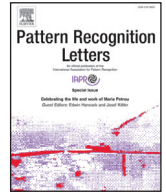




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Compressed feature vector-based effective object recognition model in detection of COVID-19

Chao Chen^a, Jinhong Mao^{b,*}, Xinzhi Liu^b, Yi Tan^a, Ghada M Abaido^c, Hamdy Alsayed^d

^a College of Communication Engineering, Chengdu University of Information Technology, Chengdu 610225, China

^b Air Force Early Warning Academy, Wuhan 430019, China

^c Department of Media and Communication Studies, Faculty of Communication, Arts and Sciences, Canadian University Dubai, Dubai, United Arab Emirates

^d Applied Science University, Al Eker, Kingdom of Bahrain

ARTICLE INFO

Article history:

Received 28 May 2021

Revised 1 December 2021

Accepted 22 December 2021

Available online 25 December 2021

Edited by: Maria De Marsico

ABSTRACT

To better understand the structure of the COVID-19, and to improve the recognition speed, an effective recognition model based on compressed feature vector is proposed.

Object recognition plays an important role in computer vision. To improve the recognition accuracy, most recent approaches always adopt a set of complicated hand-craft feature vectors and build the complex classifiers. Although such approaches achieve the favourable performance on recognition accuracy, they are inefficient. To raise the recognition speed without decreasing the accuracy loss, this paper proposed an efficient recognition model trained with a kind of compressed feature vectors. Firstly, we propose a kind of compressed feature vector based on the theory of compressive sensing. A sparse matrix is adopted to compress feature vector from very high dimensions to very low dimensions, which reduces the computation complexity and saves enough information for model training and predicting. Moreover, to improve the inference efficiency during the classification stage, an efficient recognition model is built by a novel optimization approach, which reduces the support vectors of kernel-support vector machine (kernel SVM). The SVM model is established with whether the subject is infected with the COVID-19 as the dependent variable, and the age, gender, nationality, and other factors as independent variables. The proposed approach iteratively builds a compact set of the support vectors from the original kernel SVM, and then the new generated model achieves approximate recognition accuracy with the original kernel SVM. Additionally, with the reduction of support vectors, the recognition time of new generated is greatly improved. Finally, the COVID-19 patients have specific epidemiological characteristics, and the SVM recognition model has strong fitting ability. From the extensive experimental results conducted on two datasets, the proposed object recognition model achieves favourable performance not only on recognition accuracy but also on recognition speed.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Since the outbreak of the COVID-19, the number of infections worldwide is increasing every day. The main symptoms are respiratory cough, fever, shortness of breath, difficulty breathing, etc. Severe cases have pneumonia, acute respiratory syndrome, kidney failure, and even die. There is no specific treatment for it. The manifestations of the COVID-19 patients are still not fully understood, and the virus is mutating, making the research of COVID-19 urgent.

Object recognition plays an important role in the field of computer vision and various multimedia applications. The task of object recognition is to determine whether a certain type of object

is existed in a given image. Due to the problem of blur, deformation, partial occlusion, illumination change, clutter background, etc.; object recognition becomes a challenge task. To address such kind of problems, the high dimensional hand-craft features and the classification model with high computation complexity are adopted to ensure the recognition accuracy. However, the high dimensional hand-craft features and complicated classification models make the prediction inefficient. In practice use, we demand not only on recognition accuracy but also on recognition speed. Because that the fast and accurate object recognition can provide effective support for other computer vision tasks, such as object detection and tracking. It provides powerful basis for the system's decision-making. Moreover, with the growth of the image data and the improvement of the demand of system's intelligence,

* Corresponding author.

E-mail address: jhhmao@hotmail.com (J. Mao).

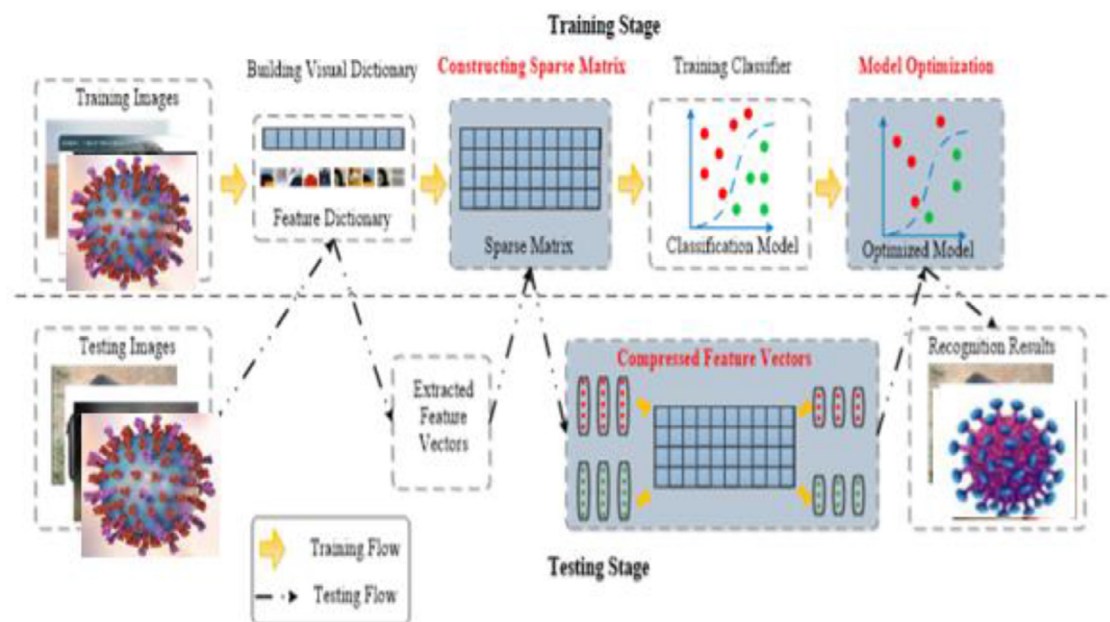


Fig. 1. Workflow of the proposed method.

developing the high-precision, real-time and well-adapted approaches have become a trend.

