IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 27, 2019, accepted January 7, 2020, date of publication January 10, 2020, date of current version January 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2965537

Detecting Object Open Angle and Direction Using Machine Learning

YIWEN ZHANG[1], YONGXIONG REN[2], GUODONG XIE[2], ZHI WANG[1], HAO ZHANG[1], TIANXU XU[1], HAOYUAN XU[1], HAO HUANG[2], CHANGJING BAO[2], ZHONGQI PAN[3], (Senior Member, IEEE), AND YANG YUE[1]

[1]Institute of Modern Optics, Nankai University, Tianjin 300350, China
[2]Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA
[3]Department of Electrical and Computer Engineering, University of Louisiana at Lafayette, Lafayette, LA 70504, USA

Corresponding author: Yang Yue (yueyang@nankai.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB0703500, Grant 2019YFB1803700, and Grant 2018YFB0504400, in part by the National Natural Science Foundation of China (NSFC) under Grant 11774181 and Grant 61775107, and in part by the Fundamental Research Funds for the Central Universities, Nankai University under Grant 63191511.

ABSTRACT Nowadays, the object detection techniques have been developed rapidly for different applications, ranging from remote sensing to autonomous vehicles. We demonstrate identification of object open angle and direction using machine learning (ML) algorithms based on received light beam's intensity profiles. Compared with previous optical orbital-angular-momentum (OAM) spectrum system and other related works, our proposed technique only uses a single-shot image, and can efficiently reduce the complexity of hardware implementation. Specifically, we verify the reliability of the simulation results experimentally for 14 open angles and 32 directions. Experimental result shows that convolutional neural network (CNN) outperforms the other traditional ML algorithms, such as decision tree (DT), k-nearest neighbor algorithm (KNN), and support vector machine (SVM). As one of the variant of CNN, MobileNet (MN) has relatively simplified iteration algorithm than VGG-like net. It reduces the computational power, while still maintaining high accuracy for identification issues.

INDEX TERMS Object detection, remote sensing, machine learning.

I. INTRODUCTION

When an object is interrogated by a light beam, its characteristics can be identified by the intensity or phase information of the deviation from the original incidental beam [1]. Orbital angular momentum (OAM) beam, due to its unique features, has enabled novel applications in multiple fields [2]. Recently, researchers started to apply OAM beams into object identification [3]–[8]. As the intensity of an OAM beam is circularly symmetric, the intensity information itself is not enough to identify the rotation characteristics. For free space object detection issues, proposed systems require additional phase information of the OAM beams, and thus complex OAM spectrum analysis is usually required [5]–[8]. In such systems, spatial light modulator (SLM) and power meter are indispensable to acquire the OAM spectrum, and the process takes a lot of time and efforts. Consequently, a simple method with reduced hardware requirement is thus highly desirable.

In the field of computer vision, object detection has been developed for many highly practical applications, such as human behavior analysis, face recognition, and autonomous driving [9]–[11]. The object's location, movement, occlusion and background conditions are all useful information to be detected. One of the popular method it to use the image captured by a camera. However, the image can be easily influenced by the environment, and be blurry due to the diffraction caused by the truncation after propagating through the object. Therefore, it is vital to capture the subtle feature of the images, and classify them with suitable methods [12].

Recently, machine learning (ML) technique, which has the ability to automatically extract features of data at multiple levels of abstraction, has already been impacting a wide range of sensing and detection work by creating nonlinear detection boundaries. Deep learning (DL) with hierarchical structure has emerged as a new field of ML research and achieved more

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Ayoub Khan.

significant results based on the structure of multi-layer [13]. It has been extensively developed for many innovated optical applications, such as computer vision, optical communication performance monitoring, face identification and augmented reality, etc., [14]–[18]. Convolutional neural network (CNN) based DL method can extract the features and recognize input data in an efficient manner, especially aiming at complicated data. Nowadays, DL has been applied to further improve the ability of computers to identify objects in texts, voice, images, even videos [19]–[23]. The accuracy, storage space, computation cost and energy consumption are critical factors to be considered for real-time detection problems. There are some methods to reduce the model size and improve the function of the whole neural network, like convolutional kernel decomposition, flattened networks, low bit networks, and bottleneck structure based squeezenet [24]–[26]. Specifically, many innovative techniques, such as the StuffNet [27], the HyperNet [28], and the MobileNet (MN) [29], which can reduce the requirement on computational resources by transforming from the traditional vanilla CNN.

In this work, we propose to use a reference Gaussian beam to simplify the OAM spectrum system for object identification mission using several ML methods. We demonstrate object parameter identification, such as open angle and direction by two CNN-based deep learning methods. Comparing with previous OAM spectrum analysis system, our work could reduce the hardware implementation complexity, and potentially serve for remote sensing and real-time object feature detection. Meanwhile, the MN can further reduce the computation of the CNN model. Our demonstrated technique has the potential to extend its applications, ranging from monitoring fan blade, propeller, to tire pressure. Furthermore, using external interrogated source could help detect cloaking objects.

