DATA AUGMENTATION APPROACHES FOR SATELLITE IMAGE SUPER-RESOLUTION

M. A. A. Ghaffar[1], A. McKinstry[1] , T. Maul[2], T. T. Vu[3]

[1] Irish Centre for High-End Computing (ICHEC), Ireland - (moataz.ahmed,alastair.mckinstry)@ichec.ie
[2] The University of Nottingham Malaysia, School of Computer Science, Malaysia - tomas.maul@nottingham.edu.my
[3] Hoa Sen University, Ho Chi Minh , Vietnam - thuy.vutuong@hoasen.edu.vn

ICWG II/III: Pattern Analysis in Remote Sensing

KEY WORDS: Deep Learning, Data Augmentation, Super Resolution, Landsat-8, Convolutional Neural Networks

ABSTRACT:

Data augmentation is a well known technique that is frequently used in machine learning tasks to increase the number of training instances and hence decrease model over-fitting. In this paper we propose a data augmentation technique that can further boost the performance of satellite image super resolution tasks. A super-resolution convolutional neural network (SRCNN) was adopted as a state-of-the-art deep learning model to test the proposed data augmentation technique. Different augmentation techniques were studied to investigate their relative importance and accuracy gains. We categorized the augmentation methods into instance based and channel based augmentation methods. The former refers to the standard approach of creating new data instances through applying image transformations to the original images such as adding artificial noise, rotations and translations to training samples, while in the latter we fuse auxiliary channels (or custom bands) with each training instance, which helps the model learn useful representations. Fusing auxiliary derived channels to a satellite image RGB combination can be seen as a spectral-spatial fusion process as we explain later. Several experiments were carried out to evaluate the efficacy of the proposed fusion-based augmentation method compared with traditional data augmentation techniques such as rotation, flip and noisy training inputs. The reconstruction quality of the high resolution output was quantitatively evaluated using Peak-Signal-To-Noise-Ratio (PSNR) and qualitatively through visualisation of test samples before and after super-resolving.

1. INTRODUCTION

Massive advances have been witnessed in Remote Sensing technologies recently. Whether improving sensors or processing techniques, these advances are reflected in terms of the 5Vs of big data (Yang et al., 2016a), i.e.: volume, velocity, variety, veracity and value. The spatial, spectral and temporal resolutions of Remote Sensing data, are core contributors to the volume factor. Moreover, very high spatial resolution (VHR) datasets are increasingly emerging as superior solution components that cover the needs of various geospatial and remote sensing applications (Romero et al., 2016). For example, submeter spatial resolution has been reached by recently launched satellites such as WorldView-4 (SIC, 2017). Nevertheless, VHR datasets are still expensive and only low to mid resolution datasets such as Landsat and Sentinel (AWS, 2016, AWS, 2017) are open for free use and download. For this reason, there is a current need for enhancing the freely available, lower resolution datasets.

In line with progress in sensor accuracy, data processing techniques are continually being developed and refined. Due to its ability to form useful hierarchical representations of complex data, deep learning has made significant contributions to the field of image processing (Huang et al., 2015). Deep learning algorithms have been used in several remote sensing applications such as classification (Liu et al., 2017) (Kussul et al., 2017), feature selection and extraction (Zhang , Wang, 2015) and super-resolution (SR) (Liebel , Körner, 2016). It is a perfect match to big data (Schmitt , Zhu, 2016) and it has shown high performance levels in classical computer vision problems such as SR (Dong et al., 2016). From a methodological perspective, SR using deep learning techniques can be achieved either using multi-frame or single-image approaches. The former relies on building a model which is capable of learning how to enhance spatial resolution through learning features from multiple scenes from different perspectives/angles of the same spot. With reference to satellite image super resolution, the multi-frame method could be interpreted as super-resolving the overlapping areas between scenes e.g. scenes 127,057 and 127,058 (see figure 1). The aforementioned method can exploit useful new information from the different perspectives and sub-pixel shifting of the same spot that could be used to form the high resolution output image (Bevilacqua, 2014). The single-image approach, on the other hand, relies on recovering a high resolution image from a single lower resolution image (Dong et al., 2016). The original high-resolution image corresponds to the target and a synthesized version is generated to represent the lower resolution source by a down-sampling process applied to the target, followed by an up-sampling of the target image using the same down-sampling factor.

Unlike trivial image interpolation processes (e.g. bicubic or bilinear filtering) which rely mainly on magnifying the existing details of the input image, SR targets the prediction of details that were missing from the input low resolution image (Bevilacqua, 2014).

