


Article

Exploration in Mapping Kernel-Based Home Range Models from Remote Sensing Imagery with Conditional Adversarial Networks

Ruobing Zheng ^{1,2} , Guoqiang Wu ¹, Chao Yan ³, Renyu Zhang ⁴, Ze Luo ^{2,*} and Baoping Yan ²

¹ University of Chinese Academy of Sciences, Beijing 100049, China; zhengruobing@cnic.cn (R.Z.); wuguoqiang16@mailsucas.ac.cn (G.W.)

² e-Science Technology and Application Laboratory, Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China; ybp@cnic.cn

³ Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37240, USA; chao.yan@vanderbilt.edu

⁴ Toyota Technological Institute at Chicago, Chicago, IL 60637, USA; zhangr@ttic.edu

* Correspondence: luoze@cnic.cn

Received: 21 August 2018; Accepted: 27 October 2018; Published: 31 October 2018



Abstract: Kernel-based home range models are widely-used to estimate animal habitats and develop conservation strategies. They provide a probabilistic measure of animal space use instead of assuming the uniform utilization within an outside boundary. However, this type of models estimates the home ranges from animal relocations, and the inadequate locational data often prevents scientists from applying them in long-term and large-scale research. In this paper, we propose an end-to-end deep learning framework to simulate kernel home range models. We use the conditional adversarial network as a supervised model to learn the home range mapping from time-series remote sensing imagery. Our approach enables scientists to eliminate the persistent dependence on locational data in home range analysis. In experiments, we illustrate our approach by mapping the home ranges of Bar-headed Geese in Qinghai Lake area. The proposed framework outperforms all baselines in both qualitative and quantitative evaluations, achieving visually recognizable results and high mapping accuracy. The experiment also shows that learning the mapping between images is a more effective way to map such complex targets than traditional pixel-based schemes.

Keywords: home range; remote sensing; deep learning; generative adversarial networks; habitat mapping; Bar-headed Geese

1. Introduction

The basic concept of the home range is defined as the area traversed by the animal during its normal activities of food gathering, mating, and caring for young [1]. Estimating home range is an important part of investigating species status, analyzing habitat selection, and developing conservation strategies [2]. With the development of Geographic Information System (GIS), home ranges are now often estimated from the locational data obtained with radio-tracking techniques [3]. The Minimum Convex Polygon (MCP) is a simple but popular method which assumes the uniform use of space within the outside boundary of animal locations [4]. However, animals are unlikely to use their home range in a uniform manner in the real world. Therefore, a series of kernel-based probabilistic methods have shown the advantages in habitat studies [5,6], especially in understanding the internal structure of spatially heterogeneous environments. This type of model [2,7,8] produces a two-dimensional probability density map to represent the probability of animals occurring at each

location in the defined area. Ref. [9] demonstrates that kernel home range models would enhance the studies of animal movements, species interactions, and resource selection.

Kernel home range models are based on locational data: either the density of locations or link distance between locations. Data for mapping home ranges used to be gathered by careful observation, but nowadays such data is usually collected automatically using GPS collars placed on animals. In practice, scientists usually capture and mark a certain amount of target species and collect GPS data within the validity period of the vulnerable tracking device [10,11]. The whole process is costly and in most cases a one-time job. However, as introduced in [12], home ranges change dynamically over time, and estimating them with relocations requires sufficient and timely GPS records. Inadequate GPS data often prevents scientists from applying kernel home range models in long-term and large-scale research.

To solve this problem, we try to find alternatives. Habitat mapping studies [13–15] have demonstrated the strong connection between animal activities and environmental variables. These studies effectively leverage remote sensing imagery to map different types of habitat characteristics, which inspired us to map the home ranges from long-term support remote sensing imagery. However, most habitat mapping studies employ traditional classification or regression models on each independent pixel. This pixel-based scheme ignores the structural information in remote sensing imagery, which is an obvious defect in mapping our image-based target, the probabilistic home range map.

In this paper, we train an end-to-end deep learning framework to achieve this goal. The well-trained deep convolutional network could effectively produce the home range maps from image-based source data, without the need for animal relocations. Our main contributions can be summarized as follows:

- We propose a general-purpose framework for mapping the kernel-based home range models from time-series remote sensing imagery. We innovatively use the adversarial network as a supervised model to learn the mapping between image-based data and the target (Figure 1). Our method enables scientists to carry out their home range analysis even if the GPS data is insufficient for long-term and large-scale research. To our knowledge, this is the first exploration in mapping home range models using an image-based strategy.
- We illustrate our method in a real-world scenario for mapping the home ranges of Bar-headed Geese in Qinghai Lake area. We build a specific dataset for training the mapping model and elaborate each stage in the experiment. Our experience will assist researchers in extending their research scale in various wildlife analyses.
- We qualitatively and quantitatively compare our method against several baseline models. We analyze the strengths and drawbacks of selected baselines and further discuss why our method is suitable for this specific task.

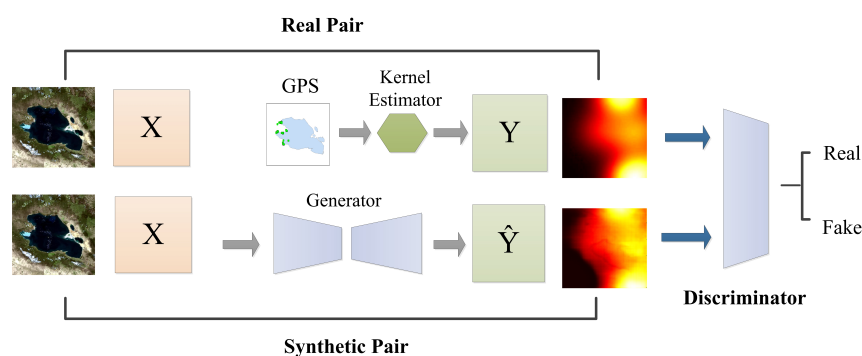


Figure 1. Using the adversarial framework to simulate the kernel-based home range estimator. X is the time-series remote sensing image, and Y is the corresponding home range map. The generator, G , learns the mapping function $X \rightarrow Y$. The discriminator D tries to classify the real and synthetic data–target pairs. Both G and D are deep convolutional neural networks.

2. Related Work

2.1. Kernel-Based Home Range Models

The original concept of the home range is introduced by [1] (1943). He constructed a map delineating the outside boundary of the animal's movement during the course of its activities. A more formal definition is the Utilization Distribution (UD) [16], which takes the form of a bivariate probability density function that represents the probability of finding an animal in a defined area [17]. Kernel UD [7,18] (also called bivariate Gaussian) is the best known home range model for constructing UD. It employs Kernel Density Estimation (KDE) [19] on animal relocations and uses a Gaussian kernel to calculate the probability on each location of the defined area. Several recent studies extended the kernel approach by using the movement patterns of wildlife, such as the Brownian bridge movement model [20], which takes the time dependence between locations into account. In summary, these kernel-based home range models produce a two-dimensional probability density map that represents the probability of the animal's occurrence.