Traditional virus classification tests take into account the morphology, serology, host range, protein, and physical and chemical characteristics, such as sensitivity to organic solvents, cell culture, structure, and molecular weight. With the development of computer algorithms, intelligent systems are widely used in identification of chemical composition and complex biological macromolecules because of their sensitivity and accuracy. Some researchers have used relevant algorithms to conduct systematic cluster analysis to identify viruses. Similar strains clustering together has important implication significance for identifying virus strains.

This paper deals with the practical problems of object recognition based on feature compression and classification model optimization. This model assists in the identification and classification of the COVID-19. Firstly, from the aspect of feature compression, the traditional approaches always involve a large number of matrix decomposition operations, such as Principal Component Analysis (PCA) [1] and Singular Value Decomposition (SVD) [2], which are inefficient. To improve the compression efficiency, a feature compression algorithm based on compressive sensing is proposed. Though this algorithm, a sparse random matrix is adopted to map the high-dimensionality feature into the low-dimensionality feature space, and the mapping just involves the operation of matrix multiplication. Secondly, although the commonly used classification model—Kernel SVM [3] as good generalization capability, the prediction cost will increase with the number of support vectors increases. Therefore, a support vector reduction algorithm is proposed to optimize the classification model, which reconstructs a simplified subset of original support vectors through the way of cyclic iterations. The optimized classification model achieves the similar generalization capability with the original model but spends less time on predicting.

Fig. 1 shows the workflow of the proposed method, which is divided into two training stage and the testing stage. High-dimensional feature vectors are extracted from the COVID-19 image first. In the training stage, the high-dimensional feature vectors are extracted from the image firstly, and then a very sparse matrix is constructed to map the high-dimensional feature vec-

tors into the low-dimensional domain. Lastly, a kernel SVM classifier is trained by the compressed features, and the kernel SVM is optimized by the support vector reduction method. In the testing phase, we adopt the same method to compress the features and put the low-dimensional features into the optimized kernel SVM for prediction.

The rest of this paper is organized as follows: In Section 2, we introduce the relevant studies about the related works. In Section 3, the basic theories and analysis regarding to feature compression and model optimization are introduced. In Section 4, experimental results and analysis are presented. In Section 5, we conclude this paper and propose future work.

2. Related works

After the virus infects the body, the body recognizes the virus, initiates an innate immune response, produces an antiviral response, secretes a large number of inflammatory cytokines, and mediates the occurrence of inflammation. The body resists viral infections by clearing the infected virus, like type I interferon can induce the expression of “Mx” gene, hinders the initial transcription process of PB2 polymerase. The joint feature and compressed dictionary learning have few applications in detection of viruses. Some scholars have applied this learning algorithm to classify cars and found, that the learning algorithm can improve the accuracy of car model recognition.

In this section, we briefly review the recent studies [4–13] in the literature on object recognition.

Recently, object recognition approaches focus on extracting discriminative features, which can be divided into two kind of frameworks: deep learning [4] and bag-of-words [5]. Although the deep learning approaches achieved excellent performance in object recognition, these approaches always involved higher computation resources and training set. The classical deep model such as AlexNet, GoogleNet and ResNet were designed for large scale image classification task, which adopt the ImageNet that is a very large image dataset contained about 14 million images for training and testing[6]. However, the deep model is not suitable for the dataset with less samples. Moreover, the deep models always trained on multi-GPUs, which is time consuming for the platform

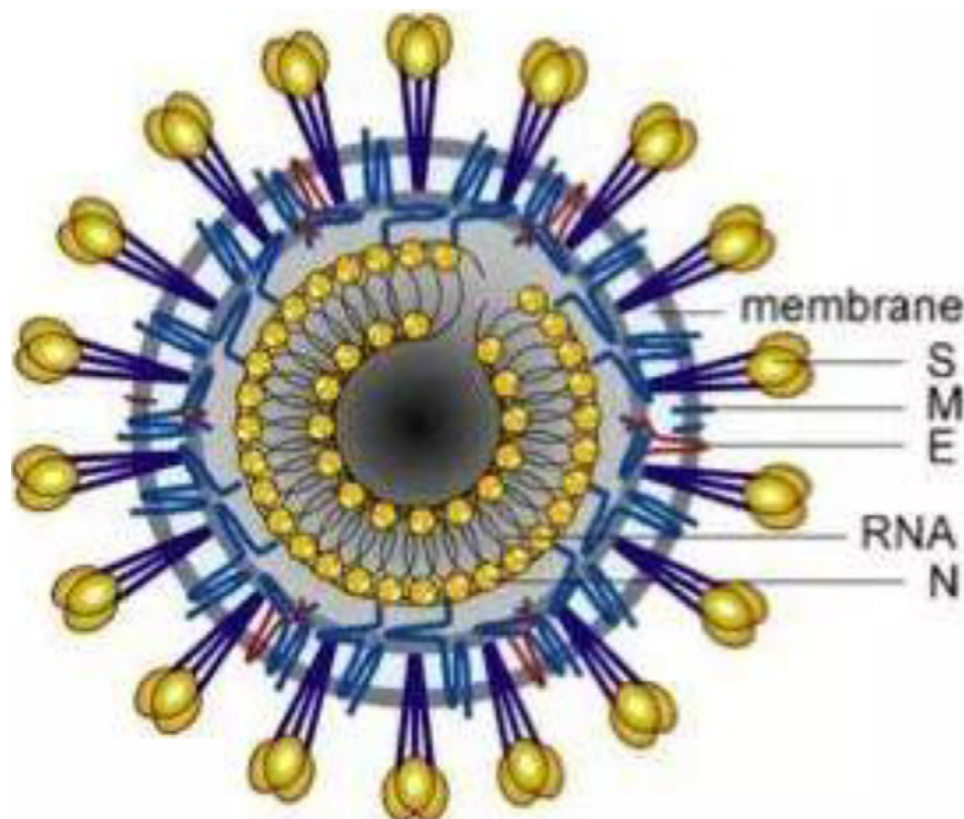


Fig. 2. Structure of coronary viruses.

with on CPU architecture[7]. Therefore, for the situation of less samples and less computing resources, the bag-of-words framework is well adoptive.

