Machine learning-based signal degradation models for attenuated underwater optical communication OAM beams

Patrick L. Neary [a],*, Abbie T. Watnik [b], K. Peter Judd [b], James R. Lindle [c], Nicholas S. Flann [a]

[a] Utah State University, Logan, UT 84322, USA
[b] Naval Research Laboratory, Washington, DC 20036, USA
[c] DCS Corporation, Alexandria, VA 22310, USA

ABSTRACT

Signal attenuation in underwater communications is a problem that degrades classification performance. Several novel CNN-based (SMART) models are developed to capture the physics of the attenuation process. One model is built and trained using automatic differentiation and another uses the radon cumulative distribution transform. These models are inserted in the classifier training pipeline. It is shown that including these attenuation models in classifier training significantly improves classification performance when the trained model is tested with environmentally attenuated images. The improved classification accuracy will be important in future OAM underwater optical communication applications.

1. Introduction

Orbital Angular Momentum (OAM) in electromagnetic waves has created a stir in communications research [1,2]. OAM communications have great potential in free space or underwater [3] environments to increase communication bandwidth by multiplexing modes together. The idea behind such schemes is to equate the distinct intensity patterns associated with each mode to a particular bit pattern. Messages can then be sent between the source (laser) and receiver (imager) provided that one can successfully recognize the image that was sent.

Orbital angular momentum beams provide a unique method to increase the capacity of underwater communications using the spatial dimension; for reference on work by others, see [4]. In this paper, an OAM underwater optical communication approach was implemented using 16 spatial modes (including the null set). The analysis provided in this paper focuses on the classification of these modes as an important step to full implementation of an underwater OAM system.

Several significant limiting factors in OAM communications exist, these include both turbulence [5] and signal attenuation [6]. Attenuation is a result of the total absorption and scattering experienced in a medium such as water. The power loss due to attenuation can be described as $P_r = P_0 exp^{-(a+b)z}$, where P_r is received power, P_0 is source power, a is the absorption coefficient, b is the scattering coefficients, and z is the range [7]. Attenuation from particulates is anticipated to

occur in any open system as natural events, such as storms, stir up sediment in the water.

While strides have been made toward developing a free space OAM communication link [8,9], a number of issues remain in the underwater domain. For example, the signal to noise ratio (SNR) between a laser source and an imager can change due to water turbidity. SNR in clear water may be initially high, but weather conditions may stir up particulates in the water column and cause it to decrease. Additionally, simply increasing the distance between source and receiver can cause a significant drop in SNR. Consequently it becomes important to investigate classifier performance and robustness in these situations.

A potential issue inherent in any classification system is how well it performs when it is given data outside of its training set. The objective becomes, how to use a single data set and extend it such that the trained classifier continues to perform well when the environment degrades beyond the conditions of the original training set. Also, given that environments can change drastically, what can be done to quickly and easily update a classification algorithm when conditions change beyond what is originally anticipated?

This work, therefore, seeks to develop a classifier capable of recognizing attenuated (noisy) OAM modes. The classifier is based on the prior work of [10] as well as that of [11]. These cited works applied deep learning to a collection of OAM images as a way to
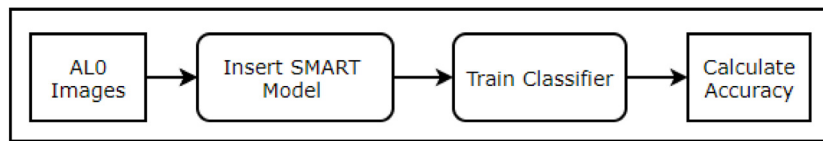
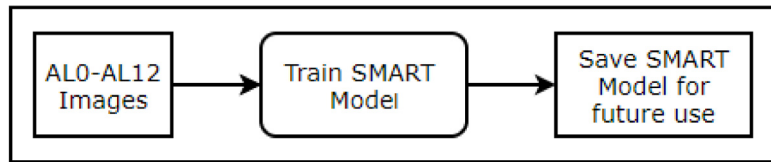**Fig. 1.** High level training process for the classification model.

Fig. 2. High level training process for a SMART model.

learn multiplexed OAM modes in a free space channel corrupted by turbulence (but no attenuation). Their images were collected on a optical bench with turbulence imparted through turbulence phase screens. Both used a custom shallow CNN as well as AlexNet [12] for OAM mode classification.

Here, their work is extended to signals attenuated in the underwater domain, and use the shallow CNN they presented as a baseline architecture. This work is specifically interested in creating machine learning-based models of the attenuated signals, and using those models during training of a classifier, with the intent of improving robustness to a low SNR environment. This paper refers to the machine learning-based model as 'SNR MAchine leaRning aTtenuation Model', or SMART Model for short. A SMART Model in this paper can be any machine learning-based model, such as a convolutional neural network (CNN).

SMART model use during the classifier training process is shown in Fig. 1. This figure shows a set of training images at the left. Batches of training images are passed to a SMART model, which will randomly apply attenuation to images within the batches. The batches of images are then passed to the classifier for training. Accuracy is calculated during the training process as depicted on the right.

The purpose of the SMART model is to act on training images and impart attenuation to those images. By imparting attenuation to training images, before they are used to train the classifier, the classifier may become more robust. In principle, the classifier will then perform better when environmentally attenuated images are presented to it. It is common practice to apply translations, rotations, flips, and other operations to images during the training process. These operations effectively extend the data set and help generalize the classifier. The SMART model is a specialized extension of this idea, where it applies attenuation operations to the image in preparation for training the classifier. The SMART models learn how to apply attenuation in several ways that are discussed in detail later.

