

# **Automatically Recognizing and Grading Spangle on the Galvanized Steels**

## **Surface Based on Convolutional Neural Network**

Haineng Ma<sup>a</sup>, Dexiang Zong<sup>b</sup>, Yingna Wu<sup>a\*</sup>, Rui Yang<sup>a</sup>

<sup>a</sup> *Center for Adaptive System Engineering, ShanghaiTech University, Shanghai 201210, China*

<sup>b</sup> *Research Institute, Baoshan Iron & Steel Co., Ltd., Shanghai 200941, China*

---

\*Corresponding author.  
E-mail address: [wuyn@shanghaitech.edu.cn](mailto:wuyn@shanghaitech.edu.cn) (Y. Wu)

### **Abstract**

Currently, quality control of zinc barrier coatings on the galvanized steels, especially the homogeneity of spangle pattern, is becoming increasingly stringent due to the higher coating performance demanded from the end users. However, manual check by naked-eyes dominates the on-line inspection process in the steel factories, which is labor-consuming and reliant on experience. In this paper, the artificial intelligence method is applied to recognize spangles based on our proposed features which are different from the ones recognized by typical texture parsing descriptors. During spangle recognition process, combination of Multi-column Convolutional Neural Network (MCNN) and Gaussian Mixture Model (GMM) is proposed to obtain features, nuclei number and average nearest neighbor distance (ANND), for grading by support vector machine (SVM). The final grading results indicate that the accuracy reaches 91.3%, much higher than 41.3%-67.3% based on the typical texture descriptors, such as LBP, Laws, Harris, Laplace and SLIC. The effects of training data scale, network structure and hyperparameters of Gaussian kernel on the errors of counting and localization are discussed.

**Keywords:** *Spangle, Recognizing, Grading, Convolutional Neural Network, Gaussian Mixture Model*

## 1. Introduction

Hot dip galvanization, offering a zinc barrier coating (10-25um thick [1]) on the steels with unique formability, excellent corrosion resistance, low cost and recyclability, is a commercially mature process since 1830s [2]. Currently, the galvanized steel sheets have been extensively applied in industrial fields, such as automobiles, electrical home appliances or construction[3, 4]. For decorative purpose, the surface of zinc coating on some galvanized steel presents the typical snowflake-like or six-fold star pattern which is visible by naked eye. The patterns with characteristics of uneven, light and dark interlacing and dense dendrites are usually called spangles, which are formed when liquid zinc adhering to a steel surface is cooled to temperatures below the melting point of zinc, approximately 419°C[5-9].

With the expanding of application, quality control for galvanized sheet steel is increasingly stringent. Especially, considering the decorative appearance and performance, homogeneity of spangle size, is becoming one of the key criteria [10-14]. Although researchers have already carried out lots of efforts on stably controlling spangle size in production process, quality degrading led by inhomogeneous spangle size still happens in production line. Therefore, on-line inspection is necessary to real-timely check uniformity of spangle size and minimize cost of rework before delivering to customers. Currently, inspection for grading spangle size is still traditional method which relies on naked-eye of inspectors to count the number of spangle petals on the 100mm line segment. Mis-counting always happens when the inspectors are tired or distracted. Automatic inspection process is required to replace this labor-consuming and low efficient inspection process.

With the development of machine learning and deep learning algorithms, artificial computation has been widely applied in daily life[15] and industry[16-18] to address the laborious or time-consuming issues. For example, in civil engineering, Artificial Neural

Networks (ANN) model was reported to describe the steel beam-column connection properties[16] and predict the compressive strength, restored stress and strain of new concrete materials[17, 18] to circumvent great labor in the experimental works. Moreover, due to the characteristics of efficient, robust, accurate and non-subjective, texture pattern analysis based on artificial algorithms has been investigated since 1960s and widely applied on medical images, quality inspection, content-based image retrieval, satellite or aerial imagery, face analysis, biometrics, object recognition, et.al[19-22]. Especially, texture analysis can be easily integrated into solutions of inspection and characterization based on images of industrial products. It was reported that microstructure morphologies, such as the graphite particles in the ductile iron [23], NiCrBSi-WC metal [24], were characterized automatically by machining learning and deep learning methods. Besides the microstructure recognition and classification, texture analysis can also be applied on the automatic defect detection and localization [25-28].

Since spangle is a visible aesthetic feature of crystallites with the characteristics of feathery, ridged or snowflake, it can also be defined as a kind of texture with both random and regular texels. In 2006, Singh and Mishra [29] developed a machine vision system for spangle classification using typical statistical methods, which were Haralic gray level cooccurrence matrix (GLCM) and Law textural analysis. Three different patterns of spangles, which were flowery, feathery and dendritic spangle, could be automatically recognized and classified. Zhen et. al. [30] developed a feature extraction method to analyze particle size which cannot be measured by manual detection after the dual-ball punching test of galvanized steel sheet. Recently, Pangang Group Company Limited filed a patent [31], which claimed that the size and number of spangle were counted through extracting spangle contours using super-pixel segmentation algorithm.

Different from the former research works on spangle recognition and classification by typical algorithms, we mainly focus on the definition of spangle intrinsic features and recognition method using convolutional neural network to automatically grade spangles

among the images with different spangle density. Based on the understanding of spangle growth mechanism, the spangle morphology is analyzed to figure out the efficient features for spangle extraction and recognition. Multi-column Convolutional Neural Network (MCNN)[32] algorithm, which has good performance on crowd counting, is applied in this paper for spangle nuclei recognition and number counting. The recognition performance of MCNN combined with Gaussian Mixture Model (GMM) is investigated through comparing with the typical texture descriptors based on the final grading results by Support Vector Machine (SVM).

## 2. Methodology

Galvanizing process was executed in Baowu Steel Group Corporation. Before galvanizing, the strip was cleaned and activated to ensure the uniform adhesion of zinc alloy coating. When the strip entered galvanized solution through the dipping roller and maintained for a while, the mutual diffusion between iron and zinc elements occurred. At last, air knife with high pressure was applied to remove extra galvanized solution to control thickness and homogeneity of coatings. After cooling and drying, the galvanized sheet was inspected by quality inspector to identify the level through real-time observation. As the standard for level evaluating, galvanized sheet steel is usually graded from level 1 to level 5, with the decrease of spangle numbers within 100 mm scale. In this paper, the typical galvanized sheets of level 2, level 3 and level 4 were chosen for algorithms investigation of recognition and classification since these three levels' galvanized sheets cover most of products in industry.

To acquire spangle images with high resolution, the industrial area scan camera (MER-2000-19U3M, DAHENG IMAGING) was fixed above the galvanized sheet by UR5 robot which can precisely adjust the distance between the camera and galvanized sheet. The parallel line light sources were placed besides the galvanized sheet to abate the effect of uneven

illumination. Considering the possible glare caused by the mirror-like surface of the galvanized sheet, background was set as black. The macro-morphological information of spangles with the resolution of 600 PPI was acquired through the above settings. As shown in Fig.2, the spangle morphologies of three levels are similar, composed of dense triangles and sectors with different brightness and size. Obviously, the triangles and sectors of level 4 spangle are bigger than the others. The significant difference of spangle morphologies between level 2 and level 3 is uniformity of size of these triangles and sectors. It is observed that the homogeneity of size in level 3 is much better than level 2, since there are several fine triangles and sectors distributed randomly among the regular ones, as shown in Fig.2 (a) and (b). It is worth mentioned that misjudgment by inspectors often occurs when they count the number of spangles of level 2 sheets, since the fine triangles and sectors are difficultly counted in, which induces the similar number of spangles between level 2 and level 3. Moreover, the micro-morphology of spangles was observed through optical microscopy (DMi8, Leica) to better analyze the spangle structure and figure out its intrinsic features. The detail discussion is introduced in Section 3.1.

A dataset of 486 spangle images with the size of 350\*350 pixels, cropped from the originally captured images, was applied in this paper for training and testing different algorithms. Specifically, the numbers of images of level 2, level 3 and level 4 are 180, 180, 126 respectively. The effects of different dataset division on the performance of MCNN were investigated through establishing the relationship between counting error and training data scale, as mentioned in Section 4.1.1.

### 3. Spangle grading algorithms

#### 3.1 Nuclei number and distribution

There are several typical algorithms based on statistical and structural methods to extract and recognize the features of texture, such as local binary pattern (LBP) [33], Laws [34], Harris [35], Laplace [36], superpixel [37]. However, spangle grading is different from common texture recognition since grading need more precise information to distinguish different spangle density among the similar morphologies. The features extracted by these typical algorithms cannot exactly reflect the difference among the different levels' images. For example, over-segmentation and under-segmentation cannot be prevented when superpixel algorithm is applied, since it is difficult to separate spangles when the boundary between two spangles is blurred. Therefore, the features reflecting spangle density for grading should be figure out firstly before recognition.

Before we define the features, the detail structure of spangle should be understood firstly. As shown in Fig.3, optical micrograph indicates that the typical spangle is hexagonal dendritic structure with coexisted shiny and dull regions, which are separated by a straight dendritic arm boundary. Six straight dendritic arms, also called primary dendrite arms, linked together by a center nucleus, as shown in Fig.3 (a) and (b). There are also secondary dendrite arms that grow laterally away from the primary dendrite arms, when morphology is zoomed, as shown in Fig.3 (c).

Although a perfect zinc spangle should display the hexagonal dendritic morphology with 60 deg angles between the primary dendritic arms, the final shape of spangles is usually divergent, which could be explained by the solidification model proposed by Strutzenberger and Faderl in 1998[6]. Zinc solidification starts from heterogeneous nuclei at the steel/zinc interface in the initial stage. Solidification occurs by rapid sideways dendritic expansion of the nucleated grains along the steel/zinc interface until the grains impinge upon one another.

Final shape of the spangles is determined in the stage of rapid sideways expansion, which leads to formation of different shaped grains with thicker and thinner sectors. High rate of nucleation leads to the formation of small grains, while low rate of nucleation leads to the growth of large grains. Obviously, compared with the spangle shape, nuclei number should be a better indicator to reflect spangle density, since nuclei form in the initial stage and nuclei number don't change no matter how different the final spangle shape is.

Strutzenberger and Faderl [6] also pointed out that the grain size is correlated with secondary dendritic arm spacing, which confirmed that interfacial energy during solidification process is a determining parameter for grain size. However, it is difficult to extract a clear secondary dendritic arm spacing using algorithms, since it is often invisible in the spangle images captured by industrial area scan cameras. From another point of view, nuclei distribution, reflecting the density difference, is easily calculated after the nuclei are recognized. Sparse nuclei with long distance between two nuclei indicate large spangles and vice versa. Therefore, nuclei distribution was figured out as the second feature for recognition, except for nuclei number. In this paper, average nearest neighbor distance (ANND) for each image was applied to quantify the nuclei distribution. Based on the above discussion, nuclei number and nuclei distribution quantified by ANND, were defined as the intrinsic features.

Since it is difficult to extract the spangle nuclei using algorithms based on typical statistical and structural methods, a kind of neural network, Multi-Column Neural Network (MCNN) was tried in this paper. Before nuclei recognition, MCNN should be trained by manually annotated spangle images, which can also help evaluate the performance of recognition and localization algorithms. During annotating nuclei for each image, positioning the adjacent nuclei should be based on clearly distinguishing the boundary of the marked one. The manual annotation process continued until all typical spangle nuclei were figured out, as shown in Fig.4. After 486 images were annotated, the distribution of the two features for each image, calculated according to the manual annotation results, is shown in Fig.5. Obviously,

both nuclei number and ANND show the difference among three levels' images, although there is a little overlap between the adjacent levels. Here, we should mention that the annotation of ground truth is one-time process, although it needs labor time. As soon as we fix the model, spangle grading in the production line is an automatic process without manual intervention.

In order to assess the effect of proposed features, the grading results based on the common features extracted by the typical algorithms, such as LBP, Laws, Harris, Laplace and superpixel, were compared with that based on proposed two features in Section 5.

### 3.2 Algorithms framework

The whole framework of the spangle grading algorithm, consisting of recognition, localization and classification, is shown in Fig.6. Overall, the ground truth for different algorithms consists of two types: 1) density maps for Multi-column Convolutional Neural Network (MCNN); 2) nuclei number, ANND and level labels for SVM classifier.

