WILDetect: An intelligent platform to perform airborne wildlife census automatically in the marine ecosystem

A new non-parametric approach, WILDetect, has been built using an ensemble of supervised Machine Learning (ML) and Reinforcement Learning (RL) techniques. Readers may recall that the first part of the paper was published in May' 24 issue of Coordinates magazine. *We present here the concluding part*

Kaya Kuru
Corresponding author
School of Engineering, University of Central Lancashire, Fylde Rd, Preston, Lancashire, PR12HE, UK

Stuart Clough
APEM Inc., 2603 NW 13th Street, 402, Gainesville, FL 32609-2835, USA[1]

Darren Ansell
School of Engineering, University of Central Lancashire, Fylde Rd, Preston, Lancashire, PR12HE, UK

John McCarthy
APEM Ltd., The Embankment Business Park, Stockport, SK4 3GN, UK[2]

Stephanie McGovern
APEM Ltd., The Embankment Business Park, Stockport, SK4 3GN, UK

5. Implementation of the methodology in splitting and counting (A.4) using the recursive RL technique

Objects can appear in different regions of the image and in different scales. In order to solve this problem, the sliding window method (Fig. 3) is used (Forsyth & Ponce, 2012). It consists of a detection window that slides over an image extracting regions and classifying them. A Gaussian pyramid (Witkin, 1984) (Fig. 3) is also applied to the image to perform a scale search to detect similar objects in different sizes.

A multi-threaded approach was established to speed up the calculations and reduce the processing time. In this multi-threaded approach, jobs are distributed among the resources in the same network, partic- ularly among the multi-core processors, with one job for each core. The user can choose one of the two processing options, either multi-threaded where powerful computing resources can be deployed to perform many tasks at once, or sequentially where operations are performed in order and results can be followed by the user per image. The multi-threaded option reduces the processing time significantly based on the power of the resources used. Some of the resources in use can be stopped to be used for other purposes, and vice versa, new resources can be incorporated into the system while the splitting or counting process is ongoing, using a novel flexible cloud computing approach built in this study.

It is worth noting that datasets are imbalanced -i.e., not uniform within surveys most of the time as mentioned earlier regarding the larger number of negative images (negative class) compared to a smaller number of positive images (positive class). This imbalance is mitigated using an ensemble of ML and RL techniques within the research in two phases of automated data analysis. The selection of the best detectors in the splitting phase is based on the features of the background to discard most of the negative images while aiming to place all the positive images in the positive folder whereas it is based on the features of the targeted objects in the counting phase to count all the objects in the images placed in the positive folder while aiming to discard all the remaining negative images placed in the positive folder during the splitting phase. Four values are measured to assess the obtained results, namely,

$$\begin{aligned} Sensitivity{=}Se{=}TruePositiveRate(TPR){=}TP/(TP+FN), \\ Specificity = Sp = TrueNegativeRate(TNR) = TN/(TN+FP), \\ Accuracy = Acc = (TP+TN)/(P+N) = (TP+TN)/(TP \\ +TN+FP+FN),\ Precision{=}Pr{=}TP/(TP+\mathbf{FP}) \end{aligned} \quad (1)$$

The first three values — *Se*, *Sp*, and *Acc* — are explained in Section 6 in detail based on the data analysis of the particular approaches. Pr is mainly employed to identify the class imbalance problem and assess how imbalanced data in favour of ‘‘negative images’’ that may lead to large FPs is influencing the results. More specifically, this assessment helps to understand (i) if the high values of *Se*, *Sp* and *Acc* are biased and most importantly (ii) if the two phases of using an ensemble of learning techniques help alleviate the bias regarding the improvement in Pr through obtaining the final counting results. The low values of Se, e.g., < 0.80, require the implementation of cost-sensitive analy- sis (CSA), as we conducted in our previous research in Kuru et al. (2013) to get more reliable improved results. In CSA, classes have different costs associated with them using weights with respect to the number of instances; the classes with fewer instances, i.e., under- represented classes (positive cases in this research) are assigned higher costs (i.e., adding cost-sensitivity, e.g., P:N = 10:1) to reduce the number of false predictions, particularly in favour of the class with less number of instances, and consequently increase the reliability of the results related to that class by assigning different penalties to misclassification of samples (Kuru et al., 2013) in which there is a trade-off between *Se* and *Pr*.

5.1. Implementation of the platform in splitting

Most of the time, more than 95% of images in a survey contain no targeted objects, and therefore this phase of the implementation aims to separate out the images with no targeted objects in a reduced overall processing time. Strictly speaking, the main objective of this phase is to perform the best splitting between negative and positive images based on the parameters specified in Section 4.2.3. The negative images are placed in the negative folder and the positive images are placed in the positive folder. Then, the images in the positive directory are analysed in detail to locate all targeted objects, which is explained in Section 5.2. The methodology selects a set of detectors for each feature extraction technique to deploy during the splitting process based on the particular characteristics and specific patterns of the images in surveys. This step is explained in Section 5.1.1. Then, how the splitting is performed is explored using these selected detectors in Section 5.1.2.

5.1.1. Pattern recognition and specification of the best feature extraction detectors for splitting using RL (A.4.1)

The methodology chooses the best detectors regarding separating negative images from positive images successfully based on the texture patterns and characteristics of the images in the surveys using the user- model-data interaction as illustrated in Fig. 3, A.4.1. The components of the recursive RL algorithm employed in this phase are demonstrated in a broader perspective in Fig. 11 and the main steps are explained as follows.

First, a very small subset of the negative images (i.e., 5–10) representing the whole of the negative images (i.e. background) in the survey is selected by the user. The characteristics of this very small set play an important role in determining the best convenient detectors. Therefore, the user is expected to choose blank images that have diverse background textures in the survey. For instance, at least a blank image taken from each camera mounted on the aeroplane and blank images taken from different time intervals should be placed in this set in order to represent the background characteristics of the whole survey. Alternatively, processing of the images from different cameras or in different time intervals – subsets of surveys – can be conducted separately, which can increase the efficacy of the platform further.

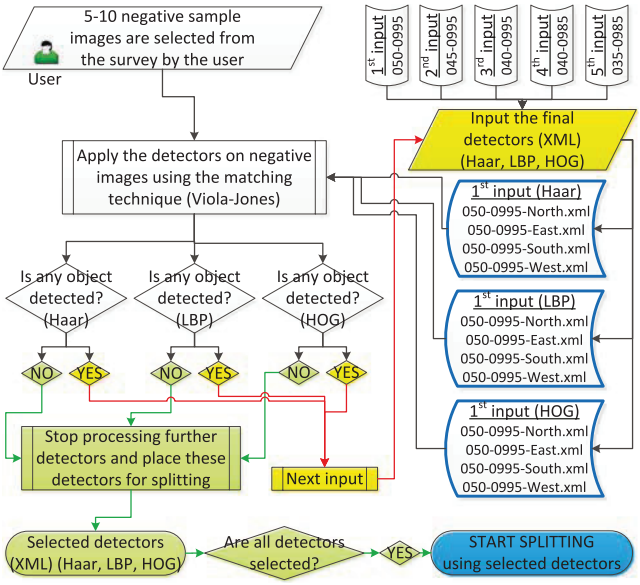


Fig. 11. Use of Reinforcement Learning (RL) for selecting the best detectors for splitting.

