

RESEARCH

Open Access



Ensemble feature learning for material recognition with convolutional neural networks

Peng Bian¹, Wanwan Li^{2,3}, Yi Jin^{2,3*} and Ruicong Zhi^{4,5}

Abstract

Material recognition is the process of recognizing the constituent material of the object, and it is a crucial step in many fields. Therefore, it is valuable to create a system that could achieve material recognition automatically. This paper proposes a novel approach named ensemble learning for material recognition with convolutional neural networks (CNNs). In the proposed method, firstly, a CNN model is trained to extract the image features. Secondly, knowledge-based classifiers are learned to get the probabilities of the test sample that belongs to different material categories. Finally, we propose three different ways to learn the ensemble features, which achieves higher recognition accuracy. The great difference from the prior work is that we combine the knowledge-based classifiers on probability level. Experimental results show that the proposed ensemble feature learning method performs better than the state-of-the-art material recognition methods and can archive a much higher recognition accuracy.

Keywords: Material recognition, Convolutional neural networks, Ensemble learning

1 Introduction

In our daily life, the scenes consist of all kinds of materials such as leather, fabric, and wood. We deal with all kinds of categories and apperceive the material categories. For example, we would try to avoid stepping on the puddles of water when we walk on the road. So material recognition plays a critical role in real world and exerts great influence on human life. Therefore, it is valuable to design a system that could achieve material recognition automatically. It could provide valuable cues to numerous applications, including product search and human-machine systems. For example, an autonomous vehicle or a mobile robot can

make decisions on whether a forthcoming terrain is asphalt, gravel, or grass. A cleaning robot can distinguish among wood, tile, or carpet.

In the past few decades, so many methods were proposed for material recognition; however, no satisfactory accuracy has been achieved. And the traditional method for material recognition is modeling the characteristic appearance of different materials or the context information. However, quite a few categories of materials would have abundant surfaces and are visually very wealthy. And the visual characteristic would perform differently due to the lighting and scene. The challenge in material recognition is largely due to the wide variety of appearances which each material may exhibit, such as plastic,

* Correspondence: yjin@bjtu.edu.cn; camus_0616@163.com

²School of Computer and Information Technology, Beijing Jiaotong University, No.3 Shangyuancun Road, Beijing 100044, China

³Beijing Key Lab of Traffic Data Analysis and Mining, No.3 Shangyuancun Road, Beijing 100044, China

Full list of author information is available at the end of the article

which may appear in a number of different colors, textures, and reflectance. There are also many other influencing factors for material recognition, such as reflection estimation, the illumination condition, and other conditions. All of them would have a huge influence on material recognition performance.

In recent years, a major breakthrough in material recognition is that large-scale databases combined with convolutional neural networks (CNNs). CNN has proved its success in vision tasks such as object detection and recognition and has recently achieved the state-of-the-art results in object classification and detection. Many advanced architectures have been introduced, such as GoogLeNet [1] and VGG [2]. It is also used for per-pixel segmentation. Farabet et al. [3] proposed multi-scale CNNs to predict category on pixel level for segmentation. Recently, a fully convolutional framework [4] has been proposed; it predicts from the image directly. The recent work includes R-CNNs [5] and overfeat [6]. On the other hand, the material recognition has been facilitated by the large-scale databases. There are already many large-scale databases; the detail introduction of the databases would be shown in the next part. The common ground of these databases is that they contain so many kinds of categories, and each of the categories contains a large number of images. And these databases may consider the illumination, angle, and other factors.

One of the major contributions of our paper is that we achieve higher recognition accuracy than the prior methods by combining some knowledge-based classifiers. The great difference from the previous work is that we make use of the category probabilities, not the category label. We combine these probabilities through three ways to increase the probability of correct category, and reduce the probability of error category, that is to say, achieve higher recognition accuracy. Meanwhile, a new algorithm is proposed for learning the weights for these knowledge-based classifiers. Experiment results show that ensemble learning for material recognition with convolutional neural network could get higher accuracy than a knowledge-based classifier. And we also do some experiments with the prior methods; the results show that the way we proposed to combine the knowledge-based classifiers would have a better performance than the prior methods.

The contributions of this new approach can be summarized as follows:

- (1) We introduce three ways of ensemble learning for material recognition to get higher accuracy.
- (2) A new algorithm is proposed to learn weights for knowledge-based classifiers.

The remainder of this paper is organized as follows. Section 2 is the related work, and Section 3 describes the ensemble learning for material recognition with CNNs. Experiment results and discussion on the material database are presented in Section 4. Section 5 draws the conclusion of this paper.

2 Related work

In this paper, we recognize material categories with CNNs and improve the recognition accuracy by ensemble learning. There are already some perfect methods for every part.

2.1 Material recognition

As described above, the material database plays an important role in the material recognition. There are already various material databases, and the material databases would have a major impact on the recognition accuracy. CURET database [7] contains 61 material samples, and each sample is captured under 205 different lighting and view conditions. Flicker Material Database (FMD) [8] contains ten categories, and each category contains 100 images, which is very choosy to contribute to the rich visual variation in each category. Although FMD has already been applied to material recognition, it is not enough for the material recognition in the real world. Bell et al. [9] released OpenSurfaces that contains more than 20,000 real-world-labeled scenes; this database is a large number material database. The Materials in Context Database (MINC) [10] contains more than 3 million patches, and these patches are classified into 23 material categories (Fig. 1). There are also many large image sets collected in the wild [11–13]. The more detailed information is shown in Table 1.