For the bag-of-words framework, the image features always extracted with dense sampling scheme, which adopted a fixed size and scale to extract a large number of local descriptions from the image. The extracted local feature descriptions have higher redundancy, which can be encoded into a feature vector, then, the feature vector is fed into classifier for classification. For example, Xu et al. [8] designed the HOG (Histogram of oriented gradient) feature and trained a linear SVM, which achieved great success in vehicle recognition. Ma et al. [9] adopted GMM (Gaussian mixture model) and LCC (Local coordinate coding) to reduce the redundant information of the densely sampled SIFT (Scale-invariant feature transform) descriptions from the image and trained a linear classifier for object recognition. Zhuang et al. [10] fused multi kind of local feature, vector quantization coding and adopted spatial pyramid match to generate the high dimensional feature vectors, then trained them by a kernel SVM for recognition. Due to the rich and hierarchical information of the images, this approach performed well on object recognition. Clearly, with the richer information of feature vector, the classifier obtained better performance. Therefore, in order to improve the recognition performance, multi kind of feature were fused, but this will cause the computation cost. Jorge et al. Dr. Lowery from the University of Ards Northern Ireland developed a "DNA fingerprinting" system. It can detect a variety of viruses including smallpox within 15 years.

[11] proposed the approach that adopted the fisher vector to encode the feature vector and then compressed the high dimensional feature vector by product quantization algorithm. With the compressed feature vectors, the trained classifier performed well on object recognition. Feature compression aims to save as much information of original feature vectors as possible, which reduce

the model training time and the storage space. The PCA and SVD are two main feature compression methods in the field of object recognition. The PCA is used to map the feature vector from the high dimensional space to the low dimensional space by orthogonal linear projection[12]. While the SVD adopts the singular value decomposition of matrix to extract the key information of original feature vectors. However, the matrix decomposition usually involves higher computation complexity.

Additionally, classification model is another important in object recognition framework. The SVM is the classical classification model and it was adopted in a lot of works, such as the studies [8–10]. It shows well generalization capability in the situation of less samples, non-linear and high dimensional problems. However, the SVM still has some limitations, such as the prediction speed will decrease with the number of the support vectors increases. Koibayashi et al. [13] think that the support vector set can be reduced by finding a simplified support vector set in feature space. Geebelen et al. [14] proposed a method that selected a typical sample set from the original training set and adopted the reduced sample set to train the prediction model.

Therefore, this work deal with the object recognition problems from two aspects: 1) designing an efficient and compressed feature to reduce the computational complexity and speed up the prediction, 2) developing a model optimization method to improve the recognition speed without the accuracy loss.

3. Materials and methods

3.1. Characteristics of coronary viruses

Corona viruses have no cell structure and can colonize living cells. They rely on living cells to synthesize protein. The main genetic material of the coronaviruses is single-stranded positive-

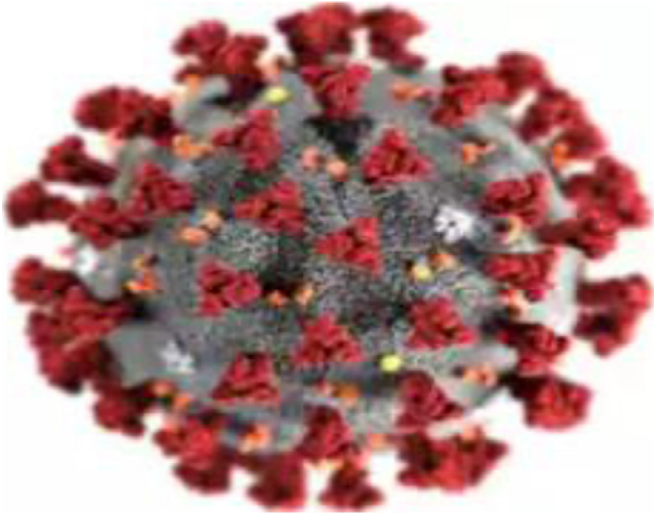


Fig. 3. The COVID-19.

stranded RNA. Its basic unit is ribonucleotides, which can be stained with Piro Red dye. They bind to ACE2 on the cell membrane to invade the cell, indicating that the cell membrane is responsible for information exchange. The offspring of the virus are discharged out of the cell through the vesicles, indicating that the biofilm has certain fluidity. Fig. 2 shows the structure of the coronavirus. S is the spike protein, E is the mantle protein, M is the membrane protein, and N is the ribonucleoprotein. Fig. 3 is COVID-19.

3.2. Feature extraction

50 cases infected with COVID-19 admitted to XX hospital were selected as the research subjects to investigate the epidemiology under their consent. The SVM model was established with the COVID-19 infection as the dependent variable, and age, gender, education level, and close contact history as the independent variables.

