the membership degree $\phi_p^{(2)}$ that is defined as:

$$\phi_p^{(2)} = \mu_c(x_\tau^{(1)}),\ p = 1, 2, \cdots, ml,\ c = 1, 2, \cdots, ml. \tag{2.2}$$

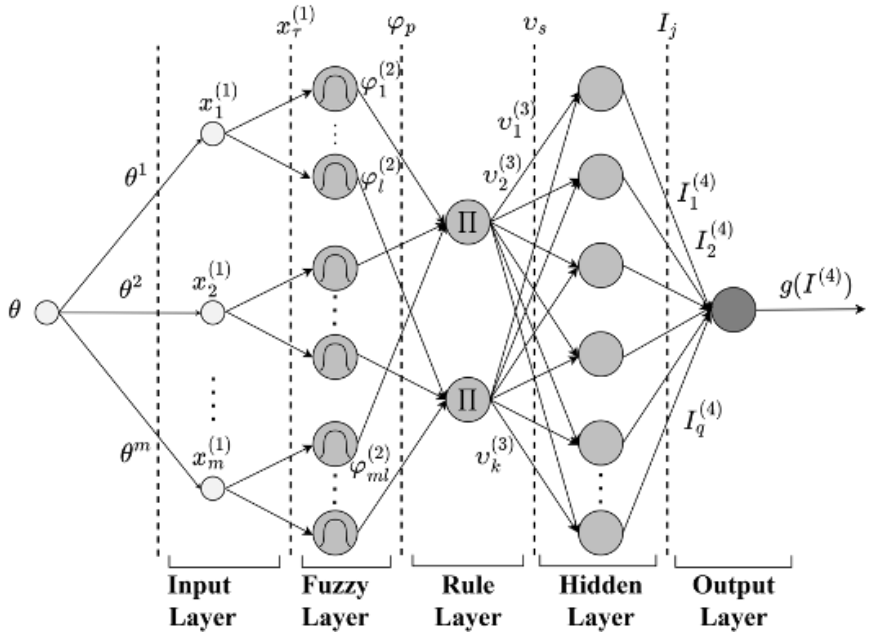


FIGURE 2.1. Sketch of layers of FNN of an example of a binary classification task.

The rule layer has k neurons, where the input is $\phi_p^{(2)}$, and the output is the result $v_s^{(3)}$ based on the membership multiplier, as shown below:

$$v_s^{(3)} = \Pi_{h=s}^{k(m-1)+s} \phi_p^{(2)}(x_l^{(1)}),\ h = s, k+s, 2k+s, \cdots, k(m-1)+s,\ s = 1, 2, \cdots, k. \tag{2.3}$$

The hidden layer has q neurons for forward propagation calculation. The input-output transformation relationship of each neuron is shown as

$$\begin{cases} I_j^{(4)} = \sum_{s=1}^{k} \omega_{js}^{(4)} v_s^{(3)}, \\ O_j^{(4)} = g(I_j^{(4)}),\ j = 1, 2, \cdots, q, \end{cases} \tag{2.4}$$

where $\omega_{js}^{(4)}$ represents the weight coefficient from the s-th neuron in the rule layer to the j-th neuron in the hidden layer.

The number of neurons in the hidden layer adopts the trial-and-error method, which is determined according to

$$q = \sqrt{\alpha + \beta} + \gamma, \tag{2.5}$$

where α is the number of input neurons, β is the number of output neurons, and γ is an integer falling within the interval $[1, 10]$, which needs to be roughly estimated and can be adjusted according to the training situation during the training process.

The input-output transformation relationship of each neuron in the last layer is shown as:

$$\begin{cases} I^{(5)} = \sum_{j=1}^{q} \omega^{(5)} O_j^{(4)}, \\ y = g(I_j^{(5)}),\ j = 1, 2, \cdots, q, \end{cases} \tag{2.6}$$

where $\omega^{(5)}$ is the activation function and could be the ReLU (rectified linear unit) function, see [11], as shown in

$$g(x) = \begin{cases} 0, x \leq 0 \\ x, x > 0. \end{cases} \tag{2.7}$$

The ReLU function described above is a piecewise linear function that zeros out all negative values while leaving positive values unchanged. ReLU is a crucial activation function in deep learning models as it introduces nonlinearity, addressing the problem of vanishing gradients. Its sparsity-inducing nature allows for improved feature identification and a better fit to the training data, making it an essential component in deep learning architectures.

3. CRFNN: contextual reasoning based FNN model

This section provides an in-depth exploration of the network's uncertain output and presents a comprehensive approach to mitigate this uncertainty through the application of contextual reasoning.

3.1. Uncertainty of network outputs.

Definition 3.1. (Uncertain outputs of networks) *Let the network output interval be* $[\epsilon, \zeta]$*, and specify another interval* (λ, ξ)*, which satisfies* $(\lambda, \xi) \subset [\epsilon, \zeta]$*. When the network output, represented as* y*, satisfies* $y \in [\epsilon, \lambda]$ *or* $y \in [\xi, \zeta]$*, the network output result is called the deterministic type, and when* $y \in (\lambda, \xi)$*, we call the network output as the uncertain type.*

For example, we consider a binary classification task where the output value y falls within the range of real numbers inside the interval $[0, 1]$. When y approaches the value of approximately 0.5, the task of definitively assigning it to one of the two classes becomes challenging, thereby introducing uncertainty into the classification task.

To address this uncertainty issue, this paper introduces a novel method that leverages contextual knowledge to aid in the decision-making process. The subsequent sections provide the relevant definitions of contextual knowledge and outline the methodology used to determine the network output.

3.2. **Contextual reasoning to reduce the uncertainty of network outputs.** Most FNN models only utilize the local information within an individual image, often ignoring the broader contextual knowledge or relationships among contextual images, which contain valuable global information. This omission of contextual information contributes to the uncertainty observed in the network's output. Therefore, it becomes imperative to integrate contextual knowledge into the object recognition process to enhance accuracy and reduce uncertainty.

Definition 3.2. (Context of the object)*: For the current object to be recognized, denoted as O_0, the previously recognized neighboring objects are called the context of the object, which can be formulated as*

$$Q = \{O_l, l = 1, 2, 3, \cdots, \sigma : O_l \text{ are neighboring objects of } O_0\} \tag{3.1}$$

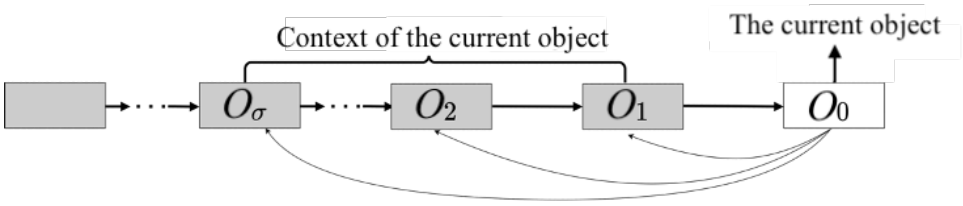


FIGURE 3.1. Context of object recognition.

The relationship between the current object O_0 and its neighboring objects O_l $(l = 1, 2, 3, \cdots, \sigma)$ is shown in Fig. 3.1. Here, the neighbors of the current object are considered as its context, with the context length denoted as σ $(\sigma = |Q|)$. To effectively manage the context of the current object, a queue data structure is well-suited. Queues exhibit the characteristic of first-in-first-out, aligning with the concept that objects farther from the current object have a weaker relationship, while those closer possess a stronger relationship. In the queue, objects that are farther away are dropped out earlier, aligning with their reduced relevance to the current object.

