

Satellite image inpainting with deep generative adversarial neural networks

Mohamed Akram Zaytar, Chaker El Amrani

Department of Computer Engineering, Faculty of Sciences and Technologies, Tangier, Morocco

Article Info

Article history:

Received Aug 20, 2020

Revised Jan 29, 2021

Accepted Feb 11, 2021

Keywords:

Air pollution

Generative adversarial nets

Image inpainting

Neural networks

Satellite imagery

ABSTRACT

This work addresses the problem of recovering lost or damaged satellite image pixels (gaps) caused by sensor processing errors or by natural phenomena like cloud presence. Such errors decrease our ability to monitor regions of interest and significantly increase the average revisit time for all satellites. This paper presents a novel neural system based on conditional deep generative adversarial networks (cGAN) optimized to fill satellite imagery gaps using surrounding pixel values and static high-resolution visual priors. Experimental results show that the proposed system outperforms traditional and neural network baselines. It achieves a normalized least absolute deviations error of $L_1 = 0.33$ (21% and 60% decrease in error compared with the two baselines) and a mean squared error loss of $L_2 = 0.15$ (29% and 73% decrease in error) over the test set. The model can be deployed within a remote sensing data pipeline to reconstruct missing pixel measurements for near-real-time monitoring and inference purposes, thus empowering policymakers and users to make environmentally informed decisions.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Mohamed Akram Zaytar

Department of Computer Engineering

Faculty of Sciences and Technologies

Route Ziaten, Tangier, Morocco

Email: m.zaytar@uae.ac.ma

1. INTRODUCTION

Climate change poses serious challenges that threaten humanity's long-term safety [1]. Addressing these challenges depends on breakthroughs in environmental policy and climate science [2]. Climate research is key to understanding the long-term effects of global warming on agriculture [3-4], food security [5], air quality [6], and weather conditions [7]. On the other hand, one of the primary data sources that empower environmental research is satellite imagery [8]. Satellite programs such as landsat [9], sentinel, aqua/terra, among others, provide a wealth of freely available data sets for the masses. This tremendous progress has unlocked many innovations that extract valuable insights [10-11] from satellite imagery using big data pipelines and advanced machine learning systems [12].

Satellite sensors are limited by their Spatio-temporal resolution. A satellite's temporal resolution represents the duration of getting information about the same point on earth. On the other hand, spatial resolution specifies the surface size of 1 pixel of information (ex. Sentinel-2 has an RGB spatial resolution of 10×10 m per pixel). Due to various reasons, most satellite imagery contains "holes" or "gaps" of missing pixel values. Clouds are the primary contributor to such noise. Satellite noise is challenging because it worsens the satellite's temporal resolution and introduces uncertainty into atmospheric monitoring pipelines. Many have resorted to using IoT sensors [13] that provide a higher-quality ground-level stream of

measurements [14]. However, ground sensors can only give information about a specific location and, as a result, do not have the geographic coverage that satellites have (most polar satellites cover the whole earth).

Remote sensors are the primary data source for large-scale atmospheric monitoring and, more specifically, air quality monitoring. Enhancing the Spatio-temporal resolution of satellite sensors is of critical importance since it enables greater visibility over the state of planet earth. For this reason, this study focuses on inpainting satellite NO₂ images. Each NO₂ pixel measures the atmospheric NO₂ vertical density (in Dobson units) over the pixel. NO₂ is a trace gas that negatively affects air quality and the climate. It is linked to road traffic and industrial activities such as fossil fuel combustion [15]. A high NO₂ concentration can cause numerous respiratory diseases [16].

This paper proposes a generative adversarial system used to fill the missing gaps in images based on the image's content (available pixel values) and high-resolution visual priors. The system's novelty lies in its use of a different data modality (higher-resolution static RGB images) encoded by a conditional layer to provide auxiliary features to the completer network. As a result, the neural system inpaints all future images and pushes the sensor's temporal resolution to its theoretical limit (nullifying the effects of clouds or sensory errors). The paper's main contributions are outlined as:

- A cGAN-based neural system for inpainting multi-spectral satellite imagery.
- A full description of the pre-inference data preprocessing pipeline.
- Case study: the method is evaluated on NO₂ pollution images for near-real-time air quality monitoring, showcasing the potential of fusing multi-modal satellite data using neural approximators.

The rest of the paper is structured as; "Related works" describes the most notable research efforts that tackle image inpainting. "Research method" introduces the neural system architecture, the training algorithm, and the data set. "Results and discussion" describes synthetic noise generation, introduces the performance metrics, and presents the final results. Finally, it provides an intuitive understanding of the effects of priors and their limitations. "Conclusion" summarizes the paper and describes future work.