To this end, several modeling approaches are explored. It is shown that when a classifier is trained with high SNR images, it performs poorly on low SNR images. To improve performance, a SMART Model is inserted in the training pipeline. This paper, therefore, explores the training of two distinct processes. First, several methods for creating SMART Models are explored as shown in Fig. 2. Second, standard CNN-based image classifier is created and its performance is improved by using the SMART Model in the training pipeline, as depicted in Fig. 1.

The most compelling such model is based on the Radon Cumulative Distribution Transform (R-CDT) [11,13,14]. The R-CDT has been shown to closely approximate the physics of image propagation [15], hence it is expected to provide a useful model for classification purposes. Indeed, it will be shown to provide the best results while using only a single attenuated image from each class. As deep neural net models often require many samples, this approach could have significant implications in other machine learning applications as well.

Concretely, contributions of this paper include the following points. It is shown that, starting from the baseline architectures of [11], when

trained on a baseline data set and then presented with attenuated data, the classifier experiences serious degradation in performance. By creating, training, and inserting a SMART Model in the classifier training pipeline, the robustness of the classifier is significantly improved. Several novel approaches to creating the SMART Models are developed.

2. Background and prior art

In this section related work in the area of image classification is discussed first. Optimal transport and the R-CDT are reviewed, followed by a discussion of their usage in communication links. Finally, an approach for estimating the attenuation model from observed imagery is described.

2.1. Related work

In the area of image classification, the impact of environmental conditions on accuracy has been studied previously [16]. A variety of methods have been published to address these problems. For example, approaches have been developed to remove haze from images based on color [17–19]. While effective, the images used in our work are grayscale, so color based methods will not work. Other approaches have developed camera calibration algorithms for underwater conditions [20]. While calibration may help to some extent, this does not resolve continuously changing conditions in an underwater environment. Other approaches [21] work to remove attenuation from the images.

All of these approaches attempt to make the image better, so an algorithm can properly classify it. Here, the problem is approached from a different perspective by adding attenuation to the images during training. This way, the images do not have to be brought up to a specific level of quality in order to be classified correctly. Rather, classifier training is improved by including attenuated images from a SMART model. With a robustly trained classifier, then the cited methods can potentially be added to further improve performance. However, that effort is beyond the scope of this work.

2.2. Optimal transport and the R-CDT

Traversal of electromagnetic waves, through a turbulent medium, can be framed as a transport problem [5]. From this perspective, the space between the transmitter source and received plane, can be viewed as a series of lenses where the index of refraction changes at each surface. This model of the medium is based on turbulence causing variations in the index of refraction.

[5] showed that given an image starting at location $\vec{x}_0$ and an image ending at $\vec{x}_z$, a physics based transformation between the two could be made. The expression equating the two is given by Eq. (1).

$$\rho(\vec{x}_z) = \rho(\vec{x}_0)|J_f(\vec{x}, z)|^{-1} \tag{1}$$

It turns out the mapping that solves (1) can be obtained using methods developed in the field of optimal transport. Specifically, it was shown in [15] that such a mapping could be obtained by solving the problem

$$f^* = \min_f \int_{\Omega_0} \|f(\vec{x}_0, Z) - \vec{x}_0\|_2^2 \rho(\vec{x}_0) d\vec{x} \qquad (2)$$

and that this solution is exact in the event that the refractive index perturbations vanish. However, it provides an accurate propagation model even in strong turbulence [15].

Models based on the minimization of (2) have found uses in a number of one and two-dimensional classification problems [22]. One can appropriately think of the mapping $f(\vec{x}_0, z)$ as a signal transform. Moreover, in the transform domain many nonlinear classification problems become linearly separable [13]. This property, combined with the physically meaningful nature of the transform, led the authors of [11] to create an OAM beam classifier based on these mappings. The cited work showed that, in the transform space between clean and turbulence-corrupted imagery, one could build a classifier that matched a deep neural net performance using only a shallow CNN.

An efficient method for computing the transform is called the Radon Cumulative Distribution Transform [5,13]. The R-CDT is used in this work to provide transforms between high and low SNR images. The R-CDT works by computing the mapping f in polar coordinates using the so-called "sliced" Wasserstein distance which effectively divides the problem of finding the two-dimensional mapping into a number of one-dimensional problems which can be combined in closed form to yield the needed transform, f.

To summarize, this is a fast, efficient means of estimating a physically meaning full model of propagating electric fields using only two observed intensity patterns. In what follows, the R-CDT and a CNN will be used to develop SMART Models for attenuated bit patterns in a communications link.

2.3. Orbital angular momentum

Orbital angular momentum in electromagnetic waves (EM) is a relatively new finding in physics. While known to exist at the atomic level, [23] initially identified OAM properties, and then [24] rediscovered and demonstrated its existence using Laguerre-Gaussian (LG) beams. Allen and his team showed that certain fields from Laguerre-Gaussian beams contain an angular component relative to the z-axis (direction of propagation), that is dependent on $exp(-i\ell\Phi)$, where ℓ is called the topological charge. When $|\ell| > 0$ the wavefront becomes a helically shaped wavefront. The beam center, which usually has a radially Gaussian intensity profile, adopts a dark hole at the center (optical void) with an intensity pattern in a ring around the void. As the value of ℓ increases, the radius of the ring also increases. The sign on ℓ determines the direction of rotation about the z-axis. R1.3

Monochromatic light from an LG beam propagating in the z direction can be expressed in cylindrical coordinates as:

$$U_{l,m}(r, \Phi, z) = A_{l,m}[\frac{W_0}{W(z)}](\frac{r}{W(z)})^l L_m^l(\frac{2r^2}{W^2(z)})exp(-\frac{r^2}{W^2(z)})$$
$$\times exp[-jkz - jk\frac{r^2}{2R(z)} \mp jl\Phi + j(l + 2m + 1)\zeta(z)], \qquad (3)$$

where r is the radius and Φ the angle defining the plane perpendicular to the axis of propagation, z. The parameters l and m are azimuthal and radial indices, $A_{l,m}$ is a constant, W_0 is the waist radius, $W(z)$ is the beam width, L_m^l represents generalized Laguerre polynomials, $R(z)$ is the radius of curvature for the wavefront at position z, and ζ is phase delay of the wavefront. The OAM dependency is given by $exp(-i\ell\Phi)$, which controls whether the EM wave experiences a helical wavefront ($|\ell| > 0$) or a plane wave ($\ell = 0$).