In this study we focus on single satellite image super resolution using a deep learning approach. Building training datasets for deep learning algorithms is a problematic process that can affect the performance and speed of classification or regression values

(Castelluccio et al., 2015, Gardner Daniel, 2017), and satellite images are no exception. It is usually difficult to build datasets of the scale required by deep learning. Data augmentation is frequently reported to mitigate the problem of data scarcity in many domains (Ding et al., 2016). The technique also reduces over-fitting, which can occur when applying the model to closely related data instances (Wang , Perez, 2017) and overcomplex models. However, unfortunately not much work has been reported in the domain of satellite image super-resolution with data augmentation techniques[1].

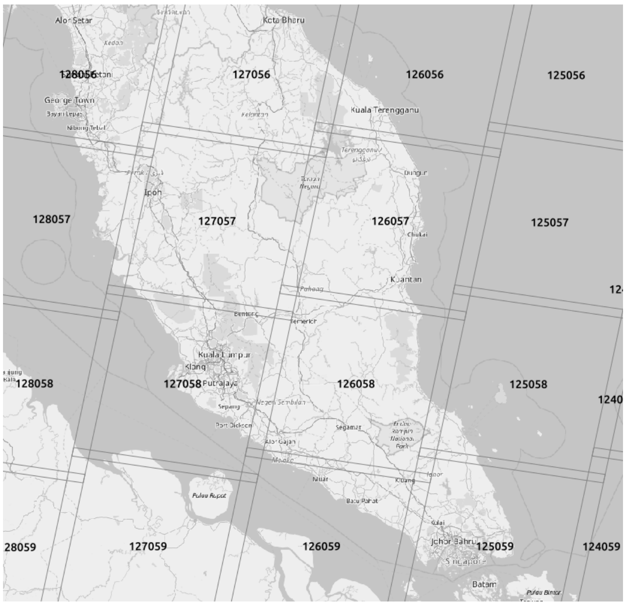

Figure 1. The overlapping areas between scenes.

In (Pouliot et al., 2018), the authors tested satellite image super resolution using deep and shallow convolutional neural networks to demonstrate the effects of tuning the training model depth on the accuracy. The authors used SRCNN methods to super resolve Landsat-8 and Landsat-5 images using Sentinel-2 data as a target/label. Although the authors explained the importance of data augmentation techniques, they stated clearly that no data augmentation approaches were used except for Landsat-5 data to cover the lack of training samples. In (Liebel , Körner, 2016), SR was applied on Sentinel-2 (AWS, 2017) 13 spectral bands without mentioning any usage of data augmentation techniques. The same applies to (Wang et al., 2018), where super-resolution convolutional neural networks (SRCNNs) were applied to super-resolving aerial images. Both papers reported relatively high levels of accuracy. In this paper, we demonstrate how to apply channel and instance augmentation/fusion techniques to super-resolution problems in order to achieve higher accuracy. Inspired by (Cirean et al., 2012) and (Ahn et al., 2015) data augmentation approaches for classification and restoration problems and (Pohl , Van Genderen, 2016) data fusion approaches, we formed a spatial-spectral fusion at feature level which can boost satellite image super resolution accuracy. We also compare the performance of the proposed fusion technique against standard instance augmentation approaches encapsulated by rotated, flipped and noisy inputs.

[1] Searching Scopus and Google Scholar databases for papers published from 2012-2018 with the phrase ”satellite image” AND ”augmentation” AND ”super-resolution” OR ”super resolution” returned only 32 results and majority of them mentioned data augmentation as a technique to increase the number of training samples.

This study aims to answer the following questions:

1. Does data fusion via channel augmentation boost accuracy of satellite image super resolution?
2. What are the most useful artificial bands to be fused with input training samples?
3. Does traditional instance augmentation based on mirror flips, rotation, and noise, add value to the model?
4. What type of noise is most useful for instance augmentation for satellite image SR?

Section 2 provides a detailed description of the methodology, including the convolutional neural network, dataset, data augmentation techniques and experimental design adopted. Section 5 summarizes the main results obtained and section 6 provides several conclusions and guidelines for future work.

2. DATA AUGMENTATION METHODS

2.1 Data Augmentation

Data augmentation aims at enlarging the dataset in order to address gaps in data representation and to minimize the problem of over-fitting. This leads to improved model performance and prevents imbalanced learning. In case of the image classification domain, both satellite and conventional images have benefited from several augmentation techniques (Ding et al., 2016, Wang , Perez, 2017). In satellite image classification, clipping, rotating, flipping, shifting and translating are the most commonly used image transformations (Ding et al., 2016, Yang et al., 2016b), whereas in the context of satellite image super-resolution, materials weren’t available to form a strong opinion on which augmentation techniques were more useful.
Moreover, we assume that the statistical properties of satellite images are sufficiently different from generic photographs, thus affecting the effectiveness of SR solutions, and justifying a study that focuses on data augmentation techniques specifically in the satellite image super resolution domain.
We categorize augmentation techniques in 2 main groups as follows:

- **Instance-based augmentation** This refers to all well-known geometric, coloring, noise, deformation, translation, brightness and smoothing image transformations which have been used in scene classification and conventional image super resolution tasks to increase the number of samples/instances in a dataset.
- **Fusion-based augmentation** This represents stacking artificial bands that might add supplementary beneficial observations (or perspectives) to the model that contribute towards improving model accuracy. In other words, we add a set of derived channels to be fused with input data which increases the dimensionality of the data rather than the number of training instances.

Since there are many types of kernels that can be applied to images to derive meaningful features, in the next section we focus on some of the most relevant to denoising, sharpening and resolution enhancement, as we believe that such features might be useful to the model. Although deep architectures traditionally aim to automate the feature extraction process, we

https://doi.org/10.5194/isprs-annals-IV-2-W7-47-2019

prove quantitatively and qualitatively that fusing manually extracted features with training samples adds useful representations to the model which can boost super resolution accuracy. The proposed added features include edge detection, unsharp masking and contrast enhancement. We argue that adding such features to the input will prevent the network from learning irrelevant kernels in addition to emphasizing the edge-focused and sharpening features. At the same time, this can help in boosting the overall model performance. A detailed description of each channel and how it gets fused with the input layer is explained in the following sections.

- **Edge Detection**
  In general, any kind of convolutional neural network attempts to find useful representations of data via edges, corners, textures and many other features. At the beginning of a neural network, kernels extract simple features such as edges while the extraction becomes much more complex and abstract when it goes deeper through the network. Since edges are very important features in CNNs in general and particularly in super resolution, it motivated us to study the effect of fusing manually generated edge features on super resolution accuracy.

  In order to derive edge features form satellite images, we converted an RGB combination to grayscale first and hence we were able to create different types of edges features. Edge kernels such as Sobel, Touzi and Gradient are the most used ones in satellite image analysis applications. However, a gradient-based approach was adopted based on several experiments as we explain later in the results section.

  Gradient edge detection uses derivatives to detect large intensity changes in an image (Katiyar , Arun, 2014) using the magnitude and direction of vectors. Magnitude and direction can be calculated through equations 1 and 2 respectively.

$$\|\bigtriangledown I\| = \sqrt{(\frac{\delta F}{\delta x})^2 + (\frac{\delta F}{\delta y})^2} \tag{1}$$

$$\theta = \tan^{-1}[\frac{\delta F}{\delta y}/\frac{\delta F}{\delta x}] \tag{2}$$

  where ***F*** refers to the image function
  x and ***y*** are directions axis

  In order to measure the effects of fusing the gradient edge feature with SRCNN, it was necessary to inspect the output feature maps of a training sample stacked with gradient features against the same training sample with only RGB color combination to evaluate the resulting representations.

  Fig. 2 shows a sentinel-2 training sample containing an aircraft (A) and the results of convolving (A) with and without fusing edge feature after 1 epoch during first layer (9x9) of the SRCNN architecture. The convolution process was followed by non-linear activation function Rectified Linear Unit (Relu). The resulting feature map (C) was strongly affected by fusing gradient edge features as clearly seen from the number of activated neurons. We hypothesised that fusing such features at an early stage of training might assist the model by presenting useful representations.

  The edge detection derived channel was fused with the RGB color combination to enforce the model to learn and extract edge-focused features in addition to preserving details of complex image structures that could be beneficial to construct the HR output. Quantitative results using Peak-Signal-to-Noise-Ratio metric show the gains of fusing such features in the results section.

- **Contrast Enhancement**
  From an image enhancement perspective, contrast enhancement is used to improve the disparity between the objects in a scene and their background (Hall, 1979). We assume that enhancing contrast can expose several hidden but relevant image details, which can help in extracting significant features for SR. Among contrast enhancement methodologies, histogram equalisation is the most used algorithm in remote sensing image manipulation (Fu et al., 2015). Histogram equalisation aims to derive the intensity mapping that will as best as possible equalise the image histogram.

  As depicted in Fig. 3, the fusion of a contrast enhanced channel using histogram equalisation algorithm with the RGB color combination helped in representing more details in (C) which were not so apparent in the original training sample (A).