2. RELATED WORKS

The existing research literature on image inpainting can be grouped into two main parts. Non-learning methods such as diffusion/patch-based algorithms, and the relatively recent work that attempts to learn inpainting by training convolutional neural network-based architectures (CNNs). This section outlines the most notable efforts from both sides.

2.1. Diffusion or patch-based methods

The early success in image inpainting is attributed to information propagation techniques through patch similarity or variational methods. Efros and Leung [17] proposed to model image textures as Markov random fields then use similarity search to fill the missing pixels. Other efforts [18-19] were directed toward inpainting images through their texture and structure using search-based guided propagation that synthesizes patterns resembling other image regions or other images within a searchable database.

Variational methods are also present in [20] that use feature extractors such as patch statistics, colors, and gradients to synthesize the missing image gaps. Lastly, out-of-sample inpainting was achieved by [21] using an extensive database of images. Its algorithm inpaints missing regions of an image by finding similar images then diffusing extracted low-level features. Unlike others, this technique can suggest multiple completions based on the chosen database item.

This class of methods works well on images that contain repeated or static patterns (examples: sand, grid, paper) but fails on images with rich semantic content. Furthermore, automatic non-learning algorithms cannot inpaint abstractions that make complex images cohesive in their content, and their use of out-of-sample information is limited due to their local dependencies.

2.2. Learning-based approaches

One of the earliest efforts to use representation learning for image inpainting proposed a multi-layer perceptron (MLP) architecture to fill missing pixels in gray-scale images by minimizing the reconstruction loss [22]. The paper established the importance of masking missing pixels and the potential of neural networks (NNs) in image completion. Furthermore, Xu *et al.* [23] used a CNN architecture to propose a general method for solving three tasks: image inpainting, denoising, and image degradation recovery.

Recently, neural networks trained using pixel-wise reconstruction error and adversarial loss reported promising results. The work of [24] introduced context encoders to fill large holes in image centers. Yang *et al.* [25] enabled high-resolution image inpainting by proposing joint content and texture losses. Xu *et al.* [26] combined local and global discriminators into one network and used convolutions and dilated convolutions to inpaint images.

Yu *et al.* [27] improved the previous architecture by dividing the generation process into two stages. The first outputs a blurry image optimized with spatial discounted \mathcal{L}_1 reconstruction loss, and the second refines and outputs the final image. The authors used the network's output as input to the global and local discriminators and chose wasserstein GANs (WGAN) to train the neural system (WGAN stabilizes the overall optimization process). Finally, Xu *et al.* [28] improved the previous architecture by incorporating contextual attention and dilated gated convolutions into both the coarse and refinement networks.

Although the mentioned neural systems provide impressive inpaintings and predict high-quality visual semantics, none have experimented with priors or extended the generator/discriminator with a conditional layer. Furthermore, all of the mentioned methods assume one source data distribution to be modeled, as shown in Table 1. This study establishes the importance of using different data modalities and fusing them through a conditional layer to solve image inpainting in general, and air quality estimation specifically.

Table 1. A Comparison of different algorithms that aim to solve image inpainting

Features\Methods	Non-Parameteric Sampler [17]	PatchMatch [29]	Mask-FC Inpainter [22]	LocalGlobal [26]	Ours
Free-Form	✓	✓	✓	✓	✓
Out-of-Sample		✓	✓	✓	✓
Semantics			✓	✓	✓
Multi-Modal					✓

3. RESEARCH METHOD

Two CNN-based network architectures were trained within a conditional adversarial framework as shown in Figure 1. The generator network, responsible for filling the missing gaps using contextual information and static priors, and an auxiliary discriminator network trained to distinguish between real and completed pollution patches. Both networks are conditioned over true-color imagery that corresponds to the region covering the input patch. The prior is encoded by a conditional layer (reducer). The input to the generator consists of a damaged image (x) and its high-resolution prior (p). The reducer network compresses p to the same size of x then stacks both for inpainting. The discriminator network takes either a healthy or a completed image with its encoded prior. The discriminator judges if an image is real or completed.