A useful property of OAM, is that topical charge modes are orthogonal to each other [25]. Consequently, modes of different topical charges can be multiplexed together, and data carried within the modes

are completely orthogonal and easily demultiplexed, at least in ideal environmental conditions.

Several methods exist for imparting OAM modes to Gaussian laser beams [26–28]. Optical vortex phase plates are an optic composed of a helical surface that is specially designed for specific wavelengths. Each phase plate is designed to impart a specific mode to the Gaussian beam.

2.3.1. OAM communications

[11] presented free space OAM mode classification with turbulence simulated by using a phase screen. They trained classifiers using OAM images and compared them to classifiers trained using patterns from the R-CDT. Their work is used as a launching point for exploring the UWOC domain. Specifically, they used a custom neural network and showed very good classification results with it. That same architecture will be used here. They used the R-CDT patterns for classification, but in this paper, the R-CDT will be used to map intensity distributions from low to high attenuation levels.

While neural networks have been able to achieve impressive results in many different applications, one concern with using them is that their predictions may be inconsistent with physical realities. This has been one of the drivers behind the idea of physics-guided training in neural networks. [29], for example, demonstrated a physics-guided neural network that was used for modeling lake temperatures. During training, it was shown that the neural network could predict values inconsistent with reality. By including physics models in the training loop, they were able to speed up and improve the model's results by using physics models to impose limits on realistic outputs.

This principle is adapted to training an OAM classification model. As it has been shown that optimal transport is able to capture the underlying physics in a system [15], it can be used to learn a physics-based SMART model for attenuation. An advantage of using the R-CDT is that it is able to capture the transform using a minimal number of samples. Once the SMART model is learned, it can be placed in the training loop. The SMART model can be used to modify a high SNR data set and transform some of the images to their low SNR counterparts. It will be shown that applying these principles greatly improves classification performance.

The rest of this paper will proceed as follows. Section 3 discusses the hardware configuration used to generate OAM data, details attenuation model creation, and explains classification network architecture with training parameters. Section 4 presents results from training and inference. Finally, Sections 5 and 6 wrap-up with discussions of the results and conclusion.

3. Experiment setup

In this Section the hardware configuration used to generate the OAM images, image set composition, SMART model architectures and training details, OAM mode classifier architecture and training details, and final test configuration are discussed.

3.1. Hardware configuration

A 1.2 meter water tank was used in the configuration for image acquisition. A Gaussian beam is generated by Bright Solutions, Onda Q-switched diode pumped solid state laser operating at 532 nm with a 5 ns pulse width. The laser and a camera are synchronized at 1 kHz. As displayed in Fig. 4, the laser beam is split into four coherent beams and is passed through vortex phase plates with different charges, where an OAM phase is imparted to each beam. Subsequently, the beams are recombined using beamsplitters, the beam passes through the water tank, and are incident on the camera. The OAM modes used for this configuration are [1, 4, −6, −8]. This combination provided the most unique patterns, given the available set of vortex plates. The range of OAM modes was kept close, as higher order modes diverge significantly more than low order modes. High order and low order
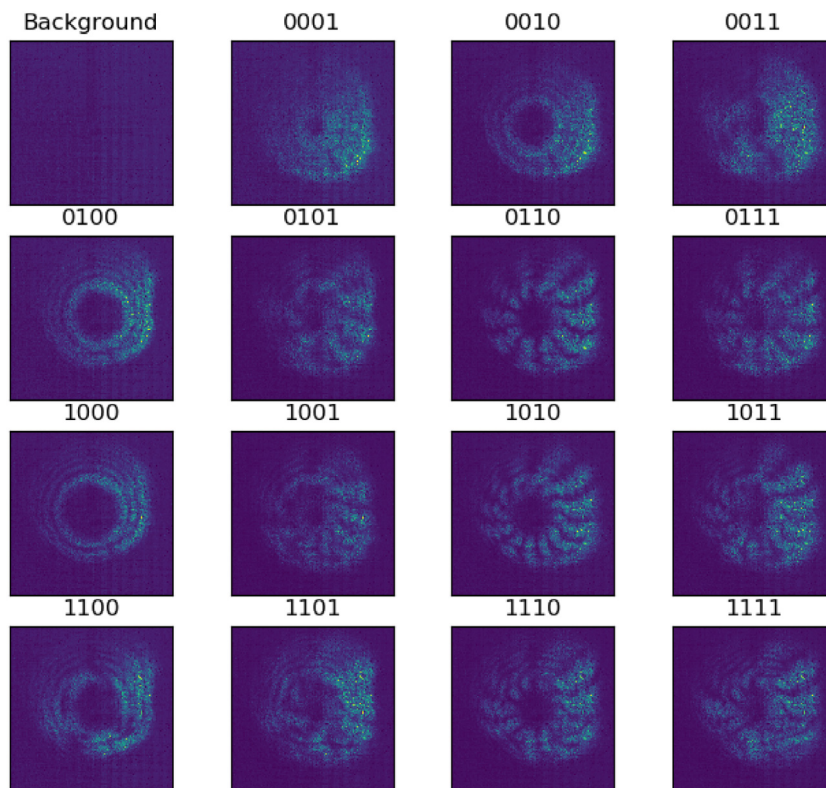
Fig. 3. Examples of intensity patterns for each OAM mode combination.