In the stage of recognition, Multi-column Convolutional Neural Network (MCNN) was proposed to recognize spangle nuclei through training based on the learning of ground truth of density map (GT density map), generated by normalized Gaussian kernel in accordance with manually annotated positions of spangle nuclei. Subsequently, a series of 350 x 350 pixels images are encoded by MCNN into the estimated density maps (ET density maps) to obtain nuclei number by the global integral. More details are explained in Section 4.1. Since the average nearest neighbor distance (ANND), the other one of intrinsic features, relies on the precise coordinate of nuclei, localization process is followed by recognition. In the stage of localization, each ET density map is divided into several independent clusters based on the connected domain features after pre-processes of average pooling and percentage filtering for suppressing intrusive noise. Subsequently, the accurate coordinate of each nucleus is output

by Gaussian Mixture Model (GMM) based on the local integral value of cluster. The details of localization are introduced in Section 4.2. In the stage of classification, support vector machine (SVM) is applied as grading model, fitted from two-dimensional features, the number of spangle nuclei and ANND, and the matched labels of spangle level. The estimated number of spangle nuclei and ANND are sent into the trained classifier to obtain the final grading results. More details are introduced in Section 4.3.

### 3.3 Spangle nuclei recognition

Based on the former discussion, effective recognition algorithm should focus on the spangle nuclei and eliminate the impact of interfering and redundant features of background. Actually, the recognition algorithm for crowd counting could be leveraged for nuclei recognition, since there are similar recognition target and challenges, such as background distraction, annotation errors and density variation both for head and nuclei.

For crowd counting, earlier approaches focused on designing hand-crafted detectors based on the assumption that the crowd consists of independent individuals [38-41], or combines variable features from multiple sources [42], such as SIFT, GLCM, etc. However, the counting accuracy of those methods is generally low due to the perspective change or crowd occlusion. Recently, researchers try to train deep convolutional neural networks on public dataset of crowd count, such as Multi-Column Neural Network (MCNN) [32], CSRnet [43], CAN [44], to achieve higher accuracy. CSRnet focuses on the significance of receptive field of crowd counting through introducing dilated convolution to increase the receptive field. It also takes VGG-16 as the front-end to get more abundant features. CAN adopts adaptive encoding to take advantage of plentiful contextual information since the estimation of crowd counting should be a consistent process. Compared with CSRnet and CAN, MCNN introduces different convolutional filters to achieve different receptive fields with relatively

simple architecture. Considering the application scenario of online spangle grading in the galvanizing line, MCNN was applied to recognize spangle nuclei in this paper.

As shown in Fig.7, three columns with different convolutional filters are designed to extract multi-scale traits in the static images, since size of spangle varies with density change. After the input image is processed by three columns of convolutional (kernel size: 3x3, 5x5, 7x7, 9x9) and pooling (kernel size: 2x2) layers with different receptive fields, feature maps with different channels is obtained. Eventually, estimated density map is output by convolution based on the merged feature maps. The estimated nuclei number is obtained through calculating discrete integral of density map. Moreover, the estimated density maps can visualize the location information of spangles and provide guidance for galvanization process improvement to achieve the uniform spangle.

In this paper, mean absolute error (MAE) and root mean squared error (RMSE) of estimated spangle number are used to evaluate the absolute error of MCNN model, which are defined as follows:

(1)

Besides, mean absolute percentage error (MAPE) and normalized root mean squared error (NRMSE), defined in the following, are used to measure the relative error.

(2)

Moreover, R-squared ( $R^2$ ) is used to judge the model fitting status and its definition is in the following:

(3)

For the formula (1), (2) and (3),  $N$  is the number of test images,  $a_i$  is the actual number of spangle nuclei in  $i$ th image, and  $\hat{a}_i$  is the estimated number of spangle nuclei in the  $i$ th image.  $\bar{a}$  represents the mean of the actual number of spangle nuclei.

In the stage of MCNN model training, GT density map, the prior information to train neural network, was acquired through generating normalized Gaussian kernel around each

annotated spangle nuclei. Two different Gaussian kernel types, fixed and variable ones, were applied respectively to generate two different density maps. The main difference between them is the grey value distribution of kernel which is defined by  $\sigma$ . The variable one is geometry-adaptive Gaussian kernel which has variable  $\sigma$  and is designed to match the different spangle density, while the fixed one has the constant value of  $\sigma$ . For each annotated spangle nuclei, represents the coordinates of the th spangle nuclei. KD-tree, constructed by spangle nuclei position, is used to calculate the distance of nearest neighbors around as . Accordingly, the average distance is . Since the probability distribution of kernel for variable kernel type is related with the variance , simple proportional relationship between the average distance and the variance is established to create variable Gaussian kernel, which can be represented by . For the variable kernel type, and k were set as 0.15 and 4 respectively in this paper, while  $\sigma$  was set as 2 for fixed one. The values setting of and k were referred to hyperparameters of MCNN[32]. The determination of  $\sigma$  was referred to the average size of variable Gaussian kernel created by and k.

As the key hyperparameters, batch size was fine-tuned to 12 considering the stability, speed and GPU maximum capacity during model training. Each batch was constructed through randomly sampling 4 images from each level to avoid the impact of unbalanced data. Corresponding to 12 of batch size, learning rate was set as  $10^{-5}$ . During the process of MCNN model training, we randomly initialized the weights of network by granting parameters which comply with the normal distribution,  $N(0, 0.01)$ . In order to obtain low MAE and RMSE, the effects of Gaussian kernel types for GT density map generation and convolution kernels with different receptive fields for ET density map generation were mainly investigated in the following sections. Moreover, the suitable training data size was analyzed since the high similarity on morphology of different spangle images is quite different from the common public dataset for crowd counting.

### 3.4 Spangle nuclei localization

As we mentioned above, nuclei number and ANND were defined as two features for grading. Based on the ET density map, coordinates of the spangle nuclei should be figured out firstly before calculating distance through searching local maximum of grey value. In this paper, Gaussian Mixture Model (GMM) was applied to realize localization of spangle nuclei, since the weak intimation of nuclei position is Gaussian-like distribution generated via the supervision of normalized Gaussian kernel.

Usually, Gaussian model (GM) is a kind of supervised model, assuming the variables were subjected to Gaussian distribution and iteratively optimizing the parameters of mean and variance via expectation maximization (EM) algorithm [45]. GMM is an extended application of GM from one to multiple dimension. To enlarge the value of real peak and suppress disturbing noise, pre-processes, such as average pooling and percentage filtering, were applied before adopting GMM. Specifically,  $3 \times 3$  kernel size was used for average pooling and pixels whose gray values are lower than 12% of the maximum value was set as 0 during percentage filtering. Subsequently, pixels of estimated density maps acquire pre-grouped labels through 8-connected components analysis. For a specified connected component with three-dimensional vector which is defined as , respectively represent the coordinate of non-zero pixels and represents the gray value of the pixels. The number of predicted spangle nuclei is , the rounding of element-wise sum of . The probability distribution after Gaussian mixture fitting can be express as:

(4)

where , component vector of , denote the mean of respectively and represents the covariance matrix. in are the coordinate of spangle nuclei. To evaluate localization accuracy, precision and recall were used, which could be defined as following:

(5)

Where TP, FP and FN represent true positive, false positive and false negative of spangle nuclei. True positive represents the number of correctly estimated spangle nuclei. False positive represents the number of missing spangle nuclei which should be estimated. False negative represents the number of erroneously estimated spangle nuclei. Moreover, MAE and RMSE were also used to criticize the calculation of the estimated ANND and reveal the relationship between calculation of ANND and localization accuracy.

### 3.5 Spangle classification

Essentially, the final spangle grading is a process of classifying images. In this paper, Support Vector Machine (SVM) was selected to realize spangle grading since SVM is suitable for small datasets with low dimension and few support vectors could perform prediction efficiently.

Support Vector Machine (SVM), a kind of classic binary classifier based on the concept of structural risk minimization through statistical learning strategy, is widely used in variable scenarios, such as face recognition[46] and images classification[47]. As the non-linearly separable binary classification, SVM learning algorithms focus on seeking the optimal separating hyperplane with respect to training data via kernel method to transform the original input space into higher dimensional feature space. The following optimization problem is vital for SVM:

(6)

where and respectively represent the slack variables of the misclassifying error and penalty factor for involving additional control. represents input vector () with labels (). The hyperplane could be expressed by , where denotes a normal vector. Here, training vectors is projected to the high-dimensional space through the kernel function .

The 2-dimensional input vector of SVM is number and ANND of spangle nuclei. The penalty factor was set as 1 and radial basis function (RBF) was used as kernel function. Since

spangle grading in this paper is a task of multi-classification, one-versus-one scheme was selected to train 3 SVM models and the final prediction result was synthesized by the results from those 3 models. Evaluation protocol of classification is the prediction accuracy, which is defined as the percentage of correct predictions for each level.

## 4. Results & discussion

### 4.1 Spangle recognition using MCNN

MCNN was applied in this paper to recognize spangle through generating density maps containing approximate location information of spangle nuclei. Evaluation protocols of MCNN models were described by MAE, RMSE, MAPE, NRMSE and  $R^2$  when the hyperparameters were adjusted. Since the morphological difference among the spangle images of each level is unobvious, training data with suitable scale is advantageous to effectively extract pivotal features of spangle and avoid overfitting. The detailed discussion for the effect of training data size on protocols was explained in Section 4.1.1. As mentioned in 3.3, MCNN model providing three columns of convolutional and pooling layers with different receptive fields and the final feature maps were obtained through merging different channels. The effects of each column and the combination of multi columns on model performance were investigated in Section 4.1.2 after the training dataset scale being fixed. Moreover, the effect of hyperparameters of Gaussian kernel applied to generate density maps of ground truth on the model performance was investigated in Section 4.1.3.

#### 4.1.1 Training dataset scale

The total amount of images used in this paper is 486 (level 2: 180, level 3: 180, level 4: 126). In detail, 186 images (level 2: 80, level 3: 80, level 4: 26) were selected as test dataset to

evaluate the model performance. Training datasets in different scale was sampled randomly from the remaining 300 images (level 2: 100, level 3: 100, level 4: 100). The number of images for each level is exactly same in training dataset to keep the dataset balanced. Based on these criteria, MCNN models were trained via seven training datasets respectively with different scales of 12, 36, 60, 108, 180, 240, 300 to figure out the suitable training dataset scale. For each training dataset scale, MCNN models were trained three times with the different sampled training dataset, except for 300. The results of protocols showed in this section are the average values. Gaussian kernel for ground truth generation is fixed with in this Section.

The counting error, MAE and RMSE, and  $R^2$  on test datasets are shown in Fig.8 with the increase of training dataset scale. The results indicate that larger scales of training data lead to the decrease of counting error and higher  $R^2$ . Both MAE and RMSE values keep higher when MCNN models were trained by data with the size of 12, 36 and 60 and decrease rapidly when training size changes from 60 to 108. MAE and RMSE change slowly from 4.4 to 3.8 and 5.8 to 5.0 respectively when data size increases from 108 to 300, which means counting error could only decrease less than 1 when the training data scale is magnified almost three times. Similar as counting error, the significant increase of  $R^2$  occurs when training data increases from 36 to 108.

Moreover, the change trends of counting error for images in different spangle levels are different with the increase of training data size, as shown in Fig.9. The trends of MAE and RMSE for level 2 images are similar as that for whole images, while the counting error keeps relatively stable and the change is less than 1 for the other two levels. Obviously, the increase of training data size only impacts the counting accuracy of level 2 images since the denser spangles consist of more complex information for level 2 images compared with level 3 and 4 images.

Furthermore, the progression of MAE and RMSE on both training and test datasets is compared in Fig.10. Different from the results on test dataset, the counting error on training

data shows continuous decrease with the increase of training data size, especially when data size is larger than 108. It indicates that overfitting occurred when MCNN models were trained by larger scale of training data, especially when the training data size is 300.

Considering the training efficiency and the impact of error on the nuclei number counting, training data size is ultimately fixed as 108 for further exploring the effect of network structure and ground truth generation on recognition accuracy, the performance of spangle localization and grading.

In the scenario of future deployment of these models trained by 108 data scale in industrial production line, long-time consumption on manual annotation and overfitting should be avoided. The proposed method is to generate additional ground truth of fresh data by above models, called as supervisor models, instead of manual annotation. Subsequently, independent training based on fresh data could circumvent overfitting. Furthermore, more expected model should be trained by dataset with the adjusted ratio between fresh and old data, due to the error between ground truth generated by the supervisor and the annotators.

#### 4.1.2 Column structure

In this paper, MCNN with different convolutional filters was designed to provide the different receptive fields and extract multi-scale traits with the spangle size varying. Therefore, the effect of different receptive fields on counting error was investigated through evaluating the counting error of model with one-column convolutional neural network (OCNN) and multi-column convolutional neural network (MCNN). Gaussian kernel for ground truth generation was also fixed with in this Section.