And in the past few decades, many material recognition methods had been proposed. The major method is modeling the characteristic appearance of different materials or the context information. And many prior methods of material recognition have focused on the classification problem, which would be divided into two categories, the first one is based on the object reflectance [14–17]. Cula and Dana [14] proposed 3D texture recognition using bidirectional feature histograms. Liu and Gu [15] proposed to use coded illumination to directly measure discriminative feature for material classification. Lombardi and Nishino [16] forced on single image that consists of multiple materials and proposed to constrain the possible solutions so that the recovered reflectance conform with those real-world materials. Zhang et al. [17] introduced a framework called reflectance hashing that modeled the reflectance disks with dictionary learning and binary hashing. Most of this

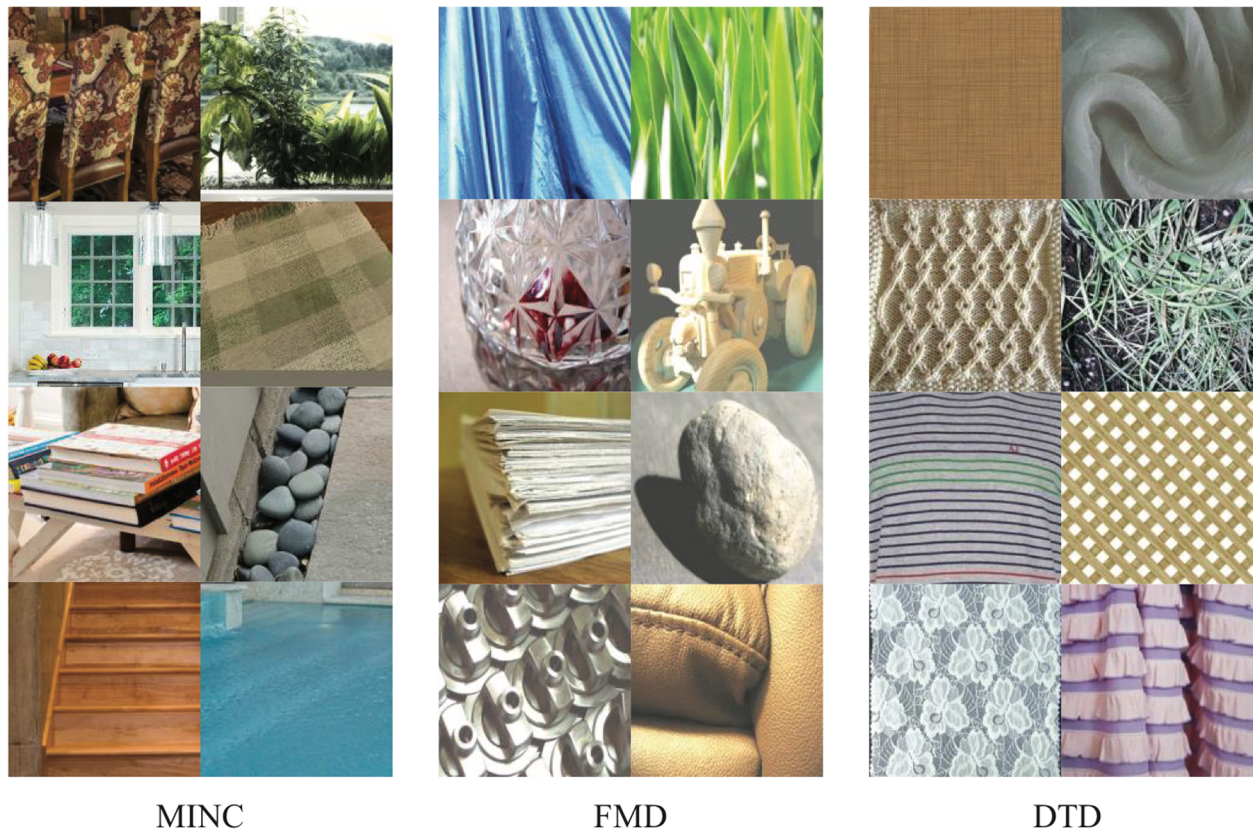


Fig. 1 Some examples in the MINC database, FMD database, and DTD database

work would require some added condition, such as the scene geometry [18, 19] or illumination [20, 21] should to be known ahead of time, or others [22]. The other kind of the prior method is extracting features from the image appearance directly. For example, Liu et al. [23] proposed a method that package a few features into reflectance-based edge feature. Hu et al. [24] proposed features based on

variances of oriented gradients. Qi et al. [25] studied the transform invariance (TI) of co-occurrence features and introduced a pairwise local binary pattern (LBP) feature. Schwartz and Nishino [26] introduced visual discriminative object-specific information. And in the recent days, Cimpoi et al. [27] achieved state-of-art on FMD by combining object descriptors and texture descriptors. Recently, many methods

Table 1 Comparison of the publicly available material databases

Database	Sample	Categories	Source	Time
CURvT	61		Unknown	1999
KTH-TIPS	11	11	Unknown	2004
FMD	100	10	Flick	2009
Open-Surface	105,000 (segmentations)	22	Flick	2013
UBO2014	84	7	Unknown	2014
Reflectance disk	190	19	Unknown	2015
MINC	3,000,000	23	Flick Houzz	2015
4D Light-field	1200	12	Unknown	2016
NISAR	100	100	Unknown	2016
GTOS	606	40	Unknown	2016

based on deep learning are proposed. For example, Xue et al. [13] propose to take a middle-ground approach for material recognition that takes advantage of both rich radiometric cues and flexible image capture. Xu et al. [28] propose a material recognition approach based on both transfer learning and a mainstream deep convolutional neural network (D-CNN) model used for foreign object debris (FOD) material recognition. Younis et al. [29] utilized the advances in deep learning to build a system for material recognition.