OAM mode patterns do not spatially overlap sufficiently to produce unique patterns [30]. A Photron FASTCAM SA-Z camera is externally triggered to be in sync with the laser. The camera lens was removed for this setup.

Since OAM modes are orthogonal, the modes can be mixed and matched to produce different light patterns. The beams for each phase plate can be blocked, so with 4 OAM modes there are 2^4 possible combinations. For this effort, 15 patterns are used and one class consists of a background reading, which is an image capture on the camera when the laser beam is completely blocked. One thousand images were collected for each class (or OAM pattern), for a total of 16,000 images per data collection. Fig. 3 shows intensity patterns, without attenuation, for each of the OAM mode combinations used in this paper.

In Fig. 5, one may notice vertical stripes in the images. They appear as a result of turning up the gain on the camera when detecting highly attenuated signals. Additionally, the image intensity patterns appear to be grainy, as opposed to being continuous. The graininess and lower pattern-intensity level is a characteristic inherent of the specific camera used in this experiment when operating with short exposures.

3.2. Image set composition

Five micron, PSP-5, polyamid seeding particles (Dantec Dynamics) were added to the water to attenuate the signal and model a turbid water column. The beads are designed to scatter light, where the amount of scattering is proportional to the number of beads present. The attenuation length was measured in the lab by using a probe laser passing through the tank, parallel to the OAM beam. The probe laser was incident on a sensor that gave a reading of the attenuation length.

A total for 4 sets of data collects were made for attenuation lengths 0, 4, 8, and 12 (AL0, AL4, AL8, AL12), where AL0 was collected without any attenuation beads present. Each set contains 16 patterns. There are 1000 captures for each pattern, so each AL set has 16,000 total images. Signal attenuation lengths are described by Eq. (4), where ρ is
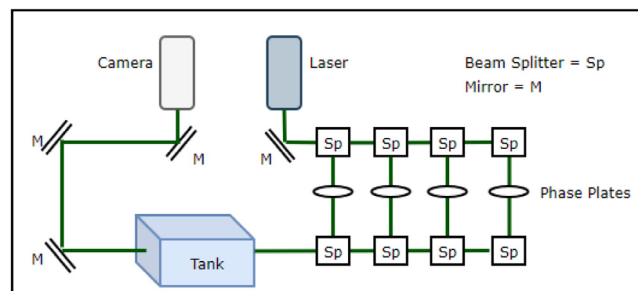

Fig. 4. Diagram of bench configuration. Beam generated by the laser, bounces off of a mirror and, goes through beam splitters. Split beams pass through phase plates where OAM modes are imparted. OAM beams are then multiplexed together, using beam splitters, and passed through the tank. The beam then bounces off several mirrors and is incident on the camera.

the attenuated signal, ρ_0 is the original signal strength, and AL is the level of attenuation. In this case AL values are set to 0, 4, 8, and 12

$$\rho = \rho_0 e^{-ALz} \tag{4}$$

Fig. 5 shows an example of six classes at each of the attenuation lengths. In OAM pattern 0001, the AL0 image shows a distinct pattern with a high SNR. As the attenuation length increases, the noise floor of the camera moves up and the SNR drops. At AL12, noise overcoming the lower intensity portions of the OAM pattern is observed. For a standard convolutional neural network trained to look for patterns, it is easy to see why there would be significant degradation of performance when presented with images with low SNR.

In generating results for this paper, the data sets were divided up in the following way. Each attenuation set was split into 70%/15%/15% for training, testing, and validation, respectively. The attenuation and classification models were trained using only the training and validation sets. Test sets were used only once, at the end to produce final
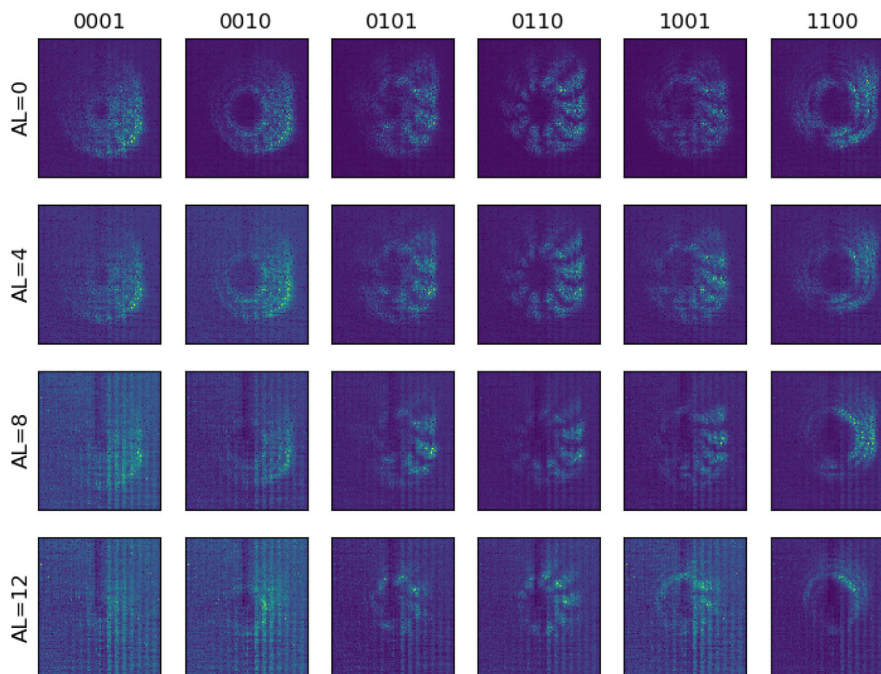
Fig. 5. Examples of attenuation in underwater environment. Each column contains patterns for one OAM mode and each row represents a specific level of attenuation. The first row contains attenuation level 0. The subsequent rows show progressively higher levels of attenuation. Note that these images are scaled to the brightest pixel in each individual image to make the OAM mode patterns easier to see.

results. Images from the camera were originally 1024×1024 grayscale images. The borders were cropped to remove pixels without intensity information and were downsampled to a final dimension of 128×128 pixels.