As shown in Table 1, OCNN-large, OCNN-medium and OCNN-small corresponds to the one-column models with the different convolutional kernel size respectively. For OCNN model, the absolute and relative errors increase with the decrease of convolutional kernel size,

which indicates that the large receptive field carried by the large convolutional kernel facilitates acquiring better counting accuracy. After three columns of convolutional filter are combined, the errors are slightly decreased compared with MAE and RMSE of OCNN-large model. Moreover, MCNN and OCNN-large have the relatively high  $R^2$ , which indicates the better model fitting status.

In order to figure out the root cause of the above phenomena, the comparison among those estimated density maps by MCNN, OCNN and the ground truth is shown in Fig.11. The discrete integral of the region selected by the red box in the ground truth is 0, while the matched values in estimated density maps of MCNN, OCNN-large, OCNN-medium and OCNN-small are 0.07, 0.08, 0.17 and 0.43 respectively. Obviously, inappropriate receptive fields induce attention of network on the essential features to be diluted by placing too much information on insignificant area in the process of spangle counting when the convolutional filter with small kernel size was applied. The combination of convolutional filter with different kernel size could eliminate these unnecessary attentions.

Moreover, the results of Table 2 also show that the counting error for level 2 images is kept higher compared with the results for the other level images, which indicates that performance of MCNN is getting worse with the increase of spangle nuclei density. As shown in Fig.12, the estimated density maps for three levels images are compared with ground truth generated by fixed Gaussian kernel. It is observed that insignificant regions could be easily figured out in the estimated density map of level 2 images since the relatively more plentiful information in fine spangle images hinders the recognition of real spangle nuclei.

#### 4.1.3 Gaussian kernel

Since GT density map is critical for MCNN model performance, Gaussian kernels with different type were investigated. As we mentioned before, fixed Gaussian kernel means that

is constant. Variable Gaussian kernel means that  $\sigma$  is corresponded with the change of average of the distances of nearest spangle nuclei. For each type of Gaussian kernel,  $\sigma$ ,  $\mu$  and  $k$ , representing the standard deviation of the Gaussian distribution, number of the nearest neighbor nuclei and proportionality factor respectively, were adjusted to understand the effect of kernel size on the model performance. MAE and RMSE of models with different Gaussian kernel are shown in Fig.13. The results show that there is no significant difference for MAE and RMSE no matter the kernel type or kernel size changed. The change ranges of MAE and RMSE are only 0.4 and 0.6 respectively. The counting error is relatively larger when Gaussian kernel is fixed with size of  $3 \times 3$ . Fig.14 shows the comparison results between ground truth and estimated density maps for level 2 images. The ground truth generated by variable Gaussian kernel shows the different grey values of spangle nuclei clusters, which are corresponded to the different nuclei density. The brighter the cluster, the denser the spangles. Fig.14 (a) and (b) provides two scenarios that the difference between ground truth and estimated density map is significant and limited respectively. Obviously, the estimated density maps trained based on the different ground truths shows similar morphologies and counting number for these two scenarios, which could explain the similar performance of models with Gaussian kernel type changed.

Since MCNN model also affects the results of the following localization and classification besides spangle recognition, the effect of Gaussian kernel on the performance of GMM and SVM which were used as localization and classification were consecutively investigated.

## 4.2 Spangle localization

As shown in Fig.6, the final grading results from the classifier depends on two features, the nuclei number and ANND. In Section 4.1, the nuclei were recognized by MCNN model and the number was calculated by discrete integration of the gray values in the estimated density maps. In this section, ANND calculated based on the estimated nuclei positions is discussed.

Since the estimated density maps output by MCNN model did not provide the precise nuclei coordinates, Gaussian Mixture Model (GMM) was used to extract the location of the nuclei. In order to more precisely position the center of kernel, pre-processing estimated density map is necessary before GMM was adopted. In this paper, average pooling and percentage filtering were applied to enlarge the value of real peak and suppress disturbing noise erroneously detected by MCNN model. Since the background noise or artifact erroneously detected by the algorithm does have impact on the accuracy of ANND, systematical investigation for pre-processing should be developed later.

MAE and RMSE of ANND results are shown in Fig.15 (a). It is observed that both MAE and RMSE increase with the increase of Gaussian kernel size, except for those at and . The results indicate that the key factor on the calculation accuracy of ANND is the size of Gaussian kernel. For the fixed Gaussian kernel, MAE and RMSE of ANND change from 5.6 and 7.5 at to 12.9 and 15.9 at . For the variable Gaussian kernel, MAE and RMSE of ANND change from 5.6 and 7.4 at to 9.5 and 11.4 at . Overall, the maximum change of MAE and RMSE with the variation of kernel size is more than 1.5 times. Compared with this dramatic change, the change of MAE and RMSE of counting is relatively limited, as shown in Fig.13. Obviously, the size of Gaussian kernel has greater impact on the calculation accuracy of ANND compared with the counting accuracy of the spangle.

1-1 matching scheme[48], through manually setting tolerance radius, was used to divide the estimated and ground truth nuclei into true positive (TP), false positive (FP) and false negative (FN). to calculate precision and recall, representing the proportion of the correctly estimated ones to all estimated nuclei and ground truth nuclei respectively. Precision and recall could reflect the accuracy of nuclei localization. Fig.15 (b) illustrates the trend of precision and recall with the change of kernel size and type. Similarly, regardless of the kernel type, the precision and recall decrease as the kernel size increase. For the fixed Gaussian kernel, the precision and recall change from 70.2% and 86.4% at  $\sigma = 1$  to 38.8% and 33.3% at  $\sigma = 3$ . For the variable Gaussian kernel, the precision and recall change from 69.4% and 86.4% at  $\sigma = 1$  to 44.4% and 37.3% at  $\sigma = 3$ . The results indicate that MAE and RMSE of ANND increase corresponded to the decrease of precision and recall.

The correlation of the precision and recall with MAE and RMSE of ANND is further explained in Fig.16, which shows the estimated results of images with different spangle density. As shown in Fig.16 (a), the spangle images with manually annotated nuclei positions represents in blue. Fig.16 (b) to (e) show the estimated nuclei positions represent in red, which are fitted from the estimated density maps output by different MCNN models trained in accordance with different Gaussian kernels. It is indicated that the accuracy of nuclei localization has great impact on ET ANND. For the top spangle image, the smallest absolute difference between ET ANND and GT ANND is 3.93 at  $\sigma = 1$ , which also corresponds to the highest ones of precision (0.86) and recall (0.78). With the increase of  $\sigma$ , the absolute difference between ET ANND and GT ANND quickly increase while precision and recall correspondingly decrease. Similar phenomena could be observed for the middle and bottom images.

Obviously, with the increase of Gaussian kernel size, the nuclei clusters in the estimated density maps are getting larger, which should have significant effect on the accuracy of nuclei localization. More precise ANND relies on MCNN model trained by smaller Gaussian kernel.

Different from the performance of MCNN on localization, the accuracy of counting keeps almost stable with the increase of Gaussian kernel size, as shown in Fig.13, since nuclei number is calculated by discrete integration of the gray value in the estimated density maps which provides similar value for the larger dark and smaller bright nuclei clusters.

Since the changes of MAE and RMSE of counting are relatively limited compared with those of ANND, the protocols of ANND and localization are mainly considered to select suitable model before grading.

### 4.3 Grading

In this paper, MCNN and GMM were used to extract two key features in spangle images and provide input for SVM. In order to have a comparison, several typical texture parsing descriptors based on the statistical and structural methods, such as Local Binary Pattern (LBP) [33], Laws [34], Harris [35], Laplace[36], superpixel [37], were also applied to extract the features for SVM grading. LBP is a simple but effective texture operator, originally proposed by Ojala et al. in 1996. It has been widely used in face description, eye localization, fingerprint recognition and so on, due to its low computation complexity and extraordinary adaptability to illumination variations. In detail, new label is calculated based on binary code encoded from comparing the candidate pixel to each of its 8 neighbors. Laws' mask is also defined as energy mask, contributed by K.I. Laws in 1979, which is widely applied in texture classification and medical images analysis. Essentially, Laws' mask is a combination of different pre-defined convolution kernel, which focuses on typical texture features including level, edge, spot, ripple, and wave. Harris is a detector that concentrates on corner detection for texture images, introduced by Chris Harris and Mike Stephens in 1988. This detector is extensively used in object recognition, motion detection and so on. For Harris detector, the corner can be interpreted as the intersection of two edges, while the edge could be translated into the region possessing significant gradient of gray value. Laplace operator is an edge

detector, named after the French mathematician Pierre-Simon de Laplace. Besides the computer vision, Laplace operator is widely used in celestial mechanics, quantum mechanics and so on. A superpixel can be defined as a group of pixels sharing common characteristics. Simple Linear Iterative Clustering (SLIC), introduced by Achanta et al in 2012, was used in this paper to extract superpixel. Similar to the preprocessing before estimating density maps in our proposed algorithm framework, global percentage filtering was also applied to obtain feature maps without insignificant information via the typical textures descriptors. The processed features were used as the input of SVM classifier.

The final grading results are listed in Table 3 and the images with pixels of interest by different features extraction algorithms are summarized in Fig.17. Obviously, our proposed method achieves much better grading results for all different levels. The grading accuracy based on the features extracting method, combination of MCNN and GMM, achieves 85.4%-91.3%, especially 91.3% at  $k=4$  and  $\beta=0.15$ . The accuracy based on the typical extracting algorithms is relatively lower, in the range of 41.3%-67.3%.

As shown in Fig.17 (b), LBP pays more attention on the local small-scale information of spangle which should be ignored, and is unable to extract the global features with different levels. Laws is more interested in the second dendrites in the large-scale spangles which is usually invisible in the small-scale spangles, as shown in Fig.17 (c). Therefore, the grading results of LBP and Laws on level 4 images are relatively higher compared with the results on level 2 and 3 images, since pixels of interest extracted by LBP and Laws are quite different for large-scale spangles from those of small-scale spangles.

In Fig.17 (d) and (e), the results show that Harris and Laplace focus on the information of corners and edges respectively. The contrast of their grading accuracy in Table 3 demonstrates that the corner features are more helpful than the edge features in the spangle images grading. Meanwhile, it is noticed that Harris achieves relatively higher grading accuracy on both level 2 and level 4 images, which are 87.5% and 84.6%, while only 41.7% accuracy for level 3

images. The reason is related to the quite difference of corner feature of spangles between level 2 and level 4 images, while limited diversity exists in level 3 spangles compared with the other two levels.

As shown in Fig.17 (f), SLIC selectively ignores smaller-scale spangles during the process of recursively merging similar pixels, which could be the key features to distinguish the spangles among level 2, 3 and 4. Therefore, the grading accuracy results of SLIC, shown in Table 3, are quite low (32.1%-56.2%) for all the three levels images.

Based on the above discussion, the similar morphology and high randomness in distribution for spangles should be the root causes for that typical texture descriptors could not capture the intrinsic features of spangle.

As shown in Fig.17 (g), our proposed method focusing on the central locations of each spangle, called as the spangle nuclei, does capture the intrinsic difference among three levels spangle, which demonstrates higher grading accuracy in Table 3. Due to the ability of graph neural network to extract nuclei and generalize nuclei distribution, our proposed method has strong adaptability for the similar spangle morphology.

## 5. Conclusions

In this paper, a novel spangle grading framework was investigated to provide an industrial solution for on-site spangle quality inspection. Different from the typical statistical and structural methods, two features characterizing the difference of spangles with different levels were proposed for recognition based on the understanding of spangle growth mechanism, which are the number and average nearest neighbor distance (ANND) of spangle nuclei. Since nuclei number corresponds to the spangle number in the image, while ANND related to the spangle density in the image, the bigger number and smaller ANND should correspond to the lower level of spangles.

Multi-column Convolutional Neural Network (MCNN), a network for crowd counting, was leveraged for nuclei recognition followed by nuclei localization via Gaussian Mixture Model (GMM). After the training data scale was investigated to extract robust features and avoid overfitting simultaneously, 108 was ultimately selected. Subsequently, the estimated nuclei number and ANND acquired by our proposed combination of MCNN and GMM, features recognized by LBP, Laws, Harris, Laplace and SLIC were all input to support vector machine (SVM) for spangle grading respectively. The main conclusions are shown in the following:

1. Compared with the grading accuracy (41.3%-67.3%) based on the input of typical texture descriptors, it achieved 91.3% based on that of our proposed combination of MCNN and GMM method since the intrinsic nuclei were precisely recognized and localized. Features recognized by the typical descriptors, such as local small-scale information by LBP, second dendrites by Laws, corners and edges by Harris and Laplace, merged pixel clusters by SLIC, did not reflect the difference of spangle morphologies with different levels.
2. During the recognition and localization processes by MCNN and GMM, it was observed that the estimated density maps generated by MCNN was critical to control error of estimated number and ANND, which was strongly related to network structure and hyperparameters of Gaussian kernel.
3. In terms of model architecture, one column structure with small convolutional kernel size provided small receptive fields which induced attention on non-spangle area and led to imprecise estimated density maps. Multi-column structure with different convolutional kernel size could eliminate the attention on non-spangle area.

