

Classification of Eddy Sea Surface Temperature Signatures Under Cloud Coverage

Evangelos Moschos, Alexandre Stegner, Olivier Schwander, Patrick Gallinari

▶ **To cite this version:**

Evangelos Moschos, Alexandre Stegner, Olivier Schwander, Patrick Gallinari. Classification of Eddy Sea Surface Temperature Signatures Under Cloud Coverage. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13, pp.3437-3447. 10.1109/JSTARS.2020.3001830 . hal-03028674

HAL Id: hal-03028674

https://hal.sorbonne-universite.fr/hal-03028674

Submitted on 27 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Classification of Sea Temperature Eddy Signatures with Strong Cloud Coverage

Evangelos Moschos, Alexandre Stegner, Olivier Schwander and Patrick Gallinari

***Abstract*—Mesoscale oceanic eddies have a visible signature on Sea Surface Temperature (SST) satellite images, portraying diverse patterns of coherent vortices, temperature gradients and swirling filaments. However, learning the regularities of such signatures defines a challenging pattern recognition task, due to their complex structure but also to the cloud coverage which can corrupt a large fraction of the image. We introduce a novel Deep Learning approach to classify sea temperature eddy signatures, even if they are corrupted by strong cloud coverage. A large dataset of SST image patches is automatically retained and used to train a CNN-based classifier. Classification is performed with very high accuracy on coherent eddy signatures and is robust to a high level of cloud coverage, surpassing human expert efficiency on this task. This methodology can serve to validate and correct detections on satellite altimetry, the standard method used until now to track mesoscale eddies.**

***Index Terms*—Mesoscale Eddies, Deep Learning, Oceanography, Remote Sensing, Computer Vision, Pattern Recognition**

I. Introduction

A. The prominence of mesoscale eddies

MESOSCALE eddies are oceanic vortices with a typical radius of the order of 20-80 kilometres which is equal or larger than the local Rossby deformation radius. They can be long-lived, with lifetimes of several months or even years. Significant advances in the resolution of both satellite altimetry measurements [1] and high resolution oceanic numerical models [2] have revealed the predominance of eddies in the global oceanic circulation. These large, coherent structures can trap and transport heat, salt, pollutants and various biogeochemical components from their regions of formation to remote areas [3]. Their dynamics can impact significantly the biological productivity at the ocean surface [4], [5], influence clouds and rainfall within their vicinity [6], modify the mixed layer [7], amplify locally the vertical motions [8] and even concentrate and transport microplastics [9]. Eddies have been demonstrated to play a prominent role in regional circulation in various areas such as the Southern Ocean [10], the Sargasso Sea [11], the Indo-Atlantic exchange [12] or the Mediterranean Sea [13]–[15]. We focus on the latter in this study.

Evangelos Moschos and Alexandre Stegner are at the LMD/IPSL, École Polytechnique, Institut Polytechnique de Paris, ENS, PSL Université, Sorbonne Université, CNRS, Palaiseau France.
Contact: evangelos.moschos // alexandre.stegner@lmd.polytechnique.fr
Olivier Schwander and Patrick Gallinari are with the LIP6, Sorbonne Université, Paris, France. Patrick Gallinari is also with Criteo AI Lab, Paris
Contact: olivier.schwander // patrick.gallinari@lip6.fr
Evangelos Moschos is a student member of the IEEE and the GRSS.
Manuscript submitted on February 28, 2020

B. Altimetric-based eddy detection and tracking algorithms

In order to detect and follow the trajectories of a very large number of mesoscale eddies on multi-satellite altimetry maps, several automatic eddy detection and tracking algorithms have been developed during the last ten years: The Okubo-Weiss parameter [16], [17], which quantifies the relative importance of rotation with respect to deformation, is used in many studies to detect and track eddies on the geostrophic surface velocities field [18]–[20]. Geometric properties of the streamlines have been used by other methods [21], [22] to identify coherent vortices without considering their intensity. Finally, a physical parameter, the local normalized angular momentum, introduced by [13] and [23], combines both the dynamical and the geometrical properties of the signature of mesoscale eddies on altimetric-based products. In this study, we follow this approach and use the Angular Momentum Eddy Detection and Tracking Algorithm (AMEDA) [23] which has shown to be very effective in locating mesoscale eddies in the Mediterranean Sea [15], [24], [25].

Despite the potential of these methods, their main drawbacks stem from the spatio-temporal heterogeneity of altimetric measurements. The creation of a daily gridded product requires an optimal spatio-temporal interpolation between the satellite track measurements. This produces low-resolution fields (1/12° in the Mediterranean Sea) with a limit on the spatial scales resolved as well as uncertainty in areas which have not been sampled by satellites. We refer to these products here on as AVISO/CMEMS altimetry maps, referring to their provider for the Mediterranean Sea.

These limitations have been quantified by [26]. They have shown that mesoscale eddies in the North Atlantic Ocean and the Mediterranean Sea could be overestimated by a 19% and 8 % respectively. Besides, according to the same study sub-mesoscale eddies, i.e. those with sizes smaller than the mesoscale, are undersampled by 94% and 84% respectively for these two regions due to the coarse resolution of the AVISO/CMEMS altimetry products.

Real-time eddy tracking on altimetry maps is also constrained, as eddies can be ”lost” by the tracking algorithms, when crossing an area at a time when it is not sampled by any satellite tracks. Similarly, they could be detected in a position prior to their real-time one, as a result of the last available measurement.

C. Why Deep Learning for eddy signature classification?

Eddy signatures are nevertheless also apparent in visible satellite imagery such as Sea Surface Temperature (SST),

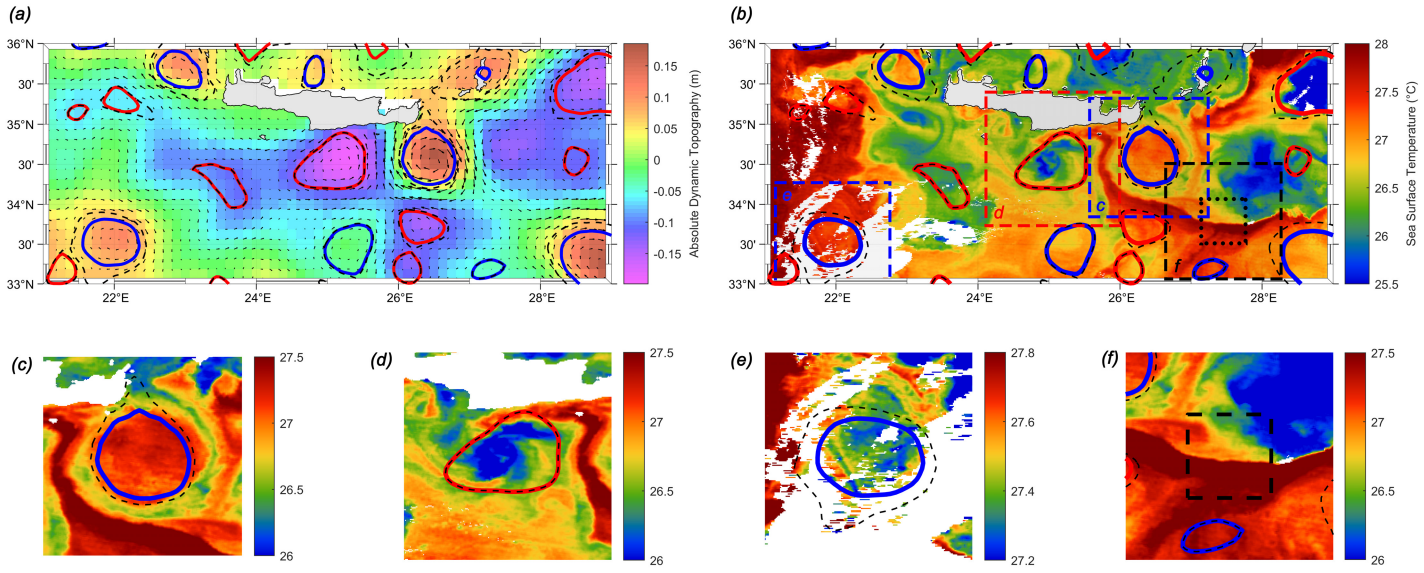


Fig. 1. (a) Absolute Dynamic Topography (altimetry) field with superimposed geostrophic velocity vectors and (b) Sea Surface Temperature field (white areas represent clouds) around Crete on the 24/08/2018. Maximum velocity and outermost contours detected by AMEDA on the velocity field are superimposed on both figures. In (b) some characteristic SST image patch selections are represented with dashed line RoI boxes: (c) A warm-core anticyclone image (d) A cold-core cyclone image (e) A cold-core anticyclone image covered by clouds (f) A non-eddy image, with the area of no-contour constrain outlined with a dashed line.