Definition 3.3. (Contextual knowledge) *Contextual knowledge (CK) is the category of the objects in the context, and can be expressed as*

$$CK = \{C : C = category(O_l), l = 1, 2, 3, \cdots, \sigma, O_l \in Q\}, \tag{3.2}$$

and CK will be used to determine the category of the object.

We consider the feature set of $\theta(O_l) = \{\theta^1, \theta^2, \cdots, \theta^m\}$, where θ represents the feature of the context object O_l. With the application of contextual reasoning, the distance between the current

Algorithm 1

Input: image sequence
Output: recognition result

Image pre-processing using image graying, binarization, and Gaussian filtering
Get the output y from FNN
if $y \notin [\lambda, \xi]$ **then**
    output y
else
    Initialize queue Q and list L
    Enqueue: Q = the contextual objects
    Insert: L = contextual objects sorted by their features from Q and $d_l < d_{lim}$
    repeat
    if $l < \sigma$ **then**
        Visiting $L[l]$ and calculate d_l
        if $d_l < d_{lim}$ **then**
            category(O_0)=category(O_l)
        else
            $l = l + 1$
    else
        fail to recognize and store results
until the object has been classified.

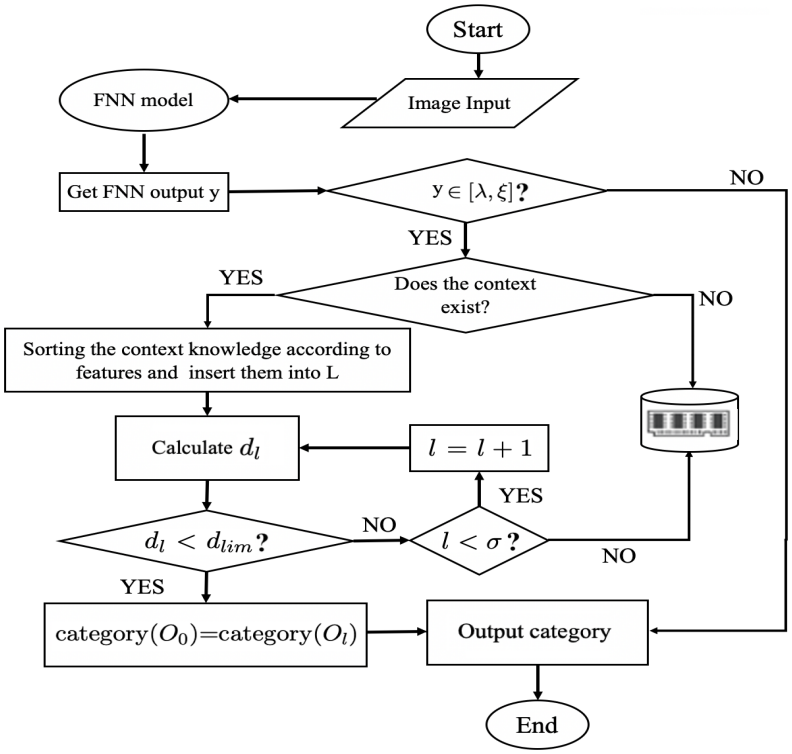


FIGURE 3.2. Flowchart of the CRFNN algorithm.

object and the objects in the context according to their normalized features are calculated as

$$d(O_0, O_l) = \sqrt{\sum_{\tau=1}^{m} \left(\frac{\theta_0^\tau - \theta_l^\tau}{\theta_{max}^\tau - \theta_{min}^\tau} \right)^2},\ l = 1, 2, \cdots, \sigma,\ \theta_0^\tau \in \theta(O_0),\ \theta_l^\tau \in \theta(O_l), \tag{3.3}$$

where θ_{min} and θ_{max} represent the minimum and maximum values of the feature in the training set. θ_0^τ represents the τ-th feature of the current object, and θ_l^τ represents the τ-th feature of the l-th object in the context. For convenience, $d(O_0, O_l)$ is abbreviated as d_l.

Definition 3.4. (Contextual reasoning)*: In situations where uncertainty arises in the output of FNNs, the algorithm utilizes contextual knowledge. The distance-based clustering method for contextual reasoning is adopted to solve the uncertainty. Specifically, if the distance d_l satisfies $d_l < d_{lim}$, where d_{lim} represents a predetermined threshold, it is determined that the current object belongs to the same category as object O_l, namely,*

$$category(O_0) = category(O_l),\ when\ d_l < d_{lim}. \tag{3.4}$$

3.3. **CRFNN algorithm.** Based on the previously described components, the CRFNN method is formulated as presented in Algorithm 1. This method comprises two parts: the FNN and the contextual reasoning, both of which have been elaborated upon in a comprehensive manner. Fig. 3.2 describes the flow chart of the process of object recognition based on CRFNN.

Algorithm 1 and the flowchart in Fig. 3.2 collectively offer a comprehensive insight into the functioning of the CRFNN algorithm. Here is a concise summary of the algorithm's operation. First, an image is input into the FNNs to obtain the output result, denoted as y. In accordance with Definition 3.1, if the network output is determined to be deterministic, the object's category is established based on the output result. However, when the network output is characterized as uncertain, it becomes necessary to judge whether the context of the object exists. If contextual information exists, we proceed to insert this contextual knowledge into the list, which is subsequently sorted according to the features. The features of the current object, along with those of the elements in the list, are then utilized to calculate distances for further measurement. Visiting each object in the list, if the distance is less than d_{lim}, the category of the current object is assigned to the same category as the object meeting this proximity criterion. When the context does not exist or all contexts in the list are not close to the current object, the algorithm fails to recognize and store results for the subsequent processing. Finally, we output the category of the object.

4. Experiments

In this section, we design related experiments and use various parameter metric/indicators to verify whether the proposed CRFNN algorithm is effective and superior to FNNs and other existing object recognition algorithms.

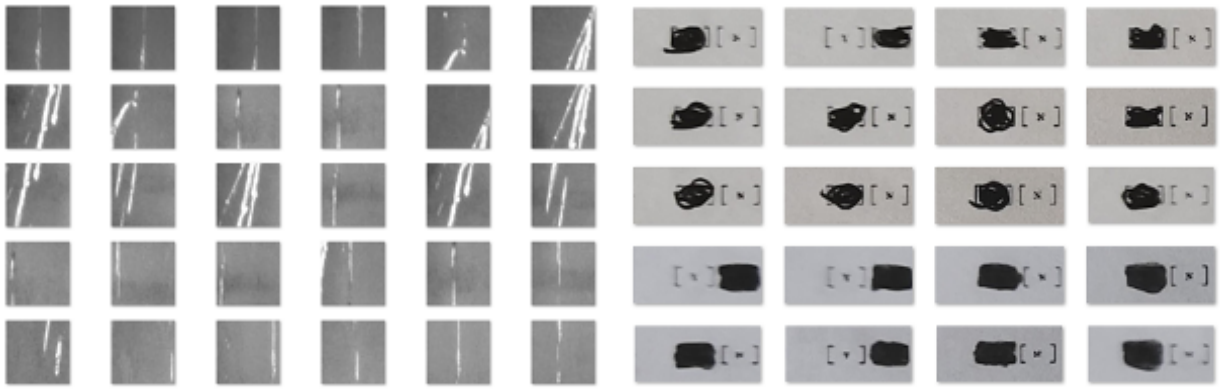
4.1. **Datasets.** We conduct experiments on a 64-bit Windows10 with an Intel i78700 CPU and 16G memory. The programming language is Python. We use two surface defect detection datasets (shown in Fig. 4.1(a) for Dataset 1 and Fig. 4.1(c) for Dataset 3), and one answer card filling point dataset (shown in Fig. 4.1(b) for Dataset 2), which are from NEU surface defect database [31], and the answer card filling point database [1], respectively. The detailed information of the datasets, including the number of training set images, test set images, etc., is shown in Table 4.1.