4. Compared with the Gaussian kernel type, the size change of Gaussian kernel had more significant effect on GT density maps which induced the obvious change of estimated ANND. Gaussian kernel size had obvious effect on searching accurate coordinate of spangle nuclei, since the density maps, encoded by MCNN based on the smaller Gaussian kernel, promoted acquiring precise nuclei position and ANND. The limited change of nuclei number with the variation of kernel size was attributed to the calculation method, the integral of estimated density map.
5. Different from the scenario of crowd counting, the variable kernel did not play the significant role to improve precision of spangle nuclei recognition and localization since the change of spangle density among different levels is inapparent compared with the crowd density.

Moreover, some improvement methods should be investigated in the near future to achieve better grading results. Firstly, based on the algorithm of MCNN, counting number is calculated through discrete integral of density map, while the calculation of ANND strongly relied on the coordinates of the spangle nuclei. Therefore, the background noise or artifact erroneously detected by the algorithm does have more significant impact on the accuracy of ANND compared with counting number. Although some pre-processes were preliminarily tried to eliminate noise in this paper, systematic analysis on the effect of different pre-processes on the accuracy of ANND should be investigated later. Moreover, recognition and localization are realized by MCNN and GMM subsequently, which makes the algorithms complex and impacts the calculation speed. MCNN model should be optimized to accomplish both recognition and localization simultaneously with high accuracy later.



## References

- [1] T. Takaki, K. Wada, K. Ishii, H. Yasuda, Two-dimensional phase-field study for spangle texture formation in hot-dip galvanizing, *Computational Materials Science*, 187 (2021) 110077.<http://dx.doi.org/10.1016/j.commatsci.2020.110077>
- [2] S.M.A. Shibli, B.N. Meena, R. Remya, A review on recent approaches in the field of hot dip zinc galvanizing process, *Surface and Coatings Technology*, 262 (2015) 210–215.<http://dx.doi.org/10.1016/j.surfcoat.2014.12.054>
- [3] M.S. Yuan, Properties and Application for Galvalume Coating Steel Sheet, *WORLD IRON & STEEL*, (3): 70–72 (2004)
- [4] T. Guo, Q. Xu, C. Ran, X. Dong, Influence of Process Conditions on Spangle Size of 55%Al-Zn Alloy Coated Steel Sheet, *Materials Science Forum*, 889 (2017) 127–132.<http://dx.doi.org/10.4028/www.scientific.net/MSF.889.127>
- [5] F.A. Fasoyinu, The solidification of hot dipped galvanized coatings on steel, University of British Columbia, Vancouver, 1989.
- [6] J. Strutzenberger, J. Faderl, Solidification and spangle formation of hot-dip-galvanized zinc coatings, *Metallurgical and Materials Transactions A*, 29 (1998) 631–646.<http://dx.doi.org/10.1007/s11661-998-0144-8>
- [7] A.R. Marder, The metallurgy of zinc-coated steel, *Progress in Materials Science*, 45 (2000) 191–271.[http://dx.doi.org/10.1016/S0079-6425\(98\)00006-1](http://dx.doi.org/10.1016/S0079-6425(98)00006-1)
- [8] A. Sémorez, Y. Durandet, M. Rappaz, EBSD characterization of dendrite growth directions, texture and misorientations in hot-dipped Al–Zn–Si coatings, *Acta Materialia*, 49 (2001) 529–541.[http://dx.doi.org/10.1016/S1359-6454\(00\)00322-0](http://dx.doi.org/10.1016/S1359-6454(00)00322-0)
- [9] L. Zhai, P. P, Y. Liu, Y. Lei, S. Deng, X.P. Su, The influence of morphology and crystal orientation of spangles on hot-dip Zn–0.5Sn alloy coating, *Journal of Mining and Metallurgy, Section B: Metallurgy*, 57 (2021) 4–4.<http://dx.doi.org/10.2298/JMMB200726004Z>
- [10] S.J. Park, H.S. Koh, J.G. Kim, M.-b. Moon, Production of Heavy Gauged HY-GALUME in Hyundai HYSCO, *Corrosion science and technology*, 10 (2011) 157–161.<http://dx.doi.org/10.14773/cst.2011.10.5.157>
- [11] S.K. Shukla, B.B. Saha, B.D. Triathi, R. Avtar, Effect of Process Parameters on the Structure and Properties of Galvanized Sheets, *Journal of Materials Engineering and Performance*, 19 (2010) 650–655.<http://dx.doi.org/10.1007/s11665-009-9527-2>
- [12] P.R. Seré, J.D. Culcasi, C.I. Elsner, A.R. Di Sarli, Relationship between texture and corrosion resistance in hot-dip galvanized steel sheets, *Surface and Coatings Technology*, 122 (1999) 143–149.[http://dx.doi.org/10.1016/S0257-8972\(99\)00325-4](http://dx.doi.org/10.1016/S0257-8972(99)00325-4)
- [13] H. Asgari, M.R. Toroghinejad, M.A. Golozar, On texture, corrosion resistance and morphology of hot-dip galvanized zinc coatings, *Applied Surface Science*, 253 (2007) 6769–6777.<http://dx.doi.org/10.1016/j.apsusc.2007.01.093>
- [14] Y. Ren, Study of the spangle size of 55% Al– Zn– Si steel, (2014)
- [15] O. Akinfala, F. Ogunwolu, C. Onyedikam, A SOFT COMPUTING APPROACH TO TRIP GENERATION ESTIMATION IN LAGOS METROPOLIS, NIGERIA, *Journal of Civil Engineering, Science and Technology*, 13 (2022) 6–22.<http://dx.doi.org/10.33736/jcest.3821.2022>
- [16] A.B.H. Kueh, Artificial neural network and regressed beam–column connection explicit mathematical moment–rotation expressions, *Journal of Building Engineering*, 43 (2021) 103195.<http://dx.doi.org/https://doi.org/10.1016/j.jobe.2021.103195>
- [17] Y.X. Tang, Y.H. Lee, M. Amran, R. Fediuk, N. Vatin, A.B. Kueh, Y.Y. Lee, in: *Sustainability*,

2022.

- [18] C.K. Ma, Y.H. Lee, A.Z. Awang, W. Omar, S. Mohammad, M. Liang, Artificial neural network models for FRP–repaired concrete subjected to pre-damaged effects, *Neural Computing and Applications*, 31 (2019) 711–717.<http://dx.doi.org/10.1007/s00521-017-3104-7>
- [19] M. Mirmehdi, Xie, X., & Suri, J. , *Handbook of texture analysis*, London: Imperial College Press, 2008.
- [20] J. Zhang, T. Tan, Brief review of invariant texture analysis methods, *Pattern Recognition*, 35 (2002) 735–747.[http://dx.doi.org/10.1016/S0031-3203\(01\)00074-7](http://dx.doi.org/10.1016/S0031-3203(01)00074-7)
- [21] E.R. Davies, *Texture analysis*, Computer Vision (Fifth Edition), Elsevier, London, 2018.
- [22] L. Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa, M. Pietikäinen, From BoW to CNN: Two Decades of Texture Representation for Texture Classification, *International Journal of Computer Vision*, 127 (2019) 74–109.<http://dx.doi.org/10.1007/s11263-018-1125-z>
- [23] K. Alrfou, A. Kordjazi, P. Rohatgi, T. Zhao, Synergy of unsupervised and supervised machine learning methods for the segmentation of the graphite particles in the microstructure of ductile iron, *Materials Today Communications*, 30 (2022) 103174.<http://dx.doi.org/https://doi.org/10.1016/j.mtcomm.2022.103174>
- [24] D. Rose, J. Forth, H. Henein, T. Wolfe, A.J. Qureshi, Automated semantic segmentation of NiCrBSi–WC optical microscopy images using convolutional neural networks, *Computational Materials Science*, 210 (2022) 111391.<http://dx.doi.org/10.1016/j.commatsci.2022.111391>
- [25] X. Xie, M. Mirmehdi, TEXEMS: Texture Exemplars for Defect Detection on Random Textured Surfaces, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (2007) 1454–1464.<http://dx.doi.org/10.1109/TPAMI.2007.1038>
- [26] S.Y. Wang, P.Z. Zhang, S.Y. Zhou, D.B. Wei, F. Ding, F.K. Li, A computer vision based machine learning approach for fatigue crack initiation sites recognition, *Computational Materials Science*, 171 (2020) 109259.<http://dx.doi.org/https://doi.org/10.1016/j.commatsci.2019.109259>
- [27] C. Hua, S. Chen, G. Xu, Y. Lu, B. Du, Defect identification method of carbon fiber sucker rod based on GoogLeNet-based deep learning model and transfer learning, *Materials Today Communications*, 33 (2022) 104228.<http://dx.doi.org/https://doi.org/10.1016/j.mtcomm.2022.104228>
- [28] C. Hu, H. Liao, T. Zhou, A. Zhu, C. Xu, Online recognition of magnetic tile defects based on UPM–DenseNet, *Materials Today Communications*, 30 (2022) 103105.<http://dx.doi.org/https://doi.org/10.1016/j.mtcomm.2021.103105>
- [29] V. Singh, R. Mishra, Developing a machine vision system for spangle classification using image processing and artificial neural network, *Surface and Coatings Technology*, 201 (2006) 2813–2817.<http://dx.doi.org/10.1016/j.surfcoat.2006.05.031>
- [30] Z. Zhen, C. Qixin, Y. Yang, X. Nianjiong, A feature extraction method for galvanized steel sheet powdering rates classification, *Imaging Science Journal – IMAGING SCI J*, 57 (2009) 94–100.<http://dx.doi.org/10.1179/174313108X384061>
- [31] L. Kai, A spangle image recognition method, 2021, CN113129290A
- [32] Y. Zhang, D. Zhou, S. Chen, S. Gao, Y. Ma, Single–Image Crowd Counting via Multi–Column Convolutional Neural Network, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27–30 June 2016, 589–597.<http://dx.doi.org/10.1109/CVPR.2016.70>
- [33] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray–scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (2002) 971–987.<http://dx.doi.org/10.1109/TPAMI.2002.1017623>
- [34] K.I. Laws, *Textured Image Segmentation*, University of Southern California, the States, 1980.

- [35] C. Harris, M. Stephens, A combined corner and edge detector, In Proc. of Fourth Alvey Vision Conference, 1988, 147–151–147–151
- [36] E. Alparslan, Componentwise edge detection by Laplacian operator masks, Signal Processing, 2 (1980) 179–183.[http://dx.doi.org/10.1016/0165-1684\(80\)90009-2](http://dx.doi.org/10.1016/0165-1684(80)90009-2)
- [37] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, SLIC Superpixels Compared to State-of-the-Art Superpixel Methods, IEEE Transactions on Pattern Analysis and Machine Intelligence, 34 (2012) 2274–2282.<http://dx.doi.org/10.1109/TPAMI.2012.120>
- [38] H. Idrees, K. Soomro, M. Shah, Detecting Humans in Dense Crowds Using Locally-Consistent Scale Prior and Global Occlusion Reasoning, IEEE Transactions on Pattern Analysis and Machine Intelligence, 37 (2015) 1986–1998.<http://dx.doi.org/10.1109/TPAMI.2015.2396051>
- [39] T. Zhao, R. Nevatia, B. Wu, Segmentation and Tracking of Multiple Humans in Crowded Environments, IEEE Trans. Pattern Anal. Mach. Intell., 30 (2008) 1198–1211.<http://dx.doi.org/10.1109/TPAMI.2007.70770>
- [40] M. Li, Z. Zhang, K. Huang, T. Tan, Estimating the Number of People in Crowded Scenes by MID Based Foreground Segmentation and Head–shoulder Detection, (2008) 1–4.<http://dx.doi.org/10.1109/ICPR.2008.4761705>
- [41] W. Ge, R. Collins, Marked point processes for crowd counting, 2012 IEEE Conference on Computer Vision and Pattern Recognition, 0 (2009) 2913–2920.<http://dx.doi.org/10.1109/CVPRW.2009.5206621>
- [42] A. Bansal, K. Venkatesh, People Counting in High Density Crowds from Still Images, Computer Vision and Pattern Recognition, abs/1507.08445 (2015).<http://dx.doi.org/10.48550/arXiv.1507.08445>
- [43] Y. Li, X. Zhang, D. Chen, CSRNet: Dilated Convolutional Neural Networks for Understanding the Highly Congested Scenes, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 18–23 June 2018, 1091–1100.<http://dx.doi.org/10.1109/CVPR.2018.00120>
- [44] W. Liu, M. Salzmann, P. Fua, Context–Aware Crowd Counting, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 15–20 June 2019, 5094–5103.<http://dx.doi.org/10.1109/CVPR.2019.00524>
- [45] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum Likelihood from Incomplete Data Via the EM Algorithm, Journal of the Royal Statistical Society: Series B (Methodological), 39 (1977) 1–22.<http://dx.doi.org/10.1111/j.2517-6161.1977.tb01600.x>
- [46] Q. Jun, H. Zhong–Shi, A SVM face recognition method based on Gabor–featured key points, 2005 International Conference on Machine Learning and Cybernetics, 18–21 Aug. 2005, 5144–5149 Vol. 5148.<http://dx.doi.org/10.1109/ICMLC.2005.1527850>
- [47] H. Wang, Y. Shi, X. Zhou, Q. Zhou, S. Shao, A. Bouguettaya, Web Service Classification Using Support Vector Machine, (2010) 3–6.<http://dx.doi.org/10.1109/ICTAI.2010.9>
- [48] C. Liu, X. Weng, Y. Mu, Recurrent Attentive Zooming for Joint Crowd Counting and Precise Localization, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 15–20 June 2019, 1217–1226.<http://dx.doi.org/10.1109/CVPR.2019.00131>