The SIFT descriptor and the dense sampling scheme is adopted in our work. Firstly, the image of COVID-19 is segmented into local patch with 16×16 pixels; each local patch is divided into 16 grids with the size of 4×4 pixels. Secondly, the gradient magnitude $g(x, y)$ and orientation (x, y) of each position (x, y) in the grid is calculated by Eq. (1), then the orientation of each position (x, y) is assigned to a specific bin and forming a histogram (which is ranged from 0 to 360° and divided into 8 bins, each of which is cover 45°). Thirdly, all of the orientation histograms of each local patch (with 4×4 grids) is aggregated and is formed into a 128D (D is short for dimension) SIFT descriptor; the detail is shown in Fig. 2. The SIFT descriptor of next local patch is calculated with spacing of 8 pixels until whole image is traversed.

$$g(x, y) = \sqrt{((I(x+1, y) - I(x-1, y)))^2 + (I(x, y+1) - I(x, y-1))^2}$$

$$\alpha(x, y) = \tan^{-1}((I(x, y+1) - I(x, y-1)) / (I(x+1, y) - I(x-1, y))) \quad (1)$$

3.3. Feature coding

The extracted 128D local features contained a lot of redundancy and are very difficult to model the semantic features. Therefore, though the bag-of-words framework, the quantified local feature descriptors can be seen as “visual word”, and the image content is expressed by the distribution of the “visual words” in the image. In our approach, we employed 200 images and extracted their local feature descriptors according to the method in Section 3.1,

then the K-Means [15] algorithm is adopted to calculate the cluster centers. We set the hyper parameter K to 1024, therefore, 1024 “visual words” were selected to build a “visual dictionary”. Each image can be encoded by the “visual dictionary” with the size of 1024D. This procedure is described in Fig. 3.

3.4. Feature compression

From compressive sensing theory [16,17], it is known that if dimensionality of feature space is extremely high, these features can be randomly projected to a low-dimensional feature space that includes enough information. Given a vector v in the low dimensionality space, $v \in \mathbb{R}^n$, a vector u in the high dimensionality feature space, $u \in \mathbb{R}^m$, and the mapping by a random matrix P is defined as Eq. (2): (Figs. 4 and 5)

$$v = P * u \quad (2)$$

where $n \ll m$.

The projection v is similar to a compressive measurement in the compressive sensing encoding stage. From the JL (Johnson-Lindenstrauss) lemma, given one signal is linear combination of only K basis [18], that signal is possible to reconstruct from a small number of random measurements.

Lemma1. (Johnson-Lindenstrauss lemma): Let D be a finite collection of d points in \mathbb{R}^m . Given $0 < \epsilon < 1$ and $\beta > 0$, let n be a positive integer such that Eq. (3):

$$n \geq \left(\frac{4 + 2\beta}{\epsilon^2/2 + \epsilon^3/3} \right) * \ln(d) \quad (3)$$

Then, for any set D of d points in \mathbb{R}^m , there a mapping $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ exists, such that for every pairs $u, v \in D$,

$$(1 - \epsilon) \|u - v\|_2^2 \leq \|f(u) - f(v)\|_2^2 \leq (1 + \epsilon) \|u - v\|_2^2 \quad (4)$$

Let f be a random matrix $P, P \in \mathbb{R}^{n \times m}$, and $P(i, j) = p_{ij}$ Eqs. (5) and (6), where,

$$p_{ij} = \begin{cases} 1, & \text{with probability } 1/2 \\ -1, & \text{with probability } 1/2 \end{cases} \quad (5)$$

Or

$$p_{ij} = \begin{cases} \sqrt{3}, & \text{with probability } 1/6 \\ 0, & \text{with probability } 2/3 \\ -\sqrt{3}, & \text{with probability } 1/6 \end{cases} \quad (6)$$

Therefore, for any two vectors $u_1, u_2 \in D$, the Eq. (4) will be satisfied with probability which exceeding $1 - d^{-\beta}$.

Liu et al. [19] adopted random Gaussian matrix and showed that sparse random measurement matrix obtained favorable results in texture classification. Actually, a very sparse measurement matrix with sparse elements can be defined as Eq. (7):

$$p_{ij} = \sqrt{x} * \begin{cases} 1, & \text{with probability } 1/2x \\ 0, & \text{with probability } 1 - 1/x \\ -1, & \text{with probability } 1/2x \end{cases} \quad (7)$$

Chen et al. [20] proved that very sparse matrix can obtain the similar results with random gaussian matrix, and x can be defined as $x = O(m)$, $x = \sqrt{m \log m}$ or even $x = m / \log_{10}^m$. In our work, we define $x = m / \log_{10}^m$ for building the very sparse measurement matrix.

3.5. Model optimization

When the optimal classification hyperplane of kernel SVM is determined, the normal vector ψ of its eigenspace (in the Eq. (8)) can be represented as a linear combination of all support vectors. Model optimization approach aims to find a simplified support vector set $\{c_i\}, 1 \leq i \leq m$, to replace the original support vector

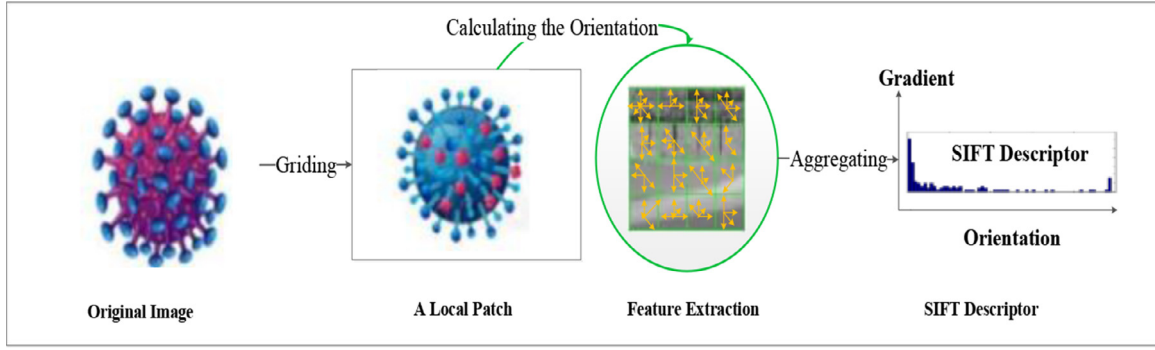


Fig. 4. The procedure of the feature extraction.