As our objective is to categorize assessments into positive and negative, we assign numerical values to the labels. To train the FNN model, which is utilized for the analysis of fuzzy attributes, we also enlist the expertise of manual evaluators to assess the objects. Specifically, for Dataset 1, we annotate 436 images as the training set and 1000 images as the test set. For Dataset 2, 200 images are marked as the training set, with 1000 images designated as the test set. In the case of Dataset 3, 186 images are categorized as the training set, and 1000 images are earmarked for the test set. On Dataset 1 and Dataset 3, the algorithm determines whether the image exhibits defects. On Dataset 2, the algorithm identifies whether the option is filled. Both of these object recognition tasks involve binary classifications. Consequently, in FNNs, the rule layer comprises two neurons to accommodate this binary classification.

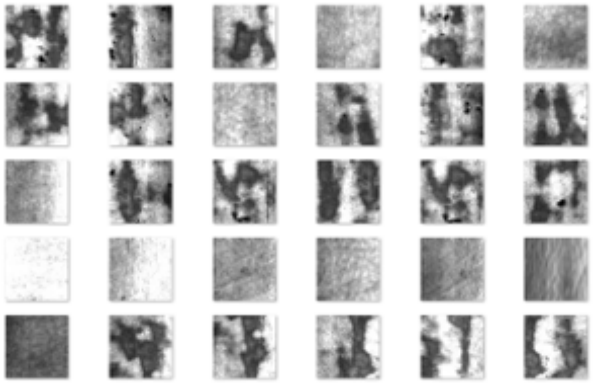
	Test set size		**Training set size**		**Resolution**
Databases	**Positive scale**	**Negative scale**	**Positive scale**	**Negative scale**	**ratio**
Dataset 1	850	150	296	140	64×64
Dataset 2	489	511	100	100	97×112
Dataset 3	835	165	126	60	64×64

TABLE 4.1. Datasets description.

4.2. **Experimental setup.** The features employed for characterizing the region of interest (ROI) include color, area, contour, texture, gray mean, contour detection, contrast detection, etc. Prior to feature extraction, it is essential to perform image preprocessing, which involves procedures like image graying and Gaussian filtering. Grayscale conversion transforms the image channels from 3



(a) Dataset 1. (b) Dataset 2.



(c) Dataset 3.

FIGURE 4.1. Some representative images from Dataset 1, Dataset 2, and Dataset 3.

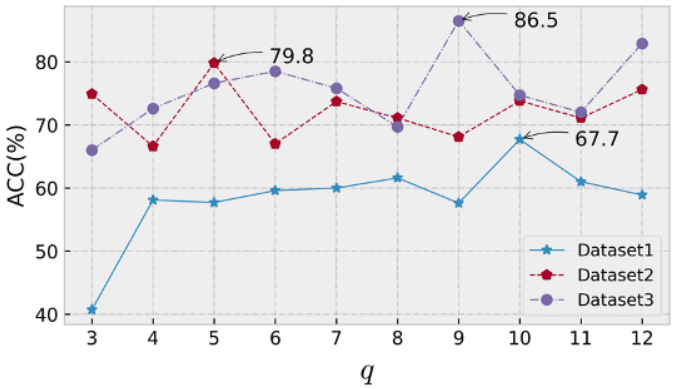


FIGURE 4.2. Accuracy of the FNNs.

to 1, simplifying subsequent image operations. Gaussian filtering serves to eliminate image noise, followed by binary conversion for facilitating feature extraction. Gray mean and area are pivotal image features in our experiments. Gray mean is extracted from grayscale images, while area is obtained from the binary image by counting the number of pixels with a value of 255 in the ROI.

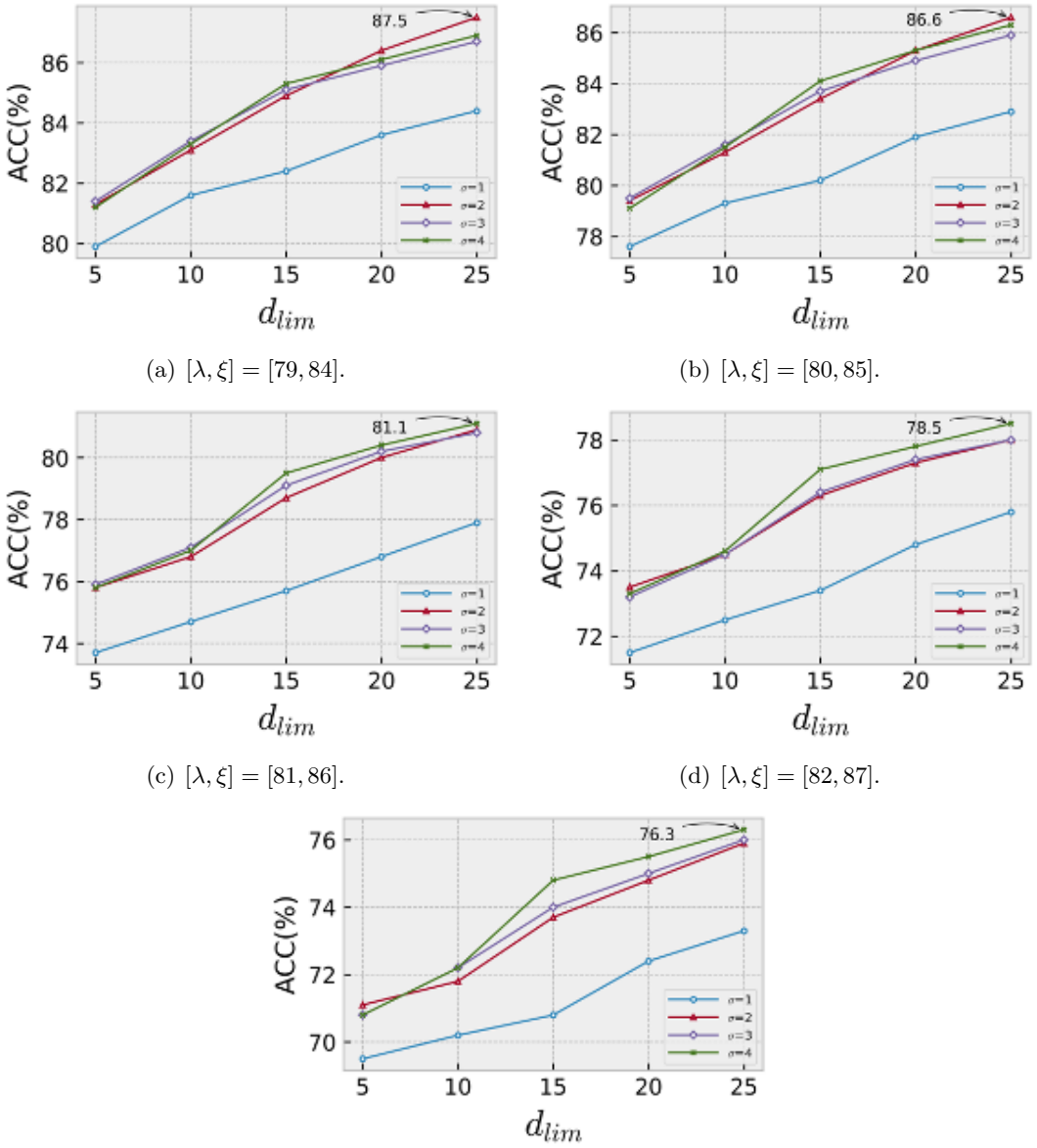
4.3. **CRFNN for the uncertainty of the output of FNNs.** The number of hidden layer neurons has a significant impact on FNN's performance. According to (2.5), in binary classification tasks, the number of hidden layer neurons in FNNs ranges from 3 to 12. To determine the optimal number of hidden layer neurons, different numbers of neurons are used to compare the performance of the neural network through the testing of target recognition accuracy, which is a key performance indicator for the algorithm. During this assessment, we record the results of object recognition as positive samples and negative samples. The number of images that are finally correctly classified as positive samples is recorded as true positives (TP), while the number of images that are correctly classified as negative samples is recorded as true negative samples (TN). Thus the accuracy index (expressed as a percentage %) is defined as