Ocean Color/Chlorophyll (CHL), or synthetic-aperture radar (SAR) images. These images have an average resolution ten times higher than that of altimetry and are not a product of interpolation. However, they are strongly affected by cloud coverage which creates missing values in the observation. This effect is especially prominent during day-time and winter-month measurements.

Several methods of eddy detection have been developed on SST images: [27] approach the problem through iso-SST pattern recognition to detect swirling fronts and gradients. In [28] the velocity field is derived from the SST field through the assumption of the thermal wind equation. Finally [29] conducted an early study training an Artificial Neural Network with gradient-based methods for eddy detection on the SST field. However, as the sign of the core surface temperature anomaly (warm core or cold core), is not always correlated to the eddy sign (anticyclonic or cyclonic), a robust method for eddy detection on SST cannot be based on the thermal wind equation and the temperature gradients.

Deep Learning has been rapidly gaining in popularity and solving problems in remote sensing [30], climate and the environment [31]. Machine learning methods have also been used in previous studies to tackle altimetric eddy detection and tracking on the SSH field. In [32] and [33] a pixel-wise segmentation approach is adopted, with the original labeling of the train set stemming from an Okubo-Weiss (OW) eddy detection method. Similarly in [34] the OW detection method is used to label training data derived from the velocity field. These studies while successfully exploring novel methods for eddy detection application, stumble upon the inherent limitation of the gridded altimetric products, on which the learning dataset is based. The measurement error will therefore propagate throughout the whole training process. In visible imagery, Deep Learning has been employed by [35]to classify eddy signatures on SAR images.

Here, following [36], we employ Convolutional Neural Networks to build a Sea Surface Temperature eddy signature classifier, a tool which can serve for validating and correcting altimetry eddy detections. This study provides contributions in automatically retaining a large dataset of SST patches with eddy signatures and constructing a CNN-based classifier of sea temperature eddy signatures. Our classifier achieves very high performance on coherent eddy signatures while being robust to high levels of cloud coverage. Our data is available under a creative commons licence.

The structure of this study is as follows: In Section II an automatic method is presented to retain a large dataset of SST image patches containing eddy signatures based on altimetry detection region proposal. In Section III the methods used to train and evaluate CNN-based classifiers are described. In Section IV the performance of the classifiers is evaluated on images containing coherent eddy signatures. Subsequently, in Section V we assess the effect of cloud coverage on the performance of the classification. Finally, in Section VI main conclusions on the given task and future prospects of Deep Learning for eddy detection are discussed.

II. Dataset Creation and Features

The task of this study consists in classifying SST images which can contain either the signature of an Anticyclonic Eddy (AE), a Cyclonic Eddy (CE) or No Eddy signature (NE). Anticyclones (cyclones) rotate in the opposite (same) direction with the earth's rotation, viz. clockwise (counter-clockwise) in the Northern Hemisphere. To this end a dataset containing such SST image patches needs to be extracted from images of larger domains. In this section a regional proposal method through

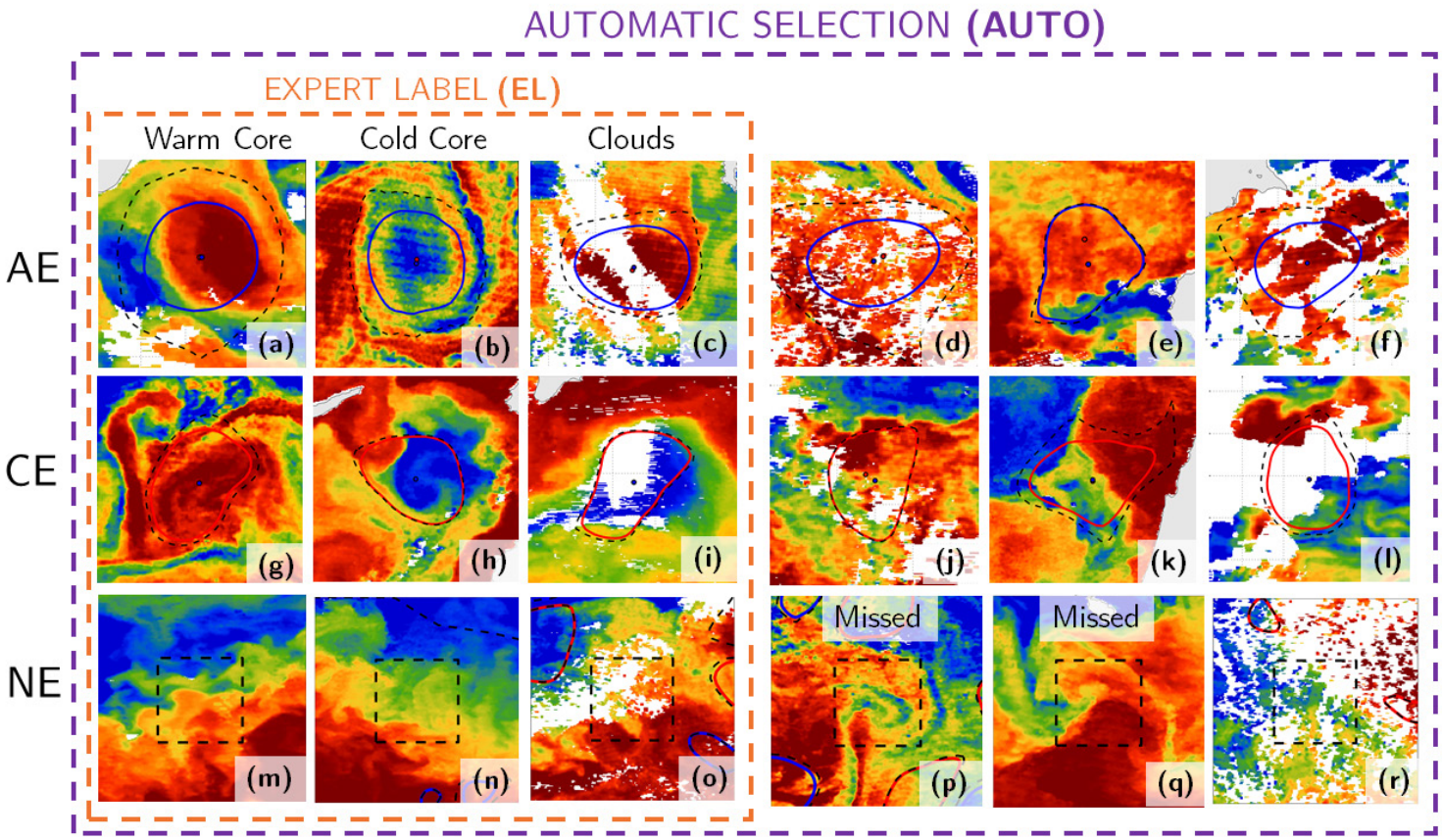


Fig. 2. Samples of images contained in the datasets. The dashed orange line box outlines coherent examples, representative of the EDDIES-EL dataset (coherent signatures), while the dashed purple line box outlines examples representative of the EDDIES-AUTO dataset (automatic selection). Row values represent the dataset labels while columns categorize coherent signature characteristics (sign of core temperature anomaly, Cloud Coverage). In the EDDIES-AUTO set, images retained through the altimetric detection regional proposal might not have a visible eddy signature on the SST, as seen in examples (d),(e) for AE, (j),(k) for CE. Similarly, images retained and labeled as NE, through the no-contour selection criterion (black dashed line box), can contain an eddy signature missed by altimetry as seen in examples (p) for an AE signature and (q) for a CE signature. Finally, examples (f),(l) and (r) represent images where validation of their label is delicate for a human expert, due to strong cloud coverage.

altimetric detection is presented, and the extracted dataset is presented.