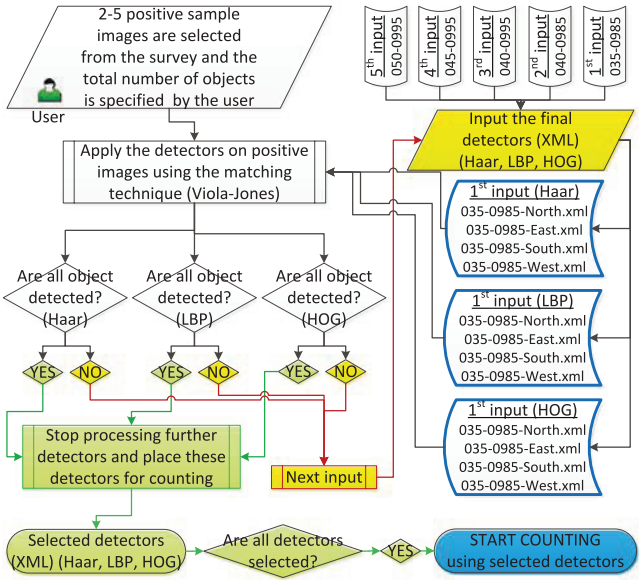


Fig. 12. Use of Reinforcement Learning (RL) for selecting best detectors for counting.

Second, the blank images selected by the user are processed by the approach to determine the best detectors for each technique based on the observed *Sp* (i.e., TN/(TN+FP)) values. In this process, a screening test is performed with preferably higher *Sp* values to increase the chance of placing images with no targeted object in the negative folder. In other words, the FP cases are reduced to a minimum resulting in very high *Sp* values with an ability to correctly place the negative images in the negative folder and this means that if an image is tagged as negative, it is a high probability that there is no object in that image. The RL algorithm makes the detectors run on the sample negative images fed by the user using the Viola–Jones matching technique to single out the successful detectors for splitting based on the characteristics of the background texture. This process starts from the detectors with the highest threshold values (i.e., 050–0995 in Table 2) that may result in many FPs reducing *Sp* whereas the background has a complicated texture. However, no FP may be obtained if the background has a clear texture. This iterative process using predetermined nominated detectors (Fig. 9) proceeds (Fig. 11) until no FP is obtained per detector where *Sp* is 1. In other words, the process stops per detector where *Sp* is 1 and the detector is selected at this stage in which a satisfactory pattern is observed and learned by the system. Otherwise, the last detectors with the smallest threshold values are processed where the *Sp* may be slightly smaller than 1 and they are selected for splitting.

Finally, the methodology determines the most suitable detectors for each technique (i.e., Haar, LBP, HOG) through the detector sets trained previously as depicted in Table 2 that are above the green line in Fig. 9. The results of the RL process for the 13 surveys regarding the selection of the detectors for splitting are explained in Section 6.1.

5.1.2. Object recognition and splitting (A.4.2.Phase1)

The detectors determined by the RL approach at the start of the splitting process as explained in Section 5.1.1 are utilised in this phase. The methodology makes these detectors run on all the images using the Viola–Jones matching technique and the images are placed in the negative directory if they are specified as negative; in other words, these are the images in which no object is detected by any of these detectors. The images are readily placed into the positive directory when an object is detected by any detector without screening the image for other objects using the remaining detectors. The main aim is to increase *Sp* by reducing FPs with respect to each technique, but to increase *Se* using 3 techniques at the same time by reducing FNs regarding the number of positive images (see Fig. 10). The higher the number of objects in an image, the more likely that the image will be put into the positive directory. The splitting phase was evaluated on several surveys (Fig. 4III) and the results (Table 3) are explained in Section 6.1.

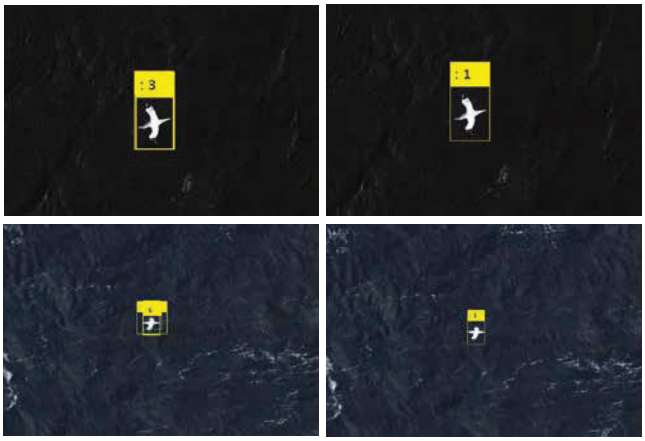


Fig. 13. Two magnified gannet objects detected by several detectors (left) and counted only once (right).

5.2. Implementation of the platform in counting objects

In the splitting phase, the application places any image into the positive directory when an object is detected without screening the image for other objects using the remaining detectors. In this way, the processing time of the splitting is reduced significantly. On the other hand, the aim of the counting phase is to detect every targeted object in images placed in the

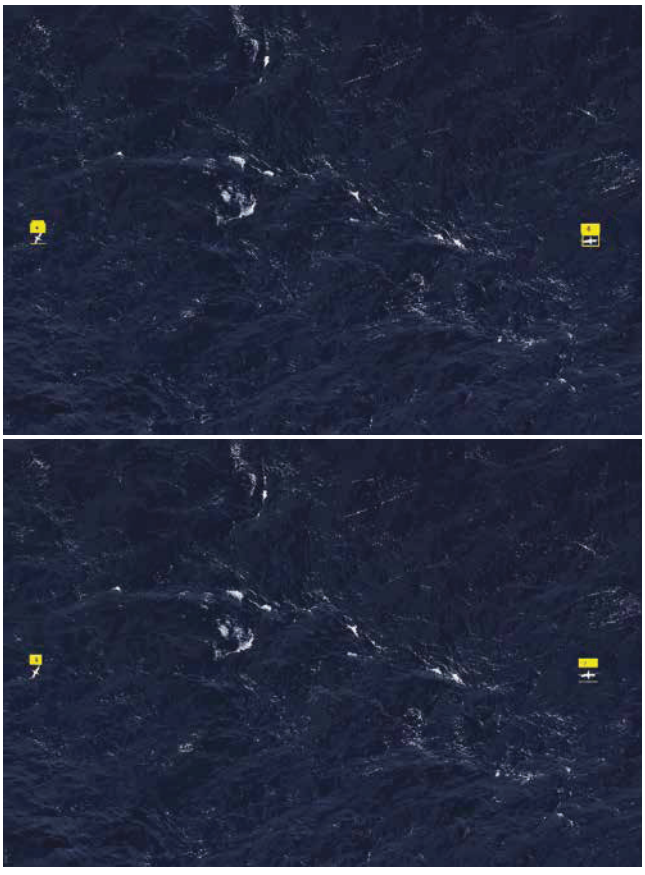


Fig. 14. Gannets detected by the convenient detectors: Multiple gannet objects detected by several detectors (top) and counted only once (bottom).

Table 3 Accuracy rates of the snag dataset based on the trained files of 4 different parameters for splitting images into positive and negative categories: all snags are recognised successfully by the training parameters, FAR = 0.50 and TPR = 0.995 with the combination of 3 techniques.