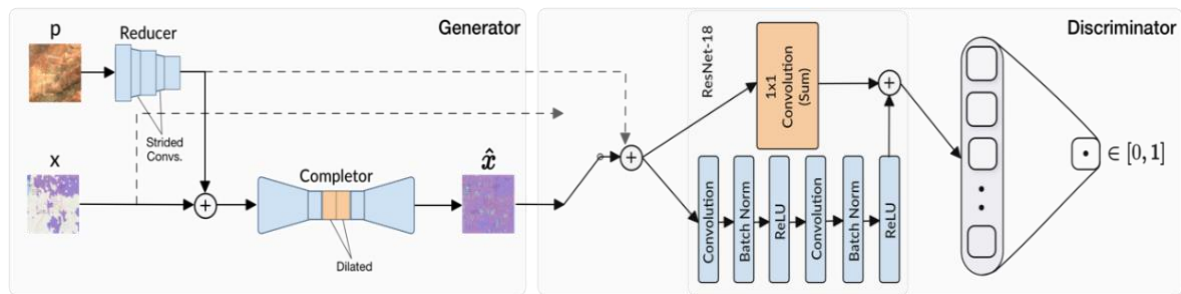


Figure 1. System overview

3.1. Convolutional neural networks

The reducer, completer, and discriminator networks are based on convolutional neural networks (CNN). CNNs are a special type of neural network that uses weight sharing to extract hierarchical visual features with minimal free parameters and maximal local connections. Kernel weights are optimized to produce activations that help in the final prediction task. CNNs are capable of progressively extracting higher-order abstractions that serve to minimize a pre-defined objective function. A specific activation is calculated using (1).

$$Y_{m,n} = \sigma\left(b + \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} K_{i,j} X_{m+i,n+j}\right) \quad (1)$$

With X representing the input, K the kernel (matrices of learnable weights), s is the kernel size, and (m, n) are the indices of the target value $Y_{m,n}$ in the activation layer. Dilated convolutional layers are also

used to simulate a larger receptive field without adding more parameters. To calculate dilated activations, one parameter is added to the previous definition: η , which is the dilation factor as (2).

$$Y_{m,n} = \sigma\left(b + \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} K_{i,j} X_{m+\eta i, n+\eta j}\right) \quad (2)$$

3.2. Conditional generative adversarial networks

Generative adversarial networks (GAN) are a class of neural networks trained in an adversarial manner. A GAN consists of two networks: a generative network $G(\cdot)$ that learns the true data distribution (the process that generated the training data) and a discriminative network $D(\cdot)$ that estimates if a sample came from the true data distribution or $G(\cdot)$. Ideally, both G and D are trained simultaneously, G 's parameters are adjusted to minimize $\text{Log}(1 - D(G(z)))$ (i.e., to fool the discriminator), and D 's parameters are tuned to maximize $\text{Log}(D(x))$ (i.e., to detect fake generated inputs). D and G play the following two-player minimax game with value function $V(G, D)$ as (3).

$$\min_G \max_D V(G, D) = E_{x \sim p_d(x)} [\log(D(x))] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (3)$$

In the context of this study, visual priors are of higher resolution than pollution patches. Downsampling visual imagery to fit pollution patches will result in losing much of its encoded information. Additionally, from the perspective of a vanilla GAN (i.e., an unconditioned GAN), there is no control over data modes during the inpainting process. Inputting pollution images without pixel-level meta-data will result in a model that mimics a general-purpose interpolator. However, by conditioning the output over its region's visual imagery, the model can produce accurate inpaintings by finding correlations between priors and pollution patches.

As a result, the generative adversarial network is extended with a conditional layer to encode the static priors. The generator and discriminator are provided with high-resolution encoded imagery (priors: p). The objective function of the two-player minimax game is updated as (4).

$$\min_G \max_D V(G, D) = E_{x \sim p_d(x)} [\log(D(x|p))] + E_{z \sim p_z(z)} [\log(1 - D(G(z|p)|p))] \quad (4)$$

In a conditional generative adversarial setup, the same condition is provided to both the generator and discriminator networks. Priors are purposely used as conditions to help the generator enhance its completions. For example, one can imagine how useful vehicle traffic density images would be for a model that predicts near-real-time NO_2 concentrations. In this case, RGB images provide low-level information about urban and greenness densities. This study argues that a visual prior could be useful to the task of predicting NO_2 densities over large regions of interest.

3.3. Completion network

The completion network takes low-resolution NO_2 images that contain the gaps to be filled, and a mask channel that indicates which pixels are missing. Each damaged patch has its corresponding high-resolution RGB image that covers the same region and provides gap-free visual information. Two networks were trained. The reducer acts as a down-sampler that intelligently resizes the high-resolution RGB image (the prior) to the same size as the damaged image. Table 2 presents its layers in successive order. The completer network is a fully convolutional network (FCN) that acts as the main inpainter. It is optimized to fill the missing gaps in the input image. Table 3 specifies its layers. The activations for both networks were passed through batch normalization and ReLU after each layer.