As one goal is to see how well the classifier performs with unseen data, the OAM mode classifier is trained using only the AL0 data set. Once the classifier has been trained, its accuracy is evaluated with all attenuated image test sets. Concretely, the AL0 test set is first passed through the classifier and its performance metrics are recorded. AL4 is then passed through the classifier and its performance metrics are recorded. AL8, and AL12 are sequentially processed in the same manner. This way, classifier performance with each group and can easily be recorded along with the performance trends as the attenuation level increases. While the sets were processed serially, they could have been processed in parallel with the right hardware.

3.3. SMART model details

The purpose of the SMART attenuation model is to learn how to mimic the SNR characteristics of the environment and hardware. Once an algorithm learns that mapping, the SMART model can be placed in the classifier training loop. In this section, details of the SMART models used to learn signal attenuation performed by the environment are discussed. Two different SMART models are discussed, a convolution based model and an R-CDT based model.

3.3.1. CNN-based SMART model

For the CNN SMART model, convolutions and feature maps are combined in a non-standard way to produce images that resemble their attenuated counterparts. Fig. 6 shows the architecture for the CNN model. Model construction is based on observations that AL4–AL12 images have a low SNR and the laser introduces grainy characteristics in the intensity patterns. The first convolution and max layers act on the original image and then the original image is subtracted from the result. There are several benefits to this combination. One, the effective SNR between background and signal is reduced. Additionally, as previously mentioned, the laser introduces a grainy characteristic

in the intensity patterns. Convolutions can potentially smooth out an image, so subtracting off the original image can reintroduce some of the grainy characteristics in the intensity patterns. This process is repeated in several stages as displayed in Fig. 6.

In training the CNN-based SMART attenuation model, all AL training data sets (AL0, AL4, AL8, and AL12) are used. The four sets contain a total of 54,400 images. During training, AL0 images are randomly selected from the training set. Once the source images have been collected, attenuated images are drawn from matching classes and paired with the source images. Loss terms are calculated using the mean-squared-error (MSE) between the model output and images from the attenuated set. Model weights are updated using automatic differentiation [31] with the calculated loss. In this way, the CNN model learns how to map from one attenuation level to another. This process is shown in Fig. 7.

Automatic differentiation is used in backpropagation, but can be generally applied to any equation. In this case, the CNN-based model is trained using automatic differentiation. Once an image is generated by the model, a difference between the two images is calculated using the MSE from pixel differences in the images. The loss calculated from the MSE is used to generate gradients, which are then applied to the weights in the model. Training of the CNN model takes place over 100 epochs with a learning rate set to $1e^{-4}$.

Fig. 8 contains an example of the output from this architecture. The image on the left is the AL0 input, the image in the middle is the AL12 reference, and the image on the right is the model output.

3.3.2. R-CDT-based SMART models

As previously discussed, the R-CDT, learns a mapping from one mass distribution to another. In this case, the initial intensity distribution is provided by image classes in the AL0 training set. The 'final' intensity patterns are provided by corresponding classes in the attenuated data sets.

Once the R-CDT has been computed, it creates a transform that enables a translation of the intensity 'mass' from one OAM AL level to another. The R-CDT enables either forward or backward transforms, so as an example, one might have an image without attenuation, apply the
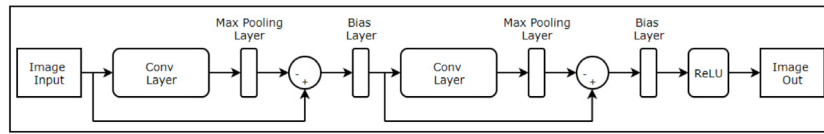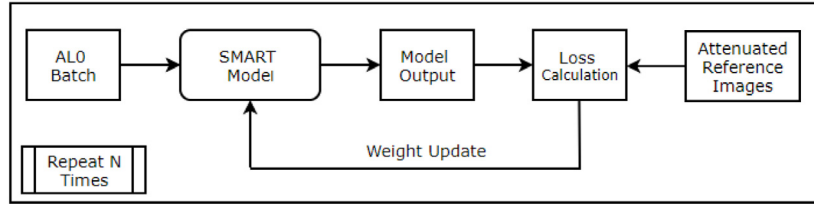
Fig. 6. CNN model architecture.

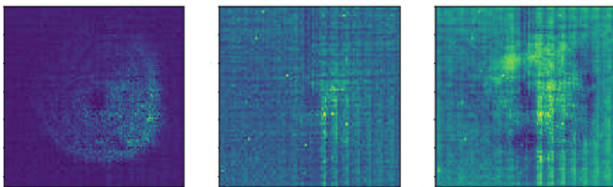**Fig. 7.** Training loop for the CNN based model.

Fig. 8. CNN AL0 input (left), AL12 reference (center), and CNN forward map output (right).
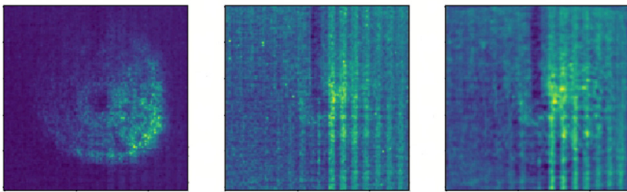

Fig. 9. R-CDT AL0 input (left), AL12 reference (center), and R-CDT forward map output (right).

Table 1
Classifier comparison.