A. Region proposal through altimetric eddy detections

The domain of the dataset of this study is the Mediterranean Sea on a 3-year time period (2016-2018). Two data sources are considered:

- **SST images** are received with a daily resolution from the Copernicus - Marine environment monitoring service. These high-resolution (1/120°) images are a product of supercollation, as described in [37] and stem from merged multisensor data, representative of nightime SST values.
- **Eddy locations and contours** are retained by applying the AMEDA on daily Adjusted Dynamic Topography and the AVISO/CMEMS surface geostrophic velocity fields with applied cyclogeostrophic corrections [24]. The AMEDA [23] detects eddies by identifying minima and maxima of the Local Normalized Angular Momentum (LNAM), computed on the surface velocity fields, and selecting closed streamlines around them. The algorithm does also dynamically track eddies backward and forward in time, as well as identifies their merging and splitting events. Eddy tracks detected by AMEDA, are labeled as AE or CE based their LNAM sign and are supplied with other metadata such as:
  - The contour of the eddy where the velocity is maximum (here on *"contour"*, shown with a bold blue line for AE and a bold red line for CE in figures). Its corresponding values of the radius R_{max} of an equal-area circular contour and of the velocity V_{max} along it.
  - The geometrical barycenter of the maximum velocity contour.
  - The outermost contour of the eddy (shown with a dashed black line in figures).

To extract image patches containing eddy signatures from daily SST maps the detections of AMEDA on corresponding daily altimetric maps are used as a regional proposal tool. Regions of Interest (RoI) are centered on the barycenters of the altimetric contours, scaled according to the physical eddy size and then interpolated to a constant pixel size. This process is illustrated in Figure 1.

TABLE I
DATASET CHARACTERISTICS

Dataset Name	Type	Year	Image No.	AE:CE:NE
EDDIES-EL(16/17)	Train	2016-17	3600	1:1:1
EDDIES-AUTO(16/17)	Train	2016-17	24000	1:1:2
EDDIES-EL(18)	Test	2018	1200	1:1:1
EDDIES-CLOUDY(18)	Test	2018	8x10x300	1:1:1

The RoI physical size corresponds to $k = \lambda * R_{max}$, where $\lambda = 5$. RoI are cropped and interpolated to a constant size of $m = \lambda * \bar{R}_{max}(km)$ where $\bar{R}_{max} = 42.5km$ is the mean maximum velocity radius of all AMEDA contours retained in this study. This results to retained rectangular image patches of side $m = 230$ pixels, labeled as AE or CE following the corresponding altimetric contour. Examples of AE and CE selections are shown in Figure 1 (c) and (d).

To extract SST image patches that do not contain an eddy signature, a box of size $m = 230$ pixels is slided along the domain of the Mediterranean Sea, with a stride of $m/2$ pixels. This way RoI are retained, on the condition that they do not contain any contour inside a centered area of side R_{max}, and labeled as NE. A NE selection example can be seen in Figure 1 (f). The no-contour centered area of side R_{max} is visualized in figures through a black dashed line box.

B. Dataset creation and labeling

Examples of images retained through the aforementioned process are given in Figure 2. These images are used to create datasets, which are used for training and testing CNN-based classifiers. The characteristics of the datasets used in this study are outlined in Table I.

Image labels received by the altimetric region proposal do not necessarily visually correspond to the SST signature depicted in them. This can be due to various reasons:

- Uncertainty of the AVISO/CMEMS altimetry maps, due to interpolation between satellite track measurements.
- Error induced by the AMEDA algorithm.
- Strong cloud coverage of the SST signature.
- Unclear SST eddy signatures due to air-sea interactions.

A large dataset named **EDDIES-AUTO**, is automatically created as described above, and contains images with labels corresponding to 3 classes $k \in AE, CE, NE$. To filter out incorrect proposals stemming from altimetric detection, only the RoI that correspond to large and intense eddies detected by altimetry are retained.

Still, the automatically retrieved dataset contains images whose visual signature does not clearly refer to their assigned label, due to a combination of the aforementioned reasons. This set is denoted as $\tilde{D}$ and contains u examples $(x, \tilde{y})$. The labels retained by the altimetry region proposal automatic selection, are denoted as $\tilde{y}$ to refer to the presence of label noise.

The effect of label noise on the EDDIES-AUTO dataset is visualized in Figure 2 : As an example, an Anticyclonic (AE) labeled image contained in this set can, have a visible signature that corresponds to its label (examples (a),(b),(c)) or one that does not (examples (d),(e)). Besides, delicate samples as the example (f), could lead experts having ambiguous opinions on its label. The same follows for the CE and NE labeled examples of Figure 2.

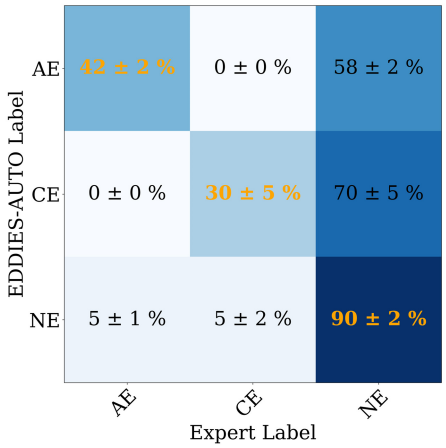


Fig. 3. Noise Matrix N_{ij} for the EDDIES-AUTO dataset, received by manually labeling 400 random samples per class, by different experts. Row values represent the labels in the EDDIES-AUTO dataset, while column values the labels assigned by experts. Cell values are normalized by the total number of sampled images per class.

Expert Labeling consists of a method which, albeit time-consuming can provide reliant and accurate labels. For the purposes of our study, oceanographic experts labeled a smaller dataset named **EDDIES-EL**, by selecting images with a coherent signature corresponding to their label. This set, denoted as $D \subset \tilde{D}$, contains images and accurate labels (x, y). The probability distribution $p(x, y)$ reflects the true distribution of the True Labels y between the three classes.

Representative examples of the coherent signatures images contained in EDDIES-EL can be seen outlined by a orange dashed line in Figure 2. Their SST signature can be distincted based on the sign of the eddy core temperature anomaly, leading to warm (examples (a),(g)) and cold core (examples (b),(h)) eddies. Both AE/CE have warm and core core examples in the EDDIES-HL set, in contrast to what is usually assumed through the thermal wind balance. Furthermore, images corrupted by clouds are included on this set (examples (c),(i),(o)), only when their signature is clearly visible to a human expert.

By defining the label-noise distribution $p(\tilde{y}|y, x)$ we can specify the level of discrepancy between the expert labeling and the noisy labels obtained by the automatic altimetry region proposal. This distribution for the EDDIES-AUTO dataset can be inferred by manually labeling a random sample of v images, with $v/u \ll 1$. We receive thus the 3 by 3 sized noise matrix of probabilities:

$$N_{ij} = p(\tilde{y} = j|y = i) \quad (1)$$

The noise matrix of the EDDIES-AUTO dataset, sampled by different experts on 400 examples of each class is shown in Figure 3. On average, 42 % of AE and 30 % of CE images are confirmed to have a humanly visible signature corresponding to their label. The rest of the images with these labels, but no humanly visible eddy signature, are allocated to the NE

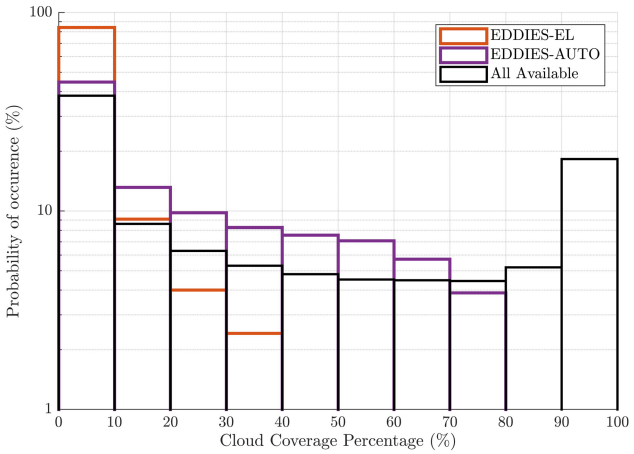


Fig. 4. Distribution of Cloud Coverage Percentages in: The EDDIES-EL dataset (Orange Line), the EDDIES-AUTO dataset (Purple Line) and all the available RoI in the sampled domain, before applying thresholds on image retainment (Black Line).

class. Likewise, out of the sampled NE-labeled images, an average 5% allocated to each of the AE and CE classes. This reflects the percentage of eddies missed by altimetric detection, corresponding to examples (p) and (q) of Figure 2. Overall, the noise matrix evaluation shows that less than half of the eddy labeled (AE,CE) images in the EDDIES-AUTO dataset have a humanly visible signature corresponding to their labels. Additionally, a small fraction of the NE labeled images contain missed eddy signatures. This discrepancy between visible signatures and labels portrays the effect of noisy labeling on this dataset, and is tackled through a transfer learning approach in the next sections.