- **Unsharp Masking** As a simple method of increasing details in an image, unsharp masking consists of adding a mask to the original image (Maheshwari , Pati, 2014) which is created by subtracting a blurred version of the original image from the original one as follows. Let *A* be the original image, B the blurred version, *mask = B - A*. The created mask is then combined with the original image to produce a sharpened version. We presume that adding such channel will highlight the deblurring features to be extracted during training. Following the same approach of visualising gradient edge feature, Fig. 4 shows the difference between applying the same convolution process on training sample (B) with manually added contrast enhanced feature (C) and without adding in (A). Although it shows less added value than the gradient edges feature this might be justified as early layers in CNNs tend to focus on edges.

In the next section, we explain how datasets were created to test the proposed augmentation methods on both Landsat-8 and Sentinel-2 products.

3. DATASET

Creating a balanced training dataset for deep learning tasks is an important prerequisite to having an unbiased model. In remote sensing, hundreds of scenes are being captured and sent to ground stations for publishing and ingestion. Starting from late 2008, Landsat-8 (AWS, 2016) data became open for download by all users free of charge and within a few hours of scene capture. The 11-band spatial resolution is 30m for visible and infrared bands and 15m for panchromatic data.
On the other hand, the European Union adopted an open data policy for the Copernicus programme (Sentinel Online FAQ, 2018). Sentinel-2 data contains 13 bands with spatial resolution varying from 10, 20 to 60 meter bands which serves a wide range of applications related to the Earth.

https://doi.org/10.5194/isprs-annals-IV-2-W7-47-2019

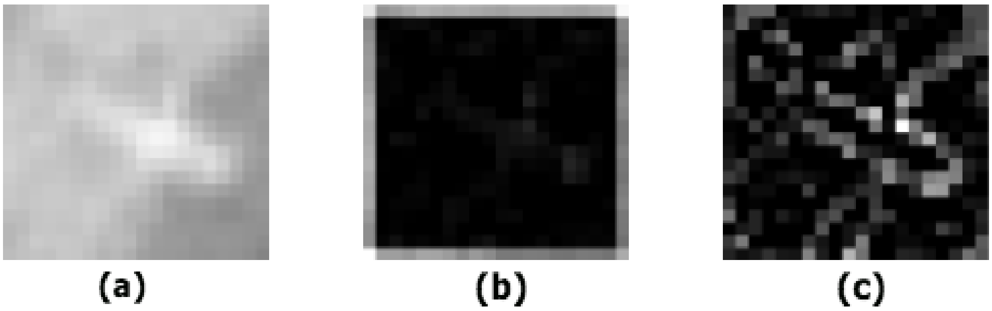

Figure 2. Difference between convolution output of manually added gradient edge feature (C) and convolution output for same training sample without adding gradient edge feature (B) while original training sample is (A).

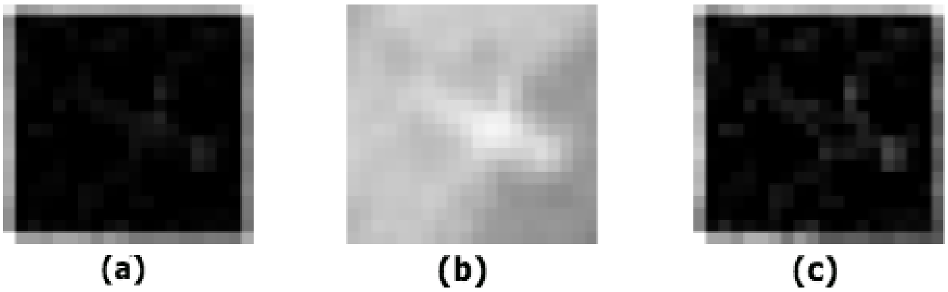

Figure 3. Difference between convolution output of manually added contrast enhanced feature (C) and convolution output for same training sample without adding contrast enhanced feature (A) while original training sample is (B).

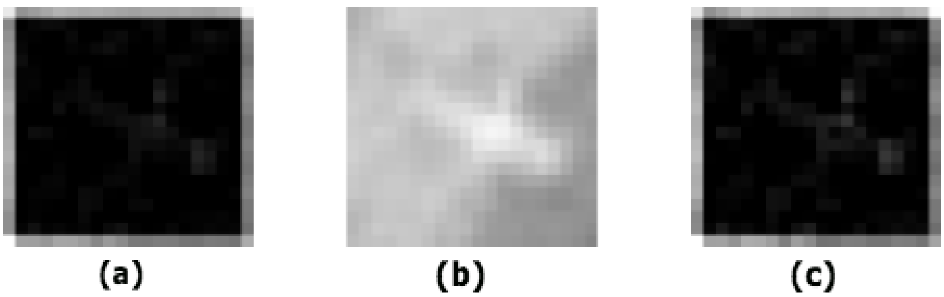

Figure 4. Difference between convolution output of manually added unsharp mask feature (C) and convolution output for same training sample without adding unsharp mask feature (A) while original training sample is (B).