Classifier	AL0	AL4	AL8	AL12	Num. Model images
No model	100	75.9	23.1	6.1	N/A
CNN model	99.6	82.6	66.5	55.5	54,400
R-CDT model 1	99.9	86.8	65.3	73.5	32
R-CDT model 2	99.9	97.3	82.6	74.2	64

forward transform, and get an image with intensity mass distribution of a high attenuation level.

In this work, two R-CDT SMART models were created. The first model is a mapping from AL0 to AL12 images, so it uses 32 total images to form the model (16 from AL0 and 16 from AL12). With this model, AL4 and AL8 are also evaluated for improved gains in accuracy similar to AL12. The second model includes mappings from AL0 to AL4, AL8, and AL12. This model is created by using a single source image from each class in AL0 and mapping them to corresponding images from the classes in the other attenuated sets. This model uses a total of 64 images to create the mappings. This is in contrast to the 54,400 (drawn from AL0–AL12) images used to train the CNN model.

Fig. 9 shows an example of the R-CDT input, reference, and output images. The left image is from the AL0 set and is given to the R-CDT forward map operation. The middle image is an example of the same class from the AL12 data set. The right image is the output from the forward operation of the R-CDT model.

3.4. Classification model details

The training classifier is a shallow convolutional network patterned after the one used in [11]. Using this same architecture provides a baseline for comparison to previous work. The shallow convolutional network is composed of six basic layers. It has a convolution layer with 16, 11×11 kernels, a convolution layer with 32, 3×3 convolution kernels, a max pooling layer, a ReLU, a flattening operation, and a dense layer with an output size of 16 (one for each class). The optimizer is Adam with an initial learning rate of $1e^{-3}$. The mini-batch size is 32 and training takes place over 1050 iterations.

The classifier training process is shown in Fig. 10. While training the classifier, training batches from AL0 are acquired. Images within the batch are then randomly selected to be transformed with the SMART model. Random selection helps broaden the variety of images the classifier sees over multiple epochs of training. The batch is then passed on to the classification model for training.

Fig. 11 shows the process for using the attenuation model in the training pipeline. At the top, it shows the different attenuation length sets. The left box indicates a test set that will be fed through the classifier. The data goes through the classifier, generates a set of predictions, and then the results are accumulated and a confusion matrix is created based on those results. The rows of the confusion matrix represent instances in a predicted mode while each column represents the instances in an actual mode. A confusion matrix with perfect classification will only have entries along the diagonal; any entries off the diagonal represent where the classifier misidentified the mode. Test sets are passed, one at a time, through the classifier and results are collected for each set.

4. Results

The CNN and R-CDT-based SMART models were trained on images from the AL0, AL4, AL8, and AL12 training sets. After training the CNN-based SMART model, it was placed in the classification training pipeline and randomly enabled to apply transformations to the high SNR images. The transformed images were then passed on as inputs to the classifier for training. Likewise, once the R-CDT-based SMART attenuation model was created, it was placed in a new classification training model pipeline and randomly enabled to apply transformations on the high SNR images.

Table 1 shows results from inference using the different SMART models in the classification architecture. The first column displays which SMART model was used in the classifier training pipeline. Columns 2–5 contain the classification results for the AL data sets. The
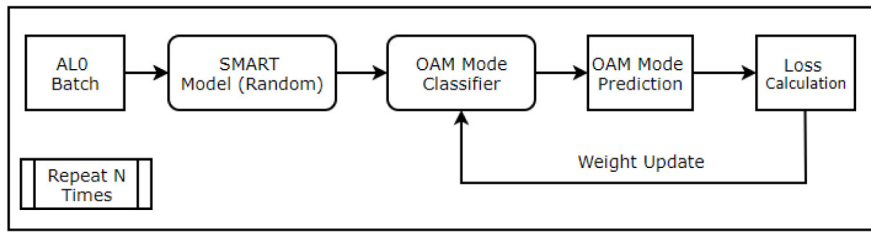
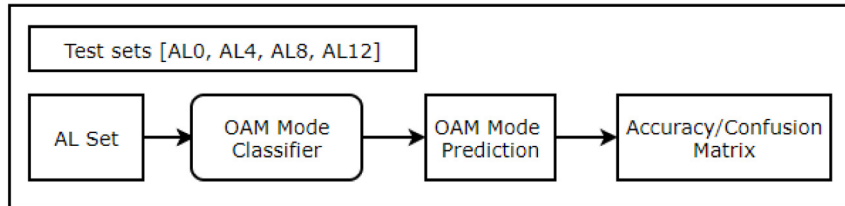**Fig. 10.** Training loop for the classification model.

Fig. 11. Test loop for the classification model.

last column displays the number of images used to create the SMART model.

The first entry shows results for the classifier that is trained without a SMART model in the training pipeline. This provides the baseline results to see how much improvement is made by adding a SMART model into the training baseline. As expected, results for AL0 are very good. AL4 sees a drop to 75.9%. AL8 and AL12 accuracies become unusable at 23.1% and 6.1%.

The second entry in Table 1 displays results for the CNN-based SMART model. AL0 experiences a small drop in accuracy, however AL4 jumps to 82.6% while AL8 and AL12 also experience significant jumps in accuracy to 66.5% and 55.5%.

Row three in the table shows results for the first R-CDT SMART model. This model contains the single mapping from AL0 to AL12. AL0 shows improved performance over the CNN model. It also shows better results for AL4 at 86.8%, slightly lower results for AL8 at 65.3%, and a significant improvement in AL12 at 73.5%.

Row four contains results from the second R-CDT SMART model, with mappings from AL0 to AL4, AL8, and AL12. This model shows the best results with AL0 at 99.9%, AL4 at 97.3%, AL8 at 82.6%, and AL12 at 74.2%.