									Selected parameters for three techniques					
									Haar		LBP		HOG	
Surveys	NegativeImages	PositiveImages	GannetsTot	#	TP	TN	Se	Sp	FAR	TPR	FAR	TPR	FAR	TPR
Survey	1	100	31	31	30	99	0.968	0.990	0.350	0.985	0.400	0.995	0.350	0.985
Survey	2	100	7	7	7	93	1.000	0.930	0.350	0.985	0.350	0.985	0.350	0.985
Survey	3	100	3	3	3	97	1.000	0.970	0.350	0.985	0.400	0.995	0.400	0.985
Survey	4	100	2	2	2	92	1.000	0.920	0.350	0.985	0.350	0.985	0.350	0.985
Survey	5	100	1	1	1	99	1.000	0.990	0.500	0.995	0.500	0.995	0.500	0.995
Survey	6	100	2	2	2	99	1.000	0.990	0.350	0.985	0.400	0.995	0.400	0.985
Survey	7	100	10	10	9	99	0.900	0.990	0.350	0.985	0.400	0.995	0.400	0.985
Survey	8	100	3	3	3	100	1.000	1.000	0.400	0.995	0.400	0.995	0.500	0.985
Survey	9	100	1	1	1	99	1.000	0.990	0.500	0.995	0.500	0.995	0.500	0.995
Survey	10	100	1	1	1	97	1.000	0.970	0.400	0.995	0.400	0.995	0.350	0.985
Survey	11	100	10	10	10	98	1.000	0.980	0.350	0.985	0.400	0.995	0.450	0.995
Survey	12	100	3	3	3	99	1.000	0.990	0.400	0.995	0.500	0.995	0.400	0.995
Survey	13	500	202	256	196	484	0.970	0.968	0.350	0.985	0.350	0.985	0.400	0.995
Split Avg							**0.988**	**0.975**						

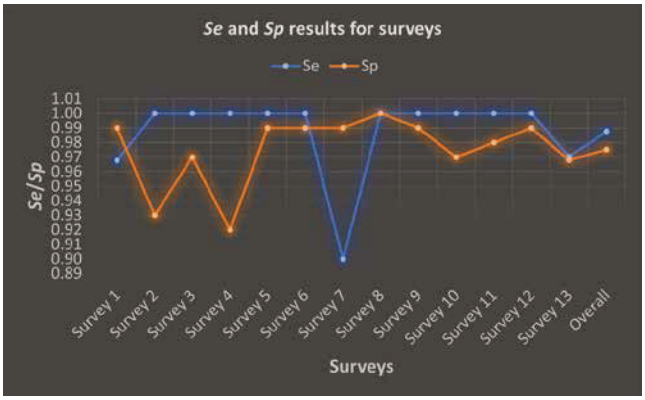


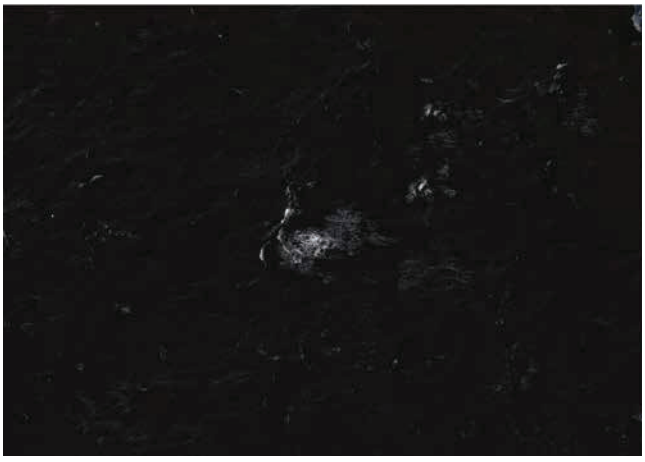
Fig. 15. Visualisation of Se and Sp results of the surveys in Table 3 for splitting images into positive and negative categories.

positive directory. New detectors are selected to complete this task using a similar recursive RL approach explored above, but differently as explained in Section 5.2.1 in order not to miss any targeted object in the positive images.

5.2.1. Pattern recognition and specification of the best feature extraction detectors for counting using RL (A.4.1)

The methodology chooses the best detectors regarding counting ob- jects in the images placed in the positive folder based on the particular patterns and characteristics of the objects in those images as illustrated in Fig. 3, A.4.1. The main objective of this phase is to detect and count all the targeted objects successfully. The components of the recursive RL algorithm employed in this phase are demonstrated in Fig. 12 and the main steps are explained as follows.

(a) Example 1: the object with very difficult background texture; Bottom centre.



(b) Example 2: the object during diving in exactly different shape; at the middle left.

Fig. 16. Examples of gannet objects in whole images not detected by the trained classifiers.

First, a very small subset of the positive images (i.e., 2–5) in the survey are selected and all the objects in these images are outlined with a bounding box along with the targeted object counts per image by the user upon the interface provided in the application during the selection of a subset of the negative samples at the start of the survey analysis as mentioned in Section 5.1. In other words, the detectors for both splitting and counting are designated by the recursive RL approach before the survey analysis starts. In this way, the methodology carries out the counting process automatically after the splitting phase is completed.

Second, these selected positive images are processed with respect to the user-specified objects by the RL approach to determine the best de- tectors for each technique based on the observed *Se* (i.e., TP/(TP+FN)) values. In this step, an object recognition test is performed with preferably higher *Se* values to increase the chance of detecting a targeted object in the positive folder. In other words, the FN cases are reduced substantially with respect to the targeted objects, preferably to zero, resulting in very high *Se* values with an ability to correctly detect the objects in the images. The RL algorithm makes the detectors run on the sample positive images using the Viola–Jones matching technique to single out the successful detectors for counting by referencing the users' object inputs from the selected positive images. This time, different from the splitting phase, the process starts from the detectors with the lowest threshold values (i.e., 035–0985 in Fig. 9) that may result in many FNs which may reduce *Se*. This iterative process proceeds until no FN is obtained per detector where *Se* is 1. In other words, the process stops per detector where *Se* is 1 and the detector is selected at this stage in which a satisfactory pattern is observed and learned by the system. Otherwise, the last detectors with the highest threshold values are processed where the *Se* may be slightly smaller than 1 and they are selected for counting. Additionally, the last detector with the highest threshold values may result in several FP where the images have complex backgrounds, which may reduce *Sp* of the system at this stage. But, the objective is to detect all targeted objects successfully with a high *Se*, preferably 1, as specified earlier even though compromising *Sp* slightly. The use of multiple designated detectors at a time in a collective way aims to ensure a high *Se* — one of the other two detectors can detect an object if it is missed by a detector.

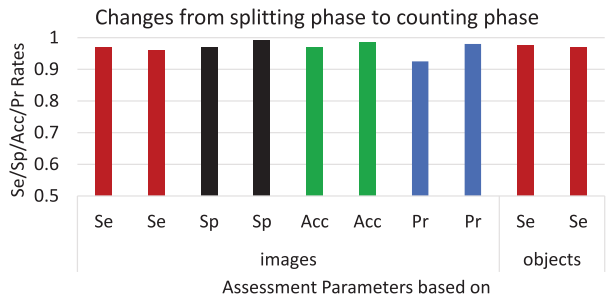


Fig. 17. Performance of the counting phase with respect to splitting regarding Survey 13 presented in Table 3.

Finally, the methodology places the most convenient detectors for each technique (i.e., Haar, LBP, HOG) through the detector sets trained previously as depicted in Table 2 that are above the green line in Fig. 9. The results of the RL process for the last survey (Table 3) regarding the selection of the detectors for counting are explained in Section 6.2 (Fig. 4III).

5.2.2. Object recognition and counting of objects in surveys (A.4.2.Phase2)

In this phase, the aim is to detect all targeted objects in the positive images with an increased Se by giving several FPs if necessary, in order not to miss any targeted objects. Every image in the positive directory is processed by the Viola–Jones technique using each designated detector and objects are tagged wherever they are detected and coordinates for one or more rectangular ROI (coloured bounding box around the recognised object (e.g., Fig. 13)) are returned. These coordinates are mainly utilised for both counting each detected object once using the non-maximum suppression technique (Fig. 3), as explained in the following paragraph, and cropping the tagged objects automatically for further analysis.

Due to the fact that detection windows overlap each other, the same object can be counted more than once. The main reason for this is that 12 detectors are applied for detecting objects in any direction, which may detect and specify an object several times. For instance, a gannet object is detected by 3 detectors and consequently counted 3 times and likewise, another gannet object is detected by 5 detectors and counted 5 times in Fig. 13(left). The non-maximum suppression technique, in which windows with a local maximum classifier response suppress nearby windows (Forsyth & Ponce, 2012), is employed to count the same object only once as shown in Fig. 13(right). Two gannets are located by the detectors several times and they are counted as 2 objects in a whole image in Fig. 14 using the non-maximum suppression technique.