2.2 Ensemble learning

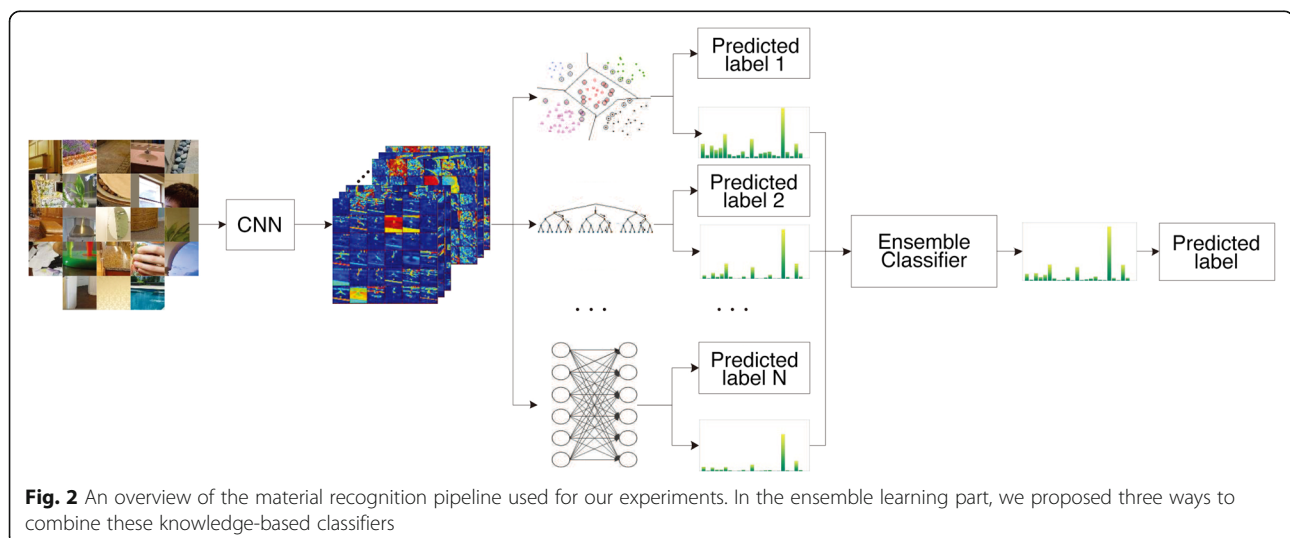
Ensemble learning [30] is the machine learning that combines various classifiers to achieve better prediction performance. That is to say, the ensemble learning approaches attempt to predict some results and package the results to generate a new final result [31]. The prior methods show that the ensemble performed better than the knowledge-based classifiers. Generally speaking, the ensemble learning methods are constructed in two steps. First, build some knowledge-based classifiers, and then combine these classifiers. The main algorithms are divided into boosting, bagging, and stacking. Bagging is the method that involves different types of classifiers. Each knowledge-based classifier has its own training set that generated from using random draw method with replacement [32]. Build a model for each classifier after all the training sets are generated. The final prediction result is combined by voting. This ensemble learning method reduces the overfitting problem, and it will be more effective on unstable learning algorithms. Boosting also has a superior performance

on improving the prediction accuracy of some machine learning models. In all boosting algorithms, during each learning phase, the instances are reweighted [33]. The wrong classified instances would be selected in the next step, so that they could be classified correctly in this step. All the results of the classifiers are combined by the majority voting [34]. Stacking is to make the predicted results of the classifiers as new features and train the final model on the new training set. Last but not the least, be sure to choose the different knowledge-based classifiers for your ensemble learning.

3 Proposed method

We start with three proposed ways to combine these knowledge-based classifiers on the probability level. Next, we propose a new algorithm for learning the weights for these knowledge-based classifiers. Then, we describe the application of the proposed methods to material recognition. Finally, we provide recognition accuracy computational analysis, which is a proof that the ensemble learning classifier can achieve higher recognition accuracy than a knowledge-based classifier.

Figure 2 shows an overview of our ensemble learning method for recognizing materials. Given a CNN that can extract feature for training these knowledge-based classifiers, so these knowledge-based classifiers are trained for ensemble learning. Then, we use the proposed methods to combine these knowledge-based classifiers. Specifically, we use a new algorithm for learning the weights of these knowledge-based classifiers in the third method. The biggest distinction from the previous methods is that in our method we realize the



Algorithm 1 Learning process for the weights

Input: The predicted labels of knowledge-based classifiers L_z ($z=1,2,\dots,n$), and the correct test labels L .

Output: The weights of these base classifiers.

- 1: Initialize the weights w_z ($z=1,2,\dots,n$); And set the step ;
 - 2: Let w_z ($z=1,2,\dots,n$) add the step, and w_j ($j \neq z$) reduce the step/ $n-1$ when $L_z = L$ and L_j ($j \neq z$) $\neq L$;
 - 3: The weights will be maintained in other cases;
 - 4: Normalize the weights;
 - 5: return w_z ($z=1,2,\dots,n$).
-

ensemble learning on the probability level and we would recalculate the probabilities of the test sample that belongs to the 23 material categories. The label corresponding to the maximum value of the 23 predicted probabilities is the predicted label for the test sample. The results show that the ensemble learning would get a higher accuracy than the knowledge-based classifiers, and our proposed ensemble learning ways would get a higher accuracy than the prior ways.

3.1 Training procedure for knowledge-based classifiers

First, we train our CNN model with the classifier of Softmax. Then, the CNN feature is extracted for training the knowledge-based classifiers. In the training procedure of the support vector machine (SVM) [35] classifier, we should generate the train set and test set from the original data, and preprocessed the data as

$$y = (y_{\max} - y_{\min}) \times (x - x_{\min}) \div (x_{\max} - x_{\min}) + y_{\min} \quad (1)$$

Then, we train the SVM classifier on the train set, and the decision function as

$$f(x) = \text{sgn}\left(\sum_{i=1}^n W_i \exp(-\text{gamma}\|x_i - x\|^2) + b\right) \quad (2)$$

We train the Random Forest [36] model to obey the Gini value.

$$\text{Gini} = 1 - \sum P(i)^2 \quad (3)$$

$P(i)$ is the data set of the present node. And we also train the other knowledge-based classifiers like the extreme learning machine model [37], Treebagger model.

3.2 Our proposed ensemble learning

As for the ensemble learning methods, different from the previous methods, we achieve the goal of combining

various knowledge-based classifiers on the probability level. In our approach, the predicted results of these knowledge-based classifiers could be obtained after training the knowledge-based classifiers; the predicted results would contain the predicted probabilities of the material categories and the predicted labels. Lastly, we propose three ways to compute the final probability map on the basis of the probability maps of the knowledge-based classifiers.

- The first way is to compute the mean value of the predicted probabilities of different material categories, then the final probability map would be generated. For a test sample, the label corresponding to the maximum value of each row is the predicted label of this test sample.

$$P(i, j) = \sum_{z=1}^n P_z(i, j) / n \quad (4)$$

$n = (1, 2, \dots, m)$, $P(i, j)$ means the probability of sample i belongs to material category j . z means the z th knowledge-based classifier.