Table 2. Reducer network layers

N°	Type	Kernel	Stride	Output
1	Conv.	5	1	16
2	Conv.	3	2	32
3	Conv.	3	1	64
4	Conv.	3	2	32
5	Conv.	3	1	1

Table 3. Completor network layers

Nº	Type	Kernel	Stride	Dilation	Output
1	Conv.	5	1	1	16
2	Conv.	3	2	1	32
3	Conv.	3	1	1	64
4	Conv.	3	2	1	64
5	Diconv.	3	1	2	64
6	Diconv.	3	1	3	64
7	Conv.	3	1	1	64
8	$conv^T$	4	2	1	32
9	Conv.	3	1	1	32
10	$conv^T$	4	2	1	16
11	Conv.	3	1	1	8
12	Conv.	3	1	1	1

The completor network is used to inpaint the missing regions in the input. On the other hand, the reducer network resizes the prior to the same size as the input. Without it, the model would learn the unconditioned pollution image distribution, which is not optimal for location-variant patterns. Urban, land, and other visual features serve as strong priors for the completor to generate accurate patches. The completor network was trained using mean squared error loss (MSE) averaged over the masked (gap) pixels.

3.4. Discriminator network

The discriminator network is trained to detect completed NO₂ images. A ResNet-18 [30] architecture shown in Figure 1 is used to extract the feature vector which is mapped to the probability of the input being completed (fake) or real. Reducer network weights are frozen while optimizing the discriminator since the reducer is optimized for efficient inpainting, not discrimination. The primary role of RGB images in the context of the discriminator is to provide a useful prior that is independent of whether the input image is real or not. Hence, the discriminator is optimized to estimate P(input = completed|p).

3.5. Training

The completor network is denoted: $C(x, p)$. x represents a batch of NO₂ images with masks M comprised of 0s and 1s, with 1s representing the pixels that are missing in x . p are the priors for each image in x . They consist of many RGB images for the corresponding ROIs. Similarly, $D(\bar{x}, p)$ designates the discriminator network, \bar{x} represents the pollution images (real or completed), and p the encoded priors over \bar{x} 's regions of interests (ROIs).

Mean squared error loss (\mathcal{L}_2) is an inpainting loss choice that results in blurred estimations over the gaps. It averages the squared differences between gap pixel predictions and targets as (5).

$$\mathcal{L}_2(x, \hat{x}) = \left| \left| M_c \odot (C(\bar{x}, p) - x) \right| \right|^2 \quad (5)$$

On the other hand, adversarial loss can be formulated as (6).

$$\min_c \max_D E [\log(D(x|p)) + \log(1 - D(C(\bar{x}|p)|p))] \quad (6)$$

C is the completor network, D is the discriminator, x is the input, \bar{x} is the damaged/healthy input, and p the prior. \mathcal{L}_2 and adversarial losses are combined to formalize the general optimization problem as (7).

$$\min_c \max_D E [\mathcal{L}_2(x, p, C) + \log(D(x|p)) + \log(1 - D(C(x|p)|p))] \quad (7)$$

GANs are challenging to train due to the instability between the generator and discriminator networks in the early training phase. For this reason, the training loop is balanced as described in Algorithm 1.

The method proposed in [26] is chosen as the neural baseline. Its model was trained to produce visually appealing completions for a variety of natural scene images. It serves as a good benchmark because the proposed architecture is an extension of the baseline's modular design.

Algorithm 1: T_C & T_D serve to pre-train each network separately before conducting adversarial training

Data: X, M_c, M_d, p

Hyperparameters: $T_{max} = 10,000, T_C = 100, T_D = 100, \eta = 0.001, \eta' = 0.01$

Begin

```

For  $t_i \in [0, T_{max}]$  do
  Sample minibatch  $x \subset X$ 
  Generate completer masks  $M_c$  for  $\forall x_i \in x$ 
  Generate discriminator masks  $M_d$  for  $\forall x_i \in x$ 
  If  $t_i < T_{max} - T_C$  then
     $W_C \leftarrow W_C - \eta \nabla_{W_C} L_2(x, p, C)$ 
  Else if  $t_i < T_{max} - T_D$  then
     $W_D \leftarrow W_D - \eta' \nabla_{W_D} BCE(\bar{x}, p, D)$ 
  Else
    Update  $W_C$  with joint loss
    Update  $W_D$  with binary cross entropy loss

```

3.6. Data

The European organization of the exploitation of meteorological satellites (EUMETSAT) is an international satellite agency responsible for acquiring, preprocessing, and distributing reliable weather, climate, and environmental data. Its low-orbiting satellite, MetOp, continuously delivers critical climate data. EUMETSAT also distributes data from other partners such as the national oceanographic and atmospheric administration (NOAA). Offline EUMETSAT's data products are free and available for research purposes.