C. Cloud Coverage

Cloud coverage has a direct impact on SST images and on the learning process, as it creates missing values in the sampled images, and often corrupts the signature apparent in the image. Cloud coverage is also related with $p(\tilde{y}|y,x)$: the ability to infer the True Label y of a cloud-covered signature depends on both the location and the density of the cloud pattern.

To quantify the presence of clouds in the datasets used in this study, a cloud coverage percentage (CCP) is calculated for every Region of Interest as:

$$CCP = n_{NaN}/m^2 \tag{2}$$

where n_{NaN} is the number of missing value pixels in each image, excluding the ones that represent the coast and $m = 230$ is the RoI side in pixels.

The distribution of CCP values is quantified in Figure 4. With a black line, the histogram of CCP values is plotted for all the available RoI to be retained through the regional proposal methodology. Out of them, only images under a threshold of 80% of Cloud Coverage are retained on the EDDIES-AUTO dataset (Figure 4, Purple Line). Thus images with a large degree of cloud coverage (last two black line bins of Figure 4) which completely corrupts the temperature signature are avoided. Finally, the EDDIES-EL dataset, has distribution with much lower values of CCP, due to the expert selection process: more than 80% of images in the EDDIES-HL have only $0-10\%$ of cloud-coverage (Orange Line, Figure 4).

III. Deep Learning Architecture and Training

Convolutional Neural Networks [38] have been successfully used in numerous computer vision applications, including ones of remote sensing. In this section we describe the architecture of a CNN-based classifier and the methods used in the training process. We also introduce a transfer learning scheme as well as indices of evaluation of the classification performance.

A. CNN Architecture

Due to the large size of the dataset and the complexity of the image features, a deep CNN architecture is used build a classifier. Residual networks [39] utilize skip connection between layers in order to build efficient deep architectures. Here, a ResNet18 architecture, with 18 fully-connected layers and skip connections, is used through the torchvision package of the Pytorch library.

The input layer of the network is modified so that a two channel input image can be received: The first channel represents the normalized temperature values and the second channel a semantic mask representing missing data locations. The final layer of the network is also adapted to a three-class output, normalized through the soft max equation. Training and weight update is performed through a cross-entropy loss and stochastic gradient descent with momentum.

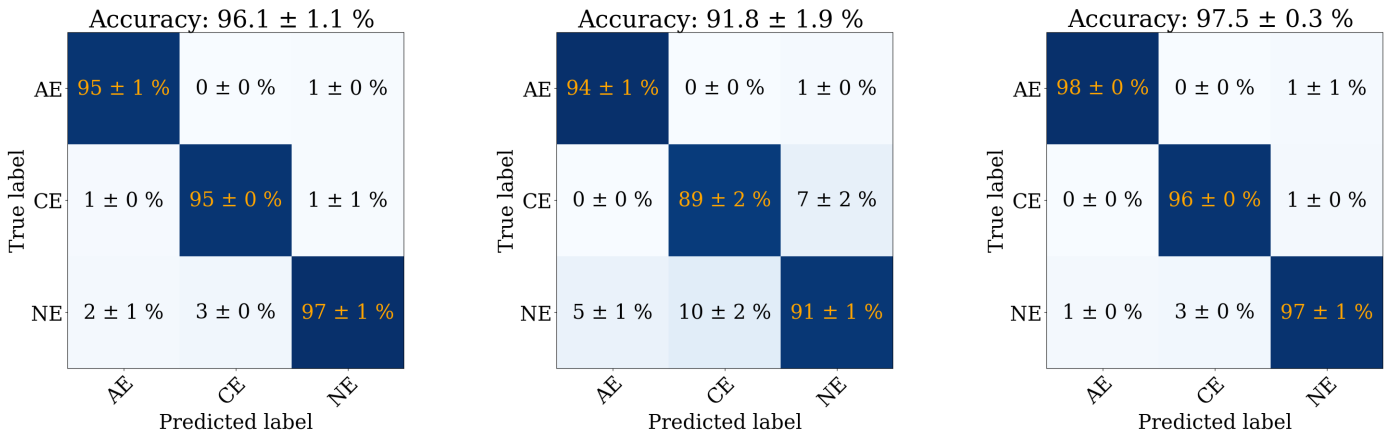
B. Training methods and transfer learning

Random orthogonal rotation is performed on input images during the training process, in order to achieve rotational invariant model training. Rotational data augmentation provides both a different geometric perspective as well as potential alternative instants of image that depict physically rotating structures.

A 5-fold cross-validation is performed in all model training runs. A different 20% of the train set, serves each time for validating the performance after every training epoch. To avoid overfitting, regularization is performed in the training process. An early stopping scheme is adopted based on the loss of the validation set.

Transfer learning aids CNN training aids by extracting features from a large dataset of images and utilizing the learned features for a more specific task. Here, we do this by pretraining CNNs on datasets of images larger than the specific task. We perform a non-zero weight initialization in our CNN training by pretraining in two different ways:

- All the ResNets trained for the purposes of this study are already pretrained on Imagenet, a large dataset of more than 14 million images. This way, weight initialization is performed with the shallower layers being able to detect common image features such as edges or gradients.
- Pretraining is also performed on the larger EDDIES-AUTO dataset, providing weight initialization for fine-tuning on the EDDIES-EL dataset. The model trained this way is referenced as $AUTO/EL$.



(a) Classifier-EL (b) Classifier-AUTO (c) Classifier-AUTO/EL

Fig. 5. Confusion Matrices on the EDDIES-EL(18) test set. Two models, trained on different datasets through a 5-fold cross-validation, are evaluated. Cell values represent the mean ± the standard deviation of the C_{ij}^{pre} of the classifiers trained on 5 different folds of the corresponding dataset.

C. Trained Classifiers and evaluation indices

We train three different classifiers, through the 5-fold cross-validation scheme:

- **Classifier-EL** is trained on the EDDIES-EL(16/17) dataset, that is on a relatively small amount of coherent signature images, with weak cloud coverage, whose labels can be directly validated by an expert.
- **Classifier-AUTO** is trained on the EDDIES-AUTO(16/17) dataset, that is on a more diverse set of images, with strong cloud coverage and presence of label noise.
- **Classifier-AUTO/EL** is first pretrained on the EDDIES-AUTO(16/17) dataset, and then finetuned on the EDDIES-EL(16/17) dataset. Finetuning is performed on all layers of the CNN. This way features from the more diverse in signatures and cloud-coverage EDDIES-AUTO dataset can be extracted, while finetuning on the EDDIES-EL dataset of coherent signatures.

Evaluation of the classifiers is performed on test sets in the form of precision normalized confusion matrices. Each cell (i, j) of the 3 by 3 sized matrix represents the precision defined as the probability of an image predicted by the classifier (y_{pred}) in class j to be labeled in the dataset (y_{true}) as class i:

$$C_{ij}^{pre} = p(y_{pred} = j | y_{true} = i) \tag{3}$$

Values of equation 3, where $i = j$, i.e. at the diagonal of the confusion matrix, are referred to as the Class Precision. In order for the CNN-classifier to be confident in the eddy signature classification task, high values of class precision are required for the AE and CE classes.

The overall evaluation of a classifier can also be performed through the Classification Accuracy, a metric robust for class-balanced test sets. The classification accuracy is defined as the percentage of images predicted correctly in the test set used for evaluation:

$$A = p(y_{pred} = y_{true}) \tag{4}$$

By performing a 5-fold cross-validation training, precision and accuracy values are provided in a mean ± standard deviation form, between the evaluation of the different training folds.