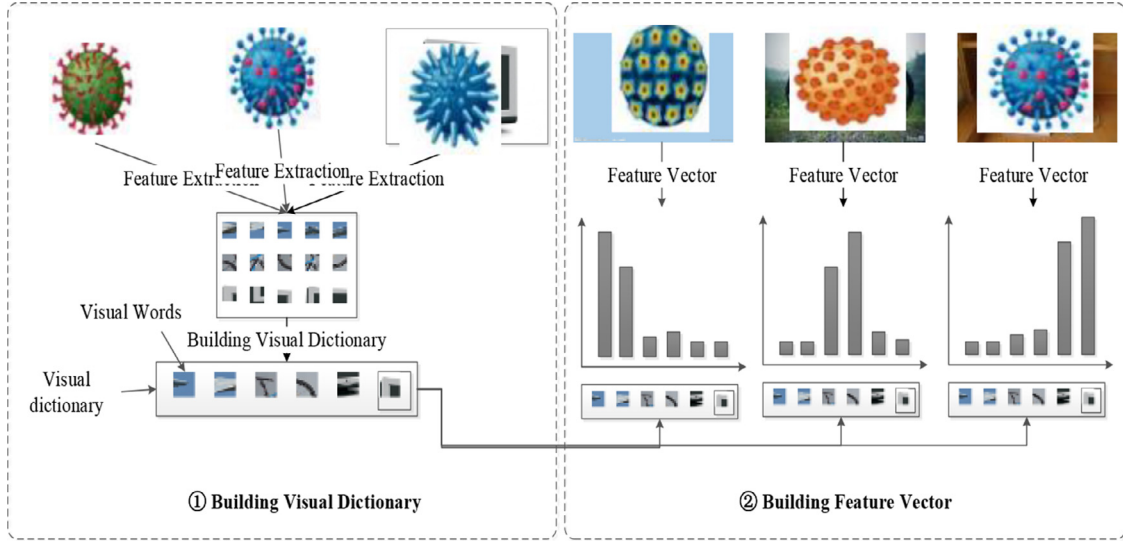


Fig. 5. The framework of bag-of-words.

set $\{s_i\}$, $1 \leq i \leq l$ of the kernel SVM model, where $m < l$. The normal vector of the simplified support set is defined as Eq. (9).

$$\Psi = \sum_{i=1}^l \alpha_i \Phi(\mathbf{c}_i), \alpha_i \in \mathbb{R} \quad (8)$$

$$\Psi' = \sum_{i=1}^l \beta_i \Phi(\mathbf{s}_i), \beta_i \in \mathbb{R} \quad (9)$$

Then, we define ε (in Eq. (10)) to be the square error of Ψ and Ψ' .

$$\begin{aligned} \varepsilon &= \|\Psi - \Psi'\|^2 = \left\| \sum_{i=1}^l \alpha_i \Phi(\mathbf{c}_i) - \sum_{i=1}^m \beta_i \Phi(\mathbf{s}_i) \right\|^2 \\ &= \left(\sum_{i=1}^l \alpha_i \Phi(\mathbf{c}_i) - \sum_{i=1}^m \beta_i \Phi(\mathbf{s}_i) \right)^T \\ &\quad \cdot \left(\sum_{i=1}^l \alpha_i \Phi(\mathbf{c}_i) - \sum_{i=1}^m \beta_i \Phi(\mathbf{s}_i) \right) \\ &= \sum_{i,j=1}^l \alpha_i \alpha_j k(\mathbf{c}_i, \mathbf{c}_j) + \sum_{i,j=1}^m \beta_i \beta_j k(\mathbf{s}_i, \mathbf{s}_j) \\ &\quad - 2 \sum_{i=1}^l \sum_{j=1}^m \alpha_i \beta_j k(\mathbf{c}_i, \mathbf{s}_j) \end{aligned} \quad (10)$$

Take the derivative of β in Eq. (10) and the result is shown in equation (11) and (12)

$$\frac{\partial}{\partial \beta_k} \|\Psi - \Psi'\|^2 = -2 \Phi(\mathbf{s}_k) \left(\Psi - \sum_{i=1}^m \beta_i \Phi(\mathbf{s}_i) \right) \quad (11) \quad \beta = (K^s)^{-1} K^{sc} \alpha \quad (12)$$

Where, $K_{i,j}^s = k(\Phi(\mathbf{s}_i), \Phi(\mathbf{s}_j))$ and $K_{i,j}^{sc} = k(\Phi(\mathbf{s}_i), \Phi(\mathbf{c}_j))$. Moreover, in order to solve the simplified support vector set, we follow the Eq. (13)

$$\mathbf{s} = \arg \min_{\mathbf{s}} \|\Phi(\mathbf{s}) - \Psi\|^2 \quad (13)$$

Take the derivative of sin Eq. (13), the result in Eq. (14)

$$\mathbf{s} = \frac{\sum_{i=1}^l \alpha_i k'(\|\mathbf{c}_i - \mathbf{s}\|^2) \mathbf{c}_i}{\sum_{i=1}^l \alpha_i k'(\|\mathbf{c}_i - \mathbf{s}\|^2)} \quad (14)$$