6. Results for splitting and counting

6.1. Results for phase 1: splitting

The methodology was evaluated on each of the 13 surveys (Fig. 4III) in which gannet objects exist to observe the success rates of splitting.

The number of gannets and the negative images along with the success rates are presented in Table 3 and Fig. 15 with respect to the surveys. The detectors selected by the system for each feature extraction tech- nique are shown in the column titled "selected parameters for three techniques" of Table 3 regarding each survey. For instance, the values FAR = 0.350 and TPR = 0.985 for survey 1 correspond to the four detectors, namely, 0.35–0.985-North.xml, 0.35–0.985-East.xml, 0.35– 0.985-South.xml, and 0.35–0.985-West.xml determined for the Haar technique.

The large images with gannets that were not detected as positive are presented in our technical report — MarineObjects_ Gannet_Supplement_4.pdf in the supplementary materials. Additionally, the blank images with no gannets that were detected as positive are pre- sented in our technical report — MarineObjects_Gannet_Supplement_ 5.pdf as well. The average *Se* of the system concerning the *Se* results of 13 surveys based on the number of images (i.e., the column titled *Se* in Table 3) is 0.988. The average *Sp* of the system concerning the *Sp* results of 13 surveys based on the number of images (i.e., the column titled *Sp* in Table 3) is 0.975.

Se – correctly-detected-positive-images/all-positive-images – shows the power of the techniques used in the paper in giving assurance that if an image is tagged as a positive image, with at least one bird, that image most probably comprises at least one bird with a belief, an average confidence level of 0.988. In other words, we can conclude that there is a chance that this image does not comprise a bird with an average confidence level of 0.012, which is significantly low in a sense of showing high confidence when a decision is given about an image that is determined as ''positive''. How the splitting process is implemented successfully can be noticed in Table 3 in the column ''TP'' compared to the column ''Positive Images''. Almost all positive images with birds are placed in the positive folder for further processing (e.g., counting). This success is clearer in Survey 13 with many negative images and positive images with multiple targeted objects. On the other hand, *Sp* – correctly-detected-negative-images/all-negative-images – shows the power of the techniques in giving assurance that if an image is tagged as a negative image, that image most probably comprises no bird with a belief, an average confidence level of 0.975. In other words, we can conclude that there is a chance that this image is not a bird-free image with an average confidence level of 0.025, which is significantly low in a sense of showing high confidence when a decision is given about an image that is determined as ''negative''.

6.2. Results for phase 2: counting

The last survey – Survey 13 – in Table 3 was used to evaluate the viability of the object recognition and counting phase (Fig. 4III). The reason for selecting this survey is that it is the largest survey and has multiple gannets in some of the images, which can help quantify the obtained results more realistically with less bias. The detectors with the parameters of 0.40–0.995, 0.45–0.995, and 0.40–0.985 were selected respectively for Haar, LBP, and HOG techniques by the recursive RL technique. These parameters are bigger than the parameters selected by the RL algorithm in phase 1 (i.e., splitting with 0.35–0.985, 0.35– 0.985, and 0.40–0.995) as explained in Section 5.2.1. This shows that different detectors may be chosen for different purposes (i.e., splitting and counting) by the same recursive RL technique using two different approaches to realise the two different objectives, higher *Sp* with a high level of splitting and higher *Se* with a high level of object detection respectively. 248 objects out of 256 objects in 202 images were tagged as positive successfully resulting in a *Se* value of 0.968 which is 0.976 during the splitting phase regarding the number of objects. 6 objects are missed during the splitting phase within 6 different positive images (Table 3) and 2 objects within 2 different images are missed here during the counting phase. The two objects not detected by the application are shown in Fig. 16. The difference in *Se*, i.e., 0.08 (0.976–0.968), is not found to be significant ($p > 0.01$ using the statistical paired t-test). *Se* – correctly-detected-positive-objects/all-positive-objects – shows the power of the techniques used in the paper in giving assurance that if an object is tagged as a positive gannet, that object most probably is a gannet with a belief, a confidence level of 0.968. In other words, we can conclude that there is a chance that this object is not a gannet with a confidence level of 0.032, which is significantly low in a sense of showing high confidence when a decision is given about an object that is determined as ''positive''.

On the other hand, it could be highly informative to compare the results between the splitting and counting phases based on the number of images rather than the number of objects for assessing how the counting phase is performing in further splitting, particularly, in handling the imbalanced data. 194 images out of 202 images were tagged as positive successfully resulting in a *Se* value of 0.960 which is 0.970 during the splitting phase. 6 positive images are missed during the splitting phase (Table 3) and 2 positive images are missed here during the counting phase. The difference in *Se*, is not found to be significant ($p > 0.01$, i.e., 0.1578). There were 4 FPs where waves were shaped similarly to the shape of gannets in the snags included in the training process. 496 out of 500 negative images are detected correctly as TN after the counting process whereas it is 484 during the splitting phase for survey 13 (Table 3). This results in a *Sp* value of 0.992 whereas it is 0.968 during the splitting phase based on the number of images. The difference, 0.024 (0.992–0.968), was found to be statistically significant ($p < 0.01$, i.e., 0.0005015) considering the number of negative images, i.e., 500, using the statistical paired t-test. The reduction of FP regarding the increased *Sp* is highly important, particularly for the surveys that are comprised of a great majority of bird-free negative images (e.g., >%95) leading to imbalanced data distribution and bias on the obtained results as elaborated above in Section 5. Moreover, *Pr* is increased slightly from the splitting *Pr*, 0.925 (TP / (TP + FP) = 196 / (196 + 16) = 0.925), to the counting *Pr*, 0.980 (196 / (196 + 4)) based on the number of images, which is statistically significant ($p < 0.01$). Finally, overall *Acc* rises from 0.969 ((196 + 484) / (202 + 500)) to 0.985 ((196 + 496) / (202 + 500)) based on the number of images, which is statistically significant ($p < 0.01$) as well. The results are presented in Fig. 17 for better visualisation. To summarise, the techniques used during the counting phase provide (i) a successful way of object detection leading to counting objects correctly, and (ii) further successful splitting leading to discarding the FP images substantially as well. The high value of *Pr* indicates that there is still a large room to perform CSA by which *Se* can be increased while compromising *Pr* slightly if *Se*, resulting from the minority positive class, is not deemed as satisfactory

(<%95 for our research) due to the imbalanced data class distribution that may cause unreliable results. These outcomes demonstrate that the two phases of using ensemble techniques proposed in this study can work successfully in performing the offshore bird censuses even without needing to perform CSA (Section 5) and most importantly, the proposed approaches can be generalised to the automated counting of broader species.

A comprehensive field test with a completely new survey has not yet been completed. The system was validated by two field experts from APEM using a completely new evaluation dataset with a decent number of example species (i.e., 20 positive images with 21 gannets and 500 negative images (Fig. 4IV) taken from other recent surveys at the end of the project at the UCLan Intelligent Systems Laboratory before a comprehensive field survey is conducted using the established system in this study. There was a single juvenile gannet in this dataset and this was not detected as positive where all other gannets (i.e., 21) were detected correctly without missing a single one and without producing any FP by excluding other types of flying birds such as terns (i.e., 2 terns) and shearwaters (i.e., 11 shearwaters) as TN. The reason for not detecting this juvenile gannet is that the features of the juvenile gannets seem significantly different from their mature ones, e.g., first-year juvenile gannets are almost black, and subsequent sub-adult plumages show increasing amounts of white (SeabirdCentre, 2017). It is noteworthy to mention that there were no juvenile gannets either in our training nugget dataset or in our surveys. We suggest the construction of new classifiers specific to juvenile gannets to increase the chance of their detection. The correct labelling of the other images with other types of species (e.g., terns and shearwaters) as TN indicates that the classifiers established for gannets perform perfectly for detecting gannets as anticipated, and particular classifiers need to be established for the other species as mentioned in Sections 4.2.1–4.2.3 to identify them. This outcome confirms that the designed techniques in this research enable the automated classification of multispecies and counting them since every targeted species has its particular classifiers.