2.2. Habitat Mapping

Habitat mapping [21,22] is a well-studied topic in the literature of remote sensing applied to ecology studies. Compared to its great success in the vegetation field [23,24], the related application to animals is limited by more complicated correspondence [25–27]. Some studies [13,14,28,29] leverage the remote sensing imagery to map the quality and extent of wildlife habitats. These studies mostly focus on the specific species and have different source data, mapping models, and final targets. Technically, they mainly employ classification or regression models on the independent multi-dimensional vector contained in each pixel of remote sensing images, to predict discrete habitat categories [30] or continuous habitat index [31]. This pixel-based scheme successfully identifies, verifies, and explains the habitat characteristic at the pixel level. However, it fails to consider the strong dependencies between pixels in highly structured remote sensing imagery.

2.3. Image-to-Image Translation

Image-to-Image translation is a class of problems emerging in computer vision literature, which aims to learn the mapping between an input image and an output image using a training set of aligned image pairs [32]. This technique has many successful applications such as generating photographs from sketches [33], image style transfer [34], and image inpainting [35]. The Fully Convolutional Networks (FCN) [36] can be seen as the embryonic form of this work, which removed the last fully connected layers in traditional Convolutional Neural Network (CNN) [37] to make the dense prediction at the full image level. Later, deep generative models [38] in a conditional setting have shown promise in this field, such as Conditional Variational Autoencoder (CVAE) [39] and Conditional Generative Adversarial Network (CGAN) [40]. Typically, [41] proposed a general-purpose solution for the image-to-image translation problem. The “pixel2pixel” framework extends CGAN and leads to a substantial boost in the quality of translated images. Several studies [32,42] have carried out this work and further discussed the solution of multi-modal and unpaired image-to-image translation. This series of studies greatly inspires us in the following work.

3. Materials and Methods

3.1. Data and Target

We annotate the time-series remote sensing imagery with corresponding home range maps to build image pairs. These image pairs are used to train the end-to-end mapping framework. The target, home range map, is calculated from GPS data with the kernel-based estimator. It is technically an image-based probability density map. The data are time-series remote sensing imagery. The type of remote sensing imagery is decided by each specific wildlife home range analysis. We align the

image-based data and target at both spatial and temporal level. More details of the pre-processing procedure are described in a specific example in Section 4.3.

3.2. Mapping Model

Assuming that we have produced a set of aligned data–target pairs, we try to find the mapping $X \rightarrow Y$, from multi-layer remote sensing images $X \in \mathbb{R}^{H \times W \times B}$ (B is the number of layers) to home range maps $Y \in \mathbb{R}^{H \times W \times 1}$. Both X and Y have the same size and continuous pixel values. We achieve this goal by the following formulations.

Formulation

The basic idea of Generative Adversarial Nets (GANs) [38] is to simultaneously train a pair of adversarial networks, a generator G , and a discriminator D . The target of G is to produce samples $G(z)$ under the distribution p_g , from a random variable z . The D is to make the generated distribution p_g close to the real data distribution p_{data} . The objective of GAN can be expressed as follows:

$$\max_D V_{GAN}(D) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p(z)} [\log (1 - D(G(z)))] \quad (1)$$

$$\min_G V_{GAN}(G) = \mathbb{E}_{z \sim p(z)} [\log (1 - D(G(z)))] \quad (2)$$

Original GANs can be extended to a conditional mode (CGAN [40]) in which both the generator and discriminator are conditioned on some extra information. This modification enables researchers to use an input image as conditional information to generate the corresponding output image. The representative *pix2pix* model took advantage of CGAN and extended the adversarial loss with the ℓ_1 loss balanced by λ . Technically, the *pix2pix* learns a mapping from a type of image A and random noise z , to another type of image B : $\{A, z\} \rightarrow B$. The objectives can be expressed as:

$$\max_D V_{p2p}(D) = \mathbb{E}_{A, B \sim p(A, B)} [(\log (D(A, B)))] + \mathbb{E}_{A \sim p(A), z \sim p(z)} [\log (1 - D(A, G(A, z)))] \quad (3)$$

$$\ell_1(G) = \mathbb{E}_{A, B \sim p(A, B), z \sim p(z)} [\|B - G(A, z)\|_1] \quad (4)$$

$$\min_G V_{p2p}(G) = \mathbb{E}_{A \sim p(A), z \sim p(z)} [\log (1 - D(A, G(A, z)))] + \lambda \ell_1(G) \quad (5)$$

In this paper, the required mapping is from the remote sensing imagery X to the home range maps Y . We adapt the adversarial loss from *pix2pix* to fit our scenario and apply the least-square loss [43] to stabilize the training procedure and expedite the convergence. The final objects are shown as:

$$\min_D V_F(D) = \frac{1}{2} \mathbb{E}_{X, Y \sim p(X, Y)} [(D(X, Y) - 1)^2] + \frac{1}{2} \mathbb{E}_{X \sim p(X), z \sim p(z)} [D(X, G(X, z))^2] \quad (6)$$

$$\min_G V_F(G) = \frac{1}{2} \mathbb{E}_{X \sim p(X), z \sim p(z)} [(D(X, G(X, z)) - 1)^2] + \lambda \ell_1(G) \quad (7)$$

3.3. Implementation

Network Architectures

Working with the above loss functions, two deep convolutional neural networks are used to implement the adversarial framework which consists of a generative network G and a discriminative network D . We adapt the network architecture from those in [41]. As seen in Figure 2, the generator G is a deep convolutional encoder–decoder [44]. The encoder extracts the high-level features from remote sensing layers while the decoder interprets these features and upsamples them to a home range map. Convolution layers help to extract the features with the consideration of structural information. U-Net connection [45] is used to share the low-level information between encoder and decoder layers [41].

The discriminator D is a traditional CNN classifier which helps G to learn more accurate mapping in the adversarial training. It is worth noting that the input for D is the data–target pair. The work of D is to determine whether the input pairs are real or synthetic. The synthetic data–target pair is formed with the real satellite image and the generated home range map. The architecture of D is shown in Figure 3.

For the generator part, the encoder is $C64 - C128 - C256 - C512 - C512 - C512$, and the decoder is $CD512 - CD512 - CD512 - C256 - C128 - C64$. Cn is a Convolution–BatchNorm–LeakyReLU layer with n feature filters. CDn denotes a Convolution–BatchNorm–Dropout–LeakyReLU layer with a 50% dropout rate. In terms of the discriminator, the convolutional layers are designed as $C64 - C128 - C256 - C512 - C512 - C512$.