The study area consists of the western side of Malaysia which is covered by paths 126, 127 and rows 57 and 58 as shown in Fig. 1 for landsat-8 Worldwide Reference System (WRS) Path/Row. Regarding Sentinel-2, the same region was covered by granules 47NQE, 47NQD, 47NRE and 47NRD.

Given an image sample size of 64x64 pixels, 12k samples were created from the selected scenes with a stride of 8 pixels. 8400 samples were allocated for training while 3600 were allocated for validation. Training samples were divided into X and Y, where Y consists of the original samples representing the high resolution target, and X refers to the corresponding synthesized low resolution version of the original samples, obtained by downscaling and subsequent upscaling (with a corresponding scaling factors).

4. EXPERIMENTAL DESIGN

This section explains the achieved experiments to evaluate the efficacy of adding fusion-based data augmentation compared to traditional data augmentation techniques. In addition, utilising the flexibility of the SRCNN 9-1-5 architecture, we evaluate using SRCNN in new domains such as denoising and dehazing using synthesised noise and haze effects.

In order to evaluate the proposed augmentation methods on Landsat-8 and Sentinel-2 datasets, a grayscale scene was created to derive the unsharp mask, contrast enhanced, and gradient edge bands, in addition to bands 4 (red), 3 (green) and 2 (blue). The selected scene channels were then stacked prior to cropping patches of size 64×64. The created patches were cropped through a grid of vector polygons designed with a stride of 8 pixels. The raw dataset of 12,000 samples was divided into 70%, 20% and 10% portions to form non-overlapping training, validation and test sets respectively for both Landsat-8 and Sentinel-2 products. A set of experiments were designed with the following core conditions:
Super-Resolution RGB with instance-based augmentation:

- Bands 4,3,2 in addition to instance augmentation including (flipping, rotating, noise and smoothing).

Super-Resolution RGB with fusion-based augmentation:

- Bands 4,3,2 fused with gradient edge detected feature
- Bands 4,3,2 fused with unsharp masking feature
- Bands 4,3,2 fused with contrast enhanced feature
- Bands 4,3,2 in addition to near and shortwave infrared bands
- Bands 4,3,2 fused with gradient edge detected, contrast enhanced and unsharp masking features

Moreover, additional augmentation methodologies were tested in different domains such as denoising and dehazing as follows:
Denoising:

- Bands 4,3,2 with noise effects (salt-and-pepper).

Dehazing:

- Bands 4,3,2 with synthesised haze effects.

In the next section, we demonstrate how the results of SRCNN are influenced by our proposed augmentation method compared with traditional instance-based augmentation. Moreover, the SRCNN was tested against denoising and dehazing tasks. As an example of synthesised noise, salt-and-pepper effect was added as augmented samples. In addition, to simulate the haze effect, red, green and blue bands were split to modify the contrast of every channel separately. The effect of haze is obviously noticed in blue bands and less in red and green bands as mentioned in (Qin et al., 2018). Although haze effects are usually not following a uniform distribution especially that it depends on atmospheric scattering models (Qin et al., 2018), this experiment aimed to approach the haze problem from contrast deterioration perspective. The assessment of image restoration can be measured through Peak Signal to Noise Ratio (PSNR) metric. PSNR can be denoted as:

$$\mathrm{PSNR} = 10 \cdot \log_{10} \frac{C^2}{\mathrm{MSE}} \tag{3}$$

where C^2 refers to the maximum value a pixel can have. MSE refers to the Mean Squared Error which can be obtained using equation 4.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} ||\mathbf{Y}_i - \mathbf{X}_i||^2 \tag{4}$$

where ***Y*** represents the predicted high resolution output ***X*** represents the ground truth sample.

5. RESULTS

Taking advantage of the flexibility of the SRCNN design, various channels have been merged to test the effectiveness of fusion-based augmentation methodologies compared to instance-based ones. In addition, we introduce the SRCNN architecture to new problems such as denoising and dehazing.[2] The results are shown in table 1 and figures 7, 9 and 8, where we can observe the following: (i) insignificant performance improvement when instance augmentation including flipping and rotating have been applied to the RGB channels; (ii) it is noteworthy that gradient edge detection provides the most useful fused feature compared to contrast enhanced and unsharp mask features; (iii) adding edge detection, contrast enhancement and unsharp masking channels to the visual bands improved the performance more than any other augmentation method we experimented with SRCNN; (iv) RGB is improved to some extent by instance augmentation and infrared channels, but more significantly by the proposed channel augmentation. On the other hand, figures 5 and 6 show the denoised and dehazed samples. The proposed augmentation approaches to simulate noise and haze effects didn't work as expected and the results show spectral distortions throughout test samples.