MetOp is a series of 3 polar-orbiting meteorological satellites developed by the European space agency (ESA) and operated by EUMETSAT. MetOp takes 90 minutes to orbit the earth, totaling 14 times a day. Having three satellites enhances the temporal resolution of MetOp as a data provider. MetOp carries a payload of 11 scientific instruments. After transferring the data, it gets preprocessed into multiple levels and fed into numerical simulators for weather forecasting and environmental monitoring. Many of its available data products provide vertical density measurements.

Gas traces were acquired from the global ozone monitoring experiment-2 (GOME-2) instrument, a scanning spectrometer that provides global monitoring coverage. The near real-time total column (NTO) product provides concentration measurements for four types of atmospheric trace gases: O_3 , NO_2 , SO_2 , and $HCHO$. The product is operational since 01/12/2007, has a spectral resolution of $0.26 - 0.51\text{nm}$, and provides global geographic coverage with a spatial resolution of $40 \times 40\text{km}$ (MetOp-A).

On the other hand, Meteosat is a series of geostationary meteorological satellites operated by EUMETSAT. Meteosat second generation (MSG) provides images of the full earth disc and data for weather forecasts. It has a temporal resolution of 15 minutes. The spinning enhanced visible and infrared imager (SEVIRI) instrument captures the true-color images in a spatial resolution of $1 \times 1\text{km}$. The prior's (p) imagery is collected and preprocessed from MSG's SEVIRI instrument.

MSG covers the ROI of Morocco. The tiles were clipped using the region of interest and merged through pixel-averaging to store a single (mosaic) high-quality image over the ROI.

In this study, Morocco was chosen as a region of interest. All images were filtered to be in the bounding box $[(-5.39, 35.54), (-5.34, 35.54), (-5.34, 35.59), (-5.39, 35.59), (-5.39, 35.54)]$ in (latitude, longitude) coordinates. All SEVIRI tiles were taken from 06/2017 to 01/2018. For pollution images, tiles were acquired and filtered for the same ROI that range from 03/2018 to 09/2018. The priors were sampled from a previous timeframe because they will be used to inpaint future patches. Synthetic mask generation is explained in the "RESULTS AND DISCUSSION" section.

T is denoted as the set of acquired tiles for the region of interest. for each $t_i \in T$, it is processed as follows:

- t_i is projected into a pre-defined static spatial grid to normalize pixel positions.
- t_i is split into N 64×64 patches, denoted x_i .
- For each image $x_{ij} \in x_i$, its 256×256 RGB prior (p_{ij}) is collected. It covers the same region as x_{ij} .
- $\{x_{ij}, p_{ij}\}$ are split into two sets. The first contains healthy pollution patches (all pixel values are available). The second set holds the damaged patches.
- Pixel values were normalized and standardized between $[0, 1]$ for both input patches x_i and priors p_i .

The final data set was clipped at 1M NO_2 patches. 90% of the data set was used to train the model and the remaining 100,000 images for testing. A time-series train/test split was conducted where all of the 10% test samples came from a future time-interval.

4. RESULTS AND DISCUSSION

4.1. Noise generation

As opposed to the task of natural image inpainting, where not much consideration is given to the shape of the gap mask, the geometry of missing regions in satellite imagery follows strong patterns (clouds, holes, and lines). The mask generator should reproduce these patterns in training time. As a result, a separate GAN G_n was trained using the isolated noised patches to learn the natural noise distribution. G_n serves as a noise mask generator that is used to sample artificial gaps during training. For every healthy patch $x_i \in X$, a 64×64 gap mask is sampled from G_n , a copy of x_i is purposely damaged using the mask. The mask is stacked on top of the damaged image; the final 2-channel image represents the input. The original image x_i becomes the target used in loss calculation.

4.2. Results

The proposed model was benchmarked against two algorithms, PatchMatch [29], which represents the classical state-of-the-art method, and the LocalGlobal [26] neural inpainter. For training, input images of size 64×64 pixels and prior images of size 256×256 pixels were used as input to the model. The training and testing datasets were extracted using a time-series split.