The paper is organized as following: in section 2, we present the concept of our proposed object identification system, and provide the comparison with previous OAM spectrum analysis system. In section 3, we first describe the experimental setup and the algorithms implemented. Next, we display the simulated and experimentally captured images of light beams truncated by various objects. In section 4, we achieve object identification with various open angles and direction angles. The advantage of CNN, especially the MN, is explained in details. In section 5, we sum up the work and provide additional outlook.

II. CONCEPT

In general, Fig. 1(a) depicts the common concept of previously demonstrated object identification techniques using SLMs and power monitor [6]. In such systems, complex OAM spectrum analysis is needed to monitor rotation objects by acquiring additional phase information of the interrogated OAM beam. The complex OAM spectrum of a beam forms a Fourier pair with its spatial distribution in the azimuthal direction, in which an opening in the object would lead to a *Sinc* function of its OAM intensity spectrum. Furthermore,

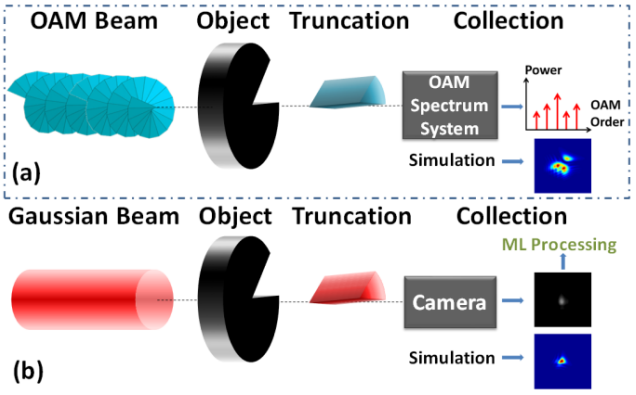

FIGURE 1. (a) Concept of using the OAM spectrum to measure an object's parameters in Ref. [6]. (b) Concept of open angle and direction identification using Gaussian beam and ML in our work.

additional data analysis is needed to determine the object's characteristics. To improve the detection accuracy, some works used the interference of two OAM beams, which is even more complicated [7]. Overall, none of these techniques can identify images in bulk.

Here, we propose a method that effectively simplifies the system's hardware, as displayed in Fig. 1(b). It reduces at least two SLMs and a power monitor, by introducing a camera. Machine learning methods can achieve extremely fast processing, and even batch identification.

After the object, the truncated beam is captured as intensity images and further processed by ML algorithms. The amplitude of a Gaussian beam, propagating along the z-axis, is given by:

$$U_{(r,\varphi,z)} = exp(-r^2/w_o^2)exp(ik_z z) \tag{1}$$

where w_o is the waist size and r is the radial distance of the beam from the central axis.

When the beam is truncated by the object in the x-y plane, the amplitude is given by:

$$U_{(r,\varphi,z)} = \Lambda\,(x, y, \alpha, \theta)\, exp(-r^2/w_o^2)exp(ik_z z) \tag{2}$$

where $\Lambda\,(x, y, \alpha, \theta)$ is a deformation of the Heaviside step function in the polar coordinate system, which is used to emulate the object.

III. EXPERIMENTAL SETUP

Figure 2(a) illustrates our experimental setup, in which we use a Gaussian beam as the probe. The SLM is used to emulate objects with various features of different open angles and directions. Different ML techniques, such as decision tree (DT), k-nearest neighbor (KNN), support vector machine (SVM), VGG-like (visual geometry group network-like) net and MN are trained to identify the images collected by the camera.

Figure 3 shows the object's features, together with the simulated and experimental images of the truncated optical beam. We generated a dataset containing 1345 collected images with 14 open angles and 32 directions (open angle θ: $\frac{2\pi}{2}$, $\frac{2\pi}{3}$, ..., $\frac{2\pi}{15}$ and direction

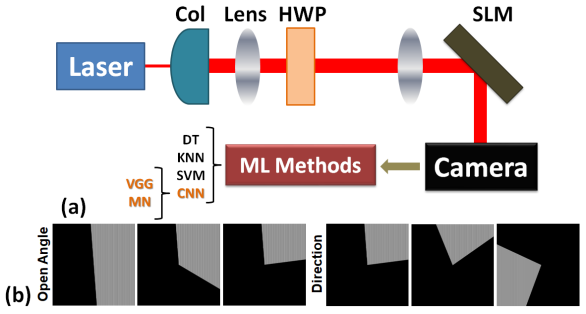

FIGURE 2. (a) Experimental setup. Col: collimator; HWP: half-wave plate; SLM: spatial light modulator. (b) The phase masks implemented to the SLM. A free-space Gaussian beam emitted from the laser source passes through lenses and HWP. The SLM is used to emulate the objects with different open angle and direction. The truncated beam can be captured by the camera, and the results are processed by different ML methods.