- The second way is maximum value algorithm. The first step is comparing these predicted probabilities of each knowledge-based classifier, and build a new probability map that contains the maximum predicted probabilities of each knowledge-based classifier. The predicted label is the label corresponding to the maximum value of these maximum predicted probabilities for this test sample.

$$P(i) = \max(\max(P_z(i, :))) \quad (5)$$

z means the z th knowledge-based classifier. $P(i, :)$ means the predicted probabilities that the sample i belongs to the material categories.

- The third way is setting weights w_z for every knowledge-based classifier. A new learning algorithm is proposed to learn the weights of the knowledge-based classifiers. The algorithm is described as follows:
 - 1) The main idea of this algorithm is choosing one of the knowledge-based classifiers as the original classifier and improving recognition accuracy by making use of the other knowledge-based classifiers;
 - 2) Set the initial weights w_z for each knowledge-based classifier, and set the step;
 - 3) Compare the predicted labels to the correct test labels. When the predicted label of the knowledge-based classifier C_z is right, the predicted labels of the other knowledge-based classifiers are wrong, the weight w_z adds the step and the weight $w_j (j \neq z)$ reduces the step $n - 1$. The weights will be maintained in other cases;
 - 4) Normalize the weights. The whole method of the proposed learning process is summarized in Algorithm 1.

The final predicted probability map is the sum of the product of the knowledge-based classifiers and their corresponding weights. The label corresponding to the maximum value of these predicted probabilities is the predicted label of this test sample.

$$P(i, j) = \sum_{z=1}^n w_z P_z(i, j) \quad (6)$$

$z = (1, 2, \dots, n)$ means the z th knowledge-based classifier, $P(i, j)$ means the probability of sample i belongs to material category j .

4 Results and discussion

In this section, we evaluate the effect of many different parameters for training knowledge-based classifiers. Then, the best performance parameters are chosen to

train the knowledge-based classifiers, which are prepared for the ensemble learning. And we also evaluate our proposed ensemble learning methods and compare with the prior ensemble learning methods, including the stacking and voting. All of our experiments were carried on MINC [10] database.

4.1 Training procedure

We choose the MINC [10] for this database contains 3 million patches that split into 23 categories. And each category contains the same number of samples. And we train our CNN model with the AlexNet network framework [38], the champion of the challenge on ImageNet image classification on 2012. Like Bell et al. [10], when training AlexNet [38], we use stochastic gradient descent with batch size 128, dropout rate 0.5, momentum 0.9, and a base learning rate of 10^{-3} that decreases by a factor of 0.25 every 50,000 iterations. The network is showed in Fig. 3. After the training process, we would get the predicted labels of Softmax classifier. And we extract the CNN feature and get the predicted probabilities of the test sample that belongs to the 23 categories when recognized by the Softmax classifier.

We choose 2125 patches form each category as the train set and 250 patches from each category as the test set. The feature is the output of the first fully connected layer (fc6) of the CNN model, and the features are used for training the knowledge-based classifiers, such as SVM [35] and Random Forest [36]. In the training procedure of the SVM classifier, the train set and test set should be normalized firstly, then the SVM classifier trained with the LIBSVM [39] toolbox, and the acquisition of optimum parameter within a possible range is attributed the success to the cross validation (CV) method. Finally, the SVM classifier is trained on the train set.

In order to train the Random Forest classifier, we use the Random Forest open source toolbox, which

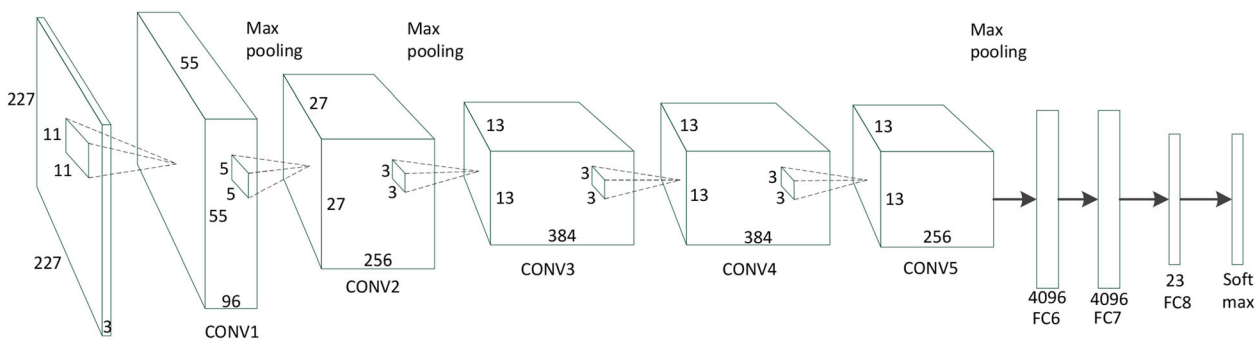


Fig. 3 The illustration of the architecture of AlexNet we used. The kernel size of the three max pooling is 3. The last layer is the Softmax; the output of the network is the predicted label. And we choose the output of the layer FC6 as the CNN feature

was proposed by Abhishek Jaialtil of the University of Colorado Boulder, to train the model. It is worth mentioning that in the classification and recognition area, CART (classification and regression tree) obeys the Gini value. And we also train the extreme learning machine classifier, Treebagger classifier, and other knowledge-based classifiers.

4.2 Ensemble learning

In the ensemble learning part, some classifiers are chosen as the knowledge-based classifiers. In our experiments, we choose the Softmax classifier, SVM classifier, and Random Forest classifier as the knowledge-based classifiers. Softmax has a very wide application in machine learning; it is a lay of convolutional neural networks and always follows with fully connected layer to achieve the problem of classification. Softmax classifier is a simple and convenient algorithm, and it has a good performance. SVM is a common classifier, and it used to be a hot research filed. The core idea of the SVM is to structure the hyperplanes to separate the samples. And it is robust. SVM has a great advantage in solving the high-dimensional problem.