As opposed to natural scene completion, where MAE and MSE are not considered good metrics because many completions can be conceptually possible, in the case of pollution images, sensor measurements are unique targets. Hence, the model was evaluated using two regression metrics; least absolute deviations loss as (8).

$$\mathcal{L}_1 = \frac{1}{n} \sum_{i=1}^n |M_c \odot (C(\bar{x}_i, p_i) - x_i)| \quad (8)$$

\mathcal{L}_1 averages the absolute differences between the target and predicted gap pixels. Mean squared error (\mathcal{L}_2) is also reported. Both metrics measure error as the distance between the model's predictions and the ground-truth normalized targets.

The model achieves a least absolute deviations error of $\mathcal{L}_1 = 0.33$ (21% and 60% decrease in error with respect to the two baselines) and an MSE of $\mathcal{L}_2 = 0.15$ (29% and 73% decrease in error, respectively) over the test set. These results represent a significant performance increase over the PatchMatch and LocalGlobal inpainters. When converted to Dobson Units (DU), we get $\mathcal{L}_1 = 0.56$ DU and $\mathcal{L}_2 = 0.41$ DU. 1 *dobson unit* corresponds to a column density of $2.8 \times 10^{16} \text{ cm}^{-2}$, the raw measurements range between 0 and 2.5 DU.

4.3. Discussion

The main contributor to the increase in performance is the conditional layer. An ablation study was conducted by removing the conditional layer and training the GAN without visual priors. The resulting \mathcal{L}_1 and \mathcal{L}_2 scores were slightly worse than LocalGlobal [26], last column of Table 4.

Table 4. Benchmarks of mean \mathcal{L}_1 and mean squared loss on the test set.

Metrics\Methods	PatchMatch [29]	LocalGlobal [26]	Ours	Ours (<i>no priors</i>)
\mathcal{L}_1	0.83	0.42	0.33	0.44
\mathcal{L}_2	0.55	0.27	0.15	0.28

Figure 2 showcases the effects that visual priors can have on pollution predictions. The model predicts higher NO_2 concentrations in the city of Casablanca without processing high-density neighboring pixels, top row in Figure 2. It predicted high NO_2 concentrations by relying solely on the RGB prior. Such patterns are noticeable in other urban and industrial cities in Morocco. However, at the bottom row, the model failed to predict a high pollution concentration over the region of Taourirt. That could have been the result of the temporal nature of industrial activities and the noise that is inherent to sensory measurements. The average increase in performance indicates that the model learned to associate certain visual features to high/low pollution densities. This study showcases the benefits of multi-modal learning in computer vision.

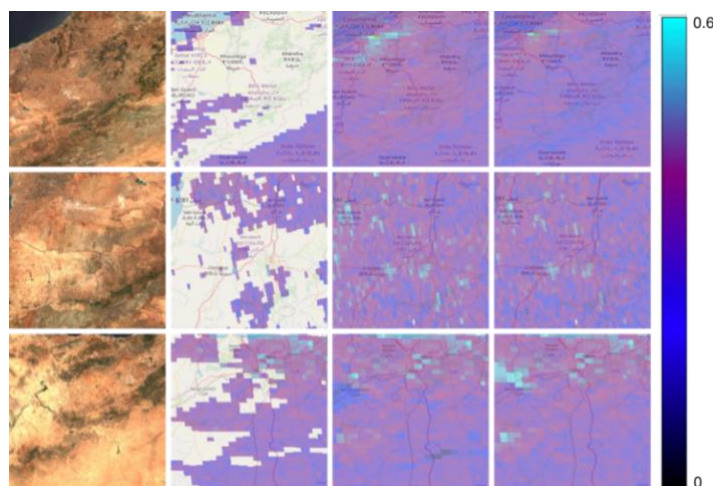


Figure 2. Sample: normalized predictions over northern Morocco. From left to right: prior, input image, generated image, and ground-truth image (higher values represent greater concentrations)

Despite the prior not providing population, time-dependent industry activity estimates, or traffic information, it increased the overall performance by simply providing high-quality visual information. The reducer also played a critical role in transforming the priors in a way that was useful to inpaint the missing regions. The proposed model can also be used as a data enhancer for downstream tasks by removing cloud effects. Downstream applications include object detection [31], landcover generation [32], landuse classification [33], change detection, crop monitoring, field and urban mapping among others.

5. CONCLUSION

This paper proposed a neural image inpainter capable of filling measurement gaps using neighboring pixel values and static priors. It described the preprocessing data pipeline, the CNN-based sub-networks, and the training process. The neural system was successfully trained to fill pollution gaps in the region of Morocco. It outperformed two prior-less baselines and showed the potential of data fusion for satellite image inpainting. The described neural system can be deployed within a remote sensing pipeline to fill incoming satellite patches, resulting in greater near-real-time visibility over weather, atmospheric, and climate conditions. The system intelligently nullifies the effects of sensor perturbations and cloud effects. However, one limitation of the system is that temporal resolution enhancement is not enough to offer a competitive alternative to IoT-based devices for small-scale monitoring, the limited spatial resolution of satellite sensors remains the most important open challenge in remote sensing. However, the model can be modified to tackle super-resolution through prior learning. Such system can enhance the satellite's spatial resolution by using low-resolution source imagery and high-resolution static priors.