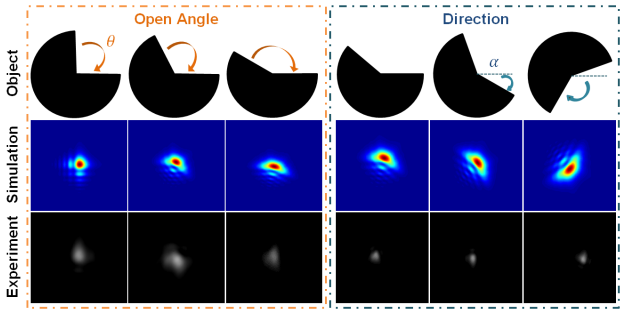

FIGURE 3. Some simulation and experiment results of the light beam truncated by various objects with different open angle (θ) or direction angle (α) relative to the x axis.

angle α: $\frac{\pi}{64}, \frac{2\pi}{64}, \ldots, \frac{32\pi}{64}$). Using data argumentation, the dataset has been expanded by 6 times. Images are divided into training and testing sets by randomly selecting 80% and 20% of the overall images.

The pseudo codes of vanilla VGG-like CNN and the ML methods we used for identification are shown in Fig. 4. In the proposed technique, the vanilla CNN typically includes three convolutional (Conv) layers in a hierarchical manner for parameter prediction, pooling layers, and fully-connected (FC) layers. After preprocessing and data argumentation, the input images are preprocessed to a size of $224\times224\times3$. The first convolutional layer has three channels with sample input. We used different filter in each layer and set rectified linear unit (ReLU). Max pooling layer with a stride of 2×2 can reduce the data size by half. The fully connected layer can flatten and memorize the data. Linear regression and soft-max algorithms are typically applied. Backpropagation and stochastic gradient descent (SGD) are used to train the CNN, adding dropouts to handle overfitting and gradient vanishing problems. Each layer chooses specific elements to learn, such as the curves or edges in the image.

We used a few traditional machine learning methods, including DT, KNN and SVM. The corresponding pseudo codes are shown in Fig. 5. DT is a flowchart-like structure in which each internal node represents a "test" on an attribute. This method can directly reflect the characteristics of the data

```
input picture size [3, 224, 224];
// first convolutional layer parameters
hidden C1 [60, 224, 224] from Picture convolve {
inputshape = [224, 224];
kernelshape = [3, 3];
stride = [1, 1];
padding = [T, T];
mapcount = 60;
}
// first pooling layer parameters
hidden P1 [60, 112, 112] from C1 max pool {
inputshape = [60, 224, 224];
kernelshape = [1, 2, 2];
stride = [1, 2, 2];
}
// second convolutional layer parameters
hidden C2 [120, 112, 112] from P1 convolve {
inputshape = [60, 112, 112];
kernelshape = [3, 3];
stride = [1, 1];
padding = [F, T, T];
mapcount = 120;
}
// second pooling layer parameters
hidden P2 [120, 56, 56] from C2 max pool {
inputshape = [120, 112, 112];
kernelshape = [1, 2, 2];
stride = [1, 2, 2];
}
// third convolutional layer parameters
hidden C2 [240, 56, 56] from P1 convolve {
inputshape = [120, 56, 56];
kernelShape = [3, 3];
stride = [1, 1];
padding = [F, T, T];
mapcount = 240;
}
// third pooling layer parameters
hidden P2 [240, 28, 28] from C2 max pool {
inputshape = [240, 56, 56];
kernelshape = [1, 2, 2];
stride = [1, 2, 2];
}
hidden H3 [640] from P2 all;
hidden H4 [448] from H3 all;
output result [14] softmax from H4 all;
//optimizer = optim.SGD(model.parameters(),
lr=0.0009, //momentum=0.95)
//criterion = nn.CrossEntropyLoss()
//epoch from 1 to 300

SGD: stochastic gradient descent; Lr: learning rate
```

(a)