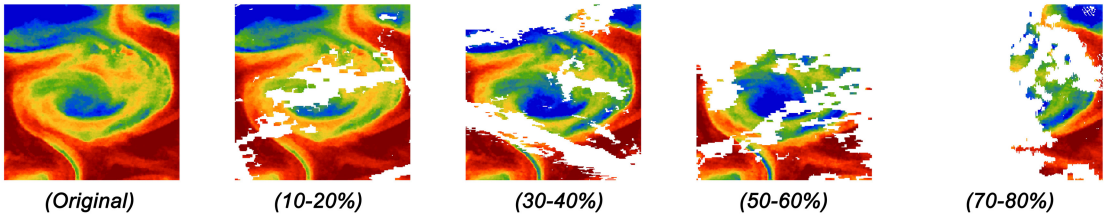
IV. Classification of Coherent Signatures

The classification performance is firstly evaluated on images containing coherent signature with a small or no amount of corruption due to cloud coverage. To this end, the EDDIES-EL(18) test set is used in order to evaluate classifier performance. The three trained classifiers (EL,AUTO,AUTO/EL) are inter-compared based on the precision normalized confusion matrices. The confusion matrices of Figure 5 show the precisions C_{ij}^{pre} for each of the given cells.

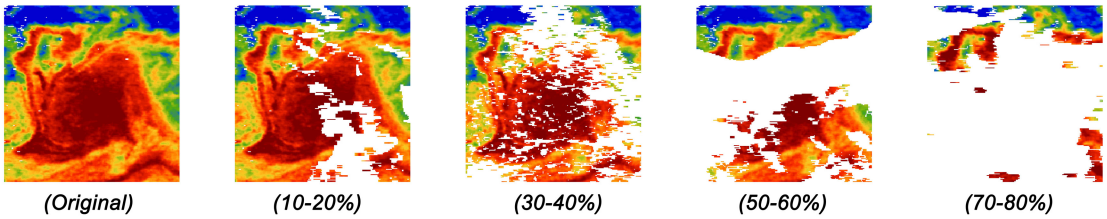
All classifiers show a robust performance on the EDDIES-EL test set, with mean classification accuracies of $91.8 \pm 1.9\%$ (Classifier-AUTO), $96.1 \pm 1.1\%$ (Classifier-EL) and $97.5 \pm 0.3\%$. (Classifier-AUTO/EL). The high classification accuracy achieved by the Classifier-EL shows that by training on a small dataset of coherent signature images, as is the EDDIES-EL(16/17) train set, a classifier with robust performance on these type of examples can be constructed.

The effect of noisy labeling of the EDDIES-AUTO(16/17) set in the training process can also be seen here: The Classifier-AUTO achieves the lowest classification accuracy between the three classifiers, when evaluating on a dataset of coherent signature images (Figure 5b). However, by finetuning it on the EDDIES-EL dataset, the received Classifier-AUTO/EL achieves the best performance between the three by increasing the mean and reducing the standard deviation of the classification accuracy (Figure 5c).

Nevertheless, the experiment presented here is evaluated on a dataset containing signatures which are much more clear that the ones existing in the whole domain of application. The robustness of classification on examples with strong cloud coverage corrupting the SST signature, is evaluated on the next section.



(a) Anticyclonic Eddy Example



(b) Cyclonic Eddy Example

Fig. 6. Examples of Cloud Data Augmentation of an AE (top line) and a CE (bottom line) from the EDDIES-CLOUDY set. The original image from the EDDIES-HL(18) test set, along with multiple examples with different levels of cloud coverage percentages are visualized. The corruption is performed by superimposing random cloudy masks from an auxiliary EDDIES-AUTO(18) set. All of the examples in this figure were correctly predicted as AE/CE correspondigly by the Classifier AUTO/EL. Colour range is on the $5_{th} - 95_{th}$ percentiles of the non-missing pixels.

V. Classification of Cloud-Covered Signatures

Cloud coverage is present in automatically sampled images from the domain of application. Strong cloud coverage can partially or completely corrupt the SST signature apparent in the sampled image, rendering the classification task delicate even when manually performed by an oceanographic expert. In this section, the robustness of a CNN-based classification on images corrupted by different degrees of cloud coverage is examined, providing an assessment on its performance on samples encountered in the real domain of application.

A. Cloud Data Augmentation

The EDDIES-AUTO dataset has a distribution with higher cloud coverage values than the EDDIES-EL dataset (Figure 4), and is therefore more depictive of the application domain, albeit being limited by the 80% threshold on CCP. Nevertheless, the noisy labeling of the EDDIES-AUTO dataset creates a discrepancy between visible signatures and image labels. Therefore, using this dataset to test the CNN-based classifier, does not allow for a confident evaluation of their robustness to cloud coverage. To tackle this issue, a test set representative of cloud values is constructed based on the coherent signature images contained in EDDIES-EL(18) test set, whose labels have been validated by experts.

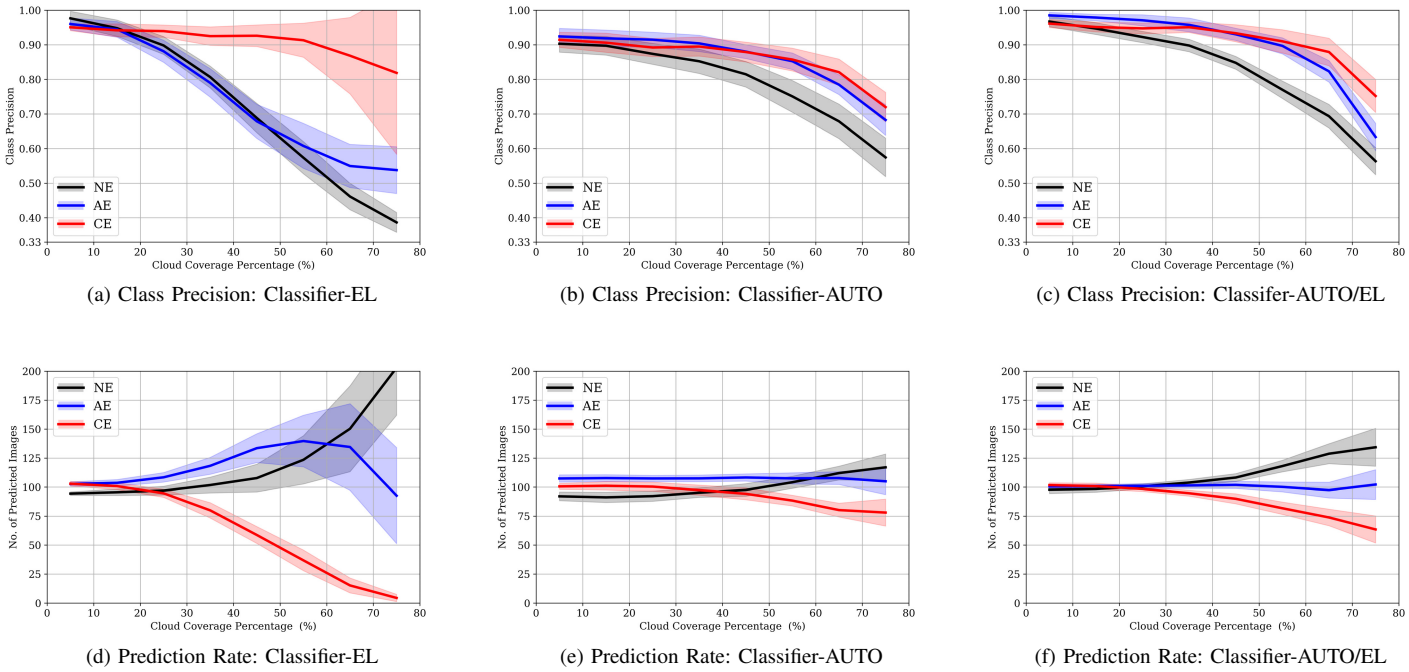
The produced augmented test set, named **EDDIES-CLOUDY** here on, is created by randomly adding to the images contained in the EDDIES-EL test set, cloud masks which are retrieved from the EDDIES-AUTO dataset. This way a test set of images with expert-validated labels is produced, which is also corrupted by realistic cloud patterns, effectively simulating samples from the domain of application of the classifier. The cloud masks are extracted from images corresponding to the year 2018, so that the same cloud patterns appearing in the images used for training the classifiers, corresponding to years 2016/2017, are not repeated in the EDDIES-CLOUDY test sets.