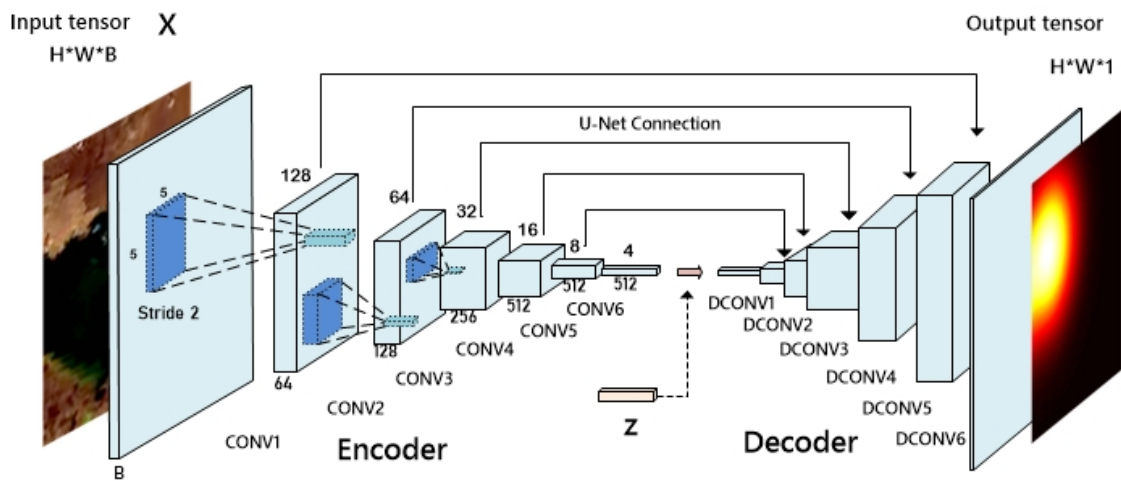


Figure 2. Architecture of the generator G in the adversarial framework. G generates samples $G(X, z)$ from auxiliary information X and random noise z . “CONV/DCONV, stride = 2” denotes a convolutional/deconvolutional layer with two strides on each move. The number k under each tensor and the number d on the top stand for the size of the tensor is $d \times d \times k$.

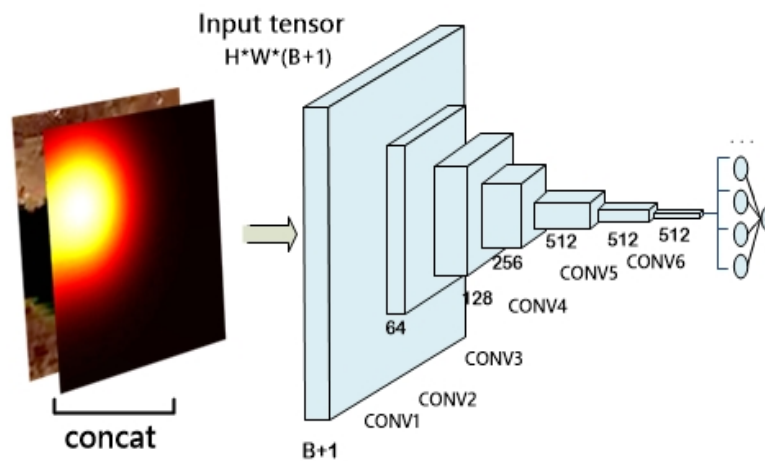


Figure 3. Architecture of the discriminator D in the adversarial framework. The number under each tensor stands for the number of features.

4. Experiment

In this section, we illustrate our method in a real-world scenario for mapping the home ranges of Bar-headed Geese in Qinghai Lake area. We also compare our method against several baselines with both qualitative and quantitative evaluation. The experiment was conducted in the desktop environment (Intel i5-6600K, NVIDIA Geforce GTX 1070). The home range maps were estimated using an R package (adehabitatHR [18]) in R v3.5.1. The deep learning networks were implemented by Tensorflow [46] version 1.7 in Python 2.7 environment.

4.1. Study Area and Field Knowledge

As shown in Figure 4, the study area (96.6° and 102.4° E, and 34.2° and 38.8° N) mainly covers Qinghai Lake, Gyaring Lake, Ngoring Lake, and Donggi Conag Lake, in Qinghai Province, China. These lakes, as well as the surrounding wetlands and estuaries, serve as a critical breeding ground and migratory staging area for many kinds of migratory waterfowl, especially the Bar-headed Goose (*Anser indicus*). This special species gained political and scientific attention following the large outbreak of highly pathogenic H5N1 avian influenza at Qinghai Lake area in the spring of 2005 [47,48]. This single event is the first large-scale outbreak of avian influenza. It caused the death of nearly 5% of the global population of Bar-headed Geese [49], which sparked a global debate on the role that wild birds play in the spread of H5N1 [50].

4.2. Source Data

4.2.1. GPS Data

We select five Bar-headed Geese captured and GPS-collared in Qinghai Lake area in 2007. These water birds are equipped with a 45 g solar-powered portable transmitter terminal. We recorded the GPS locations for each bird during the breeding season in 2007 and 2008. The details of the radio-tracking data are shown in Table 1.

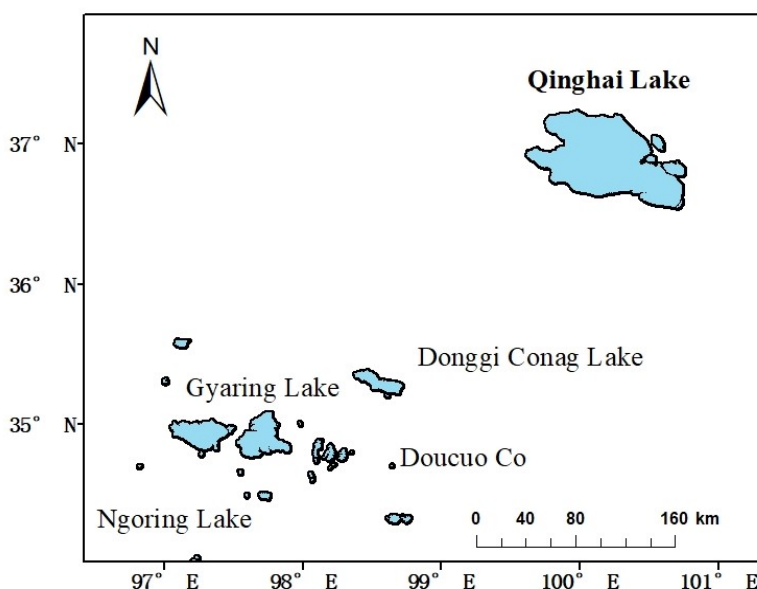


Figure 4. The study area includes Qinghai Lake, Ngoring Lake, Gyaring Lake, Donggi Conag Lake, and several wetlands and estuaries. These places serve as a critical breeding ground and migratory staging area for Bar-headed Geese.

Table 1. ID number, sex, capture time, and number of GPS locations of five selected Bar-headed Geese.