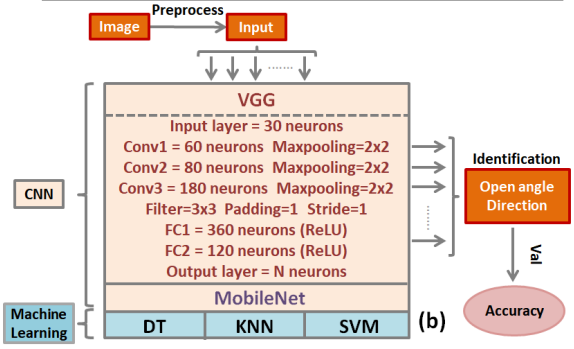

FIGURE 4. (a) The pseudo codes of the VGG net. (b) ML methods for classification and the structure of VGG. N: the total number of categories. Val: validation.

```
input picture size [3, 224, 224];
input x_train, y_train, x_test, y_test;
//from sklearn.tree import DecisionTreeClassifier;
//from sklearn.model_selection import train_test_split;
for md = 100 to 500 do
    for mf = 100 to 1000 do
    dt = DecisionTreeClassifier(max_depth = 100,
                                max_features = 1200)
    dt.fit(x_train,y_train);
    score=dt.score(x_test, y_test);
    end for
```

(a)

```
input picture size [3, 224, 224];
input x_train, y_train, x_test, y_test;
//from sklearn.neighbors import KNeighborsClassifier;
//from sklearn.model_selection import train_test_split ;
for k = 1 to 20 do
    M = KNeighborsClassifier (n_neighbors=k);
    knn.fit(X_train, y_train);
    score = knn.score(x_test, y_test);
end for
```

(b)

```
input picture size [3, 224, 224];
input x_train, y_train, x_test, y_test;
//extract HOG features
cellsize = [20, 25];
for i = 1 to number of the training images
    imagetrain = imresize(imageTrain,imageSize);
    featurestrain(i,:) =
    extractHOGFeatures(imagetrain,'cellsize',[20,25]);
end for
//input trainlabels and testlabels
classifer = fitcecoc(featurestrain,trainlabels);
for i = 1 to number of the test images do
   imagetest = imresize(imagetest,imagesize);
   featuretest = extractHOGFeatures(imagetest,'cellsize',[20,25]);
   [predictIndex,score] = predict(classifer,featureTest);
   f (i) = predictIndex;
end for
for i=1 to number of the test images
    fff = testLabels;
    if f (i) == fff (i) then
       accuracy(i) = 1;
    else then
      accuracy (i) = 0;
    end if
end for
```

(c)

FIGURE 5. The pseudo codes of the (a) DT; (b) KNN; (c) SVM methods.

and not susceptible to extreme data but easy to go overfitting. We categorized different feature vectors into the corresponding branches with the largest gain to train the model using the "sklearn.tree.DecisionTreeClassifier" module of python with the binary recursive segmentation algorithm of classification and regression tree (CART), as Fig. 5(a) shows. KNN is an algorithm especially suitable for mutiparameter problems, but it cannot be applied when the samples are unbalanced. KNN can divide the feature space into some subspaces for the test point to find the subspace to which it belongs. The number of the nearest neighbor was optimized after the dimensionality reduction for the KNN method by the "KNeighborsClassifier" module in python, which is

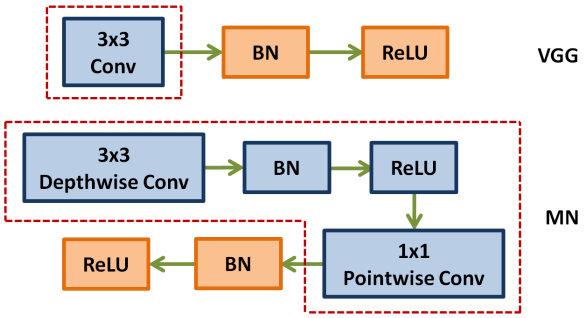

FIGURE 6. Different kernel structures (in the dashed boxes) of the VGG and MN can transform to each other. Conv: convolutional kernel; BN: batch normalization; ReLU: rectified linear unit; Depthwise Con: depthwise convolutional kernel; Pointwise Cov: pointwise convolutional kernel.

displayed in Fig. 5(b). SVM is an algorithm to minimize structural risk by constructing the optimal decision surface, which can achieve nonlinear classification by kernel method. It has lower error rate and strong generalization ability, but it is sensitive to missing data and puts high requirement on space and time. Combing with the oriented gradient (HOG) features, the parameters in SVM were set to find the optimum location of the decision surface using the "fitcecoc" function in the MATLAB, as Fig. 5(c) illustrates.

We further used MN, which is a lightweight model with less parameters and computations, based on depthwise separable convolutions. For VGG, its convolution core is used on all the input channels, while depthwise separable convolution based MN uses different convolution cores, depthwise convolutions and pointwise convolutions, for each input channel. We used depthwise convolutions to apply a single filter to the input channel of each layer. Pointwise convolution is a 1×1 convolution, which is used to create a linear combination of the output of the depthwise layer, as Fig. 6 illustrates. The number of the two-dimensional convolution kernels is the same as the number of the input channels.