Algorithm 1: Cloud Data Augmentation

```
Input: Datasets: EL{Contains 300 uncorrupted
       images}, AUTO {Contains cloud masks}
Output: CLOUDY {80 sets of 300 images}
initialization;
for cbin = 10-20 to 70-80% {Loop over CCP bins} do
    for rep = 1 to 10 {Repeat different masks} do
        for img = 1 to 300 {Repeat for img in EL} do
            Get uncorrupted img from EL;
            Compute CCP of img;
            while CCP outside of cbin do
                Get random mask from AUTO;
                Apply random mask on uncor. img;
                Compute CCP of corrupted img;
            end
            Save corrupted img to CLOUDY;
        end
    end
end
```

Masks are randomly added to each of 300 images selected from the EDDIES-EL test set, in order to create corrupted images falling in 8 different bins of cloud percentages ($0-10\%$ to $70-80\%$). 10 random corruption realizations for each original uncorrupted image are performed for each of the 8 cloud range bins, creating 80 class-balanced test sets of 300, for a total of 24000 images (see Table I) Algorithm 1 describes the iterative process followed for the test data augmentation.

An example from the EDDIES-CLOUDY dataset is given in Figure 6. An AE (Fig. 6a) and a CE (Fig 6b) example from the EDDIES-EL(18) test set are corrupted with different levels



(a) Class Precision: Classifier-EL (b) Class Precision: Classifier-AUTO (c) Class Precision: Classifer-AUTO/EL

(d) Prediction Rate: Classifier-EL (e) Prediction Rate: Classifier-AUTO (f) Prediction Rate: Classifier-AUTO/EL

Fig. 7. Classifier performance the EDDIES-CLOUDY dataset. The y-axis of figures on the top line represents the Class Precision $C_{i=j}^{pre}$ and on the bottom line the number of predicted images per class. Bold lines and envelopes represent respectively the mean and standard deviation of experiment runs (5-fold training and 10 test sets per CCP bin). Colours represent the performance over the three different classes (black for NE, blue for AE, red for CE). The x-axis represents the mean Cloud Coverage Percentage (CCP) range of the test set (0-10% to 70-80%). Figures in different columns show the performance of different classifiers.

of cloud coverage, here shown for the 10-20%, 30-40%, 50-60% and 70-80% bins. All of the visualized image samples were correctly predicted by the best performing classifier. The level of corruption caused by the cloud coverage on the eddy signature, is not only affected by the percentage of the missing values but also by their positioning on the image. In all of the examples, a part of the information important for the correct classification of the image (core anomaly signature, swirling filaments) is still visible, despite the image corruption, and exploited by the CNN to achieve correct classification.

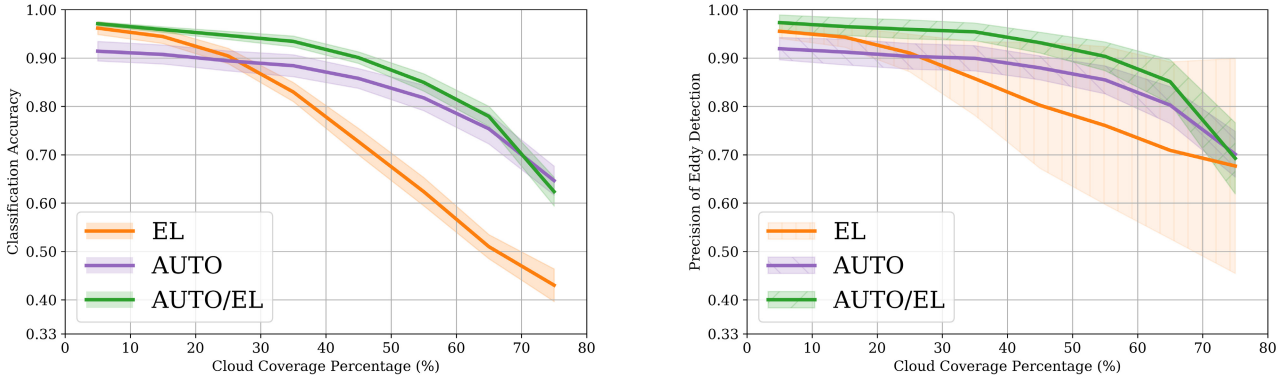
B. Experimental results

The CNN-based classifiers, previously evaluated on the EDDIES-EL test set, are now assesed on their ability to correctly predict the label of cloud-corrupted images contained in the EDDIES-CLOUDY test sets. This is evaluated by computing the Class Precision $C_{i=j}^{pre}$ (i.e. the values corresponding to the diagonal of the normalized confusion matrices) for each of the three classes. For each of the 8 cloud range bins, the values of $C_{i=j}^{pre}$ are calculated by running the 5-fold corss-validated models on each of the 10 test set repetitions. A mean and a standard deviation of the 50 (5x10) received class precision values is thus received, and plotted in the top-line of Figure 7 as the thick line and the envelope respectively, for each of the three classes. A high mean precision on eddy-signature images means that a high fraction of images predicted as AE or CE will have a signature corresponding to their predicted label. A thinner envelope shows convergence between different test realizations.

On the bottom line of Figure 7 the number of predicted images per class is plotted on the y-axis. As before, the thick line and the envelope, represents respectively the mean and standard deviation of experiment runs. As each test set of 300 images is class balanced, 100 images per class suggest a balanced prediction, although that doesn't directly imply that these images were correctly predicted. To assess the performance of each classifier the information of Precision is combined with that of the Predicted numbers.

The Classifier-EL, trained on coherent signature samples, while performing a high precision on test sets with small amounts of CCP, proves incapable of correctly predicting eddy signatures corrupted with strong levels of cloud coverage. This is depicted in Figures 7a in which the initial high precision on AE images in the bin $0-10\%$ of cloud coverage, drops rapidly for increasing values of CCP. The high precision on CE images for high values of CCP is caused by the large drop in the amount of images predicted as CE (Fig 7d). This is also visualized by the large spread of the envelope in the CE precision. However, the EDDIES-EL train set used here, contains images with CCP up to 40% (Figure 4). Nevertheless, the AE class scores an above-random precision (ranging from 70% to 55%) for images with CCP of 40-80%. This demonstrates the ability of the classifier to generalize learning on treating missing values, as it has not encountered images with more than 40% of CCP during the training process.

The Classifier-AUTO, trained on a wider variety of samples with up to 80% of cloud coverage, shows a more robust performance on the EDDIES-CLOUDY test sets. Starting from



(a) Accuracy (b) Precision of Eddy Detection

Fig. 8. Intercomparison of classifier performance on the EDDIES-CLOUDY test sets: Classifier-EL (organge), Classifier-AUTO (purple), Classifier-AUTO/EL (green). The classifiers are compared based on their (a) Classification Accuracy (ratio of correctly predicted images in the EDDIES-CLOUDY test sets) and (b) Precision of Eddy Detection (mean of AE and CE precision). The best performing Classifier AUTO/EL shows very high accuracy and precision (> 0.90) for images with up to 50% of CCP while also being robust to images with even higher amounts of cloud coverage.

the same point of high precisions for the $0-10\%$ cloud coverage bin, this classifier sustains high values of precisions for increasing values of CCP (Figure 7b), while prediction numbers remain almost class balanced (100 images per class), up until the $40-50\%$ cloud coverage bin (Figure 7e). For higher values of CCP, the balance of predicted CE rates drops in favour of more NE predictions.

Precision on the EDDIES-CLOUDY test sets is furtherly augmented by Classifier-AUTO/EL (Figure 7c), which consists of the previous classifier finetuned on EDDIES-EL. When compared to the precisions of Classifier-AUTO, Classifier-AUTO/EL shows a common behaviour on the test set, with yet an increased mean precision of 0.05 on the eddy-classes (AE and CE), and a thinner envelope for the NE class, up until the $40-50\%$ cloud coverage bin. The balance of predicted image numbers (Figure 7f) is also stable (80-120 images per class) up until the $40-50\%$ bin, above which there is likewise a drop in CE and a gain in NE predictions.