Table 2 shows a confusion matrix generated from the classifier using second R-CDT SMART model. This table provides some insight into where the majority of the problems exist for the AL12 data set. AL12 was selected because it performed the worst and may provide the most insight into ways to improve. In this table, '0' corresponds to the 'Background' image from Fig. 3, '1' corresponds to '0001', and so on. From the table, it can be seen that images 6 (0110), 7 (0111), 11 (1011), 13 (1101), and 15 (1111) performed poorly.

5. Discussion

As shown in Table 1, the base classifier was unable to perform at the AL8 and AL12 attenuation levels, while having moderate results at the AL4 attenuation. Once the SMART models were introduced, accuracy for all attenuation levels jumped and AL12 experienced the greatest improvement by jumping from 6.1% to 74.2%.

It is interesting to note the significant improvement in performance of the R-CDT-based SMART model over the CNN SMART model. While the CNN-based model used all available training data, the R-CDT model used only one image from each of the classes in the AL12 based set. One of the difficulties that is experienced in machine learning applications is the collection and labeling of training data sets. Indeed, in many

Table 2

Confusion matrix for AL12 classification with R-CDT model 2.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	246	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	237	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	232	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	64	0	169	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	219	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	227	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	22	15	1	0	36	0	8	0	0	0	158	0
7	0	0	0	0	1	187	0	44	0	0	0	0	0	0	1	0
8	0	0	0	0	0	0	0	0	212	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	20	208	0	0	0	0	0	21
10	0	0	0	0	0	0	0	0	4	0	237	0	0	0	1	0
11	0	0	0	0	0	0	0	0	20	29	87	78	2	0	11	3
12	0	0	0	0	7	0	0	0	49	0	0	0	104	0	77	0
13	0	0	0	0	35	1	0	0	10	46	0	0	31	93	8	23
14	0	0	0	0	0	0	0	0	0	0	4	0	3	0	222	0
15	0	0	0	0	8	0	0	0	23	33	2	3	0	7	151	4

applications where groups desire to apply machine learning, the collection and labeling of large data sets for training represent a significant hurdle. The fact that the R-CDT SMART model was able to generate such a significant improvement in performance with so few examples is a powerful result. This may have some potential implications in the areas of one-shot and few-shot learning.

Implications for the usefulness of the optimal transport and the R-CDT warrant some additional commentary. The value of creating a SMART model, when data for the attenuated conditions is already in possession, may be questioned. Why not just train the classifier with all available data and be done? In reality, the likelihood of simulated environments capturing characteristics of all real environments is low. What the R-CDT SMART model approach provides, is a manageable way to update the classifier when new conditions arise. Succinctly, when conditions change, a total of 16 new images can be taken to capture the new environment. The R-CDT model can be updated to include those new images and the classifier can be retrained with the SMART model in the loop. Rather than collecting thousands of new images to add to the training set, a single image for each class can be captured and the original AL0 training set can be used to retrain the classifier. This provides a streamlined and efficient way to update the classifier when new conditions arise.

From the confusion matrix in Table 2, a number of classes with particularly poor performance is observed. The poor performance is very likely due to attenuated signals whose patterns are very similar

to each other. It shows that patterns 6, 7, 11, 12, 13, and 15 have the majority of the problems. Looking more closely at 6 (0110), for example, shows that most of the misclassifications are directed at 14 (1110). Inspecting the two patterns visually, it is apparent that the two closely resemble each other. In future efforts, this information can be used to help select OAM modes and patterns with characteristics that differ more under high attenuation environments.

6. Conclusion

From the outcomes presented in Section 4, a significant improvement in OAM classification results was observed under heavy signal attenuation. Without an attenuation model, the baseline classifier experienced a 6.1% accuracy level for the AL12 data set. With the best SMART model in the training loop, the accuracy level was raised to 74.2%.

Significant improvement in accuracy was shown for attenuation levels the classifier had not previously seen. This was accomplished by inserting a SMART model into the classifier training pipeline. The best performing model was based on the R-CDT, which has strong fundamental ties to the underlying physics of the optimal transport of photons.

A significant characteristic of the R-CDT based models is the low number of samples required. The first model required only 32 images to perform the mapping between AL0 OAM patterns and AL12 OAM patterns. The second model required only a few more at 64 images to map from AL0 to AL4, AL8, and AL12. The ability of the R-CDT to create such a robust SMART model with so few images may have significant implications in machine learning, especially for applications dependent on one-shot or few-shot learning. That is an area of application that warrants future research.

The confusion matrix for the second R-CDT model is shown in Table 2. It provides future direction as to which patterns to consider changing to provide greater pattern discrimination in high attenuation environments.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We would like to acknowledge support for this project from the US Naval Research Laboratory through a 6.2 base program for underwater communications.