Bird	Sex	Capture	GPS Records	
			2007	2008
BH07_67582	F	03/25/07	924	80
BH07_67690	F	03/27/07	311	23
BH07_67695	M	03/29/07	333	571
BH07_67698	M	03/30/07	642	11
BH07_74898	M	03/31/07	864	690
Sum			3074	1375

4.2.2. Remote Sensing Imagery

Moderate Resolution Imaging Spectroradiometer (MODIS) [51] Land Products are used in this application. We select environmental factors based on the field survey and a review of the literature (Table 2). The Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) are used to determine the food availability. The Normalized Difference Water Index (NDWI) is used to determine the access to water. MODIS land cover type is used to examine the shelter conditions. We use the involved MODIS reflectance bands instead of the original factor maps as input. We leverage the deep neural network to approximate the band math and avoid the overlapped bands. We obtained the MODIS land reflectance bands from MOD09Q1 and MOD09A1 8-day L3 products.

4.3. Preprocessing

In the pre-processing stage, we annotate the MODIS reflectance data with the home range maps calculated from the kernel UD [7,18] estimator. To build the image-based data–target pairs, we align the home range maps with remote sensing images at both the spatial and temporal level. On the temporal dimension, we group bird GPS data by every 8 days to match the time interval of selected MODIS Land Products. For the spatial resolution, all raster data are transformed into the same projected coordinate system (EPSG: 4326: WGS84) and resampled to the 250 m resolution. Then, we use the R package (adehabitatHR [18]) to estimate the probability in each pixel on the remote sensing image. We slice the MODIS image as well as the probability map into numerous 256×256 tiles and pair them, as shown in Figure 5. We produce a total of 832 image pairs with the GPS data in 2007 and 2008. We randomly select 100 pairs as the test set, 132 pairs as the validation set and 600 training pairs.

Table 2. The selected environmental factors and corresponding MODIS land reflectance bands used in this application. We also list the relevant waterfowl study which used the same factor. RED represents the wavelength of 620–670 nm for MODIS Land Bands; NIR covers the wavelength of 841–876 nm; BLUE covers the wavelength of 459–479 nm, and the GREEN covers the wavelength of 545–565 nm.

Factors	Formula	Involved Bands	Reference
NDVI	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$	Band 1, 2	[52] 2013
EVI	$2.55 (\text{NIR} - \text{RED}) / (\text{NIR} + 6\text{RED} - 7.5\text{BLUE} + 1)$	Band 1, 2, 3	[53] 2018
NDWI	$(\text{GREEN} - \text{NIR}) / (\text{GREEN} + \text{NIR})$	Band 2, 4	[54] 2010
Land Cover	MODIS land cover classification algorithm (MLCCA)	Band 1–7	[10] 2010

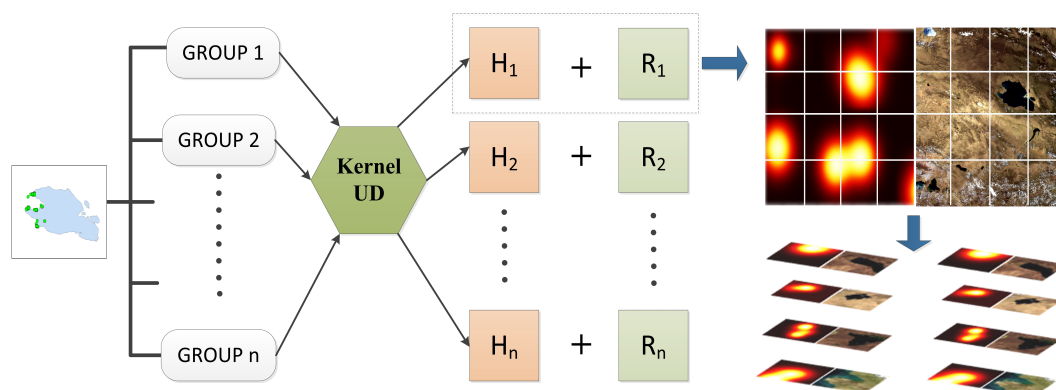


Figure 5. The pre-process procedure of building the image-based data–target pairs from source data. The GPS data are divided into time-series groups which are used to estimate home range maps via the kernel UD estimator. Then we pair each home range map (H_n) with the corresponding remote sensing image (R_n) to form the image pairs.

4.4. Training Details

We initialize all convolutional kernels with a Gaussian distribution $\mathcal{N}(0, 0.02)$. The ReLU unit is leaky with slope 0.2 and the parameter λ_{ℓ_1} is set to 10. We search the best value of these hyperparameters on the validation set. Following the experience in [43] in which the non-momentum optimizers perform better in very nonstationary problems, we use RMSProp optimizer [55] with a learning rate of 0.0002 instead of the default Adam optimizer [56]. The improvement of training stability can be found in Section 4.5. Considering that our self-made dataset is relatively small compared to common datasets, we employ data enlargement operations in the training procedure. Mirroring and rotation for the input image are used before each training batch. Image random jitter [41] is also applied by resizing the 256×256 input to 275×275 resolution and then random cropping back to original size.

4.5. Training Stability

Unstable training is a common problem [41] for adversarial frameworks. Because the two adversarial components have opposing objectives in the simultaneous training procedure, their loss curves are usually highly-fluctuated. As mentioned in Section 4.2, we use the least-squares loss and RMSProp optimizer to improve the training stability. Here, we present the loss curves for both the generator and discriminator of our model in Figure 6. We also compared our model against two different combinations.

We find that the loss curves of the Cross-Entropy (CE) loss and Adam optimizer are highly-fluctuated, especially for the discriminator. They also have a slower convergence than two least-squares models. The second combination shows that applying the Adam optimizer on least-squares loss is not a good choice even though least-squares (LS) loss can still lead to a fast convergence. Our model (LS loss + RMSProp) achieves a fast convergence and more stable training for both the generator and discriminator, which confirms the previous report [43] that the optimizers without momentum perform better in the LS loss.

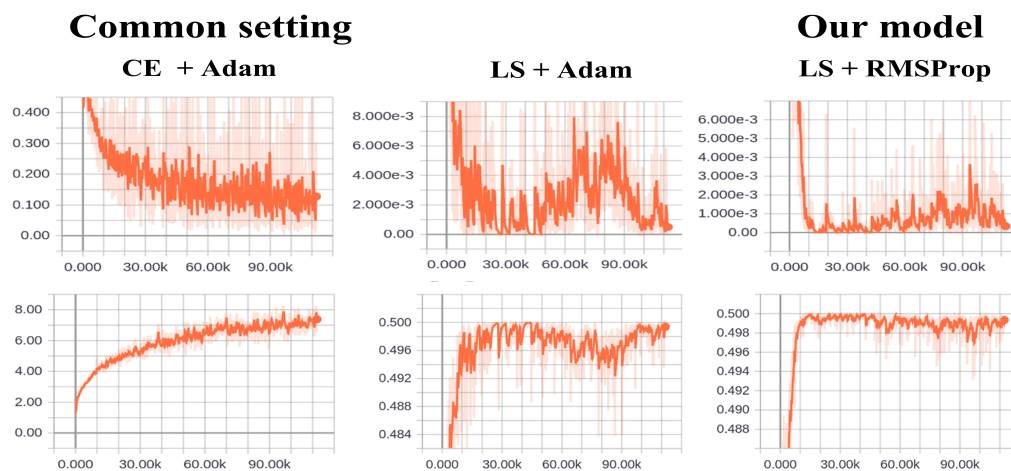


Figure 6. The loss curves of our model and two different combinations (loss + optimizer). The first row presents the loss of the discriminator and the second row presents the loss of the generator. The X axis represents the number of training steps, and the Y axis represents the loss values. Our model has a fast convergence and smooth loss curves for both the generator and discriminator among all combinations.