The inter-comparison of the three classifiers is more precisely depicted in Figure 8: The precision of eddy detection, that is the mean between the precisions of the red and blue lines in the top line of Figure 7 is shown in Figure 8b. The higher robustness of Classifier-AUTO to Classifier-EL is depicted here. The first has a higher mean and a lower standard deviation of eddy detection precision as values of CCP increase. A further difference in precision of eddy detection of 0.05 is obtained by the Classifier-AUTO/EL up until the $40-50\%$ cloud coverage bin, after which it narrows down to zero.

An inter-model comparison of the classification accuracies (Eq. 4) for increasing CCP ranges on Figure 8a shows essentially the same behaviour: training on the EDDIES-AUTO dataset proves more robust to cloud coverage than training on the EDDIES-EL dataset, while pretraining on EDDIES-AUTO and finetuning on EDDIES-EL, furtherly improves classification accuracy on clear signature images.

Overall, the best performing Classifier-AUTO/EL, achieves a considerable precision of more than 90% for the AE and CE classes and more than 80% for the NE class, for images with up to 50% of cloud coverage. It still shows robust performance for images with up to 80% of cloud coverage, although with a lower precision, with a minimum of 70% mean precision of eddy detection. Robustness on classification of cloud covered eddy signature images is higher for Anticyclonic than Cyclonic signatures, shown by the stable number of AE predictions (Figure 7f). This depicts the fact that cyclones have a more complex, and difficult to classify, signature on the SST.

The pretraining methodology followed here, allows for feature extraction from a large, automatically retained dataset with a high variety of signatures, corrupted by missing values and with presence of label noise.. By finetuning a classifier trained on such images on a smaller subset of coherent signatures with accurate labels, we show that coherent signature, uncorrupted cases can be classified with almost no error, while maintaining a robust performance on images corrupted by missing values.

The Deep Learning approach also achieves a performance which exceeds that of an oceanographic human expert in classifying eddy signatures with strong cloud coverage: the Classifier AUTO/EL proves able to correctly classify eddy signatures with up to 80% of cloud coverage with an increasing amount of error as CCP increases. However, when asked to perform the same task, human experts only selected images with up to 40% of CCP (Figure 4) to assign them as coherent eddy signatures. Such an approach can therefore aid not only in automating a time-costly task but also in achieving a superior performance.

VI. Conclusion

In this study a novel Deep Learning approach is presented to validate the detection of mesoscale eddies from standard altimetry products, using Sea Surface Temperature images. An SST image CNN-based classifier is trained, showing potential to detect eddy signatures, even if the images are corrupted by

a high level of cloud coverage. Such a classifier can be used as a tool to validate or correct standard eddy detections based on altimetry products, which are often uncertain due the to interpolation between satellite track measurements.

A methodology to automatically retain a large dataset of SST image samples, based on altimetric detection region proposal, is first presented. However, a dataset retrieved this way contains a large number of noisy labels, due to complex eddy signatures or to a significant amount cloud coverage. On the other hand, a smaller subset of coherent signature images is labeled by oceanographic expert, in order to extract a reference dataset with coherent eddy signature images.

The best performing SST eddy signature classifier is constructed by pretraining a ResNet18 CNN on a large dataset of automatically retained images, and then fine tuning it on a smaller subset of coherent signature, expert labeled ones. A mean classification accuracy of 97.5% is achieved on a test set containing coherent eddy signatures.

Our classifier achieves significant performance on cloud-covered eddy-signature images, with a precision larger than 90% on Anticyclonic and Cyclonic signature predictions for images having up to 50% of cloud coverage. Furthermore it shows robust performance on images with 80% of cloud coverage, reaching a minimum mean precision of 70% on eddy detection.

It is demonstrated thus that a CNN-based classifier can successfully exploit the high-resolution information available on visible imagery such as the SST, while being robust to strong cloud coverage. From an oceanographic point of view, our classifier can provide an automatic validation of altimetric eddy detections by processing the information in SST images. Moreover, the Deep Learning approach followed here, exceeds the performance of human experts on correctly classifying such images when they are corrupted by a large amount of cloud coverage. Besides, the classification tool can also be exploited to furtherly characterize the complex surface temperature signatures of oceanic eddies.

From a machine learning point of view, a task is presented where pretraining on a large set of complex and corrupted images and finetuning on a set of coherent signature ones, provides a robust training strategy. The ability of a CNN-based classifier to generalize the treatment of missing data is also assessed by corrupting coherent signature images with masks of existing missing value patterns.

The advantages of utilizing high-resolution visible satellite imagery for eddy-signature classification can be extended by using a multi-modal image input. The pattern information contained in all visible satellite imagery such as SST, CHL and SAR can thus be combined. Eventually, object detection and tracking CNN-based methods such as RCNN [40] or YOLO [41] can be employed to construct an independent Deep Learning eddy detection and tracking algorithm on satellite imagery. Besides, future advances in satellite altimetry and imagery, will provide with increasing information of mesoscale and submesoscale eddy signatures.

ACKNOWLEDGMENT

This study has been conducted using E.U. Copernicus Marine Service Information.
We provide data of this study under a creative commons license through https://github.com/evangelos-mos.

REFERENCES

[1] Dudley B Chelton, Michael G Schlax, and Roger M Samelson, "Global observations of nonlinear mesoscale eddies," *Progress in Oceanography*, vol. 91, no. 2, pp. 167–216, 2011.
[2] Zhan Su, Jinbo Wang, Patrice Klein, Andrew F Thompson, and Dimitris Menemenlis, "Ocean submesoscales as a key component of the global heat budget," *Nature communications*, vol. 9, no. 1, pp. 1–8, 2018.
[3] Zhengguang Zhang, Wei Wang, and Bo Qiu, "Oceanic mass transport by mesoscale eddies," *Science*, vol. 345, no. 6194, pp. 322–324, 2014.
[4] Peter Gaube, Dudley B Chelton, Peter G Strutton, and Michael J Behrenfeld, "Satellite observations of chlorophyll, phytoplankton biomass, and ekman pumping in nonlinear mesoscale eddies," *Journal of Geophysical Research: Oceans*, vol. 118, no. 12, pp. 6349–6370, 2013.
[5] Dennis J McGillicuddy Jr, "Mechanisms of physical-biological-biogeochemical interaction at the oceanic mesoscale," *Annual Review of Marine Science*, vol. 8, pp. 125–159, 2016.
[6] Ivy Frenger, Nicolas Gruber, Reto Knutti, and Matthias Münnich, "Imprint of southern ocean eddies on winds, clouds and rainfall," *Nature geoscience*, vol. 6, no. 8, pp. 608–612, 2013.
[7] Shinya Kouketsu, Hiroyuki Tomita, Eitarou Oka, Shigeki Hosoda, Taiyo Kobayashi, and Kanako Sato, "The role of meso-scale eddies in mixed layer deepening and mode water formation in the western north pacific," in *New Developments in Mode-Water Research*, pp. 59–73. Springer, 2011.
[8] Patrice Klein and Guillaume Lapeyre, "The oceanic vertical pump induced by mesoscale and submesoscale turbulence," *Annual review of marine science*, vol. 1, pp. 351–375, 2009.
[9] Laurent Brach, Patrick Deixonne, Marie-France Bernard, Edmée Durand, Marie-Christine Desjean, Emile Perez, Erik van Sebille, and Alexandra ter Halle, "Anticyclonic eddies increase accumulation of microplastic in the north atlantic subtropical gyre," *Marine pollution bulletin*, vol. 126, pp. 191–196, 2018.
[10] Matthew R Mazloff, Patrick Heimbach, and Carl Wunsch, "An eddy-permitting southern ocean state estimate," *Journal of Physical Oceanography*, vol. 40, no. 5, pp. 880–899, 2010.
[11] DJ McGillicuddy, AR Robinson, DA Siegel, HW Jannasch, R Johnson, TD Dickey, J McNeil, AF Michaels, and AH Knap, "Influence of mesoscale eddies on new production in the sargasso sea," *Nature*, vol. 394, no. 6690, pp. 263–266, 1998.
[12] Rémi Laxenaire, Sabrina Speich, and Alexandre Stegner, "Evolution of the thermohaline structure of one agulhas ring reconstructed from satellite altimetry and argo floats," *Journal of Geophysical Research: Oceans*, 2019.
[13] Nadia Mkhinini, Andre Louis Santi Coimbra, Alexandre Stegner, Thomas Arsouze, Isabelle Taupier-Letage, and Karine Béranger, "Long-lived mesoscale eddies in the eastern mediterranean sea: Analysis of 20 years of aviso geostrophic velocities," *Journal of Geophysical Research: Oceans*, vol. 119, no. 12, pp. 8603–8626, 2014.
[14] Romain Escudier, Lionel Renault, Ananda Pascual, Pierre Brasseur, Dudley Chelton, and Jonathan Beuvier, "Eddy properties in the western mediterranean sea from satellite altimetry and a numerical simulation," *Journal of Geophysical Research: Oceans*, vol. 121, no. 6, pp. 3990–4006, 2016.
[15] Artemis Ioannou, Alexandre Stegner, Briac Le Vu, Isabelle Taupier-Letage, and Sabrina Speich, "Dynamical evolution of intense ierapetra eddies on a 22 year long period," *Journal of Geophysical Research: Oceans*, vol. 122, no. 11, pp. 9276–9298, 2017.
[16] Akira Okubo, "Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences," in *Deep sea research and oceanographic abstracts*. Elsevier, 1970, vol. 17, pp. 445–454.
[17] John Weiss, "The dynamics of enstrophy transfer in two-dimensional hydrodynamics," *Physica D: Nonlinear Phenomena*, vol. 48, no. 2-3, pp. 273–294, 1991.
[18] Jordi Isern-Fontanet, Emilio García-Ladona, and Jordi Font, "Identification of marine eddies from altimetric maps," *Journal of Atmospheric and Oceanic Technology*, vol. 20, no. 5, pp. 772–778, 2003.