REFERENCES

- [1] C. D. Thomas, A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. Erasmus, M. F. De Siqueira, A. Grainger, L. Hannah et al., "Extinction risk from climate change," *Nature*, vol. 427, no. 6970, pp. 145–148, 2004. <https://doi.org/10.1038/nature02121>.
- [2] E. Shove, "Beyond the abc: climate change policy and theories of social change," *Environment and Planning A*, vol. 42, no. 6, pp. 1273–1285, 2010. <https://doi.org/10.1068/a42282>.
- [3] M. Dimiyati, K. Kustiyo, and R. D. Dimiyati, "Paddy field classification with modis-terra multi-temporal image transformation using phenological approach in java island," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 2, pp. 1346–1358, 2019. DOI: <http://doi.org/10.11591/ijece.v9i2.pp1346-1358>.
- [4] S. B. Jadhav, "Convolutional neural networks for leaf image-based plant disease classification," *IAES International Journal of Artificial Intelligence (IJAI)*, vol. 8, no. 4, p. 328, 2019. DOI: <http://doi.org/10.11591/ijai.v8.i4.pp328-341>.
- [5] O. Olaniyi, E. Daniya, J. Kolo, J. Bala, and A. Olanrewaju, "A computer vision-based weed control system for low-land rice precision farming," *International Journal of Advances in Applied Sciences (IJAAS)*, vol. 9, no. 1, pp. 51–61, 2020. DOI: <http://doi.org/10.11591/ijaas.v9.i1.pp51-61>.

- [6] S. M. Bernard, J. M. Samet, A. Grambsch, K. L. Ebi, and I. Romieu, "The potential impacts of climate variability and change on air pollution-related health effects in the united states," *Environmental health perspectives*, vol. 109, no. Suppl 2, pp. 199–209, 2001. <https://doi.org/10.2307/3435010>.
- [7] W. Suparta and W. S. Putro, "Comparison of tropical thunderstorm estimation between multiple linear regression, Dvorak, and anfis," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 6, no. 2, pp. 149–158, 2017. DOI: 10.11591/eei.v6i2.648.
- [8] E. C. Barrett and L. F. Curtis, "Introduction to environmental remote sensing," *Psychology Press*, 1999.
- [9] B. C. Naik and B. Anuradha, "Extraction of water-body area from high-resolution Landsat imagery," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 6, pp. 4111–4119, 2018. DOI: <http://doi.org/10.11591/ijece.v8i6.pp4111-4119>.
- [10] S. Dhingra and D. Kumar, "A review of remotely sensed satellite image classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 3, pp. 1720–1731, 2019. DOI: <http://doi.org/10.11591/ijece.v9i3.pp1720-1731>.
- [11] M. Dimiyati, A. Fauzy, and A. S. Putra, "Remote sensing technology for disaster mitigation and regional infrastructure planning in an urban area: a review," *Telkomnika Telecommunication Computing Electronics and Control*, vol. 17, no. 2, pp. 601–608, 2019. DOI: <http://dx.doi.org/10.12928/telkomnika.v17i2.12242>.
- [12] B.-R. B. Semlali, C. El Amrani, and S. Denys, "Development of a java-based application for environmental remote sensing data processing," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 3, pp. 1978–1986, 2019. DOI: <http://doi.org/10.11591/ijece.v9i3.pp1978-1986>.
- [13] I. M. Shirbhate and S. S. Barve, "Solar panel monitoring and energy prediction for smart solar system," *International Journal of Advances in Applied Sciences (IJAAS)*, vol. 8, no. 2, pp. 136–142, 2019. DOI: <http://doi.org/10.11591/ijaas.v8.i2.pp136-142>.
- [14] H. F. Hawari, A. A. Zainal, and M. R. Ahmad, "Development of real-time internet of things (IoT) based air quality monitoring system," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 13, no. 3, pp. 1039–1047, 2019. DOI: <http://doi.org/10.11591/ijeecs.v13.i3.pp1039-1047>.
- [15] J. J. Quackenboss, J. D. Spengler, M. S. Kanarek, R. Letz, and C. P. Duffy, "Personal exposure to nitrogen dioxide: relationship to indoor/outdoor air quality and activity patterns," *Environmental science & technology*, vol. 20, no. 8, pp. 775–783, 1986. <https://doi.org/10.1021/es00150a003>.
- [16] A. Chauhan, M. Krishna, A. Frew, and S. Holgate, "Exposure to nitrogen dioxide (no2) and respiratory disease risk." *Reviews on environmental health*, vol. 13, no. 1-2, pp. 73–90, 1998.
- [17] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling" in *Proceedings of the seventh IEEE international conference on computer vision*, IEEE, vol. 2. 1999, pp. 1033–1038. DOI: 10.1109/ICCV.1999.790383.
- [18] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, pp. 341–346, 2001. <https://doi.org/10.1145/383259.383296>.
- [19] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on image processing*, vol. 13, no. 9, pp. 1200–1212, 2004. DOI: 10.1109/TIP.2004.833105.
- [20] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE transactions on image processing*, vol. 10, no. 8, pp. 1200–1211, 2001. DOI: 10.1109/83.935036.
- [21] J. Hays and A. A. Efros, "Scene completion using millions of photographs," *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, p. 4, 2007. <https://doi.org/10.1145/1276377.1276382>.
- [22] R. Kohler, C. Schuler, B. Scholkopf, and S. Harmeling, "Mask-specific inpainting with deep neural networks" in *German Conference on Pattern Recognition*. Springer, 2014, pp. 523–534.
- [23] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *Advances in neural information processing systems*, 2014, pp. 1790–1798.
- [24] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2536–2544.
- [25] C. Yang, X. Lu, Z. Lin, E. Shechtman, O. Wang, and H. Li, "High-resolution image inpainting using multi-scale neural patch synthesis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6721–6729.
- [26] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent image completion," *ACM Transactions on Graphics (ToG)*, vol. 36, no. 4, p. 107, 2017. <https://doi.org/10.1145/3072959.3073659>.
- [27] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5505–5514. DOI: 10.1109/CVPR.2018.00577.
- [28] Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., & Huang, T. S., "Free-form image inpainting with gated convolution," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 4471–4480.
- [29] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein, "The generalized patchmatch correspondence algorithm," in *European Conference on Computer Vision*. Springer, 2010, pp. 29–43. DOI: https://doi.org/10.1007/978-3-642-15558-1_3.
- [30] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90