7. Discussion

Prevention of regional and global extinction of species during industrial developments and environmental changes (e.g., climate change, habitat loss with rapid urbanisation and coastal disturbance, toxic pesticide use) is a social responsibility from a conservationist point of view. In this sense, a species whose population is in decline needs to be identified urgently and should be protected with higher priority before it is too late. Data science is considered by Gibert et al. (2018) as the multidisciplinary field that combines data analysis with data processing methods and domain expertise, transforming data into understandable and actionable knowledge relevant to informed decision-making. Interdisciplinary efforts will help precipitate the shift towards increased use of computer-automated aerial photographic species census techniques (Chabot & Francis, 2016). Within this context, this study by bringing domain expertise and data scientists together in a fruitful collaborative team aims to develop a novel environmental platform for monitoring the marine ecosystem and performing bio censuses in an automated manner at regular intervals to track changes in a particular species population. Birds are sensitive indicators of biological richness, environmental health, ecosystem integrity, and environmental trends and fulfil many key ecological functions; they contribute to our understanding of natural processes (Bibby et al., 1998; Burger & Gochfeld, 2004; Morelli, 2015). Extinction of the passenger pigeon (Ectopistes migratorius), once likely the most numerous bird on the planet, provides a poignant reminder that even abundant species can go extinct rapidly (Rosenberg et al., 2019). Continuously, automated monitoring of species is of paramount importance which requires the use of advanced tools equipped with effective intelligent surveillance techniques. In this sense, a new non-parametric platform composed of an ensemble of supervised ML and RL techniques, WILDetect, is built to segment, split and count maritime species, in particular, birds for performing automated censuses in a highly dynamic maritime environment. Typically, parameter selection to mitigate the variations in datasets and obtain the best possible outcome in an intelligent autonomous system are carried out by users based on several predictions and trials and the success rates of the systems are highly associated with the wisdom of this assumption and implementation of trials correctly, which is a non-trivial task, specifically for ordinary users. Furthermore, there is no single best approach that suits every type of problem space based on the changing characteristics of datasets (e.g., quantity, quality, attributes) and many other environmental dynamics (e.g., different seasons and time zones, different weather conditions, different settings and types of image-capturing technologies). It can be concluded based on the preliminary tests, as elaborated in Section 4.1, and the current research attempts in the literature to count species and classify multispecies, as elaborated in Sections Section 2, that (i) there is no computer-automated study that analyses datasets of small species acquired from the photogrammetry settings using small aeroplanes to survey very large areas in shorter times compared to the other on-ground, ships or UAS platforms, (ii) The most popular learning technique, the so-called DNN, yield the precision values ranging from 60% to 97.66% for bird detection using the aforementioned platforms, (iii) Large data samples with distinctive features (e.g., species that contrast distinctively from image backgrounds) may result in high accuracy rates in using DNN, (iv) The inner states of the DNN approaches are accepted as black boxes by the research community and these approaches do not let the researchers intervene in their inner states which may help increase their efficacy if they do not produce desired outcomes, (v) The misclassification of multispecies is high using DNN and clustering techniques if data instances in different groups resemble each other too closely as seen in bird species. In the proposed intelligent platform – within a dynamic approach

that adjust its parameters according to the features of targeted objects, their background and the targeted accuracy rate – the best possible parameters, resulting in the best outcome, are chosen by the platform itself through the automated selection of pre-trained models, in which the parameters are instilled, using the user-model-data interaction solution that is implemented within a new recursive RL technique for mitigating the highly dynamic characteristics of the maritime ecosystem as well as the concerns mentioned with the aforementioned approaches. Additionally, the use of multiple trained models at a time, focusing on different features, ensures a high accuracy rate where one of the other two detectors/models can detect an object if it is missed by the other detector/model in use as elaborated in Section 4.2.2.

The validation of the platform, as summarised in Fig. 4, has been performed on several aerial maritime domains resulting in successful empirical evidence for the viability of the model. During the splitting phase, a positive image is most likely to be placed in a positive folder if there are several targeted objects in that image. Strictly speaking, there is a very high probability that one of the objects in an image will be detected by at least one of the three techniques using 12 detectors regarding the orientation of the objects during the splitting phase. Therefore, the more targeted objects in images, the higher the success rate of splitting. We would like to emphasise that the success rates are very high even though there is mostly only one gannet object in images in the surveys in this study (Table 3 and Fig. 15). The main reason for not detecting 2 of the gannet objects depicted in Fig. 16 in the second phase (i.e., recognition and counting) is that one of them does not look like the shape of a gannet in the training set, because, it is in the diving position, while the other one was not detected because of the very complex background texture behind the gannet. The training snag set should have more similar object types to be able to represent the real-world better and in this way, these types of objects are not missed by the trained detectors.

The trained files established for the gannets do not detect other types of birds as TP, such as common gulls, shearwaters, or terns. Therefore, if the objective is to count other types of birds as well, all bird types should be trained independently as explained in Section 4.2.2 to increase the accuracy of the system. In this way, the classification of other bird species becomes possible using the specific classifiers trained for these types of species. The methodology developed for the detection, splitting, and counting of birds, particularly gannets, in large-scale aerial images may be used for the UK marine gannet census since the most important nesting ground for northern gannets is in the UK with about half of the world's population (55.6%) (JNCC, 2015). Furthermore, multiple types of species of interest can be classified and counted at once using the methodology (as concluded in Section 6.2) with the multiple classifiers that can be obtained as explained in Section 4. It is worth mentioning that the methodology can be expandable with more feature extraction techniques in addition to the feature extraction techniques (i.e., Haar, LBP, and HOG) that we employ in this study.

Given the current pace of global environmental change, quantifying change in species abundances is essential to assess ecosystem impacts. Evaluating the magnitude of declines requires effective long-term monitoring of population sizes and trends, data that are rarely available for most species (Rosenberg et al., 2019). Models perform better as they are attributed to the results of more realistic/recent-data analysis on particular domains. With the proposed platform, current labour-intensive and costly censuses of species conducted in longer time intervals can be replaced with cost-effective and more robustly automated computerised systems and can be repeated in an automated manner at regular intervals. Hence, cycles of the census can be conducted more frequently in shorter intervals over time, and incorporation of near-real-time results along with the prior results (e.g., population fluctuations) attributed to shorter intervals into these models paves the way for developing more effective ecological environmental models with realistic data trends and future projections. This, in turn, can boost the decision-making and prediction abilities of these data-driven simulation models, particularly, about the ecological footprint of human activities on the environment, specifically, on areas/offshores that are being turned into industrial zones, for both assessing the likely impact of the industrial developments on nature (e.g., habitat associations) and constraining/alleviating their potential damaging effects.