Table 1. Performance of models with different structure

Model	MAE	MSE	RMSE	NRMSE	MAPE	R <sup>2</sup>
MCNN	4.4	32.5	5.7	16.6%	14.3%	0.79
OCNN-large	4.4	33.0	5.7	16.7%	14.5%	0.78
OCNN-medium	4.7	41.9	6.1	17.7%	14.9%	0.67
OCNN-small	5.2	47.1	6.9	19.9%	16.4%	0.69

Table 2. MAE, RMSE for models with different structure

Model	MAE				RMSE			
	total	level 2	level 3	level 4	total	level 2	level 3	level 4
MCNN	4.2	5.8	3.6	1.3	5.3	7.3	4.5	1.7
OCNN-large	4.3	5.5	4.0	1.5	5.4	7.0	5.0	1.9
OCNN-medium	4.5	6.0	4.1	1.5	5.8	7.6	5.2	2.0
OCNN-small	5.2	6.7	4.5	2.4	6.6	8.6	5.7	3.4

Table 3. Classification results of different algorithms

Features Extraction	Accuracy (%)			
	total	level 2	level 3	level 4
MCNN ( $\sigma=0.5$ ) + GMM	85.4	88.1	82.8	96.9
MCNN ( $\sigma=2$ ) + GMM	89.8	91.6	84.6	100.0
MCNN ( $k=2, \beta=0.05$ ) + GMM	89.7	89.5	56.6	100.0
MCNN ( $k=4, \beta=0.15$ ) + GMM	91.3	91.7	88.5	98.7
LBP	50.9	35.8	56.2	80.8
Laws	62.7	87.5	31.3	83.3
SLIC	42.8	32.1	56.2	34.6
Harris	67.3	87.5	41.7	84.6
Laplace	41.3	50.0	40.0	20.5

Fig.1 Camera settings for capturing morphology of spangle on galvanized sheet

Fig.2 Typical spangle morphologies of different levels acquired by area-scan camera (a) level 2 (b) level 3 (c) level 4

Fig.3 Optical morphologies of level 2 spangles

Fig.4 Original spangle image (a) and image with annotated spangle nuclei and corresponded edge (b)

Fig.5 Statistics of nuclei number and average nearest neighbor distance of total 486 images

Fig.6 Prediction algorithm framework of spangle grading

Fig.7 MCNN model framework

Fig.8 Progression of counting error on test dataset with the increase of training data scale

Fig.9 Progression of MAE (a) and RMSE (b) for each level images on test dataset with the increase of training data scale

Fig.10 Progression of MAE (a) and RMSE (b) on both training and test datasets with the increase of training data scale

Fig.11 Density maps of ground truth for level 2 image (a) and estimated by MCNN (b), OCNN-large (c) OCNN-medium (d) and OCNN-small

Fig.12 Comparison of ground truth and estimated density maps for different level images (a) level 2 image (b) level 3 image (c) level 4 image

Fig.13 MAE and RMSE of MCNN models with different kernels

Fig.14 Density maps of ground truth generated by fixed and variable Gaussian kernels and corresponded estimated density maps for level 2 images

Fig.15 (a) MAE and RMSE of ANND and (b) precision and recall of localization with the change of Gaussian kernel

Fig.16 Spangle images with manual annotated nuclei positions (a) and estimated density maps and nuclei positions when fixed Gaussian kernels were  $\sigma=0.5$  (b)  $\sigma=2$  (c)  $\sigma=4$  (d) and  $\sigma=8$  (e)

Fig.17 Original spangle images (a) and images with pixels of interest by algorithms of LBP (b), Laws (c), Harris (d), Laplace (e), SLIC (f) and combination of MCNN and GMM (g)

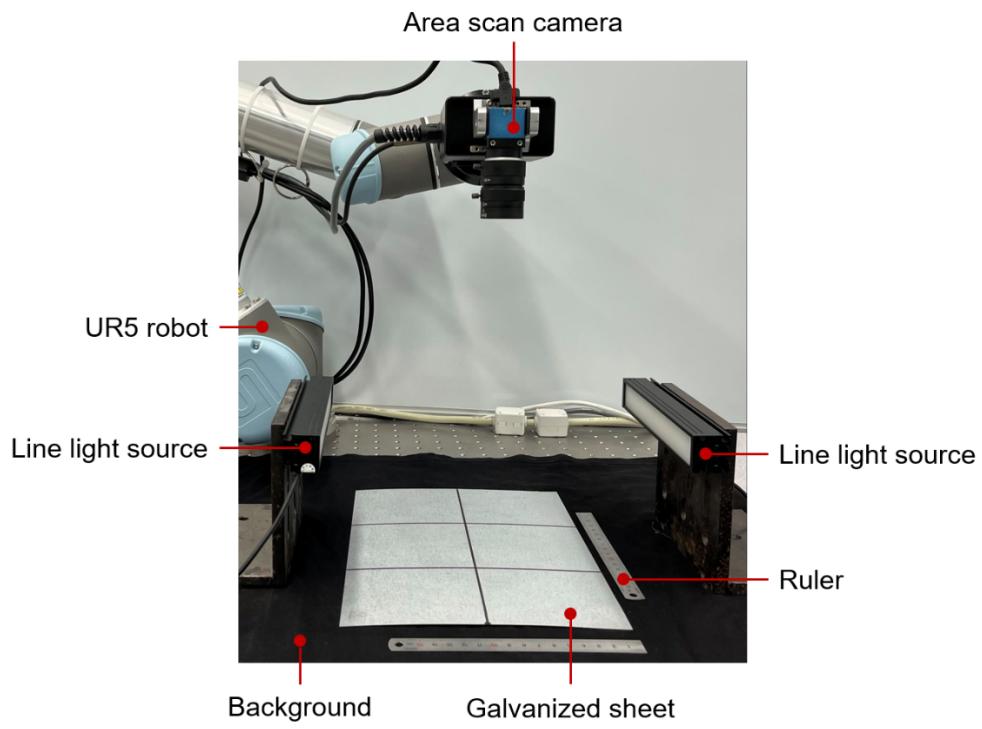


Fig.1 Camera settings for capturing morphology of spangle on galvanized sheet

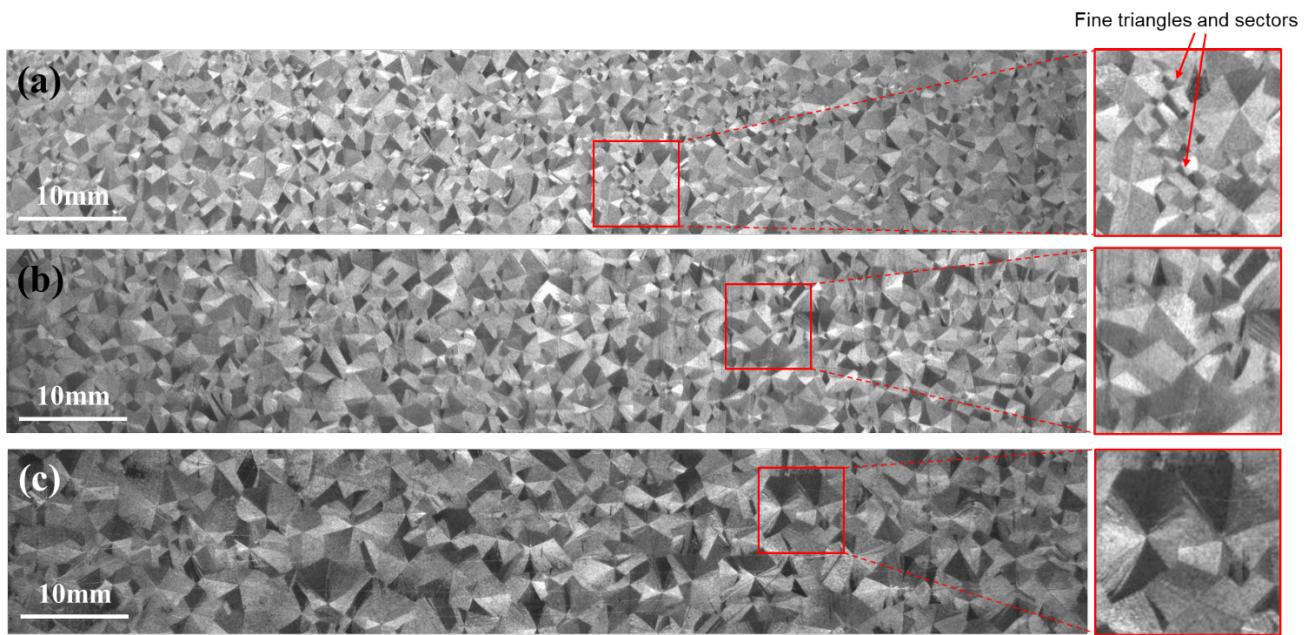


Fig.2 Typical spangle morphologies of different levels acquired by area-scan camera (a) level 2 (b) level 3 (c) level 4