6. CONCLUSIONS

In this paper, we presented data augmentation methods that help boost super resolution accuracy in the satellite image domain. The proposed data augmentation methodology integrates

[2] Due to unsatisfying results, only qualitative results for the denoising and dehazing experiments are included.

https://doi.org/10.5194/isprs-annals-IV-2-W7-47-2019

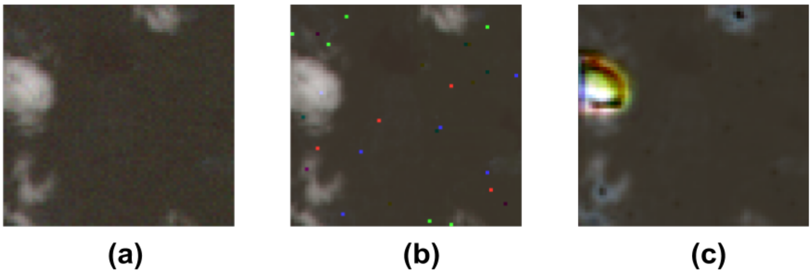

Figure 5. Example of denoise process where (a) represents the original sample, (b) represents the added salt-and-pepper noise and (c) is the result.

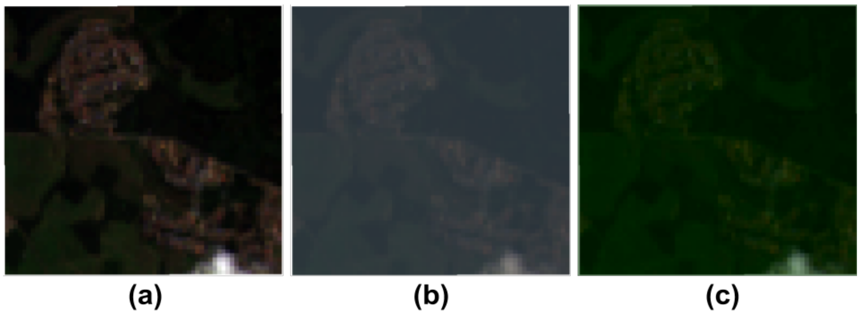

Figure 6. Example of dehaze process where (a) represents the original sample, (b) represents the synthesised haze and (c) is the result.

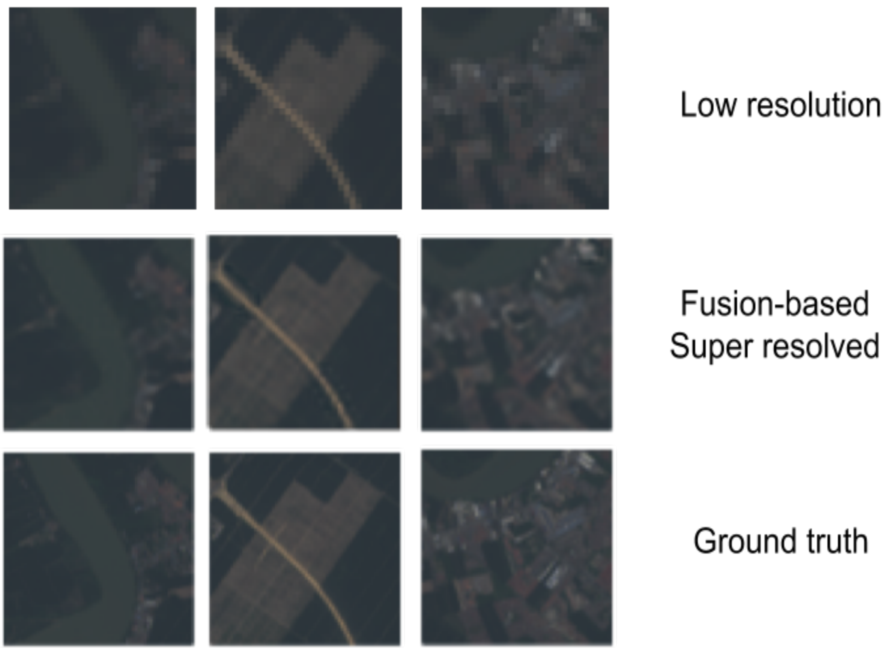

Figure 7. Examples from test set results of the proposed fusion-based augmentation compared with ground truth samples.