4.6. Results

We select six representative test samples which have obvious and differently shaped home ranges, as shown in Figure 7. We observe that the synthesized home range maps from our model successfully, capturing the primary distribution of the real target, even though there still exist some artifacts and noises. This result demonstrates that our model has the ability to implement the mapping between remote sensing imagery and kernel-based home range models.

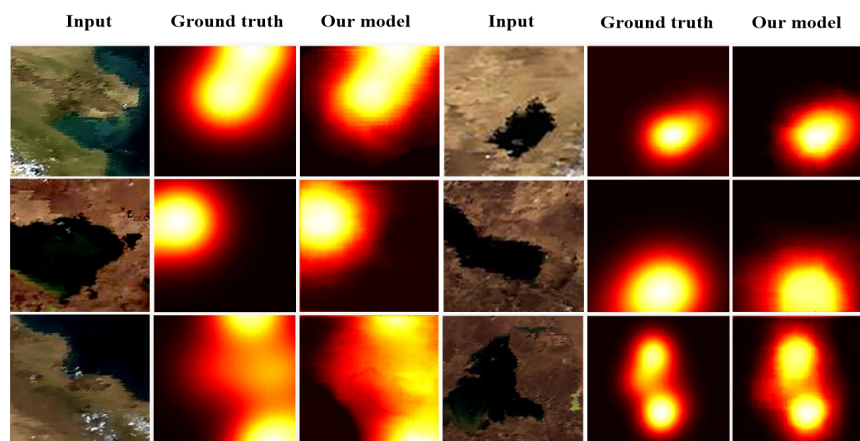


Figure 7. The mapping result of the selected samples in the test dataset. In each set, the first image is the true color composite (band 1,4,3) of MODIS land products, which represents the input remote sensing imagery. The last image (ground truth) is estimated by the kernel UD estimator with GPS data. The middle one is the synthesized home range map which is directly mapped from remote sensing images using our end-to-end model. We colorized the original probability map with the hot colormap.

4.6.1. Baselines

Although we focus on a specific scenario, we investigate several potential solutions in both habitat mapping and computer vision literature.

- **kNN:** k-Nearest Neighbors algorithm [57] is a non-parametric method used for both classification and regression. In the regression mode, the output value is the average of the values of its k nearest neighbors.

- **Decision Tree:** Decision Tree [58] is a non-parametric supervised learning method used for both classification and regression. Classification and Regression Tree (CART) has been used in mapping the extent and quality of wildlife habitat in many studies [31,59]. We first test the decision tree as a regression model to map our target using a pixel-based scheme.
- **Random Forest:** Random Forest regressor [60] fits a number of classification and decision trees on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. It has also been used as an advanced model instead of the decision tree in several studies [61,62]. Here, we examined it as a pixel-based baseline to investigate whether the improvement at the model level can overcome the limitation of the pixel-based scheme.
- **CNN + ℓ_2 loss:** CNN with ℓ_2 loss is the most straightforward way to predict a continuous target using deep learning. The ℓ_2 loss is the general choice in image processing tasks [63,64]. Here, we used the same encoder–decoder as our model to avoid the impact of network architecture. This baseline is actually training the proposed generator network in ℓ_2 loss.
- **Conditional VAE:** Deep generative models have achieved good performance in the image-to-image translation. Besides CGAN, another well-established generation model, CVAE [65], has also shown promise in similar studies [65,66]. Different from GAN, VAE makes strong assumptions concerning the posterior and prior distribution of hidden variable and target data. It also tries to approximate these distributions with neural networks.

4.6.2. Metrics

- **Regression Metrics:** To quantitatively evaluate the prediction of the continuous values on home range maps, we employ Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 to measure the mapping accuracy from the perspective of regression.
- **SSIM Loss:** Considering that our target is also a structured image like nature images, we used the Structural Similarity Index (SSIM) [67] to measure the structural similarity between the synthesized home range maps and the real target. SSIM is an established metric to measure quality and similarity between images [68,69]. The SSIM for pixel y_i is:

$$SSIM(y_i) = \frac{2\mu_{y_i}\mu_{\hat{y}_i} + C_1}{\mu_{y_i}^2 + \mu_{\hat{y}_i}^2 + C_1} \cdot \frac{2\sigma_{y_i\hat{y}_i} + C_2}{\sigma_{y_i}^2 + \sigma_{\hat{y}_i}^2 + C_2} = l(y_i) \cdot cs(y_i) \quad (8)$$

where the \hat{y}_i and y_i are the predicted and real value of the i th pixel respectively. $l(y_i)$ is the luminance comparison and $cs(y_i)$ is the contrast and structure comparison, and the constants C_1 and C_2 are used to avoid instability. The means and standard deviations are computed with a Gaussian filter on y_i . The SSIM loss for an image with N pixels can be expressed as:

$$\mathcal{L}^{SSIM} = \frac{1}{N} \sum_{i=1}^N 1 - SSIM(y_i) \quad (9)$$

4.6.3. Qualitative Evaluation

As shown in Figure 8, we observe that the kNN regressor fails to produce recognizable home range maps. The other two pixel-based regression models (DT and RF) produce relatively noisy results compared to other baselines. Although they successfully predict the high probability area in the right place, they also bring a large number of noises scattered in the full image field. Examining the results reveals that pixel-based models are limited when it comes to mapping an image-based target. Concerning the two deep learning models, we find that both CNN + ℓ_2 and CVAE eliminate random noise but suffer the problem of blurring. The CNN + ℓ_2 produces fuzzy images and a significantly larger positive area than the ground truth. However, compared with the pixel-based methods,

the image-based models produce more explicit and recognizable home range maps. Our model achieves the clearest and most visually realistic result among all baselines.

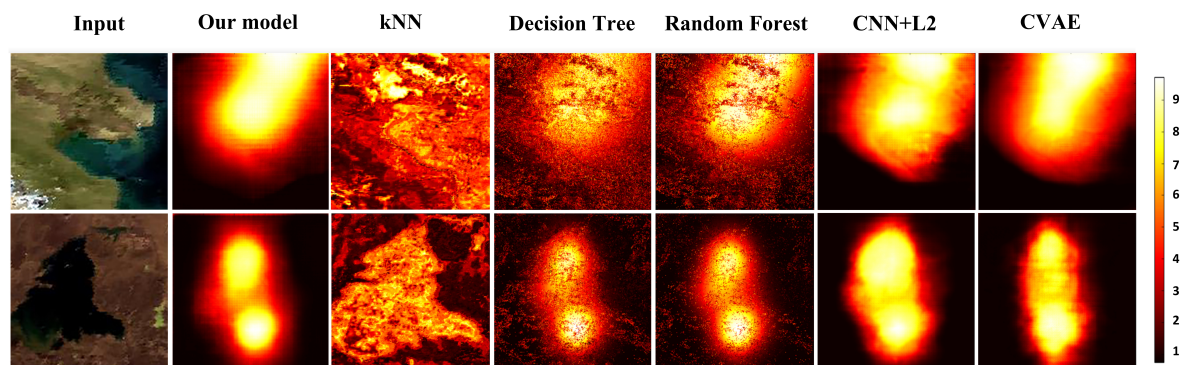


Figure 8. We compared the performance of our model and baselines on two test samples. We colorized the original probability map with the hot colormap which represents the probability value from low-to-high using dark-to-bright colors.