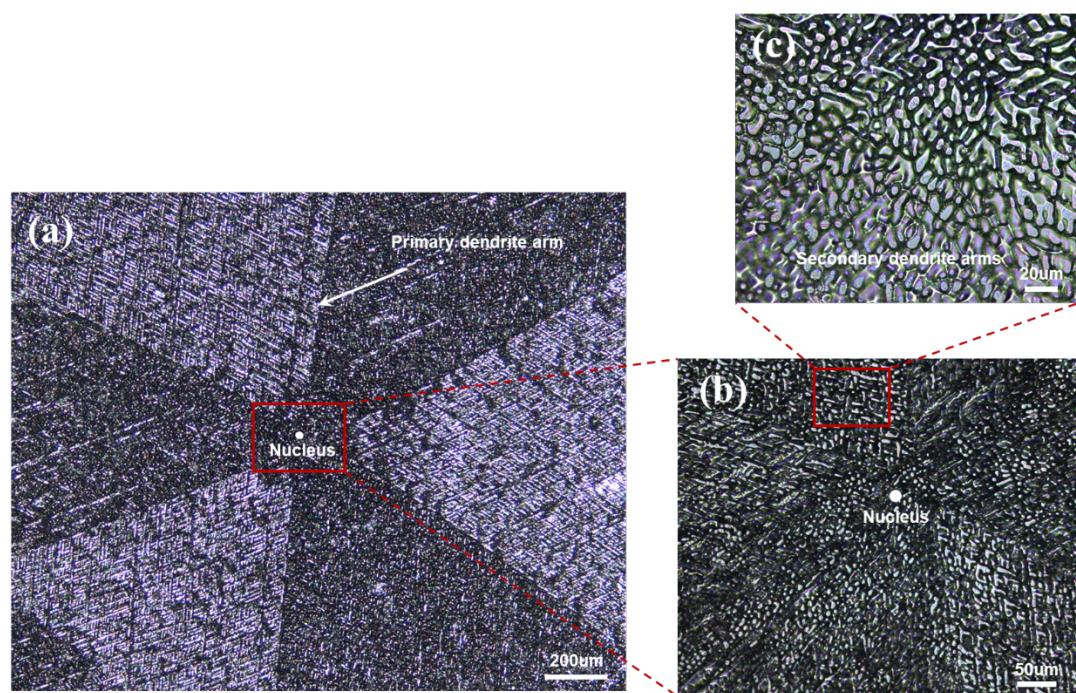


Fig.3 Optical morphologies of level 2 spangles

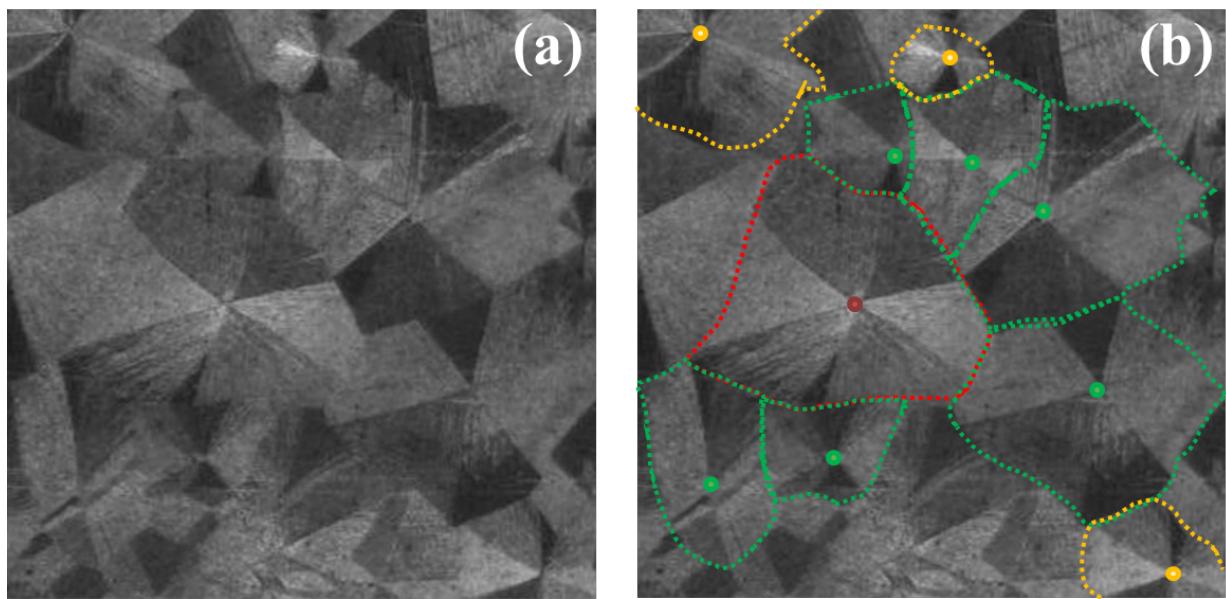


Fig.4 Original spangle image (a) and image with annotated spangle nuclei and corresponded edge (b)

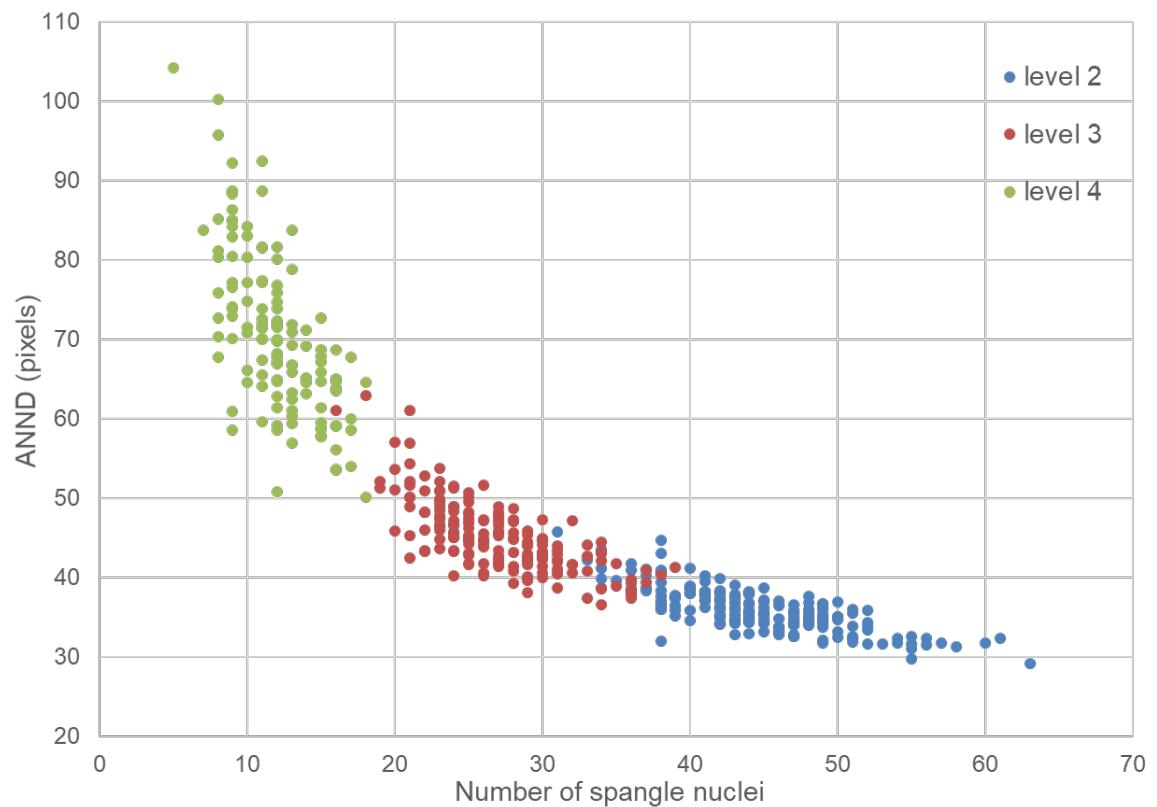


Fig.5 Statistics of nuclei number and average nearest neighbor distance of total 486 images

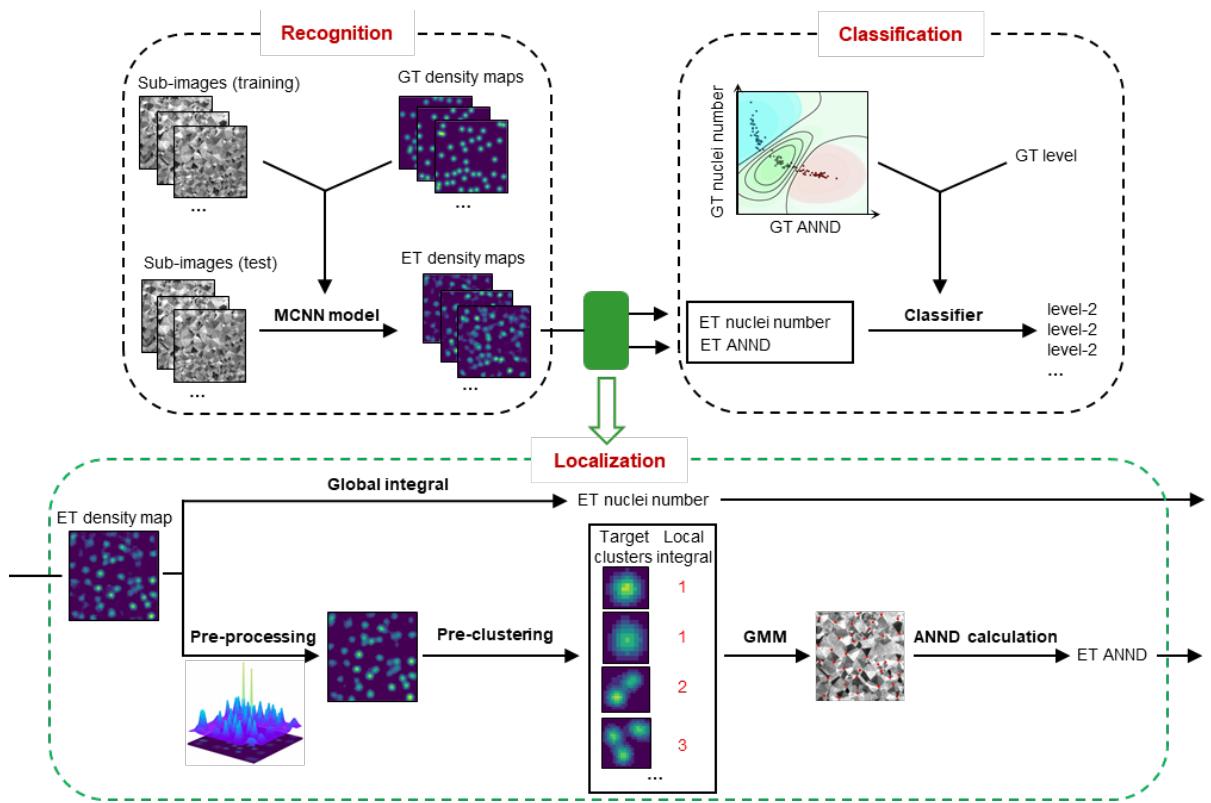


Fig.6 Prediction algorithm framework of spangle grading

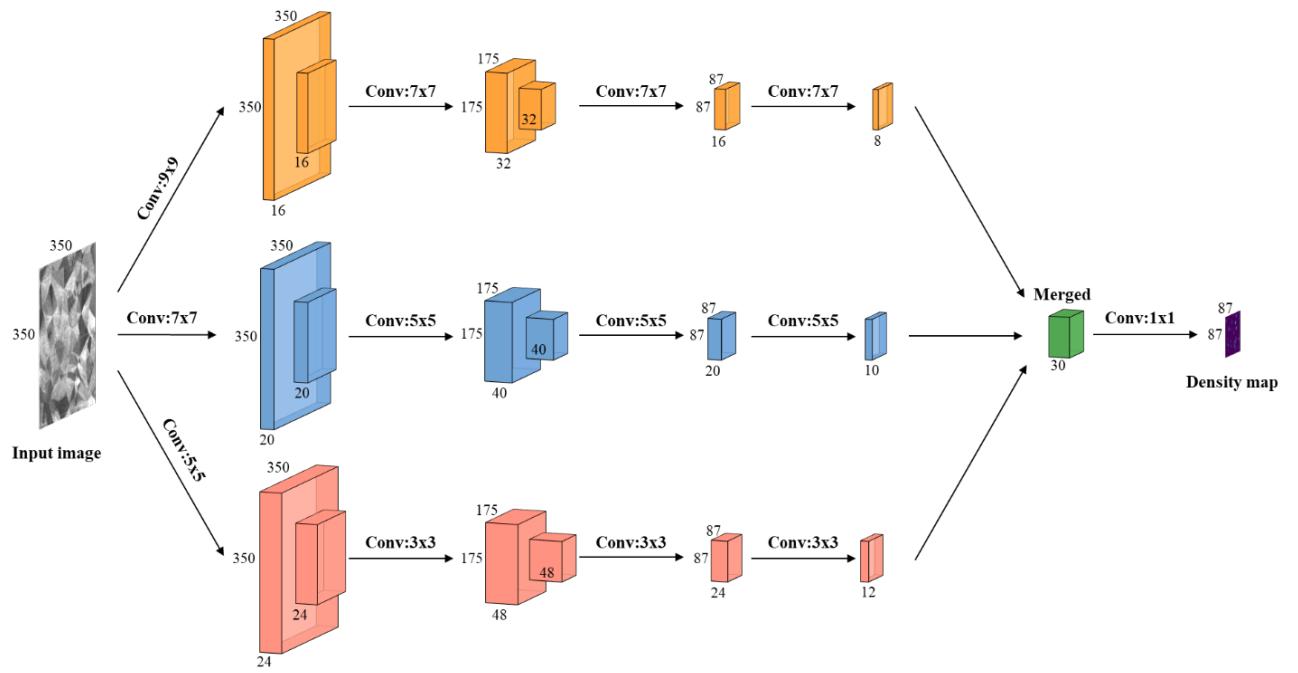


Fig.7 MCNN model framework

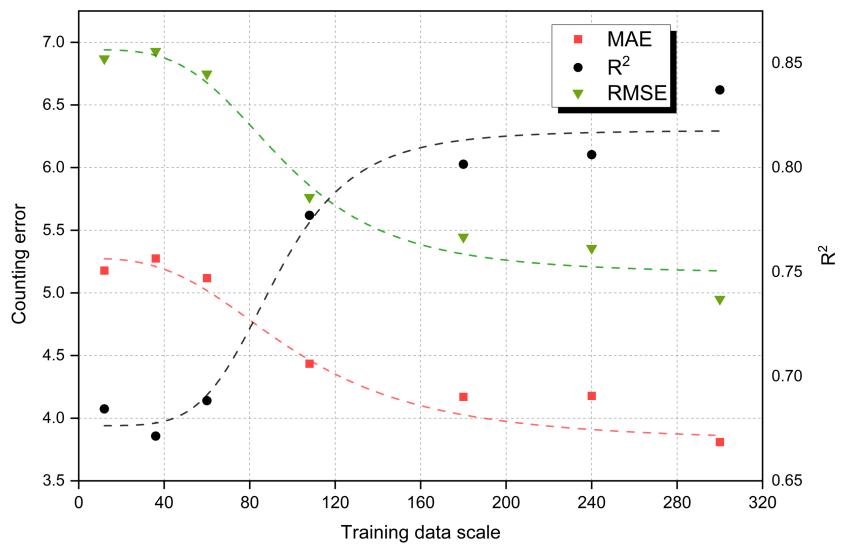


Fig.8 Progression of counting error on test dataset with the increase of training data scale