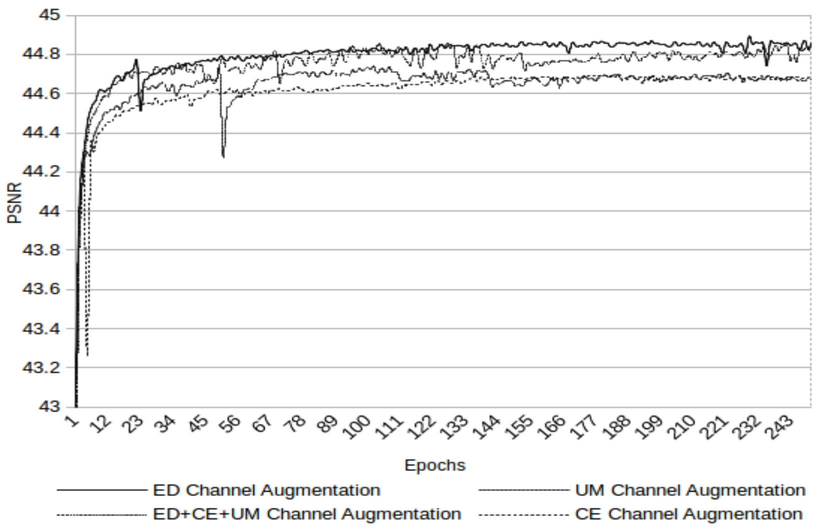

Figure 8. Comparison between the learning convergence of different channel augmentations added to RGB channels.

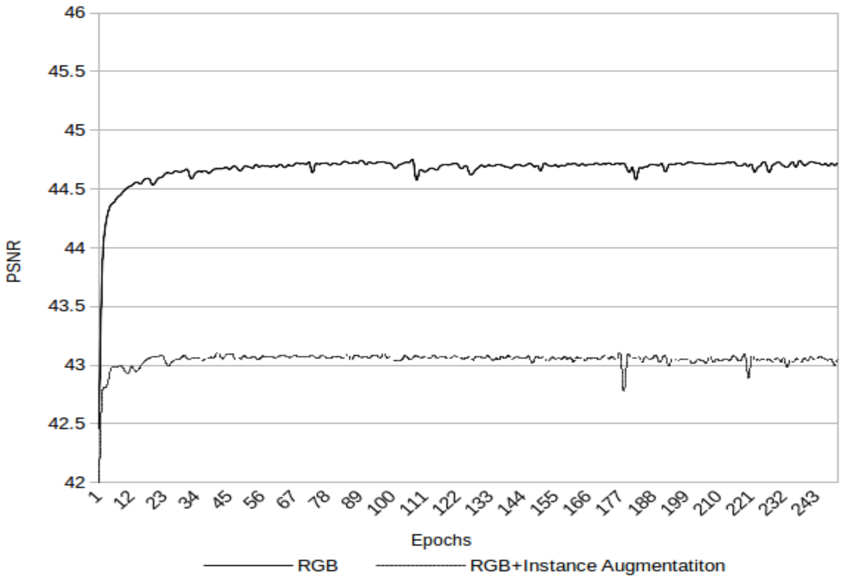

Figure 9. Comparison between RGB channels with and without instance augmentation learning convergence.

Table 1. A comparison between different augmentation approaches testing results where CE refers to Contrast Enhanced, UM Unsharp Mask and ED Edge Detection.

Training strategy	**Average PSNR in dB**
RGB	**44.768**
RGB with instance augmentation	**45.224**
RGB fused with CE	45.254
RGB fused with UM	45.023
RGB fused with ED	45.605
RGB fused with (CE,UM,ED)	**45.788**
RGB + NIR + SWIR I + SWIR II	45.125

the power of SRCNN for learning nonlinear mappings between low and high resolution images, with the flexibility to incorporate additional feature representations via fusing auxiliary input channels with input data. A selection of artificial channels was experimented with, that provided an alternative training strategy, which contributed to the construction of more accurate high resolution outputs. Following the rule of thumb of evaluating image structure similarity, the PSNR metric was used to numerically assess the super resolved images by comparing them against the original ground truth high resolution images. The experimental results showed the effectiveness of our proposed fusion-based augmentation approach compared with standard instance augmentation methods. However, limitations were highlighted when augmentation was introduced to different problem domains such as denoising and dehazing. Since it is out of the scope of this paper, in future work we aim to explore the required changes to adapt the model for denoising and dehazing problems.

REFERENCES

Ahn, B., Cho, N.I., 2015. Self-Commmittee Approach for Image Restoration Problems using Convolutional Neural Network. *unpublished*, 14, 1–5.

AWS, 2016. Landsat on AWS. https://aws.amazon.com/public-datasets/landsat/. Accessed: 2017-11-20.

AWS, 2017. Sentinel-2 on AWS.https://aws.amazon.com/public-datasets/sentinel-2/. Accessed:2017-11-20.