$$ACC(\%) = \frac{TP + TN}{Total} \times 100\%. \tag{4.1}$$

In Fig. 4.2, we present the relationship between the accuracy and the number of hidden layer neurons, where the horizontal axis is q and the vertical axis is the accuracy percentage. The test results show that in Dataset 1, the highest accuracy rate is 67.7%; in Dataset 2, the highest accuracy rate is 79.8%; and in Dataset 3, the highest accuracy rate is 86.5%.

To evaluate the impact of CRFNN on improving accuracy, we select the network that currently achieves the highest accuracy as the baseline for further analysis and comparison. Notably, comparing the highest accuracy achieved by this reference network with the accuracy attained by our proposed CRFNN method provides a more compelling demonstration of whether CRFNN can enhance performance accuracy.

In Fig. 4.3, Fig. 4.4, and Fig. 4.5, we show the ACC obtained when using the CRFNN algorithm for the three datasets. During the testing process of each dataset, we perform fine-tuning on the context interval, the threshold d_{lim}, and σ, with the goal of optimizing accuracy and obtaining the best possible results as well as investigating the impact of these parameters on accuracy. In each subfigure, the horizontal axis represents d_{lim}, and the vertical axis represents ACC. For each dataset, we choose five intervals and increase d_{lim} from 5 to 25 to study the trade-off between ACC and σ. It is evident from the results that the accuracy (ACC) of CRFNN significantly outperforms that of FNN. For Dataset 1, the best accuracy achieved by CRFNN is 87.5%, surpassing the best accuracy achieved by FNN (67.7%) by 19.8%. For Dataset 2, CRFNN achieves a best accuracy of 87.5%, exceeding the best accuracy achieved by FNN (79.8%) by 9.8%. For Dataset 3, CRFNN achieves the highest accuracy of 97.7%, which is 11.2% higher than the best accuracy achieved by FNN (86.5%).



(a) $[\lambda, \xi] = [79, 84]$.

(b) $[\lambda, \xi] = [80, 85]$.

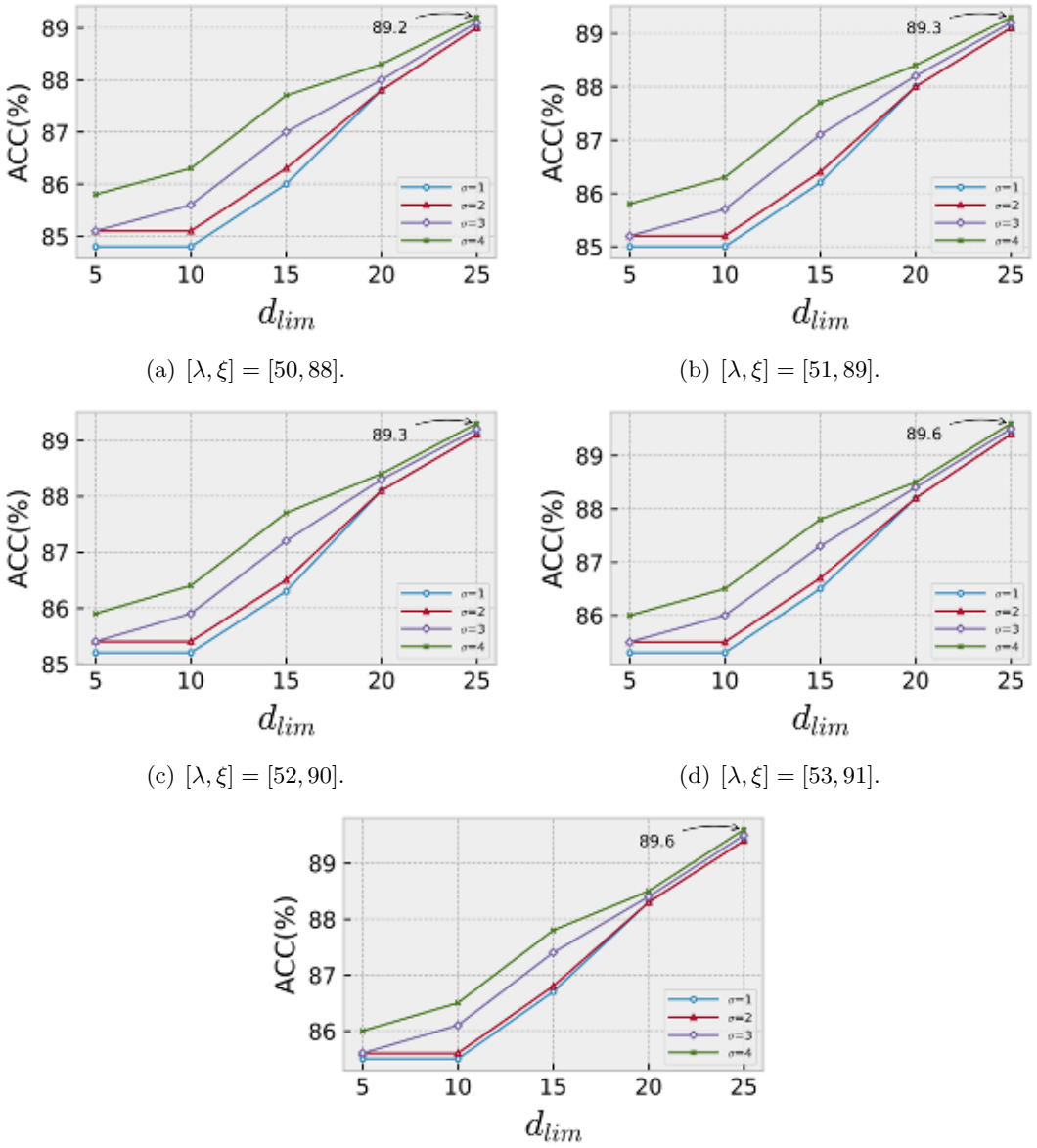
(c) $[\lambda, \xi] = [81, 86]$.

(d) $[\lambda, \xi] = [82, 87]$.

(e) $[\lambda, \xi] = [83, 88]$.

FIGURE 4.3. Parameters of the uncertain interval, d_{lim} and σ vs. ACC on Dataset 1.

Additionally, it is worth noting that Fig. 4.3 and Fig. 4.4 illustrate that the result distributions for Dataset 1 and Dataset 2 exhibit similar patterns, which is quite different with the observed trend in Dataset 3. To delve into the underlying reasons, we choose one characteristics, namely the "area", of these three datasets as an example. In Fig. 4.6, we present box plots from the training set, revealing that the Inter Quartile Range (IQR) of the three datasets is 1265.5, 3034.25, and 410, respectively. This indicates that Dataset 3 exhibits the smallest degree of variation, implying greater stability in the "area" characteristic compared to the other two datasets. This may be a



(a) $[\lambda, \xi] = [50, 88]$.