After inputting our test set to these knowledge-based classifiers, the predicted results could be obtained, which contains two kinds of data. The first one is the predicted label for every test sample, and the other one is the probability maps that represent the probabilities of each test sample that belongs to the different material categories. In our experiments, each test sample has 23 predicted probabilities that mean the probabilities belong to these 23 material categories. After the test set was imported into the CNN model and knowledge-based classifiers, three 5175*23 probability maps are generated. Then, the new final probability map is computed on the three probability maps with proposed methods. And we also do some experiments with the prior methods, including the stacking and voting.

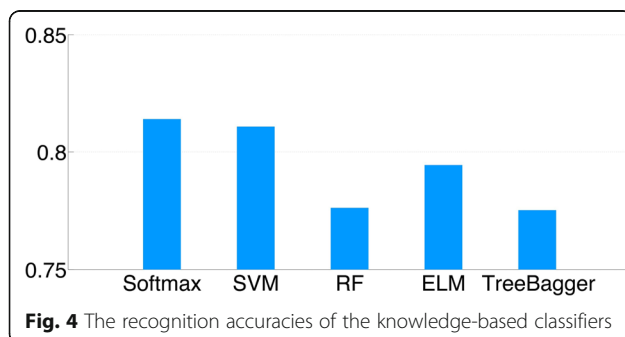


Fig. 4 The recognition accuracies of the knowledge-based classifiers

Table 2 Varying neuron number. Train EML model with different number of neurons

The number of neuron	Accuracy (%)
100	78.22
300	79.04
400	79.09
500	79.04
1000	78.59
2000	77.81
3000	77.45
4000	77.13

4.3 Material recognition accuracy

- The recognition accuracies of the knowledge-based classifiers

We train the knowledge-based classifiers, including the Softmax classifier and SVM classifier. The accuracies of these knowledge-based classifiers are shown in Fig. 4. And the recognition accuracies showed in this figure are the best performance on accuracy of these knowledge-based classifiers. These knowledge-based classifiers showed in this figure are Softmax, SVM, Random Forest, ELM, and Treebagger. The Softmax classifier has a better performance than others.

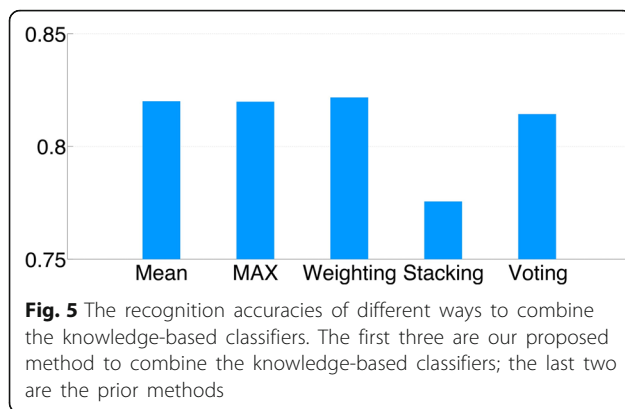
In the training procedure of the ELM classifier, different numbers of neuron had been set to get the state-of-the-art result. The recognition accuracies of the ELM classifier with different neuron numbers are showed in Table 2.

Similarly, we also train the Treebagger classifier with different decision tree numbers. The recognition accuracies are shown in Table 3.

From Table 2 and Table 3, we can find that as the neuron number or the decision tree number grows, the recognition accuracy will be improved firstly, but when it comes to a certain recognition accuracy, the recognition accuracy will be reduced. Meanwhile, the training time will increase. So, how to choose the

Table 3 Varying decision tree number. Train the Treebagger classifier with different number of decision tree

The number of tree	Accuracy (%)
200	73.49
500	74.72
700	74.90
1000	75.44
1500	75.08
2000	75.17



number of the neuron or the number of the decision tree is still a difficult topic.

- The recognition accuracies of the ensemble classifiers

In this part, we compare the recognition accuracies of our proposed three ways to combine these knowledge-based classifiers. And we also compare the recognition accuracies of our methods with the prior methods, including the stacking and voting. The recognition accuracies of the five methods are described in Fig. 5. From the figure, we could see that our methods to combine the knowledge-based classifiers have a better performance than the prior methods.

And the 23 categories recognition accuracies of the five ensemble learning methods are showed in Fig. 6. We can see that there is a large range among the 23 recognition accuracies.

The results show that our proposed methods obviously perform better than the prior methods in most number of the material categories. And these specific values of the 23 categories of the five ensemble classifiers are showed in Table 4.

- Ensemble learning methods with VGG16

In order to further prove the superiority of our proposed methods. We have carried experiments on the same train set and test set with VGG16. Firstly, the CNN features are extracted from the trained VGG16 model. Then, we train these knowledge-based classifiers and combine them with our proposed methods and the prior methods. The recognition accuracies of these knowledge-based classifiers are shown in Fig. 7.

These knowledge-based classifiers are the Softmax classifier, SVM classifier, and Random Forest classifier. We can see that the recognition accuracy of Softmax classifier is still the highest, but the recognition accuracy of Random Forest classifier improves obviously. After these knowledge-based classifiers are trained, we combine them with our proposed methods. And to prove the superiority of our proposed methods, we also carried experiments with the prior methods. The accuracies of these ensemble learning methods are shown in Fig. 8.

From Fig. 8, we can see that the recognition accuracies of our proposed methods are higher than the prior ensemble learning methods. And the 23 categories recognition accuracies of the three proposed ensemble learning methods are showed in Fig. 9. There also is a large range among the 23 recognition accuracies.