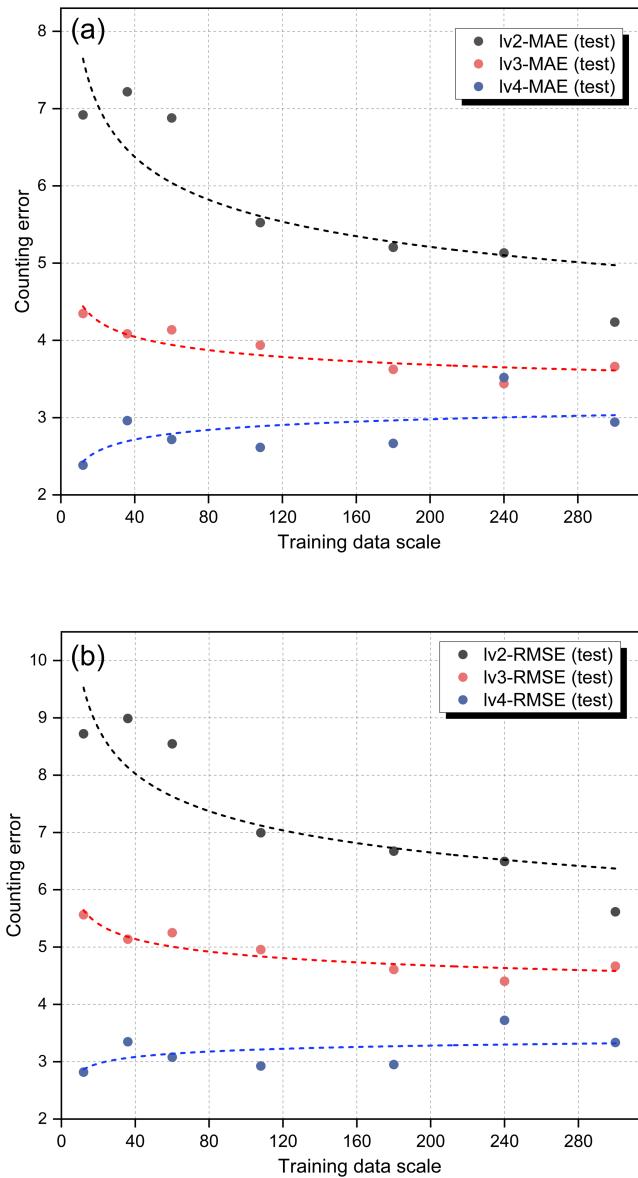


Fig.9 Progression of MAE (a) and RMSE (b) for each level images on test dataset with the increase of training data scale

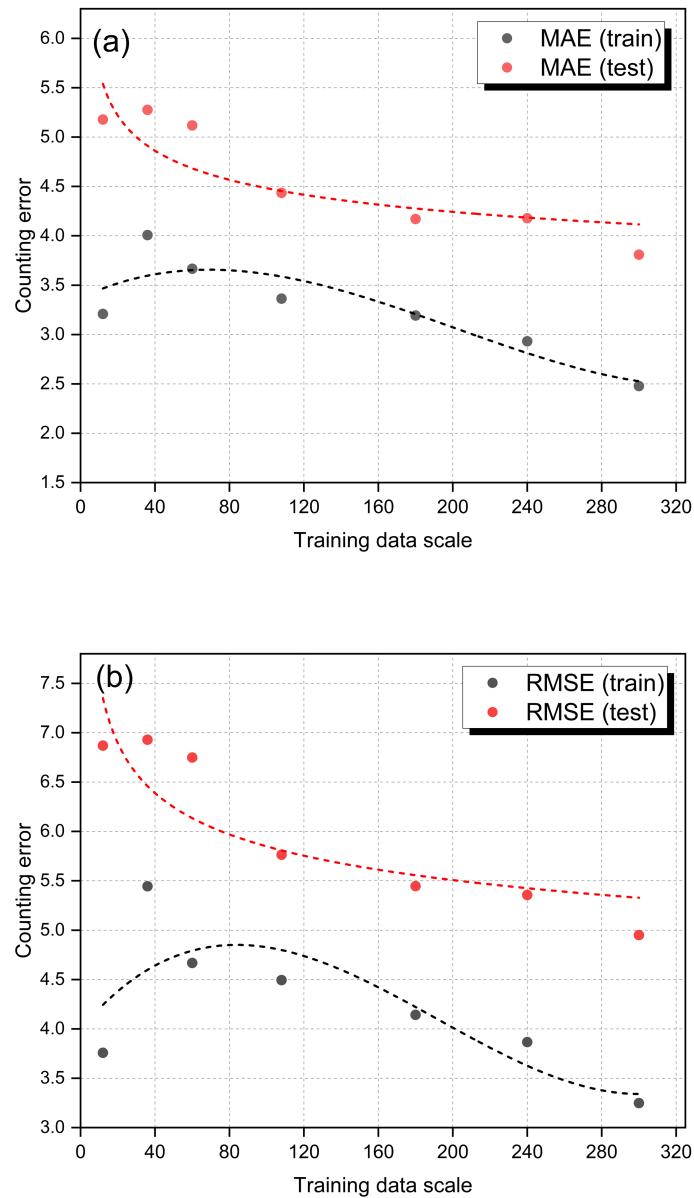


Fig.10 Progression of MAE (a) and RMSE (b) on both training and test datasets with the increase of training data scale

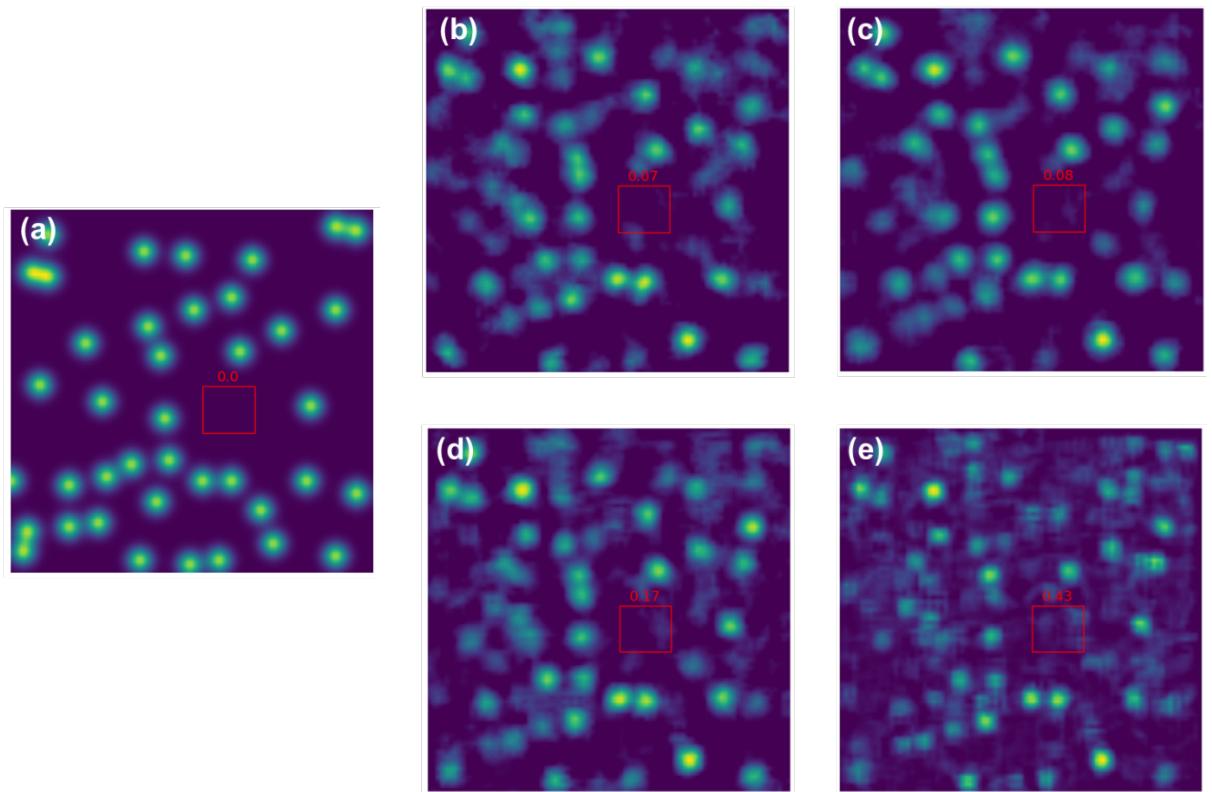


Fig.11 Density maps of ground truth for level 2 image (a) and estimated by MCNN (b),  
OCNN-large (c) OCNN-medium (d) and OCNN-small

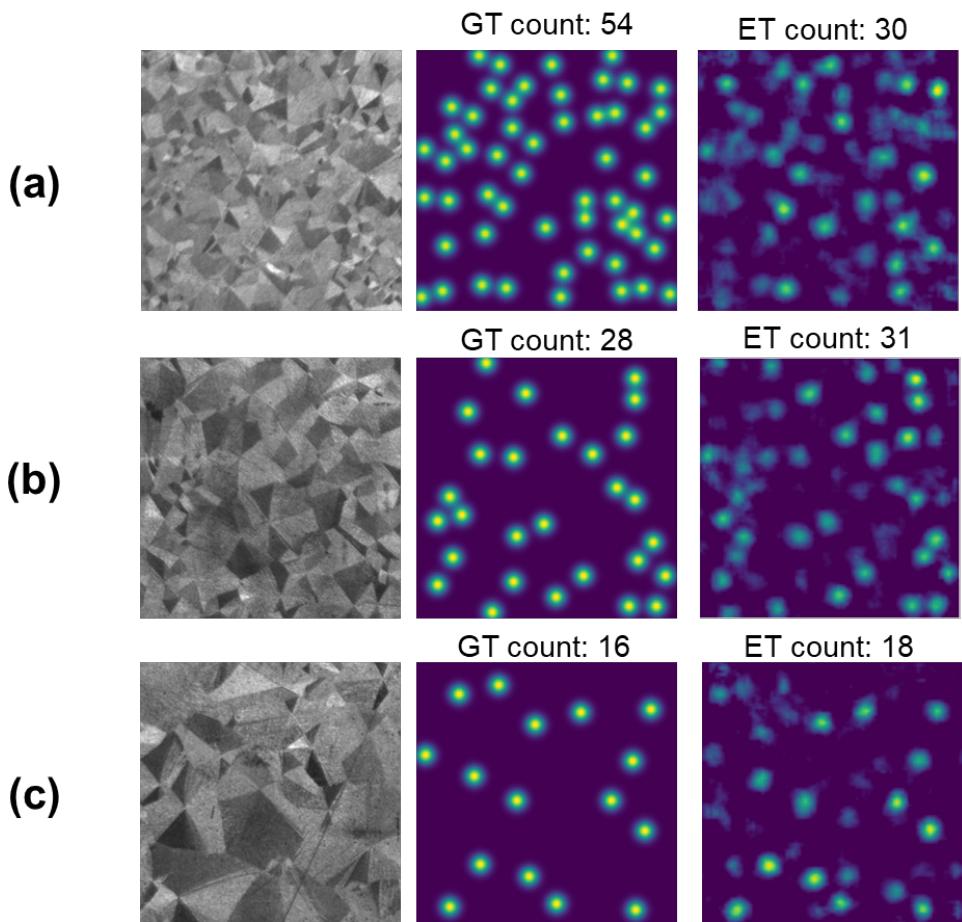


Fig.12 Comparison of ground truth and estimated density maps for different level images (a) level 2 image (b) level 3 image (c) level 4 image