(b) $[\lambda, \xi] = [51, 89]$.

(c) $[\lambda, \xi] = [52, 90]$.

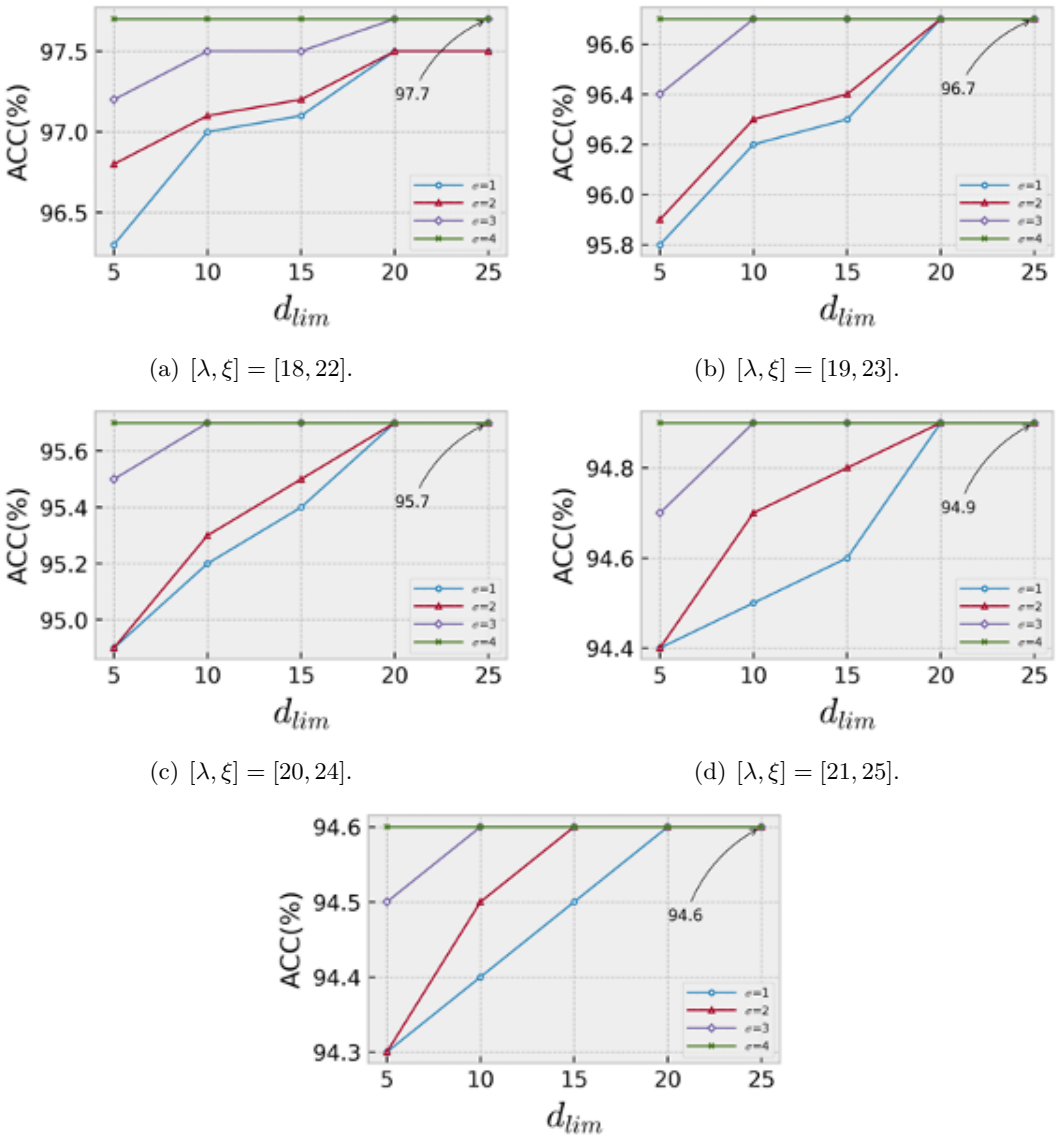
(d) $[\lambda, \xi] = [53, 91]$.

(e) $[\lambda, \xi] = [54, 92]$.

FIGURE 4.4. Parameters of the uncertain interval, d_{lim} and σ vs. ACC on Dataset 2.

factor in the smaller increase for Dataset 3 compared to Dataset 1 and Dataset 2, as the accuracy achieved with contextual reasoning for Dataset 3 is already excellent.

Further comprehensive comparison of these three datasets are given in Fig. 4.7, which displays the "area" characteristic values along with the number and frequency count of samples in the training set. The results reveal that the distributions of Dataset 1 and 2 are also exhibit similarities, while the distribution of Dataset 3 is very close to the Gaussian distribution. This finding suggests a potential link between the accuracy of recognition and the distribution of image features. A



(a) $[\lambda, \xi] = [18, 22]$.

(b) $[\lambda, \xi] = [19, 23]$.

(c) $[\lambda, \xi] = [20, 24]$.

(d) $[\lambda, \xi] = [21, 25]$.

(e) $[\lambda, \xi] = [22, 26]$.

FIGURE 4.5. Parameters of the uncertain interval, d_{lim} and σ vs. ACC on Dataset 3.

plausible hypothesis that arises is that when the feature distribution closely resembles a Gaussian distribution, it may lead to longer context lengths and faster convergence of accuracy to the maximum. To confirm this hypothesis, further data analysis and more tests are required, which will be implemented in our future research.

To assess the classification performance of the proposed CRFNN model, we employ the Receiver Operating Characteristic (ROC) curves, a widely recognized method given in [13]. The computations of ROC curves need the parameters of "True Positive Rate" (TPR) and "False Positive Rate"