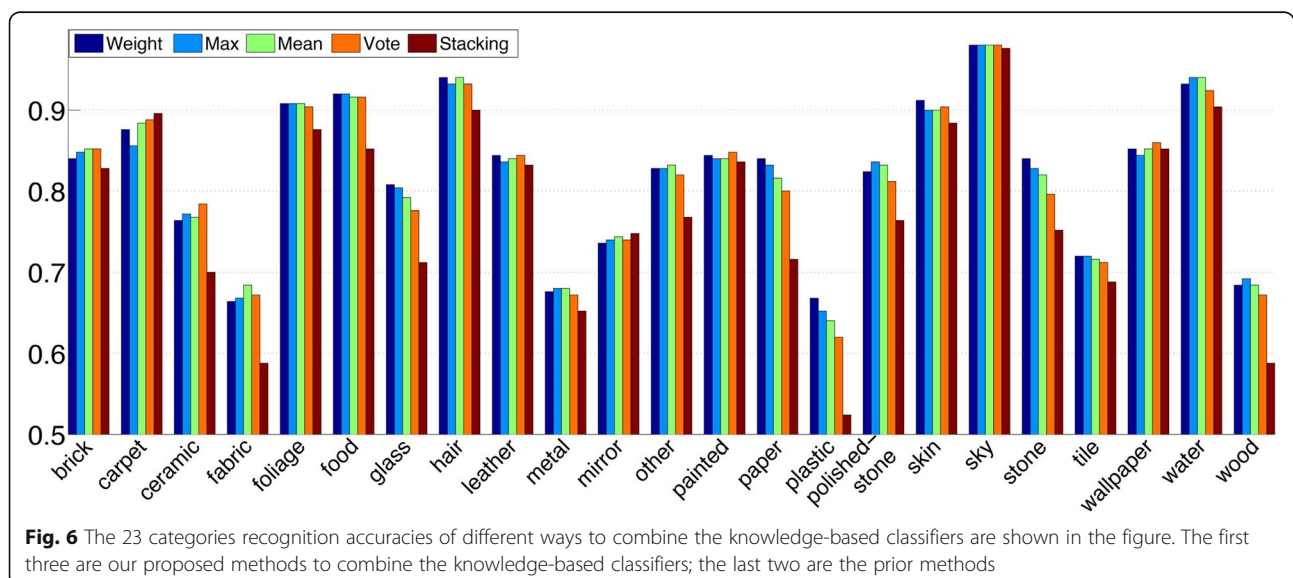


Table 4 The recognition accuracies of the 23 categories for these five methods to combine these knowledge-based classifiers

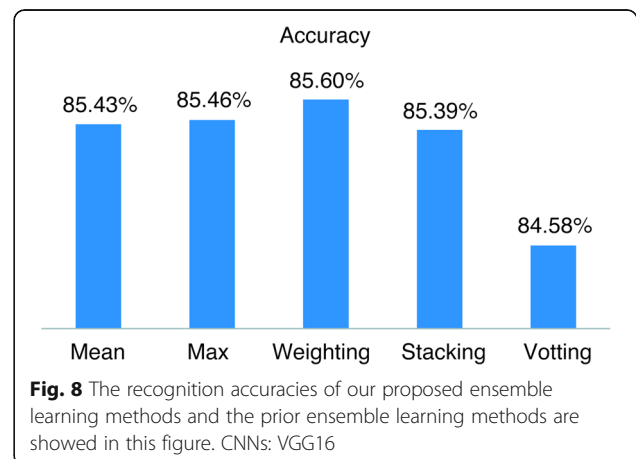
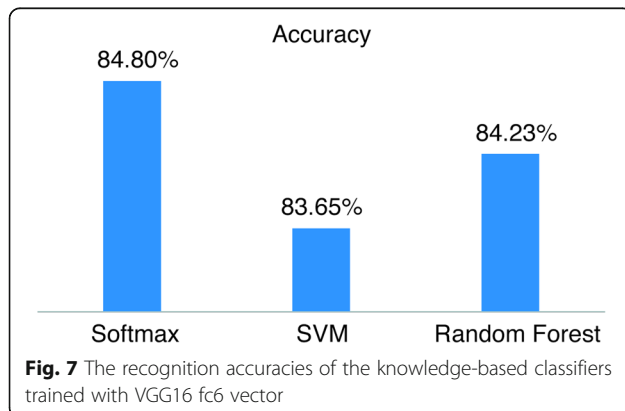
	Category							
	Brick	Carpet	Ceramic	Fabric	Foliage	Food	Glass	
Vote	0.852	0.888	0.784	0.672	0.904	0.916	0.776	
Stacking	0.828	0.896	0.700	0.588	0.876	0.852	0.712	
Mean	0.852	0.884	0.768	0.684	0.908	0.916	0.792	
Max	0.848	0.856	0.772	0.668	0.908	0.920	0.804	
Weight	0.840	0.876	0.764	0.664	0.908	0.920	0.808	
	Category							
	Hair	Leather	Metal	Mirror	Other	Painted	Paper	Plastic
Vote	0.932	0.844	0.672	0.740	0.820	0.848	0.800	0.620
Stacking	0.9	0.832	0.652	0.748	0.768	0.836	0.716	0.524
Mean	0.94	0.840	0.680	0.744	0.832	0.840	0.816	0.640
Max	0.932	0.836	0.680	0.740	0.828	0.840	0.832	0.652
Weight	0.94	0.844	0.676	0.736	0.828	0.844	0.840	0.668
	Category							
	Polished stone	Skin	Sky	Stone	Tile	Wallpaper	Water	Wood
Vote	0.812	0.904	0.980	0.796	0.712	0.860	0.924	0.672
Stacking	0.764	0.884	0.976	0.752	0.688	0.852	0.904	0.588
Mean	0.832	0.900	0.980	0.820	0.716	0.852	0.940	0.684
Max	0.836	0.900	0.980	0.828	0.720	0.844	0.940	0.692
Weight	0.824	0.912	0.980	0.840	0.720	0.852	0.932	0.684

- Our proposed weight learning method vs. prior method

This paper also proposes an arithmetic for learning the weights of these knowledge-based classifiers. In order to prove the superiority of our proposed method, we have conducted experiment with the prior weight learning method. The recognition accuracies of our proposed method and normalize weights arithmetic are shown in Table 5. From this table, we can see our proposed arithmetic is superior to the normalize weights arithmetic.