IV. RESULTS AND DISCUSSION

For each identification issue, we selected different parameters to be variable or fixed. Accuracies of these two CNN methods are very close to each other for all the classification issues studied. In the following accuracy line charts, they are thus expressed by a single orange line (CNN). Accuracy curve of different open angle θ (14 sets) over specific direction angles α is shown in Fig. 7(a). One can see that all the four methods have $>75\%$ accuracy for open angle θ classification under direction angle α from $2\pi/64$ to $33\pi/64$. The collected images are blurred due to the object truncation or other environmental factors. One can see that traditional ML methods can't judge small features accurately.

As Fig. 7(b) shows, we depicted the accuracy curve of different direction angle α (32 sets) over specific open angles. SVM performs slight better than DT and KNN, but still worse than CNN. CNN outperforms other methods significantly, even more improvement than the open angle identification issue.

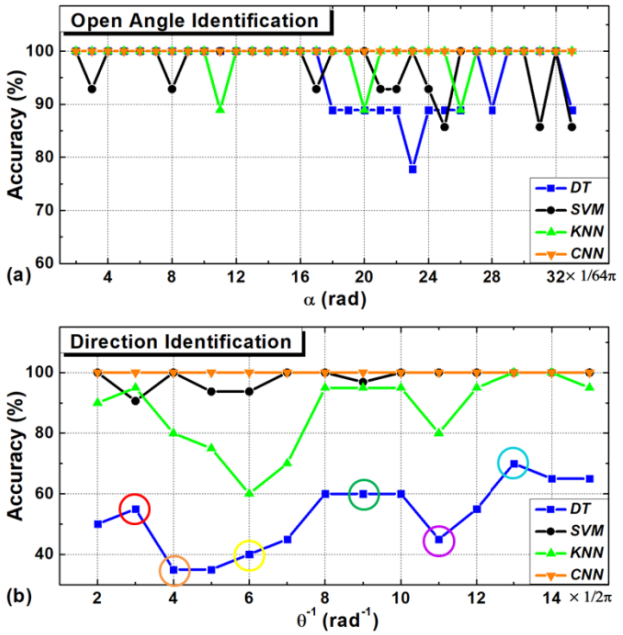

FIGURE 7. Accuracy curves of object identification using different ML methods when one parameter is fixed. (a) Open angle identification under different direction. (b) Direction identification under different open angle.

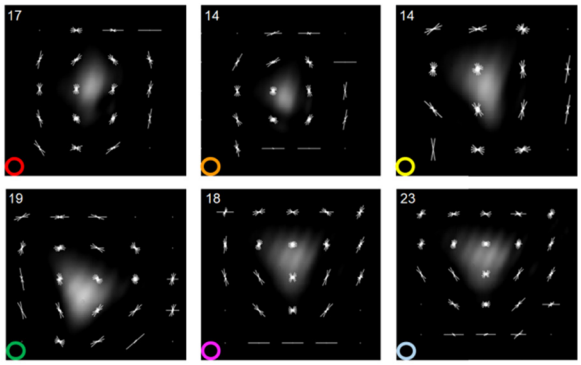

FIGURE 8. HOG features of some specific turning point. (The efficient amount of the white gradient symbols is marked on the top left corner of each image.)

One can see that the trends of the DT and SVM methods are similar. We divided the images into small cells, calculated the gradient of each pixel in the image, and got the histogram of the HOG for each cell with 20×25 pixels. For intuitive understanding, we extracted some HOG features of several turning points, which have been circled with various colors in Fig. 7(b). As shown in Fig. 8, the features can be represented by the complexity and the efficient amount of the gradient symbols.

The amount and complexity of the features altogether determine the accuracy of different ML algorithms. When the features are blurred and complex, the difference between CNN and traditional ML methods is more pronounced. CNN can achieve higher accuracy due to extensive interconnections among a large number of neurons.

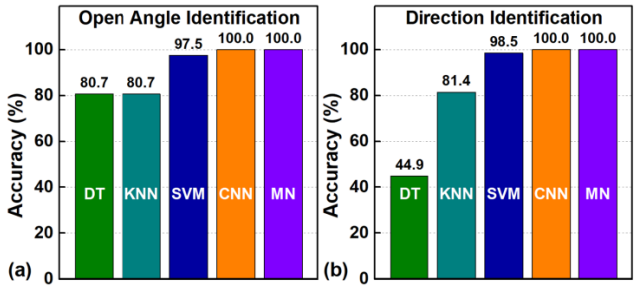

FIGURE 9. Accuracy histograms of identification using different ML methods. (a) Open angle identification under all directions. (b) Direction identification under all open angles.

Furthermore, by expanding the dataset, we identified the open angles under all directions and the directions under all open angles, respectively.