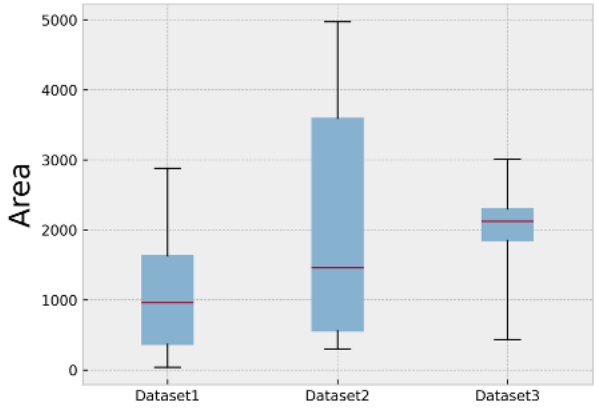
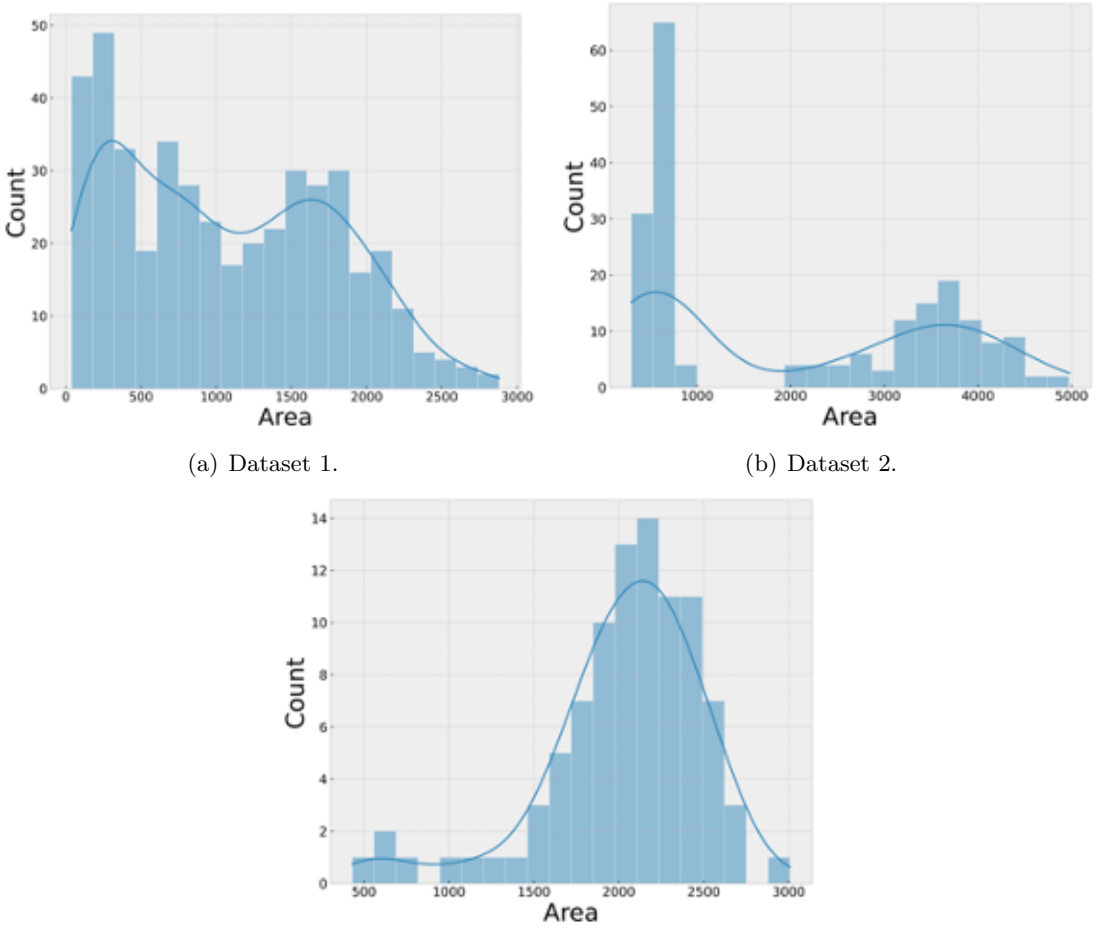


FIGURE 4.6. Box plots presenting the area distribution of the compared three datasets.

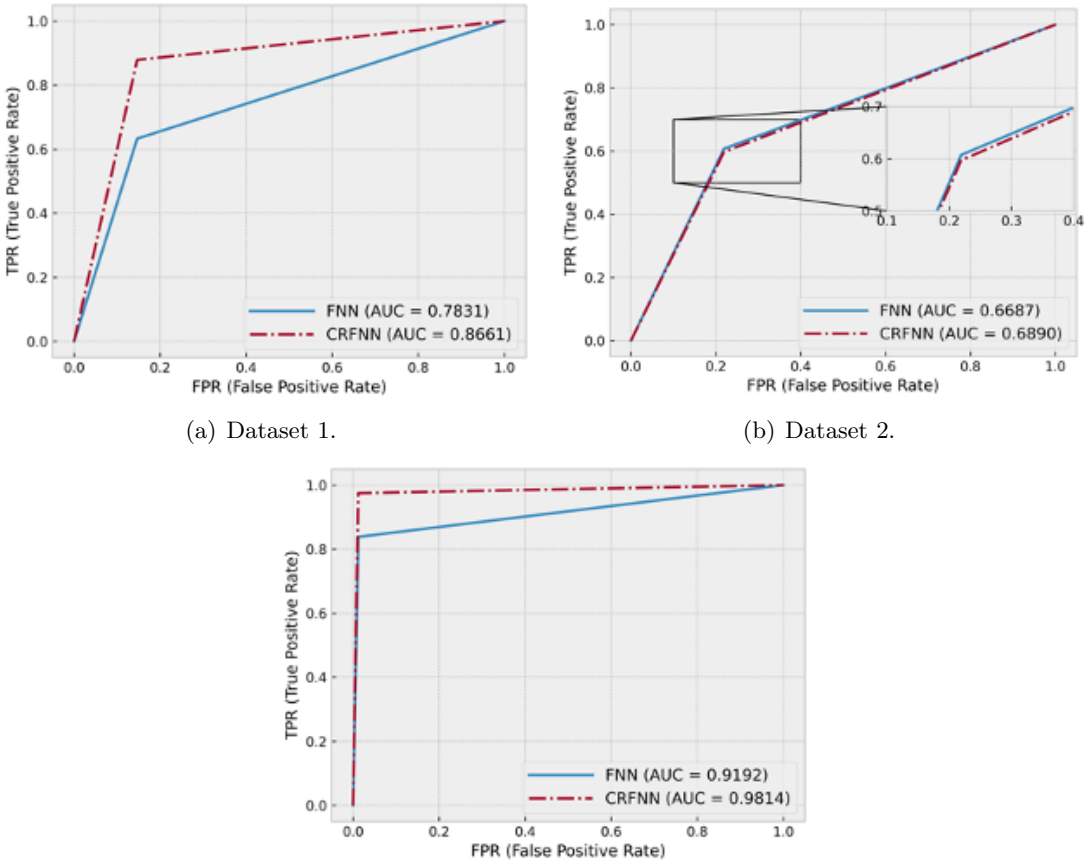


(a) Dataset 1.

(b) Dataset 2.

(c) Dataset 3.

FIGURE 4.7. Area distribution.



(a) Dataset 1.

(b) Dataset 2.

(c) Dataset 3.

FIGURE 4.8. ROC curve of the algorithm.

(FPR), which are calculated by

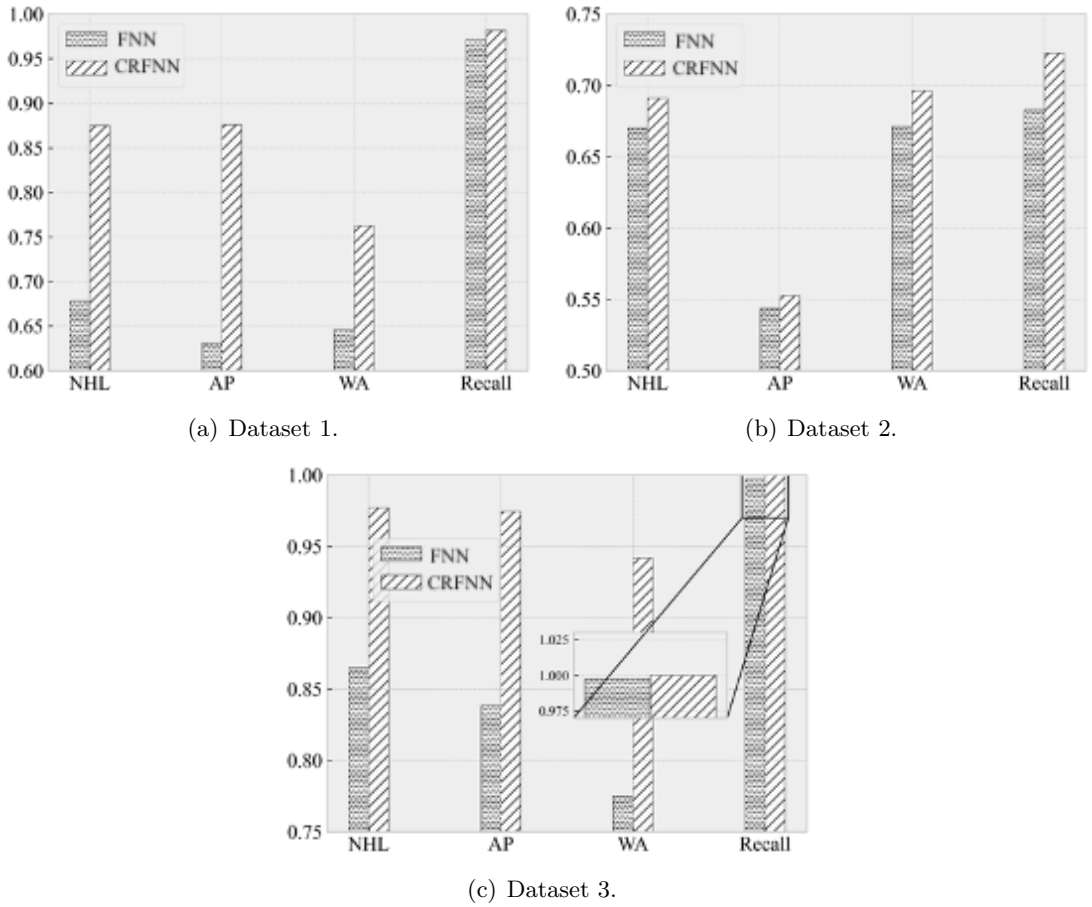
$$True\,Positive\,Rate = \frac{TP}{(TP+FN)}, \tag{4.2}$$