5 Conclusions

Material recognition is a long-standing and challenging problem. In this paper, we introduce the ensemble learning for material recognition with convolutional neural networks (CNNs). We combine these trained knowledge-based classifiers on the probability level, and we introduce a new arithmetic for weight learning. In our new method, ensemble learning-based convolutional neural networks (CNNs) is proposed to improve the accuracy of material



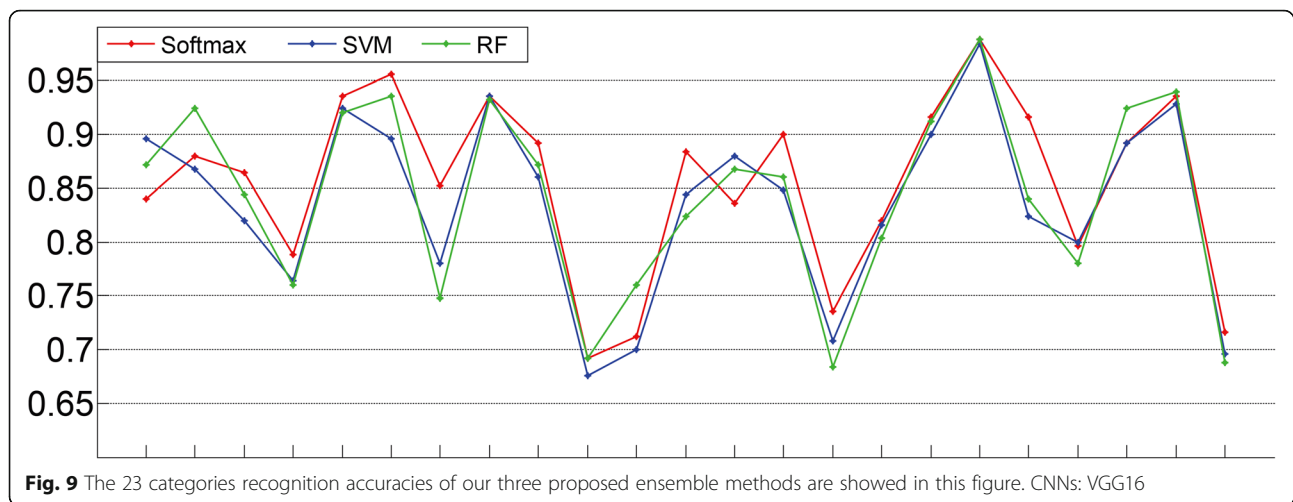


Table 5 Weight learning experiments. The recognition accuracies of our proposed method and the prior method

Method	Best performance (%)
Normalize weights	82.0
Our method	82.17

recognition. Then, a new weight learning method is applied to learn the weights of these knowledge-based classifiers. The great difference from the prior work is that we learn the ensemble classifier on probability level. The experimental results show that the ensemble classifier can achieve higher recognition accuracy than the knowledge-based classifiers, and our proposed ensemble learning methods are superior to the prior methods. Meanwhile, our proposed arithmetic for weight learning is superior to the other methods. Many future avenues of work remain. One of the future works is that we will try to propose a new deep feature learning method for the pixel-level material recognition task so that better recognition accuracy can be achieved.

Abbreviations

CNN: Convolutional neural network; FMD: Flicker Material Database; LBP: Local binary pattern; MINC: Materials in Context Database

Acknowledgements

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

Ethical approval and consent to participate

Approved.

Funding

This work was supported in part by a grant from the National Natural Science Foundation of China (No. 51505004, No. 61673052), a grant from the China Scholarship Council (No. 201708110041), a grant from the Fundamental Research Funds for the Central Universities (No.2018JBM017), and a grant from the CES-Kingfar Excellent Young Scholar Joint Research Funding.

Availability of data and materials

We can provide the data.

Authors' contributions

All authors take part in the discussion of the work described in this paper. The author YJ wrote the first version of the paper. The author WL did part of the experiments of the paper. PB and RZ revised the paper in different versions of the paper. All authors read and approved the final manuscript.

Authors' information

Peng Bian received the PhD degree in industrial design from Beijing Institute of Technology in 2012. He is currently an assistant professor in College of Mechanical and Material Engineering, North China University of Technology. In 2018, he visits University of North Carolina at Charlotte as a visiting scholar. His research interests include product design, computer aided design, and human-computer interaction E-mail: bianpeng@ncut.edu.cn.

Wanwan Li received the bachelor degree in computer science and technology from Yanbian University in 2016. She is now a master student in Beijing Jiaotong University. From 2012 to now, she has been the recipient of more than ten awards, including the National Scholarship. Her research interests include pattern recognition, material recognition and semantic segmentation. E-mail: 16125180@bjtu.edu.cn.

Yi Jin received the Ph.D. degree in Signal and Information Processing from the Institute of Information Science, Beijing Jiaotong University, Beijing, P.R. China, in 2010. She is currently an Associate Professor in the School of Computer Science and Information Technology, Beijing Jiaotong University. She has been a visiting scholar in School of Electrical and Electronic Engineering, Nanyang Technological University of Singapore (2013-2014). Her research interests include computer vision, pattern recognition, image processing and machine learning. E-mail: yjin@bjtu.edu.cn (*corresponding author).

Ruicong Zhi received the PhD degree in signal and information processing from Beijing Jiaotong University in 2010. From 2016-2017, she visited the University of South Florida as a visiting scholar. She visited the Royal Institute of Technology (KTH) in 2008 as a joint PhD. She is currently an associate professor in School of Computer and Communication Engineering, University of Science and Technology Beijing. She has published more than 50 papers, and has six patents. She has been the recipient of more than ten awards, including the National Excellent Doctoral Dissertation Award nomination. Her research interests include facial and behavior analysis, artificial intelligence, and pattern recognition. E-mail: zhirc@ustb.edu.cn

Consent for publication

Approved.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹College of Mechanical and Material Engineering, North China University of Technology, No.5 Jinyuanzhuang Road, Beijing 100144, China. ²School of Computer and Information Technology, Beijing Jiaotong University, No.3 Shangyuancun Road, Beijing 100044, China. ³Beijing Key Lab of Traffic Data Analysis and Mining, No.3 Shangyuancun Road, Beijing 100044, China. ⁴School of Computer and Communication Engineering, University of Science and Technology Beijing, No.30 Xueyuan Road, Beijing 100083, China. ⁵Beijing Key Laboratory of Knowledge Engineering for Materials Science, No.30 Xueyuan Road, Beijing 100083, China.