In our approach, we define $k(\mathbf{c}_i, \mathbf{c}_j) = \exp(-\frac{\|\mathbf{c}_i - \mathbf{c}_j\|^2}{2\sigma^2})$ (the Gaussian kernel), the Eq. (14) can be define as Eq. (15)

$$\mathbf{s} = \frac{\sum_{i=1}^l \alpha_i \exp(-\frac{\|\mathbf{s} - \mathbf{c}_i\|^2}{2\sigma^2}) \mathbf{c}_i}{\sum_{i=1}^l \alpha_i \exp(-\frac{\|\mathbf{s} - \mathbf{c}_i\|^2}{2\sigma^2})} \quad (15)$$

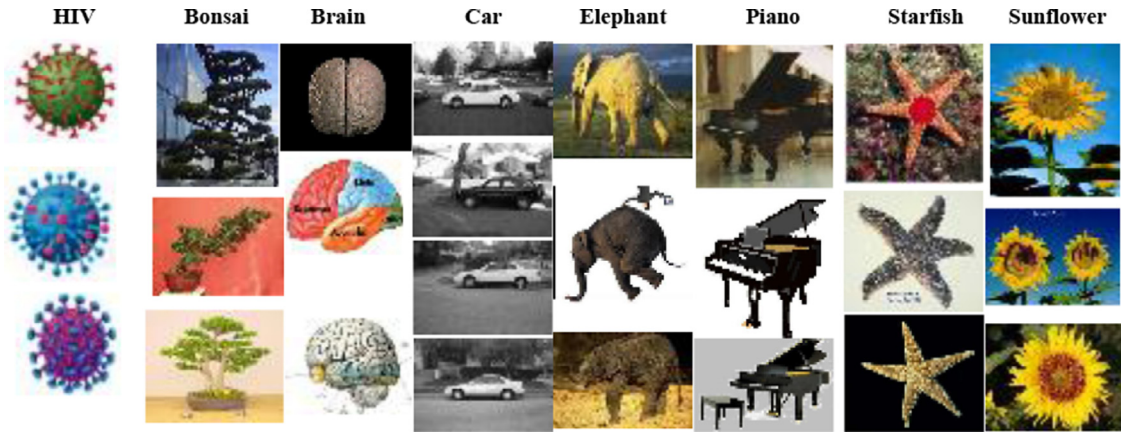


Fig. 6. Some examples of each categories.

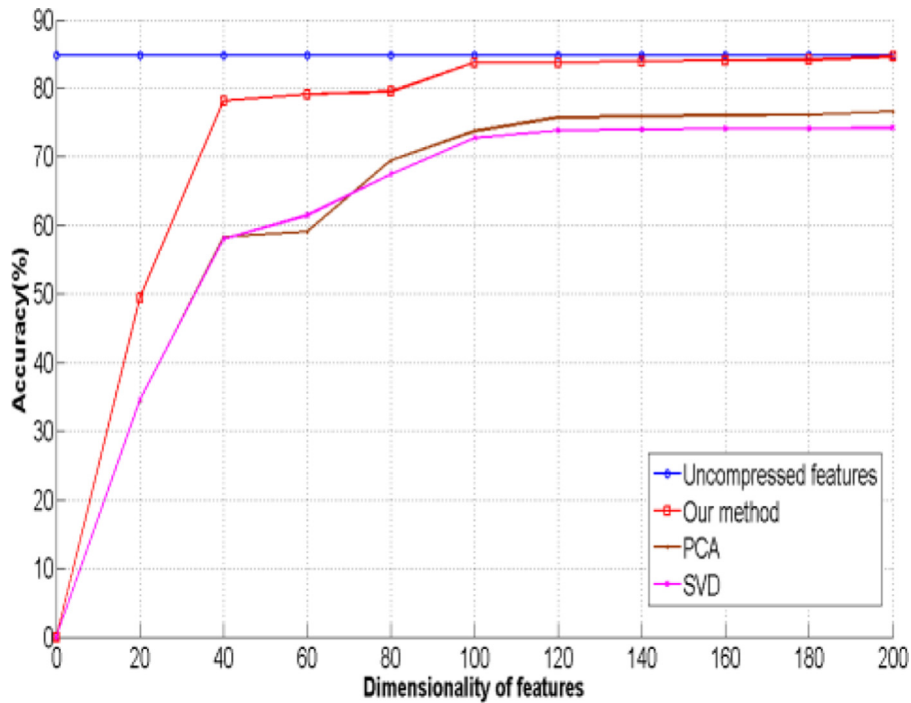


Fig. 7. The compression performance of different methods.

According to the Eq. (12) and (14), we can iteratively solve the β_m and s_m (shown in Eq. (16)), where,

$$\Psi_{m+1} := \Psi_m - \beta_m \Phi(s_m) \tag{16}$$

The condition for the iteration to terminate is that $\|\Psi_{m+1}\|$ less than a threshold or m reaches the specified number of support vectors.

4. The experimental results

In order to validate the performance of our approach, we tested it on Caltech 101 dataset. In the experiment, all approaches were programmed in Matlab 2014a and were run on a PC with an CPU (Intel Core i5 and 2.5 GHz) and 12 GB memory.

The Caltech 101 dataset is a popular dataset in the field of computer vision. Each category of this includes from 31 to 1100 images. These images are medium resolution with the size of 500×800 . In our experiments, about 775 images of 8 sub-categories were selected for training and testing. Fig. 6 and Table 1 respectively shows the distribution and some examples of each sub-categories.