$$False\,Positive\,Rate = \frac{FP}{(FP+TN)}. \tag{4.3}$$

TPR represents the proportion of true positive results among all positive samples during the test, while FPR represents the proportion of false positive results among all negative samples. The ROC curve plotted in Fig. 4.8 displays the relations of these two values, where the horizontal axis represents the FPR, and the vertical axis represents the TPR. Based on the ROC curves, another parameter, denoted as AUC, is defined as

$$AUC = \int_0^1 ROC(t)dt, \tag{4.4}$$

which represents the 2D area under the ROC curve from $(0, 0)$ to $(1, 1)$. The AUC value ranges from 0 to 1, and the values closer to 1 indicate accurate predictions and the effectiveness of classification



(a) Dataset 1.

(b) Dataset 2.

(c) Dataset 3.

FIGURE 4.9. Model classification performance evaluation index.

algorithms. It is evident that the AUC values (shown in Fig. 4.8) for CRFNN surpass those of the FNN method, demonstrating superior performance. It is worth noting that for Dataset 2, the AUC values of the two ROC curves obtained by FNN and CRFNN are relatively similar. This can be attributed to the fact that the image features in Dataset 2 are more obvious, which are illustrated in Fig. 4.7(b) where many image features are concentrated within the range of $0 \sim 1000$. Nevertheless, it is crucial to emphasize that the CRFNN algorithm consistently outperforms the FNNs in terms of AUC values across all three datasets.

In addition to the ROC curve, we also utilize four additional evaluation indicators [23] to assess the algorithm. These indicators include Hamming loss (HL), average precision (AP), weighted accuracy (WA), and recall. For the indicator HL, we use its complementary measure, $1-$HL, referred as NHL. In all cases, higher values of these four indicators consistently indicate better model performance. The comparison between FNNs and CRFNNs using the above four indicators is shown in Fig. 4.9, where the horizontal axis represents the above four indicators (NHL, AP, WA,

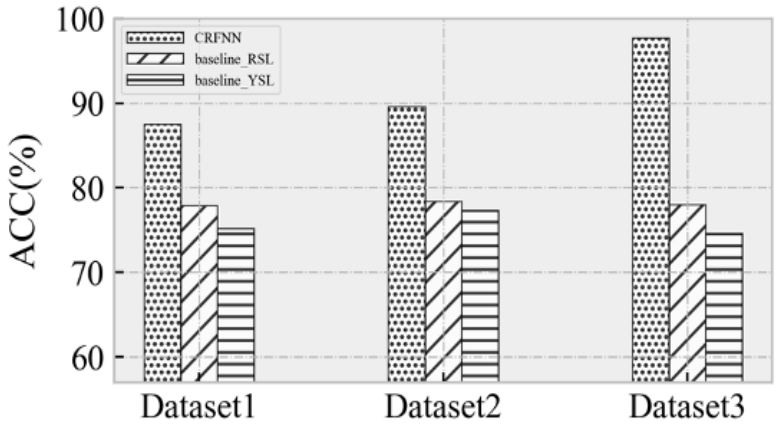


FIGURE 4.10. Comparison with the YSL and RSL algorithm.

Recall), ranging from 0 to 1. It is evident that CRFNN outperforms FNN in all indicators. Across the three datasets, there is an average improvement of 10.9% for NHL, 12.9% for AP, 10.3% for WA, and 10.8% for Recall.

Finally, we assess the accuracy performance of the proposed CRFNN model by comparing it to the two algorithms, YSL [38] and RSL [29], which are currently considered the most advanced, state-of-art methods for object recognition. We undertake this comparative analysis for the following two reasons. First, YSL stands as the most recent advancement for filling recognition, and it has exhibited notable superiority when contrasted with other existing algorithms. Second, YSL uses the weighted average gray level of the object for recognition, and RSL uses the area of the object for recognition, while our CRFNN also incorporates these features. As two kinds of representative algorithms, YSL and RSL can be aptly considered as representative benchmarks for our assessments. In Fig. 4.10, the ACC of three algorithms are compared. It is remarkable that the performance of RSL is better than that of YSL, which may be attributed to the significant influence of the area feature in these recognition tasks. Importantly, we emphasize that, across all examined datasets, the accuracy performance of CRFNN consistently and significantly outperforms the other two methods, which unequivocally demonstrates the superior effectiveness of the CRFNN model.

5. Conclusions

In this study, we design a novel neural network model, the context reasoning-based Fuzzy Neural Network, abbreviated as CRFNN. Built upon a fuzzy neural network framework, the CRFNN model facilitates the calculation of object similarity by combining feature distances, thus incorporating contextual information. It also integrates a local search strategy to enhance recognition accuracy, resulting in a substantial reduction of output uncertainty in the FNN. Detailed test results are

meticulously presented. The CRFNN model attains significantly improved accuracy, exceeding that of the traditional FNN by 19.8%, 9.8%, and 11.2% across the three datasets, respectively. Furthermore, a comprehensive set of evaluation metrics, including ACC, HL, AP, WA, recall, and AUC, is employed to assess the performance of the proposed algorithm. This evaluation underscores the superior performance of the CRFNN model compared to existing baseline algorithms, which confirms that the CRFNN stands as a dependable solution to the challenge of uncertain output within the FNNs. The future research directions include extending the proposed algorithm to more complex recognition scenarios, exploring potential integration with contextual reasoning and context prediction in intricate contexts, and evaluating its applicability in industrial and military domains, with a specific emphasis on real-time object recognition.

References

[1] Filling point database. *Retrieved from https://github.com/ruan-wang/Context-and-Fuzzy/dataset*, 2022.

[2] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi. A review of uncertainty quantification in deep learning: Techniques, applications, and challenges. *Information Fusion*, 76:243–297, 2021.

[3] D. Akmaykin and V. Grinyak. Fuzzy logic identifying of air targets with wwo-coordinate radars. *2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon).*

[4] G. Alpaydin. An adaptive deep neural network for detection, recognition of objects with long range auto surveillance. *Proceedings - 12th IEEE International Conference on Semantic Computing, ICSC 2018*, 92:316–317, 2018.