4.6.4. Quantitative Evaluation

We qualify the mapping accuracy of our model and baselines using two types of metrics. As seen in Table 3, three pixel-based models have higher RMSE, MAE, and lower R^2 in the test dataset, which confirms our previous assessment in the qualitative evaluation. The random forest regressor outperforms the decision tree on both RMSE and MAE but has nearly the same SSIM loss. This result reveals that the defect of the pixel-based scheme in image-based tasks cannot be overcome by the improvement of regression models. As for image-based baselines, they generally achieve better performance than pixel-based baselines. The CNN + ℓ_2 achieves a relatively lower RMSE due to its objective being also to minimize the ℓ_2 norm of errors. The blurry samples from CNN and CVAE lead to higher SSIM loss. In general, our model shows a promising result on both the three regression metrics and the structure similarity, which demonstrates that the adversarial loss and convolutional network architecture make a significant contribution to producing accurate and high-quality results.

Table 3. The quantitative evaluation employed in our model and baselines, with the metrics of RMSE, MAE, R^2 and \mathcal{L}^{SSIM} .

Method	Metrics			
	RMSE	MAE	R^2	\mathcal{L}^{SSIM}
Our model	17.363	13.028	0.865	0.368
kNN Regressor	52.82	41.29	0.167	0.998
Decision and Regression Tree	28.357	18.867	0.702	0.974
Random Forest Regressor	22.442	16.435	0.771	0.934
CNN + ℓ_2 loss	18.605	19.481	0.821	0.495
Conditional VAE	20.311	17.871	0.788	0.511

5. Discussion

In this paper, we proposed a novel end-to-end deep learning framework to simulate kernel home range models by learning a mapping between image pairs. Instead of defining a new habitat model and explaining its ecological meaning, we focused on extending the applicability of the existing home range estimators. This work explores a novel way to solve the specific problem in animal ecology. We hope that the proposed approach can benefit both remote sensing and ecology communities.

Let us review the traditional habitat mapping studies [14,29,62], which virtually assume that each pixel in remote sensing images is an independent vector in a multi-dimensional environmental space. They mainly employ traditional classification or regression models in the data mining and machine learning field, to make predictions at the pixel level. This assumption is reasonable when scientists want to identify, verify, and explain the habitat characteristic pixel by pixel. However, remote sensing images are highly structured, and their pixels exhibit strong dependencies. The neglect of structural information will significantly reduce the mapping accuracy. Our experiment has confirmed this statement. So far, structural information has been less considered in habitat mapping studies.

Next, we would like to delve into the substance of our end-to-end deep learning framework. We interpret our model from two parts: the convolutional encoder–decoder, and the adversarial framework. Briefly, the convolutional encoder–decoder is the main body to implement the mapping, and the adversarial framework is a superior training strategy. Similar to the original CNN [37] which combines feature extraction, feature selection, and classifier into one network, the convolutional encoder–decoder merges feature extraction, feature selection, latent code interpretation, and image reconstruction into one end-to-end model and trains them together. On the contrary, traditional habitat studies often carry out every stage separately. The end-to-end deep learning models have achieved great success in computer vision [70,71] and remote sensing applications [72]. We believe that this type of model could provide a new solution to habitat mappings as well. Regarding the adversarial framework, its key feature is the adversarial loss constructed by both the generator and discriminator. The adversarial loss can be viewed as a high-level goal which covers many low-level losses [41], and therefore brings a better result. In our supervised model, the adversarial loss is actually a superior objective to train the convolutional encoder–decoder. We effectively implement the image-based mapping by combining the advanced network architecture with a high-level training objective, which leads to the success in mapping the complicated targets from remote sensing imagery.

6. Conclusions

In conclusion, we propose a general-purpose framework to simulate kernel-based home range estimators. The experiment demonstrates that our framework could produce visually recognizable and highly accurate results from remote sensing imagery. Our approach could be generalized to map other types of habitat models as well, such as habitat suitability models [73] and habitat potential models [29]. The deep neural network could help to seek the relationship between animal habitat and environmental factors, instead of the GPS data used in these models. Our approach still has some limitations. One important issue is that the selection of input layers mainly relies on expert knowledge. Our framework could hardly provide an explicit ranking for input layers due to its deep convolutional architecture. In future work, we will attempt to incorporate more effective selection strategies into our framework to improve the mapping performance.

Author Contributions: Conceptualization, R.Z. (Ruobing Zheng) and G.W.; Methodology, R.Z. (Ruobing Zheng); Software, R.Z. (Ruobing Zheng); Validation, G.W., R.Z. (Renyu Zhang); Formal Analysis, R.Z. (Ruobing Zheng), G.W.; Investigation, R.Z. (Ruobing Zheng); Resources, Z.L.; Data Curation, Z.L.; Writing—Original Draft Preparation, R.Z. (Ruobing Zheng); Writing—Review and Editing, C.Y., G.W., R.Z. (Ruobing Zheng); Visualization, R.Z. (Renyu Zhang); Supervision: B.Y.; Project Administration, Z.L., B.Y.; Funding Acquisition, Z.L., B.Y.

Funding: Funding was provided by the Natural Science Foundation of China [61361126011, 90912006]; the National R&D Infrastructure and Facility Development Program of China, “Fundamental Science Data Sharing Platform” [DKA2018-12-02-XX]; the Strategic Priority Research Program of the Chinese Academy of Sciences [Grant No. XDA19060205], Around Five Top Priorities of “One-Three-Five” Strategic Planning, CNIC [Grant No. CNIC_PY-1408, PY-1409]; the Special Project of Informatization of Chinese Academy of Sciences [XXH13505-03-205, XXH13506-303, XXH13506-305]; Science and Technology Service Network Initiative, CAS [Y82E01]; Open Research Fund Program of State Key Laboratory of Hydrosience and Engineering [sklhse-2017-B-03].

Acknowledgments: For logistics and field support, we are grateful to the following groups and individuals: the Institute of Automation, Chinese Academy of Sciences (S. Xiang), the Qinghai Lake National Nature Reserve staff (Z. Xing, D. Zhang), Qinghai Forestry Bureau (S. Li).