8. Conclusions and future work

Advanced tools, enabling effective monitoring of species, are needed to observe and predict the likely effects of environmental changes on species, mostly caused by indispensable industrial developments to take urgent proper actions, e.g., rebuilding natural habitats to maintain/increase species counts. Birds have been demonstrated to serve as good indicators of biodiversity and environmental change and as such can be used to make strategic conservation planning decisions for the wider environment (Bibby et al., 1998). Based on the literature reviewed in Chabot and Francis (2016), a major shift to computer- automated aerial photographic bird censusing is not yet underway and investigators are encouraged to study for potential approaches to automate animal detection and enumeration in aerial images. In this study, a novel supervised ML platform supported by a new recursive RL approach using several off-the-shelf feature extraction techniques and a matching algorithm were developed to conduct marine bird censuses in an automated manner. In the proposed approach, the uncertainties within a highly dynamic maritime environment and inconsistencies/variations in the characteristics of datasets attributed to the diverse sets of image-capturing technologies used in the maritime ecosystem have been mitigated using the recursive RL technique with the user-model-data interaction. In this technique,

the most available parameters based on the characteristics of the dataset to be analysed are selected within the platform by the direction of the user at the start of the analysis to result in the best possible outcome. In this way, the developed approach adapts itself to the characteristics of the dataset concerning targeted objects and background and the environmental dynamics, which leads to resulting a desired solution to the current problem space in hand. The methodology has been evaluated and validated by field experts using several surveys and datasets that are independent of the dataset used in the training phase as outlined in Fig. 4. Experimental results on many aerial surveys demonstrate that the proposed methodology is effective and efficient in the detection and segmentation of targeted objects in the maritime ecosystem. The efficacy of the proposed approach can be increased as the techniques are trained with larger datasets for particular species.

The outcome of the study is expected to benefit the entire environmental modelling community. In particular, the proposed techniques can shed light on similar object detection implementations in finding the best possible parameters for analysis in an automated manner by employing the user-model-data interaction solution. Moreover, the platform can be employed to detect all types of birds after these species are pre-processed and trained, as mentioned in Sections 4.2.1–4.2.3. The outcomes elaborated in Section 6 demonstrate that the proposed approaches can be generalised to the automated counting of a broader number of species in a given area and these automated approaches can help track population changes of particular species at different specific locations on a regular basis with a true picture. Strictly speaking, it can be primarily deployed by environmentalists, researchers, authorities, and policymakers to monitor the marine ecosystem for fulfilling their goals effectively.

Within a holistic view, we aim to study other bird species and other marine species (e.g., turtles) as well as man-made maritime objects (Kuru et al., 2022) to be able to observe the bio marine ecosystem with the possible environmental footprint in the short, mid, and long-term. Moreover, the automatic classification of maritime ecosystems based on a variety of species will be in our future plans to support all types of environmental models with near-real-time information with multiple species.

9. Limitations of the study

The established environmental platform can work for other bird species, but using the specific detectors that can be trained for each species as explained in Sections 4.2.1 and 4.2.2. The higher the quality of the datasets representing the real environment, the higher the accuracy rates. We aim to share our results with other papers about our ongoing research on multispecies census of other species such as shearwaters, terns, gulls, scooters, fulmars.

CRediT authorship contribution statement

Kaya Kuru: Conceptualization, Methodology, Software, Writing – original draft. Stuart Clough: Validation, Resources, Data curation. Darren Ansell: Investigation, Project administration. John McCarthy: Validation, Resources, Data curation. Stephanie McGovern: Valida- tion, Resources, Data curation.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank APEM Ltd for providing the maritime aerial surveys used in this study. Anyone who wants to use any of the images either in this manuscript or in our technical reports in the supplementary materials should request permission from APEM Ltd, which retains the copyright and intellectual property contained within. The authors would like to thank the anonymous reviewers for their constructive comments which doubtlessly improved the quality of the paper significantly.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eswa.2023.120574.

Endnotes

[5] MATLAB and Simulink are registered trademarks of the MathWorks, Inc. See www.mathworks.com/trademarks.

[6] https://uk.mathworks.com/help/vision/ug/train- a- cascade- object- detector.html.

References

Akçay, H. G., Kabasakal, B., Aksu, D., Demir, N., Öz, M., & Erdoğan, A. (2020). Automated bird counting with deep learning for regional bird distribution mapping. *Animals*, *10*(7), http://dx.doi.org/10.3390/ani10071207, URL: https://www.mdpi. com/2076-2615/10/7/1207.

Alqaysi, H., Fedorov, I., Qureshi, F. Z., & O'Nils, M. (2021). A temporal boosted YOLO- based model for birds detection around wind farms. *Journal of Imaging*, *7*(11), http://dx.doi.org/10.3390/jimaging7110227, URL: https://www.mdpi.com/2313- 433X/7/11/227.

Bibby, C., Jones, M., & Marsden, S. (1998). *Expedition field techniques: Bird surveys.* London: Royal Geographical Society.

BOEM (2022). Ocean wind 1 offshore wind farm draft environmental impact statement. URL: https://www.boem.gov/sites/default/files/documents/renewable- energy/state-activities/OceanWind1-DEIS.pdf.

Borowicz, A., Le, H., Humphries, G., Nehls, G., Höschle, C., Kosarev, V., & Lynch, H. J. (2019). Aerial-trained deep learning networks for surveying cetaceans from satellite imagery. *PLOS ONE*, *14*(10), Article e0212532. http://dx.doi.org/10.1371/journal. pone.0212532.

Burger, J., & Gochfeld, M. (2004). Marine birds as sentinels of environmental pollution. *EcoHealth*, *1*(3), http://dx.doi.org/10.1007/s10393-004-0096-4.

Caswell, N., Kuru, K., Ansell, D., Jones, M., Jon Watkinson, B., Leather, P., Lancaster, A., Sugden, P., Briggs, E. et al (2020) Patient Engagement in Medical Device Design: Refining the Essential Attributes of a Wearable, Pre-Void, Ultrasound Alarm for Nocturnal Enuresis. Pharmaceutical Medicine, 34 (2). pp. 1-19. https://doi.org/10.1007/s40290-019-00324-w.

Chabot, D., Dillon, C., & Francis, C. M. (2018). An approach for using off-the-shelf object-based image analysis software to detect and count birds in large volumes of aerial imagery. *Avian Conservation and Ecology*, *13*(1), http://dx.doi.org/10.5751/ ace-01205-130115.

Chabot, D., & Francis, C. M. (2016). Computer-automated bird detection and counts in high-resolution aerial images: a review. *Journal of Field Ornithology*, *87*(4), 343–359. http://dx.doi.org/10.1111/jofo.12171, URL: https://onlinelibrary.wiley.com/ doi/abs/10.1111/jofo.12171.

Clements, N., & Robinson, W. (2022). A re-survey of winter bird communities in the Oregon Coast Range, USA, initially surveyed in 1968–1970. *Biodiersity Data Journal*, http://dx.doi.org/10.3897/arphapreprints.e91575.

Cruz, J. E. C., Shiguemori, E. H., & Guimaraes, L. N. F. (2015). A comparison of Haar-like, LBP and HOG approaches to concrete and asphalt runway detection in high resolution imagery. *The Journal of Computational Interdisciplinary Sciences*, *6*(3), 121–136.

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proceedings of the 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) - Volume 1 - Volume 01* CVPR '05, (pp. 886–893). Washington, DC, USA: IEEE Computer Society, http://dx.doi.org/10.1109/CVPR. 2005.177.

Davis, K. L., Silverman, E. D., Sussman, A. L., Wilson, R. R., & Zipkin, E. F. (2022). Errors in aerial survey count data: Identifying pitfalls and solutions. *Ecology and Evolution*, *12*(3), http://dx.doi.org/10.1002/ece3.8733.

Delhez, A. (2022). *Multiple flying bird and bird keypoint detection toolbox for processing bird videos* (Master's thesis), Belgium: University of Louvain.

Dujon, A. M., Ierodiaconou, D., Geeson, J. J., Arnould, J. P. Y., Allan, B. M., Katse- lidis, K. A., & Schofield, G. (2021). Machine learning to detect marine animals in UAV imagery: effect of morphology, spacing, behaviour and habitat. *Remote Sensing in Ecology and Conservation*, *7*(3), 341–354. http://dx.doi.org/10.1002/rse2.205, URL: https://zslpublications.onlinelibrary.wiley.com/doi/abs/10.1002/rse2.205.

Forsyth, D., & Ponce, J. (2012). *Computer vision* (2nd ed.). London: Pearson.