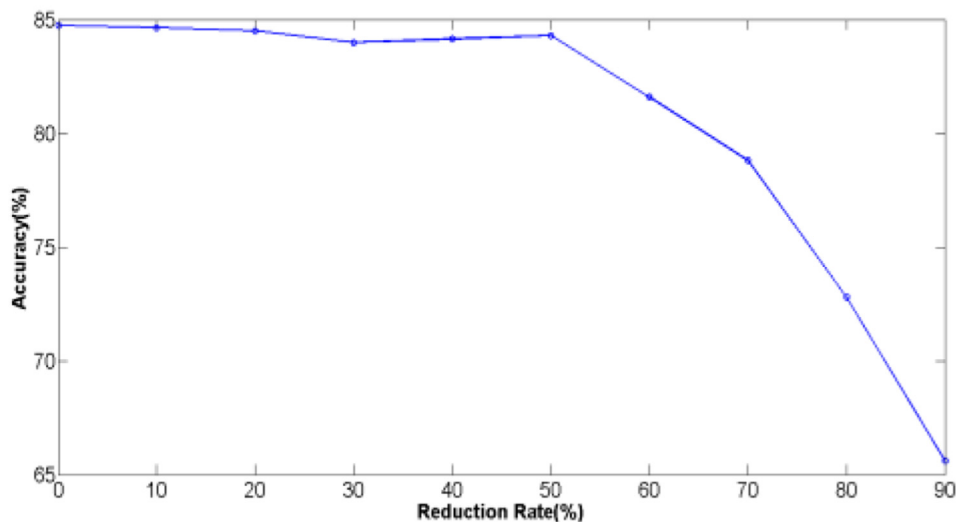
Table 1
Distribution of 8 sub-categories.

Categories	Image size(pixels)	Number
Butterfly	About 300×200	91
Bonsai	About 300×200	128
Brain	About 300×250	98
Car	About 300×200	123
Elephant	About 300×250	64
Piano	About 300×280	99
Starfish	About 300×250	86
Sunflower	About 200×300	86

(1) Recognition accuracy

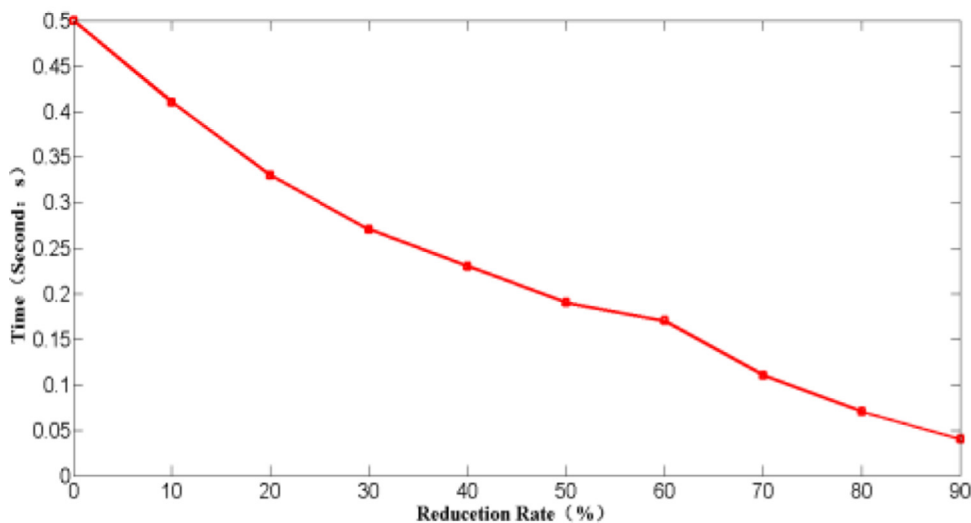
In the training stage, firstly, we randomly selected 10 images from each sub-categories and extracted the SIFT descriptors with the method mentioned in Section 3.1, and then the K-Means algorithm (K was set to 1024) was adopted to build a “visual dictionary” with the size of 1024 (described in Section 3.2). The rest images were randomly divided into 1:1 and adopted for model train-

Relationship between reduction rate and accuracy



(a)

Relationship between reduction rate and speed



(b)

Fig. 8. Relationship between reduction rate and accuracy, speed.

ing and testing. Secondly, the feature vectors were compressed to 100D by a very sparse matrix of size 100×1024 (which was built based on the compressive sensing method mentioned in Section 3.3). The compressed feature vectors were fed into the Kernel SVM for model training. Finally, the trained model was optimized by the method in Section 3.4. About 50% of support vectors were reduced in the experiments.

In the testing stage, we firstly extracted the feature vector from each testing image; secondly, the feature vectors were adopted the same sparse matrix (in training stage) for feature compression. Thirdly, the compressed features were fed into the optimized model for prediction. We compared the proposed method

with other two methods as follows, which employed the uncompressed feature vectors and unoptimized classification model. The tests were conducted 5 times and the comparison results show in Table 2. extracting SIFT descriptors and building 1024D feature vector with the same “visual dictionary” of our method, then adopting linear SVM for prediction extracting SIFT descriptors and building 1024D feature vector with the same “visual dictionary” of our method, then adopting Kernel SVM for prediction from the results in Table 2, the kernel SVM achieved the best performance but consumed most time. Our proposed method obtained the approximate recognition performance with kernel SVM, but the recognition speed is much faster than kernel SVM. The linear SVM achieved the

Table 2
Comparison results of three method.

Categories	SIFT + linear SVM (%)	SIFT + Kernel SVM (%)	Our method (%)
Butterfly	71.50±2.13	72.30±1.96	71.80±1.93
Bonsai	75.87±3.12	76.18±3.40	76.16±3.35
Brain	85.33±3.75	87.27±3.75	87.33±3.75
Car	72.09±5.71	72.55±5.19	72.55±5.19
Elephant	81.88±4.79	86.47±4.71	86.47±4.71
Piano	95.35±1.12	98.73±1.10	98.73±1.09
Starfish	85.61±4.04	87.09±4.36	87.25±4.33
Sunflower	92.38±5.82	98.25±6.43	97.85±6.31
Mean acc. (%)	82.50±3.81	84.86±3.79	84.75±3.83
Speed	0.12s	0.25s	0.10s

worst recognition performance and the worse recognition speed than our method. Therefore, the proposed method achieved the balance of accuracy and speed.