The results are shown in Fig. 9. Although the change of the images is too tiny to be recognized, CNN can still achieve 100% prediction accuracy by updating the model using dropout and backpropagation. On the contrary, DT is inevitable to be over-fitted and to ignore the correlation between attributes. KNN typically has bad estimation accuracy when the features are not obvious. Using the optimized number of support vectors, the accuracy of SVM is still slightly lower than CNN and need longer testing time.

Two CNN-based deep learning methods both have high accuracy for the above identification problems with different complexity. To show the comparison, the parameters and computations of the VGG and MN are as follows. For VGG, the sizes of the input F, the traditional kernel, and the output feature map G are (N_k, N_k, M), (N, N, M, K), and (N_g, N_g, K), respectively. The number of the input channels is M, and the number of the output channels is K. A vanilla convolution layer has the computational of $N \times N \times M \times K \times N_k \times N_k$. For MN, after splitting vanilla convolution, we get a deep convolution kernel with a size of $(N, N, 1, M)$ and a point by point convolution kernel with a size of $(1, 1, M, K)$. In sum, the computation of the whole depthwise convolution is $N \times N \times 1 \times M \times N_k \times N_k + 1 \times 1 \times M \times K \times N_k \times N_k$.

By comparing these two methods, the computation reduction of MN over VGG is:

$$\frac{N \times N \times M \times N_k \times N_k + M \times K \times N_k \times N_k}{N \times N \times M \times K \times N_k \times N_k} = \frac{1}{K} + \frac{1}{N^2}. \tag{3}$$

In our MN structure, batch normalization and ReLU nonlinearities are added to the output part. The depthwise and the pointwise convolution parts are treated as two independent modules. We further calculated the computation of some layers in the two CNN networks, which is shown in TABLE 1. In comparison, the computation cost reductions of the four convolutional layers using MN are 74.44%, 87.22%, 90.42%, and 95.74%, respectively. In sum, compared with VGG, MN has negligible precision loss, while reducing computation and storage space significantly.

TABLE 1. Computational cost of each layer.

Filter /Input Size	Net	Computation	Million Parameters	Reduce
3×3×30 224×224×3	MN [a]	224×224×3×3×3×1+ 224×224×3×1×1×30	103.86	74.44%
	VGG	224×224×3×3×3×30	406.43	
3×3×60 112×112×30	MN	112×112×30×3×3×1+ 112×112×30×1×1×60	25.97	87.22%
	VGG	112×112×30×3×3×60	203.21	
3×3×80 56×56×60	MN	56×56×60×3×3×1+ 56×56×60×1×1×80	12.98	90.42%
	VGG	56×56×60×3×3×80	135.48	
3×3×180 28×28×80	MN	28×28×80×3×3×1+ 28×28×80×1×1×180	4.33	95.74%
	VGG	28×28×80×3×3×180	101.61	

[a]MN = MobileNet.

As dynamic image processing is an arduous task in all the fields, there are many other parameters and algorithms that need to be optimized. For real-time and complex problems, we could choose less computationally intensive deep learning techniques, such as MN, with little compromise on the prediction accuracy.

Comparing with previous techniques using OAM beams, our demonstrated technique can significantly reduce the system complexity and simplify the subsequent data analysis process. From the complex characteristics of the OAM beams, one can see that the intensity images of the OAM beams would provide more features for machine learning methods to identify, and thus potentially achieving higher prediction accuracy. Nevertheless, for the above object open angle and direction identification problem, experimental results show that the prediction accuracy using Gaussian beam has already achieved 100% with CNN methods.

V. CONCLUSION

In sum, CNN-based deep learning method has high accuracy and can stay stable when detecting the open angle and the direction of the object. CNN performs better than the other traditional ML methods, and it can achieve 100% accuracy, even under extremely slight change of the object. Compared with other works, our method can significantly reduce the hardware implementation complexity, and it only requires a single-shot measurement. We also find that MN, a lightweight model with less parameters and computations, performs much better than VGG on reducing computation, while still maintaining high prediction accuracy of the identification issues. Meanwhile, the computation can be reduced significantly by >95%. At present, our demonstrated technique targets at relatively regular and static objects. In the future, more complex identification problems, such as multi-opening, irregular or moving objects, will be further explored. One might need to leverage OAM beam or multiple beams for complex feature extraction. Furthermore, unique reflection features for objects using different materials, such as stealth materials, could impact the detection techniques. Accordingly, we will also look for more suitable algorithms.

ACKNOWLEDGMENT

(Yiwen Zhang, Yongxiong Ren, and Guodong Xie contributed equally to this work.)

REFERENCES

[1] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation Through Random Media*, vol. 152. Bellingham, WA, USA: SPIE Press, Sep. 2005.

[2] M. J. Padgett, "Orbital angular momentum 25 years on," *Opt. Express*, vol. 25, no. 10, pp. 11265–11274, May 2017.