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Burt, W.H. Territoriality and home range concepts as applied to mammals. *J. Mammal.* **1943**, *24*, 346–352. [[CrossRef](#)]
2. Katajisto, J.; Moilanen, A. Kernel-based home range method for data with irregular sampling intervals. *Ecol. Model.* **2006**, *194*, 405–413. [[CrossRef](#)]
3. Kenward, R.E. *A Manual for Wildlife Radio Tagging*; Academic Press: Cambridge, MA, USA, 2000.
4. White, G.C.; Garrott, R.A. *Analysis of Wildlife Radio-Tracking Data*; Elsevier: Amsterdam, The Netherlands, 2012.
5. Marzluff, J.M.; Millsaugh, J.J.; Hurvitz, P.; Handcock, M.S. Relating resources to a probabilistic measure of space use: Forest fragments and Steller’s jays. *Ecology* **2004**, *85*, 1411–1427. [[CrossRef](#)]
6. Seaman, D.E.; Powell, R.A. An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology* **1996**, *77*, 2075–2085. [[CrossRef](#)]
7. Worton, B.J. Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* **1989**, *70*, 164–168. [[CrossRef](#)]
8. Getz, W.M.; Fortmann-Roe, S.; Cross, P.C.; Lyons, A.J.; Ryan, S.J.; Wilmers, C.C. LoCoH: Nonparametric kernel methods for constructing home ranges and utilization distributions. *PLoS ONE* **2007**, *2*, e207. [[CrossRef](#)] [[PubMed](#)]
9. Marzluff, J.M.; Knick, S.T.; Millsaugh, J.J. High-tech behavioral ecology: Modeling The distribution of animal activities to better understand wildlife space use and resource selection. In *Radio Tracking and Animal Populations*; Elsevier: Amsterdam, The Netherlands, 2001; pp. 309–326.
10. Takekawa, J.Y.; Newman, S.H.; Xiao, X.; Prosser, D.J.; Spragens, K.A.; Palm, E.C.; Yan, B.; Li, T.; Lei, F.; Zhao, D.; et al. Migration of waterfowl in the East Asian flyway and spatial relationship to HPAI H5N1 outbreaks. *Avian Dis.* **2010**, *54*, 466–476. [[CrossRef](#)] [[PubMed](#)]
11. Harris, S.; Cresswell, W.; Forde, P.; Trehella, W.; Woollard, T.; Wray, S. Home-range analysis using radio-tracking data—A review of problems and techniques particularly as applied to the study of mammals. *Mamm. Rev.* **1990**, *20*, 97–123. [[CrossRef](#)]
12. Powell, R.A.; Mitchell, M.S. What is a home range? *J. Mammal.* **2012**, *93*, 948–958. [[CrossRef](#)]
13. Andréfouët, S. Coral reef habitat mapping using remote sensing: A user vs. producer perspective. Implications for research, management and capacity building. *J. Spat. Sci.* **2008**, *53*, 113–129. [[CrossRef](#)]
14. Maleki, S.; Soffianian, A.R.; Koupaei, S.S.; Saatchi, S.; Pourmanafi, S.; Sheikholeslam, F. Habitat mapping as a tool for water birds conservation planning in an arid zone wetland: The case study Hamun wetland. *Ecol. Eng.* **2016**, *95*, 594–603. [[CrossRef](#)]
15. Nagendra, H.; Lucas, R.; Honrado, J.P.; Jongman, R.H.; Tarantino, C.; Adamo, M.; Mairota, P. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecol. Indic.* **2013**, *33*, 45–59. [[CrossRef](#)]
16. Van Winkle, W. Comparison of several probabilistic home-range models. *J. Wildl. Manag.* **1975**, *39*, 118–123. [[CrossRef](#)]
17. Ford, R.G.; Krumme, D.W. The analysis of space use patterns. *J. Theor. Biol.* **1979**, *76*, 125–155. [[CrossRef](#)]
18. Calenge, C. *Home Range Estimation in R: The adehabitatHR Package*; Office National De La Classe Et De La Faune Sauvage: Auffargis, France, 2011.
19. Epanechnikov, V.A. Non-parametric estimation of a multivariate probability density. *Theory Probab. Appl.* **1969**, *14*, 153–158. [[CrossRef](#)]
20. Bullard, F. Estimating the Home Range of an Animal: A Brownian Bridge Approach. Ph.D. Thesis, Johns Hopkins University, Baltimore, MD, USA, 1999.
21. Guisan, A.; Zimmermann, N.E. Predictive habitat distribution models in ecology. *Ecol. Model.* **2000**, *135*, 147–186. [[CrossRef](#)]
22. Barry, S.; Elith, J. Error and uncertainty in habitat models. *J. Appl. Ecol.* **2006**, *43*, 413–423. [[CrossRef](#)]

23. Varela, R.D.; Rego, P.R.; Iglesias, S.C.; Sobrino, C.M. Automatic habitat classification methods based on satellite images: A practical assessment in the NW Iberia coastal mountains. *Environ. Monit. Assess.* **2008**, *144*, 229–250. [[CrossRef](#)] [[PubMed](#)]
24. Haest, B.; Thoonen, G.; Borre, J.V.; Spanhove, T.; Delalieux, S.; Bertels, L.; Kooistra, L.; Múcher, C.; Scheunders, P. An object-based approach to quantity and quality assessment of heathland habitats in the framework of Natura 2000 using hyperspectral airborne AHS images. In Proceedings of the GEOBIA 2010 Conference, Ghent, Belgium, 29 June–2 July 2010.
25. Lucas, R.; Medcalf, K.; Brown, A.; Bunting, P.; Breyer, J.; Clewley, D.; Keyworth, S.; Blackmore, P. Updating the Phase 1 habitat map of Wales, UK, using satellite sensor data. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 81–102. [[CrossRef](#)]
26. Beutel, T.; Beeton, R.; Baxter, G. Building better wildlife-habitat models. *Ecography* **1999**, *22*, 219. [[CrossRef](#)]
27. Drew, C.A.; Wiersma, Y.F.; Huettmann, F. *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*; Springer Science & Business Media: Berlin, Germany, 2010.
28. Hyde, P.; Dubayah, R.; Walker, W.; Blair, J.B.; Hofton, M.; Hunsaker, C. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sens. Environ.* **2006**, *102*, 63–73. [[CrossRef](#)]
29. Lee, S.; Lee, S.; Song, W.; Lee, M.J. Habitat Potential Mapping of Marten (*Martes flavigula*) and Leopard Cat (*Prionailurus bengalensis*) in South Korea Using Artificial Neural Network Machine Learning. *Appl. Sci.* **2017**, *7*, 912. [[CrossRef](#)]
30. Chegoonian, A.; Mokhtarzade, M.; Valadan Zoej, M. A comprehensive evaluation of classification algorithms for coral reef habitat mapping: Challenges related to quantity, quality, and impurity of training samples. *Int. J. Remote Sens.* **2017**, *38*, 4224–4243. [[CrossRef](#)]
31. Pastick, N.J.; Jorgenson, M.T.; Wylie, B.K.; Nield, S.J.; Johnson, K.D.; Finley, A.O. Distribution of near-surface permafrost in Alaska: Estimates of present and future conditions. *Remote Sens. Environ.* **2015**, *168*, 301–315. [[CrossRef](#)]
32. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv* **2017**, arXiv:1703.10593.
33. Sangkloy, P.; Lu, J.; Fang, C.; Yu, F.; Hays, J. Scribbler: Controlling deep image synthesis with sketch and color. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; Volume 2.
34. Gatys, L.A.; Ecker, A.S.; Bethge, M. Image style transfer using convolutional neural networks. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 2414–2423.
35. Pathak, D.; Krahenbuhl, P.; Donahue, J.; Darrell, T.; Efros, A.A. Context encoders: Feature learning by inpainting. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 2536–2544.
36. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
37. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
38. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.
39. Sohn, K.; Lee, H.; Yan, X. Learning structured output representation using deep conditional generative models. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 3483–3491.
40. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.
41. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. *arXiv* **2016**, arXiv:1611.07004.