Gibert, K., Horsburgh, J. S., Athanasiadis, I. N., & Holmes, G. (2018). Environmen- tal data science. *Environmental Modelling & Software*, *106*, 4–12. http://dx.doi. org/10.1016/j.envsoft.2018.04.005, URL: http://www.sciencedirect.com/science/ article/pii/S1364815218301269. Special Issue on Environmental Data Science. Applications to Air quality and Water cycle.

Groom, G., Stjernholm, M., Nielsen, R. D., Fleetwood, A., & Petersen, I. K. (2013). Remote sensing image data and automated analysis to describe marine bird distributions and abundances. *Ecological Informatics*, *14*, 2–8. http://dx.doi.org/10. 1016/j.ecoinf.2012.12.001, URL: https://www.sciencedirect.com/science/article/ pii/S1574954112001185. The analysis and application of spatial ecological data to support the conservation of biodiversity.

Hayes, M. C., Gray, P. C., Harris, G., Sedgwick, W. C., Crawford, V. D., Chazal, N., Crofts, S., & Johnston, D. W. (2021). Drones and deep learning produce accurate and efficient monitoring of large-scale seabird colonies. *Ornithological Applications*, *123*(3), http://dx.doi.org/10.1093/ornithapp/duab022.

Hollings, T., Burgman, M., van Andel, M., Gilbert, M., Robinson, T., & Robinson, A. (2018). How do you find the green sheep? A critical review of the use of remotely sensed imagery to detect and count animals. *Methods in Ecology and Evolution*, *9*(4), 881–892. http://dx.doi.org/10.1111/2041-210X.12973, URL: https: //besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12973.

Hong, S.-J., Han, Y., Kim, S.-Y., Lee, A.-Y., & Kim, G. (2019). Application of deep- learning methods to bird detection using unmanned aerial vehicle imagery. *Sensors*, *19*(7), http://dx.doi.org/10.3390/s19071651, URL: https://www.mdpi.com/1424- 8220/19/7/1651.

Jaderberg, M., Czarnecki, W. M., Dunning, I., Marris, L., Lever, G., Castañeda, A. G., Beattie, C., Rabinowitz, N. C., Morcos, A. S., Ruderman, A., Sonnerat, N., Green, T., Deason, L., Leibo, J. Z., Silver, D., Hassabis, D., Kavukcuoglu, K., & Graepel, T. (2019). Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science*, *364*(6443), 859–865. http://dx.doi.org/10.1126/ science.aau6249.

Jayadevan, P., Nameer, P. P. O., Jha, A., Aravind, A., Dilip, K. G., Karuthedathu, D., Tom, G., Mavelikara, H., Mannar, H., Palot, M., Johnson, J., Raveendran, J., Rodrigues, M., Mujeeb, M., Namassivayan, L., Payyeri, N., Nesrudheen, P., Narayanan, S. P., Prasanth, S., & Yathumon, M. A. (2022). Kerala Bird Atlas 2015– 20: features, outcomes and implications of a citizen-science project. *Current Science*, *122*, 298–309. http://dx.doi.org/10.18520/cs/v122/i3/298-309.

JNCC (2015). Northern Gannet Morus bassanus. URL: http://jncc.defra.gov.uk/page- 2875.

JNCC (2022). Seabirds count – the fourth breeding seabird census. URL: https://jncc. gov.uk/our-work/seabirds-count/.

Johnston, D. W. (2019). Unoccupied aircraft systems in marine science and conserva- tion. *Annual Review of Marine Science*, *11*(1), 439–463. http://dx.doi.org/10.1146/ annurev-marine-010318-095323, PMID: 30020850.

Kellenberger, B., Veen, T., Folmer, E., & Tuia, D. (2021). 21 000 Birds in 4.5 h: efficient large-scale seabird detection with machine learning. *Remote Sensing in Ecology and Conservation*, *7*(3), 445–460. http://dx.doi.org/10.1002/rse2.200, URL: https://zslpublications.onlinelibrary.wiley.com/doi/abs/10.1002/rse2.200.

Kuru, K., Ansell, D., Hughes, D., Watkinson, B. J., Gaudenzi, F., Jones, M., Lunardi, D., Caswell, N., Montiel, A. R., Leather, P., Irving, D., Bennett, K., McKenzie, C., Sugden, P., Davies, C., & Degoede, C. (2024). Treatment of Nocturnal Enuresis Using Miniaturised Smart Mechatronics With Artificial Intelligence. In IEEE Journal of Translational Engineering in Health and Medicine (Vol. 12, pp. 204–214). https://doi.org/10.1109/jtehm.2023.3336889.

Kuru, K., Clough, S., Ansell, D., McCarthy, J., & McGovern, S. (2023). Intelligent airborne monitoring of irregularly shaped man-made marine objects using statistical Machine Learning techniques. In Ecological Informatics (Vol. 78, p. 102285). Elsevier BV.

https://doi.org/10.1016/j.ecoinf.2023.102285

Kuru, K., Clough, S., Ansell, D., McCarthy, J., & McGovern, S. (2023). WILDetect: An intelligent platform to perform airborne wildlife census automatically in the marine ecosystem using an ensemble of learning techniques and computer vision. In Expert Systems with Applications (Vol. 231, p. 120574). https://doi.org/10.1016/j.eswa.2023.120574

Kuru, K., Girgin, S., Arda, K., & Bozlar, U. (2013). A novel report generation approach for medical applications: The SISDS methodology and its applications. *International Journal of Medical Informatics*, *82*(5), 435–447. http://dx.doi.org/10. 1016/j.ijmedinf.2012.05.019, URL: https://www.sciencedirect.com/science/article/ pii/S138650561200113X.

Kuru, K., & Khan, W. (2018). Novel hybrid object-based non-parametric clustering approach for grouping similar objects in specific visual domains. *Applied Soft Computing*, *62*, 667–701. http://dx.doi.org/10.1016/j.asoc.2017.11.007.

Kuru, K., Pinder, J. M., Watkinson, B. J., Ansell, D., Vinning, K., Moore, L., Gilbert, C., Sujit, A., & Jones, D. (2023). Toward Mid-Air Collision-Free Trajectory for Autonomous and Pilot-Controlled Unmanned Aerial Vehicles. In IEEE Access (Vol. 11, pp. 100323–100342). https://doi.org/10.1109/access.2023.3314504 .

Kuru, K., & Yetgin, H. (2019). Transformation to advanced mechatronics systems within new industrial revolution: A novel framework in automation of everything (AoE). *IEEE Access*, *7*, 41395–41415. http://dx.doi.org/10.1109/ACCESS.2019.2907809.

Li, Y., Yang, J., & Wen, J. (2021). Entropy-based redundancy analysis and information screening. *Digital Communications and Networks*, http://dx.doi.org/10.1016/j.dcan. 2021.12.001.

Li, Y., Yang, J., Zhang, Z., Wen, J., & Kumar, P. (2023). Healthcare data quality assessment for cybersecurity intelligence. *IEEE Transactions on Industrial Informatics*, *19*(1), 841–848. http://dx.doi.org/10.1109/TII.2022.3190405.

McConnell, R. (1986). Method of and apparatus for pattern recognition. URL: http: //www.google.co.uk/patents/US4567610. US Patent 4, 567, 610.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. http://dx.doi.org/10.1038/nature14236.

Moravčík, M., Schmid, M., Burch, N., Lisý, V., Morrill, D., Bard, N., Davis, T., Waugh, K., Johanson, M., & Bowling, M. (2017). DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, *356*(6337), 508–513. http://dx.doi. org/10.1126/science.aam6960.

Morelli, F. (2015). Indicator species for avian biodiversity hotspots: Combination of specialists and generalists is necessary in less natural environments. *Journal for Nature Conservation*, *27*, 54–62. http://dx.doi.org/10.1016/j.jnc.2015.06.006, URL: https://www.sciencedirect.com/science/article/pii/S1617138115300078.