[3] M. P. J. Lavery, F. C. Speirits, S. M. Barnett, and M. J. Padgett, "Detection of a spinning object using light's orbital angular momentum," *Science*, vol. 341, no. 6145, pp. 537–540, Aug. 2013.

[4] N. Uribe-Patarroyo, A. Fraine, D. S. Simon, O. Minaeva, and A. V. Sergienko, "Object identification using correlated orbital angular momentum states," *Phys. Rev. Lett.*, vol. 110, no. 4, Jan. 2013, Art. no. 043601.

[5] N. Cvijetic, G. Milione, E. Ip, and T. Wang, "Detecting lateral motion using light's orbital angular momentum," *Sci. Rep.*, vol. 5, Oct. 2015, Art. no. 15422.

[6] G. Xie, H. Song, Z. Zhao, G. Milione, Y. Ren, C. Liu, R. Zhang, C. Bao, L. Li, Z. Wang, K. Pang, D. Starodubov, B. Lynn, M. Tur, and A. E. Willner, "Using a complex optical orbital-angular-momentum spectrum to measure object parameters," *Opt. Lett.*, vol. 42, no. 21, pp. 4482–4485, Nov. 2017.

[7] Z. Yang, O. S. Magaña-Loaiza, M. Mirhosseini, Y. Zhou, B. Gao, L. Gao, S. M. H. Rafsanjani, G.-L. Long, and R. W. Boyd, "Digital spiral object identification using random light," *Light: Sci. Appl.*, vol. 6, no. 7, Jul. 2017, Art. no. e17013.

[8] G. Milionem, T. Wang, J. Han, and L. Bai, "Remotely sensing an object's rotational orientation using the orbital angular momentum of light," *Chin. Opt. Lett.*, vol. 15, no. 3, 2017, Art. no. 030012.

[9] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019.

[10] D. Forsyth, "Object detection with discriminatively trained part-based models," *Computer*, vol. 47, no. 2, pp. 6–7, Feb. 2014.

[11] P. N. Druzhkov and V. D. Kustikova, "A survey of deep learning methods and software tools for image classification and object detection," *Pattern Recognit. Image Anal.*, vol. 26, no. 1, pp. 9–15, Jan. 2016.

[12] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.

[13] Y. LeCun, Y. Bengio, and G. Hioton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[14] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, no. 1, pp. 1–21, 2015.

[15] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," *IEEE Commun. Surveys Tutr.*, vol. 21, no. 2, pp. 1383–1408, 2nd Quart., 2019.

[16] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, Jun. 2014.

[17] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016.

[18] Y. Zhang, Z. Pan, Y. Yue, Y. Ren, Z. Wang, B. Liu, H. Zhang, S.-A. Li, Y. Fang, H. Huang, and C. Bao, "Eye diagram measurement-based joint modulation format, OSNR, ROF, and skew monitoring of coherent channel using deep learning," *J. Lightw. Technol.*, vol. 37, no. 23, pp. 5907–5913, Dec. 1, 2019.
[19] P. Yildirim and D. Birant, "The relative performance of deep learning and ensemble learning for textile object classification," in *Proc. 3rd Int. Conf. Comput. Sci. Eng. (UBMK)*, Sep. 2018, pp. 22–26.
[20] I. Sevo and A. Avramovic, "Convolutional neural network based automatic object detection on aerial images," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 5, pp. 740–744, May 2016.
[21] W. Ouyang, X. Wang, X. Zeng, S. Qiu, P. Luo, Y. Tian, H. Li, S. Yang, Z. Wang, C.-C. Loy, and X. Tang, "DeepID-Net: Deformable deep convolutional neural networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 2403–2412.
[22] J. Han, D. Zhang, G. Cheng, N. Liu, and D. Xu, "Advanced deep-learning techniques for salient and category-specific object detection: A survey," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 84–100, Jan. 2018.
[23] J. Kim and V. Pavlovic, "A shape-based approach for salient object detection using deep learning," in *Proc. ECCV*, Sep. 2016, pp. 455–470.
[24] G. E. Hinton, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
[25] Y. Zhang, K. Sohn, R. Villegas, G. Pan, and H. Lee, "Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction," in *Proc. CVPR*, 2015, pp. 249–258.
[26] J. Xue, J. Li, and Y. Gong, "Restructuring of deep neural network acoustic models with singular value decomposition," in *Proc. 14th Annu. Conf. Int. Speech Commun. Assoc. Interspeech*, Aug. 2013, pp. 2365–2369.
[27] S. Brahmbhatt, H. I. Christensen, and J. Hays, "StuffNet: Using 'Stuff' to improve object detection," in *Proc. WACV*, Mar. 2017, pp. 934–943.
[28] T. Kong, "HyperNet: Towards accurate region proposal general and joint object detection," in *Proc. CVPR*, 2016, pp. 845–853.
[29] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," Apr. 2017, *arXiv:1704.04861*. [Online]. Available: https://arxiv.org/abs/1704.04861

YIWEN ZHANG received the B.S. degree in optical information science and technology from the Dalian University of Technology, Dalian, Liaoning, China, in 2018. She is currently pursuing the M.S. degree in optical engineering with the Institute of Modern Optics, Nankai University, Tianjin, China.