[5] X. Bai, Y. Zhang, H. Liu, and Y. Wang. Intuitionistic center-free FCM clustering for MR brain image segmentation. IEEE Journal of Biomedical and Health Informatics. *IEEE Journal of Biomedical and Health Informatics*, 23:2039–2051, 2022.

[6] B. Bhanu and T. L. Jones. Image understanding research for automatic target recognition. *IEEE Aerospace and Electronic Systems Magazine*, 8:15–23, 1993.

[7] J. J. Buckley and Y. Hayashi. Fuzzy neural networks: A survey. *Fuzzy Sets and Systems*, 66:1–13, 1994.

[8] R. Chandrasekharan and M. Sasikumar. Fuzzy Transform for Contrast Enhancement of Nonuniform Illumination Images. *IEEE Signal Processing Letters*, 25:813–817, 2018.

[9] C. J. Daniels and F. A. Gallagher. Unsupervised Segmentation of 5D Hyperpolarized Carbon-13 MRI Data Using a Fuzzy Markov Random Field Model. *IEEE Transactions on Medical Imaging*, 37:840–850, 2018.

[10] I. Despotovic, E. Vansteenkiste, and W. Philips. Spatially coherent fuzzy clustering for accurate and noise-robust image segmentation. *IEEE Signal Processing Letters*, 20:295–298, 2013.

[11] S. Dittmer, E. J. King, and P. Maass. Singular Values for ReLU Layers. *IEEE Transactions on Neural Networks and Learning Systems*, 31:3594–3605, 2020.

[12] C. Doersch, A. Gupta, and A. A. Efros. Unsupervised Visual Representation Learning by Context Prediction. *Proceedings of the IEEE international conference on computer vision*, pages 1422–1430, 2015.

[13] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27:861–874, 2006.

[14] L. M. G. Fonseca, L. M. Namikawa, and E. F. Castejon. Digital image processing in remote sensing. *Tutorials of SIBGRAPI 2009 - 22nd Brazilian Symposium on Computer Graphics and Image Processing*, pages 59–71, 2009.

[15] H. C. Fu and J. J. Shann. A fuzzy neural network for knowledge learning. *Int J. Neural Syst.*, 5:13–22, 1994.

[16] J. Gai and Y. Hu. Research on Fault Diagnosis Based on Singular Value Decomposition and Fuzzy Neural Network. *Shock and Vibration*, page 8218657, 2018.

[17] M. Gong, Z. Zhou, and J. Ma. Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering. *IEEE Transactions on Image Processing,*, 21:2141–2151, 2012.

[18] M. Kaushal, B. S. Khehra, and A. Sharma. Soft Computing based object detection and tracking approaches: State-of-the-Art survey. *Applied Soft Computing*, 70:423–464, 2018.

[19] B. Kosko. Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. *Prentice-Hall, Englewood Cliffs*, 1992.

[20] A. Kulkarni and N. Kulkarni. Fuzzy Neural Network for Pattern Classification. Procedia Computer Science. *Shock and Vibration*, 167:2606–2616, 2020.

[21] J. Lindblad and N. Sladoje. Linear time distances between fuzzy sets with applications to pattern matching and classification. *IEEE Transactions on Image Processing*, 23:126–136, 2014.

[22] Q. Lu, S. Lee, and L. Chen. Image-driven fuzzy-based system to construct as-is IFC BIM objects. Automation in Construction. *Automation in Construction*, 92:68–87, 2018.

[23] O. Maimon and L. Rokach. *Data Mining and Knowledge Discovery Handbook*. Series in Solid-State Sciences. Springer US, 2010.

[24] P. Maji and S. Mahapatra. Circular Clustering in Fuzzy Approximation Spaces for Color Normalization of Histological Images. *IEEE Transactions on Medical Imaging*, 39:1735–1745, 2020.

[25] T. C. Ng and S. K. Choy. Variational Fuzzy Superpixel Segmentation. *IEEE Transactions on Fuzzy Systems*, 30:14–26, 2022.

[26] T. V. Nguyen and D. Choi. Context Reasoning Using Contextual Graph. *IEEE 8th International Conference on Computer and Information Technology Workshops*, pages 488–493, 2008.

[27] S. Schulte, V. de Witte, M. Nachtegael, D. van der Weken, and E. E. Kerre. Fuzzy two-step filter for impulse noise reduction from color images. *IEEE Transactions on Image Processing*, 15:3567–3578, 2006.

[28] C. H. Seng, A. Bouzerdoum, M. G. Amin, and S. L. Phung. Probabilistic fuzzy image fusion approach for radar through wall sensing. *IEEE Transactions on Image Processing*, 22:4938–4951, 2013.

[29] L. R. Shao. Research on information extraction and recognition technology in online marking. *University of Electronic Science and Technology of China*, 2010.

[30] A. F. Sheta, A. Baareh, and M. Al-Batah. 3D object recognition using fuzzy mathematical modeling of 2D images. *Proceedings of 2012 International Conference on Multimedia Computing and Systems, ICMCS 2012*, pages 278–283, 2012.

[31] G. Song, K. Song, and Y. Yan. EDRNet: Encoder-Decoder Residual Network for Salient Object Detection of Strip Steel Surface Defects. *IEEE Transactions on Instrumentation and Measurement*, 69:9709–9719, 2020.

[32] Y. Song and H. Yan. Image Segmentation Techniques Overview. *AMS 2017 - Asia Modelling Symposium 2017 and 11th International Conference on Mathematical Modelling and Computer Simulation*, page 103–107, 2018.

[33] H. Tonekabonipour, A. Emam, M. Teshnelab, and M. A. Shoorehdeli. Comparison of neuro-fuzzy approaches with artificial neural networks for the detection of ischemia in ECG signals. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, page 4045–4048, 2010.

[34] R. J. Wai and Y. W. Lin. Adaptive moving-target tracking control of a vision-based mobile robot via a dynamic petri recurrent fuzzy neural network. *IEEE Transactions on Fuzzy Systems*, 21:688–701, 2013.

[35] L. Wang, Z. Chen, and G. Yang. An interval uncertainty analysis method for structural response bounds using feedforward neural network differentiation. *Applied Mathematical Modelling*, 82:449–468, 2020.

[36] X. Wang, F. Xie, W. Liu, S. Tang, and J. Yan. Robust small infrared target detection using multi-scale contrast fuzzy discriminant segmentation. *Expert Systems with Applications*, page 118813, 2023.

[37] Y. Wang and Y. Shang. Fuzzy clustering RBF neural network applied to signal processing of the imaging detection. *International Conference on Measuring Technology and Mechatronics Automation*, 2:321–324, 2010.

[38] S.-L. Yao, S.-R. Wang, G. Meng, and W. Zhen. A Novel Recognition Algorithm of Objective Questions for Exam Answer Sheets. *IEEE Transactions on Image Processing*, 32:135–145, 2019.

[39] F. Ye, W. Luo, M. Dong, D. Li, and W. Min. Content-Based Remote Sensing Image Retrieval Based on Fuzzy Rules and a Fuzzy Distance. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022.

[40] M. T. Yildirim, A. Başẗurk, and M. E. Yuksel. Impulse noise removal from digital images by a detail-preserving filter based on type-2 fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 16:920–928, 2008.

[41] L. A. Zadeh. Fuzzy Sets. *Information and Control*, 8:338–353, 1965.

[42] D. Zhang, M. Guo, J. Zhou, D. Kang, and J. Cao. Context reasoning using extended evidence theory in pervasive computing environments. *Future Generation Computer Systems*, 26:207–216, 2010.