Bevilacqua, M., 2014. Algorithms for super-resolution of images and videos based on learning methods. Theses, Uni-versité Rennes 1.

Castelluccio, M., Poggi, G., Sansone, C., Verdoliva, L., 2015. Land Use Classification in Remote Sensing Images by Convolutional Neural Networks. *CoRR,* abs/1508.00092. http://arxiv.org/abs/1508.00092.

Cirean, D., Meier, U., Schmidhuber, J., 2012. Multi-column Deep Neural Networks for Image Classification. *International Conference of Pattern Recognition*, 3642–3649.

Ding, J., Chen, B., Liu, H., Huang, M., 2016. Convolutional Neural Network With Data Augmentation for SAR Target Recognition. *IEEE Geoscience and Remote Sensing Letters*, 13, 364-368.

Dong, C., Loy, C.C., He, K., Tang, X., 2016. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38, 295-307.

Fu, X., Wang, J., Zeng, D., Huang, Y., Ding, X., 2015. Remote Sensing Image Enhancement Using Regularized-Histogram Equalization and DCT. *IEEE Geoscience and Remote Sensing Letters*, 12, 2301-2305.

Gardner, D., Nichols, D., 2017. Multi-label Classification of Satellite Images with Deep Learning. *CoRR*. http://cs231n.stanford.edu/reports/2017/pdfs/908.pdf.

Hall, Ernest, 1979. Computer Image Processing and Recognition.

Huang, W., Xiao, L., Wei, Z., Liu, H., Tang, S., 2015. A New Pan-Sharpening Method With Deep Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, 12, 1037–1041. http://ieeexplore.ieee.org/lp-docs/epic03/wrapper.htm?arnumber=7018004.

Katiyar, S.K., Arun, P.V., 2014. Comparative analysis of com mon edge detection techniques in context of object extraction. *IEEE TGRS*, 50, 68–79. https://arxiv.org/pdf/1405.6132.pdf.

Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*,14,778–782. http://ieeexplore.ieee.org/document/7891032/.

Liebel, L, Körner, M, 2016. Single-image Super Resolution for Multispectral Remote Sensing Data using Convolutional Neural Networks. *The International Arhives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 883–890.

Liu, P., Zhang, H., Eom, K.B., 2017. Active Deep Learning for Classification of Hyperspectral Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10, 712–724.http://ieeexplore.ieee.org/document/7568999/.

Maheshwari, S., Pati, U.C., 2014. Mosaicing of images using unsharp masking algorithm for interes point detection. *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies*, IEEE,1431–1434.

Pohl, C., Van Genderen, J., 2016. *Remote sensing image fusion: A practical guide*. Crc Press.

Pouliot, D., Latifovic, R., Pasher, J., Duffe, J., 2018. Landsat Super-Resolution Enhancement Using Convolution Neural Networks and Sentinel-2 for Training. *Remote Sensing*, 10. http://www.mdpi.com/2072-4292/10/3/394.

Qin, M., Xie, F., Li, W., Shi, Z., Zhang, H., 2018. Dehazing for Multispectral Remote Sensing Images Based on a Convolutional Neural Network With the Residual Architecture. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11, 1645-1655.

Romero, A., Gatta, C., Camps-Valls, G., 2016. Unsupervised Deep Feature Extraction for Remote Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54, 1349 – 1362.

Schmitt, M., Zhu, X., 2016. Data Fusion and RemoteSensing An Ever-Growing Relationship. 4, 6-23.

Sentinel Online FAQ, 2018. ht-tps://sentinel.esa.int/web/sentinel/faq.

SIC, 2017. Worldview-4 satellite sensor (0.31m). https://www.satimagingcorp.com/satellite-sensors/geoeye-2/. Accessed: 2017-11-20.

Wang, J., Perez, L., 2017. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. *unpublished*. http://cs231n.stanford.edu/reports/2017/pdfs/300.pdf.

Wang, T., Sun, W., Qi, H., Ren, P., 2018. Aerial Image Super Resolution via Wavelet Multiscale Convolutional Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, 1–5.

Yang, C., Huang, Q., Li, Z., Liu, K., Hu, F., 2016a. Big Data and cloud computing: innovation opportunities and challenges. *International Journal of Digital Earth*, 10, 13–53. https://www.tandfonline.com/doi/full/10.1080/17538947.2016.1239771.

Yang, N., Tang, H., Sun, H., Yang, X., 2016b. Dropband: a convolutional neural network with data augmentation for scene classification of VHR satellite images.

Zhang, T., Wang, Q., 2015. Deep Learning Based Feature Selection for Remote Sensing Scene Classification. *IEEE Geoscience and Remote Sensing Letters*, 12, 2321–2325.