YONGXIONG REN received the B.E. degree in communications engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2008, the M.S. degree in radio physics from Peking University (PKU), Beijing, in 2011, and the Ph.D. degree in electrical engineering from the University of Southern California (USC), Los Angeles, CA, USA, in 2016.

He has authored or coauthored more than 130 research articles with a Google Scholar citation number of more than 5500. His publications include 57 peer-reviewed journal articles, 76 international conference proceedings, one book chapter, and three patents. His main research focuses on high-capacity free-space and fiber optical communications, millimeter-wave communications, space division multiplexing, orbital angular momentum multiplexing, atmospheric optics, and atmospheric turbulence compensation.

GUODONG XIE, photograph and biography not available at the time of publication.

ZHI WANG received the Ph.D. degree in optics from Nankai University, Tianjin, China, in 2005. He is currently a Professor with the Institute of Modern Optics, Nankai University. His research interests include photonic crystal fiber/multimode fiber mode control theory, micro/nanostructured fiber sensing technology, ultrafast fiber laser technology, nonlinear fiber optics, and nonlinear space-time dynamics of multimode fiber.

HAO ZHANG received the Ph.D. degree in optics from Nankai University, Tianjin, China, in 2005. He is currently a Professor with the Institute of Modern Optics, Nankai University. His research interests include micro/nanostructured fiber devices, novel fiber sensors, and fiber lasers.

TIANXU XU received the master's degree in optical engineering from Zhengzhou University, Zhengzhou, China, in 2018. She is currently pursuing the Ph.D. degree with the Institute of Modern Optics, Nankai University, Tianjin, China. Her research interest includes three dimensional reconstruction based on the depth camera in computer vision.

HAOYUAN XU is currently pursuing the bachelor's degree in optoelectronic information science and engineering with Nankai University, Tianjin, China. His current research interest is computer vision.

HAO HUANG received the B.S. degree from Jilin University, Changchun, China, in 2006, the M.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2009, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2014. He is currently working as a Hardware Engineer with Lumentum Operations LLC. He has coauthored more than 100 publications, including peer-reviewed journals articles and conference proceedings. His research areas include optical communication systems and components, optical sensing systems, and digital signal processing. He is a member of the Optical Society of America (OSA).

CHANGJING BAO received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2017. He has authored or coauthored more than 100 journal articles and conference proceedings. His research interests include optical communications, nonlinear optics, and integrated optics.

ZHONGQI PAN (Senior Member, IEEE) received the B.S. and M.S. degrees from Tsinghua University, China, and the Ph.D. degree from the University of Southern California, Los Angeles, CA, USA, all in electrical engineering.

He is currently a Professor with the Department of Electrical and Computer Engineering. He also holds BORSF Endowed Professorship in electrical engineering II and BellSouth/BoRSF Endowed Professorship in telecommunications. His research is in the area of photonics, including photonic devices, fiber communications, wavelength-division-multiplexing (WDM) technologies, optical performance monitoring, coherent optical communications, space-division-multiplexing (SDM) technologies, and fiber sensor technologies. He has authored or coauthored 160 publications, including five book chapters and 18 invited presentations/articles. He also has five U.S. patents and one China patent. He is a Senior Member of OSA.

YANG YUE received the B.S. and M.S. degrees in electrical engineering and optics from Nankai University, Tianjin, China, in 2004 and 2007, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2012.

He is currently a Professor with the Institute of Modern Optics, Nankai University. He has published more than 150 peer-reviewed journal articles and conference proceedings, two edited book, one book chapter, more than ten invited articles, more than 30 issued or pending patents, and more than 80 invited presentations. His current research interests include intelligent photonics, optical communications and networking, optical interconnect, detection, imaging and display technology, integrated photonics, and free-space and fiber optics.

Dr. Yue is a member of the IEEE Communications Society (ComSoc), the IEEE Photonics Society (IPS), the International Society for Optical Engineering (SPIE), the Optical Society of America (OSA), and the Photonics Society of Chinese-American (PSC). He is the Editor Board Member of two scientific journals. He served as the Guest Editor for three journal special issues, a Committee Member and the Session Chair for ∼30 international conferences, and a Reviewer for more than 50 prestigious journals and the OSA Centennial Special Events Grant 2016.