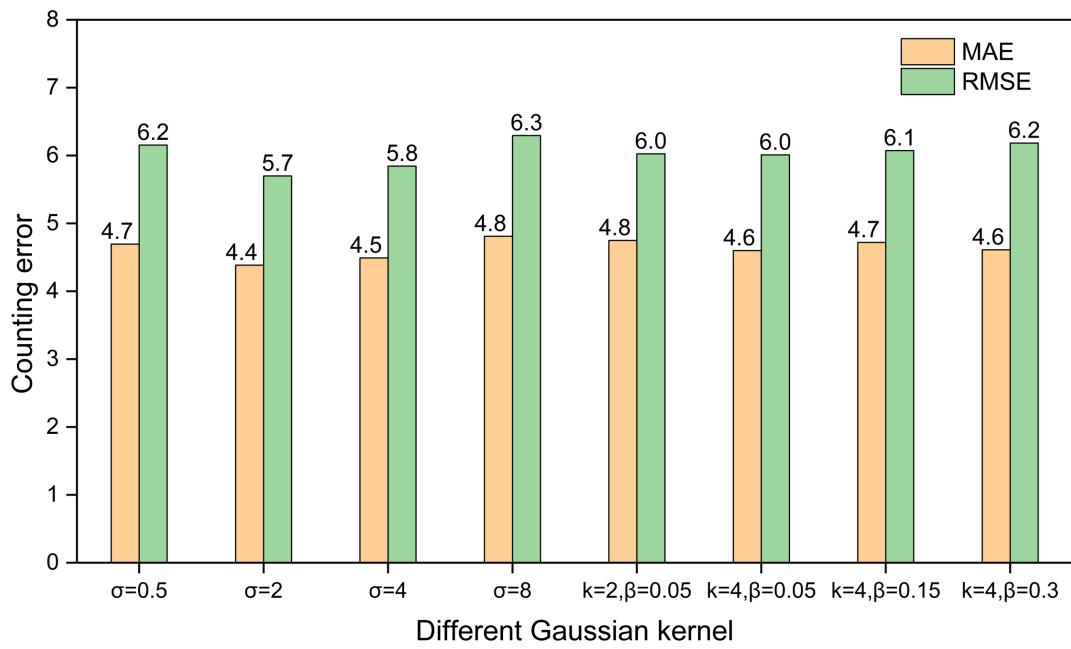


Fig.13 MAE and RMSE of MCNN models with different kernels

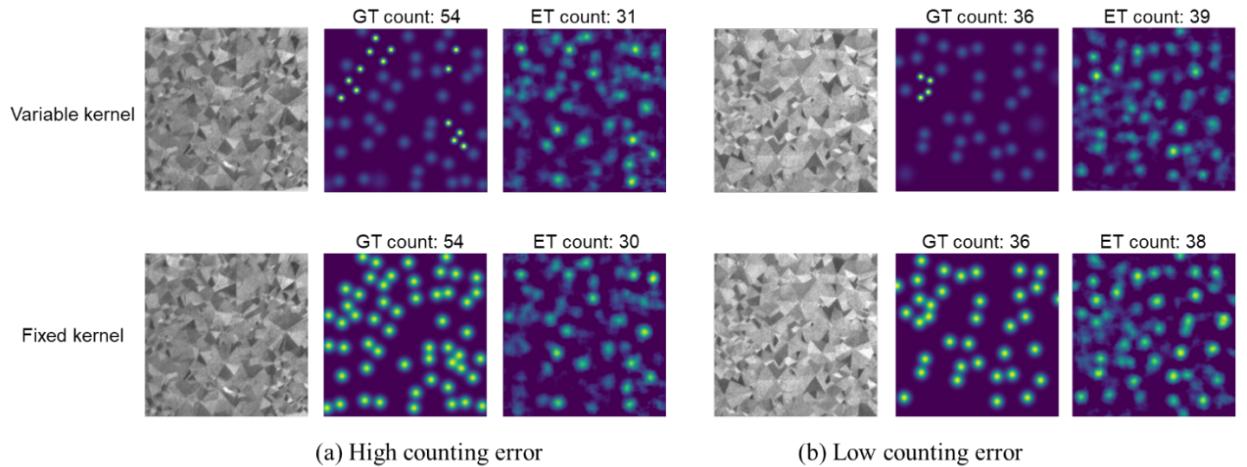


Fig.14 Density maps of ground truth generated by fixed and variable Gaussian kernels and corresponded estimated density maps for level 2 images

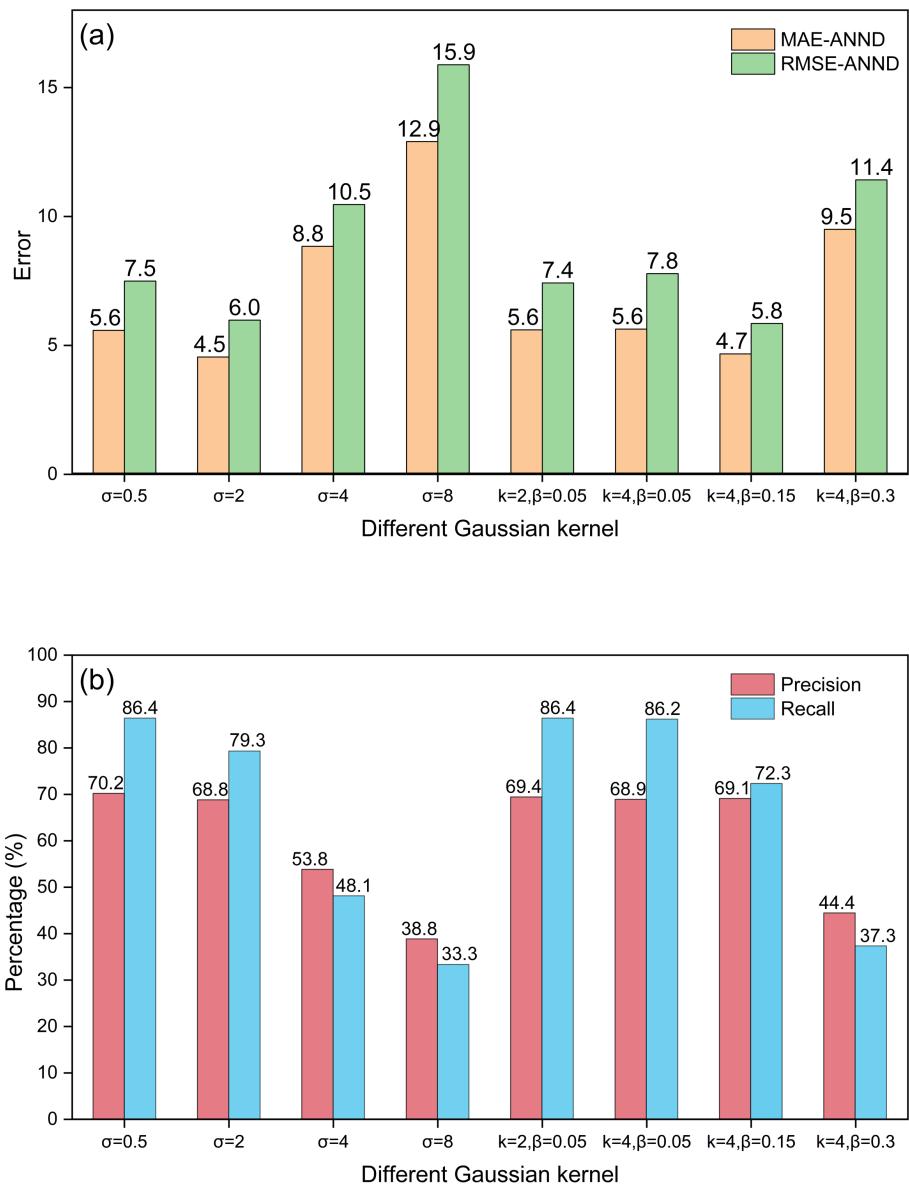


Fig.15 (a) MAE and RMSE of ANND and (b) precision and recall of localization with the change of Gaussian kernel

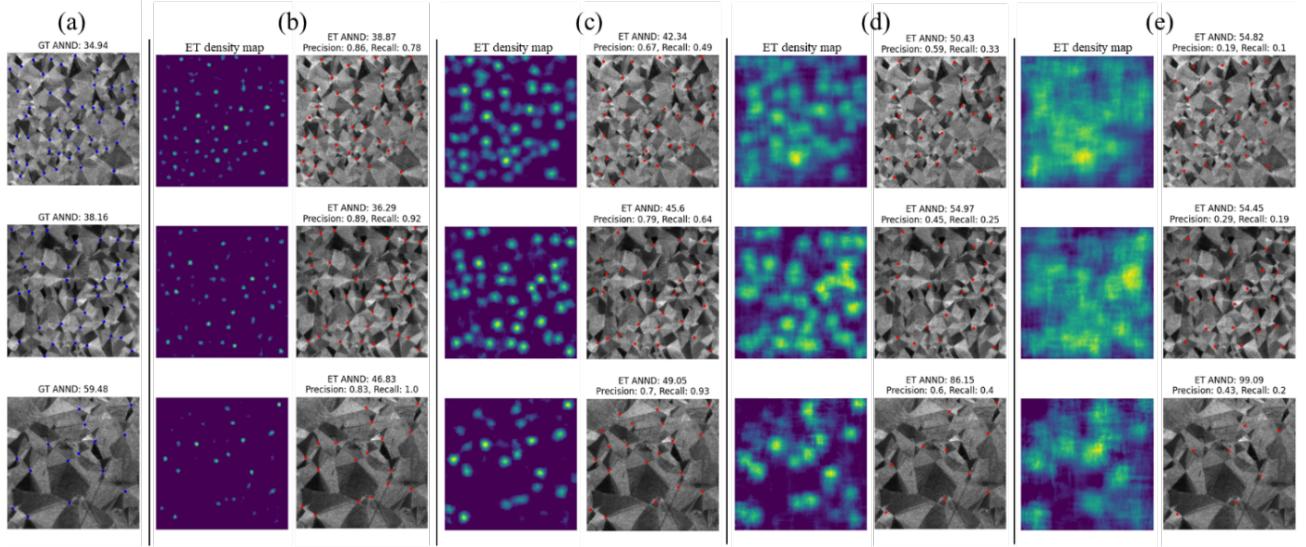


Fig.16 Spangle images with manual annotated nuclei positions (a) and estimated density maps and nuclei positions when fixed Gaussian kernels were  $\sigma=0.5$  (b)  $\sigma=2$  (c)  $\sigma=4$  (d) and  $\sigma=8$  (e)

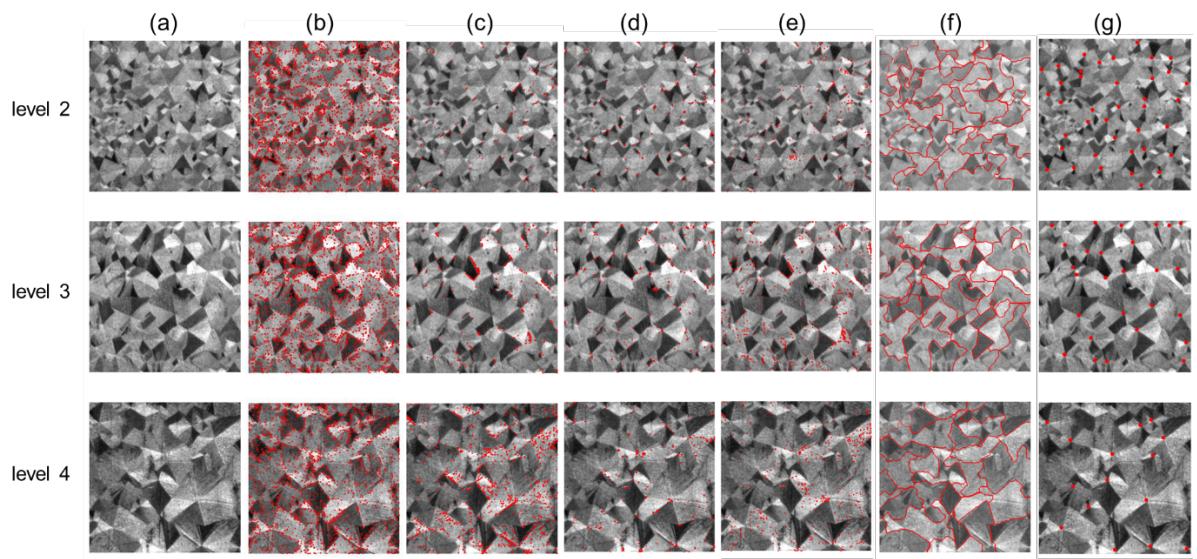


Fig.17 Original spangle images (a) and images with pixels of interest by algorithms of LBP (b), Laws (c), Harris (d), Laplace (e), SLIC (f) and combination of MCNN and GMM (g)