Murray, S., Harris, M. P., & Wanless, S. (2014). An aerial survey of Northern Gannets Morus bassanus on Scar Rocks, southwest Scotland, in 2014. *Seabird*, *27*(1), 104–109.

Murray, S., Harris, M. P., & Wanless, S. (2015). The status of the gannet in Scotland in 2013-14. *Scottish Birds*, *35*(1), 3–18.

Murray, S., Smith, I., & Smith, A. (2014). Gannet and Guillemot breeding on Rockall, North Atlantic. *Scottish Birds*, *34*(1), 13–15.

Ojala, T., Pietikainen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proceedings of 12th international conference on pattern recognition, Vol. 1* (pp. 582–585). http://dx.doi.org/10.1109/ICPR.1994.576366.

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 971–987. http://dx.doi.org/10. 1109/TPAMI.2002.1017623.

Paleczny, M., Hammill, E., Karpouzi, V., & Pauly, D. (2015). Population trend of the world's monitored seabirds, 1950–2010. *PLOS ONE*, *10*(6), Article e0129342. http://dx.doi.org/10.1371/journal.pone.0129342.

Papageorgiou, C. P., Oren, M., & Poggio, T. (1998). A general framework for object detection. In *Sixth international conference on computer vision (IEEE Cat. No.98CH36271)* (pp. 555–562). http://dx.doi.org/10.1109/ICCV.1998.710772.

Reyes, E. (2014). *A comparison of image processing techniques for bird detection* (Master's thesis), the Faculty of California Polytechnic State University.

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., & Marra, P. P. (2019). Decline of the North American avifauna. *Science*, *366*(6461), 120–124. http://dx.doi.org/10. 1126/science.aaw1313, URL: https://www.science.org/doi/abs/10.1126/science. aaw1313. arXiv:https://www.science.org/doi/pdf/10.1126/science.aaw1313.

RSBP (2015). Gannet. URL: https://www.rspb.org.uk/birds-and-wildlife/wildlife- guides/bird-a-z/gannet/.

SeabirdCentre (2017). Gannet. URL: https://seabird.org/wildlife/seabirds/gannet/12/ 26/47.

Shi, Y., Shen, C., Fang, H., & Li, H. (2017). Advanced control in marine mechatronic systems: A survey. *IEEE/ASME Transactions on Mechatronics*, *22*(3), 1121–1131. http://dx.doi.org/10.1109/TMECH.2017.2660528.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. http: //dx.doi.org/10.1038/nature16961.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T. P., Simonyan, K., & Hassabis, D. (2017). Mastering chess and shogi by self-play with a general reinforcement learning algorithm. ArXiv, abs/1712.01815.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, *362*(6419), 1140–1144. http://dx.doi.org/10.1126/ science.aar6404, arXiv:https://www.science.org/doi/pdf/10.1126/science.aar6404.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359. http://dx.doi.org/10. 1038/nature24270.

Smyser, T. J., Guenzel, R. J., Jacques, C. N., & Garton, E. O. (2016). Double-observer evaluation of pronghorn aerial line-transect surveys. *Wildlife Research*, *43*(6), 474. http://dx.doi.org/10.1071/wr16006.

Stefano, L. D., Mattoccia, S., & Mola, M. (2003). An efficient algorithm for exhaustive template matching based on normalized cross correlation. In *12th international conference on image analysis and processing, 2003. Proceedings* (pp. 322–327). Mantova, Italy: IEEE.

Thompson, D. (2021). Seabird survey methods for offshore installations: Black-legged kittiwakes. URL: https://assets.publishing.service.gov.uk/government/uploads/ system/uploads/attachment_data/file/974338/Kittiwake_survey_advice_v2.1.pdf.

Trefný, J., & Matas, J. (2010). Extended set of local binary patterns for rapid object detection. In *Computer vision winter workshop 2010*. Czech Society for Cybernetics and Informatics.

Vállez, N., Deniz, O., & Bueno, G. (2015). Sample selection for training cas- cade detectors. *PLOS ONE*, *10*(7), 1–16. http://dx.doi.org/10.1371/journal.pone. 0133059.

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, Vol. 1* (pp. I–511–I–518). http://dx.doi.org/10.1109/CVPR.2001.990517.

Wang, L., & He, D.-C. (1990). Texture classification using texture spectrum. *Pattern Recognition*, *23*(8), 905–910.

Wang, D., Shao, Q., & Yue, H. (2019). Surveying wild animals from satellites, manned aircraft and unmanned aerial systems (UASs): A review. *Remote Sensing*, *11*(11), http://dx.doi.org/10.3390/rs11111308, URL: https://www.mdpi.com/2072-4292/ 11/11/1308.

Witkin, A. (1984). Scale-space filtering: A new approach to multi-scale description. In *ICASSP '84. IEEE international conference on acoustics, speech, and signal processing, Vol. 9* (pp. 150–153). http://dx.doi.org/10.1109/ICASSP.1984.1172729.

Yang, Z. (2010). Fast template matching based on normalized cross correlation with centroid bounding. In *2010 international conference on measuring technology and mechatronics automation, Vol. 3* ICMTMA, (pp. 224–227). Changsha, China: IEEE.

Zhang, L., Chu, R., Xiang, S., & Liao, S. (2007). *Lecture notes in computer science, Facts and fallacies: A book of definitive mistakes and misguided predictions in advances in biometrics*. Berlin Heidelberg: Springer.

Kuru, K. (2014). A Novel Hybrid Clustering Approach for Unsupervised Grouping of Similar Objects. In Lecture Notes in Computer Science (pp. 642–653). https://doi.org/10.1007/978-3-319-07617-1_56

Kuru, K. (2014). Optimization and enhancement of H&E stained microscopical images by applying bilinear interpolation method on lab color mode. In Theoretical Biology and Medical Modelling (Vol. 11, Issue 1). https://doi.org/10.1186/1742-4682-11-9

Kuru, K. (2021). Planning the Future of Smart Cities With Swarms of Fully Autonomous Unmanned Aerial Vehicles Using a Novel Framework. In IEEE Access (Vol. 9, pp. 6571–6595). https://doi.org/10.1109/access.2020.3049094.

Kuru, K. (2022). TrustFSDV: Framework for Building and Maintaining Trust in Self-Driving Vehicles. In IEEE Access (Vol. 10, pp. 82814–82833). https://doi.org/10.1109/access.2022.3196941

Kuru, K. (2021). Management of geo-distributed intelligence: Deep Insight as a Service (DINSaaS) on Forged Cloud Platforms (FCP). In Journal of Parallel and Distributed Computing (Vol. 149, pp. 103–118). https://doi.org/10.1016/j.jpdc.2020.11.009

Kuru, K. (2023) Definition of Multi-Objective Deep Reinforcement Learning Reward Functions for Self-Driving Vehicles in the Urban Environment. IEEE Transactions on Intelligent Transportation Systems. (Submitted)

Kuru, K. (2023). MetaOmniCity: Toward Immersive Urban Metaverse Cyberspaces Using Smart City Digital Twins. In IEEE Access (Vol. 11, pp. 43844–43868). https://doi.org/10.1109/access.2023.3272890

Kuru, K. (2021). Conceptualisation of Human-on-the-Loop Haptic Teleoperation With Fully Autonomous Self-Driving Vehicles in the Urban Environment. In IEEE Open Journal of Intelligent Transportation Systems (Vol. 2, pp. 448–469). https://doi.org/10.1109/ojits.2021.3132725

Kuru, K. (2023). Sensors and Sensor Fusion for Decision Making in Autonomous Driving and Vehicles. MDPI.

Kuru, K., Erogul, O., & Xavier, C. (2021). Autonomous Low Power Monitoring Sensors. Sensors, 21.