- [31] J. Shermeyer and A. Van Etten, "The Effects of Super-Resolution on Object Detection Performance in Satellite Imagery," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1432-1441, 2019. DOI: 10.1109/CVPRW.2019.00184.
- [32] Jayanetti, J. M., Meedeniya, D. A., Dilini, M. D. N., Wickramapala, M. H., & Madushanka, J. H., "Enhanced land cover and land use information generation from satellite imagery and foursquare data," In *Proceedings of the 6th International Conference on Software and Computer Applications*, pp. 149-153, 2017. <https://doi.org/10.1145/3056662.3056681>.
- [33] Meedeniya, D. A., Jayanetti, J. M., Dilini, M. D. N., Wickramapala, M. H., and Madushanka, J. H., "Land-Use Classification with Integrated Data," *Machine Vision Inspection Systems: Image Processing, Concepts, Methodologies and Applications*, 2020, ch. 1, pp. 1-36. <https://doi.org/10.1002/9781119682042.ch1>.

BIOGRAPHIES OF AUTHORS



Mohamed Akram Zaytar is a Ph.D. candidate at the faculty of sciences and technologies Tangier, working at the intersection of representation learning and environmental science. He obtained his Master's Degree in Information Systems and Networking from the Faculty of Sciences and Technologies (Tangier) in 2017. He is interested in machine learning, representation learning, and climate change. He has research experience in weather forecasting, air quality monitoring, precision agriculture, and satellite imagery applications. His work has been featured on international venues such as ICLR 2020, Deep Learning Indaba 2018, and EGU 2018.



Professor Chaker El Amrani is a Doctor in Mathematical Modelling and Numerical Simulation from the University of Liege, Belgium (2001). He lectures in distributed systems and promotes HPC education at the University of Abdelmalek Essaadi. His research interests include Cloud Computing, Big Data Mining, and Environmental Science. Dr. El Amrani has served as an active volunteer in IEEE Morocco. He is currently Vice-Chair of IEEE Communication and Computer Societies Morocco Chapter, and advisor of the IEEE Computer Society Student Branch Chapter at Abdelmalek Essaadi University. He is the NATO Partner Country Director of the real-time remote sensing initiative for early warning and mitigation of disasters & epidemics in Morocco.