[19] Dudley B Chelton, Michael G Schlax, Roger M Samelson, and Roland A de Szoeke, "Global observations of large oceanic eddies," *Geophysical Research Letters*, vol. 34, no. 15, 2007.
[20] Alexis Chaigneau, Arnaud Gizolme, and Carmen Grados, "Mesoscale eddies off peru in altimeter records: Identification algorithms and eddy spatio-temporal patterns," *Progress in Oceanography*, vol. 79, no. 2-4, pp. 106–119, 2008.
[21] I Ari Sadarjoen and Frits H Post, "Detection, quantification, and tracking of vortices using streamline geometry," *Computers & Graphics*, vol. 24, no. 3, pp. 333–341, 2000.
[22] Francesco Nencioli, Changming Dong, Tommy Dickey, Libe Washburn, and James C McWilliams, "A vector geometry–based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the southern california bight," *Journal of Atmospheric and Oceanic Technology*, vol. 27, no. 3, pp. 564–579, 2010.
[23] Briac Le Vu, Alexandre Stegner, and Thomas Arsouze, "Angular momentum eddy detection and tracking algorithm (ameda) and its application to coastal eddy formation," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 4, pp. 739–762, 2018.
[24] Artemis Ioannou, Alexandre Stegner, Alexandre Tuel, Briac Levu, Franck Dumas, and Sabrina Speich, "Cyclostrophic corrections of aviso/duacs surface velocities and its application to mesoscale eddies in the mediterranean sea," *Journal of Geophysical Research: Oceans*, 2019.
[25] P Garreau, F Dumas, S Louazel, A Stegner, and B Le Vu, "High-resolution observations and tracking of a dual-core anticyclonic eddy in the algerian basin," *Journal of Geophysical Research: Oceans*, vol. 123, no. 12, pp. 9320–9339, 2018.
[26] Angel Amores, Gabriel Jordà, Thomas Arsouze, and Julien Le Sommer, "Up to what extent can we characterize ocean eddies using present-day gridded altimetric products?," *Journal of Geophysical Research: Oceans*, vol. 123, no. 10, pp. 7220–7236, 2018.
[27] Davide D'Alimonte, "Detection of mesoscale eddy-related structures through iso-sst patterns," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 2, pp. 189–193, 2009.
[28] Changming Dong, Francesco Nencioli, Yu Liu, and James C McWilliams, "An automated approach to detect oceanic eddies from satellite remotely sensed sea surface temperature data," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 6, pp. 1055–1059, 2011.
[29] M Castellani, "Identification of eddies from sea surface temperature maps with neural networks," *International journal of remote sensing*, vol. 27, no. 8, pp. 1601–1618, 2006.
[30] Liangpei Zhang, Lefei Zhang, and Bo Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, 2016.
[31] David Rolnick, Priya L Donti, Lynn H Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, et al., "Tackling climate change with machine learning," *arXiv preprint arXiv:1906.05433*, 2019.
[32] Redouane Lguensat, Miao Sun, Ronan Fablet, Pierre Tandeo, Evan Mason, and Ge Chen, "Eddynet: A deep neural network for pixel-wise classification of oceanic eddies," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018, pp. 1764–1767.
[33] Katharina Franz, Ribana Roscher, Andres Milioto, Susanne Wenzel, and Jürgen Kusche, "Ocean eddy identification and tracking using neural networks," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018, pp. 6887–6890.
[34] Mohammad D Ashkezari, Christopher N Hill, Christopher N Follett, Gaël Forget, and Michael J Follows, "Oceanic eddy detection and lifetime forecast using machine learning methods," *Geophysical Research Letters*, vol. 43, no. 23, pp. 12–234, 2016.
[35] Dongmei Huang, Yanling Du, Qi He, Wei Song, and Antonio Liotta, "Deepeddy: A simple deep architecture for mesoscale oceanic eddy detection in sar images," in *2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)*. IEEE, 2017, pp. 673–678.
[36] Evangelos Moschos, Olivier Schwander, Alexandre Stegner, and Patrick Gallinari, "Deep-sst-eddies: A deep learning framework to detect oceanic eddies in sea surface temperature images," in *ICASSP 2020-45th International Conference on Acoustics, Speech, and Signal Processing*, 2020.
[37] B Buongiorno Nardelli, C Tronconi, A Pisano, and R Santoleri, "High and ultra-high resolution processing of satellite sea surface temperature data over southern european seas in the framework of myocean project," *Remote Sensing of Environment*, vol. 129, pp. 1–16, 2013.
[38] Yann LeCun et al., "Generalization and network design strategies," *Connectionism in perspective*, vol. 19, pp. 143–155, 1989.
[39] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
[40] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.
[41] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.



Evangelos Moschos is PhD student at the École Polytechnique - IP Paris, Laboratoire de Météorologie Dynamique (France). He holds an Engineering Diploma from the Athens Polytechnic (Greece) and a Master in Earth and Environmental sciences from the Paris-Saclay University. His current research is focused on the detection and characterization of oceanic eddies through Machine Learning and Data Analysis approaches. Previously, he conducted research on marine renewable energy resources and coastal processes.



Alexandre Stegner is CNRS researcher and associate Professor at the Ecole Polytechnique. He had a Master in Theoretical Physics (1993), a PhD in geophysical fluid dynamics (1997) at the Ecole Normale (France) and was post-doc at Cambridge University. He his now at the Laboratoire de Météorologie Dynamique (LMD, IPSL). His current research is focused on coastal oceanography using both theoretical, numerical and experimental approaches in addition to remote sensing data analysis. The last years he conducted several experimental projects and took part to four oceanic surveys in Antartica, the North-Atlantic Sea and the Mediterranean Sea. He is the PI of the DYNED-Atlas project which is focused on operational eddy detection and their 3D characterization from remote-sensing and in-situ data-sets.

Olivier Schwander is Assistant Professor at the Sorbonne Université (Paris, France). Since 2015, he is with the LIP6 laboratory, in the Machine Learning for Information Access (MLIA) team. Previously, he was with the L2S at Centrale Supélec (Gif, France) and with the CVML group at the University of Geneva (Switzerland). He completed his PhD thesis in 2013 at the École Polytechnique (Palaiseau, France). His research is focused on the fields of machine learning and signal processing: he is particularity interested in the relationships between statistical learning, Riemaninan geometry and information geometry.

Patrick Gallinari is Professor at the Sorbonne Université and Distinguished Researcher at the Criteo AI Lab. Previously, he was Director of the LIP6 laboratory and responsible for the Machine-Learning and Information Access (MLIA) team. He received his Ph.D. degree in computer science from the University of Compiegne (France). He teaches statistical learning and directs doctoral students on key topics in this field. His work focuses on algorithmic aspects of machine learning, as well as application domains that involve the analysis of semantic data. Gallinari also works on the modeling of dynamic systems and their application in high-stakes areas, such as climate sciences.