42. Zhu, J.Y.; Zhang, R.; Pathak, D.; Darrell, T.; Efros, A.A.; Wang, O.; Shechtman, E. Toward multimodal image-to-image translation. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 465–476.
43. Mao, X.; Li, Q.; Xie, H.; Lau, R.Y.; Wang, Z.; Smolley, S.P. Least squares generative adversarial networks. *arXiv* **2016**, arXiv:1611.04076.
44. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [[CrossRef](#)] [[PubMed](#)]
45. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin, Germany, 2015; pp. 234–241.
46. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv* **2016**, arXiv:1603.04467.
47. Chen, H.; Smith, G.; Zhang, S.; Qin, K.; Wang, J.; Li, K.; Webster, R.; Peiris, J.; Guan, Y. Avian flu: H5N1 virus outbreak in migratory waterfowl. *Nature* **2005**, *436*, 191–192. [[CrossRef](#)] [[PubMed](#)]
48. Liu, J.; Xiao, H.; Lei, F.; Zhu, Q.; Qin, K.; Zhang, X.W.; Zhang, X.L.; Zhao, D.; Wang, G.; Feng, Y.; et al. Highly pathogenic H5N1 influenza virus infection in migratory birds. *Science* **2005**, *309*, 1206. [[CrossRef](#)] [[PubMed](#)]
49. Liu, D.; Zhang, G.; Qian, F.; Hou, Y.; Dai, M.; Jiang, H.; Lu, J.; Xiao, W. Population, distribution and home range of wintering bar-headed goose along Yaluzan, bu River. *Tibet. Acta Ecol. Sin.* **2010**, *30*, 4173–4179.
50. Butler, D. Doubts hang over source of bird flu spread. *Nature* **2006**, *439*, 772. [[CrossRef](#)] [[PubMed](#)]
51. Justice, C.O.; Vermote, E.; Townshend, J.R.; Defries, R.; Roy, D.P.; Hall, D.K.; Salomonson, V.V.; Privette, J.L.; Riggs, G.; Strahler, A.; et al. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1228–1249. [[CrossRef](#)]
52. Dong, Z.; Wang, Z.; Liu, D.; Li, L.; Ren, C.; Tang, X.; Jia, M.; Liu, C. Assessment of habitat suitability for waterbirds in the West Songnen Plain, China, using remote sensing and GIS. *Ecol. Eng.* **2013**, *55*, 94–100. [[CrossRef](#)]
53. Zhang, W.; Li, X.; Yu, L.; Si, Y. Multi-scale habitat selection by two declining East Asian waterfowl species at their core spring stopover area. *Ecol. Indic.* **2018**, *87*, 127–135. [[CrossRef](#)]
54. Cappelle, J.; Girard, O.; Fofana, B.; Gaidet, N.; Gilbert, M. Ecological modeling of the spatial distribution of wild waterbirds to identify the main areas where avian influenza viruses are circulating in the Inner Niger Delta, Mali. *EcoHealth* **2010**, *7*, 283–293. [[CrossRef](#)] [[PubMed](#)]
55. Tieleman, T.; Hinton, G. RMSprop Gradient Optimization. Available online: http://www.cs.toronto.edu/tijmen/csc321/slides/lecture_slides_lec6.pdf (accessed on 30 October 2018).
56. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
57. Altman, N.S. An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* **1992**, *46*, 175–185.
58. Rokach, L.; Maimon, O.Z. *Data Mining with Decision Trees: Theory and Applications*; World Scientific: Singapore, 2008; Volume 69.
59. Kobler, A.; Adamic, M. Identifying brown bear habitat by a combined GIS and machine learning method. *Ecol. Model.* **2000**, *135*, 291–300. [[CrossRef](#)]
60. Ho, T.K. Random decision forests. In Proceedings of the Third International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp. 278–282.
61. Garzon, M.B.; Blazek, R.; Neteler, M.; De Dios, R.S.; Ollero, H.S.; Furlanello, C. Predicting habitat suitability with machine learning models: The potential area of *Pinus sylvestris* L. in the Iberian Peninsula. *Ecol. Model.* **2006**, *197*, 383–393. [[CrossRef](#)]
62. Vincenzi, S.; Zucchetta, M.; Franzoi, P.; Pellizzato, M.; Pranovi, F.; De Leo, G.A.; Torricelli, P. Application of a Random Forest algorithm to predict spatial distribution of the potential yield of *Ruditapes philippinarum* in the Venice lagoon, Italy. *Ecol. Model.* **2011**, *222*, 1471–1478. [[CrossRef](#)]
63. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a deep convolutional network for image super-resolution. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 184–199.

64. Wang, Y.Q. A multilayer neural network for image demosaicking. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 30 October 2014; pp. 1852–1856.
65. Walker, J.; Doersch, C.; Gupta, A.; Hebert, M. An uncertain future: Forecasting from static images using variational autoencoders. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 835–851.
66. Xue, T.; Wu, J.; Bouman, K.; Freeman, B. Visual dynamics: Probabilistic future frame synthesis via cross convolutional networks. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 91–99.
67. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
68. Ridgeway, K.; Snell, J.; Roads, B.; Zemel, R.S.; Mozer, M.C. Learning to generate images with perceptual similarity metrics. *arXiv* **2015**, arXiv:1511.06409.
69. Zhao, H.; Gallo, O.; Frosio, I.; Kautz, J. Loss functions for image restoration with neural networks. *IEEE Trans. Comput. Imaging* **2017**, *3*, 47–57. [[CrossRef](#)]
70. Huang, L.; Yang, Y.; Deng, Y.; Yu, Y. Densebox: Unifying landmark localization with end to end object detection. *arXiv* **2015**, arXiv:1509.04874.
71. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 91–99.
72. Maggiori, E.; Tarabalka, Y.; Charpiat, G.; Alliez, P. Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 645–657. [[CrossRef](#)]
73. Calenge, C.; Darmon, G.; Basille, M.; Loison, A.; Jullien, J.M. The factorial decomposition of the Mahalanobis distances in habitat selection studies. *Ecology* **2008**, *89*, 555–566. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).