(2) Feature compression

We compared the performance of the feature compression, the classical PCA and SVD were employed for comparison. In addition to the method of feature compression, we adopted the same settings for training and testing stage. The original 1024D feature vectors were compressed to different dimensions ($D=20, 40, 60, 80, 100, 120, 140, 160, 180, 200$). The results are shown in Fig. 7.

(3) Model optimization

Additionally, we validated the performance under different reduction rate (it means the proportion of the reduced support vectors to the total number of support vectors). Fig. 8 illustrate the results. From the aspect of accuracy, the optimized model keeps the performance approximately unchanged with reduction rate from 10% to 50%, when the reduction rate is raised to 60% or higher, the performance becomes decrease. From the aspect of recognition speed, with the reduction rate increases, the recognition speed increases linearly.

5. Conclusions

In this paper, we focus on dealing with the practical problem in the task of object recognition. In order to improve the recognition efficiency of the object recognition workflow, we proposed the method to improve the stage of feature compression and classification. Traditionally, high dimensional features are compressed and extracted the main information by PCA or SVD. To improve the compression efficiency, we adopted a very sparse matrix to map the high dimensional features to the low dimensional space by the theory of compressive sensing. Moreover, to improve the recognition speed further, we proposed a model optimization method, which accelerates the classification speed of the model by reducing the number of the support vector of the kernel SVM. From the experimental results, our proposed model achieved favourable recognition accuracy and speed.

Our future work will focus on developing more powerful features. In this work, although the local feature and yield favourable results, we think there is room for improvement. Moreover, applying this method into other computer vision tasks (such as object detection and object tracking) is very interesting as well.

Declaration of Competing Interest

The authors declared that they have no conflicts of interest to this work.

Acknowledgments

Thanks are due to Dawei Liu for assistance with the experiments and to Lin Fu for valuable discussion.

References

- [1] P. Amicic, L. Tommy, H. Fouad, Structured Sparse Principal Components Analysis with the TV-Elastic Net penalty, *IEEE Trans. Med. Imaging* 37 (2) (2016) 396–407.
- [2] V. Klema, A. Laub, The singular value decomposition: Its computation and some applications, *IEEE Trans. Autom. Control* 25 (1980) 164–176.
- [3] S. Yin, J. Yin, Accelerated max-margin multiple kernel learning, *Inf. Sci/ 370* (2016) 92–102.
- [4] S. Neha, J. Vibhor, M. Anju, An Analysis Of Convolutional Neural Networks For Image Classification, *Procedia Comput. Sci.* 132 (2018) 377–384.
- [5] X. Peng, L. Wang, X. Wang, Y. Qiao, Bag of visual words and fusion methods for action recognition: Comprehensive study and good practice”, *Comput. Vision Image Understanding* 150 (2016) 109–125.
- [6] R. Khan, X. Zhang, R. Kumar, Analysis of ResNet and GoogleNet models for malware detection, *J. Comput. Virol. Hacking Techniques* 15 (2019) 29–37.
- [7] Z. Lu, Zhang, X. Jiang, C. Alex, Deep Coupled ResNet for Low-Resolution Face Recognition, *IEEE Signal Process Lett.* 25 (4) (2018) 526–530.
- [8] Y. Xu, G. Yu, Y. Wang, A hybrid vehicle detection method based on viola-jones and HOG + SVM from UAV images, *Sensors* 16 (8) (2016) 1325.
- [9] B. Ma, H. Hu, J. Shen, et al., Generalized Pooling for Robust Object Tracking, *IEEE Trans. Image Process.* 25 (9) (2016) 4199–4208.
- [10] D. Zhuang, D. Zhang, J. Li, Binary feature from intensity quantization and weakly spatial contextual coding for image search, *Inf. Sci.* 302 (2015) 94–107.
- [11] S. Jorge, R. Javier, Exponential family Fisher vector for image classification, *Pattern Recognit. Lett.* 59 (2015) 26–32.
- [12] M. Emmanuel, T. Yuliya, C. Guillaume, A. Pierre, Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification, *IEEE Trans. Geosci. Remote Sens.* 55 (2) (2017) 645–657.
- [13] T. Kobayashi, N. Otsu, Efficient reduction of support vectors in kernel-based methods, *IEEE Int. Conf. Image Process.* (2009) 2077–2080.
- [14] D. Geebelen, J. Suykens, J. Vandewalle, Reducing the number of support vectors of SVM classifiers using the smoothed separable case approximation, *IEEE Trans. Neural Netw. Learn. Syst.* 23 (4) (2012) 682–688.
- [15] A. Manjunath, M. Pavithra, A comprehensive of transforms, Gabor filter and k-means clustering for text detection in images and video, *Appl. Comput. Inform/ 12* (2) (2016) 109–116.
- [16] A. Youness, K. Naima, g. Hassan, et al., A performance comparison of measurement matrices in compressive sensing, 2018, *Int. J. Commun. Syst.* (2018) e3576 1–18.
- [17] E. Nasser, A. Ali, Compressive Sensing Image Restoration Using Adaptive Curvelet Thresholding and Nonlocal Sparse Regularization, *IEEE Trans. Image Process.* 25 (7) (2016) 3126–3140.
- [18] S. Dasgupta, A. Gupta, An elementary proof of a theorem of johnson and lindenstrauss, *Random Struct. Algorithms* 22 (1) (2003) 60–65.
- [19] L. Liu, P. Fieguth, Texture Classification from Random Features, *IEEE Trans. Pattern Anal. Mach. Learn.* 34 (2012) 574–586.
- [20] C. Chen, C. Vong, C. Wong, et al., Efficient extreme learning machine via very sparse random projection, *Soft Comput.* 22 (2018) 3563–3574.