References

[1] R. Chen, H. Zhou, M. Moretti, X. Wang, J. Li, Orbital angular momentum waves: Generation, detection and emerging applications, IEEE Commun. Surv. Tutor. (2019) 1.
[2] A.E. Willner, Z. Zhao, Y. Ren, L. Li, G. Xie, H. Song, C. Liu, R. Zhang, C. Bao, K. Pang, Underwater optical communications using orbital angular momentum-based spatial division multiplexing, Opt. Commun. 408 (2018) 21–25.
[3] J. Baghdady, K. Miller, K. Morgan, M. Byrd, S. Osler, R. Ragusa, W. Li, B.M. Cochenour, E.G. Johnson, Multi-gigabit/s underwater optical communication link using orbital angular momentum multiplexing, Opt. Express 24 (9) (2016) 9794–9805.
[4] Y. Ren, L. Li, Z. Wang, S.M. Kamali, E. Arbabi, A. Arbabi, Z. Zhao, G. Xie, Y. Cao, N. Ahmed, Y. Yan, C. Liu, A.J. Willner, S. Ashrafi, M. Tur, A. Faraon, A.E. Willner, Orbital angular momentum-based space division multiplexing for high-capacity underwater optical communications, Sci. Rep. 6 (1) (2016) 33306.
[5] J.M. Nichols, T.H. Emerson, L. Cattell, S. Park, A. Kanaev, F. Bucholtz, A. Watnik, T. Doster, G.K. Rohde, Transport-based model for turbulence-corrupted imagery, Appl. Opt. 57 (16) (2018) 4524–4536.
[6] B. Cochenour, K. Dunn, A. Laux, L. Mullen, Experimental measurements of the magnitude and phase response of high-frequency modulated light underwater, Appl. Opt. 56 (14) (2017) 4019–4024.
[7] B. Cochenour, K. Morgan, K. Miller, E. Johnson, K. Dunn, L. Mullen, Propagation of modulated optical beams carrying orbital angular momentum in turbid water, Appl. Opt. 55 (31) (2016) C34–C38.
[8] T. Doster, A.T. Watnik, Measuring multiplexed OAM modes with convolutional neural networks, in: Lasers Congress 2016 (ASSL, LSC, LAC), Optical Society of America, 2016, p. LTh3B.2, http://dx.doi.org/10.1364/LSC.2016.LTh3B.2, URL: http://www.osapublishing.org/abstract.cfm?URI=LSC-2016-LTh3B.2.
[9] S. Lohani, E.M. Knutson, M. O'Donnell, S.D. Huver, R.T. Glasser, On the use of deep neural networks in optical communications, Appl. Opt. 57 (15) (2018) 4180–4190.
[10] T. Doster, A.T. Watnik, Machine learning approach to OAM beam demultiplexing via convolutional neural networks, Appl. Opt. 56 (12) (2017) 3386–3396.
[11] S.R. Park, L. Cattell, J.M. Nichols, A.T. Watnik, T. Doster, G.K. Rohde, De-multiplexing vortex modes in optical communications using transport-based pattern recognition, Opt. Express 26 (4) (2018) 4004–4022.
[12] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: F. Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25, Curran Associates, Inc., 2012, pp. 1097–1105, URL: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.
[13] S. Kolouri, S.R. Park, G.K. Rohde, The radon cumulative distribution transform and its application to image classification, IEEE Trans. Image Process. 25 (2) (2016) 920–934.
[14] S.R. Park, S. Kolouri, S. Kundu, G.K. Rohde, The cumulative distribution transform and linear pattern classification, Appl. Comput. Harmon. Anal. 45 (3) (2018) 616–641.
[15] J.M. Nichols, A.T. Watnik, T. Doster, S. Park, A. Kanaev, L. Cattell, G.K. Rohde, An optimal transport model for imaging in atmospheric turbulence, 2017, arXiv: 1705.01050.
[16] Y. Pei, Y. Huang, Q. Zou, H. Zang, X. Zhang, S. Wang, Effects of image degradations to CNN-based image classification, 2018, CoRR abs/1810.05552.
[17] K. He, J. Sun, X. Tang, Single image haze removal using dark channel prior, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2341–2353.
[18] Q. Zhu, J. Mai, L. Shao, A fast single image haze removal algorithm using color attenuation prior, IEEE Trans. Image Process. 24 (11) (2015) 3522–3533.
[19] D. Berman, T. Treibitz, S. Avidan, Non-local image dehazing, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1674–1682.
[20] T. Yau, M. Gong, Y.-H. Yang, Underwater Camera Calibration using wavelength triangulation, in: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, in: CVPR '13, IEEE Computer Society, USA, 2013, pp. 2499–2506, http://dx.doi.org/10.1109/CVPR.2013.323.
[21] N. Wang, B. Zheng, H. Zheng, Z. Yu, Feeble object detection of underwater images through LSR with delay loop, Opt. Express 25 (19) (2017) 22490–22498.
[22] S. Kolouri, S.R. Park, M. Thorpe, D. Slepcev, G.K. Rohde, Optimal mass transport: Signal processing and machine-learning applications, IEEE Signal Process. Mag. 34 (4) (2017) 43–59.
[23] R.A. Beth, Mechanical detection and measurement of the angular momentum of light, Phys. Rev. 50 (2) (1936) 115–125.
[24] L. Allen, M.W. Beijersbergen, R.J.C. Spreeuw, J.P. Woerdman, Orbital angular momentum of light and the transformation of laguerre-Gaussian laser modes, Phys. Rev. A 45 (11) (1992) 8185–8189.
[25] W. Cheng, W. Zhang, H. Jing, S. Gao, H. Zhang, Orbital angular momentum for wireless communications, IEEE Wirel. Commun. 26 (1) (2019) 100–107.
[26] N.R. Heckenberg, R. McDuff, C.P. Smith, A.G. White, Generation of optical phase singularities by computer-generated holograms, Opt. Lett. 17 (3) (1992) 221–223.
[27] M.W. Beijersbergen, L. Allen, H.E.L.O. van der Veen, J.P. Woerdman, Astigmatic laser mode converters and transfer of orbital angular momentum, Opt. Commun. 96 (1) (1993) 123–132.
[28] M.W. Beijersbergen, R.P.C. Coerwinkel, M. Kristensen, J.P. Woerdman, Helical-wavefront laser beams produced with a spiral phase plate, Opt. Commun. 112 (5) (1994) 321–327.
[29] A. Karpatne, W. Watkins, J. Read, V. Kumar, Physics-guided neural networks (PGNN): An application in lake temperature modeling, 2017, CoRR abs/1710.11431.
[30] S. Avramov-Zamurovic, A.T. Watnik, J.R. Lindle, K.P. Judd, Designing laser beams carrying OAM for a high-performance underwater communication system, J. Opt. Soc. Amer. A 37 (5) (2020) 876–887.
[31] A.G. Baydin, B.A. Pearlmutter, A.A. Radul, J. Siskind, Automatic differentiation in machine learning, J. Mach. Learn. Res. (2017).