Received: 11 April 2018 Accepted: 29 June 2018

Published online: 30 July 2018

References

1. C Szegedy, W Liu, Y Jia, P Sermanet, S Reed, D Anguelov, D Erhan, V Vanhoucke, A Rabinovich, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Going deeper with convolutions (2015)
2. K Simonyan, A Zisserman, *Very deep convolutional networks for large-scale image recognition*. *arXiv preprint arXiv:1409* (2014), p. 1556
3. C Farabet, C Couprie, L Najman, Y LeCun, Learning hierarchical features for scene labeling. *IEEE Trans Pattern Anal Mach Intell (PA- MI)* **35**(8), 1915–1929 (2013)
4. J Long, E Shelhamer, T Darrell, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Fully convolutional networks for semantic segmentation (2015)
5. R Girshick, J Donahue, T Darrell, J Malik, in *Computer Vision and Pattern Recognition (CVPR)*. Rich feature hierarchies for accurate object detection and semantic segmentation (2014)
6. P Sermanet, D Eigen, X Zhang, M Mathieu, R Fergus, YLC Overfeat, in *International Conference on Learning Representations (ICLR)*. CBLs. Integrated recognition, localization and detection using convolutional networks (2014)
7. KJ Dana, B Van Ginneken, SK Nayar, JJ Koenderink, Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics (TOG)* **18**(1), 1–34 (1999)
8. L Sharan, R Rosenholtz, E Adelson, Material perception: what can you see in a brief glance? *J. Vis.* **9**(8), 784–784 (2009)
9. S Bell, P Upchurch, N Snavely, K Bala, Opensurfaces: a richly annotated catalog of surface appearance. *ACM Transactions on Graphics (TOG)* **32**(4), 111 (2013)
10. S Bell, P Upchurch, N Snavely, K Bala, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Material recognition in the wild with the materials in context database (2015)
11. Bell S, Upchurch P, Snavely N, et al. Material recognition in the wild with the materials in context database[C]. *IEEE conference on computer vision and pattern recognition IEEE computer Society*(2015), 3479–3487.
12. TC Wang, JY Zhu, E Hiroaki, et al, A 4D light-field dataset and CNN architectures for material recognition [C]. *European conference on computer vision, Lecture Notes in Computer Science*, vol 9907. (Springer Cham, 2016), p. 121–138.
13. J Xue, H Zhang, K Dana, et al, Differential angular imaging for material recognition [J], *eprint arXiv:1612.02372*. (2016), p. 6940–6949.
14. OG Cula, KJ Dana, 3D texture recognition using bidirectional feature histograms. *Int. J. Comput. Vis.* **59**(1), 33–60 (2004)
15. C Liu, J Gu, Discriminative illumination: per-pixel classification of raw materials based on optimal projections of spectral BRDF. *IEEE Trans Pattern Anal Mach Intell* **36**(1), 86–98 (2014)
16. S Lombardi, K Nishino, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Single image multimaterial estimation (2012)
17. H Zhang, K Dana, K Nishino, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Reflectance hashing for material recognition (2015)
18. S Lombardi, K Nishino, Reflectance and natural illumination from a single image. In *European Conference on Computer Vision volume VI*, 582528–541595 (2012)
19. S Lombardi, K Nishino, Single image multimaterial estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 238–245 (2012)
20. G Oxholm, K Nishino, Shape and reflectance from natural illumination. *European Conference on Computer Vision*. Part of the Lecture Notes in Computer Science book series. LNCS. **7572**, 528528–541541 (2012)
21. G Oxholm, K Nishino, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Multiview shape and reflectance from natural illumination (2014)
22. Ting-ChunWang, Jun-YanZhu, EbiHiroaki, ManmohanChandraker, AlexeiAEfros, and Ravi Ramamoorthi. A 4D light-field dataset and CNN architectures for material recognition. In *European Conference on Computer Vision (ECCV)*. 121–138. (2016).
23. C.Liu, L.Sharan, E. H.Adelson, and R. Rosenholtz.Exploringfeaturesinabayesian framework for material recognition. In *CVPR*, pages 239 - 246. IEEE, (2010).
24. D. Hu, L. Bo, and X. Ren. Toward robust material recognition for everyday objects. In *BMVC*, 1 - 11. Citeseer, (2011).
25. X Qi, R Xiao, J Guo, L Zhang, in *ECCV*. Pairwise rotation invariant co-occurrence local binary pattern (Springer, 2012), pp. 158–171
26. G Schwartz, K Nishino, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*. Visual material traits: recognizing per-pixel material context (2013)
27. M Cimpoi, S Maji, A Vedaldi, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Deep filter banks for texture recognition and segmentation (2015)
28. Xu H, Han Z, Feng S, et al. Foreign object debris material recognition based on convolutional neural networks [J]. *Eurasip Journal on Image and Video Processing*, (1):21 (2018).
29. Younis K S, Ayyad W, Al-Ajlony A. Embedded system implementation for material recognition using deep learning [C]. *IEEE Jordan conference on applied electrical engineering and computing technologies*. IEEE, 1–6 (2017).
30. Z-H Zhou, *Ensemble Learning* (Springer, In, 2015)
31. T Hastie, R Tibshirani, J Friedman, *The Elements of Statistical Learning*, 2nd edn. (Springer, New York, NY, USA, 2009)
32. D Opitz, R Maclin, Popular ensemble methods: an empirical study. *J. Artif. Intell. Res.* **11**, 169–198 (1999)
33. M Pandey, S Taruna, A comparative study of ensemble methods for students performance modeling. *Int. J. Comput. Appl.* **103**(8), 0975–8887 (2014)
34. R.Polikar,Ensemble learning, in: *Ensemble Machine Learning*.Springer.1–34.(2012).
35. C Cortes, V Vapnik, Support-vector network. *Mach. Learn.* **20**, 273–297 (1995)
36. L Breiman, Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
37. GB Huang, What are extreme learning machines? Filling the gap between Frank Rosenblatt's dream and John von Neumann's puzzle. *Cogn. Comput.* **7**(3), 263–278 (2015)
38. A Krizhevsky, I Sutskever, GE Hinton, in *Advances in Neural Information Processing Systems*. Imagenet classification with deep convolutional neural networks (2012), pp. 1097–1105
39. CC Chang, CJ Lin, LIBSVM: a library for support vector machines. **2**(3), 1–27 (2001)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)